**І.В. Бейко**
**П.М. Зінько**
**О.Г. Наконечний**

# Задачі, методи і алгоритми оптимізації



Начальний посібник

Рекомендовано Міністерством освіти і науки України
як навчальний посібник для студентів вищих навчальних закладів

Рівне – 2011

УДК 519.9

Рецензенти: **А.О. Чикрій**, професор, доктор фіз.-мат. наук, член-кореспондент НАН України (Інститут кібернетики імені В.М. Глушкова НАН України)
**В.Я. Данилов**, професор, доктор технічних наук (ННК «Інститут прикладного системного аналізу» НТУУ «Київський політехнічний інститут»)

Гриф надано Міністерством освіти і науки України,

лист № 1/11-7429 від 6 серпня 2010 р.



Бейко І.В., Зінько П.М., Наконечний О.Г.
**Задачі, методи і алгоритми оптимізації:** Навчальний посібник. – Рівне: НУВГП, 2011. – 624 с.

В навчальному посібнику викладено сучасні методи і алгоритми для розв'язування задач оптимізації, що виникають в різних областях науки і техніки, у сфері управління економічними, соціальними, екологічними, технічними та іншими процесами. Розглянуто задачі оптимізації в багатокритеріальних та ієрархічних системах, задачі одновимірної оптимізації, задачі безумовної оптимізації (диференційованої та недиференційованої), задачі лінійного програмування, задачі нелінійного та стохастичного програмування, задачі оптимізації в нескінченновимірних просторах. Структура викладу матеріалу уніфікована: формулювання задачі, припущення, опис ідеї методу, покроковий алгоритм, теореми збіжності, зауваження і практичні поради щодо використання алгоритму.

Для студентів, аспірантів, науковців, котрі спеціалізуються в області теорії і практики прийняття оптимальних рішень, планування, прогнозування, проектування, виробництва й експлуатації систем різної природи.

ISBN _______________

# Зміст

Позначення і символи ………………………………………………………… 13
Передмова ……………………………………………..……………………. 16

***Вступ.* Типові задачі оптимізації. Умови оптимальності і алгоритми оптимізації …………………………………………...……………………..** 18

0.1. Приклади типових задач оптимізації ………………………………… 18
0.2. Методи повного перебору та цiленаправленого перебору. Алгоритми відтинань та мінорант …….......................................................................... 25
1. Методи повного перебору ……………………………………..……… 25
2. Методи відтинань та алгоритми медіан і половинного поділу ……………… 26
3. Метод мінорант …………………………………………………..….. 28
0.3. Необхідна умова мінімуму диференційованої функції і градієнтні алгоритми оптимізації ……………………………….......................... 30
1. Необхідна умова мінімуму ……………………………….…………… 30
2. Екстремальні розв'язки задачі оптимізації. Алгоритм обчислення екстремального розв'язку ……………………………….……………… 31
3. Метод Ньютона і метод січних для уточнення екстремального розв'язку …… 32
0.4. Умови оптимальності і методи мінімізації диференційованих функцій багатьох змінних ………..………………………….......................... 33
1. Методи Монте–Карло …………………….……………………………… 33
2. Необхідна умова мінімуму диференційованої функції багатьох змінних і алгоритм найшвидшого спуску …………………………………….……… 35
3. Методи послідовних наближень ……………………………………… 38
4. Градієнтні методи, алгоритми допустимих напрямків і теореми збіжності …. 39
5. Необхідні і достатні умови оптимальності для задачі мінімізації опуклої функції. Алгоритм узагальнених градієнтів ……………………… 44
6. Методи відтинань для мінімізації опуклих функцій. Метод еліпсоїдів ……… 47
7. Умови оптимальності для задачі мінімізації неперервної функції з кусково-неперервними градієнтами ……………………………………… 49
8. Методи стохастичних градієнтів ……………………………………… 51
9. Методи штрафів для задач оптимізації з обмеженнями ……………………… 53
0.5. Прискорені методи оптимізації ………………………………………...... 55
1. Алгоритм спряжених градієнтів для мінімізації квадратичних функцій …..… 55
2. Методи другого порядку. Метод Ньютона ……………………………….… 56
3. Квазіньютонівські методи …………………………………….……… 57
4. Прискорені методи оптимізації із розтягуванням простору ……………… 58
5. $r$-алгоритми мінімізації з розтягуванням простору у напрямку різниці субградієнтів .................................................................................................. 60
0.6. Умови оптимальності для задач оптимізації із обмеженнями …………… 61
1. Теорема Каруша–Куна–Таккера для задачі опуклої оптимізації ……………… 61
2. Метод внутрішньої точки в лінійному програмуванні ................................ 62
3. Необхідні умови оптимальності для задач оптимізації з диференційованими функціями обмежень ........................................................................ 64
0.7. Методи оптимізації в задачах з обмеженнями ………………………….. 64

1. Методи штрафів для задач оптимізації з обмеженнями ……………….....… 64
2. Алгоритми двоетапної оптимізації ……………………………………… 66
0.8. Методи оптимізації в задачах варіаційного числення та оптимального керування …………………………………………………………………... 67
1. Задачі варіаційного числення ……………………………………………… 67
2. Задачі оптимального керування ........................................................... 74
3. Методи розв'язування типових задач оптимального керування ................ 76
3.1. Лінійна задача Майєра …………...…………………………..……. 76
3.2. Побудова керування оптимального по швидкодії …..................... 79
3.3. Двоїста задача Майєра. Мінімізація амплітуди керування ……. 80
3.4. Мінімізація квадратичного функціоналу ……………………….... 82
3.5. Узагальнена задача мінімізації квадратичного функціоналу ….. 85
3.6. Градієнтні методи побудови оптимального керування …..……. 86
3.7. Чисельні методи побудови оптимального керування при наявності обмежень на керування …………..………………………... 87
3.8. Чисельний метод побудови оптимального керування за принципом максимуму ………………………………………..….. 88
3.9. Розв'язуючі оператори для оптимального керування складними граф-операторними системами ……………………..…...……. 89
3.10. Використання методів оптимізації для побудови математичних моделей ………………………………………………. 91

***Розділ 1.* Алгоритми прийняття рішень і оптимізації в багатокритеріальних та ієрархічних системах** ……..……………………. 94

1.1. Прийняття рішень (вибір) і оптимізація ………………………………... 94
1. Критеріальний підхід до вибору альтернатив …………………………… 94
2. Нормалізований мультикритерій ………………........................... 95
3. Суперкритерій ………………………………………………………… 96
4. Метод головного частинного критерію ……………………………….. 97
5. Метод послідовних поступок ………………………......................... 97
6. Пошук альтернативи із заданими властивостями ……………………... 98
7. Метод бажаної точки ………………………………………………… 99
8. Прийняття рішень в умовах невизначеності ………….......................... 100
9. Приклади постановок задач багатокритеріальної оптимізації ……………. 102
9.1. Задача проектування оптимального програмного комплексу . 102
9.2. Трьохрівнева задача керування гнучким автоматизованим виробництвом ……………………………..…………. 103
1.2. Оптимізація в системах з ієрархічною структурою ………………. 104
1. Тривіальний випадок …………………………………………….. 105
2. Загальний випадок ……………………………………………….. 105
2.1. Принцип гарантованого результату ……………………………. 106
2.2. Прийняття рішень в умовах доброзичливості ………………... 106
3. Приклади ієрархічних систем керування …………………………… 107
3.1. Задача розподілу ресурсів ……………………………..……. 107
3.2. Задача нормування шкідливих викидів ………….............. 108
110

3.3. Задача управління економічною системою за допомогою штрафів і доплат (заохочень) …………………………....... 115
Запитання та вправи для самостійної роботи до розділу 1 ………………....... 116

*Розділ 2.* **Методи одновимірної оптимізації** ……………………………… 116

2.1. Методи Фібоначі ……………………………………………………… 116
1. Основний алгоритм …………………………………………………….. 116
2. Модифікація методу Фібоначі ………………………………………… 117
2.2. Метод золотого перерізу …………………………………………….... 120
2.3. Методи типу Ньютона …………………………......................... 124
1. Метод Ньютона ………………………………………………………… 124
2. Метод січних ……………………………………………………………… 126
2.4. Методи дотичних ………………………………............................ 127
1. Випадок диференційованої функції …………………………………… 127
2. Випадок недиференційованої функції ………………………………… 128
2.5. Методи глобального пошуку ………………………………………… 129
1. Метод глобального пошуку …………………………………………… 129
2. Рандомізований метод глобального пошуку ……………………….. 131
2.6. Методи пошуку інтервалу найбільших значень багатоекстремальних функцій ……………………………………………………………………….. 132
2.7. Методи пошуку глобального мінімуму, які використовують стохастичні автомати ……………………………………………………………………… 135
1. Метод, який використовує модель Буша–Мостеллера ……………… 135
2. Метод, який використовує усереднені значення функції ...................... 137
2.8. Адаптивні методи ……………………………………………………….. 138
1. Метод Кіфера–Вольфовиця ……………………………………………… 138
2. Простий перебір ……………………………………......................... 139
Запитання та вправи для самостійної роботи до розділу 2 ………………… 143

*Розділ 3.* **Методи оптимізації диференційованих функцій в задачах без обмежень** ………………………………………………………………….…... 145

3.1. Аналітичні методи ……………………………………………………….. 145
1. Необхідна умова локальної оптимальності ………………………………. 145
2. Достатня умова локальної оптимальності ……………………………….. 145
3.2. Градієнтні методи ……………………………………………………….. 150
1. Метод найшвидшого спуску …………………………………………….... 151
2. Модифікований метод найшвидшого спуску ……………………….…. 153
3. Основний варіант градієнтного методу ……………………………..…. 155
4. Модифікація основного варіанту градієнтного методу (метод з подрібненням кроку) …………………………………………………..……………… 156
5. Градієнтний метод з постійним кроковим множником …..……………… 158
6. Варіант градієнтного методу із матрицею прискорення збіжності ..………… 160
7. Модифікований градієнтний метод, який не потребує обчислення похідних . 161
3.3. Методи типу Ньютона ………………………………………………..…. 164
1. Метод Ньютона–Канторовича ……………………………………...…… 165
2. Узагальнений метод Ньютона–Канторовича ………………………… 168

3. Модифікація узагальненого методу Ньютона–Канторовича …………….. 171
4. Модифікований узагальнений метод Ньютона–Канторовича, який не вимагає обчислення матриці других похідних ……………………………………. 173
3.4. Методи спряжених градієнтів ……………………………………… 175
1. Загальна схема алгоритмів спряжених градієнтів ……………………….. 175
2. Метод спряжених градієнтів з відновленням ……………………………. 176
3. Модифікації методів спряжених градієнтів ……………………………….. 181
4. Метод спряжених градієнтів для мінімізації квадратичних функцій ……… 184
5. Реалізовна модифікація методу зі змінною метрикою …………..………. 185
3.5. Методи спряжених напрямків ……………………………………….. 186
1. Метод спряжених напрямків із відновленням матриці ……........................ 186
2. Метод спряжених напрямків без відновлення матриці ………………… 188
3. Мінімізація квадратичних функцій за допомогою методу спряжених напрямків ……………………………….......................................... 189
4. Модифікований метод спряжених напрямків, який не потребує обчислення похідних …………………………………..………………………….. 194
3.6. Методи псевдообернених операторів ………..……………………… 196
1. Основний метод ……....................................................................................... 196
2. Стійке псевдообернення погано обумовлених матриць ………………... 200
3.7. Стохастичні квазіградієнтні методи…………………………….…… 203
1. Загальний стохастичний квазіградієнтний метод для детермінованих задач ... 203
2. Загальний стохастичний квазіградієнтний метод для стохастичних задач …. 205
3. Стохастичний квазіградієнтний метод з процедурою переривання ………… 206
4. Стохастичний квазіградієнтний метод з постійним кроковим множником …. 208
3.8. Метод локальних варіацій ……………………........................... 209
3.9. Методи випадкового пошуку ………………………………………….. 210
1. Метод випадкового пошуку в опуклих задачах мінімізації ………………… 210
2. Адаптивний метод випадкового пошуку ……………………………… 212
Запитання та вправи для самостійної роботи до розділу 3 ………………... 213

***Розділ 4.* Методи оптимізації недиференційованих функцій і відшукання сідлових точок в задачах без обмежень ………………………...………… 215**

4.1. Означення узагальненого градієнта (субградієнта). Приклади обчислення субградієнтів ………………………..……………………….. 215
4.2. Опуклі функції та їх властивості ……………………………………. 219
4.3. Методи узагальненого градієнтного спуску .. …..………………………. 222
1. Метод зі сталим кроковим множником ..………………………….…… 222
2. Основний метод ……………………..…………………………….. 223
3. Модифікація основного методу …..……………………………….. 228
4. Перший метод із спеціальним вибором крокового множника ………… 229
5. Другий метод із спеціальним вибором крокового множника …………… 230
6. Метод, що використовує апріорне значення мінімуму функції ………… 232
7. Метод, що стійкий до похибок обчислень ……………………………….. 233
8. Багатокроковий метод узагальненого градієнтного спуску …………….. 234
9. $\varepsilon$-субградієнтний метод ……..…………………………………… 235
4.4. Методи градієнтного типу з розтягненням простору …........................ 237

1. Метод градієнтного типу з розтягненням простору в напрямку майже градієнта ……………………………..……………………………… 238
2. Методи градієнтного типу з розтягненням простору в напрямку різниці двох послідовних майже градієнтів ( $r(\alpha)$ -алгоритм) ……………..… 241
4.5. Методи локального випадкового пошуку ……………………….….. 243
1. Метод локального випадкового пошуку з парною пробою ……..…..… 243
2. Метод локального випадкового пошуку з поверненням при невдалому кроці ………………............................................................................... 243
3. Метод локального випадкового пошуку з лінійною екстраполяцією …. 244
4. Метод випадкового пошуку по найкращій пробі з накопиченням …….. 244
5. Метод статистичного градієнта ……………………………………… 245
4.6. Квазіградієнтні методи ……………………………………..……… 246
1. Квазіградієнтний метод мінімізації слабоопуклої донизу функції ……… 246
2. Стохастичний квазіградієнтний метод мінімізації слабоопуклої функції .. 247
4.7. $\varepsilon-$квазіградієнтні методи ………………………………………… 248
1. $\varepsilon-$квазіградієнтний метод мінімізації опуклих функцій ……………… 248
2. $\varepsilon-$квазіградієнтний метод мінімізації слабоопуклих функцій ………….. 250
4.8. Методи узагальнених майже градієнтів для мінімізації функцій, які задовольняють локальну умову Ліпшиця ……………………………….. 251
1. Детермінований випадок ……………………………………………….. 251
2. Стохастичний випадок…………………………………………………... 252
4.9. Метод усереднення напрямків спуску для мінімізації функцій, що задовольняють умову Ліпшиця ……………………………………….. 254
4.10. Методи послідовних наближень для розв'язування дискретних мінімаксних задач …………………......................................................... 255
1. Перший метод послідовних наближень ……………………………… 255
2. Модифікація першого методу послідовних наближень ……………….. 258
3. Другий метод послідовних наближень ………………………………… 260
4. Модифікація другого методу послідовних наближень ………………… 262
4.11. Методи Ерроу–Гурвиця розв'язування неперервних мінімаксних задач 263
1. Детермінований метод Ерроу–Гурвиця ……………………………… 263
2. Стохастичний метод Ерроу–Гурвиця ………………………………… 266
4.12. Квазіградієнтні методи розв'язування дискретних мінімаксних задач стохастичного програмування …………………………………………. 267
1. Мінімізація функції $E_\omega \max_{i\in\Im} \varphi_i(x,\omega)$ ……………………………. 267
2. Мінімізація функції $\max_{i\in\Im} E_\omega \phi_i(x,\omega)$ …………………………………….. 269
Запитання та вправи для самостійної роботи до розділу 4 ………………… 270

***Розділ 5.*** **Методи розв'язування задач лінійного програмування ………** 273

5.1. Приклади задач лінійного програмування ……………………………... 273
5.2. Графічний спосіб розв'язування задач лінійного програмування ……… 275
1. Задача ЛП з максимізацією цільової функції …………………………….. 275
2. Задача ЛП з мінімізацією цільової функції ……………………………….. 279
3. Додаткові змінні в задачах лінійного програмування …………………….. 281
4. Особливі випадки розв'язування задач лінійного програмування ……….. 283

5.3. Графічний аналіз чутливості …………………………………… 285
1. Аналіз чутливості коефіцієнтів цільової функції …………………… 285
2. Аналіз чутливості коефіцієнтів функцій обмежень. Вартість ресурсів …… 288
5.4. Загальна задача лінійного програмування ……………………… 292
5.5. Стандартна задача лінійного програмування. Базис, базисний розв'язок 293
5.6. Канонічна задача лінійного програмування. Метод Жордана–Гауса перебору вершин допустимої області ……………………………….. 297
1. Канонічна задача лінійного програмування ……………………………... 297
2. Метод Жордана–Гауса перебору вершин допустимої області ……......... 299
3. Табличний симплекс-метод ……………………………..…………… 300
5.7. Симплекс-метод та його варіанти ………………………..………… 304
1. Розв'язування СЗЛП з $n$ змінними графічним способом в просторі $R^2$ … 304
2. Симплекс-метод розв'язування невиродженої стандартної задачі лінійного програмування ……………………………………………………... 307
3. Методи пошуку початкового базису ……………………………… 310
4. Симплекс-метод розв'язування виродженої стандартної задачі лінійного програмування ………………………………………………… 319
5. Модифікований симплекс-метод ……………………………… 320
5.8. Двоїстий симплекс-метод ……………………………………… 323
1. Основний метод ……………………………………………………… 323
2. Методи пошуку початкового опорного розв'язку спряженої задачі ……. 326
2.1. Метод пошуку опорного розв'язку спряженої задачі по відомому допустимому розв'язку ………………………………………… 326
2.2. Метод пошуку оптимального розв'язку спряженої задачі без попереднього обчислення допустимого розв'язку …………… 329
3. Правило вибору вектора $a^s$, який вводиться в базис, що гарантує від зациклення у виродженому випадку ……………………………….. 330
5.9. Методи послідовного скорочення нев'язок …..………………… 335
1. Метод послідовного скорочення нев'язок ………………………… 335
2. Метод двосторонніх оцінок ……………………………………….. 337
5.10. Методи блочного програмування …..……………………………. 339
1. Метод розкладання Данцига–Вульфа ……………………………….. 340
1.1. Випадок обмеженості множини $X^1$ …………………………... 340
1.2. Випадок необмеженості множини $X^1$ ……………………………. 343
2. Метод розкладання розв'язування задач лінійного програмування з блочно-діагональною матрицею ……………………………………….. 344
3. Метод, який використовує узагальнений градієнтний спуск ……………. 349
5.11. Модифікований симплекс-метод для розв'язування задачі лінійного програмування з двосторонніми обмеженнями …………………...…… 350
1. Основний метод ……..……………………………………………… 351
2. Правило вибору індексу $r$ (який визначає вектор $a^{i_r}$, що виводиться із базису) для запобігання зациклювання ……………………………….. 355
5.12. Модифікований симплекс-метод для розв'язування загальної задачі лінійного програмування ……………........................................................ 356
5.13. Ітераційні методи ……………....................................................................... 362
1. Ітераційний метод Петшиковського ……………………………………… 362

2. Ітераційний метод, який використовує модифіковану функцію Лагранжа .... 364
3. Ітераційний метод Федоренка .................................. 365
4. Алгоритм «Заєць» розв'язування прямої і двоїстої задач лінійного програмування ................................ 369
5. Ітераційний метод, який використовує модифіковану функцію Лагранжа для розв'язання двоїстої пари задач лінійного програмування ................ 372
5.14. Методи параметричного програмування ................................ 375
1. Випадок наявності параметра в цільовій функції ................ 375
2. Випадок наявності параметра в правих частинах обмежень ................ 379
Запитання та вправи для самостійної роботи до розділу 5 ................ 383

***Розділ 6.* Методи розв'язування задач нелінійного і стохастичного програмування ................................** 387

6.1. Методи проекції градієнта ................................ 387
1. Загальний метод ................................ 387
2. Метод проекції градієнта для мінімізації функції при лінійних обмеженнях 389
3. Гібридний метод проекції градієнта для мінімізації функції при нелінійних обмеженнях ................................ 394
4. Метод проекції градієнта за наявності збурень ................ 396
6.2. Загальний метод штрафних функцій ................................ 398
6.3. Методи зовнішніх штрафних функцій ................................ 401
1. Задачі з обмеженнями у вигляді нерівностей ................ 401
2. Задачі з обмеженнями у вигляді рівностей ................ 404
3. Модифікований метод з процедурою переривання ................ 405
4. Метод зовнішніх штрафних функцій для мінімізації недиференційованих функцій ................................ 408
6.4. Методи внутрішніх штрафних функцій ................................ 410
1. Загальна схема ................................ 410
2. Реалізовна схема з процедурою переривання ................ 412
3. Методи внутрішньої точки із застосуванням $Q$-функцій ................ 413
6.5. Комбіновані методи штрафних функцій ................................ 419
6.6. Методи можливих напрямків ................................ 420
1. Методи можливих напрямків розв'язування задач мінімізації з обмеженнями типу нерівностей ................................ 420
2. Методи можливих напрямків розв'язування задач мінімізації з обмеженнями змішаного типу ................................ 432
3. Метод можливих напрямків із квадратичним пошуком ................ 437
4. Модифікований метод можливих напрямків Зойтендейка ................ 439
5. Аналог методу можливих напрямків у задачах мінімізації майже диференційованих функцій ................................ 441
6. Стохастичний аналог методу можливих напрямків ................ 443
7. Методи можливих напрямків для відшукання точок локальних мінімумів неопуклої функції на неопуклій множині ................ 445
6.7. Метод центрів ................................ 447
1. Основний варіант ................................ 448
2. Модифікований метод центрів ................................ 448

3. Реалізація модифікованого методу центрів з використанням методу золотого перерізу …………………………………………………… 450
4. Реалізація модифікованого методу центрів з використанням функцій переривання …………………………………………………………. 451
6.8. Методи типу Ньютона ……………………………………………... 452
1. Метод типу Ньютона з регулюванням кроку ………………………. 452
2. Метод типу Ньютона при наявності збурень ………………………. 454
3. Квазіньютоновські методи ……………………………………………. 455
6.9. Методи лінеаризації ………………………………………………….. 458
1. Обмеження типу нерівностей ………………………………………. 458
2. Обмеження типу рівностей ………………………........................ 461
3. Обмеження змішаного типу …………………………………………. 463
4. Конструктивний метод лінеаризації ……………………………………. 464
5. Аналог методу лінеаризації в детермінованих задачах мінімізації майже диференційованих функцій ………………………………………. 465
6. Аналог методу лінеаризації в стохастичних задачах мінімізації майже диференційованих функцій ………………………………………….. 467
7. Стохастичний метод лінеаризації ………………………………………. 468
6.10. Методи відсікання ………………………………………………….. 469
1. Лінійний випадок ……………………………………………………… 469
2. Загальний випадок ……………………………………………………… 471
3. Метод відсікання з розтягненням простору для розв’язування задач опуклого програмування ………………………………………………….. 472
6.11.Методи, які використовують функцію Лагранжа ………………….. 476
1. Аналітичний метод ……………………………………………………… 476
2. Градієнтний метод для задач із обмеженнями типу нерівностей ………… 481
3. Градієнтний метод для задач з обмеженнями типу рівностей ………… 483
4. Метод квадратичної апроксимації для задач з обмеженнями типу рівностей 484
5. Двоїстий метод для задач з обмеженнями типу рівностей ………………. 485
6. Метод Ньютона для задач з обмеженнями типу рівностей …..………….. 485
6.12. Методи, які використовують модифіковані функції Лагранжа ……….. 487
1. Градієнтний метод ……………………………………………………… 487
2. Метод, який використовує штрафні функції експоненціального типу ….. 489
6.13. Методи навантаженого функціоналу ……………………………… 491
1. Загальний випадок ……………………………………………………… 492
2. Опуклий випадок ………………………………………………………… 493
3. Наближена схема ……..………………………..………………………… 495
6.14. Методи штрафних оцінок ……………………………………………… 497
1. Детермінований випадок ……………………………………………… 497
2. Стохастичний випадок ……………………..…………………………….. 499
3. Метод штрафних оцінок для задач опуклого програмування …………….. 501
6.15. Методи проектування узагальненого градієнта ………………………… 503
1. Основний метод …………………………………………………..……… 503
2. Багатокроковий метод узагальненого градієнтного спуску …………….. 504
3. Методи проекції узагальненого градієнта для розв'язування граничних екстремальних задач ……………………………………..……….. 506
6.16. Методи умовного градієнта ……………………………………... 507

1. Реалізований метод умовного градієнта ………………………………… 507
2. Алгоритм Франка–Вульфа ……………………………………………. 509
3. Прискорений алгоритм Франка–Вульфа ………………………………… 509
6.17. Метод спряжених градієнтів ……………………………………………. 510
1. Метод спряжених градієнтів ……………………………………………… 510
2. Стохастичний аналог методу спряжених градієнтів ………………… 512
6.18. Методи покоординатного спуску ………………………………………... 514
1. Детермінований покоординатний спуск ………………………………… 514
2. Випадковий покоординатний спуск ……………………………………… 516
6.19. Стохастичні квазіградієнтні методи ……………………………………. 517
1. Приклади обчислення стохастичних квазіградієнтів ……………….…… 518
2. Метод проектування стохастичних квазіградієнтів …………………….… 520
3. Стохастичний метод зменшення нев'язок у детермінованих задачах …….. 523
4. Метод зменшення нев'язок у задачах стохастичного програмування ……. 526
5. Гібридний стохастичний метод ……………………………………….…. 527
6.20. Комбінований метод стохастичних градієнтів і штрафних функцій … 529
6.21. Методи усереднення напрямків спуску ………………………………... 531
1. Детермінована задача ………………………………………………… 531
2. Стохастична задача ………………………………………………….… 533
6.22. Прямий метод розв'язування задач стохастичного програмування ….. 535
6.23. Метод випадкового пошуку в опуклих задачах мінімізації …………… 537
6.24. Методи розв'язання задач оптимізації з нескінченним числом обмежень 538
1. Загальний метод ………………………………………………………… 539
2. Послаблений метод ……………………………………………………….. 540
6.25. Методи розв'язування задач квадратичного програмування ………… 541
1. Метод спряжених градієнтів для мінімізації квадратичної функції на підпросторі ……………………………………………………………… 541
2. Метод спряжених градієнтів для загальної задачі квадратичного програмування …………………………………………………………….. 544
3. Метод спряжених градієнтів для задачі квадратичного програмування з простими обмеженнями ……………………………………………………….. 547
4. Модифікація методу спряжених напрямків для розв'язування задач квадратичного програмування великої розмірності ………………… 548
5. Стійкий метод для розв'язування задач квадратичного програмування ….. 551
Запитання та вправи для самостійної роботи до розділу 6 ………………… 556

***Розділ 7.* Спеціальні методи розв'язування мінімаксних задач і відшукання сідлових точок ………………………………………………...** 560

7.1. Методи послідовних наближень розв'язування дискретних мінімаксних задач ……………………………………………………………………… 560
1. Мінімаксна задача з обмеженнями простої структури ………………….. 560
2. Перший метод послідовних наближень розв'язування мінімаксної задачі з обмеженнями типу нерівностей ……………………………………… 561
3. Другий метод послідовних наближень розв'язування мінімаксної задачі з обмеженнями типу нерівностей …………………………………………….. 564
4. Модифікація другого методу послідовних наближень ………………… 565

7.2. Узагальнений безпараметричний метод зовнішньої точки розв'язування дискретних мінімаксних задач …… 566
1. Основний метод …… 566
2. Модифікація основного методу …… 567
7.3. Сітковий метод послідовних наближень розв'язування неперервних мінімаксних задач …… 568
7.4. Методи стохастичного квазіградієнта в задачі пошуку максиміну …… 569
1. Основний метод …… 570
2. Модифікація основного методу …… 572
3. Опуклий випадок …… 574
7.5. Квазіградієнтні методи розв'язування неперервних мінімаксних задач стохастичного програмування …… 575
1. Стохастичний квазіградієнтний метод …… 576
2. Модифікований стохастичний квазіградієнтний метод …… 577
7.6. Градієнтні методи відшукання сідлових точок …… 578
1. Основний метод …… 579
2. Градієнтний метод відшукання сідлових точок із постійним кроковим множником …… 582
3. Узагальнений градієнтний метод …… 583
Запитання та вправи для самостійної роботи до розділу 7 …… 584

***Розділ 8.* Методи оптимізації в нескінченновимірних просторах** …… 586

8.1. Диференціали, субдиференціали та їх властивості …… 586
8.2. Теореми про існування екстремумів. Необхідні умови екстремумів …… 593
8.3. Застосування методів оптимізації в нескінченновимірних просторах 601
1. Проблема мінімізації квадратичних функціоналів та варіаційні рівняння … 601
2. Багатокритеріальні задачі для варіаційних рівнянь …… 602
8.4. Рівновага за Штакельбергом …… 607
8.5. Наближені методи розв'язування екстремальних задач …… 608
Запитання та вправи для самостійної роботи до розділу 8 …… 616

Література …… 619

# ПОЗНАЧЕННЯ І СИМВОЛИ

$R-$ множина дійсних чисел.

$\alpha,\ \beta,\ \gamma,\ \delta,\ \varepsilon,\ \lambda,\ \mu,\ \nu,\ \rho,\ \sigma,\ \tau-$ дійсні числа.

$i,\ j,\ k,\ l,\ m,\ n,\ s-$ цілі числа.

$R^n$ – $n$-вимірний евклідовий простір, тобто множина векторів $x=(x_1,x_2,\ldots,x_n)^T$, всі компоненти яких $x_1,x_2,\ldots,x_n$ є дійсними числами.

$a,b,c,d,x,y,z,\ldots$ і $a^i,b^i,a^k,x^k,\ldots-$ вектори, тобто $a\in R^n$, $a^i\in R^k,\ldots$

$A,\ B,\ C,\ H-$ матриці.

$a_{ij}$ – елементи $i$-го рядка та $j$-го стовпця матриці $A$.

$\Im,\ K,\ L-$ множина індексів (цілих чисел).

$(x,y)-$ скалярний добуток векторів $x$ та $y$, тобто $(x,y)=\sum_{i=1}^{n}x_i y_i$.

$\|x\|-$ евклідова норма вектора $x$, тобто $\|x\|=(x,x)^{1/2}$.

$f_i$ – скалярна функція.

$f_i : X\to Y$ – функція, яка визначена на множині $X$ і приймає значення із множини $Y$.

$f : X\to R^n$ – $n$-вимірна вектор-функція $f=(f_1,f_2,\ldots,f_n)$.

$A^T$ – матриця, транспонована до матриці $A$.

$A^{-1}$ – матриця, обернена до матриці $A$.

$\nabla f_i(x)-$ градієнт скалярної функції $f_i$ в точці $x$, тобто:

$$\nabla f_i(x)=\left(\frac{\partial f_i(x)}{\partial x_1},\ldots,\frac{\partial f_i(x)}{\partial x_n}\right)^T.$$

$\nabla f(x)-$ матриця Якобі для вектор-функції $f$ в точці $x$, тобто елемент $i$-го рядка та $j$-го стовпця даної матриці дорівнює частинній похідній $\partial f_i(x)/\partial x_j$.

$\nabla^2_{xx} f_i(x)-$ матриця Гессе для скалярної функції $f_i$ в точці $x$, тобто елемент $k$-го рядка та $j$-го стовпця даної матриці дорівнює $\partial^2 f_i(x)/\partial x_k\partial x_j$.

$\hat{\nabla} f_i(x)-$ узагальнений градієнт функції $f_i$ в точці $x$.

$\tilde{\nabla} f_i(x)-$ квазіградієнт функції $f_i$ в точці $x$.

$G\triangleq Q-$ $G$ дорівнює $Q$ за визначенням ($G$ еквівалентно $Q$).

$G\supset Q-$ $G$ містить $Q$.

$G\subset Q-$ $G$ міститься в $Q$.

$G\bigcup Q-$ об’єднання $G$ і $Q$.

$G\bigcap Q-$ перетин $G$ і $Q$.

$G \times Q -$ декартовий добуток $G$ і $Q$.

$\{x | S\} -$ множина всіх $x$, для яких виконується твердження $S$.

$x \neq y -$ $x$ не дорівнює $y$.

$x \in Q -$ $x$ належить $Q$.

$x \notin Q -$ $x$ не належить $Q$.

$\varnothing -$ порожня множина.

$I -$ одинична матриця.

$(\alpha; \beta) -$ відкритий інтервал $\{x | \alpha < x < \beta\}$.

$(\alpha; \beta] -$ півінтервал $\{x | \alpha < x \leq \beta\}$.

$[\alpha; \beta] -$ замкнутий інтервал $\{x | \alpha \leq x \leq \beta\}$.

$\max\limits_{i \in \Im} -$ максимум по всім $i$, які належать множині $\Im$.

$R_{+}^{n} -$ множина $\{x | x \geq 0, \ x \in R^{n}\}$.

$\operatorname{diam} Y -$ діаметр множини $Y$.

$\operatorname{int} X -$ внутрішність множини $X$.

$\operatorname{tr} A -$ слід матриці $A$.

$\operatorname{rang} A -$ ранг матриці $A$.

$\operatorname{Ent}(t) -$ ціла частина числа $t$.

$R(A) -$ образ матриці $A$.

$[1:n]$ або $\overline{1,n}$ – множина цілих чисел від 1 до $n$ включно.

$\operatorname{co} X$ – опукла оболонка множини $X$.

$\det A -$ визначник матриці $A$.

$O(\alpha) -$ величина порядку $\alpha$.

$o(\alpha) -$ величина нескінченно мала порівняно з $\alpha$.

$(b \vdots A) -$ матриця, яка утворена стовпцем $b$ і стовпцями матриці $A$.

$\overset{\text{м.н.}}{=} -$ дорівнює майже напевно.

$[t]_{+} - [t]_{+} = \max\{0, t\}$.

$(\beta_{ij})_{i=\overline{1,m}}^{j=\overline{1,n}} -$ матриця розмірності $m \times n$, $(i, j)$-м елементом якої є $\beta_{ij}$.

$\{x, y, \ldots, z\} -$ множина, яка складається із елементів $x, y, \ldots, z$.

$\max\{t_1, t_2, \ldots, t_n\} -$ максимальний елемент із множини $\{t_1, t_2, \ldots, t_n\}$.

$P\{$подія $S\} -$ ймовірність події $S$.

$\forall x \in X -$ для всіх елементів $x$ із множини $X$.

$\exists x \in X -$ існує такий елементів $x$ в множині $X$.

$\arg\min\limits_{x \in X} f_i(x) -$ той елемент $x^*$ із множини $X$, на якому досягається найменше значення $f_i(x^*)$ функції $f_i$ на множині $X$, тобто:

$$\forall x \in X \quad f_i(\arg\min_{x \in X} f_i(x)) \le f_i(x) .$$

$\mathrm{Arg}\min_{x \in X} f_i(x) -$ множина всіх елементів $\arg\min_{x \in X} f_i(x)$.

$\arg\max_{x \in X} f_i(x) -$ елемент $x^{**}$ із множини $X$, на якому досягається найбільше значення $f_i(x^{**})$ функції $f_i$ на множині $X$, тобто:

$$\forall x \in X \quad f_i\left(\arg\max_{x \in X} f_i(x)\right) \ge f_i(x) .$$

$X^* -$ множина розв'язків задачі оптимізації.

$E\xi -$ математичне сподівання випадкової величини $\xi$.

$E(\xi / x) -$ умовне математичне сподівання при заданому $x$.

$D\xi$ – дисперсія випадкової величини $\xi$.

$\Omega -$ простір елементарних подій $\omega$.

## Передмова

Методи оптимізації в скінченновимірних просторах, зокрема методи лінійного та нелінійного програмування, методи багатокритеріальної та багатоетапної оптимізації, а також і методи оптимізації в нескінченно вимірних просторах, методи варіаційного числення, оптимального керування та диференціальних ігор, використовуються в різних сферах науки і виробництва й увійшли до навчальних програм підготовки бакалаврів і магістрів різних спеціальностей.

Сучасний математично-комп'ютерний інструментарій оптимізації використовується для підвищення якості та надійності нових виробів, для підвищення продуктивності праці, для мінімізації ризиків, для підвищення ефективності управлінських рішень на кожному етапі від проектування і до завершення кожної комплексної системної програми оптимізації, від оптимізації окремих технологічних процесів до оптимізації загальнодержавних соціально-екологічних і фінансово-економічних процесів на рівнях державного управління.

Широкому впровадженню методів оптимізації та оптимального керування сприяла та обставина, що прикладні оптимізаційні задачі в більшості зводяться до розв'язаних типових задач оптимізації, а зростання потужності оновлюваних комп'ютерних засобів сприяє розширенню сфер успішного застосування методів оптимізації у розв'язанні все складніших задач все більших розмірностей.

Беручи до уваги, що опанування математичними методами оптимізації вимагає високого рівня математичної підготовки, у вступному розділі посібника висвітлено необхідні для теорії оптимізації математичні поняття на прикладах розв'язання елементарних типових задач оптимізації. Для опанування поглибленими знаннями у процесі самостійної роботи та при розв'язанні задач оптимізації великої розмірності на допомогу студентам пропонуються детально розроблені алгоритми, описані за єдиною зручною для практичного використання схемою, що включає:

1) математично чітке формулювання оптимізаційної задачі;
2) перелік умов застосовності вибраного алгоритму;
3) опис основної ідеї, реалізованої в алгоритмі;
4) деталізований опис кожного кроку реалізації алгоритму;
5) теореми теоретичного обґрунтування алгоритму;
6) рекомендації щодо практичного використання алгоритму при розв'язанні практичних задач.

Автори сподіваються, що запропонований зручний для практичного використання уніфікований опис кожного алгоритму сприятиме підвищенню ефективності самостійної роботи студентів в опануванні методами оптимізації та в розробках модифікованих математично-комп'ютерних алгоритмів для розв'язання задач оптимізації великої розмірності. Наявність ґрунтовних бібліографічних посилань на першоджерела робить даний посібник зручним путівником у області конструктивної теорії оптимізації також і для фахівців з розробки математично-комп’ютерних методів і алгоритмів для розв’язування прикладних задач оптимізації великої розмірності.

У *першому* розділі описано 9 алгоритмів багатокритеріальної оптимізації, у *другому* – 18 алгоритмів одновимірної оптимізації, в *третьому* – понад 30 алгоритмів безумовної оптимізації диференційованих функцій, у *четвертому* розділі описано близько 35 алгоритмів безумовної оптимізації недиференційованих функцій, у *п’ятому* розділі міститься 30 алгоритмів розв’язування задач лінійного програмування, в *шостому* – 80 алгоритмів розв’язування задач нелінійного і стохастичного програмування, в *сьомому* – біля 20 алгоритмів розв’язування мінімаксних задач і задач відшукання сідлових точок, у *восьмому* розділі – 18 алгоритмів оптимізації в нескінченновимірних просторах. Всі алгоритми супроводжуються теоретичними обґрунтуваннями з відповідними теоремами збіжності, що робить даний посібник корисним також і для фахівців з теорії оптимізації. Більшість алгоритмів апробовано на тестових прикладах. Висвітлені в даному посібнику задачі, методи й алгоритми оптимізації допоможуть студентам і аспірантам опанувати поглибленими знаннями та навичками побудови математично-комп’ютерних алгоритмів для розв’язання прикладних задач оптимізації великої розмірності, а також будуть корисними і для фахівців, які використовують математично-комп’ютерні засоби оптимізації для розв’язання прикладних задач.

Автори виражають глибоку подяку співробітникам кафедри системного аналізу та теорії прийняття рішень Київського національного університету імені Тараса Шевченка (особлива подяка аспіранту Зіньку Т.П.) і співробітникам кафедри математичної фізики Національного технічного університету «Київський політехнічний інститут» за цінні пропозиції та допомогу при підготовці даного посібника до опублікування.

**В с т у п**

# ТИПОВІ ЗАДАЧІ ОПТИМІЗАЦІЇ. УМОВИ ОПТИМАЛЬНОСТІ І АЛГОРИТМИ ОПТИМІЗАЦІЇ

## 0.1. Приклади типових задач оптимізації

З а д а ч а  1. Для заданої функції $f: R \to R$ знайти число $x^*$, яке задовольняє нерівності $f(x^*) \le f(x)$ для всіх $x \in R$.

З а д а ч а  2. Для заданої функції $f: R^n \to R$ знайти вектор $x^* = (x_1^*, x_2^*, ..., x_n^*)^T \in R^n$, який задовольняє нерівності $f(x^*) \le f(x)$ для всіх $x \in R^n$.

З а д а ч а  3. Для заданої множини $X$ і для заданої функції $f: X \to R$ знайти у множині $X$ елемент $x^*$, який задовольняє нерівності $f(x^*) \le f(x)$ для будь-якого елемента $x$ із множини $X$.

**П р и к л а д  1.** Знайти розв'язок $x^*$ задачі 1 для функції $f(x) = (x-2)^2$, $x \in R$.

Покажемо, що у прикладі 1 $x^* = 2$. Дійсно, із рівності $f(2) = (2-2)^2 = 0$ і справедливої для всіх $x \ne 2$ нерівності $f(x) = (x-2)^2 > 0$ випливає, що число 2 задовольняє нерівності $f(2) \le f(x)$ при будь-якому $x$. Це означає, що розв'язком задачі 1 для функції $f(x) = (x-2)^2$ є число $x^* = 2$.

**П р и к л а д  2.** Знайти розв'язок задачі 1 для функції $f(x) = x+5$, $x \in R$.

Очевидно, для функції $f(x) = x+5$ не існує такого числа $x^*$, щоб для всіх чисел $x$ виконувалась нерівність $f(x^*) \le f(x)$. Це означає, що задача 1 не має розв'язку для функції $f(x) = x+5$.

**П р и к л а д  3.** Знайти розв'язок задачі 3 для функції $f(x) = x+5$ і для множини $X = [1;5]$.

Із рівності $f(1) = 1+5 = 6$ та справедливої для будь-якого числа $x > 1$ нерівності $f(x) = x+5 > 6$ випливає, що нерівність $f(1) \le f(x)$ виконується для всіх чисел $x \in [1;5]$. Це означає, що число $x^* = 1$ є розв'язком задачі 3 для функції $f(x) = x+5$ і для множини $X = [1;5]$.

**П р и к л а д  4.** Знайти розв'язок задачі 3 для функції $f(x) = x+5$ і для множини $X = (2;5]$.

Очевидно для кожного числа $\overline{x}$ із множини $X=(2;5]$ є число $x=(\overline{x}+2)/2$, яке також належить множині $X=(2;5]$, але не задовольняє нерівність $f(\overline{x})\le f(x)$. Це означає, що у множині $X=(2;5]$ немає такого числа $x^*\in(2;5]$, яке задовольняє нерівність $f(x^*)\le f(x)$ для всіх $x\in X=(2;5]$. Отже, задача 3 не має розв'язку для функції $f(x)=x+5$ і для множини $X=(2;5]$.

У теорії оптимізації використовують наступні терміни та символьні позначення.

*Оптимальний розв'язок* $x^*$ задачі 1 та задачі 2 позначають символом $\arg\min\limits_x f(x)$ і називають ***мінімізатором функції*** $f$. Отже,

$$\text{для } f(x)=(x-2)^2 \text{ маємо } \arg\min_x f(x)=2,$$

$$\text{для } f(x)=x \text{ числа } \arg\min_x f(x) \text{ не існує,}$$

$$\text{а числом } \arg\min_x (x-x) \text{ є будь-яке число } x\in R.$$

Множину всіх оптимальних розв'язків $x^*$ позначають символом $Arg\min\limits_x f(x)$. Отже,

$$\text{для } f(x)=(x-2)^2 \text{ маємо } Arg\min_x f(x)=\{2\},$$

$$\text{для } f(x)=x+2 \text{ маємо } Arg\min_x f(x)=\varnothing,$$

а множина $Arg\min\limits_x (x-x)$ є інтервалом $(-\infty;\infty)$, $Arg\min\limits_x (x-x)=R$.

Множина $Arg\min\limits_x ((x-2)^2\cdot(x-1)^2)$ складається з двох елементів 2 і 1, тобто $Arg\min\limits_x ((x-2)^2\cdot(x-1)^2)=\{2,1\}$. Дійсно, для будь-якого $x$ маємо нерівність $f(x)\ge 0$ і тому значення $f(2)=(2-2)^2\cdot(2-1)^2=0$ і значення $f(1)=(1-2)^2\cdot(1-1)^2=0$ є найменшими значеннями функції $f(x)=(x-2)^2\cdot(x-1)^2$, а для будь-якого іншого числа $x$ маємо $f(x)>0$.

*Оптимальний розв'язок* $x^*$ задачі 3 позначають символом $\arg\min\limits_{x\in X} f(x)$ і називають ***мінімізатором*** функції $f$ на множині $X$ або ***точкою мінімуму*** функції $f$ на множині $X$. Множину всіх мінімізаторів позначають символом $Arg\min\limits_{x\in X} f(x)$. Отже, маємо:

$$\arg\min_{x\in[0;5]}(x-2)^2=2,\ \arg\min_{x\in[0;2]}(x-2)^2=2,\ \arg\min_{x\in[0;1]}(x-2)^2=1,$$

$$Arg\min_{x\in[0;2]}(x-2)^2=\{2\},\ Arg\min_{x\in[0;2)}(x-2)^2=\varnothing,\ Arg\min_{x\in[0;2)}(x-x)=[0;2),$$

$$Arg\min_{x\in[1;2)}((x-2)^2\cdot(x-1)^2)=\{1\},\ Arg\min_{x\in(1;2]}((x-2)^2\cdot(x-1)^2)=\{2\},$$

$$Arg\min_{x\in[1;2]}((x-2)^2\cdot(x-1)^2)=\{1;2\},\ Arg\min_{x\in(1;2)}((x-2)^2\cdot(x-1)^2)=\varnothing.$$

Значення $\bar{x}$, яке мінімізує функцію $f$ у своєму околі, називають ***точкою локального мінімуму*** функції $f$ або *локальним* розв'язком задачі оптимізації, а мінімізатор $x^* = \arg\min_x f(x)$ називають ***точкою глобальною мінімуму*** функції $f$. Іншими словами, точка $\bar{x}$, яка задовольняє нерівності $f(\bar{x}) \le f(x)$ для $x$ із деякого околу $\bar{x}$, є точкою локального мінімуму функції $f$.

Аналогічно значення $\bar{x} \in X$ називають ***локальним розв'язком*** задачі 3, або ***точкою локального мінімуму функції $f$ на множині $X$***, якщо для всіх значень $x \in X$ із околу точки $\bar{x}$ виконуються нерівності $f(\bar{x}) \le f(x)$.

Наприклад, точка $x = 5$ є точкою глобального мінімуму функції $f(x) = -(x-1)^2$ на множині $X = [0;5]$, а точка $x = 0$ є точкою локального мінімуму функції $f(x) = -(x-1)^2$ на множині $X = [0;5]$. Для функції $f(x) = \frac{1}{3}x^3 - \frac{3}{2}x^2 + 2x + 1$ точки глобального мінімуму не існує, проте є точка $\bar{x} = 2$ її локального мінімуму.

У деяких задачах оптимізації потрібно знайти не мінімізатор $x^*$, а максимізатор $x^{**}$ функції $f$ на множині $X$, який для будь-якого $x \in X$ задовольняє нерівність $f(x^{**}) \ge f(x)$. Наприклад, максимізатором функції $f(x) = -(x-1)^2$ є значення $x^{**}$ =1.

Максимізатор функції $f$ на множині $X$ позначають символом $\arg\max_{x \in X} f(x)$, а множину всіх максимізаторів функції $f$ на множині $X$ позначають символом $Arg\max_{x \in X} f(x)$. Отже, маємо:

$$\arg\max_{x \in [0;5]} (1 - (x-1)^2) = 1,$$

$$Arg\max_{x \in [0;5]} (1 - (x-1)^2) = \{1\},$$

$$Arg\max_{x \in [0;5]} 1 = [0;5].$$

Очевидно, для будь-якої функції $f$ її максимізатор $x^{**} = \arg\max_{x \in X} f(x)$ на множині $X$ є мінімізатором $x^* = \arg\min_{x \in X}(-f(x))$ функції $(-f)$ на множині $X$, тобто $\arg\max_{x \in X} f(x) = \arg\min_{x \in X}(-f(x))$.

Серед багатьох виділимо наступні *типові задачі оптимізації*.

**Задача лінійного програмування**. ***Задачею лінійного програмування*** називають задачу знаходження вектора $x = (x_1, x_2, ..., x_n)^T$, який максимізує лінійну функцію

$$f(x) = c_1 x_1 + c_2 x_2 + ... + c_n x_n \equiv (c, x)$$

на множині $X = \{x \mid Ax = b, Bx \le d\}$, заданій матрицями $A, B$ і векторами $b, d$.

**Задача квадратичного програмування**. ***Задачею квадратичного програмування*** називають задачу знаходження вектора $x = (x_1, x_2, ..., x_n)^T$, який мінімізує квадратичну функцію

$$f(x) = \sum_{j=1}^{n} c_j x_j + \sum_{k=1}^{n} \sum_{j=1}^{n} d_{kj} x_k x_j \equiv (c, x) + (Dx, x)$$

на множині $X = \{x \mid Ax = b, Bx \le d\}$. Матриці $A, B, D$ і вектори $c, b, d$ є заданими.

**Задача нелінійного програмування**. ***Задачею нелінійного програмування*** називають задачу знаходження вектора $x = (x_1, x_2, ..., x_n)^T$, який мінімізує функцію $f : R^n \to R$ на заданій функціями $f_i : R^n \to R$ множині:

$$X = \{x \mid f_i(x) \le 0,\ i \in \overline{1, m},\ f_i(x) = 0,\ i \in \overline{m+1, r}\},$$

де хоча б одна серед функцій $f$, $f_i$, $i \in \overline{1, r}$ є нелінійною функцією.

**Задача опуклого програмування.** ***Задачею опуклого програмування*** називають задачу нелінійного програмування з опуклою функцією $f$, опуклими функціями $f_i$ для $i = \overline{1, r}$.

**Задача динамічного програмування.** Прикладом ***задачі динамічного програмування*** є задача відшукання вектора $x = (x_1, x_2, ..., x_n)^T$, який мінімізує адитивну функцію

$$f(x_1, x_2, \dots, x_n) = \sum_{i=1}^{n} f_i(x_i)$$

або мультиплікативну функцію

$$f(x_1, x_2, \dots, x_n) = \prod_{i=1}^{n} f_i(x_i)$$

на допустимій множині

$$X = \{x \mid \sum_{j=1}^{n} a_j x_j \le b,\ x_j \ge 0,\ j = \overline{1, n}\}.$$

**Задача параметричної оптимізації.** У ***задачі параметричної оптимізації*** потрібно знайти функцію $x^* : P \to X$ мінімізаторів $x^*(p) = \arg\min_{x \in X} f(x, p)$ функції $f(x, p)$ на множині $X$. У більш загальній постановці задачі параметричної оптимізації множина $X$ також може залежати від параметра $p$.

**Задача мінімаксної оптимізації.** У ***задачах мінімаксної оптимізації*** мінімізують функцію $f$, значення якої визначається за формулою

$$f(x) = \max_y \overline{f}(x, y).$$

Отже, розв'язком задачі мінімаксної оптимізації є значення $x^* = \arg\min_x \max_y \overline{f}(x, y)$, яке для всіх значень $x$ задовольняє нерівність

$$\max_y \overline{f}(x^*, y) \le \max_y \overline{f}(x, y).$$

**Задача оптимізації з неповними даними.** ***Задачі з неповними даними*** є задачами параметричної оптимізації, у яких мінімізована функція $f(x, p)$ і (або) допустима множина $X(p)$ залежать від невідомого параметра $p \in P$, який належить відомій множині $P$.

Мінімізатор $x^*_{\max} = \arg\min_{x \in X} \max_{p \in P} f(x, p)$ функції $\max_{p \in P} f(x, p)$ на допустимій множині $X$ називають ***мінімаксним*** (або ***гарантованим***) розв'язком задачі мінімізації функцій з неповними даними. Мінімаксний розв'язок $x^*_{\max}$ називають ***розв'язком песиміста*** на відміну від ***розв'язку оптиміста*** $x^*_{\min}$, де $x^*_{\min} = \arg\min_{x \in X} \min_{p \in P} f(x, p)$.

**Задача стохастичної оптимізації.** У ***задачах стохастичної оптимізації*** значення $f(x, p)$ залежить від випадкової величини $p$. Якщо $p$ приймає значення із допустимої множини $P$ у відповідності із заданою мірою $\mu$, то оптимальний розв'язок $x^*$ часто визначають як мінімізатор $x^* = \arg\min_x f(x)$ математичного сподівання $f(x) = \int_P f(x, p) d\mu(p)$.



В інших постановках задач стохастичної оптимізації оптимальний розв'язок $x^*$ визначають як мінімізатор математичного сподівання $f(x) = \int_P f(x, p) d\mu(p)$ або на множині:

$$X = \{x \mid \int_{P_i} f_i(x, p) d\mu_i(p) \le 0,\ i = \overline{1, m},\ \ x \in X^1\},$$

або на множині:

$$X_\varepsilon = \{x \mid P(x \in X(p)) \ge 1 - \varepsilon\},$$

де $P(A) -$ ймовірність події $A$, а $\varepsilon \in [0;1) -$ задане значення параметра задачі стохастичної оптимізації.

На даний час відомо багато алгоритмів для пошуку оптимального розв'язку $x^*$ також і для тих важливих для практики випадків, коли міра $\mu$ є невідомою, але можна лише обчислювати значення реалізацій випадкової величини $p$.

**Задача багатоетапної оптимізації.** ***Задача багатоетапної оптимізації*** є частинним випадком задачі оптимізації з неповними даними для цільової функції $f(x, p)$, коли вибір певної частини компонент вектора $x$ дозволяється перенести на пізніший термін до того моменту уточнення множини $P$ допустимих значень невідомого параметра $p$.

Наприклад, у ***задачі двоетапної оптимізації*** шуканий вектор $x$ розбивають на дві частини, $x=(x^1,x^2)^T$, і на першому етапі вибирають лише значення $x^1$, а вибір значень $x^2$ відкладають до того часу, коли буде уточнена множина $P$ допустимих значень невідомого параметра $p$. Функцію корисності $f(x,p)$ будують з урахуванням додаткових витрат на уточнення множини невизначеності $P$ та витрат, пов'язаних із перенесенням терміну вибору $x^2$. Якщо в результаті розв'язання задачі двоетапної оптимізації виявиться, що всі додаткові витрати не можуть бути компенсовані за рахунок отримання додаткового прибутку від уточнення множини невизначеності $P$, то в такому випадку вибір оптимальних значень $x^2$ здійснюють на першому етапі разом із вибором оптимальних значень $x^1$.

**Задача багатокритеріальної оптимізації.** У ***задачах багатокритеріальної оптимізації*** замість одного критерію оптимальності $f$ є декілька критеріїв оптимальності $f_i(x),\ i=\overline{1,m}$, які бажано мінімізувати.

Проте тут виникає парадоксальна ситуація, пов'язана із пошукам неіснуючого мінімізатора $x^*$ із допустимої множини $X$, який для всіх $x\in X$ і для всіх критеріїв задовольняв би нерівностям:

$$f_i(x^*)\le f_i(x),\ i\in\overline{1,m}.$$

Очевидно, що такого елементу $x^*$, як правило, не існує, і через це задачу багатокритеріальної оптимізації часто зводять до задачі однокритеріальної оптимізації, наприклад, за допомогою деякого агрегованого критерію $f$. Прикладом такого критерію є зважена сума

$$f(x,c)=c_1f_1(x)+c_2f_2(x)+\ldots+c_mf_m(x)$$

з ваговими коефіцієнтами $\sum_{i=1}^{m}c_i=1,\ c_i>0,\ i=\overline{1,m}$. У цьому випадку на множині $X^*$ мінімізаторів $x^*(c)=\arg\min_{x\in X}f(x,c)$:

$$X^*=\{x\mid x=x^*(c),\ \sum_{i=1}^{m}c_i=1,\ c_i>0,\ i=\overline{1,m}\}$$

остаточно вибирають оптимальне значення $x^*$ за допомогою можливо і неформальних додаткових критеріїв.

**Задача варіаційного числення.** ***Задачею варіаційного числення*** називають задачу відшукання мінімізатора функціоналу $f:X\to R$ на відкритій множині $X$ функцій $x$ із деякого функціонального простору.

Розв'язком задачі варіаційного числення є функція $x^* \in X$, яка для всіх функцій $x \in X$ задовольняє нерівності

$$f(x^*) \le f(x).$$

Прикладом задачі варіаційного числення є задача відшукання на відрізку $t \in [a;b]$ функції $x(t)$, яка мінімізує функціонал $f(x) = \int_a^b w(x'(t), x(t), t)dt$ на множині $X$ неперервно диференційованих функцій з крайовими умовами $x(a) = x^1$, $x(b) = x^2$.

**Задача оптимального керування.** ***Задачею оптимального керування*** називають задачу відшукання функції $u^*$, яка на заданій у функціональному просторі множині допустимих керувань $U$ мінімізує функціонал $f(u,x)$ на множині розв'язків $x$ диференціальних рівнянь $\varphi(x', x, u) = 0$.

Прикладом задачі оптимального керування є відшукання вектор-функції $u^* : [t_0;T] \to U(\cdot)$, яка в кожен момент часу $t \in [t_0;T]$ належить замкнутій множині допустимих керувань $U$ і переводить керовану систему

$$x'(t) = \varphi(x(t), u(t), t)$$

із заданої початкової точки $x(t_0) = x^0$ у задану кінцеву точку $x(T) = x^1$ при мінімальному значенні функціоналу $f(u,x) = \int_{t_0}^T f_0(x(t), u(t), t)dt$.

У випадку функції $f_0(x(t), u(t), t) \equiv 1$ та нефіксованого кінцевого моменту часу $T$ оптимальне керування $u^*$ переводить керовану систему у точку $x(T) = x^1$ за мінімальний час $f(u,x) = T - t_0$. Таке керування називають ***оптимальним за швидкодією***.

При наявності додаткових фазових обмежень $x(t) \in U_x \subset R^n$ задачу оптимального керування називають ***задачею з фазовими обмеженнями.***

**Задача ієрархічно керованої оптимізації.** У ***задачах ієрархічно керованої оптимізації*** приймають участь різні конфліктуючі сторони (гравці) із своїми різними критеріями оптимальності $f_i$. Наприклад, перед гравцем *А* стоїть завдання вибрати елемент $x$ із допустимої множини $X$, який мінімізує значення $f_a(x,y)$, а гравець *В* має завдання вибрати значення $y$ із допустимої множини $Y$, яке мінімізує значення $f_b(x,y)$. Функцію $f_a$ називають *цільовою функцією гравця А*, а функцію $f_b$ називають *цільовою функцією гравця В*. Якщо гравець $B$ вибирає значення $y$ перед вибором гравцем $A$ значення $x$, то кажуть, що «перший хід робить гравець $B$». Аналогічно, якщо гравець $A$ вибирає значення $x$ перед вибором значення $y$ гравцем $B$, то в цьому випадку кажуть, що «перший хід робить гравець $A$». Різні задачі ієрархічно керованої оптимізації характеризуються як різними правилами вибору першого та наступних ходів, так і різною поінформованістю гравців про допустимі множини та про цільові функції інших гравців. Велике різноманіття практичних задач ієрархічно керованої оптимізації є потужним джерелом постановок нових все більш ускладнених математичних задач оптимізації в конфліктних ситуаціях.

## 0.2. Методи повного перебору та ціленаправленого перебору. Алгоритми відтинань та мінорант

### 1. Методи повного перебору

Для знаходження мінімального значення $\min\limits_{x \in [a;b]} f(x)$ функції $f$ на інтервалі $[a;b]$ і для знаходження її мінімізатора $x^* = \arg\min\limits_{x \in [a;b]} f(x)$ методом повного перебору потрібно вибрати достатньо велику кількість точок $x^1, x^2, ..., x^N$, які густо покривають інтервал $[a;b]$, і після обчислення усіх значень $f(x^1), f(x^2), ..., f(x^N)$ знайти серед них мінімальне значення $\min\limits_i f(x^i)$ і знайти мінімізатор $x^{i^*}$, $f(x^{i^*}) = \min\limits_i f(x^i)$.

Отже, в методі повного перебору вибирають кількість точок $N$, обчислюють відстань між ними $h = (b-a)/(N-1)$, обчислюють точки $x^i = a + (i-1)h$, обчислюють значення $f(x^i)$ і знаходять серед цих значень найменше значення $\min\limits_i f(x^i)$. Цей простий метод реалізований у наступному алгоритмі повного перебору.

***Алгоритм повного перебору***

I. Вибрати параметр алгоритму $N$, обчислити $h = (b-a)/(N-1)$ і покласти $x = a$, $y = f(a)$, $\bar{x} = x$, $\bar{y} = y$.

II. Якщо $x \ge b$, то перейти на крок V.

III. Покласти $x = x + h$, $y = f(x)$.

IV. Якщо $y < \bar{y}$, то покласти $\bar{y} = y$, $\bar{x} = x$ і перейти на крок II.

V. Зупинити обчислення.

Обчислені значення $\bar{x}$ і $\bar{y}$ є наближеними значеннями оптимальних розв'язків $x^* = \arg\min\limits_{x \in [a;b]} f(x)$ і $f(x^*)$. Їх похибка по аргументу $|\bar{x} - \arg\min\limits_{x \in [a;b]} f(x)|$ та похибка по функції $|\bar{y} - \min\limits_{x \in [a;b]} f(x)|$ залежать від числа

$N$. Очевидно, для неперервної на відрізку $[a;b]$ функції $f$ ці похибки прямують до нуля із зростанням $N$ до нескінченності.

**2. Методи відтинань та алгоритми медіан і половинного поділу**

Якщо функція $f: R \to R$ квазіопукла, то її мінімізатор $x^* = \arg\min_{x \in X} f(x)$ на множині $X = \{x^1, x^2, ..., x^N\}$ можна знайти за допомогою *методу відтинань* без обчислення значень функції $f$ у всіх точках $x^1, x^2, ..., x^N$. Дійсно, для квазіопуклої функції $f$ справедливе твердження: *«Якщо $f(x^i) \le f(x^k)$, то у випадку $x^i \le x^k$ існує мінімізатор $x^* \le x^k$, а у випадку $x^i \ge x^k$ існує мінімізатор $x^* \ge x^k$».* Отже, для медіани $x^{N/2}$ у випадку $f(x^{N/2}) \le f(x^{N/2+1})$ існує мінімізатор $x^* \le x^{N/2+1}$ у підмножині $\{x^1, x^2, ..., x^{N/2+1}\}$, а у випадку $f(x^{N/2}) \ge f(x^{N/2+1})$ існує мінімізатор $x^* \ge x^{N/2}$ у підмножині $\{x^{N/2}, x^{N/2+1}, ..., x^N\}$ (у випадку двозначної медіани $x^i, x^{i+1}$, ці два значення використовуються аналогічно значенням $x^{N/2}, x^{N/2+1}$ у випадку єдиної медіани $x^{N/2}$). Тобто, в обох випадках за допомогою перевірки лише одної нерівності знаходимо ту половину елементів множини $X$, яка містить мінімізатор і, отже, іншу половину можна вилучити із множини $X$. Продовження такого половинного поділу та відтинання зайвих підмножин приводить до швидкого відшукання мінімізатора $x^*$. Такий *метод відтинань* реалізований у наступному алгоритмі медіан.

***Алгоритм медіан***

I. Упорядкувати усі елементи множини $X_N$ за зростанням $x^1 < x^2 < ... < x^N$. Покласти $k = 1$, $l = N$. Якщо $k = l$, то покласти $i^* = k$ і перейти на крок VIII.

II. Якщо $f(x^k) < f(x^{k+1})$, то покласти $i^* = k$ і перейти на крок VIII.

III. Якщо $f(x^l) < f(x^{l-1})$, то покласти $i^* = l$ і перейти на крок VIII.

IV. Якщо $l - k = 1$, то у випадку $f(x^k) < f(x^l)$ покласти $i^* = k$ і перейти на крок VIII.

V. Якщо $l - k = 1$, то у випадку $f(x^l) \le f(x^k)$ покласти $i^* = l$ і перейти на крок VIII.

VI. Обчислити $i = \mathrm{Ent}((l-k)/2)$ (ціла частина від числа $(l-k)/2$).

VII. Якщо $f(x^i) < f(x^{i+1})$, то покласти $l = i$ (цією дією відтинаємо від множини $\{k, ..., l\}$ її підмножину $\{i+1, ..., l\}$, яка не містить значення

$i^* = \arg\min_{i \in [1;N]} f(x^i))$. Якщо $f(x^i) \geq f(x^{i+1})$, то покласти $k = i+1$. Перейти на крок II.

VIII. Покласти $x^* = x^{i*}$ і зупинити обчислення.

Аналогічний метод відтинань реалізований у наступному алгоритмі половинного поділу для мінімізації квазіопуклої функції $f$ на множині $X = [a;b]$.

***Алгоритм половинного поділу відрізка***

I. Вибрати значення $a$ і $b$ (інтервал $[a;b]$) і покласти $\bar{a} = a$, $\bar{b} = b$. Вибрати значення $\varepsilon > 0$ (точність обчислення наближеного розв'язку $\bar{x}$).

II. Якщо $\bar{b} - \bar{a} \leq \varepsilon$ то перейти на крок IV.

III. Покласти $c^1 = (\bar{a} + \bar{b} - \varepsilon)/2, c^2 = (\bar{a} + \bar{b} + \varepsilon)/2$. Якщо $f(c^1) < f(c^2)$, то покласти $\bar{b} = c^2$. Якщо $f(c^1) \geq f(c^2)$, то покласти $\bar{a} = c^1$. Перейти на крок II.

IV. Покласти $\bar{x} = (\bar{b} + \bar{a})/2$ і зупинити обчислення.

Очевидно, похибка обчисленого за даним алгоритмом наближеного розв'язку $\bar{x}$ не перевищує $\varepsilon/2$, $|\bar{x} - \arg\min_{x \in [a;b]} f(x)| \leq \varepsilon/2$.

Отже, у методах відтинань на $k$-му кроці знаходять підмножину $\breve{X}_k$ допустимої множини $X$, яка не містять оптимального розв'язку $x^*$, і цю підмножину відтинають від множини $X$. З отриманою в результаті «меншою» підмножиною $X \setminus \breve{X}_k$, яка містить оптимальний розв'язок $x^*$, чинять аналогічно й у такому процесі послідовних відтинань отримують все менші підмножини $X \setminus \breve{X}_1 \setminus \breve{X}_2 \setminus ... \setminus \breve{X}_k$, які можуть збігатися до множини із одним елементом, яким, очевидно, є оптимальний розв'язок $x^* = \arg\min_{x \in X} f(x)$.

Оптимальними методами відтинань є методи, які забезпечують отримання оптимального розв'язку із заданою точністю за найменшу кількість відтинань $k$. З огляду на це ефективнішим за метод половинного поділу є метод відтинань з використанням золотого перетину, де відтинання здійснюють за допомогою обчислення лише одного значення $f(c)$ у точці золотого перетину $c$ (замість обчислення двох значень $f(c^1)$ і $f(c^2)$ у алгоритмі половинного поділу відрізка). ***«Золотим перерізом» відрізка*** називають перетин, у якому відношення довжини більшої частини до меншої дорівнює відношенню довжини всього відрізка до довжини його більшої частини. Отже, золотим перетином відрізка $[0;1]$ є відрізки $[0;c]$ і $[c;1]$, які задовольняють співвідношенню $\frac{1-c}{c} = \frac{1}{1-c}$, тобто $c = (3 - \sqrt{5})/2$.

За допомогою алгоритму золотого перетину, який описаний у розділі 2, можна обчислити з будь-якою високою точністю мінімізатор квазіопуклої функції на заданому відрізку. Проте оптимальним за кількістю обчислень функції $f$ для досягнення заданої точності наближеного розв’язку $x^*$ є метод Фібоначі, у якому реалізовано оптимальний поділ відрізка за допомогою послідовності чисел Фібоначі, побудованих за формулами $F_1 = F_2 = 1$, $F_{j+2} = F_{j+1} + F_j$. Алгоритм методу Фібоначі описаний у розділі 2.

**3. Метод мінорант**

Якщо функція $F$ є мінорантою для функції $f$ на множині $X$, тобто, якщо для всіх $x$ із множини $X$ виконується нерівність $F(x) \le f(x)$, то, очевидно, виконується нерівність $\min_{x\in X} f(x) \ge \min_{x\in X} F(x)$. З іншого боку, для будь-якого значення $\bar{x} \in X$ маємо нерівність $\min_{x\in X} f(x) \le f(\bar{x})$. Отже, мінімальне значення $\min_{x\in X} f(x)$ належить інтервалу $[\min_{x\in X} F(x); f(\bar{x})]$. Це означає, що середньо інтервальне значення $\bar{f} = (\min_{x\in X} F(x) + f(\bar{x}))/2$ відрізняється від шуканого мінімального значення $\min_{x\in X} f(x)$ не більше ніж на половину довжини інтервалу $[\min_{x\in X} F(x); f(\bar{x})]$, тобто $|\bar{f} - \min_{x\in X} f(x)| \le (f(\bar{x}) - \min_{x\in X} F(x))/2$, і тому для обчислення наближеного до $\min_{x\in X} f(x)$ значення $\bar{f}$ з точністю $\varepsilon$, тобто для виконання нерівності $|\bar{f} - \min_{x\in X} f(x)| \le \varepsilon$ достатньо знайти таку точку $\bar{x}$ і таку міноранту $F$, які задовольняють нерівності $f(\bar{x}) - \min_{x\in X} F(x) \le 2\varepsilon$. Аналогічно для будь-якої підмножини $X^N = \{x^1, x^2, ..., x^N\} \subset X$ і для будь-якої міноранти $F_N$ для функції $f$ на множині $X$ маємо $\min_{x\in X} f(x) \in [\min_{x\in X} F_N(x); \min_{i\in[1:N]} f(x^i)]$, і тому для обчислення з точністю $\varepsilon$ наближеного до $\min_{x\in X} f(x)$ значення $\tilde{f} = (\min_{x\in X} F_N(x) + \min_{i\in[1:N]} f(x^i))/2$ достатньо знайти таку підмножину $X^N = \{x^1, x^2, ..., x^N\} \subset X$ і таку міноранту $F_N$, щоб задовольнити нерівність

$$\min_{i\in[1:N]} f(x^i) - \min_{x\in X} F_N(x) \le 2\varepsilon .$$

На цій основі у методі мінорант для обчислення мінімального значення $\min_{x\in X} f(x)$ будують таку послідовність $\{x^N\}_{N=1}^{\infty}$ і таку послідовність мінорант $F_N$, щоб із зростанням $N$ різниця $\min_{i\in[1:N]} f(x^i) - \min_{x\in X} F_N(x)$ стрімко прямує до нуля.

Наприклад, у випадку ліпшицевої функції $f$ з відомою константою Ліпшиця $L$ для всіх значень $\bar{x}, \tilde{x} \in X$ виконується нерівність $|f(\bar{x}) - f(\tilde{x})| \le L|\bar{x} - \tilde{x}|$, і тому потрібною мінорантою $F_N$ є функція

$$F_N(x) = \max_{i\in[1:N]}\left(f(x^i) - L|x - x^i|\right),$$

а у випадку ліпшицевої похідної $|f'(\bar{x}) - f'(\tilde{x})| \le L|\bar{x} - \tilde{x}|$, такою мінорантою є функція $F_N(x) = \max_{i\in[1:N]}(f(x^i) + f'(x^i)(x - x^i) - L(x - x^i)^2)$ тощо. Послідовність $\{x^N\}_{N=1}^{\infty}$ обчислюють за ітераційною формулою

$$x^{N+1} = \arg\min_{x\in X} F_N(x).$$

Метод мінорант для обчислення значення $\min_{x\in X} f(x)$ багатоекстремальної функції $f$ реалізований у наступному алгоритмі мінорант.

***Алгоритм мінорант***

I. Вибрати параметр $\varepsilon > 0$ точності обчислення наближеного розв'язку, вибрати початкове значення $x^1 \in X$ і покласти $X^1 = \{x^1\}$, $N = 1$, $\bar{f} = f(x^1)$.

II. Обчислити $x^{N+1} = \arg\min_{x\in X} F_N(x)$ і покласти $\bar{f} = \min\left\{\bar{f}, f(x^{N+1})\right\}$.

III. Якщо виконується нерівність $\bar{f} - F_N(x^{N+1}) \le 2\varepsilon$, то покласти $\bar{f} = (\bar{f} + F_N(x^{N+1}))/2$ і перейти на крок V.

IV. Покласти $X_{N+1} = X_N \cup \{x^{N+1}\}$, $N = N + 1$ і перейти на крок II.

V. Зупинити обчислення.

Похибка отриманого за алгоритмом 1 наближеного значення $\bar{f}$ для шуканої величини $\min_{x\in X} f(x)$ не перевищує $\varepsilon$, тобто $|\bar{f} - \min_{x\in X} f(x)| \le \varepsilon$. Послідовність $\left\{x^k\right\}_{k=1}^{\infty}$ мінімізує функцію $f$, $\lim_{N\to\infty} f(x^N) = \inf_{x\in X} f(x)$, а у випадку компактної множини $X$

$$\lim_{N\to\infty} x^N = x^* = \arg\min_{x\in X} f(x).$$

## 0.3. Необхідна умова мінімуму диференційованої функції і градієнтні алгоритми оптимізації

### 1. Необхідна умова мінімуму

Функцію $f: R \to R$ називають ***диференційованою у точці*** $\bar{x}$, якщо існує границя $\lim\limits_{x \to \bar{x}} \dfrac{f(x) - f(\bar{x})}{x - \bar{x}}$. Цю границю позначають символом $f'(\bar{x})$ (або $\dfrac{df(\bar{x})}{dx}$) і називають похідною функції $f$ у точці $\bar{x}$. Для диференційної у точці $\bar{x}$ функції $f$ має місце важлива асимптотична рівність:

$$f\left(x\right) = f\left(\bar{x}\right) + f'(\bar{x})(x - \bar{x}) + o(x - \bar{x}), \qquad (0.1)$$

де $o(x - \bar{x})$ – функція, що задовольняє рівності $\lim\limits_{x \to \bar{x}} \dfrac{o(x - \bar{x})}{x - \bar{x}} = 0$. Дійсно, беручи до уваги справедливу в околі точки $\bar{x}$ рівність $\lim\limits_{x \to \bar{x}} \dfrac{f(x) - f(\bar{x})}{x - \bar{x}} = f'(\bar{x})$, із (0.1) отримуємо:

$$\lim_{x \to \bar{x}} \frac{o(x - \bar{x})}{x - \bar{x}} = \lim_{x \to \bar{x}} \frac{f(x) - f(\bar{x}) - f'(\bar{x})(x - \bar{x})}{x - \bar{x}} =$$

$$= \lim_{x \to \bar{x}} \frac{f(x) - f(\bar{x})}{x - \bar{x}} - \lim_{x \to \bar{x}} \frac{x - \bar{x}}{x - \bar{x}} f'(\bar{x}) = f'(\bar{x}) - f'(\bar{x}) = 0.$$

***Необхідна умова мінімуму.*** *Якщо значення $x^*$ мінімізує (або максимізує) диференційну у точці $x^*$ функцію $f$, то $f'(x^*) = 0$, тобто значення $x^*$ є розв'язком рівняння*

$$f'(x) = 0. \qquad (0.2)$$

Для доведення рівності $f'(x^*) = 0$ покажемо, що нерівність $f'(x^*) \neq 0$ суперечить оптимальності точки $x^*$. Дійсно, якщо для точки $x^*$, яка мінімізує функцію $f$, мала би місце нерівність $f'(x^*) = C \neq 0$, то точка $x = x^* - \lambda f'(x^*)$ задовольняла б асимптотичній рівності (0.1)

$$f\left(x^* - \lambda f'(x^*)\right) = f\left(x^*\right) + f'(x^*)(-\lambda f'(x^*)) + o(-\lambda f'(x^*)) =$$

$$= f\left(x^*\right) - \lambda C^2 + o(-\lambda C).$$

Із існування границі $\lim\limits_{\lambda \to 0} \left(o(-\lambda C)/(-\lambda C)\right) = 0$ випливає існування числа $\lambda_1$, при якому виконується нерівність $|o(-\lambda_1 C)/(-\lambda_1 C)| < |C|/2$, тобто $|o(-\lambda_1 C)| < |\lambda_1| C^2/2$. Отже, маємо:

$$f\left(x^* - \lambda_1 f'(x^*)\right) \leq f\left(x^*\right) - \lambda_1 C^2 + o(-\lambda_1 C) <$$

$$< f\left(x^*\right) - \lambda_1 C^2 / 2 < f\left(x^*\right).$$

Отримана нерівність $f\left(x^* - \lambda_1 f'(x^*)\right) < f\left(x^*\right)$ суперечила би справедливій для всіх $x$ нерівності $f(x^*) \le f(x)$. Отже, для мінімізатора $x^*$ нерівність $f'(x^*) \ne 0$ є неможливою і цим доведено рівність $f'(x^*) = 0$.

Аналогічно доводиться також і рівність $f'(x^{**}) = 0$ у точці $x^{**}$, яка максимізує функцію $f$ і в якій існує похідна $f'(x^{**})$.

**Приклад 1.** Знайти мінімізатор функції $f(x) = ax + b$.

*Розв'язування.* Для функції $f(x) = ax + b$ маємо $f'(x) = a$. Це означає, що рівняння (0.2) має розв'язок лише за умови $a = 0$. Отже, якщо $a = 0$, то кожна точка $x$ задовольняє рівнянню (0.2) і є мінімізатором функції $f$, тобто $Arg\min_x f(x) = R$. А якщо $a \ne 0$, то рівняння (0.2) не має розв'язків і функція $f(x) = ax + b$ не має найменшого значення, тобто $Arg\min_x f(x) = \varnothing$.

**Приклад 2.** Знайти мінімізатор для функції $f(x) = ax^2 + bx + c$.

*Розв'язування.* Для функції $f(x) = ax^2 + bx + c$ маємо $f'(x) = 2ax + b$. Це означає, що:

- у випадку $a \ne 0$ рівняння (0.2) має єдиний розв'язок $x^* = -b/(2a)$, $\arg\min_x f(x) = -b/(2a)$ якщо $a > 0$, і $Arg\min_x f(x) = \varnothing$ якщо $a < 0$;
- у випадку $a = 0$ і $b = 0$ розв'язком рівняння (0.2) є будь-яке значення $x \in R$, оскільки у кожній точці $x$ функція $f(x) = ax^2 + bx + c$ має найменше значення с і $Arg\min_x f(x) = R$;
- у випадку $a = 0$ і $b \ne 0$ рівняння (0.2) не має жодного розв'язку, множина мінімізаторів є порожньою, $Arg\min_x f(x) = \varnothing$, тобто функція $f(x) = ax^2 + bx + c$ не має найменшого значення.

### 2. Екстремальні розв'язки задачі оптимізації. Алгоритм обчислення екстремального розв'язку

Якщо число $\bar{x}$ задовольняє необхідній умові мінімуму (0.2), тобто задовольняє рівнянню $f'(\bar{x}) = 0$, то $\bar{x}$ називають ***екстремальним*** або ***стаціонарним*** розв'язком задачі мінімізації функції $f : R \to R$.

Якщо значення $\bar{x}$ задовольняє нерівність $|f'(\bar{x})| < \varepsilon$, то $\bar{x}$ називають ***$\varepsilon$-екстремальним*** розв'язком задачі мінімізації функції $f$.

Якщо функція $f$ має неперервну похідну $f'$ і для деяких точок $a$ та $b$ задовольняє нерівності $f'(a) f'(b) \le 0$, то відрізок $[a;b]$ містить стаціонарний розв'язок $\bar{c}$, $f'(\bar{c}) = 0$.

Для обчислення $\varepsilon$-екстремального розв'язку можна скористатися *алгоритмом половинного поділу*.

***Алгоритм половинного поділу***

I. Вибрати значення параметру $\varepsilon > 0$, яким визначають бажану точність шуканого $\varepsilon$-екстремального розв'язку. Знайти точки $a, b$, які задовольняють нерівність $f'(a)f'(b) \le 0$.

II. Якщо $f'(a) = 0$, то покласти $c = a$ і перейти на крок VIII.
Якщо $f'(b) = 0$, то покласти $c = b$ і перейти на крок VIII.

III. Покласти $c = (a+b)/2$. Якщо $f'(c) = 0$, то перейти на крок VIII.

IV. Якщо $f'(c)f'(a) > 0$, то покласти $a = c$.

V. Якщо $f'(c)f'(b) > 0$, то покласти $a = b$.

VI. Якщо $|b-a| > \varepsilon$, то перейти на крок III.

VII. Покласти $c = (a+b)/2$.

VIII. Покласти $\tilde{c} = c$. Зупинити обчислення.

Очевидно, обчислене значення $\tilde{c}$ відхиляється від екстремального значення $\bar{c}$, $f'(\bar{c}) = 0$, не більше ніж на величину $\varepsilon / 2$.

Якщо у алгоритмі на кроці VI умову $|b-a| > \varepsilon$ замінити умовою $|f'(c)| < \varepsilon$, то отримаємо алгоритм для відшукання $\varepsilon$-екстремального розв'язку $c$.



### 3. Метод Ньютона і метод січних для уточнення екстремального розв'язку

У методі Ньютона екстремальний розв'язок $\bar{c}$ задачі мінімізації функції $f$ обчислюємо як розв'язок рівняння $f'(\bar{c}) = 0$. Для цього вибираємо довільне початкове число $c^1$ і обчислюємо послідовність $c^1$, $c^2, \ldots, c^k$, $c^{k+1}, \ldots$ за формулою

$$c^{k+1} = c^k - f'(c^k) \big/ f''(c^k)\,.$$

Дану формулу отримуємо як наближений розв'язок $c^{k+1}$ рівняння

$$f'(c^{k+1}) = f'(c^k + \Delta c) = 0,$$

якщо замість функції $f'$ скористатися її лінійною апроксимацією $f'(c^k + \Delta c) \cong f'(c^k) + f''(c^k)\Delta c$ і замість розв'язку $\Delta c$ нелінійного рівняння $f'(c^k + \Delta c) = 0$ скористатися розв'язком $\Delta c = -f'(c) \big/ f''(c)$ лінійного рівняння $f'(c^k) + f''(c^k)\Delta c = 0$.

Доведено, що при виконанні нерівностей:

$$0 \le 1 - \left[ f_0'\left(c^k\right) \Big/ f_0''\left(c^k\right) \right]' \le \gamma < 1,$$

існує число $\beta > 0$, що для всіх $k$ виконуються нерівності:

$$\left|c^{k+1}-\overline{c}\right| \le \beta \left|c^{k}-\overline{c}\right|^{2}.$$

Якщо значення другої похідної $f''(c^k)$ обчислювати наближено за формулою $f''(c^k)=\dfrac{f'(c^k)-f'(c^{k-1})}{c^k-c^{k-1}}$, то в околі екстремального розв'язку послідовність $c^1, c^2, \ldots, c^k, c^{k+1}, \ldots$ обчислюється за наближеною формулою

$$x^{k+1}=x^{k}-\frac{f_0'\left(x^{k}\right)\left(x^{k}-x^{k-1}\right)}{f_0'\left(x^{k}\right)-f_0'\left(x^{k-1}\right)}$$

і збігається до екстремального значення $\overline{c}$. Цей метод обчислення екстремального значення називають ***методом січних***.

## 0.4. Умови оптимальності і методи мінімізації диференційованих функцій багатьох змінних

Для знаходження мінімального значення функції $f(x_1, x_2, ..., x_n)$ від $n$ змінних $x_1, x_2, ..., x_n$ на допустимій множині $x_i \in [a_i; b_i]$, $i=\overline{1,n}$ часто використовують метод перебору, а саме: для різних значень $x_i^j = a_i + jh_i$, $j=\overline{0, N_i}$, $h_i = \dfrac{b_i - a_i}{N_i}$, $i=\overline{1,n}$ обчислюють значення функції $f(x_1, x_2, ..., x_n)$, серед яких знаходять потрібне найменше значення. Проте такий метод знаходження мінімального значення функції $f$ часто є нездійсненним. Наприклад, у випадку $n=10$ і $N_i = 50$ будемо мати $10^{50}$ різних векторів $x$ і тому обчислення всіх значень функції $f(x)$ є неможливим. Навіть надпотужний суперкомп'ютер, який би за одну секунду обчислював значення $f(x)$ у 1 000 000 000 000 різних точках, не зможе за мільярди років обчислити значення $f(x)$ в усіх $10^{50}$ точках $x$. Тому для мінімазації $f(x_1, x_2, ..., x_n)$ при великих значеннях $n$ використовуються спеціалізовані алгоритми, які не потребують повного перебору усіх значень $x$. Такими алгоритмами є описані нижче алгоритми Монте–Карло та алгоритми градієнтного спуску.

#### 1. Методи Монте–Карло

У методах Монте–Карло функцію $f$ обчислюють у випадково вибраних точках $x^1, x^2, ..., x^N$. Точки $x^1, x^2, ..., x^N$ вибираються за допомогою комп'ютерного генератора випадкових значень $x$, як реалізацію випадкового процесу, де розподіл випадкової величини для обчислення наступного значення $x^{k+1}$ визначається отриманими на попередніх кроках значеннями $x^1, f(x^1), x^2, f(x^2), ..., x^k, f(x^k)$.

Вражаючим прикладом успішного використання методів Монте–Карло є високоточне попереднє оцінювання результатів голосування на президентських чи парламентських виборах, де за допомогою вибіркового опитування лише декількох тисяч випадково вибраних виборців (серед десятків мільйонів усіх виборців), досягається висока точність прогнозування результатів з похибкою, що не перевищує декількох відсотків. Такий метод мінімізації $f(x_1, x_2, ..., x_n)$ на множині $X = \{x \mid x_i \in [a_i; b_i],\ i = \overline{1,n}\}$ реалізований у наступному алгоритмі методу Монте–Карло.

***Алгоритм Монте–Карло***

I. Вибрати параметри алгоритму $N > 1$ і $M$ та побудувати генератор для обчислення реалізації $G(a,b)$ рівномірно розподіленого на інтервалі $[a;b]$ випадкового числа. Для всіх $i = \overline{1,n}$ покласти $\bar{x}_i = G(a_i, b_i)$ і обчислити значення $\bar{f} = f(\bar{x}_1, \bar{x}_2, ..., \bar{x}_n)$. Покласти $k = 1$.

II. Якщо $\bar{f} \le M$, то перейти на крок V.

III. Для всіх $i = \overline{1,n}$ покласти $\tilde{x}_i = G(a_i, b_i)$ і обчислити значення $\tilde{f} = f(\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_n)$. Якщо $\tilde{f} < \bar{f}$, то для всіх $i = \overline{1,n}$ покласти $\bar{x}_i = \tilde{x}_i$, $\bar{f} = \tilde{f}$.

IV. Покласти $k = k + 1$. Якщо $k < N$, то перейти на крок II.

V. Зупинити обчислення.

Після завершення алгоритму значення вектора $\bar{x} = (\bar{x}_1, \bar{x}_2, ..., \bar{x}_n)^T$ або буде задовольняти нерівність $f(\bar{x}) < M$, або буде мінімізувати значення функції $f$ на множині всіх $N$ реалізацій рівномірно розподіленого на множині $X$ випадкового вектора $x$.

У *адаптивних* алгоритмах Монте–Карло використовуються отримані значення $x^1, f(x^1), x^2, f(x^2), ..., x^k, f(x^k)$ для адаптації на кожній ітерації генератора випадкової величини $x^{k+1}$ з метою збільшення ймовірності прискореного настання бажаної події $f(x^{k+1}) < \min_{x \in X} f(x) + \varepsilon$.

Прикладом такого генератора, реалізованого в наступному адаптивному алгоритмі Монте–Карло для мінімізації функції $f: R^n \to R$ на заданій множині $X \subset R^n$, є обчислення значення $x^{k+1} \in R^n$ як першої реалізації випадкової величини, рівномірно розподіленої у кулі деякого радіуса $r$ з центром у точці $x^{i^*} = \arg\min_{i \in [1:k]} f(x^i)$, яка задовольняє нерівність $f(x^{k+1}) < F_k \triangleq \min_{i \in [1:k]} f(x^i)$.

***Адаптивний алгоритм Монте–Карло***

I. Вибрати параметри алгоритму: $r>0$ (радіус), $N>1$ (кількість ітерацій), $M$ (бажане значення $f(x)$). Побудувати генератор для обчислення реалізації $G(\tilde{x},r)$ випадкового вектора рівномірно розподіленого у кулі радіуса $r$ з центром у точці $\tilde{x}$. Вибрати довільний початковий вектор $x^1$ із допустимої множини $X$. Покласти $\bar{x}=G(x^1,r)$ і обчислити значення $\bar{f}=f(\bar{x}_1,\bar{x}_2,...,\bar{x}_n)$. Покласти $k=1$.

II. Якщо $\bar{f}\le M$, то перейти на крок VI.

III. Покласти $\tilde{x}=G(\bar{x},r)$ і покласти $k=k+1$. Якщо $k>N$, то перейти на крок VI.

IV. Якщо $\tilde{x}\notin X$, то перейти на крок III.

V. Обчислити значення $\tilde{f}=f(\tilde{x})$. Якщо $\tilde{f}<\bar{f}$, то покласти $\bar{x}=\tilde{x}$, $\bar{f}=\tilde{f}$ і перейти на крок II (з метою прискореного зменшення $\tilde{f}$ замість $\bar{x}=\tilde{x}$ беруть значення $(1-\lambda^*)\bar{x}+\lambda^*\tilde{x}$, де $\lambda^*$ є наближеним розв'язком задачі одновимірної оптимізації $\min\limits_{\lambda>0} f((1-\lambda)\bar{x}+\lambda\tilde{x})$).

VI. Зупинити обчислення.

**2. Необхідна умова мінімуму диференційованої функції багатьох змінних і алгоритм найшвидшого спуску**



*Означення 1.* Якщо для функції $f:R^n\to R$ від $n$ змінних $x=(x_1,x_2,...,x_n)^T$ існує границя $\lim\limits_{x_i\to\bar{x}_i}\dfrac{f(\bar{x}_1,\bar{x}_2,...,\bar{x}_{i-1},x_i,\bar{x}_{i+1},...,\bar{x}_n)-f(\bar{x})}{x_i-\bar{x}_i}$, то функцію $f$ називають ***диференційованою у точці $\bar{x}$ по змінній $x_i$***. Значення цієї границі позначають символом $\dfrac{\partial f(\bar{x})}{\partial x_i}$, або спрощеним символом $f'_{x_i}(\bar{x})$, і називають ***частинною похідною функції $f$ у точці $\bar{x}$ по змінній $x_i$***.

*Означення 2.* Впорядкований вектор частинних похідних

$$\nabla f(\bar{x})=\left(\frac{\partial f(\bar{x})}{\partial x_1},\frac{\partial f(\bar{x})}{\partial x_2},...,\frac{\partial f(\bar{x})}{\partial x_n}\right)^T,$$

називають ***градієнтом функції $f$ у точці $\bar{x}$***, а функцію $f$, для якої існує градієнт $\nabla f(\bar{x})$, тобто, для якої існують похідні $\dfrac{\partial f(\bar{x})}{\partial x_i}$ по всіх змінних $x_i$, $i=\overline{1,n}$, називають ***диференційованою у точці $\bar{x}$***.

Диференційована у точці $\bar{x}$ функція $f$ задовольняє в околі точки $\bar{x}$ асимптотичній рівності

$$f(x) = f(\bar{x}) + \left(\nabla f(\bar{x}), x - \bar{x}\right) + o\left(\left\|x - \bar{x}\right\|\right).$$

***Необхідна умова мінімуму.*** Якщо у точці $x^*$, яка мінімізує функцію $f$, існує градієнт $\nabla f(x^*)$, то $\nabla f(x^*) = 0$.

Дійсно, точка $x = x^* - \lambda \nabla f(x^*)$ у околі точки $x^*$ (при достатньо малому значенні $\lambda$) задовольняє асимптотичній рівності

$$f\left(x^* - \lambda \nabla f(x^*)\right) = f\left(x^*\right) - \left(\nabla f(x^*), \lambda \nabla f(x^*)\right) + o\left(\left\|\lambda \nabla f(x^*)\right\|\right).$$

Якщо б виконувалася нерівність $\nabla f(x^*) \neq 0$, то із існування границі $\lim\limits_{\lambda \to 0} o\left(\left\|\lambda \nabla f(x^*)\right\|\right) / \lambda = 0$ випливало б існування числа $\lambda_1 > 0$, для якого виконувалась би нерівність $o\left(\left\|\lambda_1 \nabla f(x^*)\right\|\right) / \lambda_1 < \left(\nabla f(x^*), \nabla f(x^*)\right) / 2$. Але остання нерівність приводить до нерівностей:

$$f\left(x^* - \lambda_1 \nabla f(x^*)\right) = f\left(x^*\right) - \left(\nabla f(x^*), \lambda_1 \nabla f(x^*)\right) + o\left(\left\|\lambda_1 \nabla f(x^*)\right\|\right) <$$
$$< f\left(x^*\right) - \left(\nabla f(x^*), \lambda_1 \nabla f(x^*)\right) + \lambda_1 \left(\nabla f(x^*), \nabla f(x^*)\right) / 2 <$$
$$< f\left(x^*\right) - \left(\nabla f(x^*), \lambda_1 \nabla f(x^*)\right) / 2 < f\left(x^*\right),$$

які суперечать оптимальності $x^*$ і цим завершується доведення рівності $\nabla f(x^*) = 0$.

**Приклад 1.** Знайти вектор $x^* \in R^n$, який мінімізує квадратичну функцію $f(x) = (Ax, x) + (b, x)$.

*Розв'язування.* Беручи до уваги, що у даному випадку $\nabla f(x) = \left(A + A^T\right)x + b$, шукана точка $x^*$ є розв'язком лінійної системи

$$\left(A + A^T\right)x + b = 0.$$

Якщо ця система не має розв'язку, то не існує мінімізатора $x^* = \arg\min\limits_x \left[\left(Ax, x\right) + \left(b, x\right)\right]$ і квадратична функція $f(x) = (Ax, x) + (b, x)$ не має найменшого значення.

**Твердження 1.** Оптимальне значення $x^* = \arg\min\limits_{x \in X} f(x)$ належить підмножині $\breve{X} = \{x \in X \mid \nabla f(x) = 0\} \cup \bar{X}$, де $\bar{X}$ – множина граничних точок множини $X$, тобто, $\arg\min\limits_{x \in X} f(x) = \arg\min\limits_{x \in \breve{X}} f(x)$.

*Означення 3.* Якщо у точці $x^k \in R^n$ для вектора $z \in R^n$ існує границя

$$f'\left(x^k, z\right) = \lim_{\lambda \to 0+} \left[\left(f\left(x^k + \lambda z\right) - f\left(x^k\right)\right) / \lambda\right],$$

то цю границю називають ***похідною у напрямку*** $z$ ***функції*** $f$ ***в точці*** $x^k$.

Якщо функція $f$ є диференційованою в точці $x^k$, то у точці $x^k$ існує похідна $f'\left(x^k, z\right)$ за напрямком $z$ і при цьому $f'\left(x^k, z\right) = \left(\nabla f\left(x^k\right), z\right)$.

*Означення 4.* Вектор $\overline{z}\left(x^k\right)=\arg\min\limits_{\|z\|=1} f'\left(x^k,z\right)$ називають ***вектором найшвидшого спуску*** функції $f$ у точці $x^k$.

**Твердження 2.** Якщо функція $f$ диференційована у точці $x^k$, то

$$\overline{z}\left(x^k\right)=-\nabla f\left(x^k\right).$$

Це означає, що у випадку $\nabla f\left(x^k\right)\neq 0$ точка $x^{k+1}$, обчислена за формулою $x^{k+1}=x^k-\lambda\nabla f(x^k)$, $\lambda=\arg\min\limits_{\lambda>0} f\left(x^k+\lambda\overline{z}\left(x^k\right)\right)$, задовольняє нерівність $f(x^{k+1})<f(x^k)$, і тому за умови $\min\limits_x f(x)>-\infty$ так обчислена послідовність $\left\{f(x^k)\right\}_{k=1}^{\infty}$ є обмеженою знизу монотонно спадною збіжною до екстремального значення послідовністю. Метод мінімізації функції $f$ за допомогою побудови такої послідовності $\{x^k\}_{k=1}^{\infty}$ реалізований у наступному алгоритмі *найшвидшого спуску.*

***Алгоритм найшвидшого спуску***

I. Вибрати довільну початкову точку $x^0\in R^n$ і покласти $k=0$.

II. Якщо $\nabla f\left(x^k\right)=0$, то покласти $x^*=x^k$ і зупинити обчислення.

III. Обчислити $\rho_k=\arg\min\limits_{\rho\geq 0} f\left(x^k-\rho\nabla f\left(x^k\right)\right)$, $x^{k+1}=x^k-\rho_k\nabla f\left(x^k\right)$.

IV. Покласти $k=k+1$ і перейти на крок II.

Знайдений наближений розв'язок $x^k$ системи $P(x)\triangleq\nabla f(x)=0$ можемо уточнити також за формулою $x^{k+1}=x^k+\delta$, якщо поправку $\delta$ обчислити із рівняння $P(x^k+\delta)=0$ або із близького до нього рівняння

$$P(x^k+\delta)\cong P(x^k)+P'(x^k)\delta=0.$$

Маємо: $\delta=-[P'(x^k)]^{-1}P(x^k)$ і тому $\overline{x}^{k+1}=x^k-[P'(x^k)]^{-1}P(x^k)$.

Більш ефективний метод обчислення уточненого значення $x^{k+1}$ здійснюється у львівському алгоритмі, де береться до уваги, що значення $x^{k+1}$ повинно мінімізувати також і функцію $F(x)\triangleq\left(P(x),P(x)\right)$ і тому може бути обчислене також і за іншою формулою:

$$\overline{\overline{x}}^{k+1}=x^k-\rho_k F'\left(x^k\right)=x^k-\rho_k P'^T\left(x^k\right)P\left(x^k\right),$$

$$\rho_k=\arg\min_{\rho>0} F(x^k-\rho P'^T(x^k)P(x^k)).$$

На цій основі у львівському алгоритмі уточнене значення $x^{k+1}$ обчислюють за формулами:

$$x^{k+1}=\overline{x}^{k+1}+\lambda_k(\overline{\overline{x}}^{k+1}-\overline{x}^{k+1}),$$

$$\lambda_k=\arg\min_{\lambda} F(\overline{x}^{k+1}+\lambda_k(\overline{\overline{x}}^{k+1}-\overline{x}^{k+1}))$$

або за наближеними формулами:

$$u^k = x^k - a_k \left[Q_k\right]^{-1} P(x^k), \qquad v^k = x^k - \beta_k Q_k^T P(x^k),$$
$$x^{k+1} = \arg\min_{\lambda} f(u^k + \lambda(v^k - u^k)), \;\; \alpha \in (0;1],$$

де наближена матриця $Q_k$ обчислюється за формулами:

$$(Q_k)_j = \frac{P(x^k + h_k e^j) - P(x^k)}{h_k}, \quad j = \overline{1,n}.$$

**3. Методи послідовних наближень**

При розв'язуванні задачі мінімізації заданої функції $f$ на заданій множині $X$ методами послідовних наближень обчислюють послідовність $\{x^k\}_{k=1}^{\infty}$ розв'язків допоміжних задач оптимізації:

$$x^{k+1} = \arg\min_{x \in X_k} F_k(x),$$

де замість функції $f$ використовують:

або її лінійну апроксимацію $F_1$ в околі точки $x^k$,

$$F_1(x^k, x) \triangleq f(x^k) + (f'(x^k), x - x^k),$$

або її квадратичну апроксимацію $F_2$,

$$F_2(x^k, x) \triangleq f(x^k) + (f'(x^k), x - x^k) + \frac{1}{2}(f''(x^k)(x - x^k), x - x^k),$$

або апроксимації більш високих порядків (для методів прискореної збіжності), а множину $X$ заміняють в околі точки $x^k$ деякою із наступних її локальних підмножин:

$$X_1(x^k, \lambda_k) = \{x \in X \mid \|x - x^k\| \le \lambda_k\},$$
$$X_2(x^k, \lambda_k) = \{x \mid \|x - x^k\| \le \lambda_k\},$$
$$X_k(x^k, \lambda_k, J) = \{x \mid |x_i - x_i^k| \le \lambda_k, \; i \in J \subset \overline{1,n}\}.$$

Вибором різних пар ($F_k, X_k$) та вибором різних способів обчислення наближених значень $\arg\min_{x \in X_k} F_k(x)$ визначаються різні алгоритми $A_k$ для обчислення послідовності $\{x^k\}_{k=1}^{\infty}$:

$$\begin{aligned} x^{k+1} = A_k(&x^k, f(x^k), f'(x^k), f''(x^k), \ldots, f^{(n_k)}(x^k), \\ &x^{k-1}, f(x^{k-1}), f'(x^{k-1}), f''(x^{k-1}), \ldots, f^{(n_{k-1})}(x^{k-1}), \ldots, \\ &x^{k-s_k}, f(x^{k-s_k}), f'(x^{k-s_k}), f''(x^{k-s_k}), \ldots, f^{(n_{k-s_k})}(x^{k-s_k}), p^k) \end{aligned}$$

за допомогою покрокового обчислення значень:

$$x^k, f(x^k), f'(x^k), f''(x^k), \ldots, f^{(n_k)}(x^k).$$

*Означення 5.* Якщо $\lim\limits_{n\to\infty} f(x^n) = \inf\limits_{x\in R^n} f(x)$, то послідовність $\{x^n\}_{n=1}^{\infty}$ називають ***мінімізуючою*** для функції $f$. Якщо $\lim\limits_{n\to\infty} x^n = \arg\min\limits_{x\in X} f(x)$, то послідовність $\{x^n\}_{n=1}^{\infty}$ називають ***збіжною до оптимального розв'язку*** $x^* = \arg\min\limits_{x\in X} f(x)$. Якщо $\lim\limits_{n\to\infty} \nabla f(x^n) = 0$, то послідовність $\{x^n\}_{n=1}^{\infty}$ називають ***збіжною до стаціонарного*** розв'язку задачі оптимізації.

*Означення 6.* Якщо існує таке число $q < 1$, що для всіх $n = 1,2,...$ виконуються нерівності:

$$\left\| x^{n+1} - x^* \right\| \le q \left\| x^n - x^* \right\|,$$

то послідовність $x^1, x^2, ...$ називають ***збіжною із швидкістю геометричної прогресії*** або збіжною з ***лінійною швидкістю***.

Якщо на кожному кроці $n$ виконується нерівність $\left\| x^{n+1} - x^* \right\| \le q_n \left\| x^n - x^* \right\|$, то:

- у випадку $q_n \to 0$ при $(n \to \infty)$ ітераційний метод називають ***збіжним із надлінійною швидкістю***;
- у випадку $q_n \le C \left\| x^n - x^* \right\|$ – ***збіжним з квадратичною швидкістю***;
- у випадку $q_n \le C \left\| x^n - x^* \right\|^{k-1}$ – ***збіжним зі швидкістю $k$-го порядку.***

Різні алгоритми $A_k$ для обчислення послідовності $x^1, x^2, ...$ забезпечують різну швидкість збіжності у залежності від способів використання отриманих на попередніх кроках значень $x^k, f(x^k), f'(x^k), f''(x^k), ..., f^{(n_k)}(x^k)$. У алгоритмах з квадратичною швидкістю збіжності використовують значення похідних другого порядку $f''(x^n)$, а у алгоритмах прискореної збіжності вищих порядків використовуються значення похідних вищих порядків або їх різницевих аналогів. Ефективність ітеративних алгоритмів підвищують також за допомогою адаптивного оцінювання і використання геометричних властивостей функції $f$ з метою виконання нерівностей $f(x^{n+1}) < f(x^n) - C_n$ з якомога більшими значеннями $C_n > 0$.

#### 4. Градієнтні методи, алгоритми допустимих напрямків і теореми збіжності

Із справедливої для диференційованої функції $f$ асимптотичної рівності в околі точки $x^n$

$$f(x) = f(x^n) + (\nabla f(x^n), x - x^n) + o\left(\left\| x - x^n \right\|\right)$$

випливає, що напрямок найшвидшого спадання функції $f$ є протилежним до напрямку градієнта $\nabla f(x^n)$ і значення $x^{n+1}$, яке задовольняє нерівності:

$$f(x^{n+1}) < f(x^n),$$

може бути обчислене при деякому малому значенні множника $\lambda_n$ за формулою:

$$x^{n+1} = x^n - \lambda_n \nabla f(x^n),$$

тобто за *градієнтним алгоритмом* $A_n(x^n, \lambda_n) = x^n - \lambda_n \nabla f(x^n)$.

Різні модифікації градієнтних алгоритмів визначаються різними способами обчислення кроків $\lambda_n$ та різними способами обчислення наближених до антиградієнта $(-\nabla f(x^n))$ напрямків $z^n$.

Якщо напрямок $z^n$ задовольняє нерівності $(z^n, \nabla f(x^n)) \le \delta < 0$, то для лінійного наближення $F_1(x) = f(x^n) + (\nabla f(x^n), x - x^n)$ виконується нерівність:

$$F_1(x^n + z^n) = f(x^n) + (\nabla f(x^n), x + z^n - x^n) = F_1(x^n) + (\nabla f(x^n), z^n) < F_1(x^n),$$

і тому обчислене значення:

$$x^{n+1} = x^n + \lambda_n z^n,$$

з достатньо малим значенням $\lambda_n$ буде задовольняти бажаній нерівності:

$$f(x^n + \lambda_n z^n) < f(x^n).$$

Напрямки $z^n$, які задовольняють нерівності $(z^n, \nabla f(x^n)) \le \delta < 0$ називають ***допустимими напрямками,*** а методи обчислення послідовності $x^1, x^2, \dots$ за формулами $x^{n+1} = x^n + \lambda_n z^n$ з допустимими напрямками $z^n$ називають ***методами допустимих напрямків*** для мінімізації функції $f$.

Проте, вибір значень $\lambda_n$ за умовою $f(x^n + \lambda_n z^n) < f(x^n)$ не гарантує ні збіжності $\lim\limits_{n\to\infty} x^n = x^* = \arg\min\limits_{x \in X} f(x)$, ні збіжності $\lim\limits_{n\to\infty} f(x^n) = \inf\limits_x f(x)$ і навіть не гарантує збіжності $\lim\limits_{n\to\infty} \| \nabla f(x^n) \| = 0$. Наприклад, для функції $f(x) \equiv (x+1)^2$ послідовність $x_n = 1/n$ задовольняє умову $f(x^{n+1}) < f(x^n)$, але не збігається до оптимального розв'язку $x^* = -1$.

Значення $\lambda_n$, які можуть забезпечити потрібну збіжність $\lim\limits_{n\to\infty} \| \nabla f(x^n) \| = 0$, можуть обчислюватися, наприклад, як мінімізатори $\lambda_n = \arg\min\limits_{\lambda \ge 0} f\left(x^n + \lambda z^n\right)$. Проте, практичне обчислення цього мінімізатора є трудомістким у випадках трудомісткого обчислення значення $f\left(x^n + \lambda z^n\right)$. У таких випадках обчислюють $\lambda_n = \arg\min\limits_{\lambda \ge 0} F\left(\lambda\right)$, де $F\left(\lambda\right)$ є параболою $a_n\lambda^2+b_n\lambda+c$, що проходить через обчислені значення $f\left(x^n\right)$, $f\left(x^n+\lambda_{n-1}z^n\right)$ та $f\left(x^n+2^p\lambda_{n-1}z^n\right)$. Якщо $f\left(x^n+\lambda_{n-1}z^n\right)<f\left(x_n\right)$, то $p=1$, інакше $p=-1$.

***Алгоритм 1***

Початок. Задаємо довільні значення параметрів алгоритму $s>0$, $s_1>2$, $\lambda_1=1$.

Основний цикл. Обчислюємо значення крокового множника $\lambda_n$ за формулою $\lambda_n=2^{2-k_n}\lambda_{n-1}$, де $k_n$ – мінімальне натуральне число, при якому виконується нерівність:

$$f\left(x^n+\lambda_n z^n\right)<f\left(x_n\right)-s\lambda_n^2.$$

**Лема 1.** Якщо $\lambda_n$ вибирають таким чином, щоб забезпечувалось виконання нерівності:

$$f(x^{n+1})<f(x^n)-\varphi(a_n)$$

для деякої неспадної додатної при $a_n>0$ функції $\varphi$, то при $\inf\limits_x f\left(x\right)>-\infty$:

- у випадку $a_n=\left\|\nabla f(x^n)\right\|$ забезпечується збіжність $\lim\limits_{n\to\infty}\|\nabla f(x^n)\|=0$;
- у випадку $a_n=\left\|x^n-x^*\right\|$ забезпечується збіжність
$$\lim_{n\to\infty}x^n=\arg\min_{x\in X}f(x);$$
- у випадку $a_n=f\left(x^n\right)-\inf\limits_x f\left(x\right)$ забезпечується збіжність
$$\lim_{n\to\infty}f(x^n)=\inf_x f(x).$$

Дійсно, якщо припустити, що $\lim\limits_{n\to\infty}a_n\neq 0$, то із нерівності $f(x^{n+1})<f(x^n)-\varphi(a_n)$ отримаємо суперечливу нерівність:

$$\lim_{n\to\infty}f\left(x^{n+1}\right)=\lim_{n\to\infty}\left(f\left(x^{n+1}\right)-f\left(x^n\right)-f\left(x^{n-1}\right)+\ldots+f\left(x^2\right)-f\left(x^1\right)\right)<$$

$$<\lim_{n\to\infty}\sum_{i=1}^{n}\left(-\varphi\left(a_i\right)\right)+f\left(x^1\right)=-\infty.$$

Наслідок 1. Якщо для функції $f$ існує така неспадна функція $\varphi\left(a\right)>0$ при $a>0$, що

$$\left\|\nabla f\left(x\right)\right\|\geq\varphi\left(f\left(x\right)-\inf_x f\left(x\right)\right),$$

то із збіжності $\lim_{n\to\infty} \|\nabla f(x^n)\| = 0$ випливає збіжність $\lim_{n\to\infty} f(x^n) = \inf_x f(x)$, а якщо $\left\|\nabla f\left(x\right)\right\| \geq \varphi\left(\left\|x - x^*\right\|\right)$, то із збіжності $\lim_{n\to\infty} \|\nabla f(x^n)\| = 0$ випливає збіжність $\lim_{n\to\infty} x^n = x^* = \arg\min_{x\in X} f(x)$.

***Теорема 1.*** *Якщо градієнт* $\nabla f\left(x\right)$ *задовольняє на множині* $\left\{x \middle| f\left(x\right) \leq f\left(x^1\right)\right\}$ *умову Ліпшиця з константою* $L$, *то для градієнтного методу* $x^{n+1} = x^n - \lambda_n \nabla f(x^n)$ *існує функція* $\varphi$, *яка задовольняє нерівності* $f(x^{n+1}) < f(x^n) - \varphi(a_n)$ *при* $\alpha_n = \left\|\nabla f\left(x^n\right)\right\|$. *Такою функцією є*

$$\varphi\left(\alpha\right) = \min\left\{s_1\lambda_{n-1}/2, \frac{\alpha}{2\left(L+s\right)}\right\}.$$

Для доведення теореми 1 достатньо перевірити, що нерівностям $f(x^{n+1}) < f(x^n) - \varphi(a_n)$ задовольняє функція

$$\varphi\left(\alpha\right) = \min\left\{s_1\lambda_{n-1}/2, \frac{\alpha}{2\left(L+s\right)}\right\}.$$



Таким чином, з використанням алгоритму 1 для обчислення значення крокового множника $\lambda_n$ в градієнтному методі $x^{n+1} = x^n - \lambda_n \nabla f(x^n)$ в умовах теореми 1 при $\inf_x f\left(x\right) > -\infty$, забезпечується збіжність $\lim_{n\to\infty} \|\nabla f(x^n)\| = 0$.

У описаних вище градієнтних методах значення $\lambda_n$, які забезпечували збіжність, обчислювалися за отриманими значеннями $f\left(x^{n+1}\right), f\left(x^n\right),\ldots$ Проте збіжність градієнтного методу можуть забезпечувати також і з використанням послідовності $\lambda_1, \lambda_2, \ldots$ незалежної від значень $f\left(x^1\right), f\left(x^2\right),\ldots$. Один клас таких послідовностей $\lambda_1, \lambda_2, \ldots$, визначає теорема 2 та її наслідок.

***Теорема 2.*** *Якщо на послідовності* $x^1, x^2, \ldots$ *градієнти* $\nabla f\left(x^n\right)$ *задовольняють умову Ліпшиця з константою* $L$ *і функція* $f$ *обмежена знизу,* $\inf_x f\left(x\right) > -\infty$, *то метод (який називають* ***методом можливих напрямків****)*

$$x^{n+1} = x^n + \lambda_n z^n, \quad (z^n, \nabla f(x^n)) \leq \delta < 0$$

*задовольняє нерівність*

$$\min_{i\in[1:n]} \|\nabla f(x^i)\| \leq (f(x^1) - \inf_x f(x) - 2^{-1} L \sum_{j=1}^{n} \lambda_j^2) / \delta \sum_{j=1}^{n} \lambda_j .$$

Для доведення теореми 2 достатньо скористатися справедливою для всіх $n$ нерівністю $f(x^{n+1}) \le f(x^n) - \lambda_n \delta \| \nabla f(x^n) \| + 2^{-1} L \lambda_n^2$.

Наслідок 2. Якщо вибрана послідовність $\lambda_1, \lambda_2, \ldots$ задовольняє умовам:

$$\lambda_i \ge 0, \ \sum_{i=1}^{\infty} \lambda_i = \infty, \ \sum_{i=1}^{\infty} \lambda_i^2 / \sum_{i=1}^{\infty} \lambda_i = 0,$$

то в умовах теореми 2 $\underline{\lim}_{n\to\infty} \left\| \nabla f\left(x^n\right) \right\| = 0$.

Для опуклої функції $f$ вибір $\lambda_n$ спрощується за наступною теоремою.

***Теорема 3.*** *Якщо послідовність $\lambda_1, \lambda_2, \ldots$ задовольняє умови:*

$$\lambda_n > 0, \ \lambda_n \to 0, \ \sum_{i=1}^{n} \lambda_n = \infty$$

*і опукла функція $f$ досягає найменшого значення у точці $x^*$, $x^* = \arg\min_x f(x)$, то на послідовності $\{x^n\}_{n=1}^{\infty}$, яка обчислена за формулами:*

$$x^{n+1} = x^n + \lambda_n z^n, \ z^n \to \nabla f\left(x^n\right),$$

*існує мінімізуюча підпослідовність $\{x^{n_i}\}_{i=1}^{\infty}$, $\underline{\lim}_{i\to\infty} f\left(x^{n_i}\right) = f\left(x^*\right)$.*



Дійсно, якщо такої підпослідовності не існує, то це означало б існування такого числа $\varepsilon > 0$, що для всіх $n$ виконується нерівність $f(x^n) - f(x^*) > \varepsilon$. У цьому випадку із справедливої у кожній точці $x^i$ для опуклої функції $f$ нерівності $f(x^i) - f(x^*) \le (\nabla f(x^i), x^i - x^*)$ випливає нерівність $(\nabla f(x^i), x^i - x^*) > \varepsilon$, а із справедливої для всіх $n$ нерівності

$$\left\| x^{n+1} - x^* \right\|^2 = \left\| x^n + \lambda_n z^n - x^* \right\|^2 \le \left\| x^n - x^* \right\|^2 - 2\lambda_n (z^n, x^n - x^*) + \lambda_n^2 \left\| z^n \right\|^2$$

в умовах $\lambda_n > 0, \ \lambda_n \to 0,$ випливає існування такого числа $m \ge 1$, що для всіх $n \ge m$ виконується нерівність

$$\left\| x^{n+1} - x^* \right\|^2 \le \left\| x^n - x^* \right\|^2 - \lambda_n (z^n, x^n - x^*),$$

із якої випливають справедливі для всіх $n \ge 1$ нерівності:

$$\left\| x^{m+n} - x^* \right\|^2 \le \left\| x^m - x^* \right\|^2 - \sum_{i=m}^{m+n} \lambda_i \left( \nabla f(x^i), x^i - x^* \right) - \left( \nabla f(x^i) - z^i, x^i - x^* \right) <$$

$$< \left\| x^m - x^* \right\|^2 - \sum_{i=m}^{m+n} \left( \lambda_i \nabla f(x^i), x^i - x^* \right) / 2,$$

$$\left\| x^{m+n} - x^* \right\|^2 \le \left\| x^m - x^* \right\|^2 - \sum_{i=m}^{m+n} \lambda_i \varepsilon / 2.$$

Проте остання нерівність суперечить рівності $\sum_{i=1}^{n} \lambda_n = \infty$, і цим завершується доведення теореми 3.

Оскільки умову $\lambda_n \to 0$ неможливо реалізувати на комп'ютерах з обмеженими мантисами чисел, то у практичних розрахунках використовуються також і методи з достатньо малим постійним кроковим множником $\lambda_n = \lambda, \lambda > 0$.

***Теорема 4.*** *Для методу можливих напрямків*

$$x^{n+1} = x^n + \lambda z^n, \quad (z^n, \nabla f(x^n)) \le \delta < 0$$

*в умовах теореми 2 виконується нерівність*

$$\min_{i=\overline{1,n}} \left\| \nabla f(x^i) \right\| \le \left( f(x^1) - \inf_x f(x) \right) / \delta\lambda n + L\lambda / 2\delta .$$

Для доведення теореми 4 достатньо скористатися нерівністю теореми 2 за умови $\lambda_n = \lambda$.

Наслідок 3. За $N$ ітераційних кроків, $N \le \left( f(x^1) - \inf_x f(x) \right) / \lambda^3$, градієнтний метод з постійним кроковим множником $x^{n+1} = x^n + \lambda z^n$ забезпечує відшукання $\varepsilon$-екстремального розв'язку для $\varepsilon = \lambda\left(\lambda + L/2\right)/\delta$.

Дійсно, використовуючи теорему 4 отримаємо, що з постійним $\lambda$ нерівність $\left\|\nabla f(x^i)\right\| \le \lambda\left(\lambda + L/2\right)/\delta$ буде виконана не більше ніж за $N$ ітераційних кроків.



### 5. Необхідні і достатні умови оптимальності для задачі мінімізації опуклої функції. Алгоритм узагальнених градієнтів

*Означення 7.* Якщо для будь-яких двох точок $x^1, x^2 \in R^n$ і для будь-якого числа $\lambda \in [0;1]$ виконується нерівність

$$f\left(\lambda x^1 + (1-\lambda)x^2\right) \le \lambda f\left(x^1\right) + \left(1-\lambda\right) f\left(x^2\right),$$

то функцію $f$ називають ***опуклою*** (***опуклою донизу***). Якщо ж виконується строга нерівність

$$f\left(\lambda x^1 + (1-\lambda)x^2\right) < \lambda f\left(x^1\right) + \left(1-\lambda\right) f\left(x^2\right),$$

то функцію $f$ називають ***строго опуклою***, а якщо при деякому $\alpha > 0$ виконуються нерівності

$$f\left((x^1 + x^2)/2\right) \le \left( f\left(x^1\right) + f\left(x^2\right) \right) / 2 - \alpha \left\| x^1 - x^2 \right\|,$$

то функцію $f$ називають ***сильно опуклою***.

*Означення 8.* Якщо для будь-яких точок $x^1, x^2$ із множини $X \subset R^n$ усі точки $x_\lambda = \lambda x^1 + (1-\lambda)x^2$ із відрізка $\{x \mid x = \lambda x^1 + (1-\lambda)x^2, \lambda \in [0;1]\}$

належать множині $X$, то множину $X$ називають ***опуклою***. Якщо при всіх $\lambda \in (0;1)$ точки $x_\lambda$ є внутрішніми точками множини $X$, то множину $X$ називають ***строго опуклою***, а якщо для кожного $\lambda \in \left(0;1\right)$ знайдеться число $\alpha > 0$, для якого відкрита множина $\left\{ x \middle\| \left\| x - x_\lambda \right\| < \alpha \left\| x^1 - x^2 \right\| \min\left\{\lambda, 1-\lambda\right\} \right\}$ належить множині $X$, то множину $X$ називають ***сильно опуклою***.

Для будь-яких опуклих функцій $f_i$ і для будь-якої опуклої множини $Y$ множина $X^1 = \{x \mid f_i(x) \le 0,\ i = \overline{1,m},\ x \in Y\}$ є опуклою.

Важливою властивістю опуклої функції $f$ є існування в кожній точці $\bar{x}$ із її області визначення такого вектора $c(\bar{x})$, який для всіх $x \in R^n$ задовольняє нерівність

$$f(x) \ge f(\bar{x}) + (c(\bar{x}), x - \bar{x}).$$

Вектор $c(\bar{x})$ називають ***узагальненим градієнтом функції*** $f$ у точці $\bar{x}$. Важливою властивістю узагальненого градієнта є те, що у випадку нерівності $f(\bar{x}) > f(x^*)$ точка $\bar{x} - \lambda c(\bar{x})$ наближається до множини мінімізаторів $x^* = \arg\min\limits_x f\left(x\right)$ при зростанні числа $\lambda$ в околі значення 0.

Якщо $x^* = \arg\min\limits_x f(x)$, то із нерівності $f(x^*) < f(\bar{x})$, випливають нерівності:

$$(c(\bar{x}), x^* - \bar{x}) \le f(x^*) - f(\bar{x}) < 0,$$

і тому між вектором $(-c(\bar{x}))$ і вектором $x^* - \bar{x}$ є гострий кут. Це означає, що при достатньо малому значенні $\lambda$ відстань від точки $\bar{x} - \lambda c(\bar{x})$ до точки $x^*$ буде зменшуватися із зростанням числа $\lambda$ в околі $\lambda = 0$ і тому послідовність $\{x^k\}_{k=0}^{\infty}$, яка обчислена за формулою

$$x^{k+1} = x^k - \lambda_k\, c(x^k) \big/ \left\| c(x^k) \right\|,$$

може при розумному виборі значень $\lambda_k$ збігатися до мінімізаторів $x^*$. На цій основі побудований описаний нижче алгоритм узагальнених градієнтів для мінімізації опуклої функції $f$.

### *Алгоритм узагальнених градієнтів*

I. Вибрати довільну початкову точку $x^0 \in R^n$ і покласти $k = 0$. Вибрати числову послідовність $\left\{\rho^k\right\}_{k=0}^{\infty}$, яка задовольняє умовам $\rho_k \ge 0;\ \sum\limits_{k=0}^{\infty} \rho_k = \infty,\ \lim\limits_{k\to\infty} \rho_k = 0$ (такою послідовністю є, наприклад, $\rho^k = L/k$ при будь-якому значенні $L > 0$).

II. Якщо $c\left(x^k\right)=0$, то покласти $x^*=x^k$ і зупинити обчислення.

III. Обчислити $x^{k+1}=x^k-\rho_k c\left(x^k\right)/\|c\left(x^k\right)\|$.

IV. Покласти $k=k+1$ і перейти на крок II.

Доведено, що у випадку обмеженої множини $X^*=Arg\min\limits_x f(x)\neq\varnothing$ усіх мінімізаторів $x^*$ опуклої функції $f$ послідовність $\left\{x^k\right\}_{k=0}^{\infty}$, яка побудована за алгоритмом узагальнених градієнтів, збігається з будь-якої початкової точки $x^0$ до множини мінімізаторів $X^*$ і справедливі рівності:

$$\lim_{k\to\infty}\min_{x\in X^*}\left\|x^k-x\right\|=0;\lim_{k\to\infty}f\left(x^k\right)=\min_{x\in R^n}f\left(x\right).$$

**Твердження 3.** Якщо опукла функція $f$ є двічі диференційованою, то наступні умови є еквівалентними:

(1) $f\left(x^2\right)-f\left(x^1\right)\geq\left(\nabla f\left(x^1\right),x^2-x^1\right)$;

(2) $\left(\nabla f\left(x+\lambda h\right),h\right)$ – неспадна функція $\lambda$;

(3) $\left(\nabla^2_{xx}f\left(x\right)h,h\right)\geq 0 \quad \forall x,\ h\in R^n$.

**Твердження 4.** Якщо опукла функція $f$ є двічі неперервно диференційованою, то умова сильної опуклості еквівалентна умові $\left(\nabla^2_{xx}f\left(x\right)h,h\right)\geq\varepsilon\left\|h\right\|^2$, $\varepsilon>0,\ h\in R^n$. Сильно опукла $f$ досягає на замкненій множині свого мінімального значення $f_*$. Строго опукла функція може досягати свого мінімального значення на опуклій множині $X$ лише в єдиній точці.

*Означення 9.* Множину $K\left(x^1\right)$ називають ***конусом допустимих напрямів для множини*** $X$ у точці $x^1$, якщо її елементами є усі вектори $h$, для яких існує число $\varepsilon>0$, що задовольняє умові $x^1+h\cdot\varepsilon\in X$.

*Означення 10.* Множину

$$K^*(\overline{x})=\{y\,|\,(y,z)\geq 0, z\in K(\overline{x})\}$$

називають ***спряженим конусом у точці*** $\overline{x}$.

Множина $X^*$ всіх розв'язків задачі мінімізації опуклої функції $f$ на опуклій множині $X$ є множиною опуклою.

**Твердження 5.** Якщо функція $f$ і множина $X$ є опуклими і функція $f$ є неперервно диференційованою, то необхідною і достатньою умовою оптимальності точки $\overline{x}$ є включення $\nabla f(x)\in K^*(\overline{x})$, і множина $X^*=Arg\min\limits_{x\in X}f(x)$ усіх оптимальних розв'язків $x^*=\arg\min\limits_{x\in X}f(x)$ задовольняє співвідношенню

$$X^*=\{x\,|\,\nabla f(x)\in K^*(x), x\in X\}.$$

Н а с л і д о к 4. $X^* = \{x \mid (\nabla f(x), z - x) \geq 0, z \in X\} \bigcap X$, а якщо $X = R^n$, то $X^* = \{x \mid \nabla f(x) = 0\}$.

*Означення 11.* Вектор $\hat{\nabla} f(\bar{x}) \in R^n$ називають ***субградієнтом*** або ***узагальненим градієнтом функції*** $f$ ***в точці*** $\bar{x}$, якщо для всіх $x$ виконується нерівність

$$f(x) \geq f(\bar{x}) + (\hat{\nabla} f(\bar{x}), x - \bar{x}).$$

Множина $M(x)$ всіх субградієнтів неперервної опуклої функції $f$ в точці $x$ є опуклою, замкнутою і обмеженою.

**Твердження 6.** Якщо $M(\bar{x})$ є множиною субградієнтів в точці $\bar{x}$ для опуклої функції $f$, то значення похідної $f'(\bar{x}, z)$ від функції $f$ у точці $\bar{x}$ у будь-якому напрямку $z$ обчислюється за формулою

$$f'(\bar{x}, z) = \max_{y \in M(\bar{x})} (z, y).$$

Н а с л і д о к 5. Напрямок $z(\bar{x})$ найшвидшого спадання функції $f$ в точці $\bar{x}$ є розв'язком оптимізаційних задач:

$$z(\bar{x}) = \operatorname*{argmin}_{\|z\|=1} f'(\bar{x}, z) = \operatorname*{argmin}_{\|z\|=1} \max_{y \in M(\bar{x})} (z, y).$$

Н а с л і д о к 6. Якщо функція $f$ є опуклою, то для оптимальності точки $x^* = \arg\min_x f(x)$ необхідно і достатньо, щоб множина $M(x^*)$ субградієнтів функції $f$ у точці $x^*$ містила точку $x = 0$, тобто виконувалася умова $0 \in M(x^*)$.

Важливий клас прикладних задач оптимізації пов'язаний із мінімізацією функції максимуму $f(x) = \max_{y \in Y} \bar{f}(x, y)$ та відшуканням мінімізатора $x^* = \arg\min_x \max_y \bar{f}(x, y)$, для якого при всіх $x$ виконується нерівність

$$\max_y \bar{f}(x^*, y) \leq \max_y \bar{f}(x, y).$$

До такої задачі мінімаксної оптимізації зводиться, наприклад, відшукання гарантованого розв'язку $x^*_{\max} = \arg\min_{x \in X} \max_{y \in Y} \bar{f}(x, y)$ в задачах оптимізації з неповними даними $y \in Y$.

Якщо функції $\bar{f}(x, y)$ є опуклими по $x$, то функція $f(x) = \max_{y \in Y} \bar{f}(x, y)$ також є опуклою і у кожній точці $\bar{x}$ вектор $g_f(\bar{x}) = \nabla_x \bar{f}(\bar{x}, y^*)$, обчислений на будь-якому максимізаторі $y^* = \arg\max_{y \in Y} \bar{f}(x, y)$ є субградієнтом функції $f$, тобто для всіх $x$ задовольняє нерівність

$$f(x) - f(\bar{x}) \geq (g_f(\bar{x}), x - \bar{x}),$$

і тому послідовність $\{x^k\}_{k=0}^{\infty}$, обчислена за формулами:

$$x^{k+1} = x^k - h_k\, g_f(x^k) \big/ \left\| g_f(x^k) \right\|, \; h_k > 0; \quad \lim_{k\to\infty} h_k = 0; \quad \sum_{k=1}^{\infty} h_k = +\infty,$$

задовольняє рівностям:

$$\lim_{k\to\infty} \min_{y\in X^*} \left\| x^k - y \right\| = 0, \; \lim_{k\to\infty} \max_{y\in Y} \overline{f}(x^k, y) = \min_{y\in R^n} \max_{y\in Y} \overline{f}(x, y) = f^*.$$

**6. Методи відтинань для мінімізації опуклих функцій. Метод еліпсоїдів**

Беручи до уваги, що для визначеної у точці $\overline{x} \in R^n$ опуклої функції $f : R^n \to R$ існує вектор $c(\overline{x}) \in R^n$, який для всіх $x \in R^n$ задовольняє нерівності $f(x) \ge (c(\overline{x}), x - \overline{x}) + f(\overline{x})$, приходимо до висновку, що лінійна функція

$$F_{\overline{x}f}(x) \equiv (c(\overline{x}), x - \overline{x}) + f(\overline{x})$$

є мінорантою для функції $f$. Із нерівності $\left(c(\overline{x}), x - \overline{x}\right) > 0$ випливає нерівність $f(x) > f(\overline{x})$. Це означає, що мінімізатор $x^*$ функції $f$ належить множині $\{x \mid \left(c(\overline{x}), x - \overline{x}\right) \le 0\}$.

На цій основі побудовано *метод еліпсоїдів* для відшукання мінімізатора опуклої функції $f$. Якщо мінімізатор $x^*$ опуклої функції $f$ належить кулі $S(x^1, R_1) = \{x \mid \left\| x^1 - x \right\| \le R_1\}$ радіуса $R_1$ з центром в точці $x^1$, то $x^*$ належить півкулі

$$S_1(x^1, R_1) = S(x^1, R_1) \cap \left\{ x \mid \left( c(x^1), x - x^1 \right) \le 0 \right\}.$$

Якщо тілесний гіпереліпсоїд мінімального об'єму, який містить півкулю $S_1(x^1, R_1)$, розтягнути лінійним перетворенням до кулі, то радіус $R_2 = R_1(1 - 1/2n)^{-1/2}$ цієї нової кулі з новим центром $x^2$ буде меншим від радіуса $R_1$ попередньої кулі. Після відтинання від нової кулі її половини у півпросторі $\left\{ x \mid \left( c(x^2), x - x^2 \right) > 0 \right\}$, який не містить мінімізатора $x^*$, отримаємо новий ще менший гіпереліпсоїд, який містить $x^*$, і повторивши операцію його розтягування отримаємо кулю із ще меншим радіусом $R_3 = R_2(1 - 1/2n)^{-1/2}$ та новим центром у точці $x^3$. Повторюючи відтинання півкуль, які не містять мінімізатора $x^*$ отримаємо послідовність $\left\{ x^k \right\}_{k=1}^{\infty}$ центрів куль із все меншими радіусами $R_k = R_{k-1}(1 - 1/2n)^{-1/2}$, яким належить шуканий розв'язок $x^*$. Доводиться, що на кожному $k$-му кроці виконується нерівність $\left\| x^{k+1} - x^* \right\| \le q \left\| x^k - x^* \right\|$ із незалежним від функції $f$

множником $q \approx 1 - k^2/2 < 1$. Це означає, що границя послідовності $\left\{x^k\right\}_{k=1}^{\infty}$ є мінімізатором $x^*$, $\lim_{k \to \infty} x^k = x^*$.

Зауважимо, що методом еліпсоїдів можна розв'язувати також і задачу мінімізації опуклої функції $f$ на опуклій множині $X = \{x \mid f_i(x) \le 0,\, i = \overline{1,m}\}$, заданій за допомогою опуклих функцій $f_i$. Для цього достатньо у методі еліпсоїдів функцію $f$ замінити допоміжною функцією $\bar{f}$, визначеною за формулами:

$$\bar{f}(x) = \begin{cases} f(x), & x \in X, \\ f_{i^*}(x), i^* = \arg\max_{i \in [1:m]} f_i(x), & x \notin X. \end{cases}$$

Беручи до уваги, що для будь-якої множини точок $x^1, x^2, ..., x^N$ мінімізатор опуклої функції $f$ належить перетину $X_N = \bigcap_{i=1}^{N} \left\{x \mid \left(c(x^i), x - x^i\right) \le 0, \quad i = \overline{1,n}\right\}$, задачу мінімізації функції $f$ на множині $X$ можемо замінити задачею мінімізації функції $f$ на «меншій» множині $X \cap X_N$. Якщо точки $x^1, x^2, ..., x^N$ вибирати таким чином, щоб діаметр підмножини $X \cap X_N$ досяг малого значення $\varepsilon$, то, очевидно, будь-яка точка $\bar{x}$ із множини $X \cap X_N$ є наближеним розв'язком задачі мінімізації функції $f$ на множині $X$ з похибкою, що не перевищує $\varepsilon$, $\| \bar{x} - x^* \| < \varepsilon$.

Формально до методів відтинань можна віднести також і методи локалізації розв'язків на тих підмножинах елементів $x \in X$, які задовольняють описаним нижче необхідним умовам оптимальності.

### 7. Умови оптимальності для задачі мінімізації неперервної функції з кусково-неперервними градієнтами

Прикладом неперервної функції з кусково-неперервними градієнтами є функція максимуму

$$f_0(x) = \max_{i \in \Im} f_i(x), \; \Im = \overline{1,m}.$$

**Твердження 7**. Якщо точка $\bar{x}$ є розв'язком задачі оптимізації $\bar{x} = \arg\min_x \max_{i \in \Im} f_i(x)$ для диференційованих функції $f_i$, $i \in \Im$, то існують такі числа $u_i \ge 0$ $\sum_{i \in \Im} u_i = 1$, які задовольняють рівнянням:

$$\sum_{i \in \Im} u_i \nabla f_i(\bar{x}) = 0, \; u_i(f_i(\bar{x}) - \max_{i \in \Im} f_i(\bar{x})) = 0, \; i \in \Im.$$

*Означення 12.* Неперервну функцію $f$ називають ***слабо опуклою***, якщо у кожній точці $x$ існує не порожня множина $G_f(x)$, яка для $\forall h \in G_f(x)$, $\forall y \in R^n$, задовольняє умову:

$$f(y) - f(x) \geq (h, y-x) + r(x, y),$$

де значення $r(x, y)/\|x-y\|$ наближається рівномірно по $x$ до нуля при $\| x-y \| \to 0$ в кожній обмеженій замкненій множині $Q \ni x$.

Відомо, що властивістю неперервної опуклої функції є існування у кожній точці $x$ множини субградієнтів $M(x)$, для якої виконується рівність:

$$f'(x, z) = \max_{y \in M(x)} (y, z),$$

де $f'(x, z)$ є похідною функції $f$ за напрямком $z$ у точці $x$. Аналогічну властивість мають також і квазідиференційовані неопуклі функції.

*Означення 13.* Неперервну функцію $f$ називають ***квазідиференційованою***, якщо для кожної точки $x$ існує замкнута опукла множина $M_f(x)$, яка задовольняє для всіх $z \in R^n$ рівності:

$$f'(x, z) = \max_{y \in M_f(x)} (y, z),$$

де $f'(x, z)$ є похідною функції $f$ за напрямком $z$ у точці $x$. Множину $M_f(x)$ називають ***множиною квазіградієнтів***.

Зауважимо, що всі опуклі функції є квазідиференційованими, оскільки множиною квазіградієнтів для опуклої функції є множина її субградієнтів. Квазідиференційованими є також всі слабо опуклі функції $f$ та всі функції $f_0$ вигляду $f_0(x) = \max_{y \in Y} f(x, y)$ з диференційованими функціями $f$ і замкненою обмеженою множиною $Y$. Для таких функцій множиною квазіградієнтів є множина $M_f(x) = \mathrm{co}\{z \,|\, z = \nabla_x f(x, y^*)$, $y^* = \arg\max_{y \in Y} f(x, y)\}$. До квазідиференційованих належать також і функції вигляду $f_0(x) = \max_{a \in A} f(x, a)$, де $f(x, a)$ – неперервні по $a$ і слабо опуклі по $x$ при будь-якому фіксованому $a$ з компактного топологічного простору $A$. В останньому випадку $M_{f_0}(x) = \mathrm{co} \bigcup_{a \in A} G_{f(\cdot, a)}(x)$.

**Твердження 8**. Для того, щоб точка $\overline{x} \in R^n$ була мінімізатором квазідиференційованої функції $f$, необхідно, щоб $0 \in M_f(\overline{x})$.

**Твердження 9.** Вектор $z(x^1) = \arg\min_{\|z\|=1} \max_{y \in M_f(x^1)} (y, z)$ є напрямком найшвидшого спуску функції $f$ в точці $x^1$.

В роботах [29, 30] дано більш загальне визначення квазідиференційованої функції, а саме: функція $f : R^n \to R$ називається ***квазідиференційованою в точці*** $\bar{x}$, якщо вона в цій точці диференційована по напрямку і якщо існують такі опуклі компакти $\partial_1 f(\bar{x}) \subset R^n$ і $\partial^1 f(\bar{x}) \subset R^n$, що для будь-якого $z \in R^n$ виконується рівність

$$f'(\bar{x}, z) = \max_{y \in \partial_1 f(\bar{x})} (z, y) + \min_{y \in \partial^1 f(\bar{x})} (z, y).$$

Множина таких функцій замкнута відносно алгебраїчних операцій та операцій взяття максимуму і мінімуму від обмеженого числа функцій.

**Твердження 10.** Якщо $\bar{x} = \arg\min_x f(x)$, то $-\partial^1 f(\bar{x}) \subset \partial_1 f(\bar{x})$; а якщо $\bar{x} = \arg\max_x f(x)$, то $-\partial_1 f(\bar{x}) \subset \partial^1 f(\bar{x})$.

**8. Методи стохастичних градієнтів**

*Означення 14.* ***Стохастичним градієнтом*** функції $f : R^n \to R$ у точці $x^k$ називають випадковий вектор $g_\omega\left(x^k\right)$, математичне сподівання якого $M_\omega g_\omega\left(x^k\right) = g_f\left(x^k\right)$ співпадає з градієнтом $g_f\left(x^k\right)$ функції $f\left(x\right)$, $M_\omega g_\omega\left(x^k\right) = g_f\left(x^k\right)$, або з узагальненим градієнтом $g_f\left(x^k\right)$ для опуклої функції $f$, який для всіх $x$ задовольняє нерівність

$$f(x) \ge f(x^k) + (g_f(x^k), x - x^k).$$

Якщо $f(x)$ дорівнює математичному сподіванню $f(x) \triangleq M_p F(x, p)$ випадкової величини $F(x, p)$, визначеної диференційованою функцією $F : R^n \times P \to R$ з випадковою величиною $p \in P$, то $g_\omega\left(x^k\right) = \nabla_x F(x^k, p^k)$, де $p^k$ – реалізація випадкової величини $p$. У методах узагальнених стохастичних градієнтів для мінімізації функції $f$ обчислюють послідовність $x^{k+1} = x^k - h_k\left(x^k\right) g_\omega\left(x^k\right)$ (а для мінімізації математичного сподівання $f(x) \triangleq M_p F(x, p)$ обчислюють $x^{k+1} = x^k - \lambda_k \nabla_x F(x^k, p^k)$), яка в умовах наступної теореми збігається з ймовірністю 1 до мінімізатора $x^*$ функції $f$.

***Теорема 5.*** *Якщо* $x^*$ *єдина точка мінімуму функції* $f$, *кроки* $h_k$ *задовольняють умовам* $\sum_{k=1}^{\infty} h_k\left(x^k\right)=+\infty$, $h_k\left(x^k\right)>0$, $\sum_{k=1}^{\infty} h_k^2\left(x^k\right)<\infty$ *і математичне сподівання узагальненого стохастичного градієнту обмежене* $M_{\omega}\left\|g_{\omega}\left(x^k\right)\right\|^2<c$, *то з ймовірністю 1 послідовність* $x^{k+1}=x^k-h_k\left(x^k\right)g_{\omega}\left(x^k\right)$ *збігається до мінімізатора* $x^*$, *тобто* $\lim\limits_{k\to\infty}\left\|x^k-x^*\right\|=0$.

***Доведення.*** Із очевидних рівностей:

$$\left\|x^{k+1}-x^*\right\|^2=\left\|x^k-h_k\left(x^k\right)g_{\omega}\left(x^k\right)-x^*\right\|^2=$$

$$=\left\|x^k-x^*\right\|^2-2h_k\left(x^k\right)\left(g_{\omega}\left(x^k\right),x^k-x^*\right)+h_k^2\left(x^k\right)\left\|g_{\omega}\left(x^k\right)\right\|^2$$

випливає рівність

$$M_{\omega}\left\|x^k-x^*\right\|^2=M_{\omega}\left\|x^k-x^*\right\|^2-2h_k\left(x^k\right)\left(M_{\omega}g_{\omega}\left(x^k\right),x^k-x^*\right)+h_k^2M_{\omega}\left\|g_{\omega}\left(x^k\right)\right\|^2.$$

Беручи до уваги рівність $M_{\omega}g_{\omega}\left(x^k\right)=g_f\left(x^k\right)$ та нерівність $\left(g_f\left(x^k\right),x^k-x^*\right)\geq 0$, маємо нерівність

$$M_{\omega}\left(\left\|x^{k+1}-x^*\right\|^2 \mid x^k\right)\leq\left\|x^k-x^*\right\|^2+ch_k^2\left(x^k\right).$$

Це означає, що послідовність $\left\{z^k\right\}_{k=1}^{\infty}$, $z_k=\left\|x^k-x^*\right\|^2+ch_k^2\left(x^k\right)$, задовольняє нерівності $M\left(z^{k+1}\mid z^k,...,z^1\right)\leq z^k$ і є півмартингалом. Із теорії ймовірностей відомо, що півмартингали мають границю. Отже, послідовність $\left\{z^k\right\}_{k=1}^{\infty}$ має деяку границю $z^*$ із ймовірністю 1. Із рівності $\lim\limits_{k\to\infty}\sum_{s=k}^{\infty}h_s^2=0$ випливає, що послідовність $\left\{\left\|x^{k+1}-x^*\right\|^2\right\}$ також збігається до цієї границі $z^*$ із ймовірністю 1. Якщо допустити, що $\lim z^k\neq z^*$, то це означатиме існування таких значень $\varepsilon>0$, $\delta>0$, що із ймовірністю $\delta$ існує підпослідовність $x^{k_1}(\omega),...,x^{k_i}(\omega)...$, $k_1<k_2<...$, яка задовольняє нерівностям $\left\|x^{k_i}(\omega)-x^*\right\|\geq\varepsilon$. Це означало б виконання рівності

$$\sum_{k=1}^{\infty}h_k\left(x^k\right)\left(M_{\omega}g_{\omega}\left(x^k\right),x^k-x^*\right)=+\infty,$$

яка суперечить співвідношенню:

$$M_\omega \left\| x^{k+1} - x^* \right\|^2 = \left\| x_0 - x^* \right\|^2 - 2\sum_{s=0}^{k} h_s\left(x^s\right)\left(M_\omega g_\omega\left(x^s\right), x^s - x^*\right) +$$

$$+\sum_{s=0}^{k} h_s^2\left(x^s\right) M_\omega \left\| g_\omega\left(x^s\right)\right\|^2.$$

Отримане протиріччя завершує доведення теореми.

Похибки обчислення узагальненого градієнту $g_f\left(x^k\right)$ та відповідного обчислення наближеного значення $\bar{x}^k$ замість точного значення $x^k$ можуть не порушувати збіжності, а саме: якщо для опуклої функції $f$ існує мінімізатор $M^* \triangleq Arg\min_x f(x) \neq \varnothing$, то при будь-якій початковій точці $x^0$ для обчисленої за формулами $x^{k+1} = x^k - h_k \dfrac{g_f\left(\bar{x}^k\right)}{\left\| g_f\left(\bar{x}^k\right)\right\|}$ послідовності $\left\{x^k\right\}_{k=1}^{\infty}$, $\left\|\bar{x}^k - x^k\right\| \le \delta_k$, $\lim_{k\to\infty}\delta_k = 0$, $h_k > 0$, $\sum_{k=0}^{\infty} h_k\delta_k < \infty$, $\sum_{k=0}^{\infty} h_k^2 < \infty$, $\sum_{k=0}^{\infty} h_k = \infty$, справедлива рівність

$$\lim_{k\to\infty} x^k = x^* \in M^*.$$



**9. Методи штрафів для задач оптимізації з обмеженнями**

За допомогою методу штрафів задачу відшукання мінімізатора $\arg\min_{x\in X} f_0(x)$ функції $f_0 : R^n \to R$ на множині $X$, заданій системою рівнянь та нерівностей

$$X \triangleq \{x \,|\, f_i(x) = 0,\, i \in \Im^0,\; f_i(x) \le 0,\, i \in \Im^-\},$$

заміняють простішою задачею відшукання мінімізатора $\arg\min_x f(x)$ без обмежень на $x$. При цьому функцію $f$ вибирають таким чином, щоб зменшити «похибку по функції» $|\min_x f(x) - \min_{x\in X} f_0(x)|$ і/або зменшити «похибку по аргументу» $\|\arg\min_x f(x) - \arg\min_{x\in X} f_0(x)\|$. Часто допоміжну функцію $f$ вибирають як суму функції $f_0$ із функцією штрафу $h(x)$. Значення штрафу $h(x)$ повинно дорівнювати нулю у точках $x$ на множині $X$ і повинно бути дуже великим, якщо $x$ не належить множині $X$. Наприклад, якщо виберемо функцію $f(x) \equiv f_0(x) + h(x)$ із наступною функцією $h$,

$$h(x) \triangleq \begin{cases} 0, \text{ якщо } x \in X, \\ \infty, \text{ якщо } x \notin X, \end{cases}$$

то, очевидно, будемо мати бажану рівність $\arg\min_x f(x) = \arg\min_{x \in X} f_0(x)$. Проте, для такої функції штрафу $h$ важко обчислити шукане значення $\arg\min_x f(x)$. Тому у практичних розрахунках намагаються вибрати таку функцію штрафу $h$, яка полегшує обчислення мінімізатора $\arg\min_x f(x)$.

Часто функцію $h$ вибирають у вигляді:

$$h(x) \triangleq \sum_{i \in \Im^0} f_i^2(x) + \sum_{i \in \Im^-} (f_i^+(x))^2, \quad f_i^+(x) \triangleq \begin{cases} f_i(x), & f_i(x) > 0, \\ 0, & f_i(x) \le 0, \end{cases}$$

і функцію $f$ вибирають у вигляді суми $f(x) \triangleq f_0(x) + Ch(x)$ із великим числом $C > 0$. У цьому випадку із зростанням числа $C$ мінімізатор $x^*(C) = \arg\min_x f(x)$ наближається до оптимального розв'язку, тобто

$$\lim_{C \to \infty} x^*(C) = x^* = \arg\min_{x \in X} f_0(x).$$

Якщо в якості $h(x)$ вибрати недиференційовану функцію вигляду

$$h_1(x) = C\left( \sum_{i \in \Im^0} |f_i(x)| + \sum_{i \in \Im^-} f_i^+(x) \right),$$

то в багатьох випадках можна отримати оптимальне значення $x^*(C) = \arg\min_{x \in X} f_0(x)$ при використанні порівняно невеликих значень константи $C$. Наприклад, справедливе твердження 1.

**Твердження 11.** Якщо функції $f_0$, $f_i$ – опуклі, то з використанням штрафної функції $h_1(x) = C\left( \sum_{i \in \Im^0} |f_i(x)| + \sum_{i \in \Im^-} f_i^+(x) \right)$, виконується рівність

$$x^*(C) = x^* = \arg\min_{x \in X} f_0(x)$$

при будь-якому значенні $C \ge \max_i u_i$, де $u_i$ – множники Лагранжа для функції $f_0(x) + \sum_i u_i f_i(x)$.

Дійсно, якщо функції $f_0$, $f_i$ – опуклі, то $(x^*, u^*)$– сідлова точка для функції Лагранжа $f_0(x) + \sum_i u_i f_i(x)$, тобто маємо нерівності:

$$f_0(x^*) \le f_0(x) + \sum_i u_i^* f_i(x) \le f_0(x) + \sum_i u_i^* f_i^+(x) \le h_1(x) + f_0(x),$$

які разом з рівністю $h_1(x^*) = 0$ дають $x^*(C) = x^*$.

## 0.5. Прискорені методи оптимізації

### 1. Алгоритм спряжених градієнтів для мінімізації квадратичних функцій

Метод спряжених градієнтів використовують для мінімізації квадратичної функції

$$f(x) = \left(1/2\right)(Ax, x) + (g, x)$$

із заданою додатньо визначеною матрицею $A$ і заданим вектором $g$. Вектори $p^1, p^2, ..., p^n$ називають ***спряженими*** відносно симетричної, додатньо визначеної матриці $A$, якщо для всіх $i \neq j$ виконуються рівності $(Ap^i, p^j) = 0$ і для всіх $i$ виконуються нерівності $(Ap^i, p^i) \neq 0$.

У методі спряжених градієнтів послідовно мінімізують функцію $f(x)$ по кожному напрямку $p^1, p^2, ..., p^n$, тобто, починаючи з довільної точки $x^1$, послідовно обчислюють значення $x^1, x^2, ..., x^n$ за ітераційними формулами:

$$\lambda_k = \arg\min_{\lambda} f(x^k + \lambda p^k), \; x^{k+1} = x^k + \lambda_k p^k, \; k = \overline{1, n}.$$

**Твердження 1.** Метод спряжених градієнтів мінімізує квадратичну функцію $f(x) = \left(1/2\right)(Ax, x) + (g, x)$ за $n$ ітерацій, тобто

$$x^n = x^* = \arg\min_{x}\left((1/2)(Ax, x) + (g, x)\right).$$

Дійсно, якщо ввести скалярний добуток $(x, y)_1 = (Ax, y)$, то метод спряжених градієнтів виявляється еквівалентним мінімізації функції $f(x) = \left(1/2\right)(x, x)_1 + (s, x)_1$ по координатним (ортогональним) напрямкам.

Застосовуючи цю ідею до мінімізації нелінійної функції $f(x)$, будуємо такі наближення до спряжених векторів $p^1, p^2, ..., p^n$:

$$p^1 = -\nabla f(x^1), \; p^{k+1} = \|\nabla f(x^{k+1})\|^2 \; p^k / \|\nabla f(x^k)\|^2 - \nabla f(x^{k+1})$$

(або при $p^1 = -\nabla f(x^1)$, $B_1 = I$, $p^{k+1} = -B_{k+1}\nabla f(x^{k+1})$,

$q^n = \nabla f(x^{n+1}) - \nabla f(x^n)$, $s^n = x^{n+1} - x^n$, $B_{n+1} = B_n - \dfrac{B_n q^n q^{nT} B_n^T}{q^{nT} B_n q^n} + \dfrac{s^n s^{nT}}{s^{nT} q^n}$).

**Твердження 2.** Для сильно опуклої функції з ліпшицевим гесіаном метод:

$$\lambda_k = \arg\min_{\lambda} f(x^k + \lambda p^k), \; x^{k+1} = x^k + \lambda_k p^k, \; k = \overline{1, n},$$

$$p^1 = -\nabla f(x^1), \; p^{k+1} = \|\nabla f(x^{k+1})\|^2 \; p^k / \|\nabla f(x^k)\|^2 - \nabla f(x^{k+1})$$

і метод:

$$\lambda_k = \arg\min_{\lambda} f(x^k + \lambda p^k), \; x^{k+1} = x^k + \lambda_k p^k, \; k = \overline{1, n},$$

$$p^1 = -\nabla f(x^1), \; B_1 = I, \; p^{k+1} = -B_{k+1}\nabla f(x^{k+1}),$$

$$q^n = \nabla f(x^{n+1}) - \nabla f(x^n), s^n = x^{n+1} - x^n,$$

$$B_{n+1} = B_n - \frac{B_n q^n q^{nT} B_n^T}{q^{nT} B_n q^n} + \frac{s^n s^{nT}}{s^{nT} q^n})$$

забезпечують квадратичну швидкість збіжності до єдиної точки мінімуму $x^*$.

Якщо $\lambda_n$ обчислювати наближено, то $p^{n+1}$ рекомендується обчислювати за формулами:

$$p^{n+1} = \left[ \frac{(\nabla f(x^{n+1}), p^n)}{\left\| \nabla f(x^n) \right\|^2} - 1 \right] \nabla f(x^{n+1}) + \frac{\left\| \nabla f(x^{n+1}) \right\|^2}{\left\| \nabla f(x^n) \right\|^2} p^n .$$

Матриці $B_n$ можна обчислювати також за більш загальним формулами:

$$B_{n+1} = B_n - \frac{B_n q^n q^{nT} B_n^T}{q^{nT} B_n q^n} + \frac{s^n s^{nT}}{s^{nT} q^n} + C v^n v^{nT}, \ v^n = \frac{s^n}{q^n s^{nT}} - \frac{B_n q^n}{q^{nT} B_n q^n}, \ C \in R.$$

Дані методи можна інтерпретувати як варіанти квазіньютонівських методів.



**2. Методи другого порядку. Метод Ньютона**

Чим більше відхиляється функція $f$ від її лінійної апроксимації

$$F_1(x) = f(x^n) + (\nabla f(x^n), x - x^n)$$

в околі точки $x^n$, тим меншими можуть бути значення крокових множників $\lambda_n$, при яких градієнтний метод $x^{n+1} = x^n - \lambda_n \nabla f(x^n)$ забезпечує виконання бажаної нерівності $f(x^{n+1}) < f(x^n) - s\lambda_n^2$ і, відповідно, тим більше ітерацій потрібно здійснити для отримання оптимального розв'язку з потрібною точністю. Для прискорення збіжності та підвищення практичної ефективності ітераційних методів використовуються більш точні апроксимації нелінійної функції $f$. Зокрема у наступному методі Ньютона, замість лінійної апроксимації $F_1(x) = f(x^n) + (\nabla f(x^n), x - x^n)$ використовують квадратичну у околі точки $x^n$ апроксимацію $F_2(x) = F_1(x) + (C(x - x^n), x - x^n)$.

У методі Ньютона наступне $(n+1)$-ше наближення $x^{n+1}$ обчислюють як мінімізатор $x^{n+1} = \arg\min_x F_2(x)$ квадратичної апроксимації

$$F_2(x) \equiv f(x^n) + (\nabla f(x^n), x - x^n) + (1/2)(x - x^n)^T \nabla_{xx}^2 f(x^n)(x - x^n)$$

для функції $f$ в околі точки $x^n$. У відповідності із твердженням 1, значення $x^{n+1}$ є розв'язком лінійної системи $\nabla F_2(x) = 0$. Беручи до уваги рівність $\nabla F_2(x) = \nabla f(x^n) + \nabla^2_{xx} f(x^n)(x - x^n) = 0$, значення $x^{n+1}$ обчислюють за *методом Ньютона* як розв'язок $x$ лінійної системи

$$\nabla^2_{xx} f(x^n)(x^{n+1} - x^n) = -\nabla f(x^n).$$

Якщо побудована за методом Ньютона послідовність $\{x^k\}_{k=0}^{\infty}$ *збігається до точки* $x'$, то $\nabla f_0(x') = 0$.

Якщо в околі розв'язку $x^*$ гесіан $\nabla^2_{xx} f(x)$ невироджений і початкове наближення $x^1$ вибрано достатньо близько до точки $x^*$, то метод Ньютона забезпечує квадратичну швидкість збіжності на відміну від геометричної швидкості збіжності градієнтних методів. Однак через необхідність розв'язувати на кожному кроці систему лінійних алгебраїчних рівнянь $\nabla^2_{xx} f(x^n)(x - x^n) = -\nabla f(x^n)$ для обчислення значення $x^{n+1}$, трудомісткість методу Ньютона може суттєво перевищувати трудомісткість обчислення $x^{n+1}$ за градієнтним методом $x^{n+1} = x^n - \lambda_n \nabla f(x^n)$. Інші труднощі реалізації методу Ньютона пов'язані з необхідністю обирати початкове наближення $x^1$ недалеким від розв'язку $x^*$, щоб внаслідок відхилення функції $f$ від її локальної квадратичної апроксимації $F_2$ не отримати небажану нерівність $f(x^{n+1}) > f(x^n)$ і втратити збіжність обчислюваної послідовності до розв'язку $x^*$.



Для забезпечення бажаної нерівності $f(x^{n+1}) < f(x^n)$ при довільно вибраному початковому значенні $x^1$ і для забезпеченням збіжності послідовності $\{x^n\}_{n=1}^{\infty}$ до розв'язку $x^*$, наступне наближення $x^{n+1}$ уточнюють за формулою $x^{n+1} = x^n + \lambda_n(\bar{x}^n - x^n)$ із вибором крокового множника $\lambda_n$ як мінімізатора функції $f(x^n + \lambda(\bar{x}^n - x^n))$ однієї змінної $\lambda$.

Для мінімізації квадратичної функції

$$F_2(x) \equiv f(x^n) + (\nabla f(x^n), x - x^n) + (1/2)(x - x^n)^T \nabla^2_{xx} f(x^n)(x - x^n)$$

можна скористатися методом спряжених градієнтів, описаним вище для мінімізації квадратичної функції $F(x) = \dfrac{1}{2}(Ax, x) + (g, x)$. У даному випадку $A = \nabla^2_{xx} f(x^n)$, $g = \nabla f(x^n) - \nabla^2_{xx} f(x^n) x^n$.

**3. Квазіньютонівські методи**

Щоб уникнути трудомістких у методі Ньютона обчислень, які пов'язані із розв'язуванням на кожній ітерації системи

$$\nabla^2_{xx} f(x^n)(x - x^n) = -\nabla f(x^n),$$

у модифікованих методах Ньютона обчислюють гесіан $\nabla^2_{xx} f(x^n)$ не у кожній точці послідовності $\{x^n\}_{n=1}^{\infty}$, а лише у точках деякої підпослідовності $\{x^{n_i}\}_{i=1}^{\infty}$, залишаючи гесіан без змін на проміжних ітераціях $n_i \le n < n_{i+1}$. В інших випадках пропонується взагалі не обчислювати $\nabla_{xx}^2 f(x^n)$, а замість цього обчислювати $x^{n+1}$ за формулою:

$$x^{n+1} = x^n - \lambda_n B_n \nabla f(x^n),$$

де $B_n$ – деяка послідовність матриць, що в границі наближається до гесіана, $B_n \to \nabla_{xx}^2 f(x^n)$. Для таких модифікованих методів побудовано формули послідовного обчислення значень $\lambda_n$ та $B_n$, які забезпечують квадратичну швидкість збіжності методів $x^{n+1} = x^n - \lambda_n B_n \nabla f(x^n)$. Такі методи називають ***квазіньютонівськими***. Очевидно, до квазіньютонівських методів належить також і метод спряжених градієнтів:

$$\lambda_k = \arg\min_{\lambda} f(x^k + \lambda p^k),\ x^{k+1} = x^k + \lambda_k p^k,\ k = \overline{1,n},$$

$$p^1 = -\nabla f(x^1),\ B_1 = I,\ p^{k+1} = -B_{k+1} \nabla f(x^{k+1}),$$

$$q^n = \nabla f(x^{n+1}) - \nabla f(x^n), s^n = x^{n+1} - x^n,\ B_{n+1} = B_n - \frac{B_n q^n q^{nT} B_n^T}{q^{nT} B_n q^n} + \frac{s^n s^{nT}}{s^{nT} q^n}$$

та його модифікації із наближеними обчисленнями значень $\lambda_n$, векторів

$$p^{n+1} = \left[ \frac{(\nabla f(x^{n+1}), p^n)}{\left\| \nabla f(x^n) \right\|^2} - 1 \right] \nabla f(x^{n+1}) + \frac{\left\| \nabla f(x^{n+1}) \right\|^2}{\left\| \nabla f(x^n) \right\|^2} p^n$$

та матриць $B_n$ за формулами:

$$B_{n+1} = B_n - \frac{B_n q^n q^{nT} B_n^T}{q^{nT} B_n q^n} + \frac{s^n s^{nT}}{s^{nT} q^n} + C v^n v^{nT},\ v^n = \frac{s^n}{q^n s^{nT}} - \frac{B_n q^n}{q^{nT} B_n q^n},\ C \in R.$$

Різні алгоритми реалізації квазіньютонівських методів розроблені також і для тих практично важливих задач оптимізації, у яких аналітичний вираз функції $f$ є невідомим, і тому для обчислення наступного значення $x^{n+1}$ використовується обчислені на попередніх кроках значення функції $f(x^n)$, $f(x^{n-1})$, .... В основі таких методів лежать ті або інші способи обчислення невідомих градієнтів $\nabla f(x^n)$ та гесіанів $\nabla_{xx}^2 f(x^n)$ за допомогою їх різницевих апроксимацій.

#### 4. Прискорені методи оптимізації із розтягуванням простору

У розроблених академіком Н.З. Шором методах з розтягуванням простору на $n$-й ітерації «розтягують простір» в напрямку $z^n$, який є або напрямком субградієнту $g_f(x^n)$ у точці $x^n$, або напрямком різниці двох послідовних субградієнтів. Відповідно на кожному кроці коректується матриця $B_n$ за формулою:

$$B_{n+1} = B_n R_{1/\alpha_{n+1}}(z^n),$$

де $R_{1/\alpha_{n+1}}(z^n)$ – матриця розтягування простору в напрямку $z^n$ з коефіцієнтом розтягування $\alpha_{n+1}$. Елементи $r_{ij}$ матриці $R_{1/\alpha}(z^n)$ обчислюються за формулами:

$$r_{ij} = (\alpha - 1) z_i^n z_j^n, i \neq j, \; r_{ij} = (\alpha - 1) z_i^2 + 1, \; i = j.$$

Отже, для мінімізації функції $f(x)$ методом оптимізації із розтягуванням простору у напрямі субградієнта вибирається довільне початкове значення $x^0 \in R^n$ і початкові одиничні матриці $B_0 = A_0^{-1} = I_n$. На $k$-му кроці основного циклу за обчисленими на попередньому кроці значеннями $x^k$, $B_k = A_k^{-1}$, обчислюється $g_f(x^k)$ (якщо $g_f(x^k) = 0$, то процес зупиняється), обчислюється узагальнений градієнт $\tilde{g}^k$ для функції $\varphi_k(y) = f(B_k y)$ в розтягнутому просторі $y^k = A_k x^k$, $\tilde{g}^k = g_{\varphi_\kappa}(y^k) = B_k^* g_f(x^k)$ і обчислюється:

$$\xi^k = \tilde{g}^k / \left\| \tilde{g}^k \right\|; \quad x^{k+1} = x^k - h_{k+1} B_k \xi^k,$$

$$B_{k+1} = A_{k+1}^{-1} = B_k R_{\beta_{k+1}}(\xi^k), \quad \beta_{k+1} = 1/\alpha_{k+1}.$$

Збіжність цього методу забезпечується вибором коефіцієнтів розтягування $\alpha_k$ і крокових множників $h_k$. Відомо, що за умов існування чисел $N > 0, M > 0$, які в околі $S_d$ розв’язку $\left\| x - x^* \right\| \leq d$ задовольняють умовам:

$$N(f(x) - f(x^*)) \leq (g_f(x), x - x^*) \leq M(f(x) - f(x^*)),$$

при виборі:

$$x^0 \in S_d, \quad h_{k+1} = \frac{2MN}{M+N} \frac{f(x^k) - f(x^*)}{\left\| \tilde{g}_k \right\|}, \; 1 < \alpha_{k+1} \leq \frac{(M+N)}{(M-N)},$$

послідовність $x^k$ задовольняє нерівностям $\left\| A_k (x^k - x^*) \right\| \leq d.$

Отже, мінімізатор $x^*$ належить еліпсоїду $\Phi_k$ із центром в точці $x^k$, а коефіцієнт розтягування визначає відношення об'ємів еліпсоїдів $\Phi_{k+1}$ і $\Phi_k$:

$$\frac{vol(\Phi_{k+1})}{vol(\Phi_k)} = \beta_k = \frac{M-N}{M+N}.$$

Для квадратичної додатньо визначеної функції виконується: $M = N = 2$, а для кусково-лінійної функції, надграфік якої є конусом із вершиною в точці $(x^*, f^*)$, $M = N = 1$. Для цих випадків $\beta_k = \beta = 0$ і тому метод призводить до розв’язку за $k \leq n$ ітерацій.

Це означає, що розв’язок невиродженої системи

$$(a^i, x) + b_i = 0,\ i = \overline{1, n}$$

можна обчислити як мінімізатор функції $f(x) = \max_{i \in [1:n]} \left| (a^i, x) + b_i \right|$ із значенням $\beta_k = 0$. Аналогічно можна знаходити розв'язок системи нелінійних рівнянь $f_i(x) = 0, i = \overline{1, n}$, як мінімізатор функції $f(x) = \max_{i \in [1:n]} \left| f_i(x) \right|$. Якщо на розв'язку $x^*$ функції $f_i(x)$ неперервно диференційовані і якобіан системи не дорівнює нулю, то для довільного $\delta > 0$ знайдеться окіл $S_d(x^*)$, для якого значення $M$ і $N$ визначаються умовами:

$$M = 1 + \delta;\ N = 1 - \delta;\ \beta = \frac{M - N}{M + N} = \delta .$$

Якщо після кожних $n$ ітерацій оновлювати розрахунки, то забезпечується квадратична збіжність при $\beta = 0$.

### 5. $r$-алгоритми мінімізації з розтягуванням простору у напрямку різниці субградієнтів

У $r$-алгоритмі для мінімізації опуклої функції $f(x)$ вибирають довільне початкове наближення $x^0$, $B_0 = I_n$ і обчислюють $x^1 = x^0 - h_0 g_f(x^0)$, де кроковий множник $h_0$ вибирається із умови існування у точці $x^1$ субградієнта $g_f(x^1)$, який задовольняє нерівності $(g_f(x^1), g_f(x^0)) \le 0$.

На $k$-му кроці обчислюють: субградієнт $g_f(x^k)$, вектор $r^k = B_k^T (g_f(x^k) - g_f(x^{k-1}))$ різниці двох послідовних субградієнтів в перетвореному просторі, нормований вектор $\xi^k = r^k / \left\| r^k \right\|$, значення $\beta_k$ обернене до коефіцієнта розтягування простору $\alpha_k$, оператор $R_{\beta_k}(\xi^k)$ розтягування простору та оновлену матрицю $B_{k+1} = B_k R_{\beta_k}(\xi^k)$, субградієнт $\tilde{g}^k = B_{k+1}^T g_f(x^k)$ функції $\varphi_{k+1} = f(B_{k+1} y)$ у точці $y^{k+1} = A_{k+1} x^k$ $(B_{k+1} = A_{k+1}^{-1})$ та нове значення

$$x^{k+1} = x^k - h_k B_{k+1} \tilde{g}^k / \left\| \tilde{g}^k \right\| .$$

Значення $h_k \ge \arg\min_{h > 0} f(x^k - h B_{k+1} \tilde{g}^k / \left\| \tilde{g}^k \right\|)$ вибирають із умови, щоб напрямок субградієнта в точці $x^{k+1}$ утворював гострий кут із напрямком спуску із точки $x^k$.

Для задач гладкої оптимізації $r$-алгоритм є формально близьким до наступних алгоритмів квазіньютонівського типу із змінною метрикою.

## 0.6. Умови оптимальності для задач оптимізації із обмеженнями

### 1. Теорема Каруша–Куна–Таккера для задачі опуклої оптимізації

Розв’язком $x^*$ задачі опуклої оптимізації називають мінімізатор опуклої функції $f_0$ на множині $X=\{x \mid f_i(x)\le 0,\ i=\overline{1,m},\ x\in Y\}$, яка задана опуклими функціями $f_i$ та опуклою множиною $Y$.

Необхідні і достатні умови оптимальності розв’язку задачі опуклої оптимізації дає теорема Каруша–Куна–Таккера.

**Теорема Каруша–Куна–Таккера.** *Якщо всі функції $f_i$ і множина $Y$ є опуклими і якщо множина $X^1=\{x \mid f_i(x)\le 0,\ i=\overline{1,m},\ x\in Y\}$ задовольняє умову Слейтера, тобто $\exists x^1\in Y, \forall i\in\overline{1,m},\ f_i(x^1)<0$, то необхідною і достатньою умовою оптимальності точки $\overline{x}$ для задачі мінімізації опуклої функції $f_0$ на множині $X^1$ є існування розв’язку $\overline{u}_1,\ldots,\overline{u}_m$ системи:*

$$f_0(\overline{x})+\sum_{j=1}^{m}u_j f_j(\overline{x})\le f_0(x)+\sum_{j=1}^{m}u_j f_j(x), \forall x\in Y,$$
$$u_j f_j(\overline{x})=0,\quad u_j\ge 0,\quad \forall j\in\overline{1,m}.$$

*Означення 1.* Значення $\overline{u}_1,\ldots,\overline{u}_m$ називають ***множниками Лагранжа***, а функцію $\varphi(x,u)\triangleq f_0(x)+\sum_{j=1}^{m}u_j f_j(x)$ називають ***функцією Лагранжа***. Пара $(\overline{x},\overline{u})$ є сідловою точкою функції $\varphi$ на множині $Y\times R_+^n$, $\overline{x}\in Y, \overline{u}\ge 0$, тобто, задовольняє нерівностям:

$$\varphi(\overline{x},u)\le\varphi(\overline{x},\overline{u})\le\varphi(x,\overline{u}), \forall x\in Y, u\ge 0.$$

Отже, теорему Каруша–Куна–Таккера можемо сформулювати так: *якщо виконується умова Слейтера, то необхідною і достатньою умовою оптимальності точки $\overline{x}$ в задачі опуклої оптимізації є існування такого вектора $\overline{u}\ge 0$, при якому пара $(\overline{x},\overline{u})$ є сідловою точкою функції Лагранжа $f_0(x)+\sum_{j=1}^{m}u_j f_j(x)$ на множині $Y\times R_+^n$.*

Для загального випадку недиференційованих опуклих функцій справедливе твердження 1.

**Твердження 1**. Якщо в задачі опуклої оптимізації виконується умова Слейтера, то необхідною і достатньою умовою оптимальності точки $\overline{x}$ є існування таких субградієнтів $\hat{\nabla} f_i(\overline{x})\in M_i(\overline{x})$ і $u$, які задовольняють умовам: $\hat{\nabla} f_0(\overline{x})=\sum_{i=1}^{m}u_i\hat{\nabla} f_i(\overline{x})$, $u_i f_i(\overline{x})=0,\ u_i\ge 0,\ i=\overline{1,m}$.

**2. Метод внутрішньої точки в лінійному програмуванні**

***Задача лінійного програмування (ЗЛП)***. Знайти вектор $x=(x_1,x_2,...,x_n)^T$, який максимізує функцію $f(x)=\sum_{i=1}^{n} c_i x_i$ і задовольняє нерівностям $\sum_{j=1}^{n} a_{ij}x_j \le b_i$, $i=\overline{1,m}.$

Розв'язок ЗЛП можна обчислити за таким елементарним алгоритмом:

I. Вибрати довільний початковий вектор $x^1$ і покласти $k$=1.

II. II. Обчислити $j^* = \arg\max\limits_{j\in[1;m]}((a^j,x^k)-b_j)$, $a^j=\left(a_{j1},a_{j2},\dots,a_{jn}\right)$.

III. Якщо $(a^{j^*},x^k)>b_{j^*}$, то покласти $x^{k+1}=x^k-a^{j^*}\dfrac{(a^{j^*},x^k)-b_j}{(a^{j^*},a^{j^*})}$.

IV. Якщо $(a^{j^*},x^k)\le b_{j^*}$, то покласти $x^{k+1}=x^k+\lambda_k c.$

V. Якщо $k<K$, то перейти на крок II; інакше зупинити обчислення.

Відомо, що для будь-якої константи $M>0$ при виборі кроку $\lambda_k=\dfrac{M}{k}$ границею послідовності $\{x^k\}_{k=1}^{\infty}$ є розв'язок сформульованої вище задачі лінійного програмування.

***Двоїстою задачею*** називають задачу мінімізація лінійної функції

$$b_1y_1+b_2y_2+\cdots+b_my_m$$

при лінійних обмеженнях:

$$\begin{aligned}
&a_{11}y_1+a_{21}y_2+\cdots+a_{m1}y_m\ge c_1,\\
&a_{12}y_1+a_{22}y_2+\cdots+a_{m2}y_m\ge c_2,\\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\\
&a_{1n}y_1+a_{2n}y_2+\cdots+a_{mn}y_m\ge c_n.\\
&\quad y_1\ge 0, y_2\ge 0,\dots,y_m\ge 0.
\end{aligned}$$

Беручи до уваги, що при будь-якому допустимому значенні $x$ і при будь-якому допустимому значенні $y$ виконується нерівність $(c,x)\le(b,y)$, бачимо, що будь-які вектори $x^0,y^0$, для яких виконується рівність $(c,x^0)=(b,y^0)$, є оптимальними розв'язками. Дійсно, для всіх допустимих $x$ та $y$ виконуються нерівності $(c,x)\le(c,x^0)$, $(b,y^0)\le(b,y)$ і допустимі розв'язки $x$ та $y$ є оптимальними, якщо для кожного $j$ виконується рівність

$$(a_{1j}y_1+a_{2j}y_2+\cdots+a_{mj}y_m-c_j)x_j=0,\quad j=\overline{1,n}.$$

Для розв'язування задач лінійного програмування великої розмірності використовують *метод внутрішньої точки,* при якому одночасно розв'язуються пряма і двоїста ЗЛП.

Очевидно, будь-яку задачу лінійного програмування можна звести (за допомогою введення додаткових змінних) до задачі лінійного програмування в стандартній векторно-матричній формі:

$$\text{максимізувати } c^{\mathrm{T}} x$$

$$\text{при обмеженнях: } Ax = b,\ x \geq 0.$$

*Двоїстою є задача*: мінімізувати $b^{\mathrm{T}} y$

$$\text{при обмеженнях: } A^{\mathrm{T}} y + z = c, z \geq 0.$$

Із теореми Каруша–Куна–Таккера випливає, що роз'язком прямої і двоїстої задачі є розв'язок нелінійної системи:

$$Ax - b = 0, A^{\mathrm{T}} y + z - c = 0, x \geq 0, z \geq 0,$$

$$ZXe = 0,$$

де елементами $\bar{z}_{ij}$ матриці $Z \in R^{n\times n}$ є $\bar{z}_{ij} = \delta_{ij} z_i$, а елементами $\bar{x}_{ij}$ матриці $X \in R^{n\times n}$ є $\bar{x}_{ij} = \delta_{ij} x_i$ ($\delta_{ij} = 0$ якщо $i \neq j$, $\delta_{ii} = 1$), $e = (1,1,\ldots,1)^T$, який можна обчислити за методом Ньютона. Для цього спочатку вибирають довільні допустимі початкові значення $(x^1, y^1, z^1)$:

$$Ax^1 = b, x^1 > 0, A^{\mathrm{T}} y^1 + z^1 = c, z^1 > 0,$$

а далі на кожній ($k$-й) ітерації обчислюють наближений розв'язок $(\delta x, \delta y, \delta z)$ нелінійної системи:

$$A\left(x^k + \delta x\right) - b = 0,$$

$$A^T\left(y^k + \delta y\right) + z^k + \delta z - c = 0,$$

$$\left(x_i^k + \delta x_i\right)\left(z_i^k + \delta z_i\right) = 0,\ i = \overline{1,n},$$

$$x^k + \delta x \geq 0, z^k + \delta z \geq 0$$

як розв'язок лінеаризованої системи:

$$A \cdot \delta x = r_p, A^T \cdot \delta y + \delta z = r_d,\ Z \cdot \delta x + X \cdot \delta z = r_a,$$

з відомими значеннями:

$$r_p = b - Ax^k, r_d = c - z^k - A^T y^k, r_a = -XZe.$$

Виконавши прості перетворення отримаємо:

$$A\left(X^{-1}Z\right)^{-1} A^T \cdot \delta y = r_p + \left(X^{-1}Z\right)^{-1}(r_d - X^{-1} r_a),$$

$$\delta x = \left(X^{-1}Z\right)^{-1} A^T \cdot \delta y - \left(X^{-1}Z\right)^{-1}(r_d - X^{-1} r_a), \delta z = r_d - A^T \cdot \delta y.$$

Для обчислення уточненого допустимого значення $\left(x^{k+1}, y^{k+1}, z^{k+1}\right)$ перераховують розв'язок лінеаризованої системи із підкоректованим значенням $r_a$.

**3. Необхідні умови оптимальності для задач оптимізації з диференційованими функціями обмежень**

**Твердження 2**. Якщо $\overline{x}^*$ є точкою локального мінімуму функції $f_0$ на множині $X=\{x| f_i(x)\le 0, i\in\Im^-, f_i(x)=0, i\in\Im^0\}$, то існує нетривіальний розв'язок $u_0, u_i, i\in\Im^-, v_i, i\in\Im^-$ системи:

$$u_0\nabla f_0(\overline{x}^*)+\sum_{i\in\Im^-}u_i\nabla f_i(\overline{x}^*)+\sum_{i\in\Im^0}v_i\nabla f_i(\overline{x}^*)=0,$$

$$u_0\ge 0, u_i\ge 0, i\in\Im^-, u_i f_i(\overline{x}^*)=0, i\in\Im^-.$$

*Означення 2.* Точку локального мінімуму $\overline{x}^*$ називають ***регулярною точкою мінімуму***, якщо є лінійно незалежними градієнти $\nabla f_i(\overline{x}^*)$ тих функцій $f_i$, які у точці $\overline{x}^*$ дорівнюють нулю, $f_i(\overline{x}^*)=0$.

Наслідок 1. Якщо точка $\overline{x}^*$ є регулярною точкою мінімуму, то можна покласти $u_0=1$; решта множників $u_i$, $v_i$ визначаються однозначно.

Наступне твердження дає необхідні умови оптимальності другого порядку.

**Твердження 3**. Якщо функції $f_i$, $i\in\Im^0\cup\Im^-\cup\{0\}$, двічі неперервно диференційовані і точка $\overline{x}$ є регулярною точкою мінімуму, то існують такі числа $u_i$, $i\in\Im^-$, $v_i$, $i\in\Im^0$, що для кожного допустимого напрямку $h$, виконуються співвідношення:

$$(\nabla f_i(\overline{x}),h)\le 0,\ i\in\{i| f_i(\overline{x})=0,\ u_i\le 0,\ i\in\Im^-\},$$

$$(\nabla f_i(\overline{x}),h)=0,\ i\in\{i|u_i>0,\ i\in\Im^-\}\bigcup\Im^0\ ,$$

$$\nabla_x\varphi(\overline{x},u,v)=0,\ u_i\ge 0,\ u_i f_i(\overline{x})=0,\ i\in\Im^-,\left(\nabla_{xx}{}^2\varphi(\overline{x},u,v)h,h\right)\ge 0,$$

де $\varphi(x,u,v)\triangleq f_0(x)+\sum_{i\in\Im^-}u_i f_i(x)+\sum_{i\in\Im^0}v_i f_i(x),\ u=\{u_i\}_{i\in\Im^-},\ v=\{v_i\}_{i\in\Im^0}$.

## 0.7. Методи оптимізації в задачах з обмеженнями

**1. Методи штрафів для задач оптимізації з обмеженнями**

За допомогою методу штрафів задачу відшукання мінімізатора $\arg\min_{x\in X} f_0(x)$ функції $f_0: R^n\to R$ на множині $X$, заданій системою рівнянь та нерівностей

$$X\triangleq\{x|f_i(x)=0,\ i\in\Im^0,\ f_i(x)\le 0,\ i\in\Im^-\},$$

заміняють простішою задачею відшукання мінімізатора $\arg\min_x f(x)$ без обмежень на $x$. При цьому функцію $f$ вибирають таким чином, щоб зменшити «похибку по функції» $|\min_x f(x) - \min_{x\in X} f_0(x)|$ і (або) зменшити «похибку по аргументу» $\|\arg\min_x f(x) - \arg\min_{x\in X} f_0(x)\|$. Часто допоміжну функцію $f$ вибирають як суму функції $f_0$ із функцією штрафу $h(x)$. Значення штрафу $h(x)$ повинно дорівнювати нулю у точках $x$ на множині $X$ і повинно бути дуже великим, якщо $x$ не належить множині $X$. Наприклад, якщо виберемо функцію $f(x) \equiv f_0(x) + h(x)$ із наступною функцією $h$:

$$h(x) \triangleq \begin{cases} 0, \text{ якщо } x \in X, \\ \infty, \text{ якщо } x \notin X, \end{cases}$$

то, очевидно, будемо мати бажану рівність $\arg\min_x f(x) = \arg\min_{x\in X} f_0(x)$. Проте, для такої функції штрафу $h$ важко обчислити шукане значення $\arg\min_x f(x)$. Тому при практичних розрахунках намагаються вибрати таку функцію штрафу $h$, яка полегшує обчислення мінімізатора $\arg\min_x f(x)$.

Часто функцію $h$ вибирають у вигляді

$$h(x) \triangleq \sum_{i\in\Im^0} f_i^2(x) + \sum_{i\in\Im^-} (f_i^+(x))^2,$$

$$f_i^+(x) \triangleq \begin{cases} f_i(x), & \text{якщо } f_i(x) > 0, \\ 0, & \text{якщо } f_i(x) \le 0, \end{cases}$$

і функцію $f$ вибирають у вигляді суми $f(x) \triangleq f_0(x) + Ch(x)$ із великим числом $C > 0$. У цьому випадку із зростанням числа $C$ мінімізатор $x^*(C) = \arg\min_x f(x)$ наближається до оптимального розв'язку, тобто

$$\lim_{C\to\infty} x^*(C) = x^* = \arg\min_{x\in X} f_0(x).$$

Якщо в якості $h(x)$ вибрати недиференційовану функцію вигляду

$$h_1(x) = C\left(\sum_{i\in\Im^0} |f_i(x)| + \sum_{i\in\Im^-} f_i^+(x)\right),$$

то в багатьох випадках можна отримати оптимальне значення $x^*(C) = \arg\min_{x\in X} f_0(x)$ при використанні порівняно невеликих значень константи $C$. Наприклад, справедливе твердження 1.

**Твердження 1.** Якщо функції $f_0$, $f_i$ – опуклі, то з використанням штрафної функції $h_1(x) = C\left(\sum_{i\in\Im^0} |f_i(x)| + \sum_{i\in\Im^-} f_i^+(x)\right)$, виконується рівність

$$x^*(C) = x^* = \arg\min_{x \in X} f_0(x)$$

при будь-якому значенні $C \geq \left\{ \max_{i \in \mathfrak{I}^0} |u_i|, \max_{i \in \mathfrak{I}^-} |v_i| \right\}$, де $u_i$, $i \in \mathfrak{I}^0$, $v_i$, $i \in \mathfrak{I}^0$ – множники Лагранжа для функції $f_0(x) + \sum_{i \in \mathfrak{I}^0} u_i f_i(x) + \sum_{i \in \mathfrak{I}^-} v_i f_i(x)$.

**2. Алгоритми двоетапної оптимізації**

Розглянемо задачу вибору на першому етапі частини $x^1$ від вектора $x = (x^1, x^2)$, та вибору другої частини $x^2$ на другому етапі, на якому стає відомою додаткова інформація щодо уточнення множини невизначеності $P(x^1)$ параметра $p$ у функції $f(x^1, x^2, p)$.

У даному випадку оптимальне значення $x^1$ визначається як мінімізатор функції $\max_{p \in P} f(x^1, \psi(x^1), p)$, де

$$\psi(x^1) \triangleq \arg \min_{x_2 | (x^1, x^2) \in X} \max_{p \in P(x^1)} f(x^1, x^2, p).$$

В загальному випадку двоетапної оптимізації функцію корисності $f(x^1, x^2, p)$ вибирають з урахуванням витрат як на отримання додаткової інформації щодо уточнення множини невизначеності $P(x^1)$, так і витрат пов'язаних з відстрочкою вибору $x^2$.

Практично важливою задачею двоетапної стохастичної оптимізації є задача вибору на першому етапі значення $x^1$ із допустимої множини $X^1 \subset R^{n_1}$ і вибору на другому етапі значення $x^2$ із допустимої множини $X^2 \subset R^{n_2}$ після того як стане відомим значення $p$, з метою максимізувати математичне сподівання $M_p f(x^1, x^2, p)$ при обмеженнях $g_i(x^1, x^2, p) \leq 0$, $i = \overline{1, m}$. Якщо на першому етапі вибір значення $x^1$ здійснюється при невідомому значенні випадкової величини $p$, а вибір значення $x^2$ здійснюється на другому етапі після того, коли стала відомою реалізація величини $p$, то оптимальним розв'язком даної задачі двоетапної стохастичної оптимізації є:

$$\bar{x}^2(x^1, p) = \arg \max_{x^2 \in G(x^1, p)} f(x^1, x^2, p),$$

$$G(x^1, p) = \{x^2 \mid x^2 \in X^2, \ g_i(x^1, x^2, p) \leq 0, \ i = \overline{1, m}\},$$

$$x^1 = \arg \max_{x^1 \in X^1} M_p f(x^1, \bar{x}^2(x^1, p), p).$$

## 0.8. Методи оптимізації в задачах варіаційного числення та оптимального керування

### 1. Задачі варіаційного числення

У задачах варіаційного числення потрібно знаходити функцію $x^* = \arg\min_{x\in X} f(x)$, яка мінімізує значення $f(x)$ заданого функціоналу $f : X \to R$ на множині $X$ допустимих функцій $x$.

У прикладних задачах допустимою множиною $X$ буває множина поліномів, множина сплайнів, функції скінченних елементів, тригонометричні функції ряду Фур'є тощо. В узагальнених задачах множиною $X$ часто є гільбертовий простір функцій $x, y : [a;b] \to R^n$ із скалярним добутком

$$(x, y) \triangleq \int_a^b (x(t), y(t))dt \equiv \int_a^b \sum_{i=1}^n x_i(t) y_i(t)dt \,.$$

Якщо існує така функція $z$, що для всіх функцій $x$ із околу функції $\bar{x}$ виконується асимптотична рівність

$$f\left(x\right) = f\left(\bar{x}\right) + \left(z, x - \bar{x}\right) + o\left(\left\|x - \bar{x}\right\|\right), \tag{0.3}$$

то функціонал $f$ називають ***диференційованим у точці*** $\bar{x}$ функціонального простору $X$, а функцію $z$ називають ***градієнтом*** функціоналу $f$ у точці $\bar{x}$. Градієнт функціоналу $f$ у точці $\bar{x}$ позначають символом $\nabla f(\bar{x})$ або $f'(\bar{x})$.

Для практичного розв'язування задач варіаційного числення часто використовується наступне твердження: *якщо у точці* $x^* = \arg\min_x f(x)$ *існує градієнт* $\nabla f(x^*)$*, то*

$$\nabla f(x^*) = 0 \,. \tag{0.4}$$

Для доведення рівності (0.4) запишемо рівність (0.3) для точки $x = x^* - \lambda \nabla f(x^*)$ із околу точки $x^*$ при малих значеннях $\lambda$:

$$f\left(x^* - \lambda\nabla f(x^*)\right) = f\left(x^*\right) + \left(\nabla f(x^*), (-\lambda\nabla f(x^*))\right) + o\left(\left\|\lambda\nabla f(x^*)\right\|\right).$$

Якщо б виконувалася нерівність $\nabla f(x^*) \neq 0$, то із існування границі $\lim_{\lambda\to 0} o\left(\left\|\lambda\nabla f(x^*)\right\|\right) / \lambda = 0$ випливало б існування числа $\lambda_1 > 0$, для якого виконуються нерівності:

$$o\left(\left\|\lambda_1\nabla f(x^*)\right\|\right) / \lambda_1 < \left(\nabla f(x^*), \nabla f(x^*)\right) / 2 \,,$$

$$f\left(x^* - \lambda_1\nabla f(x^*)\right) = f\left(x^*\right) - \left(\nabla f(x^*), \lambda_1\nabla f(x^*)\right) + o\left(\left\|\lambda_1\nabla f(x^*)\right\|\right) <$$

$$< f\left(x^*\right) - \left(\nabla f(x^*), \lambda_1\nabla f(x^*)\right) + \lambda_1\left(\nabla f(x^*), \nabla f(x^*)\right) / 2 <$$

$$< f\left(x^*\right) - \left(\nabla f(x^*), \lambda_1 \nabla f(x^*)\right)/2 < f\left(x^*\right).$$

Отримана нерівність $f\left(x^* - \lambda_1 \nabla f(x^*)\right) < f(x^*)$ суперечить рівності $x^* = \arg\min_x f(x)$. Це означає, що нерівність $\nabla f(x^*) \neq 0$ є неможливою і цим доведено рівність (0.4).

Наведемо приклади типових задач варіаційного числення, які розв'язуються за допомогою рівності (0.4).

З а д а ч а 1. Знайти функцію $x:[a;b] \to R$, яка задовольняє крайовим умовам $x(a) = x^1, x(b) = x^2$ і мінімізує значення функціоналу $f(x) \triangleq \int_a^b w(x'(t), x(t), t)dt$ для заданої функції $w: R^3 \to R$ (через $x'$ позначено похідну від функції $x$).

*Розв'язування.* Знаходимо градієнт $\nabla f(x)$ функціоналу $\int_a^b w(x'(t), x(t), t)dt$ як функцію $z$, яка для всіх варіацій $\Delta$ задовольняє еквівалентній до (0.3) асимптотичній рівності

$$f\left(x+\Delta\right) = f\left(x\right) + \left(z, \Delta\right) + o\left(\|\Delta\|\right), \tag{0.5}$$

тобто рівності

$$\int_a^b w(x'(t)+\Delta'(t), x(t)+\Delta(t), t)dt =$$

$$= \int_a^b w(x'(t), x(t), t)dt + \int_a^b (z(t), \Delta(t))dt + o(\|\Delta\|).$$

Із цієї рівності для диференційованої функції $w$ з похідними $w'_x, w'_{x'}$ випливає, що градієнт $\nabla f(x) = z$ є функцією

$$z(t) \equiv w'_x(x'(t), x(t), t) - \frac{d}{dt} w'_{x'}(x'(t), x(t), t), \ t \in [a;b].$$

Згідно рівняння (0.2) розв'язок задачі 1 є розв'язком диференціального рівняння (рівняння Ейлера)

$$w'_x(x'(t), x(t), t) - \frac{d}{dt} w'_{x'}(x'(t), x(t), t) = 0, \ t \in [a;b]$$

із крайовими умовам $x(a) = x^1$, $x(b) = x^2$.

З а д а ч а 2. Знайти функцію $x:[a;b] \to R$, яка мінімізує функціонал

$$f(x) \triangleq \int_a^b w(x(t), x'(t), \ldots, x^{(n)}(t), t)dt$$

при обмеженнях:

$$x^{(i)}(a) = x_1^{(i)}, \ x^{(i)}(b) = x_2^{(i)}, \ i = \overline{0, n-1} \quad \left(x^{(0)} \triangleq x, \ x^{(1)} \triangleq x'\right).$$

*Розв'язування.* У випадку диференційованої функції $w$ із асимптотичної рівності (0.5) на інтервалі часу $t \in [a;b]$ знаходимо, що для функціоналу $f(x) \triangleq \int_a^b w(x(t), x'(t), ..., x^{(n)}(t), t)dt$ градієнт $\nabla f(x) = z$, де

$$z(t) = w'_x(x(t), x'(t), ..., x^{(n)}(t), t) + \sum_{i=1}^{n} (-1)^i \frac{d^i}{dt^i} w'_{x^{(i)}}(x(t), x'(t), ..., x^{(n)}(t), t),$$

і тому розв'язок задачі 2 є, згідно (0.4), розв'язком диференціального рівняння (рівняння Ейлера-Пуассона)

$$w'_x(x(t), x'(t), ..., x^{(n)}(t), t) + \sum_{i=1}^{n} (-1)^i \frac{d^i}{dt} w'_{x^{(i)}}(x(t), x'(t), ..., x^{(n)}(t), t) = 0$$

із крайовими умовами $x^{(i)}(a) = x_1^{(i)}$, $x^{(i)}(b) = x_2^{(i)}$, $i = \overline{0, n-1}$.

З а д а ч а 3. Знайти вектор-функцію $x : [a;b] \to R^n$,

$$x(t) \triangleq (x_1(t), x_2(t), ..., x_n(t)), \; x(a) = x^1, x(b) = x^2,$$

яка мінімізує функціонал

$$f(x) \triangleq \int_a^b w(x_1(t), x_2(t), ..., x_n(t), x'_1(t), x'_2(t), ..., x'_n(t), t)dt.$$

*Розв'язування.* Із рівності (0.5) знаходимо, що градієнтом $\nabla f(x)$ даного функціоналу є вектор-функція $z : [a;b] \to R^n$, визначена рівнянням

$$z(t) \equiv \nabla_x w(x(t), x'(x), t) - \frac{d}{dt} \nabla_{x'} w(x(t), x'(x), t).$$

Згідно (0.4), розв'язок $x$ задачі 3 є розв'язком системи диференціальних рівнянь (рівнянь Ейлера-Лагранжа)

$$\nabla_x w(x(t), x'(x), t) - \frac{d}{dt} \nabla_{x'} w(x(t), x'(x), t) = 0$$

з крайовими умовами $x(a) = x^1$, $x(b) = x^2$.

З а д а ч а 4. Знайти вектор-функцію $x(t) = (x_1(t), x_2(t), ..., x_n(t))$, $t \in [a;b]$, яка для заданих функціоналів

$$f_i(x) \triangleq \int_a^b w_i(x_1(t), x_2(t), ..., x_n(t), x'_1(t), x'_2(t), ..., x'_n(t), t)dt, \; i = \overline{0, m},$$

мінімізує значення $f_0(x)$ і задовольняє рівнянням $f_i(x) = 0$, $i = \overline{1, m}$.

*Розв'язування.* Беручи до уваги, що, по-перше, розв'язок $x$ задачі 4 в умовах лінійно незалежних градієнтів $\nabla f_i(x)$, $i = \overline{1, m}$ є розв'язком системи Лагранжа

$$\nabla f_0(x) = \sum_{i=1}^{m} \lambda_i \nabla f_i(x), \; f_i(x) = 0, \; i = \overline{1, m}$$

і, по-друге, для диференційованої функції $w_i$ градієнт $\nabla f_i(x)$ є вектор-функцією $z^i : [a;b] \to R^n$, яка визначається рівнянням

$$z^i(t) \equiv \nabla_x w_i\left(x(t), x'(t), t\right) - \frac{d}{dt}\nabla_{x'} w_i\left(x(t), x'(t), t\right),\ i = \overline{1,m},$$

робимо висновок, що розв'язок $x$ задачі 4 є розв'язком системи інтегро-диференціальних рівнянь:

$$\nabla_x w_0\left(x(t), x'(t), t\right) - \frac{d}{dt}\nabla_{x'} w_0\left(x(t), x'(t), t\right) =$$

$$= \sum_{i=1}^{m} \lambda_i \left( \nabla_x w_i\left(x(t), x'(t), t\right) - \frac{d}{dt}\nabla_{x'} w_i\left(x(t), x'(t), t\right) \right),$$

$$\int_a^b w_i\left(x_1(t), x_2(t), ..., x_n(t), x_1'(t), x_2'(t), ..., x_n'(t), t\right) dt = 0,\ i = \overline{1,m}.$$

Задача 5. Знайти вектор-функцію $x(t) = (x_1(t), x_2(t), ..., x_n(t))$, $t \in [a;b]$, яка для заданих функціоналів

$$f_i(x) \triangleq \int_a^b w_i(x_1(t), x_2(t), ..., x_n(t), x_1'(t), x_2'(t), ..., x_n'(t), t) dt,\ i = \overline{0,m},$$

мінімізує значення $f_0(x)$ і задовольняє нерівностям $f_i(x) \le 0$, $i = \overline{1,m}$.

*Розв'язування.* Беручи до уваги, що розв'язок $x$ задачі 5 в умовах лінійно-незалежних градієнтів

$$\nabla f_i(x) \equiv \nabla_x w_i(x(t), x'(x), t) - \frac{d}{dt}\nabla_{x'} w_i(x(t), x'(x), t),\ i = \overline{1,m}$$

є розв'язком системи Каруша-Лагранжа

$$\nabla f_0(x) = \sum_{i=1}^{m} \lambda_i \nabla f_i(x),\ \lambda_i f_i(x) = 0,\ i = \overline{1,m},$$

робимо висновок, що розв'язок $x$ задачі 5 є розв'язком системи інтегро-диференціальних рівнянь:

$$\nabla_x w_0\left(x(t), x'(t), t\right) - \frac{d}{dt}\nabla_{x'} w_0\left(x(t), x'(t), t\right) =$$

$$= \sum_{i=1}^{m} \lambda_i \left( \nabla_x w_i\left(x(t), x'(t), t\right) - \frac{d}{dt}\nabla_{x'} w_i\left(x(t), x'(t), t\right) \right),$$

$$\lambda_i \int_a^b w_i(x_1(t), x_2(t), ..., x_n(t), x_1'(t), x_2'(t), ..., x_n'(t), t) dt = 0,\ i = \overline{1,m}.$$

Задача 6. Знайти вектор-функції $x(t), u(t)$, $t \in [t_0;T]$, які для узагальнених функціоналів [1]:

$$f_i(x,u) \triangleq B_i^0(x(t_0), x(T), u, \int_{t_0}^{T} B_i^1\left(\dot{x}(t), x(t), u(t), t,\right.$$

[1] Бейко І.В. Розвиток методів розв'язуючих та асимптотично-розв'язуючих операторів для побудови оптимальних та асимптотично-оптимальних математичних моделей / І.В. Бейко // Вісник Київського університету. – Серія: Кібернетика. – 2002. – Вип.3. – С. 10-15.

$$\int_{t_0}^{t} B_i^2(\dot{x}(s), x(s), u(s), s, \dot{x}(t), x(t), u(t), t)ds)dt)\,,\ i=\overline{0,m}\,,$$

мінімізують значення $f_0(x,u)$ і задовольняють нерівностям $f_i(x,u) \le 0$, $i=\overline{1,m}$, та алгебро-інтегро-диференціальним зв'язкам

$$A(x,u) \triangleq f^0\left( x(t_0), u, \dot{x}(t), x(t), u(t), t, \int\limits_{D(t,u,x)} f^1\Big(x(s), u(s), s, x(t), u(t), t\Big)ds \right) = 0\,.$$

*Розв'язування.* В умовах неперервних по $u$ функцій $f^0$, $f^1$, $B_i^0$, $B_i^1$, $B_i^2$ з ліпшицевими похідними по змінній $x$ для широкого класу множин $D(t,u,x)$ побудовано асимптотично-розв'язуючий оператор $\Psi_i(u,\bar{u})$, який для всіх $u$ із околу функції $\bar{u} \in \Omega$ задовольняє асимптотичну рівність

$$f_i(x(u),u) = \Psi_i(u,\bar{u}) + O\,\|\, u - \bar{u}\,\|^2$$

на множині розв'язків $x = x(u)$ системи:

$$A(x,u) = 0\,,\ f_i(x,u) = 0\,,\ i = \overline{1,m}\,.$$

В умовах існування градієнта $\nabla\Psi_i(u,\bar{u})$ функціоналу $\Psi_i(u,\bar{u})$ у точці $\bar{u}$ існує градієнта $\nabla f_i(x(\bar{u}),\bar{u})$ функціоналу $f_i(x(u),u)$, який визначається рівністю $\nabla f_i(x(\bar{u}),\bar{u}) = \nabla\Psi_i(\bar{u})$. За допомогою градієнтів $\nabla f_i(x(\bar{u}),\bar{u})$, $i=\overline{0,m}$ обчислення екстремалі $(u^*,x^*)$ функціоналу $f_0(x,u)$ при зв'язках $A(x,u)=0$, $f_i(x,u) \le 0$, $i=\overline{1,m}$ здійснюється в опуклих задачах із використанням градієнтного методу:

$$\alpha_k = \max_{i=\overline{1,m}} f_i(x(u_k),u_k)\,,$$

$$z_k = \begin{cases} \nabla\Psi_i(u_k,u_k), \text{ якщо } f_i(x(u_k),u_k) = \alpha_k,\ \alpha_k > 0, \\ \nabla\Psi_0(u_k,u_k), \text{ якщо } \alpha_k \le 0, \end{cases}$$

$$u_{k+1} = u_k - \lambda_k z_k / \|\, z_k \,\|,\ \lambda_k > 0,\ \lim_{k\to\infty}\lambda_k = 0,\ \sum_{k=1}^{\infty}\lambda_k = \infty.$$

Задача 7. На множині $D \subset [t_1,t_2]\times[s_1^1,s_2^1]\times[s_1^2,s_2^2]\times\ldots\times[s_1^{n_s},s_2^{n_s}]$ знайти вектор-функції $x: D \to R^{n_1}$ і $u: D \to R^{n_2}$, які мінімізують функціонал

$$\bar{f}_0(x,u) \triangleq \iint\limits_D f_0(t,s,u(t,s),F^{f_0}(x,t,s))dtds$$

і на заданих $j$-параметричних підмножинах $D_j^i(x,u) \subset D$, $j=\overline{1,m+1}$, $i=\overline{1,i_j}$ задовольняють умовам:

$$\bar{f}_{ij}^k(t,s,x,u) \triangleq f_{ij}^k(t,s,u(t,s),F^{f_{ij}^k}(x,t,s)) = 0,\ \ (t,s)\in D_j^i(x,u)\,,\ k=\overline{1,k_{ij}}\,,$$

$$\overline{g}_{ij}^{l}(t,s,x,u) \triangleq g_{ij}^{l}(t,s,u(t,s),F^{g_{ij}^{l}}(x,t,s)) \le 0,\ \ (t,s)\in D_j^i(x,u),\ l=\overline{1,l_{ij}}$$

із заданими функціями $f_{ij}^k$, $g_{ij}^l$ та заданими композиціями $F^{f_{ij}^k}$, $F^{g_{ij}^l}$ операторів $F_1$, $F_2$ і $F_3$, де:

$F_1$ – оператор диференціювання:

$$F_1(x,t,s,0) \triangleq \frac{\partial}{\partial t}x(t,s),\ F_1(x,t,s,i) \triangleq \frac{\partial}{\partial s_i}x(t,s),\ \ i=\overline{1,n_s}\,;$$

$F_2$ – оператор, який визначений на дискретних множинах $\Omega$ (наприклад, $\Omega=\{t^i,s^i\}_{i=1}^{n_\Omega}$, $\Omega_t=\{t^i(t),s^i(t)\}_{i=1}^{n_\Omega}$, $\Omega_{x,t,s}=\{t^i(x,t,s),s^i(x,t,s)\}_{i=1}^{n_\Omega}$ ):

$$F_2(x,t,s,\Omega,\overline{F}_1,i) \triangleq (\overline{F}_1(x,t+t^i(x,t,s),s+s^i(x,t,s))),\ \ i=\overline{1,n_\Omega}\,;$$

$F_3$ – оператор інтегрування на підмножинах $\tilde{\Omega}(t,s,x,u)\subset D$ за заданою функцією $\varphi$:

$$F_3(x,u,t,s,\varphi,\tilde{\Omega}(t,s,x,u),\overline{F}_1) \triangleq$$
$$\triangleq \iint\limits_{\tilde{\Omega}(t,s,x,u)} \varphi(t,s,\overline{F}_1^{x}(x,t+\tau,s+\sigma),\overline{F}_1^{u}(u,t+\tau,s+\sigma))d\tau d\sigma\,.$$

За допомогою множин $\Omega(\alpha)$ тих функцій $(x,u)$, які при $\overline{h}_{ij}^k \triangleq -\overline{f}_{ij}^k$ для всіх $j=\overline{0,m+1}$, $i=\overline{1,i_j}$ задовольняють нерівності:

$$\overline{f}_{ij}^k(t,s,x,u)\le\alpha_k,\ \overline{h}_{ij}^k(t,s,x,u)\le\alpha_k,\ (t,s)\in D_j^i(x,u),\ k=\overline{1,k_{ij}}\,,$$
$$\overline{g}_{ij}^l(t,s,x,u)\le 0,\ \ (t,s)\in D_j^i(x,u),\ l=\overline{1,l_{ij}}\,,$$

узагальнений розв'язок узагальненої задачі оптимізації визначається як послідовність $\{(x_r,u_r)\}_{r=1}^{\infty}$, $(x_r,u_r)\in\Omega(\alpha_r)$, для якої виконуються нерівності $\overline{f}_0(x_r,u_r)\le \inf\limits_{(x,u)\in\Omega(\alpha_r)} \overline{f}_0(x,u)+\alpha_r$, на деякій монотонно спадній збіжній послідовності $\alpha_r\to 0$. Побудова узагальненого розв'язку чисельними методами [[2]] здійснюється у вкладених множинах $X^{n_x(r)}\subset X^{n_x(r+1)}$, $U^{n_u(r)}\subset U^{n_u(r+1)}$ параметричних функцій:

$$(x_{n_x(r)}(p_r,t,s),u_{n_u(r)}(q_r,t,s))\in X^{n_x(r)}\times U^{n_u(r)},$$

де при кожному $\alpha>0$ існує число $r$, для якого параметри $p_r,q_r$ задовольняють нерівностям:

[2] Бейко І.В. Математичне та комп'ютерне моделювання: Зб. наук. пр. – серія: технічні науки // Інститут кібернетики імені В.М.Глушкова НАНУ, Кам'янець-Подільський національний університет імені Івана Огієнка.– Кам'янець-Подільський: 2010. – вип.4. – С.24-32.

$$\max_{(t,s)\in D_j^i(x,u)} \overline{f}_{ij}^k(t,s,x_{n_x(r)}(p_r,\cdot,\cdot),u_{n_u(r)}(q_r,\cdot,\cdot)) \le \alpha,\ k=\overline{1,k_{ij}}$$

$$\max_{(t,s)\in D_j^i(x,u)} \overline{h}_{ij}^k(t,s,x_{n_x(r)}(p_k,\cdot,\cdot),u_{n_u(r)}(q_k,\cdot,\cdot)) \le \alpha,\ k=\overline{1,k_{ij}}$$

$$\max_{(t,s)\in D_j^i(x,u)} \overline{g}_{ij}^l(t,s,x_{n_x(r)}(p_k,\cdot,\cdot),u_{n_u(r)}(q_k,\cdot,\cdot)) \le 0,\ l=\overline{1,l_{ij}}\ ,$$

$$\overline{f}_0(x_{n_x(r)}(p_k,\cdot,\cdot),u_{n_u(r)}(q_k,\cdot,\cdot)) \le \inf_{(x,u)\in\Omega(\alpha)} \overline{f}_0(x,u)+\alpha\ .$$

***Теорема 1.*** *Якщо для вибраних послідовностей вкладених множин* $X^r$, $U^r$, $r=\overline{1,\infty}$, *для опуклих по* $(p,q)$ *функціоналів*

$$\overline{f}_0(x_r(p,\cdot,\cdot),u_r(q,\cdot,\cdot)),\ \overline{g}_{ij}^k(t,s,x_r(p,\cdot,\cdot),u_r(q,\cdot,\cdot))$$

*та лінійних по* $(p,q)$ *функціоналів* $\overline{f}_{ij}^k(t,s,x_r(p,\cdot,\cdot),u_r(q,\cdot,\cdot))$, $k=\overline{1,k_{ij}}$ *при кожному* $\alpha>0$ *знайдеться таке натуральне число* $r$, *для якого множина параметрів* $p\in R^r$ *і* $q\in R^r$, *які задовольняють нерівностії:*

$$\max_{(t,s)\in D_j^i(x,u)} \overline{f}_{ij}^k(t,s,x_r(p,\cdot,\cdot),u_r(q,\cdot,\cdot)) \le \alpha\ ,\ k=\overline{1,k_{ij}}\ ,$$

$$\max_{(t,s)\in D_j^i(x,u)} \overline{h}_{ij}^k(t,s,x_r(p,\cdot,\cdot),u_r(q,\cdot,\cdot)) \le \alpha\ ,\ k=\overline{1,k_{ij}}\ ,$$

$$\max_{(t,s)\in D_j^i(x,u)} \overline{g}_{ij}^k(t,s,x_r(p,\cdot,\cdot),u_r(q,\cdot,\cdot)) \le \alpha\ ,\ l=\overline{1,l_{ij}}\ ,$$

$$\overline{f}_0(x_r(p,\cdot,\cdot),u_r(q,\cdot,\cdot)) \le \inf_{(x,u)\in\Omega(\alpha)} \overline{f}_0(x,u)+\alpha\ ,$$

*має відкриту підмножину, то узагальнений розв'язок міститься у послідовності* $\{x_{n_x(r)}(p_r,\cdot,\cdot),u_{n_u(r)}(q_r,\cdot,\cdot)\}_{r=1}^{\infty}$, *яка обчислена за формулами:*

$$p_{r+1}=p_r-h_r v_r/\|v_r\|,\ q_{r+1}=q_r-h_r w_r/\|w_r\|,$$

$$(v_r,w_r)=\begin{cases}\hat{\nabla}_{(p,q)}\overline{f}_{ij}^k(t,s,x_{n_x(r)}(p_r,\cdot,\cdot),u_{n_u(r)}(q_r,\cdot,\cdot)),\ \textit{якщо} \\ \qquad \overline{f}_{ij}^k(t,s,x_{n_x(r)}(p_r,\cdot,\cdot),u_{n_u(r)}(q_r,\cdot,\cdot))=z, \\ \hat{\nabla}_{(p,q)}\overline{h}_{ij}^k(t,s,x_{n_x(r)}(p_r,\cdot,\cdot),u_{n_u(r)}(q_r,\cdot,\cdot)),\ \textit{якщо} \\ \qquad \overline{h}_{ij}^k(t,s,x_{n_x(r)}(p_r,\cdot,\cdot),u_{n_u(r)}(q_r,\cdot,\cdot))=z, \\ \hat{\nabla}_{(p,q)}\overline{g}_{ij}^k(t,s,x_{n_x(r)}(p_r,\cdot,\cdot),u_{n_u(r)}(q_r,\cdot,\cdot)),\ \textit{якщо} \\ \qquad \overline{g}_{ij}^k(t,s,x_{n_x(r)}(p_r,\cdot,\cdot),u_{n_u(r)}(q_r,\cdot,\cdot))=z, \\ \hat{\nabla}_{(p,q)}\overline{f}_0(x_{n_x(r)}(p_r,\cdot,\cdot),u_{n_u(r)}(q_r,\cdot,\cdot)),\qquad \textit{якщо}\ \ z\le 0,\end{cases}$$

$$z=\max\{\max_{j=\overline{0,m+1}}\max_{i=\overline{1,i_j}}\max_{k=\overline{1,k_{ij}}}\max_{(t,s)\in D_j^i(x,u)} \overline{f}_{ij}^k(t,s,x_{n_x(r)}(p_r,\cdot,\cdot),u_{n_u(r)}(q_r,\cdot,\cdot)),$$

$$\max_{j=\overline{0,m+1}}\max_{i=\overline{1,i_j}}\max_{k=\overline{1,k_{ij}}}\max_{(t,s)\in D_j^i(x,u)} \overline{h}_{ij}^k(t,s,x_{n_x(r)}(p_r,\cdot,\cdot),u_{n_u(r)}(q_r,\cdot,\cdot)),$$

$$\max_{j=\overline{0,m+1}}\max_{i=\overline{1,i_j}}\max_{k=\overline{1,k_{ij}}}\max_{(t,s)\in D_j^i(x,u)} \overline{g}_{ij}^k(t,s,x_{n_x(r)}(p_r,\cdot,\cdot),u_{n_u(r)}(q_r,\cdot,\cdot))\},$$

$$\lim_{r\to\infty} h_r = 0,\ \lim_{r\to\infty} n_x(r) = \infty,\ \lim_{r\to\infty} n_u(r) = \infty,\ \sum_{r=1}^{\infty} h_r = \infty,\ h_r > 0.$$

**2. Задачі оптимального керування**

З а д а ч а 1. Знайти керування $u:[t_0;T]\to\Omega\subset R^r$, яке максимізує функціонал

$$F(x,u) = \int_{t_0}^{T} h(x(t),t)dt$$

на траєкторіях $x:[t_0;T]\to R^n$ керованої системи

$$\frac{dx(t)}{dt} = f\left(x(t),u\left(t\right),t\right),\ t\in[t_0;T] \tag{0.6}$$

при заданому початковому значенні $x(t_0) = x^0$.

Типовими множинами допустимих керувань є:

$$\Omega_1 = \left\{u\middle| -1\le u_i \le 1, i=\overline{1,r}\right\},\ \Omega_2 = \left\{u\middle| -\alpha_i\left(t\right)\le u_i \le \beta_i\left(t\right), i=\overline{1,r}\right\},$$

$$\Omega_3 \triangleq \left\{u\middle|\sum_{i=1}^{r} u_i^2 \le \rho^2\right\},\ \Omega_4 = \{u\in R^r \mid A(t)u\le b(t)\},$$

$$\Omega_5 = \{u\in R^r \mid \varphi(u,t)\le 0\},$$

а також множини із інтегральними обмеженнями:

$$\Omega_6 = \{u(t)\mid \int_{t_0}^{T} \| u(t)\| dt \le P\},\ \Omega_7 = \{u(t)\mid \int_{t_0}^{T}(K(t)u(t),u(t))dt\le P\},$$

$$\Omega_8 = \{u(t)\mid \int_{t_0}^{T}\gamma(u(t),t)dt\le P\}.$$

Множина $\Omega$ може складатися також із різних комбінацій перетинів та об’єднань таких множин $\Omega_1 - \Omega_8$.

Труднощі практичної побудови оптимального керування пов’язані також і з тим, що задача 1 може не мати розв’язку (оптимального керування) навіть і у випадку обмежених з ліпшицевими похідними функцій $h: R^n\times R\to R$, $f: R^n\times R^r\times R\to R^n$ та з обмеженою замкненою множиною $\Omega$. Наприклад, може не існувати оптимального керування у випадку неопуклої множини $\overline{f}\left(x,\Omega,t\right)\triangleq\{v\mid v=f\left(x,u,t\right),\ u\in\Omega\}$. Для таких випадків чисельно будується узагальнене оптимальне керування як максимізатор $\overline{x}^0$ функціоналу $\int_{t_0}^{T} h(\overline{x}(t),t)dt$ на замкненій множині $\overline{X}$:

$$\overline{X}\triangleq\{\overline{x}(t)\mid \frac{d\overline{x}(t)}{dt}\in co\overline{f}\left(\overline{x}(t),\Omega,t\right),\ t\in[t_0;T]\},$$

де через $\overline{cof}\left(x,\Omega,t\right)$ позначено замкнену опуклу оболонку множини $\overline{f}\left(x,\Omega,t\right)$.

Важливою особливістю узагальненого розв'язку $\overline{x}^0$ є те, що для кожного числа $\alpha>0$ існує вимірна вектор-функція $u^\alpha:[t_0;T]\to R^r$, для якої розв'язок $x^\alpha$ задачі Коші:

$$\frac{dx^\alpha(t)}{dt}=f\left(x^\alpha(t),u^\alpha\left(t\right),t\right),\ t\in[t_0;T],\ x^\alpha(t_0)=x^0$$

задовольняє нерівність $\max\limits_{t\in[t_0;T]}\|x^\alpha(t)-\overline{x}^0(t)\|\le\alpha$. У цьому зв'язку послідовність $\{u^{\alpha_k}\}_{k=1}^\infty$, яка визначається збіжною послідовністю $\alpha_k>0$, $\lim\limits_{k\to\infty}\alpha_k=0$, називаємо ***узагальненим розв'язком*** задачі максимізації функціоналу $F(x,u)$ на траєкторіях $x:[t_0;T]\to R^n$ задачі Коші:

$$\frac{dx(t)}{dt}=f\left(x(t),u\left(t\right),t\right),\ t\in[t_0;T],\ x(t_0)=x^0\in R^n$$

які породжуються множиною керувань $u:[t_0;T]\to\Omega\subset R^r$. Таке означення узагальненого розв'язку переноситься і на більш загальні задачі оптимізації, де разом із пошуком оптимальної вектор-функції $u$ потрібно знайти також і оптимальні значення $t_0$, $x(t_0)$, $T$, $x(T)$.

Задача 2. Знайти керування $u\in\Omega$, знайти моменти часу $t_0$, $T$ і знайти початковий $x(t_0)$ та кінцевий $x(T)$ стани керованої системи

$$\frac{dx(t)}{dt}=f\left(x(t),u\left(t\right),t\right),$$

які мінімізують значення функціоналу

$$F_0(x,u,t_0,T)\triangleq\int\limits_{t_0}^{T}h_0(x(t),u(t),t)dt+g_0(t_0,T,x(t_0),x(T))$$

при обмеженнях:

$$F_i(x,u,t_0,T)\triangleq\int\limits_{t_0}^{T}h_i(x(t),u(t),t)dt+g_i(t_0,T,x(t_0),x(T))\le 0,\ i=\overline{1,m}.$$

***Теорема 2 (принцип максимуму).*** *Якщо $(x^0,u^0,t_0^0,T^0)$ є оптимальним розв'язком задачі 2, функції $f$, $h_i$, $i=\overline{0,m}$ та їх похідні по $x$ є неперервними по $x$, а функції $g_i$, $i=\overline{0,m}$ є неперервно диференційованими по всіх аргументах, то на нетривіальному розв'язку $\psi^0(t)$, $\lambda_0^0\ge 0$, $\lambda_i^0$, $i=\overline{0,m}$ системи:*

$$\frac{d\psi^0(t)}{dt} = -\psi^0(t) f_x'(x^0(t), u^0(t), t) + \sum_{i=0}^{m} \lambda_i h_{ix}'(x^0(t), u^0(t), t),$$

$$\psi^0(t_0^0) = \sum_{i=0}^{m} \lambda_i^0 \nabla_{x(t_0)} g_i\left(t_0, T, x^0(t_0), x^0(T)\right),$$

$$\psi^0(T^0) = -\sum_{i=0}^{m} \lambda_i^0 \nabla_{x(T)} g_i\left(t_0, T, x^0(t_0), x^0(T)\right)$$

*оптимальне керування* $u^0(t)$ *при кожному* $t \in [t_0;T]$ *є розв'язком задачі оптимізації*

$$u^0(t) = \arg\max_{u\in\Omega} \left\{ \left(\psi^0(t), f(x^0, u, t)\right) - \sum_{i=0}^{m} \lambda_i^0 h_i(x^0, u, t) \right\}.$$

Для побудови керування $u^0$ і траєкторії $x^0$, які задовольняють принципу максимуму, можна скористатися наступним чисельним алгоритмом:

1) вибрати довільні початкові числа $\bar{\lambda}_i^0 \ge 0$, $\bar{t}_0$, $\bar{T}$, $\bar{\lambda}_i$, $i = \overline{0,m}$ та довільні вектори $\bar{x}, \bar{\bar{x}} \in R^n$ і обчислити розв'язок узагальненої задачі Коші для системи:

$$\frac{dx(t)}{dt} = f\left(x(t), u(t), t\right),$$

$$\frac{d\psi(t)}{dt} = -\psi(t) f_x'(x(t), \bar{u}(t), t) + \sum_{i=0}^{m} \bar{\lambda}_i h_{ix}'(x(t), \bar{u}(t), t),$$

$$\bar{u}(t) = \arg\max_{u\in\Omega} [(\psi(t), f(x, u, t)) - \sum_{i=0}^{m} \bar{\lambda}_i h_i(x, u, t)]$$

за початковими даними $x(\bar{t}_0) = \bar{x}$, $\psi(\bar{t}^0) = \sum_{i=0}^{m} \bar{\lambda}_i^0 \nabla_x g_i(\bar{t}_0, \bar{T}, \bar{x}, \bar{\bar{x}})$;

2) (основний цикл) за допомогою градієнтного методу обчислити мінімізатор $(\bar{t}_0, \bar{T}, \bar{x}, \bar{\bar{x}}, \bar{\lambda}^0, \bar{\lambda})$ функції

$$\bar{F}(\bar{t}_0, \bar{T}, \bar{x}, \bar{\bar{x}}, \bar{\lambda}^0, \bar{\lambda}) \triangleq \| \psi(\bar{T}) - \sum_{i=0}^{m} \bar{\lambda}_i^0 \nabla_x g_i(\bar{t}_0, \bar{T}, \bar{x}, \bar{\bar{x}}) \|^2$$

при обмеженнях:

$$\bar{t}_0 < \bar{T},\ \bar{\lambda}^0 \equiv (\bar{\lambda}_0^0, \bar{\lambda}_1^0, ..., \bar{\lambda}_m^0) \ge 0,\ \bar{\lambda} \equiv (\bar{\lambda}_0, \bar{\lambda}_1, ..., \bar{\lambda}_m) \ge 0.$$

## 3. Методи розв'язування типових задач оптимального керування

### *3.1. Лінійна задача Майєра.*

Задача 1. Знайти керування $u(t) \in R^r$, яке у кожен момент часу $t \in [t_0;T]$ належить заданій множині $V(t) \subset R^r$ $\left(u(t) \in V(t)\right)$ і максимізує значення $(c, x(T))$ на розв'язках $x(T)$ керованої системи

$$\frac{dx(t)}{dt} = A(t)x(t) + g(u(t),t)$$

для заданих моментів часу $t_0, T$, заданої матриці $A(t) \in R^n \times R^n$, заданої вектор-функції $g(u(t),t) \in R^n$ і заданого вектора $c \in R^n$.

*Розв'язування.* Із теореми принципу максимуму для даної задачі випливає, що у кожен момент часу $t$ значення $u^0(t)$ оптимального керування є максимізатором скалярного добутку $(g(u(t),t),\psi(t))$ на допустимій множині $V(t)$, де $\psi(t)$ – розв'язок задачі Коші для спряженої системи:

$$\frac{d\psi(t)}{dt} = -A^T(t)\psi(t), \; \psi(T) = c,$$

тобто для всіх $u(t) \in V(t)$ виконується нерівність «принципу максимуму»

$$(g(u^0(t),t),\psi(t)) \geq (g(u(t),t),\psi(t)).$$

Отже, метод побудови оптимального керування $u^0(t)$ реалізується за допомогою використання чисельного алгоритму для обчислення траєкторії $\psi(t)$ у зворотному часі $t$, починаючи від відомого у кінцевий момент часу $t = T$ значення $\psi(T) = c$, до початкового моменту часу $t = t_0$ із паралельним обчисленням у кожен момент часу $t$ оптимального керування

$$u^0(t) = \arg\max_{u \in V(t)}(g(u,t),\psi(t)).$$

**Приклад 1.** Знайти керування $u(t) \in R^r$, яке у кожен момент часу $t$ належить множині $V(t) = \{u \subset R^r \mid -1 \leq u_i \leq 1, \; i = \overline{1,r}\}$, тобто кожна компонента вектор-функції керування належить інтервалу $[-1;1]$ $\left(u_i(t) \in [-1;1], \; i = \overline{1,r}\right)$ і максимізує значення скалярного добутку $(c, x(T))$ на розв'язку $x(T) = (x_1(T), x_2(T))^T$ задачі Коші:

$$\frac{dx_1}{dt} = x_2, \; \frac{dx_2}{dt} = u$$

для заданих моментів часу $t_0 = 0, \; T = 1$.

*Розв'язування.* У даному прикладі маємо керовану систему $\frac{dx(t)}{dt} = A(t)x(t) + g(u(t),t)$ із матрицю $A(t) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ і вектор-функцією $g(u(t),t) = \begin{pmatrix} 0 \\ u(t) \end{pmatrix}$. Вектор-функція $\psi(t) = (\psi_1(t),\psi_2(t))^T$ є розв'язком задачі

Коші $\frac{d\psi(t)}{dt} = -A^T(t)\psi(t),\ \psi(T) = c$, і для транспонованої матриці $A^T(t) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ задовольняє рівнянням:

$$\frac{d\psi_1(t)}{dt} = 0,\ \frac{d\psi_2(t)}{dt} = -\psi_1(t),\ \psi_1(T) = c_1,\ \psi_2(T) = c_2 .$$

Розв'язком цих рівнянь є функції $\psi_1(t) = c_1,\ \psi_2(t) = c_1(T-t) + c_2$. Для даного прикладу маємо

$$\big(g(u,t), \psi(t)\big) = u(t)\psi_2(t) = u(t)(c_1(T-t) + c_2)$$

і тому оптимальне керування $u^0(t)$ обчислюємо за формулою «принципу максимуму»

$$u^0(t) = \arg\max_{|u|\le 1}(g(u,t), \psi(t)) = \arg\max_{|u|\le 1}(u\psi_2(t)) =$$
$$= \begin{cases} 1, & \text{якщо } \psi_2(t) > 0, \\ -1, & \text{якщо } \psi_2(t) < 0, \\ \text{будь-якому числу із } [-1;1], & \text{якщо } \psi_2(t) = 0. \end{cases}$$

Із рівняння $\psi_2(t) = c_1(T-t) + c_2$ випливає, що:

$$\psi_2(t) > 0 \text{ для } t < T + c_2 / c_1;$$
$$\psi_2(t) < 0 \text{ для } t > T + c_2 / c_1;$$
$$\psi_2(t) = 0 \text{ для } t = T + c_2 / c_1.$$

Це означає, що $u^0(t) = 1$ для моментів часу $t < T + c_2 / c_1$, $u^0(t) = -1$ для моментів часу $t > T + c_2 / c_1$ і $u^0(t)$ дорівнює будь-якому числу із інтервалу $[-1;1]$ у момент часу $t = T + c_2 / c_1$.

З а д а ч а 2. Знайти керування $u(t)$, яке у кожен момент часу $t$ належить множині $u(t) \in V(t) \subset R^r$ і для заданої опуклої функції $F : R^n \to R$ максимізує значення $F(x(T))$ на розв'язках $x(T)$ задачі Коші

$$\frac{dx(t)}{dt} = A(t)x(t) + g(u(t),t),\ x(t_0) = x^0 \in R^n .$$

*Розв'язування.* Розглянемо випадок, у якому оптимальне значення $x^0(T)$ не є внутрішньою точкою множини $X(T)$ значень $x(T)$, визначених усіма допустимими керуваннями $u(t) \in V(t) \subset R^r$. Якщо $c^0 \ne 0$ є таким вектором, що оптимальний розв'язок $x_c(T)$ задачі 1 при $c = c^0$ задовольняє нерівності $F(x) \le F(x_c(T)) + (c, x - x_c(T))$ для всіх $x \in R^n$, то оптимальним розв'язком задачі 2 є оптимальний розв'язок задачі 1 для вектора $c = c^0$ і при цьому вектор $c^0$ є узагальненим градієнтом функції $F$ у точці $x_c(T)$.

Це випливає із того, що множина $X(T)$ усіх значень $x_c(T)$, визначених усіма допустимими керуваннями $u(t) \in V(t) \subset R^r$, є опуклою і вектор $c$ є опорним до множини $X(T)$ у точці $x_c(T)$. Отже, шуканий вектор $c^0$ може бути обчислений як максимізатор функції $\varphi(c) \triangleq F(x_c(T))$, узагальненим градієнтом якої є вектор $\Pi(F, x_c(T), c) - c$, де $P(F, x_c(T), c)$ є проекцією вектора $c$ на множину узагальнених градієнтів функції $F$ у точці $x_c(T)$.

У випадку диференційованої функції $F$ множина опорних градієнтів складається із одного вектора $\nabla_x F(x_c(T))$ і тому для обчислення вектора $c^0$ можна скористатися ітераційною формулою

$$c^{k+1} = c^k + \lambda_k (\nabla_x F(x_c(T)) - c^k),\ k = 1, 2, \ldots,\ c^1 \in R^n$$

з вибором множника $\lambda_k$ із умови $F(x_{c^{k+1}}(T)) > F(x_{c^k}(T))$.

***3.2. Побудова керування оптимального по швидкодії.***

З а д а ч а 3. Знайти керування $u(t) \in R^r$, яке у кожен момент часу $t \in [t_0, T]$ належить заданій множині $V(t) \subset R^r$, $u(t) \in V(t)$, і переводить керовану систему

$$\frac{dx(t)}{dt} = A(t)x(t) + g(u(t), t),\ x(t_0) = x^0 \in R^n,$$

у заданий кінцевий стан $x(T) = x^1$ за мінімальний час $T - t_0$, тобто, мінімізує значення $T$.

*Розв'язування.* Із опуклості множини $X(T)$ усіх значень $x(T)$, генерованих допустимими керуваннями $u(t) \in V(t)$ випливає, що розв'язком задачі 3 є розв'язок задачі 1 для мінімального часу $T = T^0$, який задовольняє умові $x^1 \in X(T)$, і для вектора $c^0$, при якому розв'язок $x(T^0)$ задачі 1 задовольняє рівності $x(T^0) = x^1$. Для відшукання оптимальних значень $T^0, c^0$ можна скористатися наступним чисельним алгоритмом побудови збіжної до $T^0, c^0$ послідовності $\{T^k, c^k\}_{k=1}^{\infty}$, який використовує твердження леми 1.

**Л е м а 1.** Якщо на розв'язку задачі 1 для значень $T = T^k$, $c = c^k$ допоміжна функція $\varphi(t, c^k) \triangleq (\psi(t), x^1 - x(t))$ задовольняє нерівності $\varphi(t, c^k) > 0$ для всіх $t \in [t_0; T^k]$, то $T^0 > T^k$.

Справедливість леми 1 випливає із того, що множина $X(t)$ є опуклою і, отже, із нерівності $\varphi(t, c^k) > 0$ випливає, що точка $x^T$ не належить множині $X(t)$.

Наступний момент часу $T^{k+1}$, який залежить від значення вектора $c^k$, обчислюється, як найменший корінь рівняння $\varphi(t, c^k) = 0$, більший за $T^k$, а наступне значення $c^{k+1}$ обчислюється за формулою методу узагальненого градієнта

$$c^{k+1} = \psi(T^{k+1})c^k + \lambda_k(x^1 - x(T^{k+1})) .$$

Відомо, що метод узагальненого градієнта генерує збіжну до $T^0$ послідовність $T^k$ при $\lambda_k = L/k$ з будь-якою додатною константою $L$.

***3.3. Двоїста задача Майєра. Мінімізація амплітуди керування.***

Задача 4. Знайти оптимальне керування $u(t) \in R^r$, яке у кожен момент часу $t \in [t_0;T]$ задовольняє нерівностям $|u_i(t)| \le \alpha,\ i = \overline{1,r}$, і при найменшому значенні $\alpha$ переводить лінійну керовану систему $\dfrac{dx}{dt} = A(t)x(t) + B(t)u(t)$ із заданого початкового стану $x(t_0) = x^0 \in R^n$ до стану $x(T)$, що задовольняє рівності $(c, x(T)) = d$.

*Розв'язування.* Із розв'язку задачі 1 випливає, що значення оптимального керування $u^0(t)$ для задачі 4 у кожен момент часу $t$ є максимізатором функції $(B(t)u, \psi(t))$ на множині $\{u \mid |u_i(t)| \le \alpha,\ i = \overline{1,r}\}$. Легко перевірити, що у даному випадку маємо:

$$u^0(t) = \alpha sign[B^T(t)\psi(t)],$$

де $\psi(t)$ – розв'язок задачі Коші $\dfrac{d\psi(t)}{dt} = -A^T(t)\psi(t),\ \psi(T) = c$, а символом $sign[v]$ позначено вектор, отриманий із вектора $v$, координатами якого є значення функції $sign$ від відповідних координат вектора, а саме: якщо $v_i > 0$, то $sign(v_i) = 1$; якщо $v_i < 0$, то $sign(v_i) = -1$; якщо $v_i = 0$, то $sign(v_i) = 0$.

Із тотожності

$$(c, x(T)) \equiv (\psi(t_0), x(t_0)) + \int_{t_0}^{T} (\psi(t), B(t)u(t))dt ,$$

яка виконується для траєкторії $x(t)$ керованої системи і для траєкторії $\psi(t)$ спряженої системи $\dfrac{d\psi(t)}{dt} = -A^T(t)\psi(t),\ \psi(T) = c$, випливає, що на оптимальному керуванні справедливі рівності:

$$d = (\psi(t_0), x(t_0)) + \alpha\int_{t_0}^{T} (\psi(t), B(t)sign[B^T(t)\psi(t)])dt =$$

$$= (\psi(t_0), x(t_0)) + \alpha\int_{t_0}^{T} \| B^T(t)\psi(t) \|_2 \, dt ,$$

де $\| v \|_2 \triangleq \sum_i | v_i |$. Це означає, що мінімальне значення $\alpha^0$ обчислюється за формулою

$$\alpha^0 = \left(d - (\psi(t_0), x(t_0))\right) \Big/ \int_{t_0}^{T} \| B^T(t)\psi(t) \|_2 \, dt$$

і оптимальне керування для задачі 4 обчислюється за формулою

$$u^0(t) = \alpha^0 sign[B^T(t)\psi(t)] .$$

З а д а ч а 5. Знайти оптимальне керування $u(t) \in R^r$, яке у кожен момент часу $t \in [t_0;T]$ задовольняє нерівностям $|u_i(t)| \le \alpha,\ i = \overline{1,r}$, і при найменшому значенні $\alpha$ переводить лінійну керовану систему $\frac{dx}{dt} = A(t)x(t) + B(t)u(t)$ із заданого початкового стану $x(t_0) = x^0 \in R^n$ до заданого стану $x(T) = x^1$.

*Розв'язування.* Якщо $c^0 \neq 0$ такий вектор, що оптимальна траєкторія $x^0(t)$ задачі 1 при $c = c^0$ і $d = (x^1, c^0)$ задовольняє умову $x^0(T) = x^1$, то, очевидно, оптимальним керуванням для задачі 5 є оптимальне керування для задачі 1 при $c = c^0$.

Очевидно, вектор $c^0$ обчислюється як максимізатор по $c$ мінімального значення

$$\alpha^0(c) \triangleq \left((c, x^1) - (\psi(t_0), x(t_0))\right) \Big/ \int_{t_0}^{T} \| B^T(t)\psi(t) \|_2 \, dt$$

для задачі 1 при умові $d = (c^0, x^1)$, тобто, $c^0 = \arg\max_c \alpha^0(c)$ на розв'язку $\psi(t)$ задачі Коші $\frac{d\psi(t)}{dt} = -A^T(t)\psi(t),\ \psi(T) = c$. Оскільки функція $\alpha^0(c)$ є опуклою догори, для її максимізації можна скористатися будь-яким із чисельних алгоритмів максимізації опуклих догори функцій, беручи до уваги, що узагальненим градієнтом функції $\alpha^0(c)$ є вектор $x^1 - x_c(T)$, визначений за значенням $x_c(T)$ оптимального розв'язку задачі 4 для $d = (c, x^1)$.

З а д а ч а 6. Знайти оптимальне керування $u(t) \in R^r$, яке у кожен момент часу $t \in [t_0;T]$ задовольняє нерівностям $|u_i(t)| \le \alpha,\ i = \overline{1,r}$, і при найменшому значенні $\alpha$ переводить лінійну керовану систему $\frac{dx}{dt} = A(t)x(t) + B(t)u(t)$ із заданого початкового стану $x(t_0) = x^0 \in R^n$ на задану множину $X^* \subset R^n$ допустимих кінцевих станів $x(T) \in X^*$.

Метод розв'язування задачі 6 аналогічний методу розв'язування задачі 5 для $x^T = \arg\min_{x \in X^*}(c, x)$.

### 3.4. *Мінімізація квадратичного функціоналу.*

Задача 7. Знайти оптимальне керування $u(t)\in R^r$, яке мінімізує квадратичний функціонал $\int_{t_0}^{T}(K(t)u(t),u(t))dt$ і при цьому переводить лінійну керовану систему $\frac{dx}{dt}=A(t)x(t)+B(t)u(t)+f^1(t)$ із заданого початкового стану $x(t_0)=x^0\in R^n$ у заданий кінцевий стан $x(T)=x^1\in R^n$.

*Розв'язування.* Розглянемо випадок додатньо визначеної матриці $K(t)$, для якої існує обернена матриця $K^{-1}(t)$, і скористаємося відомою формулою

$$x(t)=\Phi(t,t_0)x(t_0)+\Phi(t,t_0)\int_{t_0}^{t}\Psi^T(\tau,t_0)(B(\tau)u(\tau)+f^1(t))d\tau$$

залежності траєкторії $x(t)$ від початкового стану $x(t_0)$ та від керування $u(t)$ і збурення $f^1(t)$ на інтервалі часу $t\in[t_0;T]$, де матриця Коші $\Phi\left(t,t_0\right)$ є розв'язком матричної задачі Коші

$$\frac{d\Phi(t,t_0)}{dt}=A\left(t\right)\Phi\left(t,t_0\right),\ \Phi\left(t_0,t_0\right)=I,$$

*I*-одинична матриця, а матриця $\Psi^T\left(t,t_0\right)$ є транспонованою до матриці Коші $\Psi\left(t,t_0\right)$ для спряженої системи $\frac{d\psi(t)}{dt}=-A^T(t)\psi(t)$ , тобто

$$\frac{d\Psi(t,t_0)}{dt}=-A^T\left(t\right)\Psi\left(t,t_0\right),\ \Psi\left(t_0,t_0\right)=I.$$

Важливою властивістю цих матриць є тотожність $\Psi^T\left(t,t_0\right)\equiv\Phi^{-1}\left(t,t_0\right)$.

Із сформульованої вище теореми принципу максимуму випливає, що оптимальне керування $u^0\left(t\right)$ для задачі 1 є мінімізатором для функціоналу

$$I\left(u\right)=-\left(1/2\right)\left(K\left(t\right)u,u\right)+\left(\psi\left(t\right),B\left(t\right)u\right),$$

тобто $u^0(t)=K^{-1}(t)B^T(t)\psi(t)$ і згідно з визначенням $\Psi(t,t_0)$ маємо

$$\psi(t_0)=\Psi(t,t_0)\ \psi(t_0),\quad u^0(t)=K^{-1}(t)B^T(t)\Psi(t,t_0)\psi(t_0).$$

Тому оптимальна траєкторія

$$x(t)=\Phi(t,t_0)x(t_0)+$$

$$+\Phi(t,t_0)\int_{t_0}^{t}[\Psi^T(\tau,t_0)B(\tau)(K^{-1}(\tau)B^T(\tau)\Psi(\tau,t_0)\psi(t_0)+f^1(\tau))]d\tau$$

визначається вектором $\psi(t_0)$, який обчислюється за умовою $x(T)=x^1$ із системи рівнянь

$$x^1 = \Phi(T,t_0)x^0 +$$

$$+\Phi(T,t_0)\int_{t_0}^{T}[\Psi^T(\tau,t_0)B(\tau)K^{-1}(\tau)B^T(\tau)\Psi(\tau,t_0)\psi(t_0) + f^1(\tau)]d\tau \,.$$

Якщо ввести позначення

$$M \triangleq \int_{t_0}^{T}\Psi^T(\tau,t_0)B(\tau)K^{-1}(\tau)B^T(\tau)\Psi(\tau,t_0)d\tau$$

і скористатися тотожністю $\Psi^T\left(t,t_0\right) \equiv \Phi^{-1}\left(t,t_0\right)$, то із системи рівнянь

$$M\psi(t_0) = \Psi^T(T,t_0)x^1 - x^0 - \int_{t_0}^{T}\Psi^T(\tau,t_0)f^1(\tau)d\tau$$

отримуємо

$$\psi(t_0) = M^{-1}[\Psi^T(T,t_0)x^1 - x^0 - \int_{t_0}^{T}\Psi^T(\tau,t_0)f^1(\tau)d\tau] \,.$$

Отже, оптимальне керування $u^0(t)$ для задачі 1 визначається за формулою

$$u^0(t) = K^{-1}(t)B^T(t)\Psi(t,t_0)M^{-1}[\Psi^T(T,t_0)x^1 - x^0 - \int_{t_0}^{T}\Psi^T(\tau,t_0)f^1(\tau)d\tau] \,.$$

Очевидно, існування оберненої матриці $M^{-1}$ є достатньою умовою існування оптимального керування для будь-якого початкового стану $x(t_0) = x^0$ і для будь-якого кінцевого стану $x(T) = x^1$ керованої системи.

***Алгоритм розв'язування задачі 7***

I. Паралельно обчислити усі стовпці $\psi^i(t)$, $i = \overline{1,n}$, матриці $\Psi(t,t_0)$ як розв'язки задач Коші для спряженої системи $\dfrac{d\psi(t)}{dt} = -A^T(t)\psi(t)$, $\psi^i_j(t_0) = 0$ для $i \neq j$, $\psi^i_i(t_0) = 1$ (наприклад, методом Рунге–Кутта) із одночасним обчисленням матриць $\Psi^T(t,t_0)B(\tau)K^{-1}(t)B^T(t)\Psi(t,t_0)$, вектор-функцій $\Psi^T(t,t_0)f^1(t)$ та інтегралів:

$$M = \int_{t_0}^{T}\Psi^T(\tau,t_0)B(\tau)K^{-1}(\tau)B^T(\tau)\Psi(\tau,t_0)d\tau \,,\; f = \int_{t_0}^{T}\Psi^T(\tau,t_0)f^1(\tau)d\tau \,.$$

II. Обчислити обернену матрицю $M^{-1}$.

III. Розв'язати задачу Коші для систем:

$$\frac{d\psi(t)}{dt} = -A^T(t)\psi(t) \,,\quad \frac{dx(t)}{dt} = A(t)x(t) + B(t)K^{-1}(t)B^T(t)\psi(t) \,,$$

із початковими умовами $x(t_0) = x^0$ і $u^0(t)$, $\psi(t_0) = M^{-1}\left[\Psi^T(T,t_0)x^1 - x^0 - f\right]$. У процесі чисельного розв'язування цієї задачі Коші (наприклад, з використанням методів Рунге–Кутта) паралельно обчислюються оптимальне керування $u^0(t) = K^{-1}(t)B^T(t)\psi(t)$ та оптимальна фазова траєкторія $x^0(t)$.

**Приклад 2.** Знайти керування *u*(*t*), що переводить систему

$$\frac{dx_1}{dt} = x_2\,,\ \frac{dx_2}{dt} = u$$

із початкового стану $x(0) = x^0$ до стану $x(1) = x^1$ при мінімальному значенні функціоналу $\int\limits_0^1 (u(t), u(t))dt$.

*Розв'язування.* У даному прикладі маємо матрицю $A(t) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, матрицю $B(t) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ і матрицю $K(t) = 1$, для яких отримуємо:

$$\Phi(t,0) = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix},\ \Psi(t,0) = \begin{pmatrix} 1 & 0 \\ -t & 1 \end{pmatrix},$$

$$M = \int\limits_0^1 \begin{pmatrix} 1 & -t \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -t & 1 \end{pmatrix}dt = \begin{pmatrix} \frac{1}{3} & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{pmatrix},\ M^{-1} = \begin{pmatrix} 12 & 6 \\ 6 & 4 \end{pmatrix}.$$

Оптимальне керування обчислюємо за формулою

$$u^0(t) = B^T(t)\Psi(t,0)M^{-1}[\Psi^T(1,0)x^1 - x^0] =$$
$$= (6-12t)x_1^1 + (12t-6)x_1^0 + (6t-2)x_2^1 + (6t-4)x_2^0.$$

*Означення 1.* Система $\frac{dx}{dt} = A(t)x(t) + B(t)u(t)$ називається ***повністю керованою,*** якщо для будь-яких фазових станів $x^0, x^1$ існує керування $u$, яке переводить керовану систему із початкового стану $x(t_0) = x^0$ у кінцевий стан $x(T) = x^1$. Із отриманого вище оптимального керування

$$u^0(t) = B^T(t)\Psi(t,t_0)M^{-1}[\Psi^T(T,t_0)x^1 - x^0]$$

для задачі мінімізації квадратичного функціоналу в умовах одиничної матриці $K(t)$ і функції $f^1(t) \equiv 0$ випливає, що керована система $\frac{dx}{dt} = A(t)x(t) + B(t)u(t)$ є повністю керованою, якщо існує обернена матриця для матриці $\bar{M} = \int\limits_{t_0}^{T} \Psi^T(\tau,t_0)B(\tau)B^T(\tau)\Psi(\tau,t_0)d\tau$. Умови існування оберненої матриці $\bar{M}^{-1}$, а отже, умови повністю керованої неавтономної лінійної системи формулюються у наступній теоремі [3].

***Теорема 3 (повної керованості).*** *Якщо існує момент часу $t$, у який знайдеться n лінійно незалежних векторів серед векторів $b_{s,j}(t)$, $j=\overline{1,n}$, $s=\overline{1,r}$ визначених за формулами:*

$$b_{s,j}(t)=-A(t)b_{s,j-1}(t)+\frac{d}{dt}b_{s,j-1}(t),\quad b_{s,1}(t)=b_s(t),\ j=\overline{2,n},$$

*де $b_s(t)$ є $s$-им стовпцем матриці $B(t)$, то лінійна система $\frac{dx}{dt}=A(t)x(t)+B(t)u(t)$ є повністю керованою.*

***3.5. Узагальнена задача мінімізації квадратичного функціоналу.***

Задача 8. Знайти керування $u(t)\in R^r$, $t\in[t_0;T]$, для якого траєкторія $x(t)\in R^n$ керованої системи

$$\frac{dx(t)}{dt}=A(t)x(t)+B(t)u(t),\ t\in[t_0;T]$$

задовольняє граничні умови:

$$Hx(t_0)+h=0,\ Gx(T)+g=0$$

і значення $I(x,u)$ квадратичного функціоналу

$$I(x,u)\triangleq\int_a^b\{(C(t)x(t),x(t))+(D(t)u(t),u(t))+$$
$$+(L(t)x(t),u(t))+(k(t),u(t))+(m(t),x(t))\}dt$$

досягає мінімального значення для заданих матричних функцій $A,B,C,D,L,H,G$, вектор-функцій $k,m$ та векторів $h$ і $g$ відповідних розмірностей.

*Розв'язування.* Із теореми принципу максимуму випливає, що оптимальне керування $u^*(t)$ в кожен момент часу $t$ мінімізує функцію

$$(C(t)x(t),x(t))+(D(t)u(t),u(t))+(L(t)x(t),u(t))+$$
$$+(k(t)-B^T(t)\psi(t),u(t))+(m(t),x(t)),$$

(тобто, задовольняє рівності

$$u^*(t)=\frac{1}{2}D^{-1}(t)[B^T(t)\psi(t)-L^T(t)x(t)-k(t)])$$

на розв'язках $\psi(t)$ та $x(t)$ крайових задач для систем:

[3] Бейко И.В. О некоторых методах отыскания оптимальных управлений / И.В. Бейко, М.Ф. Карпенко. – К.: Изд-во АН УССР, Научн. Совет по кибернетике, Вычислит. математ. Вып.1. 1963. – 22 с.

$$\frac{dx(t)}{dt} = A(t)x(t) + \frac{1}{2}B(t)D^{-1}(t)[B^T(t)\psi(t) - L^T(t)x(t) - k(t)],$$

$$\frac{d\psi(t)}{dt} = -(c(t) + \frac{1}{2}L(t)D^{-1}(t)L^T(t))x(t) -$$

$$-(A^T(t) - \frac{1}{2}L(t)D^{-1}(t)B^T(t))\psi(t) + m(t) - \frac{1}{2}L(t)D^{-1}(t)k(t)$$

із крайовими умовами:

$$Hx(t_0) + h = 0,\ Gx(T) + g = 0,\ H^T p + \psi(t_0) = 0,\ G^T q + \psi(T) = 0.$$

***3.6. Градієнтні методи побудови оптимального керування.***

Задача 9. Знайти керування $u(t) \in R^r$, $t \in [t_0;T]$, яке на траєкторії $x(t) \equiv x(t,u) \in R^n$ керованої системи

$$\frac{dx(t)}{dt} = f\left(x(t),u(t),t\right),\ x\left(t_0\right) = x^0 \in R^n, t \in [t_0;T]$$

максимізує функціонал

$$I\left(u\right) \triangleq F(x(T,u)) + \int_{t_0}^{T} f_0\left(x\left(t,u\right),u\left(t\right),t\right)dt,$$

для заданих неперервних разом із своїми похідними по аргументу $x$ функцій $f: R^n \times R^r \times R \to R^n$, $f_0: R^n \times R^r \times R \to R$, $F: R^n \to R$.

*Розв'язування.* За допомогою градієнтного методу мінімізації функціоналу $I$ будуємо послідовність керувань $u^1, u^2, \ldots, u^k, u^{k+1}$,

$$u^{k+1}(t) = u^k(t) + \lambda_k \nabla_u I(u^k)(t),\ t \in [t_0;T],$$

де значення $\nabla_u I(u^k)(t)$ градієнта функціоналу $I$ на керуванні $u = u^k$ обчислюються у моменти часу $t \in [t_0;T]$ за формулами;

$$\nabla_u I(u^k)(t) = \left[\frac{\partial f(x(t,u^k),u^k(t),t)}{\partial u}\right]^T y(t) + \nabla_u f_0(x(t,u^k),u^k(t),t),$$

$$\frac{dy(t)}{dt} = \left[\frac{\partial f(x(t,u^k),u^k(t),t)}{\partial x}\right]^T y(t) + \nabla_x f_0(x(t,u^k),u^k(t),t),$$

$$y(T) = \nabla F(x(T,u^k)),$$

а множник $\lambda_k$ обчислюється як максимальне число із послідовності $\{\lambda_{k-1}/2^{q-1},\ q = 1,2,\ldots\}$, для якого виконується нерівність $I(u^{k+1}) \geq I(u^k) + s\lambda_k^2$.

Доведено, що при довільно вибраному початковому керуванні $u^1$ та довільно вибраних числових значеннях параметрів алгоритму $\lambda_0 > 0$, $s > 0$

послідовність $\{u^k\}_{k=1}^{\infty}$ задовольняє рівності $\lim_{k\to\infty} \| \nabla_u I\left(u^k(t)\right) \| = 0$ і епсилон-екстремальне керування може бути обчислене за скінченну кількість кроків $k$.

***3.7. Чисельні методи побудови оптимального керування при наявності обмежень на керування.***

З а д а ч а 10. Відрізняється від задачі 9 лише наявністю обмежень $u(t) \in \Omega \subset R^r$ на значення керування $u(t)$ у моменти часу $t \in [t_0;T]$.

*Розв'язування.* В залежності від властивостей функцій $f$, $f_0$, $F$ та властивостей множини $\Omega$ використовуються різноманітні чисельні алгоритми побудови послідовностей керувань $u^1, u^2, \ldots, u^k, u^{k+1}$, які мінімізують функціонал $I$ на допустимій множині $\Omega$. Зокрема, такі послідовності будуються :

– методами *проекції градієнтів*

$$u^{k+1}(t) = \bar{u}_{\Omega}^{k+1}(t) \triangleq \Pi_{\Omega}\left(u^k(t) + \lambda_k \nabla_u I(u^k)(t)\right) \triangleq$$
$$\triangleq \arg\min_{u\in\Omega} \| u - (u^k(t) + \lambda_k \nabla_u I(u^k)(t)) \| ;$$

– *модифікованими методами проекції градієнтів* на підмножини $\Omega_k$,

$$\Omega_k \triangleq \Omega \bigcap \{u \mid \| u - u^k \| \le \lambda_k\},\ u^{k+1}(t) = \bar{u}_{\Omega_k}^{k+1}(t);$$

– методами *допустимих напрямків*

$$u^{k+1}(t) = \arg\max_{u\in\Omega_k} (\nabla_u I(u^k)(t), u);$$

– методами *умовних градієнтів*

$$u^{k+1}(t) = u^k(t) + \lambda_k (z^k(t) - u^k(t)),\ z^k(t) = \arg\max_{u\in\Omega} (\nabla_u I(u^k)(t), u);$$

– методами *локального принципу максимуму*

$$u^{k+1}(t) = \bar{u}^{k+1}(t, \Omega_k) \triangleq \arg\max_{u\in\Omega_k} ((y(t), f(x(t,u^k), u, t)) + f_0(x(t,u^k), u, t));$$

– або методами *умовного принципу максимуму*

$$u^{k+1}(t) = u^k(t) + \bar{\lambda}_k \left(\bar{u}_{\Omega}^{k+1}(t) - u^k(t)\right),$$
$$\bar{\lambda}_k = \arg\max_{\lambda} F(x(T, u^k + \lambda(z^k - u^k)),$$

які базуються на використанні розв'язку $y(t)$, $t \in [t_0;T]$ допоміжної задачі Коші:

$$\frac{dy(t)}{dt} = \left[\frac{\partial f(x(t,u^k), u^k(t), t)}{\partial x}\right]^T y(t) + \nabla_x f_0(x(t,u^k), u^k(t), t),$$
$$y(T) = \nabla F(x(T, u^k))$$

і на використанні значень $\lambda_k$ як максимальних значень послідовності $\{\lambda_k = \lambda_{k-1} / 2^{1-q},\ q = 1,2,\ldots\}$ ($\lambda_0$ – довільне додатне число), для яких виконується нерівність

$$I(u^{k+1}) \geq I(u^k) + s(\lambda_k)^2$$

або на використанні програмно заданих послідовностей:

$$\lambda_k \to 0,\ \sum_{k=1}^{\infty} \lambda_k = \infty,\ \sum_{k=1}^{\infty} \lambda_k^2 < \infty,\ \lim_{k\to\infty} \frac{\lambda_k}{\lambda_{k+1}} = 1,$$

які забезпечують збіжність підпослідовності $\{I(u^{k_i})\}_{i=1}^{\infty}$ до локального або глобального максимуму функціонала $I(u)$ на множині $\Omega$.

***3.8. Чисельний метод побудови оптимального керування за принципом максимуму.***

Чисельний метод побудови оптимального керування $u^*(t) \in \Omega$, $t \in [t_0;T]$ для задачі 10 за допомогою принципу максимуму базується на твердженні теореми принципу максимуму про те, що у кожен момент часу $t \in [t_0;T]$ оптимальне значення $u^*(t)$ є максимізатором

$$u^*(t) = \bar{u}^*(x(t), y(t), t) \triangleq \arg\max_{u\in\Omega}((y(t), f(x(t), u, t)) + f_0(x(t), u, t))$$

на розв'язку $(x(t), y(t))$, $t \in [t_0;T]$ системи диференціальних рівнянь:

$$\frac{dx(t)}{dt} = f(x(t), \bar{u}^*(x(t), y(t), t), t), \quad t \in [t_0;T],$$

$$\frac{dy(t)}{dt} = \left[\frac{\partial f(x(t), \bar{u}^*(x(t), y(t), t), t)}{\partial x}\right]^T y(t) + \nabla_x f_0(x(t), \bar{u}^*(x(t), y(t), t), t)$$

з граничними умовами:

$$x(t_0) = x^0, \quad y(T) = \nabla F(x(T)).$$

Для обчислення значення $y^*(t_0)$, яке разом із початковим значенням $x(t_0) = x^0$ визначає розв'язок $(x(t), y(t))$, що задовольняє рівності $y(T) = \nabla F(x(T))$, використовуємо чисельні алгоритми мінімізації функції $\varphi(y(t_0)) \triangleq \| y(T) - \nabla F(x(T)) \|^2$ за допомогою методу опорних градієнтів

$$y^{k+1}(t_0) = y^k(t_0) + \lambda_k z^k(t_0),$$

де опорний градієнт $z^k(t_0)$ обчислюється як розв'язок задачі Коші:

$$\frac{dz(t)}{dt} = \left[\frac{\partial f(x^k(t), \bar{u}^*(x^k(t), y^k(t), t), t)}{\partial x}\right]^T z(t) + \nabla_x f_0(x^k(t), \bar{u}^*(x^k(t), y^k(t), t), t),$$

$$z(T) = \nabla F(x^k(T)),$$

на траєкторії $(x^k(t), y^k(t))$ із початковими значеннями $x^k(t_0) = x^0$ та $y^k(t_0)$.

### *3.9. Розв'язуючі оператори для оптимального керування складними граф-операторними системами.*

Складні керовані системи складаються із взаємодіючих підсистем, для яких часто бувають невідомими математичні моделі їх взаємодії, або ж наявні моделі бувають нелінійними, ієрархічно керованими, з неповними даними чи з великою розмірністю, що затрудняє їх оптимізацію. При відшуканні оптимального керування такою складною системою за критеріями підвищення якості, надійності та ефективності її функціонування виникає необхідність розв'язувати додаткові задачі уточнення математичних моделей підсистем, оцінювання стану керованої системи та оточуючого середовища, а також і відшукання за неформальними критеріями раціональних стратегій прийняття управлінських рішень в умовах неповних даних. Це вимагає відшукання та адекватного використання різнорідної інформації про причинно-наслідкові залежності між взаємодіючими підсистемами. Математично-комп'ютерним інструментарієм для адекватного використання наявної різнорідної інформації є побудова граф-операторної моделі, вузлами якої є математичні моделі різних підсистем. У загальному випадку $s$-та підсистема $k$-го вузла описується системою алгебраїчних, інтегральних чи узагальнених операторних алгебро-інтегро-диференціальних рівнянь:

$$\overline{A}_{ks}(x_{ks}, z_{ks}, u_{ks}) = 0, \;\; z_{ks} = \varphi_{ks}(x,u) \in Z_{ks}, \;\; s = \overline{1, N_{ks}}, \;\; k = \overline{1, N_k},$$

що пов'язують стан $x_{ks} \in X_{ks}$ $ks$-ої підсистеми із її локальними керуваннями $u_{ks} \in U_{ks}$ та впливом $z_{ks} \in Z_{ks}$ на $ks$-ту підсистему інших підсистем. Складена із таких підсистем граф-операторна модель:

$$A(x,u) \triangleq \left(A_1(x,u), \ldots, A_{N_k}(x,u)\right) = 0, \quad x \triangleq (x_1, \ldots, x_{N_k}),$$

$$x_k \triangleq (x_{k1}, \ldots, x_{kN_{ks}}), \;\; u \triangleq (u_1, \ldots, u_{N_k}), \;\; u_k \triangleq (u_{k1}, \ldots, u_{kN_{ks}}),$$

$$A_s(x,u) \triangleq \left(\overline{A}_{1s}(x_{1s}, \varphi_{1s}(x,u), u_{1s}), \ldots, \overline{A}_{N_k s}(x_{N_k s}, \varphi_{N_k s}(x,u), u_{N_k s})\right),$$

разом із «підсистемою спостережень» $v = C(x,u)$ називається $B\delta$-адекватною на множині $V \times U$ за метрикою $\rho$, якщо система $A(x,u) = 0$, $C(x,u) = v$ визначає значення $w \triangleq B(x,u)$ із точністю $\delta > 0$, тобто, якщо для функції

$$d(u,v) \triangleq \max_{\overline{x}, \overline{\overline{x}} \in \overline{X}(u,v)} \rho\left(B(\overline{x},u), B(\overline{\overline{x}},u)\right)$$

виконується нерівність $\max\limits_{u \in U} \max\limits_{v \in V} d(u,v) \le \delta$, де $\overline{X}(u,v)$ є множиною можливих траєкторій $x$ модельованої системи при керуванні $u$ та спостереженні $v = C(x,u)$. Якщо граф-операторна модель $A(x,u) = 0$ із підсистемою спостереження $C(x,u)=v$ є $B\delta$-адекватною на множині $V\times U$, то очевидно існує розв'язуючий оператор $F$, який задовольняє нерівність

$$\max_{u\in U}\max_{v\in V}\max_{x\in\bar{X}(u,v)}\rho\left(F(u,v),B(x,u)\right)\le\delta .$$

За допомогою такого оператора $F$ оптимальне на множині $D\subset U$ значення функції $B(x,u)$ за умов $C(x,u)=v$, $A(x,u)=0$ визначається з точністю $\delta$ за допомогою суттєво простішої задачі оптимізації на множині $D$ значення $F(u,v)$. Отже, побудова оператора $F$ є практично важливою задачею у відшуканні оптимального керування складною системою. Побудова такого оператора $F$ здійснюється за допомогою відшукання на вибраній множині $\bar{\bar{X}}(u,v)$ оператора $\Re^{\delta}:\left(U\times V\times W_A\times W_C\right)\to W_B$, який для частинних функцій $A:(X\times U)\to W_A$, $B:(X\times U)\to W_B$, $C:(X\times U)\to W_C$ задовольняє нерівність

$$\max_{x\in\bar{\bar{X}}(u,v)}\rho\Big(\Re^{\delta}\left(u,v,A(x,u),C(x,u)\right),\ B(x,u)\Big)\le\delta,\ (x,u,v)\in\bar{\bar{X}}(u,v)\times U\times V .$$

***Теорема 4.*** *Якщо для множини $\bar{\bar{X}}(u,v)\supset\bar{X}(u,v)$ і для числа $\delta\ge 0$ існує оператор $\Re^{\delta}$, то граф-операторна модель $A(x,u)=0$, $C(x,u)=v$ є $B\delta$-адекватною і відхилення (за метрикою $\rho$) значення $\Re^{\delta}(u,v,0,v)$ від значення розв'язуючого оператора $F(v,u)$ не перевищує числа $\delta$.*

Звідси випливає, що відхилення оптимального на множині $D\subset U$ значення $B(x(u,v),u)$, де $x=x(u,v)$ є розв'язком граф-операторної моделі $C(x,u)=v$, $A(x,u)=0$, від оптимального на множині $D$ значення $\Re^{\delta}(u,v,0,v)$ не перевищує за метрикою $\rho$ числа $\delta$. Умови існування розв'язуючих операторів $F$ отримані у роботах, де будуються такі оператори.

Чисельні алгоритми побудови операторів $F$ суттєво спрощуються у випадку лінійних по $x$ граф-операторних моделей та лінійних по $x$ функцій $B$. Наприклад, для лінійної функції

$$B(x,u)\triangleq\sum_{i=1}^{N}(c_i,x_i)+f_{N+1}(u)$$

і для лінійної моделі $k$-го вузла

$$A_{k,k}x_k+A_{k,k-1}x_{k-1}+\ldots+A_{k,1}x_1+f_k(u)=0,\ k=\overline{1,N},$$

заданої лінійними операторами $A_{k,i}$, оператор $F$ будується за допомогою розв'язку $y(A,c)\triangleq(y_N,y_{N-1},\ldots,y_1)$ лінійної спряженої системи

$$c_k + \sum_{j=k}^{N} A_{j,k}^T y_k = 0,\ k = N, N-1, ..., 1,$$

у вигляді $F(u) = \sum_{i=1}^{N} (y_i, f_i(u)) + f_{N+1}(u)$ і тому задача оптимізації на множині $u \in D$ значення $B(x(u), u)$, визначеного на розв'язку $x = x(u)$ граф-операторної моделі, за умов відсутності підсистеми спостереження $v = C(x, u)$ зводиться до задачі безумовної оптимізації значення $F(u) \equiv F(u, v)$ на множині $D$. Звідси зокрема випливає, що для частинного випадку:

$$D \triangleq D_1 \times D_2 \times ... \times D_N,\ f_i(u) = f_i(u_i),\quad f_{N+1}(u) = \sum_{i=1}^{N} g_i(u_i),\quad u_i \in D_i,$$

оптимальним керуванням для $k$-го вузла є максимізатор

$$u_k^* \triangleq \arg\max_{u_k \in D_k}((y_k, f_k(u_k) + g_k(u_k)),\ k = \overline{1, N}.$$

***3.10. Використання методів оптимізації для побудови математичних моделей.***

Методи і чисельні алгоритми оптимізації є основним інструментарієм для відшукання та оптимізації математичних моделей причинно-наслідкових залежностей $x \to y$ між причинними факторами $x \in X$ та наслідковими $y \in Y$.



Наприклад, у найпростішій задачі відшукання лінійної причинно-наслідкової залежності $y = ax + b$ за даними натурних спостережень $(x_1, y_1)$, $(x_2, y_2), \dots, (x_m, y_m)$, невідомі параметри $a$ і $b$ лінійної моделі знаходять або за методом найменших квадратів, як мінімізатор функції

$$\varphi(a,b) \triangleq \sum_{i=1}^{m} (y_i - (ax_i + b))^2,\ (a,b) = \arg\min_{(a,b)} \varphi(a,b)$$

або за методом мінімаксного оцінювання, як мінімізатор максимального відхилення

$$\psi(a,b) \triangleq \max_i | y_i - (ax_i + b) |, (a,b) = \arg\min_{(a,b)} \psi(a,b).$$

Очевидно, у першому випадку параметри $(a,b)$ є розв'язком системи двох рівнянь:

$$\frac{\partial \varphi(a,b)}{\partial a} = 0,\ \frac{\partial \varphi(a,b)}{\partial b} = 0,$$

а у другому випадку мінімізатор опуклої функції $\psi$ можна обчислити за допомогою чисельних алгоритмів мінімізації опуклих функцій.

Аналогічно, значення коефіцієнтів $p_j$ лінійно параметричної моделі $y \triangleq \sum_{j=0}^{n} p_j f_j(x)$ можемо обчислити як мінімізатор функції:

$$\bar{\bar{\varphi}}(p_0, p_1, ..., p_n) \triangleq \sum_{i=1}^{m} (y_i - (\sum_{j=0}^{n} p_j f_j(x_i))^2 ,$$

який є розв'язком системи $(n+1)$ рівнянь $\nabla \bar{\bar{\varphi}}(p) = 0$, тобто системи

$$\frac{\partial \bar{\bar{\varphi}}(p_0, p_1, ..., p_n)}{\partial p_i} = 0,\ i = \overline{0, n} ,$$

а значення векторного параметра $p \triangleq (p_0, p_1, ..., p_n)$ нелінійної моделі $y \triangleq f(x, p)$ також можемо обчислити як мінімізатор функції

$$\bar{\bar{\bar{\varphi}}}(p_0, p_1, ..., p_n) \triangleq \sum_{i=1}^{m} (y_i - f(x_i, p))^2 ,$$

який є розв'язком системи $(n+1)$ рівнянь $\nabla \bar{\bar{\bar{\varphi}}}(p) = 0$.

Відповідно, за мінімаксними критеріями параметри $p \triangleq (p_0, p_1, ..., p_n)$ математичних моделей:

$$y \triangleq \sum_{j=0}^{n} p_j x^j ,\ y \triangleq \sum_{j=0}^{n} p_j f_j(x) ,\ y \triangleq f(x, p)$$

обчислюються відповідно, як мінімізатори функцій:

$$\bar{\psi}(p_0, p_1, ..., p_n) \triangleq \max_{i=\overline{1,m}} | y_i - \sum_{j=0}^{n} p_j x_i^j | ,$$

$$\bar{\bar{\psi}}(p_0, p_1, ..., p_n) \triangleq \max_{i=\overline{1,m}} | y_i - \sum_{j=0}^{n} p_j f_j(x_i) | ,$$

$$\bar{\bar{\bar{\psi}}}(p_0, p_1, ..., p_n) \triangleq \max_{i=\overline{1,m}} | y_i - f(x_i, p) | .$$

Типові математичні моделі динамічних процесів будують у вигляді явних чи неявних систем звичайних диференціальних рівнянь:

$$\frac{dx(t)}{dt} = f(x(t), t, p) ,\quad x(t) \in R^n ,\ t \in [0; T] ,$$

$$f(\frac{dx(t)}{dt}, x(t), t, p) = 0 ;$$

або у вигляді інтегральних рівнянь

$$x(t) = x(0) + \int_0^t f(t, \tau, x(\tau), p) d\tau ;$$

або у вигляді узагальнених інтегро диференціальних систем

$$f^0\left(x(t_0), \dot{x}(t), x(t), t, p, \int\limits_{D(t,x,p)} f^1\big(x(s), s, x(t), t, p\big)ds\right) = 0.$$

Значення шуканих параметрів $p$ динамічних моделей можемо обчислювати за даними натурних спостережень як мінімізатори функціоналів нев'язок:

$$\int_0^T \left\| \frac{d\overline{x}(t)}{dt} - f(\overline{x}(t), t, p) \right\|^2 dt, \quad \left\| \int_0^T f\left(\frac{d\overline{x}(t)}{dt}, \overline{x}(t), t, p\right) \right\|^2 dt,$$

$$\int_0^T \left\| \overline{x}(t) - \overline{x}(0) - \int_0^t f(t, \tau, \overline{x}(\tau), p)d\tau \right\|^2 dt,$$

$$\int_0^T \left\| f^0(\overline{x}(t_0), \dot{\overline{x}}(t), \overline{x}(t), t, p, \int\limits_{D(t,x,p)} f^1\big(\overline{x}(s), s, \overline{x}(t), t, p\big)ds) \right\|^2 dt$$

або як мінімізатори норм різниці

$$\int_0^T \left\| x(t,p) - \overline{x}(t) \right\|^2 dt, \quad \max_{t \in [0;T]} \left\| x(t,p) - \overline{x}(t) \right\|$$

між записаною при натурних спостереженнях траєкторією $\overline{x}(t)$, $t \in [0;T]$ та траєкторією $x(t,p)$, яка обчислена за математичною моделлю з параметрами $p$.

**Розділ 1**

# АЛГОРИТМИ ПРИЙНЯТТЯ РІШЕНЬ І ОПТИМІЗАЦІЇ В БАГАТОКРИТЕРІАЛЬНИХ ТА ІЄРАРХІЧНИХ СИСТЕМАХ

## 1.1. Прийняття рішень (вибір) і оптимізація

### 1. Критеріальний підхід до вибору альтернатив

Прийняття рішень (вибір) є дією, яка надає людській діяльності цілеспрямованості. Вибір реалізує підпорядкованість людської діяльності певній цілі чи сукупності певних цілей. В людській діяльності наступають моменти, коли подальші дії можуть бути різними, а значить призводити до різних результатів (добрих, поганих, задовільних, трагічних, і т.д.).

Будемо вважати, що ***прийняття рішень*** (*вибір*) – це дія над множиною альтернатив, в результаті якої отримується підмножина вибраних альтернатив (зокрема одна альтернатива). Звужувати множину альтернатив можливо, якщо існує спосіб порівняння альтернатив між собою і визначення переважаючих. Кожен такий спосіб називається ***критерієм переваг***. Зрозуміло, що процедурі прийняття рішень передує процедура формування множини альтернатив і завчасно визначені цілі, заради досягнення яких здійснюється вибір. Найбільш простим і часто вживаним є критеріальний підхід до визначення альтернатив, суть якого полягає в тому, що кожну альтернативу можна оцінювати конкретним дійсним числом, а порівняння альтернатив зводиться до порівняння відповідних їм дійсних чисел. Позначимо через $X$ множину всіх альтернатив, а її елементи через $x$. На множині $X$ задамо функцію $f(x)$, яка володіє властивістю, що якщо альтернатива $x_1$ переважає альтернативу $x_2$, то $f(x_1) > f(x_2)$ і навпаки. Таку функцію називають: ***критерієм, критерієм якості, цільовою функцією, функцією переваг, функцією корисності***. Якщо, крім цього, вважати, що значення критерію виражає оцінку наслідку вибору альтернативи $x$, тоді природно вибирати альтернативу $x^*$, яка дає найбільше значення критерію

$$x^* = \arg\max_{x \in X} f(x). \qquad (1.1)$$

Проте, як показує практика, оцінка альтернативи одним критерієм є занадто великим спрощенням ситуації. Життя вимагає оцінювати альтернативи не по одному, а по декількох критеріях, які якісно відрізняються між собою. Наприклад, при проектуванні літака конструкторам необхідно врахувати багато критеріїв, таких як:

1) технічні: висота польоту, швидкість, вантажопідйомність, необхідна довжина злітно-посадочної смуги, тривалість польоту, вага літака і т.д.;

2) економічні: затрати на виробництво, експлуатацію і обслуговування, конкурентоздатність;

3) екологічні: рівень шуму, забруднення атмосфери;

4) ергономічні: умови роботи екіпажу, рівень комфорту пасажирів і т.д.

Будемо вважати, що кожна альтернатива оцінюється $m$ критеріями: $f_i(x),\ i=\overline{1,m}$. Без обмеження загальності можна вважати, що всі частинні критерії $f_i(x)$ діють в одному і тому ж напрямку (інгредієнті), тобто виражають лише позитивні або лише негативні якості альтернативи. Цього легко добитися за рахунок зміни знаків окремих частинних критеріїв.

#### 2. Нормалізований мультикритерій

Відмітимо також, що з метою покращення вибору часто переходять до ***нормалізованого мультикритерію*** $f_i'(x),\ i=\overline{1,m}$, шляхом введення безрозмірних величин, приведення до однієї розмірності тощо. Наведемо найбільш часто вживані способи нормалізації:

а) зведення до безрозмірних величин:

$$f_i'(x) = f_i(x)/\rho(f_i(x)),$$

де $\rho(\cdot)$ – деяка функція;

б) зведення до однієї розмірності:

$$f_i'(x) = f_i(x)/\alpha(i),$$

де $\alpha(i)$ – деяка вагова функція;

в) зміна напрямку (інгредієнта):

$$f_i'(x) = -f_i(x) \text{ або } f_i'(x) = 1/f_i(x);$$

г) природній:

$$f_i'(x) = \frac{f_i(x) - f_{\min}(x)}{f_{\max}(x) - f_{\min}(x)},$$

де $f_{\min}(x) = \min\limits_{i\in\overline{1,m}} f_i(x),\ f_{\max}(x) = \max\limits_{i\in\overline{1,m}} f_i(x)$;

д) порівняння:

$$f_i'(x) = f_i(x)/\max_{x\in X} f_i(x);$$

е) усереднення:

$$f_i'(x) = f_i(x)/\sum_{i=1}^{m} f_i(x).$$

Припустимо, що в множині $X$ існує альтернатива $\overline{x}$, яка приймає найбільше значення по всіх $m$ критеріях:

$$f_i(\overline{x}) > f_i(x),\quad i=\overline{1,m},\ \text{для всіх } x\in X. \tag{1.2}$$

Тоді ця альтернатива $\overline{x}$ являється найкращою і проблеми вибору в такій ситуації не існує. На практиці такі випадки малоймовірні. Реальнішою є ситуація, коли найбільші значення критерії досягають на різних альтернативах. Розглянемо основні підходи до розв'язування багатокритеріальних задач.

**3. Суперкритерій**

Зведення багатокритеріальної задачі до однокритеріальної здійснюється введенням ***суперкритерію***, тобто скалярної функції векторного аргументу

$$f_0(x) = \overline{f}_0\left(f_1(x), f_2(x), ..., f_m(x)\right). \quad (1.3)$$

Після згортки розв'язується звичайна оптимізаційна задача:

$$\text{знайти } x^* = \arg\max_{x\in X} f_0(x) \quad \left(\text{або } x^* = \arg\min_{x\in X} f_0(x)\right).$$

Наведемо приклади найуживаніших згорток:

а) лінійна

$$f_0(x) = \sum_{i=1}^{m} \alpha(i) f_i(x);$$

б) максимізаційна

$$f_0(x) = \max_{i\in[1;m]} [\alpha(i) f_i(x) + \beta(i)];$$

в) мінімізаційна

$$f_0(x) = \min_{i\in[1;m]} [\alpha(i) f_i(x) + \beta(i)];$$

г) мультиплікативна

$$f_0(x) = \prod_{i=1}^{m} \alpha(i) f_i(x);$$

д) Кобба – Дугласа

$$f_0(x) = \prod_{i=1}^{m} [\alpha(i) f_i(x)]^{\beta(i)},$$

де $\alpha(i)$, $\beta(i)$, $i=\overline{1,m}$ – задані функції натурального аргументу $i$, $i=\overline{1,m}$.

Вкажемо на недоліки пов'язані з використанням суперкритерію:

- важко обґрунтувати використання певного типу згортки;
- існує проблема вибору параметрів згортки – функцій $\alpha(i)$, $\beta(i)$, $i=\overline{1,m}$;
- незначна зміна функції, яка визначає згортку, приводить до значних відхилень розв'язків задачі багатокритеріальної оптимізації.

**4. Метод головного частинного критерію**

Наведені вище недоліки згортання декількох критеріїв обумовили появу інших підходів до розв'язування задач багатокритеріальної оптимізації. Один із них полягає у використанні того факту, що частинні критерії нерівнозначні між собою, через те виділяють ***(основний) головний критерій*** $f_{i_0}(x)$, а решту критеріїв вважають ***додатковими.*** Головний критерій максимізують при умові, що додаткові критерії залишаються на заданих їм рівнях. Розв'язок задачі багатокритеріальної оптимізації знаходять як розв'язок задачі на умовний екстремум:

$$x^* = \arg\max_{x\in X} f_{i_0}(x) \tag{1.4}$$

при обмеженнях:

$$f_i(x) = c_i,\quad i = 1, 2, ..., i_0 - 1, i_0 + 1, ..., m. \tag{1.5}$$

В багатьох практичних задачах обмеження (1.5) на додаткові критерії формулюються не у формі рівностей, а у формі нерівностей (наприклад, якщо додаткові критерії характеризують вартість витрат, то замість фіксації витрат розумніше задавати їх верхній рівень) і в результаті приходимо до наступної задачі

$$x^* = \arg\max_{x\in X} f_{i_0}(x) \tag{1.6}$$

при обмеженнях:

$$f_i(x) \le c_i,\quad i = 1, 2, ..., i_0 - 1, i_0 + 1, ..., m. \tag{1.7}$$



**5. Метод послідовних поступок**

В методах оптимізації (1.4–1.5) і (1.6–1.7) різниця між основним і додатковими критеріями виглядає досить значною. В методі ***послідовних поступок*** (який приводиться нижче) ця різниця дещо пом'якшується. Припускається, що критерії пронумеровані у порядку їх важливості, так що $f_1(x)$ – найважливіший з критеріїв, а $f_m(x)$ – найменш важливий. На першому кроці розв'язується задача мінімізації критерію $f_1(x)$

$$x_1^* = \arg\max_{x\in X} f_1(x).$$

Нехай $f_1^{\max} = f_1(x_1^*)$. З практичних міркувань та прийнятої точності визначається поступка $\Delta_1 > 0$ (тобто величина, на яку зменшується досягнуте значення $f_1^{\max}$ найбільш важливого критерію, щоб за рахунок поступки постаратися наскільки це можливо, збільшити значення наступного по важливості критерію $f_2(x)$) і розв'язується задача оптимізації

$$x_2^* = \arg\max_{x\in X} f_2(x)$$

при обмеженнях:

$$f_1(x) \ge f_1^{\max} - \Delta_1.$$

Визначається $f_2^{\max} = f_2(x_2^*)$ і поступка $\Delta_2 > 0$.

На $k$ -му кроці розв'язується задача

$$x_k^* = \arg\max_{x \in X} f_k(x)$$

при обмеженнях:

$$f_1(x) \ge f_1^{\max} - \Delta_1,$$
$$f_2(x) \ge f_2^{\max} - \Delta_2,$$
$$\ldots$$
$$f_{k-1}(x) \ge f_{k-1}^{\max} - \Delta_{k-1}.$$

Якщо значення $x_m^*$ задовільне, то його приймають за розв'язок задачі багатокритеріальної оптимізації, інакше переходять на перший крок, змінюючи (збільшуючи) $\Delta_1$ і т.д.

**6. Пошук альтернативи із заданими властивостями**



Третій підхід багатокритеріального вибору відноситься до того випадку коли завчасно можуть бути вказані бажані значення частинних критеріїв (або їх границі), і задача полягає в тому, щоб знайти альтернативу, яка задовольняє цим вимогам, або чи встановити, що такої альтернативи у множині $X$ не існує і вказати альтернативу, яка підходить до поставленої цілі ближче всього.

Бажані значення критеріїв $\bar{f}_i,\ i = \overline{1,m}$ задають або точно, або у вигляді верхніх чи нижніх границь. Значення $\bar{f}_i,\ i = \overline{1,m}$ називають ***рівнями вимог***, а точку їх перетину $x^*$ в $m$ -вимірному просторі критеріїв ***ціллю*** (*опорною точкою* чи *ідеальною точкою*). Оскільки рівні вимог задаються без точного знання структури множини $X$, то цільова точка може лежати як всередині, так і поза множиною $X$ (досяжна чи недосяжна ціль). Оптимізаційна задача полягає в побудові послідовності альтернатив $x^k,\ k = 0,1,...,$ яка б в границі наближалася до $x^*$. Для цього вводиться числова міра близькості між альтернативою $x$ і ціллю $x^*$, тобто між векторами $f(x) = \left(f_1(x), f_2(x), ..., f_m(x)\right)$ і $\bar{f} = (\bar{f}_1, \bar{f}_2, ..., \bar{f}_m)$. Використовуються такі функції відстані:

1) $d_k(x) = \bar{d}_k\left(f(x), \bar{f}\right) = \left[\sum_{i=1}^{m} w_i \left|f_i(x) - \bar{f}_i\right|^k\right]^{1/k}$, де параметр $k \in \mathrm{N}$, $w_i,\ i = \overline{1,m}$ – вагові коефіцієнти;

2) $d(x) = \bar{d}\left(f(x), \bar{f}\right) = \min_{i \in \overline{1,m}} \alpha_i \left(f_i(x) - \bar{f}_i\right) + \alpha_{m+1} \sum_{i=1}^{m} \alpha_i \left(f_i(x) - \bar{f}_i\right),$

причому робиться припущення про невід'ємність різниці $f_i(x) - \bar{f}_i$, $i = \overline{1,m}$; $\alpha_i$, $i = \overline{1,m}$ – коефіцієнти, які призводять доданки до однакової розмірності і одночасно враховують різну важливість критеріїв; коефіцієнт $\alpha_{m+1}$ виражає відношення до того, що важливіше – зменшити близькість до цілі будь-якого із частинних критеріїв чи сумарну близькість всіх критеріїв до цільових значень.

У випадку, коли частина бажаних значень критеріїв є обмеженнями частинних критеріїв знизу:

$$f_i(x) \ge \bar{f}_i, \quad i = \overline{1, m'},$$

інша частина є обмеженням зверху:

$$f_i(x) \le \bar{f}_i, \quad i = \overline{m'+1, m''},$$

а решту обмежень виконуються як рівності:

$$f_i(x) = \bar{f}_i, \quad i = \overline{m''+1, m},$$

то функцію відстані можна вибирати в наступному вигляді:

$$d(x) = \bar{d}\left(f_i(x), \bar{f}\right) = \min_{i \in [1;m]} F\left(f_i(x), \bar{f}_i\right) + \alpha_{m+1} \sum_{i=1}^{m} F\left(f_i(x), \bar{f}_i\right),$$

де

$$F\left(f_i(x), \bar{f}_i\right) = \begin{cases} \alpha_i\left(f_i(x) - \bar{f}_i\right), & \text{якщо } 1 \le i \le m', \\ \alpha_i\left(\bar{f}_i - f_i(x)\right), & \text{якщо } m'+1 \le i \le m'', \\ \alpha_i \min\left\{f_i(x) - \bar{f}_i, \bar{f}_i - f_i(x)\right\}, & \\ & \text{якщо } m''+1 \le i \le m. \end{cases}$$

**7. Метод бажаної точки**

Для кожного критерію $f_i(x)$, $i = \overline{1,m}$ розраховується найбільше і найменше його значення на множині альтернатив $X$:

$$f_i^{\max} = \max_{x \in X} f_i(x); \; f_i^{\min} = \min_{x \in X} f_i(x), \; i = \overline{1,m}.$$

Далі переходять до нових (нормованих, безрозмірних) критеріїв $w_i(x)$, $i = \overline{1,m}$:

$$w_i(x) = \frac{f_i^{\max} - f_i(x)}{f_i^{\max} - f_i^{\min}}, \quad i = \overline{1,m}.$$

На $k$-му кроці алгоритму на основі аналізу вибираються бажані значення критеріїв:

$$h_i^k \in [f_i^{\min}, f_i^{\max}], \quad i = \overline{1,m}$$

і на цій основі розраховуються значення:

$$w_i^k = \frac{f_i^{\max} - h_i^k}{f_i^{\max} - f_i^{\min}}, \quad i = \overline{1,m}.$$

Потім обчислюють вагові коефіцієнти нових критеріїв:

$$\alpha_i^k = \frac{\prod_{j=1, j\neq i}^{m} w_j^k}{\sum_{j=1}^{m} \prod_{l=1, l\neq j}^{m} w_l^k}, \quad i = \overline{1,m}.$$

Ефективна альтернатива $x^k$ знаходиться як розв’язок однокритеріальної задачі

$$x^k = \arg \max_{x \in X} \min_{i \in [1,m]} \alpha_i^k w_i(x).$$

В знайденій точці $x^k$ обчислюється значення всіх критеріїв

$$\left(f_1(x^k), f_2(x^k), ..., f_m(x^k)\right).$$

Якщо цей вектор значень критеріїв задовольняє особу, що приймає рішення, то $x^k$ – шукана альтернатива; інакше здійснюється перехід на $(k+1)$-й крок, тобто на основі аналізу вибираються нові бажані значення критеріїв $h_i^{k+1}$, $i = \overline{1,m}$ і т.д.



### 8. Прийняття рішень в умовах невизначеності

Будемо вважати, що задана функція $\varphi(x, y)$, яка визначена на декартовому добутку множин $X$ та $Y$, де $X$ – множина альтернатив, а $Y$ – множина неконтрольованих факторів (збурень). Функція (критерій) $\varphi(x, y)$ характеризує якість альтернативи $x \in X$ при певному значенні неконтрольованих факторів $y \in Y$. Вважаємо, що критерій $\varphi(x, y)$ виражений в позитивному інгредієнті й характеризує позитивну якість стратегії $x$ (наприклад прибуток, дохід, інтегральний рівень життя тощо), тому задача прийняття рішень полягає у виборі альтернативи $x$, яка буде робити в деякому розумінні критерій “більшим” при різних $y \in Y$.

Якщо неконтрольовані фактори фіксовані, тобто множина $Y$ складається із одного елемента $y^0$, $Y = \{y^0\}$ (такі одноелементні множини називають ***синглетонами***), то пошук найкращої альтернативи – це звичайна оптимізаційна задача знаходження вектора $x^* \in X$, для якого критерій $\varphi(x^*, y^0)$ приймає максимальне можливе значення.

Якщо неконтрольованих факторів багато, тоді можливий варіант, коли найкраща альтернатива для одного неконтрольованого фактора буде найгіршою для іншого неконтрольованого фактора. В такому випадку найчастіше використовуються дві оцінки ефективності альтернатив: ***гарантовані*** $\Phi(x)$ та ***середні*** $S(x)$.

Гарантована оцінка $\Phi(x)$ ефективності альтернатив орієнтована на найгіршу дію неконтрольованих факторів

$$\Phi(x) = \min_{y \in Y} \varphi(x, y) . \tag{1.8}$$

В тому випадку, коли критерій $\varphi(x, y)$ виражений в негативному інгредієнті (характеризує, наприклад, штрафи, витрати, забруднення, тощо), то гарантована оцінка теж орієнтована на найгіршу дію неконтрольованих факторів і має вигляд

$$\Phi(x) = \max_{y \in Y} \varphi(x, y) . \tag{1.9}$$

Оптимальну гарантовану альтернативу $\bar{x}^*$ можна обчислити, розв'язуючи максимінну задачу

$$\Phi(\bar{x}^*) = \max_{x \in X} \Phi(x) = \max_{x \in X} \min_{y \in Y} \varphi(x, y)$$

у випадку (1.8), і розв'язуючи мінімаксну задачу

$$\Phi(\bar{x}^*) = \min_{x \in X} \Phi(x) = \min_{x \in X} \max_{y \in Y} \varphi(x, y)$$

у випадку (1.9).



Для знаходження середньої оцінки ефективності $S(x)$ усереднюють значення критерію ефективності $\varphi(x, y)$ по всіх значеннях неконтрольованих факторів $y$, $y \in Y$. Спочатку припустимо, що множина неконтрольованих факторів $Y$ складається із скінченого набору

$$Y = \{ y^i, \quad i = \overline{1, m} \} . \tag{1.10}$$

Тоді середня оцінка визначається формулою:

$$S(x) = \sum_{i=1}^{m} \rho_i \varphi(x, y^i), \tag{1.11}$$

де $\rho_i$, $i = \overline{1, m}$ – вагові коефіцієнти (в основному їх вибирають із таких умов:

$$\rho_i \geq 0, \ i = \overline{1, m}; \ \sum_{i=1}^{m} \rho_i = 1).$$

Якщо множина неконтрольованих факторів злічена (тобто в рівності (1.10) $m \to \infty$), тоді вагові коефіцієнти $\rho_i$, $i = 1, 2, \ldots$ вибираються таким чином, щоб ряд (1.11) збігався.

Розглянемо третій випадок, коли множина неконтрольованих факторів $Y$ є проміжком дійсної осі (скінченим чи нескінченим). Тоді середня оцінка ефективності $S(x)$ визначається таким співвідношенням:

$$S(x) = \int_{Y} \rho(y) \varphi(x, y) dy , \tag{1.12}$$

де вагова функція $\rho(y)$ задовольняє умовам:

$$\rho(y) \ge 0, \quad y \in Y; \qquad \int_Y \rho(y)dy = 1.$$

***Оптимальною у середньому*** називається альтернатива $\hat{x}^*$, яка задовольняє умову

$$S(\hat{x}^*) = \max_{x \in X} S(x).$$

Наприкінці відмітимо, що оптимальні альтернативи у середньому використовують і тоді, коли неконтрольовані фактори є випадковою величиною чи випадковим вектором. Критерій ефективності є випадковою величиною $\varphi(x, y)$, а за осереднену оцінку альтернативи $x$ приймають математичне сподівання

$$E_y \varphi(x, y) = S(x).$$

При цьому $S(x)$ задається:

– формулою (1.11) при умові, що неконтрольований випадковий фактор $y$ приймає значення $y^i$ з ймовірністю $\rho_i$;

– формулою (1.12), коли неконтрольований фактор $y$ є випадковою величиною із щільністю розподілу $\rho(y)$ на проміжку $Y$.



**9. Приклади постановок задач багатокритеріальної оптимізації**

***9.1. Задача проектування оптимального програмного комплексу.***

При проектуванні програмного комплексу (ПК) необхідно забезпечити виконання декількох вимог: зменшити вартість ПК, збільшити точність задання вхідних даних, скоротити об'єм оперативної пам'яті, зменшити час роботи ПК, зменшити завантаження каналів зв'язку між ЕОМ і зовнішніми запам'ятовуючими пристроями і т.д. Припускається, що ПК повинен реалізувати множину операцій $a = (a_1, a_2, ..., a_n)$. Під операцією розуміється, наприклад, розв'язування лінійних чи нелінійних алгебраїчних рівнянь, систем лінійних чи нелінійних диференціальних рівнянь, знаходження екстремумів функцій певного типу, сортування інформації, пошук інформації і т.д.

Кожна із операцій $a_i \in a$ може бути реалізована будь-якою програмою із заданої множини $P_i = \left(p_{i1}, p_{i2}, ..., p_{ik_i}\right)$, $i = \overline{1,n}$. Кожна програма $p_{ij}$ характеризується своїми ознаками, які впливають на вимоги до ПК. Програмний комплекс представляє собою набір програм $P = \left(p_{1j_1}, p_{2j_2}, ..., p_{nj_n}\right)$, де $p_{ij_i} \in P_i$, $i = \overline{1,n}$.

Для оцінки якості програмного комплексу $P = \left(p_{1j_1}, p_{2j_2}, ..., p_{nj_n}\right)$ вводиться векторний критерій $\varphi(P) = \left(\varphi_1(P), \varphi_2(P), ..., \varphi_m(P)\right)$.

Відображення $\varphi : P \to R^m$ задає правило, по якому кожному набору із $n$ програм відповідає векторна оцінка виконання вимог до ПК. Таким чином, задача проектування оптимального програмного комплексу є багатокритеріальною оптимізаційною задачею.

***9.2. Трьохрівнева задача керування гнучким автоматизованим виробництвом.***

Гнучке автоматизоване виробництво (ГАВ) являє собою систему, яка складається із підсистем:

- автоматизовані технологічні модулі (станки, лінії, ділянки);
- автоматизований транспорт;
- автоматизовані склади.

Керування роботою цих підсистем і здійснення зв'язків між ними забезпечує підсистема керування ГАВ (ПК ГАВ). За її допомогою здійснюється запуск, керування і контроль за роботою технологічного обладнання, синхронізація виконуваних робіт, оптимізація завантаження обладнання, формування графіку робіт транспортних засобів, автоматизованих складів і т.д. Найбільш поширеною ПК ГАВ є трьохрівнева система керування (рис.1.1).



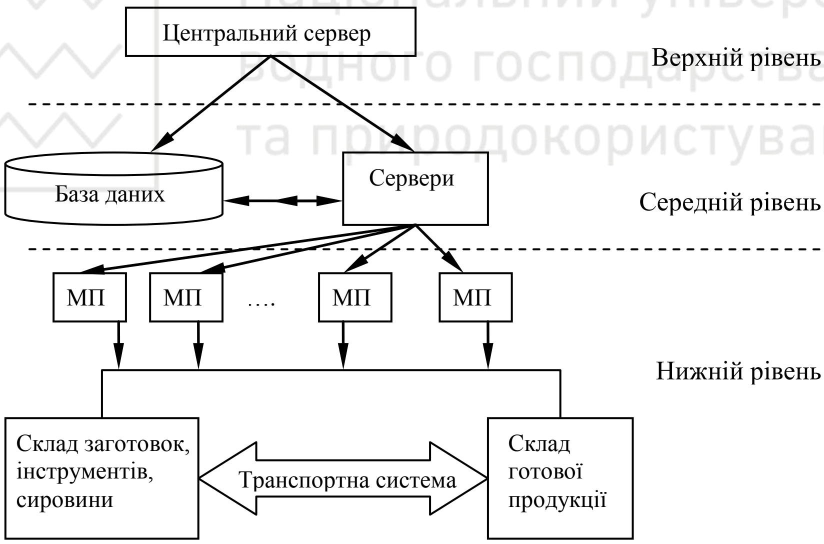



**Рис. 1.1.** Принципова схема трьохрівневої системи керування ГАВ

Верхній рівень розв'язує задачі організаційно-економічного характеру й приймає довгострокові рішення: проводить розрахунок позмінних завдань по кожному станку, завдань по технологічній підготовці ділянки; обліковує запас заготовок, інструменту і сировини на складі; накопичує інформацію для різних служб цеху.

Середній рівень здійснює контроль за роботою мікропроцесорних систем (МП); приймає оперативні рішення у відповідності з надходженням від підсистем нижнього рівня інформації; виробляє керуючі дії на ці підсистеми.

За допомогою мікропроцесорних систем нижній рівень забезпечує безпосереднє управління технологічним процесом.

Аналіз і управління роботою ГАВ потребує розв'язування значної кількості оптимізаційних багатокритеріальних задач, задач мережевого планування, транспортних задач, задач розміщення і т.д.

## 1.2. Оптимізація в системах з ієрархічною структурою

***Системами з ієрархічною структурою*** називають сукупність підсистем, які мають послідовне вертикальне розташування з встановленим пріоритетом дій і прийняття рішень, причому результати дій підсистем верхнього рівня залежать від дій підсистем нижчих рівнів. На діяльність підсистем будь-якого рівня (крім верхнього) безпосередню дію справляють підсистеми, які розміщені на вищих рівнях. Хоча така дія направлена зверху вниз, успіх дії системи в цілому і кожного рівня залежить від поведінки всіх елементів системи. Поняття пріоритету дій вказує на те, що вплив підсистем верхнього рівня передує діям більш низьких рівнів. Тому успішність роботи  підсистем вищестоящих рівнів залежить не тільки від власних дій, але й від реакцій підсистем нижніх рівнів на цей вплив.

Підсистему найвищого рівня називають ***центром***, а підсистеми нижчих рівнів називають ***елементами***. В системах керування елементам надано право виробляти певні керуючі дії, приймати рішення. Тому, поряд з ієрархією системи, кажуть про ієрархічну структуру керування. Ієрархічна структура керування в складній системі являє собою сукупність рівнів керування, які слідують один за одним в порядку певного пріоритету. Між елементами різних рівнів ієрархії існують як вертикальні, так і горизонтальні зв'язки.

Поява ієрархічної структури в системах керування і прийняття рішень обумовлена наявністю великого об'єму інформації про керовані процеси в системі, неможливістю обробки цієї інформації і прийняття рішень одним центром керування, а також існуючою в реальних системах децентралізацією процесу прийняття рішень, коли елементи, які підпорядковуються центру, виробляють керуючі дії, виходячи із вказівок центру і з урахуванням власних інтересів.

Розглянемо математичну модель *дворівневої ієрархічної системи керування*. Нехай центру $Q_0$ підпорядковані елементи системи управління $Q_1, Q_2, ..., Q_m$, які надалі будемо називати ***підсистемами***. Центр $Q_0$

виробляє керуючу дію $u=\left(u^1,...,u^m\right)$ і повідомляє її підсистемам нижчого рівня $Q_1,Q_2,...,Q_m$, які в свою чергу вибирають власні керування $v^1\left(u^1\right),v^2\left(u^2\right),...,v^m\left(u^m\right)$ із деяких множин допустимих керувань відповідно $V^1\left(u^1\right),V^2\left(u^2\right),...,V^m\left(u^m\right)$, що залежать від вибору керування центром $Q_0$. Позначимо через $U$ множину допустимих керувань центра $Q_0$. Керування $u\in U$ будемо називати ***допустимим***, якщо для будь-якого $i=\overline{1,m}$ всі множини $V_i\left(u^i\right)$, $i=\overline{1,m}$ не є порожніми.

**1. Тривіальний випадок**

Якщо для будь-якого $u\in U$ всі множини $V_i\left(u^i\right)$, $i=\overline{1,m}$ складаються із єдиних керувань, то в цьому випадку центр володіє повною інформацією про реакцію підсистем нижчого рівня на своє керування.

Нехай $\varphi_0(u,v)$ – критерій оптимальності центра $Q_0$ (тут $u=\left(u^1,u^2,...,u^m\right)$, $v=\left(v^1,v^2,...,v^m\right)$), а $\varphi_i\left(u^i,v^i\right)$, $i=\overline{1,m}$ – критерії оптимальності підсистем $Q_1,Q_2,...,Q_m$. Центр і підсистеми вибираючи свої керування, намагаються максимізувати свої критерії. У випадку, коли всі множини $V_i\left(u^i\right)$, $i=\overline{1,m}$ складаються із одного елемента, тобто $V_i\left(u^i\right)=\left\{v_i\left(u^i\right)\right\}$, центр вибирає своє керування з умови

$$\varphi_0\left(u^*,v\left(u^*\right)\right)=\max_{u\in U}\varphi_0\left(u,v\left(u\right)\right) \tag{1.13}$$

(тут $v\left(u^*\right)=\left(v^1\left(u^{1*}\right),\ldots,v^m\left(u^{m*}\right)\right)$),

а значення критеріїв підсистем будуть такими:

$$\varphi_1\left(u^{1*},v^1\left(u^{1*}\right)\right),\ \varphi_2\left(u^{2*},v^2\left(u^{2*}\right)\right),...,\ \varphi_m\left(u^{m*},v^m\left(u^{m*}\right)\right).$$

Отже, в тривіальному випадку центр, вибираючи оптимальне керування $u^*$ (з точки зору центру), однозначно визначає значення критеріїв центра і всіх підсистем.

**2. Загальний випадок**

Природно припускати, що вибір центром $Q_0$ допустимого керування $u\in U$ визначає не єдине керування кожної підсистеми, тобто кожна із множин $V_i\left(u^i\right)$, $i=\overline{1,m}$ складається більше ніж із одного елемента. Визначимо множину $G_i\left(u^i\right)$ оптимальних реакцій $i$-ої підсистеми на керування $u\in U$ центра наступним чином:

$$G_i\left(u^i\right)=\left\{v^i\in V_i\left(u^i\right)\mid\varphi_i\left(u^i,v^i\right)\ge\varphi_i\left(u^i,\bar v^i\right),\ \forall\ \bar v^i\in V_i\left(u^i\right)\right\},\quad i=\overline{1,m}.$$

Якщо множини $G_i\left(u^i\right)$, $i=\overline{1,m}$ не є одноелементними, то центр $Q_0$, приймаючи певне рішення (вибираючи керування $u \in U$ ), знаходиться в умовах невизначеності. Роблячи певні припущення про характер реакцій підсистем $Q_1, Q_2, ..., Q_m$ на керування $u$ центру $Q_0$, приходимо до різних постановок задач оптимізації в дворівневих ієрархічних системах.

***2.1. Принцип гарантованого результату.***

Будемо припускати, що у відповідь на керування $u \in U$ центру $Q_0$, підсистеми $Q_1, Q_2, ..., Q_m$ вибирають управління $v^i$, $i=\overline{1,m}$, які задовольняють нерівність:

$$\varphi_0\left(u, v\right) \le \varphi_0\left(u, \bar{v}\right) \ \ \forall \bar{v}^i \in G_i\left(u^i\right), \ \ i=\overline{1,m}, \tag{1.14}$$

тобто підсистеми $Q_1, Q_2, ..., Q_m$ діють найгіршим чином для центра $Q_0$. Тоді для центра найкращим буде вибір керування $\bar{u} \in U$ , яке задовольняє наступним співвідношенням:

$$\varphi_0\left(\bar{u}, \bar{v}\right) = \min_{v \in G\left(\bar{u}\right)} \varphi_0\left(\bar{u}, v\right) \ge \min_{v \in G\left(u\right)} \varphi_0\left(u, v\right) \quad \forall u \in U, \tag{1.15}$$

де $G\left(u\right) = \prod_{i=1}^{m} G_i\left(u^i\right)$.

Вибір центром $Q_0$ керування $\bar{u} \in U$ згідно умов (1.14)–(1.15) називають ***принципом гарантованого результату***.

***2.2. Прийняття рішень в умовах доброзичливості.***

Якщо припустити, що проявляючи доброзичливість до центру, підсистеми $Q_1, Q_2, ..., Q_m$ у відповідь на керування $u \in U$ центру вибирають найкращі для нього керування $\hat{v}\left(u^i\right)$, $i=\overline{1,m}$, тобто

$$\varphi_0\left(u, \hat{v}\left(u\right)\right) = \max_{v \in G\left(u\right)} \varphi_0\left(u, v\right). \tag{1.16}$$

Тоді природно вважати, що центр $Q_0$ буде вибирати своє керування $\hat{u}$ з умови

$$\varphi_0\left(\hat{u}, \hat{v}\left(\hat{u}\right)\right) = \max_{u \in U} \varphi_0\left(u, \hat{v}\left(u\right)\right) = \max_{u \in U} \max_{v \in G\left(u\right)} \varphi_0\left(u, v\right). \tag{1.17}$$

Оскільки, для будь-якого $u \in U$ виконується нерівність

$$\max_{v \in G\left(u\right)} \varphi_0\left(u, v\right) \ge \min_{v \in G\left(u\right)} \varphi_0\left(u, v\right),$$

то справджується також і наступна нерівність:

$$\max_{u \in U} \max_{v \in G\left(u\right)} \varphi_0\left(u, v\right) \ge \max_{u \in U} \min_{v \in G\left(u\right)} \varphi_0\left(u, v\right),$$

яка стверджує, що діючи в умовах доброзичливості, центр отримує більше значення критерію, ніж в умовах гарантованого результату.

## 3. Приклади ієрархічних систем керування

### *3.1. Приклад 1. (Задача розподілу ресурсів).*

Адміністративний центр $Q_0$ розподіляє обмежений об'єм ресурсів між підлеглими йому підрозділами $Q_1, Q_2, ..., Q_m$, які використовують цей ресурс для виробництва продукції з врахуванням власних критеріїв.

Нехай центр виділяє для $i$-го підрозділу набір ресурсів із $l$ найменувань, який позначено вектором $u^i = (u_1^i, u_2^i, ..., u_l^i)$, тобто центр вибирає систему із $m$ векторів

$$u = (u^1, u^2, ..., u^m),$$

які задовольняють умовам:

$$u^i \geq 0,\ i = \overline{1,m},\ \sum_{i=1}^{m} u^i \leq b,$$

де $b$ – вектор максимально можливих об'ємів ресурсів центра. Кожен із підрозділів $Q_i$, знаючи вибір центру $Q_0$, визначає вектор $v^i = (v_1^i, v_2^i, ..., v_n^i)$, який задовольняє нерівностям:

$$v^i A_i \leq u^i + g^i,\ v^i \geq 0,\ g^i \geq 0,\ A_i \geq 0. \qquad (1.18)$$

Тут $v^i$ інтерпретується як виробнича програма підрозділу $Q_i$ по $n$ видах виготовлюваної продукції; $A_i$ – виробнича (технологічна) матриця підрозділу $Q_i$; $g^i$ – вектор власних ресурсів підрозділу $Q_i$. Критерій центра $Q_0$ визначимо наступним чином:

$$\varphi_0(u, v) = \varphi_0\left(u^1, ..., u^m, v^1(u^1), ..., v^m(u^m)\right) = \sum_{i=1}^{m} \left\langle a^i, v^i(u^i) \right\rangle,$$

де $u = (u^1, u^2, ..., u^m)$ – керування центра $Q_0$; $v^i(u^i)$ – виробнича програма підрозділу $Q_i$, яка задовольняє нерівностям (1.18); $a^i = (a_1^i, ..., a_n^i) \geq 0$ – вектор корисності центру $Q_0$ від продукції, яка випускається підрозділом $Q_i$; $\left\langle a^i, v^i(u^i) \right\rangle$ – скалярний добуток векторів $a^i$ та $v^i(u^i)$.

Критерій $i$-го виробничого підрозділу $Q_i$ визначимо так:

$$\varphi_i\left(u^i, v^i(u^i)\right) = \left\langle c^i, v^i(u^i) \right\rangle,\quad i = \overline{1,m},$$

де $c^i = (c_1^i, c_2^i, ..., c_n^i) \geq 0$ – вектор корисності підрозділу $Q_i$ від своєї продукції. І центр і кожен $i$-й підрозділ намагаються максимізувати свій критерій.

Пропонується наступна процедура прийняття рішення. Нехай $v_*^i(u^i)$, $i = \overline{1,m}$ – розв'язок задачі параметричного програмування (параметром вважається вектор $u^i$):

$$v_*^{\,i}\left(u^i\right)=\arg\max_{v^i\in G_i\left(u^i\right)}\left\langle c^i,v^i\right\rangle, \tag{1.19}$$

де $G_i\left(u^i\right)=\left\{v^i \mid v^i\ge 0, v^iA_i\le u^i+g^i, u^i\ge 0, g^i\ge 0\right\}$, а $u_*=\left(u_*^1,u_*^2,...,u_*^m\right)$ – розв'язок задачі:

$$u_*=\arg\max_{u\in U}\sum_{i=1}^{m}\left\langle a^i, v_*^{\,i}\left(u^i\right)\right\rangle, \tag{1.20}$$

де $U=\left\{u\middle|u^i\ge 0,\ \sum_{i=1}^{m}u^i\le b\right\}$.

Неважко переконатися в справедливості таких нерівностей:

$$\varphi_0\left(u_*,v^*(u_*)\right)\ge\varphi_0\left(u,v^*(u)\right),\ \ u\in U, \tag{1.21}$$

$$\varphi_i\left(u_*^{\,i},v_*^{\,i}\left(u_*^{\,i}\right)\right)\ge\varphi_i\left(u_*^{\,i},v^i\right),\ \ v^i\in V^i\left(u_*^{\,i}\right),\ \ i=\overline{1,m}. \tag{1.22}$$

Нерівності (1.21), (1.22) вказують на те, що ні центру $Q_0$, ні підрозділам $Q_i$, $i=\overline{1,m}$ не вигідно відступати від ситуації

$$\left(u_*^{\,1},...,u_*^{\,m},v_*^{\,1}\left(u_*^{\,1}\right),...,v_*^{\,m}\left(u_*^{\,m}\right)\right).$$

Така ситуація в теорії ігор називається ***рівновагою за Нешем***.

В даному прикладі центр впливає тільки на множину допустимих керувань підлеглих підрозділів і не впливає на їх критерії.

***3.2. Приклад 2. (Задача нормування шкідливих викидів).***

Рівень шкідливих викидів в певному регіоні описується скалярною функцією:

$$q(v)=q\left(v^1,...,v^m\right)=\sum_{i=1}^{m}a^iv^i,\ \ 0\le v^i\le b^i,\ \ i=\overline{1,m},$$

де $a^i$, $i=\overline{1,m}$ – вагові коефіцієнти; $v^i$, $i=\overline{1,m}$ – об'єми викидів шкідливих речовин $i$-м підприємством. Залежність між об'ємами шкідливих викидів і затратами $i$-го підприємства на переробку не скинутих відходів задається функцією:

$$h_i\left(v^i\right)=c^i\left(b^i-v^i\right),\ \ c_i>0,\ \ i=\overline{1,m}.$$

Якщо рівень забруднень в регіоні перевищує величину $q_{\max}$, то на підприємства накладаються штрафи $s^i>0$, $i=\overline{1,m}$. Кожне підприємство намагається мінімізувати свій критерій:

$$\varphi_i\left(b^i,s^i,v^1,...,v^m\right)=\begin{cases}c^i\left(b^i-v^i\right), & \text{якщо } \sum_{i=1}^{m}a^iv^i\le q_{\max},\\ c^i\left(b^i-v^i\right)+s^i, & \text{якщо } \sum_{i=1}^{m}a^iv^i>q_{\max}.\end{cases} \tag{1.23}$$

Адміністративному центру доручено здійснювати контроль за рівнем забруднень шляхом встановлення обмежень шкідливих викидів підприємств і накладанням штрафів за забруднення, тобто встановлювати значення величин $b^1,...,b^m, s^1,...,s^m$. Центр намагається максимізувати свій критерій

$$\varphi_0\left(v^1,...,v^m\right)=\begin{cases}1, & \text{якщо } \sum_{i=1}^{m} a^i v^i \le q_{\max}, \\ 0, & \text{якщо } \sum_{i=1}^{m} a^i v^i > q_{\max}\end{cases} \qquad (1.24)$$

шляхом вибору величин $b^1,...,b^m$, $s^1,...,s^m$. Нехай об'єми викидів $v=\left(v^1,...,v^m\right)$ такі, що виконується рівність

$$\sum_{i=1}^{m} a^i v^i = Q. \qquad (1.25)$$

Тоді з (1.23), (1.24) випливає:

$$\varphi_i\left(b^i, s^i, v^1,...,v^m\right)=c^i\left(b^i - v^i\right),\ i=\overline{1,m},\ \ \varphi_0\left(v^1,...,v^m\right)=1. \qquad (1.26)$$

Дослідимо при яких значеннях $b^i, s^i$ вказана точка є точкою мінімуму функції $\varphi_i\left(b^i, s^i, v^1,...,v^m\right)$ за аргументом $v^i$. Для цього фіксуємо значення викидів $v^1,...,v^{i-1}, v^{i+1},...,v^m$ та покладемо $v^i = b^i$. Тоді отримуємо:

$$\varphi_i\left(b^i, s^i, v^1,...,v^{i-1}, b^i, v^{i+1},...,v^m\right)=s^i,\ i=\overline{1,m}.$$

Із формули (1.23), враховуючи (1.25), (1.26) для $\overline{v^i}\in\left(v^i, b^i\right)$ і $\tilde{v}^i\in\left[0, v^i\right]$, випливають такі нерівності:

$$\varphi_i\left(b^i, s^i, v^1,...,v^{i-1}, \overline{v^i}, v^{i+1},...,v^m\right) >$$
$$> \varphi_i\left(b^i, s^i, v^1,...,v^{i-1}, b^i, v^{i+1},...,v^m\right),$$
$$\varphi_i\left(b^i, s^i, v^1,...,v^{i-1}, v^i, v^{i+1},...,v^m\right) <$$
$$< \varphi_i\left(b^i, s^i, v^1,...,v^{i-1}, \tilde{v}^i, v^{i+1},...,v^m\right).$$

Отже, для того, щоб вектор $\left(v^1,...,v^m\right)$ був точкою мінімуму за кожним із аргументів, достатньо, щоб виконувалися нерівності

$$c^i\left(b^i - v^i\right) < s^i,\ i=\overline{1,m}.$$

Розв'язком задачі про нормування шкідливих викидів є будь-який вектор $\left(b^1,...,b^m, s^1,...,s^m, v^1,...,v^m\right)$, який задовольняє умовам:

$$\sum_{i=1}^{m} a^i v^i = q_{\max},$$
$$c^i\left(b^i - v^i\right) < s^i,\ \ b^i > 0,\ \ s^i > 0,\ \ i=\overline{1,m}.$$

В даному прикладі центр впливає як на область допустимих розв’язків підсистем (підприємств), так і на їх критерії (функції затрат).

***3.3. Приклад 3. (Задача управління економічною системою за допомогою штрафів і доплат (заохочень)).***

Припускаємо, що $i$-те $(i=\overline{1,m})$ підприємство виробляє продукцію $v^i$, яка задається виробничою функцією Кобба–Дугласа:

$$v^i = \alpha_i x_i^{k_i} L_i^{1-k_i}, \;\; k_i \in [0,1], \;\; i=\overline{1,m}\;, \tag{1.27}$$

де $x_i$ – об’єм фондів, $L_i$ – кількість робочої сили, $\alpha_i$, $k_i$ – деякі характеристики $i$-го підприємства. Для спрощення викладок надалі вважається, що $k_i = \frac{1}{2}$, $i=\overline{1,m}.$ Цільову функцію (критерій) $i$-го підприємства задамо в такому вигляді:

$$\varphi_i(L_i) = c_i \alpha_i x_i^{1/2} L_i^{1/2} - \omega_i L_i + s_i(v^i), \;\; i=\overline{1,m}, \tag{1.28}$$

де $c_i$ – ціна продукту $i$-го підприємства , $\omega_i$ – середня ставка заробітної плати $i$-го підприємства, $s_i(v^i)$ – додаткова доплата (або штраф), яка виплачується центром $i$-му підприємству (який $i$-те підприємство платить центру) в залежності від об’ємів випуску продукції. Будемо вважати величини $x_i$, $i=\overline{1,m}$ фіксованими, тому (1.28) можна записати так:

$$\varphi_i(L_i) = \bar{\alpha}_i L_i^{1/2} - \omega_i L_i + s_i(v^i), \;\; i=\overline{1,m}, \tag{1.29}$$

де $\bar{\alpha}_i = c_i \alpha_i x_i^{1/2}$.

Підприємства, вибираючи керування, максимізують свій критерій. Центр зацікавлений, щоби підприємства, приймаючи рішення, знали заохочення чи штрафи, тобто вигляд функції $s_i(v^i)$. Для знаходження точки максимуму функції (1.29) при кожному $i$ знайдемо її похідну та прирівняємо до нуля:

$$\frac{\partial \varphi_i(L_i)}{\partial L_i} = \frac{1}{2}\bar{\alpha}_i L_i^{-1/2} - \omega_i + \frac{ds_i}{dv^i}\cdot\frac{dv^i}{dL_i} = 0, \quad i=\overline{1,m}. \tag{1.30}$$

З рівняння (1.30) можна визначити значення $L_i^*$, на якому досягається максимум функції $\varphi_i(L_i)$. Зрозуміло, що $L_i^*$ залежить від вигляду функції доплат (штрафів) $s_i(v^i)$, тобто $L_i^*$ є фунціоналом від $s_i(v^i)$: $L_i^* = L_i^*[s_i(v^i)]$. У відповідності із формулою (1.27) оптимальний об’єм продукції $i$-го підприємства також буде функціоналом від функції $s_i(v^i): v^{i*} = v^{i*}[s_i(v^i)].$

Задача центру полягає у виборі таких функцій доплат (штрафів) $s_i(v^i)$, $i=\overline{1,m}$, які доставляють максимум критерію центра $\phi(v^1,\ldots,v^m)$. Тобто з врахуванням оптимальної поведінки підприємств критерій центра можна записати у такому вигляді

$$\phi = \phi(v^{1*}[s_1], v^{2*}[s_2], ..., v^{m*}[s_m]). \qquad (1.31)$$

Задача знаходження екстремального розв’язку функціоналу (1.31) є складною і нестандартною оптимізаційною задачею. Для її розв’язування необхідні спеціальні оптимізаційні методи.

Позначимо через $\hat{v}^i$, $i = \overline{1,m}$ об’єми продукції підприємств, на яких критерій центра $\phi = \phi(v^1, ..., v^m)$ приймає максимальне значення, і задамо функції штрафу підприємств в наступному вигляді:

$$s_i(v^i) = \lambda_i (v^i - \hat{v}^i)^2 - c_i \alpha_i x_i L_i^{1/2} + \omega_i L_i, \quad i = \overline{1,m}, \qquad (1.32)$$

де $\lambda_i$, $i = \overline{1,m}$ – довільні від’ємні числа. Із співвідношень (1.28) і (1.32) отримуємо такий критерій $i$-го підприємства:

$$\varphi_i(v^i) = \lambda_i (v^i - \hat{v}^i)^2, \quad \lambda_i < 0, \quad i = \overline{1,m}.$$

Звідси випливає, що оптимальним для $i$-го підприємства є випуск продукції $v^i = \hat{v}^i$, $i = \overline{1,m}$, тобто при виборі центром функції штрафу (1.32) інтереси центра і підприємств співпадають.

В реальних задачах величину штрафу чи заохочення або обмежують, тобто $s_i \in G_\varphi$, $i = \overline{1,m}$, де $G_\varphi$ – деяка множина; або значення критерію центра вибирають залежним також від функцій $s_i$, $i = \overline{1,m}$, тобто $\phi = \phi[v^1, ..., v^m, s_1, ..., s_m]$.

Складність оптимізації такого критерію полягає в тому, що необхідно шукати функції $s_i(v^i)$, $i = \overline{1,m}$, які залежать від фазових координат $v^i$, $i = \overline{1,m}$. Нижче розглядаються два підходи для розв’язування таких задач. Перший підхід базується на ідеї параметризації шуканих функцій, а другий – на еквівалентності розв’язків ієрархічної гри двох осіб і спеціальної задачі нелінійного програмування.

Спочатку спростимо ситуацію: вважаємо, що система керування складається із центра і одного підприємства. Центр вибирає елемент $x$, підприємство – $y$, максимізуючи свої критерії:

$$\phi(x, y) \to \max_{(x,y)}, \qquad (1.33)$$

$$\varphi(x, y) \to \max_{(x,y)}. \qquad (1.34)$$

Центр вибирає функцію $x = \psi(y)$ і повідомляє її вигляд підприємству. Припускається, що підприємство доброзичливо відноситься до центру і вибирає $y$ як розв’язок оптимізаційної задачі

$$\varphi\left(\psi(y), y\right) \to \max_{y}. \qquad (1.35)$$

В результаті розв'язування цієї задачі визначається точково-множинний оператор $y = Y[\psi(\cdot)]$. Отже, центру потрібно вибирати функцію $\psi(\cdot)$ як розв'язок задачі

$$\sup_{\psi(\cdot)} \inf_{y \in Y[\psi(\cdot)]} \phi(\psi(y), y). \tag{1.36}$$

Якщо додатково припустити, що для будь-якої функції $\psi(\cdot)$ розв'язок задачі (1.35) єдиний, то замість (1.36) центру потрібно визначати функцію $\psi(\cdot)$, яка реалізує:

$$\sup_{\psi(\cdot)} \phi(\psi(y), y),$$

де $y$ – розв'язок (єдиний) задачі (1.35) при заданій функції $\psi(\cdot)$.

**I.** *Перший підхід* до розв'язування цієї задачі полягає в параметризації функції $\psi(y)$. Задамо, наприклад, цю функцію в квадратичному вигляді:

$$\psi(y) = ay + by^2, \tag{1.37}$$

де $a$, $b$ – деякі параметри.

Використовуючи представлення (1.37), критерій (1.35) можна записати в такій формі

$$\varphi(a,b,y) \to \max_y .$$

Якщо розв'язок останньої оптимізаційної задачі при довільних параметрах $a$ і $b$ єдиний, то $y$ – це деяка функція від параметрів $a$ і $b$

$$y = y(a,b),$$

і задача (1.34) перетворюється в спеціальну задачу математичного програмування.

**II.** Опишемо реалізацію *другого підходу*. Сформулюємо спочатку таку оптимізаційну задачу

$$\varphi(x, y) \to \min_x . \tag{1.38}$$

Розв'язок $x$ цієї задачі залежить від $y : x = x^*(y)$. Тепер сформулюємо наступну оптимізаційну задачу

$$\varphi(x^*(y), y) \to \max_y . \tag{1.39}$$

Оптимальне значення функціоналу задачі (1.39) позначимо через $\varphi^*$, тобто $\varphi^* = \max_y \varphi(x^*(y), y)$.

Нарешті сформулюємо третю оптимізаційну задачу для центра:

$$\phi(x, y) \to \max_{(x,y)} \tag{1.40}$$

при обмеженні

$$\varphi(x, y) \ge \varphi^*. \tag{1.41}$$

Позначимо розв'язок задачі (1.40), (1.41) через $(x^0, y^0)$. Згідно з теоремою Гермейєра оптимальною стратегією центра буде функція $x(y)$:

$$x(y) = \begin{cases} \bar{x}^0, & \text{якшо } y = \bar{y}^0, \\ x^*(y), & \text{якщо } y \neq \bar{y}^0, \end{cases} \tag{1.42}$$

де $\bar{x}^0 = x^0$ та $\bar{y}^0 = y^0$, якщо $\varphi(x^0, y^0) > \varphi^*$; якщо ж $\varphi(x^0, y^0) = \varphi^*$, то $(\bar{x}^0, \bar{y}^0)$ такі, що $\varphi(x^0, y^0) > \varphi^*$ і $(\bar{x}^0, \bar{y}^0)$ з $\varepsilon$-точністю (яка задається центром) реалізують розв'язок задачі (1.40). Отже, якщо оптимізаційні задачі (1.38), (1.39), (1.40) розв'язані, то функція синтезу $x(y)$ виписується явно, згідно (1.42).

Задачі (1.38), (1.39), (1.40) мають простий економічний зміст. Оптимізаційну задачу (1.38) можна трактувати як задачу визначення таких дій центра на підприємство, які ставлять його в найбільш важкі умови. Оскільки ці дії знаходяться в певних рамках і кількість дій невелика, то задача (1.38) часто буває тривіальною. Наприклад, якщо центр визначає ціни, то вони повинні бути мінімальними, якщо штраф – то максимально допустимий і т.д.

Задача (1.39) – це задача вибору підприємством своєї найкращої стратегії в найгірших для нього умовах, а величина $\varphi^*$ – гарантований результат підприємства.

Задача (1.40) – це задача вибору оптимальної стратегії центра в умовах повної централізації з виконанням вимоги (1.41). Розв'язок $\left(x^0, y^0\right)$ інтерпретується як узгоджена програма: підприємству вигідно її притримуватись – воно отримує максимальне заохочення. Відступ від узгодженої програми призводить до погіршення результату.

На рис. 1.2 зображено графік функції (1.42). Функція $x(y)$ є розривною.

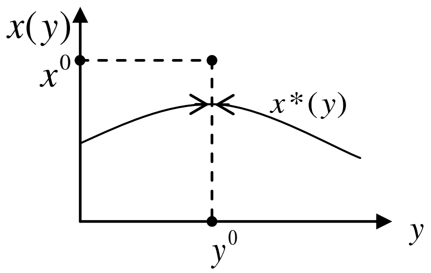



**Рис. 1.2.**

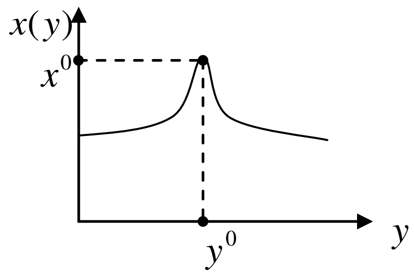



**Рис. 1.3.**

Всюди (за виключенням точки $y^0$) вона співпадає з функцією $x^*(y)$– розв'язком задачі (1.38), який визначає найгірші умови функціонування підприємства – максимальний штраф. На практиці функція може змінюватися в певному діапазоні, тому замість (1.42) користуються згладженою функцією $x(y)$ (рис. 1.3).

Застосуємо обидва підходи до розв'язування початкової задачі управління дворівневою економічною системою. Ввівши позначення

$z_i = L_i^{1/2}$, $i = \overline{1,m}$, максимізацію критерію (1.29) $i$-го підприємства з врахуванням (1.27), можна записати в такому вигляді:

$$\varphi_i(z_i) = \bar{\alpha}_i z_i - \omega_i z_i^2 + s_i(z_i) \to \max \tag{1.43}$$

при умовах $z_i \ge 0$, $s_i(z_i) \ge 0$. Розв'яжемо цю задачу першим способом (використовуючи параметризацію). Задамо функцію $s_i(z_i)$ у квадратичній формі:

$$s_i(z_i) = c_{i1} z_i + c_{i2} z_i^2, \tag{1.44}$$

де $c_{i1}$ і $c_{i2}$ – дійсні параметри. Прирівнявши похідну функції $\varphi_i$ до нуля, знайдемо розв'язок задачі (1.43):

$$z_i = (c_{i1} + \bar{\alpha}_i) / (2(\omega_i - c_{i2})). \tag{1.45}$$

Цільову функцію центра задамо в такому вигляді

$$\phi(z) = \sum_{i=1}^{m} d_i z_i - \sum_{i=1}^{m} s_i(z_i)$$

або з врахуванням рівності (1.44):

$$\phi(z) = \sum_{i=1}^{m} d_i z_i - \sum_{i=1}^{m} \left(c_{i1} z_i + c_{i2} z_i^2\right) = \sum_{i=1}^{m} \phi_i, \tag{1.46}$$

де $\phi_i = (d_i - c_{i1}) z_i - c_{i2} z_i^2$.

Підставимо в останню формулу замість $z_i$ праву частину формули (1.45):

$$\phi_i = \frac{(d_i - c_{i1})(c_{i1} + \bar{\alpha}_i)}{2(\omega_i - c_{i2})} - \frac{c_{i2}(c_{i1} + \bar{\alpha}_i)^2}{4(\omega_i - c_{i2})^2}.$$

Максимізація функції $\phi_i$ по змінних (параметрах) $c_{i1}$, $c_{i2}$ є звичайною задачею математичного програмування.

Тепер опишемо реалізацію другого підходу. Оскільки $s_i(z_i) \ge 0$, то найгіршою дією центра на підприємство буде $s_i(z_i) = 0$, тоді критерій $i$-го підприємства набуде вигляду:

$$\varphi_i(z_i) = \bar{\alpha}_i z_i - \omega_i z_i^2 \to \max.$$

Максимальне значення цієї функції досягається у вершині параболи $z_i = \bar{\alpha}_i / (2\omega_i)$ і дорівнює

$$\varphi_i^* = \bar{\alpha}_i \frac{\bar{\alpha}_i}{2\omega_i} - \omega_i \frac{\bar{\alpha}_i^2}{4\omega_i^2} = \frac{\bar{\alpha}_i^2}{4\omega_i}.$$

Третьою оптимізаційною задачею (1.40), (1.41) при кожному $i = \overline{1,m}$ буде така:

$$\phi_i = d_i z_i - s_i(z_i) \to \max_{z_i}$$

при умові $\varphi_i\left(z_i\right) \geq \bar{\alpha}_i^{\,2} / \left(4\omega_i\right)$.

Звідси отримуємо:

$$s_i = 0, \quad z_i = \bar{\alpha}_i / 2\omega_i .$$

## Запитання та вправи для самостійної роботи до розділу 1.

1. В чому полягає суть критеріального підходу до визначення альтернатив?
2. Назвіть найбільш часто вживані способи переходу до нормалізованого мультикритерію.
3. Що таке суперкритерій?
4. Наведіть приклади найуживаніших згорток.
5. Які ви знаєте недоліки, котрі пов'язані з використанням суперкритерію?
6. В чому полягає суть методу головного частинного критерію?
7. Запишіть покроково алгоритм методу послідовних поступок.
8. Наведіть приклади числових мір близькості між альтернативою $x$ та ідеальною точкою (ціллю) $x^*$.
9. Напишіть формулу переходу до нових нормованих критеріїв в методі бажаної точки.
10. Опишіть найчастіше вживані при прийнятті рішень в умовах невизначеності оцінки ефективності альтернатив: гарантовані $\Phi(x)$ та середні $S(x)$.
11. Сформулюйте математичну модель дворівневої ієрархічної системи керування.
12. Дайте означення системи з ієрархічною структурою.
13. Яке керування $u \in U$ центра називають допустимим?
14. З якої умови вибирає центр своє оптимальне керування в тривіальному випадку?
15. Запишіть співвідношення, які визначають принцип гарантованого результату для керування центра.
16. Запишіть співвідношення, згідно яких вибираються керування центром і підсистемами при прийнятті рішень в умовах доброзичливості.

**Р о з д і л 2**

# МЕТОДИ ОДНОВИМІРНОЇ ОПТИМІЗАЦІЇ

## 2.1. Методи Фібоначі

З а д а ч а 0. Знайти $\arg \min\limits_{x \in [a_0; b_0]} f_0(x)$ для заданої функції $f_0 : R^1 \to R^1$ і заданого відрізка $[a_0; b_0]$.

*Припущення 0.* Функція $f_0$ така, що на відрізку $[a_0; b_0]$ точка її локального мінімуму $x^*$ є точкою абсолютного мінімуму $f_0$ на відрізку $[a_0; b_0]$ ( $f_0 -$ унімодальна функція).

Методи Фібоначі є оптимальними для класу функцій $f_0$, які задовольняють припущенню 0 за кількістю обчислень мінімізуючої функції $f_0$ при заданій точності обчислень $x^*$.

### 1. Основний метод



***Алгоритм 1***

П о ч а т о к. I. Вибрати число $\varepsilon > 0$ – точність обчислення точки мінімуму функції $f_0$ на відрізку $[a_0; b_0]$; покласти $F_1 = F_2 = 1$.

II. Покласти $j = 1$.

III. Обчислити число $F_{j+2} = F_{j+1} + F_j$.

IV. Якщо $F_{j+1} < \frac{1}{\varepsilon}(b_0 - a_0) \le F_{j+2}$, то покласти $m = j$ і перейти на крок V; інакше покласти $j = j + 1$ і перейти на крок III.

V. Обчислити точки:

$$y_1 = a_0 + \frac{F_m}{F_{m+2}}(b_0 - a_0); \quad z_1 = a_0 + b_0 - y_1.$$

VI. Якщо $f_0(y_1) \le f_0(z_1)$, то покласти $a_1 = a_0$, $b_1 = z_1$ і перейти на крок VII; інакше покласти $a_1 = y_1$, $b_1 = b_0$ і перейти на крок VII.

VII. Покласти $k = 1$.

О с н о в н и й ц и к л. VIII. Якщо $f_0(y_k) \le f_0(z_k)$, то обчислити точку $y_{k+1} = a_k + b_k - y_k$, обчислити значення $f_0(y_{k+1})$ і перейти на крок IX; інакше покласти $y_{k+1} = z_k$, $f_0(y_{k+1}) = f_0(z_k)$ і перейти на крок X.

IX. Покласти $z_{k+1} = y_k$; $f_0(z_{k+1}) = f_0(y_k)$ і перейти на крок XI.

X. Обчислити точку $z_{k+1} = a_k + b_k - z_k$, обчислити значення $f_0(z_{k+1})$ і перейти на крок XI.

XI. Якщо $f_0(y_{k+1}) \le f_0(z_{k+1})$, то покласти $a_{k+1} = a_k$, $b_{k+1} = z_{k+1}$ і перейти на крок XII; інакше покласти $a_{k+1} = y_{k+1}$, $b_{k+1} = b_k$ і перейти на крок XII.

XII. Якщо $k < m-1$, то покласти $k = k+1$ і перейти на крок VIII; інакше покласти $\bar{x}^* = (a_m + b_m)/2$ і завершити обчислення.

***Теорема 1.*** *Якщо виконується припущення 0, то для будь-якого $\varepsilon > 0$ точка $\bar{x}^*$, яка породжена алгоритмом 1, задовольняє нерівність $|x^* - \bar{x}^*| \le \varepsilon$.*

*Зауваження 1.* Недоліком алгоритму 1 є те, що похибка в обчисленнях точок $y_k$, $z_k$ можуть настільки швидко накопичуватися, що очікувана точність розв'язку буде суттєво відрізнятися від реальної. Наступна модифікація методу Фібоначі менш чутлива до похибок обчислень.

**2. Модифікація методу Фібоначі**

***Алгоритм 2***

Початок. Кроки I – IV такі ж, як у алгоритмі 1.

V. Обчислити точки:

$$y_1 = a_0 + \frac{F_m}{F_{m+2}}(b_0 - a_0); \quad z_1 = a_0 + \frac{F_{m+1}}{F_{m+2}}(b_0 - a_0).$$

VI. Якщо $f_0(y_1) \le f_0(z_1)$, то покласти $a_1 = a_0$, $b_1 = z_1$ і перейти на крок VII; інакше покласти $a_1 = y_1$, $b_1 = b_0$ і перейти на крок VII.

VII. Покласти $k = 1$.

Основний цикл. VIII. Якщо $f_0(y_k) \le f_0(z_k)$, то обчислити точку

$$y_{k+1} = a_k + \frac{F_{m-k}}{F_{m+2}}(b_0 - a_0),$$

обчислити значення $f_0(y_{k+1})$ і перейти на крок IX; інакше покласти $y_{k+1} = z_k$, $f_0(y_{k+1}) = f_0(z_k)$ і перейти на крок X.

IX. Покласти $z_{k+1} = y_k$; $f_0(z_{k+1}) = f_0(y_k)$ і перейти на крок XI.

X. Обчислити точку $z_{k+1} = a_k + \dfrac{F_{m-k+1}}{F_{m+2}}(b_0 - a_0)$, обчислити значення $f_0(z_{k+1})$ і перейти на крок XI.

XI. Якщо $f_0(y_{k+1}) \le f_0(z_{k+1})$, то покласти $a_{k+1} = a_k$, $b_{k+1} = z_{k+1}$ і перейти на крок XII; інакше покласти $a_{k+1} = y_{k+1}$, $b_{k+1} = b_k$ і перейти на крок XII.

XII. Якщо $k < m-1$, то покласти $k = k+1$ і перейти на крок VIII; інакше покласти $\bar{x}^* = (a_m + b_m)/2$ і завершити обчислення.

Точка $\overline{x}^*$ задовольняє нерівність $|x^* - \overline{x}^*| \le \varepsilon$, якщо виконується припущення 0.

**Приклад 1.** Знайти мінімум функції $f_0(x) = 3x^3 - 5x^2 + x + 2$ на відрізку [0,5;3] з точністю $\varepsilon = 0,1$.

*Розв'язування.*

**Алгоритм 2**

I. Вибираємо $\varepsilon = 0,1$; покладемо $F_1 = F_2 = 1$ .

II–IV. Обчислимо числа Фібоначі:

$$F_3 = 2;\ F_4 = 3;\ F_5 = 5;\ F_6 = 8;\ F_7 = 13;\ F_8 = 21;\ F_9 = 34.$$

При $j = 7$ виконується умова $F_8 < (1/0,1)\cdot(3-0,5) \le F_9$, бо $F_8 = 21$; $F_9 = 34$ і справджуються нерівності 2,1<2,5<3,4.

Покладемо $m = 7,\ F_m = F_7 = 13$ і перейдемо на крок V .

V. Обчислюємо точки:

$$y_1 = 0,5 + (13/34)\cdot 2,5 = 1,45588;$$

$$z_1 = 0,5 + (21/34)\cdot 2,5 = 2,04412.$$

VI. Обчислюємо $f_0(y_1) = 2,11554;\ f_0(z_1) = 8,77560$ .

Оскільки $f_0(y_1) < f_0(z_1)$, то покладемо $a_1 = a_0, b_1 = z_1$ і переходимо на крок VII.

VII. Покладемо $k = 1$.

*1-а ітерація*:

VIII. Обчислюємо точку

$$y_2 = a_1 + \frac{F_{7-1}}{F_9}\cdot 2,5 = 0,5 + (8/34)\cdot 2,5 = 1,08824,$$

обчислюємо $f_0(y_2) = 1,03321$ і переходимо на крок IX.

IX. Покладемо $z_2 = y_1 = 1,45588,\ \ f_0(z_2) = 2,11554$ і перейдемо на крок XI.

XI. Оскільки $f_0(y_2) < f_0(z_2)$, то покладемо $a_2 = a_1 = 0,5$; $b_2 = z_2 = 1,45588$ і переходимо на крок XII.

XII. Оскільки $k = 1 < 6$, то покладемо $k = 1 + 1 = 2$ і переходимо на крок VIII.

*2-а ітерація:*

VIII. Обчислюємо точку

$$y_3 = a_2 + (F_5/F_9)\cdot 2,5 = 0,5 + (5/34)\cdot 2,5 = 0,86765,$$

обчислюємо $f_0(y_3) = 1,06311$ і переходимо на крок IX.

IX. Покладемо $z_3 = y_2 = 1,08824,\ \ f_0(z_3) = 1,03321$ і перейдемо на крок XI.

XI. Оскільки $f_0(y_3) > f_0(z_3)$, то покладемо $a_3 = y_3 = 0,86765$; $b_3 = b_2 = 1,45588$ і переходимо на крок XII .

XII. Оскільки $k = 2 < 6$, то покладемо $k = 2 + 1 = 3$ і переходимо на крок VIII.

*3-я ітерація:*

VIII. Оскільки $f_0(y_3) > f_0(z_3)$, то покладемо $y_4 = z_3 = 1,08824$; $f_0(y_4) = 1,03321$ і переходимо на крок X.

X. Обчислюємо точку

$$z_4 = a_3 + \left(F_{7-3+1}/F_9\right)\cdot 2,5 = 0,86765 + \left(5/34\right)\cdot 2,5 = 1,23530,$$

обчислюємо $f_0(z_4) = 1,26055$ і переходимо на крок XI.

XI. Оскільки $f_0(y_4) < f_0(z_4)$, то покладемо $a_4 = a_3 = 0,86765$; $b_4 = z_4 = 1,23530$ і переходимо на крок XII.

XII. Оскільки $k = 3 < 6$, то покладемо $k = 3 + 1 = 4$ і переходимо на крок VIII.

*4-а ітерація:*

VIII. Оскільки $f_0(y_4) < f_0(z_4)$, то обчислюємо точку

$$y_5 = a_4 + (F_{7-4}/F_9)\cdot 2,5 = 0,86765 + (2/34)\cdot 2,5 = 1,01471,$$

обчислюємо $f_0(y_5) = 1,00088$ і переходимо на крок IX.

IX. Покладемо $z_5 = y_4 = 1,08824$, $f_0(z_5) = 1,03321$ і перейдемо на крок XI.

XI. Оскільки $f_0(y_5) < f_0(z_5)$, то покладемо $a_5 = a_4 = 0,86765$; $b_5 = z_5 = 1,08824$ і переходимо на крок XII.

XII. Оскільки $k = 4 < 6$, то покладемо $k = 4 + 1 = 5$ і переходимо на крок VIII.

*5-а ітерація:*

VIII. Оскільки $f_0(y_5) < f_0(z_5)$, то обчислюємо точку

$$y_6 = a_5 + (F_{7-5}/F_9)\cdot 2,5 = 0,86765 + (1/34)\cdot 2,5 = 0,94118,$$

обчислюємо $f_0(y_6) = 1,01323$ і переходимо на крок IX.

IX. Покладемо $z_6 = y_5 = 1,01471$, $f_0(z_6) = 1,00088$ і перейдемо на крок XI.

XI. Оскільки $f_0(y_6) > f_0(z_6)$, то покладемо $a_6 = y_6 = 0,94118$; $b_6 = b_5 = 1,08824$ і переходимо на крок XII.

XII. Оскільки $k = 5 < 6$, то покладемо $k = 5 + 1 = 6$ і переходимо на крок VIII.

*6-а ітерація:*

VIII. Оскільки $f_0(y_6) > f_0(z_6)$, то покладемо $y_7 = z_6 = 1,01471$; $f_0(y_7) = 1,00088$ і переходимо на крок X.

X. Обчислюємо точку

$$z_7 = a_6 + \frac{F_{7-6+1}}{F_9} \cdot 2,5 = 0,94118 + (1/34) \cdot 2,5 = 1,01471,$$

обчислюємо $f_0(z_7) = 1,00088$ і переходимо на крок XI.

XI. Оскільки $f_0(y_7) = f_0(z_7)$, то покладемо $a_7 = a_6 = 0,94118$; $b_7 = z_7 = 1,01471$ і переходимо на крок XII.

XII. Оскільки $k = 6$, то покладемо

$$\overline{x}^* = (a_7 + b_7)/2 = (0,94118 + 1,01471)/2 = 0,978$$

і зупиняємо обчислення.

Наприкінці відмітимо, що мінімум функції $f_0$ на відрізку [0,5;3] досягається в точці $x^* = 1$. Отже, модифікований метод Фібоначі за шість ітерацій дав наближений розв'язок $\overline{x}^* = 0,978$ з похибкою $\Delta_\phi = |1 - 0,978| = 0,022 < \varepsilon$, причому зроблено вісім обчислень функції $f_0$ в точках $y_1, z_1, y_2, y_3, z_4, y_5, y_6, z_7$.

## 2.2. Метод золотого перерізу

З а д а ч а 1. Знайти $\arg\min\limits_{x \in [a_0;b_0]} f_0(x)$ для заданої функції $f_0 : R^1 \to R^1$ і заданого відрізка $[a_0;b_0]$.

*Припущення 1.* Функція $f_0$ така, що на відрізку $[a_0;b_0]$ точка її локального мінімуму $x^*$ є точкою абсолютного мінімуму на відрізку $[a_0;b_0]$.

**Алгоритм 1**

П о ч а т о к . І. Обчислити константу $\alpha = (3 - \sqrt{5})/2$ $(\alpha \cong 0,381966)$.

ІІ. Обчислити точки:

$$y_1 = a_0 + \alpha(b_0 - a_0); \quad z_1 = a_0 + b_0 - y_1$$

і значення $f_0(y_1)$, $f_0(z_1)$.

ІІІ. Якщо $f_0(y_1) \le f_0(z_1)$, то покласти $a_1 = a_0$, $b_1 = z_1$ і перейти на крок IV, інакше покласти $a_1 = y_1$, $b_1 = b_0$ і перейти на крок IV.

IV. Покласти $k = 1$.

О с н о в н и й ц и к л . V. Якщо $f_0(y_k) \le f_0(z_k)$, то обчислити $y_{k+1} = a_k + b_k - y_k$, $f_0(y_{k+1})$ і перейти на крок VI; інакше покласти $y_{k+1} = z_k$, $f_0(y_{k+1}) = f_0(z_k)$ і перейти на крок VII.

VI. Покласти:

$$z_{k+1} = y_k, \quad f_0(z_{k+1}) = f_0(y_k)$$

і перейти на крок VIII.

VII. Обчислити $z_{k+1} = a_k + b_k - z_k$, $f_0(z_{k+1})$ і перейти на крок VIII.

VIII. Якщо $f_0(y_{k+1}) \le f_0(z_{k+1})$, то покласти $a_{k+1} = a_k$, $b_{k+1} = z_{k+1}$ і перейти на крок IX; інакше покласти $a_{k+1} = y_{k+1}$, $b_{k+1} = b_k$ і перейти на крок IX.

IX. Обчислити $x^k = (a_{k+1} + b_{k+1})/2$.

X. Покласти $k = k+1$ і перейти на крок V.

***Теорема 1.*** *Якщо виконується припущення 1, то послідовність* $\{x^k\}_{k=1}^{\infty}$, *яка породжена алгоритмом 1, така, що:*

$$\lim_{k\to\infty} f_0(x^k) = f_0(x^*); \quad \lim_{k\to\infty} \left| x^k - x^* \right| = 0.$$

*Зауваження 1.* Довжина відрізку $[a_k;b_k]_{з.п.}$, побудованого по методу золотого перерізу, на 17% більша довжини відрізку $[a_k;b_k]_{Фіб.}$, побудованого по методу Фібоначі. Проте метод золотого перерізу має наступну перевагу, що на кожній його ітерації доводиться робити менше обчислень.

*Зауваження 1'.* Іноді на практиці комбінують обидва методи: перші кроки роблять по методу золотого перерізу, а коли оптимум достатньо близький, обраховують число *m* і переходять до методу Фібоначі.

Відмітимо також наступне. При $n \to \infty$ значення $1/F_n$ прямує до $0{,}618$, тобто методи Фібоначі та золотого перерізу асимптотично еквівалентні.

**Приклад 1.** Знайти мінімум функції $f_0(x) = 3x^3 - 5x^2 + x + 2$ на відрізку [0,5;3], виконавши шість ітерацій методу золотого перерізу.

*Розв'язування.*

**Алгоритм 1**

I. Обчислимо константу $\alpha = (3 - \sqrt{5})2 \approx 0{,}381966$.

II. Обчислюємо точки:

$$y_1 = 0{,}5 + 0{,}381966 \cdot 2{,}5 = 1{,}454915,$$

$$z_1 = 0{,}5 + 3 - 1{,}454915 = 2{,}045085$$

і значення функції $f_0$ в цих точках:

$$f_0(y_1) = 2{,}110221; f_0(z_1) = 8{,}793144.$$

III. Оскільки $f_0(y_1) < f_0(z_1)$, то покладемо $a_1 = a_0 = 0{,}5$, $b_1 = z_1 = 2{,}045085$ і переходимо на крок IV.

IV. Покладемо $k = 1$.

*1-а ітерація:*

V. Оскільки $f_0(y_1) < f_0(z_1)$, то обчислюємо точку

$$y_2 = a_1 + b_1 - y_1 = 0,5 + 2,045085 - 1,454915 = 1,09017$$

і значення функції $f_0(y_2) = 1,034722$ і переходимо на крок VI.

VI. Покладемо $z_2 = y_1 = 1,454915$, $f_0(z_2) = 2,110221$ і переходимо на крок VIII.

VIII. Оскільки $f_0(y_2) < f_0(z_2)$, то покладемо $a_2 = a_1 = 0,5$, $b_2 = z_2 = 1,454915$ і переходимо на крок IX.

IX. Обчислюємо

$$x^1 = (a_2 + b_2)/2 = (0,5 + 1,454915)/2 = 0,977.$$

X. Покладемо $k = 1 + 1 = 2$ і переходимо на крок V.

*2-а ітерація:*

V. Оскільки $f_0(y_2) < f_0(z_2)$, то обчислюємо точку

$$y_3 = a_2 + b_2 - y_2 = 0,5 + 1,444915 - 1,09017 = 0,854745$$

і значення функції $f_0(y_3) = 1,075202$ і переходимо на крок VI.

VI. Покладемо $z_3 = y_2 = 1,09017$, $f_0(z_3) = 1,034722$ і переходимо на крок VIII.

VIII. Оскільки $f_0(y_3) > f_0(z_3)$, то покладемо $a_3 = y_3 = 0,854745$, $b_3 = b_2 = 1,454915$ і переходимо на крок IX.

IX. Обчислюємо

$$x^2 = (a_3 + b_3)/2 = (0,854745 + 1,454915)/2 = 1,15483.$$

X. Покладемо $k = 2 + 1 = 3$ і переходимо на крок V.

*3-я ітерація:*

V. Оскільки $f_0(y_3) > f_0(z_3)$, то покладемо $y_4 = z_3 = 1,09017$, $f_0(y_4) = 1,034722$ і переходимо на крок VII.

VII. Обчислюємо

$$z_4 = a_3 + b_3 - z_3 = 0,854745 + 1,454915 - 1,09017 = 1,21949$$

і значення функції $f_0(z_4) = 1,224426$ і переходимо на крок VIII.

VIII. Оскільки $f_0(y_4) < f_0(z_4)$, то покладемо $a_4 = a_3 = 0,854745$, $b_4 = z_4 = 1,21949$ і переходимо на крок IX.

IX. Обчислюємо

$$x^3 = (a_4 + b_4)/2 = (0,854745 + 1,21949)/2 = 1,03712.$$

X. Покладемо $k = 3 + 1 = 4$ і переходимо на крок V.

*4-а ітерація:*

V. Оскільки $f_0(y_4) < f_0(z_4)$, то обчислюємо точку

$$y_5 = a_4 + b_4 - y_4 = 0,854745 + 1,21949 - 1,09017 = 0,984065$$

і значення функції $f_0(y_5) = 1,001004$ і переходимо на крок VI.

VI. Покладемо $z_5 = y_4 = 1,09017$, $f_0(z_5) = 1,034722$ і переходимо на крок VIII.

VIII. Оскільки $f_0(y_5) < f_0(z_5)$, то покладемо $a_5 = a_4 = 0,854745$, $b_5 = z_5 = 1,09017$ і переходимо на крок IX.

IX. Обчислюємо

$$x^4 = (a_5 + b_5)/2 = (0,854745 + 1,09017)/2 = 0,97246.$$

X. Покладемо $k = 4 + 1 = 5$ і переходимо на крок V.

*5-а ітерація:*

V. Оскільки $f_0(y_5) < f_0(z_5)$, то обчислюємо точку

$$y_6 = a_5 + b_5 - y_5 = 0,854745 + 1,09017 - 0,984065 = 0,96085$$

і значення функції $f_0(y_6) = 1,00595$ і переходимо на крок VI.

VI. Покладемо $z_6 = y_5 = 0,984065$, $f_0(z_6) = 1,001004$ і переходимо на крок VIII.

VIII. Оскільки $f_0(y_6) > f_0(z_6)$, то покладемо $a_6 = y_6 = 0,96085$, $b_6 = b_5 = 1,09017$ і переходимо на крок IX.

IX. Обчислюємо

$$x^5 = (a_6 + b_6)/2 = (0,96085 + 1,09017)/2 = 1,02551.$$

X. Покладемо $k = 5 + 1 = 6$ і переходимо на крок V.

*6-а ітерація:*

V. Оскільки $f_0(y_6) > f_0(z_6)$, то покладемо $y_7 = z_6 = 0,984065$, $f_0(y_7) = 1,001004$ і переходимо на крок VII.

VII. Обчислюємо

$$z_7 = a_6 + b_6 - z_6 = 0,96085 + 1,09017 - 0,984065 = 1,066955$$

і обчислюємо значення функції $f_0(z_7) = 1,018832$ і переходимо на наступний крок.

VIII. Оскільки $f_0(y_7) < f_0(z_7)$, то покладемо $a_7 = a_6 = 0,96085$, $b_7 = z_7 = 1,066955$ і переходимо на крок IX.

IX. Обчислюємо

$$x^6 = (a_7 + b_7)/2 = (0,96085 + 1,066955)/2 = 1,0139.$$

X. Зупиняємо обчислення.

За шість ітерацій методу золотого перерізу отримали наближений розв'язок $x^* = 1,014$, похибка якого $\Delta_{з.п.} = |1 - 1,014| = 0,014$.

Довжина відрізка $[a_7;b_7]$, який отриманий методом Фібоначі дорівнює $\delta_{Фіб.} = b_7 - a_7 = 1,01471 - 0,94118 = 0,07353$, а методом золотого перерізу $\delta_{з.п.} = 1,066955 - 0,96085 = 0,10611$, тобто є на 44% більший від відрізка методу Фібоначі.

## 2.3. Методи типу Ньютона

Задача 0. Знайти $\arg\min_{x\in[a_0;b_0]} f_0(x)$ для заданої функції $f_0 : R^1 \to R^1$ і заданого відрізка $[a_0;b_0]$.

*Припущення 0.* Функція $f_0$ двічі неперервно диференційована на $[a_0;b_0]$ і задовольняє умову

$$\left|f_0''(x) - f_0''(y)\right| \le \alpha|x-y|, \alpha < \infty.$$

### 1. Метод Ньютона

Ідея методу Ньютона полягає в тому, що функція $f_0'(x)$ лінеаризується в околі точки $x^k$ і знаходять точку $x^{k+1}$, в якій лінеаризована функція перетворюється в нуль.

***Алгоритм 1***

Початок. I. Вибрати початкове наближення $x^0 \in [a_0;b_0]$; покласти $k = 0$.

Основний цикл. II. Обчислити $f_0'(x^k)$ – першу похідну функції $f_0$ в точці $x^k$.

III. Якщо $f_0'(x^k) = 0$, то припинити обчислення (в цьому випадку точка $x^k$ є стаціонарною точкою функції $f_0$), інакше перейти на крок IV.

IV. Обчислити $f_0''(x^k)$ – другу похідну функції $f_0$ в точці $x^k$.

V. Обчислити наступне наближення

$$x^{k+1} = x^k - f_0'(x^k) \big/ f_0''(x^k).$$

VI. Покласти $k = k+1$ і перейти на крок II.

***Теорема 1.*** *Якщо виконується припущення 0 і, крім того,* $(\boldsymbol{i})-$ $f_0'(a_0) f_0'(b_0) < 0$; $(\boldsymbol{ii})-$ *для всіх* $x \in [a_0;b_0]$ $-$ $f_0''(x) > 0$; $(\boldsymbol{iii})-$ *для всіх* $x \in [a_0;b_0]$ *виконуються нерівності:*

$$0 \le 1-\left[ f_0'(x) \big/ f_0''(x) \right]' \le \gamma < 1,$$

*то послідовність* $\{x^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 1, збігається до стаціонарної точки* $\bar{x}$ *функції* $f_0$ *з квадратичною швидкістю, тобто*

$$\left|x^{k+1}-\bar{x}\right| \le \beta \left|x^k-\bar{x}\right|^2.$$

*Тут*

$$\beta = \sup_{k=0,1,\dots} \left| f_0'''\left(\theta^k\right) \cdot f_0''\left(\eta^k\right) \right| \Big/ \left[ f_0''\left(\theta^k\right) \right]^2,$$

*де* $\theta^k \in [x^k;\bar{x}]$; $\eta^k \in [\theta^k;\bar{x}]$.

**Приклад 1.** Методом Ньютона обчислити наближене значення мінімуму функції

$$f_0 = x^4 - 6x^2 + 10$$

на відрізку [1;3], виконавши чотири ітерації.

*Розв'язування*. Спочатку знайдемо першу та другу похідні функції $f_0$:

$$f_0'(x) = 4x^3 - 12x;\ \ f_0''(x) = 12x^2 - 12.$$



**Алгоритм 1**

I. Вибираємо початкове наближення $x^0 = 2 \in [1;3]$, покладемо $k=0$.

*1-а ітерація:*

II. Обчислюємо першу похідну

$$f_0'(x^0) = 4\cdot 2^3 - 12\cdot 2 = 32 - 24 = 8.$$

III. Оскільки $f_0'(2) \ne 0$, то переходимо на крок IV.

IV. Обчислюємо другу похідну

$$f_0''(x^0) = 12\cdot 2^2 - 12 = 48 - 12 = 36.$$

V. Обчислюємо наступне наближення

$$x^1 = x^0 - f_0'(x^0) / f_0''(x^0) = 2 - 8/36 = 1{,}777778.$$

VI. Покладемо $k = 0+1 = 1$ і переходимо на крок II.

*2-а ітерація:*

II. Обчислюємо

$$f_0'(x^1) = 4\cdot 1{,}777778^3 - 12\cdot 1{,}777778 = 1{,}141295.$$

III. Оскільки $f_0'(x^1) \ne 0$, то переходимо на крок IV.

IV. Обчислюємо

$$f_0''(x^1) = 12\cdot 1{,}777778^2 - 12 = 25{,}925935.$$

V. Обчислюємо наступне наближення

$$x^2 = x^1 - f_0'(x^1) / f_0''(x^1) =$$
$$= 1,777778 - 1,141295 / 25,925935 = 1,733757.$$

VI. Покладемо $k = 1 + 1 = 2$ і переходимо на крок II.

*3-я ітерація:*

II. Обчислюємо

$$f_0'(x^2) = 4 \cdot 1,733757^3 - 12 \cdot 1,733757 = 0,041009 .$$

III. Оскільки $f_0'(x^2) \neq 0$, то переходимо на крок IV.

IV. Обчислюємо

$$f_0''(x^2) = 12 \cdot 1,733757^2 - 12 = 24,070960 .$$

V. Обчислюємо наступне наближення

$$x^3 = x^2 - f_0'(x^2) / f_0''(x^2) =$$
$$= 1,733757 - 0,041009 / 24,070960 = 1,732053.$$

VI. Покладемо $k = 2 + 1 = 3$ і переходимо на крок II.

*4-а ітерація:*

II. Обчислюємо

$$f_0'(x^3) = 4 \cdot 1,732053^3 - 12 \cdot 1,732053 = 0,000053 .$$

III. Оскільки $f_0'(x^3) \neq 0$, то переходимо на крок IV.

IV. Обчислюємо

$$f_0''(x^3) = 12 \cdot 1,732053^2 - 12 = 24,000091.$$

V. Обчислюємо наступне наближення

$$x^4 = x^3 - f_0'(x^3) / f_0''(x^3) =$$
$$= 1,732053 - 0,000053 / 24,000091 = 1,73205079.$$

VI. Покладемо $k = 3 + 1 = 4$ і зупиняємо обчислення.

Відмітимо, що мінімум функції $f_0$ на відрізку [1;3] досягається в точці $x^* = \sqrt{3} \approx 1,732050808.$

Отже, методом Ньютона за чотири ітерації ми отримали наближений розв'язок $x^4 = 1,73205079$ з похибкою $\Delta = 10^{-8}$.

**2. Метод січних**

Метод січних є модифікацією методу Ньютона (алгоритму 1), в якому замість другої похідної $f''\left(x^k\right)$ використовується її різницева апроксимація.

***Алгоритм 2***

П о ч а т о к . I. Вибрати довільне початкове наближення $x^0 \in [a_0; b_0]$; покласти $k = 0$.

О с н о в н и й  ц и к л . II. Обчислити $f_0'(x_k)$ – першу похідну функції $f_0$ в точці $x^k$.

III. Якщо $f_0'(x^k) = 0$, то зупинити обчислення (в цьому випадку точка $x^k$ є стаціонарною точкою функції $f_0$), інакше перейти на крок IV.

IV. Обчислити наступне наближення

$$x^{k+1} = x^k - f_0'(x^k)(x^k - x^{k-1}) \Big/ \left( f_0'(x^k) - f_0'(x^{k-1}) \right).$$

V. Покласти $k = k+1$ і перейти на крок II.

***Теорема 2.*** *Якщо виконуються всі умови теореми 1, то послідовність* $\{x^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 2, збігається до стаціонарної точки* $\bar{x}$ *функції* $f_0$ *з надлінійною швидкістю:*

$$\lim_{k\to\infty} \frac{|x^{k+1} - \bar{x}|}{|x^k - \bar{x}|^\tau} = \left| \frac{2 f_0''(\bar{x})}{f_0'''(\bar{x})} \right|^{1/\tau},$$

*де* $\tau$ *– розв'язок рівняння* $t^2 = t + 1$ $\left(\tau = (1+\sqrt{5})/2 \cong 1{,}618\right)$.



## 2.4. Методи дотичних

З а д а ч а  0. Знайти $\arg \min\limits_{x\in[a_0;b_0]} f_0(x)$ для заданої функції $f_0 : R^1 \to R^1$ і заданого відрізка $[a_0; b_0]$.

#### 1. Випадок диференційованої функції

*Припущення 1.* (***i***) – функція $f_0$ неперервно диференційована на відрізку $[a_0; b_0]$; (***ii***) – функція $f_0$ опукла донизу на відрізку $[a_0; b_0]$; (***iii***) – $f_0'(a_0) < 0$ і $f_0'(b_0) > 0$.

***Алгоритм 1***

П о ч а т о к . I. Вибрати число $\varepsilon > 0$ – точність обчислення точки мінімуму; покласти $k = 0$.

II. Обчислити похідні $f_0'(a_0)$ і $f_0'(b_0)$ функції $f_0$ в точках $a_0$ і $b_0$, відповідно.

О с н о в н и й  ц и к л . III. Знайти точку $x^k$ – корінь рівняння

$$f_0(a_k) + f_0'(a_k)(x - a_k) = f_0(b_k) + f_0'(b_k)(x - b_k).$$

IV. Обчислити $f_0'(x^k)$.

V. Якщо $f_0'(x^k)=0$, то покласти $x^*=x^k$ і припинити обчислення; інакше перейти на крок VI.

VI. Якщо $f_0'(x^k)<0$, то покласти $a_{k+1}=x^k$, $b_{k+1}=b_k$ і перейти на крок VIII; інакше перейти на крок VII.

VII. Якщо $f_0'(x^k)>0$, то покласти $a_{k+1}=a_k$, $b_{k+1}=x^k$ і перейти на крок VIII.

VIII. Якщо $b_{k+1}-a_{k+1}\le\varepsilon$, то покласти $x^*=\left(a_{k+1}+b_{k+1}\right)/2$ і припинити обчислення; інакше перейти на крок IX.

IX. Покласти $k=k+1$ і перейти на крок III.

***Теорема 1.*** *Якщо виконується припущення 1, то для послідовності* $\left\{x^k\right\}_{k=0}^{\infty}$, *яка породжена алгоритмом 1, справедливо*

$$\lim_{\substack{k\to\infty\\ \varepsilon\to 0}} f_0(x^k)=\min_{x\in[a_0;b_0]} f_0(x).$$

*Якщо, крім того, точка мінімуму* $x^*$ *єдина, то* $\lim\limits_{\substack{k\to\infty\\ \varepsilon\to 0}} x^k=x^*$.



**2. Випадок недиференційованої функції**

*Припущення 2.* (***i***)– функція $f_0$ опукла донизу на відрізку $\left[a_0;b_0\right]$; (***ii***)– $f_0'(a_0+0)<0$ і $f_0'(b_0-0)<0$.

***Алгоритм 2***

Початок. I. Задати число $\varepsilon>0$ – точність обчислення точки мінімуму функції $f_0$; покласти $k=0$.

II. Обчислити правосторонню $f_0'(a_0+0)$ та лівосторонню $f_0'(b_0-0)$ похідні функції $f_0$ в точках $a_0$ і $b_0$, відповідно, та покласти $\gamma_0=f_0'(a_0+0)$, $\beta_0=f_0'(b_0-0)$.

Основний цикл. III. Знайти точку $x^k$ – корінь рівняння

$$f_0(a_k)+\gamma_k(x-a_k)=f_0(b_k)+\beta_k(x-b_k).$$

IV. Обчислити правосторонню $f_0'(x^k+0)$ і та лівосторонню $f_0'(x^k-0)$ похідні функції $f_0$ в точці $x^k$.

V. Якщо $f_0'(x^k-0)f_0'(x^k+0)\le 0$, то покласти $x^*=x^k$ і припинити обчислення; інакше вибрати довільне число $\delta_k$ з відрізка $\left[f_0'(x^k-0);f_0'(x^k+0)\right]$ та перейти на крок VI.

VI. Якщо $\delta_k < 0$, то покласти $a_{k+1} = x^k$, $\gamma_{k+1} = \delta_k$, $b_{k+1} = b_k$, $\beta_{k+1} = \beta_k$ і перейти на крок VII; інакше покласти $a_{k+1} = a_k$, $\gamma_{k+1} = \gamma_k$, $b_{k+1} = x^k$, $\beta_{k+1} = \delta_k$ і перейти на крок VII.

VII. Якщо $b_{k+1} - a_{k+1} \le \varepsilon$, то покласти $x^* = \left(a_{k+1} + b_{k+1}\right)/2$ і припинити обчислення; інакше перейти на крок VIII.

VIII. Покласти $k = k+1$ і перейти на крок III.

***Теорема 2.*** *Якщо виконується припущення 2, то для послідовності* $\{x^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 2, справедливе твердження теореми 1.*

## 2.5. Методи глобального пошуку

Задача 0. Знайти $\arg \min\limits_{x\in[a_0;b_0]} f_0(x)$ для заданої функції $f_0 : R^1 \to R^1$ та заданого відрізку $[a_0;b_0]$ дійсної осі $R^1$.

Приведені нижче алгоритми застосовуються для обчислення абсолютного мінімуму багатоекстремальної функції $f_0$ на відрізку $[a_0;b_0]$.



**1. Метод глобального пошуку**

***Алгоритм 1***

Початок. I. Вибрати довільну константу $\alpha > 1$.

II. Покласти $x^0 = a_0$, $x^1 = b_0$.

III. Покласти $k = 1$.

Основний цикл. IV. Розташувати точки послідовності $\{x^i\}_{i=0}^{k}$ в порядку зростання їх значень і нову послідовність позначити через $\left\{\bar{x}^i\right\}_{i=0}^{k}$, тобто

$$a_0 = \bar{x}^0 < \bar{x}^1 < \ldots < \bar{x}^k = b_0 .$$

V. Знайти максимальне абсолютне значення відносної першої різниці

$$\delta_k = \max_{1\le i\le k} |(f_0(\bar{x}^i) - f_0(\bar{x}^{i-1}))/(\bar{x}^i - \bar{x}^{i-1})| .$$

VI. Якщо $\delta_k = 0$, то покласти $\beta_k = 1$ і перейти на крок VII; інакше покласти $\beta_k = \alpha\delta_k$ і перейти на крок VII.

VII. Покласти $i = 1$.

VIII. Обчислити $\gamma(i)$ – характеристику інтервалу $(\bar{x}^{i-1};\, \bar{x}^i)$

$$\gamma(i)=\beta_k(\overline{x}^i-\overline{x}^{i-1})+\frac{(f_0(\overline{x}^i)-f_0(\overline{x}^{i-1}))^2}{\beta_k(\overline{x}^i-\overline{x}^{i-1})}-2(f_0(\overline{x}^i)+f_0(\overline{x}^{i-1})).$$

IX. Якщо $i<k$, то покласти $i=i+1$ і перейти на крок VIII; інакше перейти на крок X.

X. Знайти найменше значення $i_k\in[1:k]$, при якому виконується рівність

$$\gamma(i_k)=\max_{i\in[1;k]}\gamma(i).$$

XI. Обчислити наступне наближення

$$x^{k+1}=(\overline{x}^{i_k}+\overline{x}^{i_k-1})/2-(f_0(\overline{x}^{i_k})-f_0(\overline{x}^{i_k-1}))/2\beta_k.$$

XII. Покласти $k=k+1$ і перейти на крок IV.

***Теорема 1.*** *Нехай функція $f_0(x)$ задовольняє на $[a_0;b_0]$ умову Ліпшиця із константою $\gamma<\infty$*

$$|f_0(x')-f_0(x'')|\le\gamma|x'-x''|,\quad \forall x',\, x''\in[a_0;b_0].$$

*Тоді:* $(\boldsymbol{i})-$ *якщо функція $f_0(x)$ має на відрізку $[a_0;b_0]$ скінчене число локальних екстремумів, то будь-яка гранична точка $\overline{x}$ послідовності $\{x^k\}_{k=0}^{\infty}$, породженої алгоритмом 1, локально-оптимальна;* $(\boldsymbol{ii})-$ *якщо поряд із граничною точкою $\overline{x}$ існує інша гранична точка $\hat{x}$ послідовності $\{x^k\}_{k=0}^{\infty}$, тоді $f_0(\overline{x})=f_0(\hat{x})$;* $(\boldsymbol{iii})-$ *якщо $\overline{x}$ – гранична точка послідовності $\{x^k\}_{k=0}^{\infty}$, тоді $f_0(x^k)\ge f_0(\overline{x})$, $k=0,1,\ldots$;* $(\boldsymbol{iv})$ *– якщо на деякій ітерації алгоритму 1 справедлива нерівність $\beta_k>2\gamma$, тоді множина граничних точок послідовності $\{x^k\}_{k=0}^{\infty}$ співпадає з множиною точок абсолютного мінімуму функції $f_0$ на відрізку $[a_0;b_0]$.*

***Теорема 1'.*** *Нехай для наперед вибраної константи $\varepsilon>0$ ($\varepsilon$ – точність обчислення абсолютного мінімуму) за допомогою алгоритму 1 побудована послідовність точок $x^0,x^1,\ldots,x^{k(\varepsilon)}$, де $k(\varepsilon)$ – найменший індекс $k$, при якому виконується нерівність*

$$\overline{x}^{i_{k(\varepsilon)}}-\overline{x}^{i_{k(\varepsilon)}-1}\le\varepsilon.$$

*Тоді:* $(\boldsymbol{i})$ *– якщо на $(k(\varepsilon)+1)$-й ітерації виконується нерівність:*

$$\beta_{k(\varepsilon)}\ge\gamma(2\mu+1)/(\mu-1),$$

*де* $\mu=\min\limits_{i\in\Im}\dfrac{\overline{x}^i-\overline{x}^{i-1}}{\varepsilon}$; $\Im=\{i\,|\,\overline{x}^i-\overline{x}^{i-1}>\varepsilon,\ 1\le i\le k(\varepsilon)\}$, *то*

$$f_0(x)\ge\min_{0\le i\le k(\varepsilon)}f_0(x^i),\quad x\in[\overline{x}^{i-1};\overline{x}^i],\quad i\in\Im,$$

*тобто точка* $x^*$ *абсолютного мінімуму не може належати інтервалу, довжина якого перевищує дану точність* $\varepsilon$*;* (***ii***) *– якщо* $\alpha > \sqrt{\mu}/(\sqrt{\mu}-1)$*, то*

$$\min\{f_0(\overline{x}^i), f_0(\overline{x}^{i-1})\} > \min_{0\le j\le k(\varepsilon)} f_0(x^j), \quad i \in \Im,$$

*тобто оцінка* $\min\limits_{0\le j\le k(\varepsilon)} f_0(x^j)$ *мінімального значення функції* $f_0$ *досягається на одному із кінців інтервалу, довжина якого не перевищує точності* $\varepsilon$; (***iii***) *– для будь-якого* $\delta > 0$ *існує настільки велике значення коефіцієнту* $\alpha$*, що точки* $x^0, x^1, \ldots, x^{k(\varepsilon)}$ *утворюють* $(\varepsilon+\delta)$*-сітку в інтервалі* $[a_0;b_0]$.

*Зауваження 1.* Значення константи $\alpha$ повинно бути достатньо великим, щоб виконувалася нерівність $\beta_k > 2\gamma$, яка являється достатньою умовою збіжності алгоритму 1. Але, з іншого боку, зі збільшенням $\alpha$ зростає (наближаючись при $\alpha \to \infty$ до кількості вузлів сітки методу перебору) число обчислень функції $f_0$. У випадку, коли відома константа Ліпшиця $\gamma$, вибір значення параметру $\alpha$ не викликає ніяких труднощів. У випадку, якщо відомі лише грубі апріорні верхні оцінки константи Ліпшиця, можна скористатися описаним нижче алгоритмом.

### 2. Рандомізований метод глобального пошуку

***Алгоритм 2***

Початок. I. Вибрати нижню $\alpha_1$ і верхню $\alpha_2$ оцінки параметру $\alpha$.

II – X. Кроки II – X такі ж, як в алгоритмі 1, лише при обчисленні $\beta_k$ на кроці VI замість константи $\alpha$ потрібно взяти константу $\alpha_1$.

XI. Обчислити

$$v_1 = \beta_k(\overline{x}^{i_k} - \overline{x}^{i_k-1}) - (f_0(\overline{x}^{i_k}) + f_0(\overline{x}^{i_k-1})).$$

XII. Покласти $j_1 = i_k$, $\gamma_1 = \gamma(i_k)$.

XIII. Якщо $\delta_k = 0$, то покласти $\beta'_k = 1$ і перейти на крок XIV; інакше покласти $\beta'_k = \alpha_2\delta_k$ і перейти на крок XIV.

XIV. Покласти $i = 1$.

XV. Обчислити

$$\gamma(i) = \beta'_k(\overline{x}^i - \overline{x}^{i-1}) + \frac{(f_0(\overline{x}^i) - f_0(\overline{x}^{i-1}))^2}{\beta'_k(\overline{x}^i - \overline{x}^{i-1})} - 2(f_0(\overline{x}^i) - f_0(\overline{x}^{i-1})).$$

XVI. Якщо $i < k$, то покласти $i = i+1$ і перейти на крок XV; інакше перейти на крок XVII.

XVII. Знайти найменше значення індексу $i'_k \in [1:k]$, при якому виконується рівність $\gamma(i'_k) = \max\limits_{1 \le i \le k} \gamma(i)$.

XVIII. Обчислити
$$v_2 = \beta'_k(\overline{x}^{i'_k} - \overline{x}^{i'_k-1}) - (f_0(\overline{x}^{i'_k}) + f_0(\overline{x}^{i'_k-1})).$$

XIX. Покласти $j_2 = i'_k$, $\gamma_2 = \gamma(i'_k)$.

XX. Якщо $j_1 = j_2$, то обчислити $p = \alpha_1 /(\alpha_1 + \alpha_2)$ і перейти на крок XXI; інакше обчислити $p = (v_1 + v_2 - \gamma_2)/2v_1 + 2v_2 - \gamma_1 - \gamma_2)$ і перейти на крок XXI.

XXI. Однократно реалізувати випадковий механізм з двома виходами 1 та 2 відповідно, які мають ймовірності $p_1 = p$ і $p_2 = 1 - p$.

XXII. Якщо випадковий механізм дав результат 1, то покласти $\beta = \beta_k$, $j = j_1$ і перейти на крок XXIII; інакше покласти $\beta = \beta'_k$, $j = j_2$ і перейти на крок XXIII.

XXIII. Обчислити наступне наближення
$$x^{k+1} = (\overline{x}^{j} + \overline{x}^{j-1})/2 - (f_0(\overline{x}^{j}) - f_0(\overline{x}^{j-1}))/(2\beta).$$

XXIV. Покласти $k = k + 1$ і перейти на крок IV.

## 2.6. Методи пошуку інтервалу найбільших значень багатоекстремальних функцій



З а д а ч а 1. Знайти $\arg \max\limits_{x \in [a_0; b_0]} f_0(x)$ для заданої функції $f_0 : R^1 \to R^1$ і заданого відрізка $[a_0; b_0]$.

*Означення 1.* Зафіксуємо число $\delta > 0$ і визначимо клас $K_\delta$ функцій, що задовольняють умові: для будь-якої $\varphi(x)$ з $K_\delta$ існує такий відрізок $\Delta$ довжиною $\delta$ з $[a_0; b_0]$, що з $x \in \Delta$, $x' \notin \Delta$ випливає $\varphi(x) > \varphi(x')$. Визначення класу функцій $L_\delta$ випливає з визначення класу $K_\delta$, якщо покласти $\delta = (b_0 - a_0)/2$.

*Припущення 1.* Функція $f_0$ неперервна на $[a_0; b_0]$ і належить класу $K_\delta$.

Нижче приводиться метод для знаходження відрізка $\Delta$ для функції $f_0(x)$, що належить до класу $K_\delta$. Якщо обчислити значення функції $f_0(x)$ з класу $K_\delta$ в точках $a_0 + \delta$, $a_0 + 2\delta$, … , $a_0 + m\delta$, де
$$m = \begin{cases} (b_0 - a_0)/\delta - 1, & \text{якщо } (b_0 - a_0)/\delta - \text{ціле число;} \\ \mathrm{Ent}(b_0 - a_0)/\delta, & \text{якщо } (b_0 - a_0)/\delta - \text{неціле число,} \end{cases}$$

і серед цих точок вибрати точку $d$ з найбільшим значенням функції $f_0(d)$, то точка $d$ належить $\Delta$, а область $\Delta$ варто шукати на відрізку $[d-\delta, d+\delta]$. Таким чином, задача знаходження області $\Delta$ найбільших значень функції $f_0(x)$ із класу $K_\delta$ зводиться до задачі знаходження області $\Delta$ для функції із класу $L_\delta$. Якщо найбільше значення досягається у двох сусідніх точках $a_0+i\delta, a_0+(i+1)\delta$, то шукана область $\Delta$ збігаєтеся з інтервалом $[a_0+i\delta;\ a_0+(i+1)\ \delta]$.

Алгоритм, який приводиться нижче, для функції $f_0(x)$, що належить класу $L_\delta$ на відрізку $[d-\delta;\ \delta+d]$, за $(n-1)$ ітерацію знаходить відрізок $\Delta$ із заданою похибкою $\varepsilon>0$ за найменше можливе число ітерацій. Для початку роботи алгоритму необхідно знати деякий інтервал $[\alpha_0^+;\ \beta_0^+]$, що належить відрізку $\Delta$ і містить точку *d*.

***Алгоритм 1***

Початок. I. Вибрати число $\varepsilon>0$ – похибку обчислення відрізка $\Delta$.

II. Задати відрізок $\Delta_0=[\alpha_0^+;\ \beta_0^+]$, що містить точку $d$ і належить відрізку $\Delta$.

III. Обчислити натуральне число *n*, що задовольняє умовам:

$$F_{n-1}<(\delta-\delta_0)/\varepsilon\le F_n,$$

де $F_{n-1}$, $F_n$ – числа Фібоначі (тобто такі, що $F_0=F_1=1$, $F_i=F_{i-1}+F_{i-2},\ \ i=2, 3, \ldots$); $\delta_0=\beta_0^+-\alpha_0^+$.

IV. Покласти $k=0, G_0=\varnothing$ (тобто $G_0$ – порожня множина).

Основний цикл. V. Обчислити $\delta_k$ – довжину інтервалу $[\alpha_k^+;\ \beta_k^+]$, що належить відрізку $\Delta$, і величину $\nu_k=\delta-\delta_k$.

VI. Знайти точку $\alpha_k$, що лежить ліворуч від точки $\beta_k^+$ на відстані $\delta$, і знайти точку $\beta_k$, що лежить праворуч від точки $\alpha_k^+$ на відстані $\delta$.

VII. Якщо обидва півінтервали $[\alpha_k;\ \alpha_k^+)$ та $(\beta_k^+;\ \beta_k]$ не містять точок, у яких функція $f_0$ обчислювалася на попередніх ітераціях, то обчислити значення функції $f_0(x)$ в точках:

$$x_k'=\alpha_k^+-\nu_k F_{n-k-2}/F_{n-k};\quad x_k''=\beta_k^++\nu_k F_{n-k-2}/F_{n-k},$$

покласти $G_{k+1}=G_k\bigcup\{x_k'\}\bigcup\{x_k''\}$ і перейти на крок VIII;

якщо півінтервал $[\alpha_k;\ \alpha_k^+)$ містить точку $y_k'$, у якій функція $f_0$ обчислювалася на попередніх ітераціях, і півінтервал $(\beta_k^+;\ \beta_k]$ не містить точок, у яких функція $f_0$ обчислювалася на попередніх ітераціях, то обчислити функцію $f_0$, у точці $x_k'' = \beta_k - (y_k - \alpha_k)$, покласти $G_{k+1} = G_k \bigcup \{x_k''\}$ і перейти на крок VIII;

якщо півінтервал $(\beta_k^+;\ \beta_k]$ містить точку $y_k''$, у якій функція $f_0$ обчислювалася на попередніх ітераціях, а півінтервал $[\alpha_k;\alpha_k^+)$ не містить таких точок, то обчислити функцію $f_0$ в точці $x_k' = \alpha_k + (\beta_k - y_k'')$, покласти $G_{k+1} = G_k \bigcup \{x_k'\}$ і перейти на крок VIII.

VIII. Використовуючи наявну на $k$-й ітерації множину точок $G_{k+1}$, значення функції $f_0$ в точках множини $G_{k+1}$ і інтервал $[\alpha_k^+;\ \beta_k^+]$, що належить відрізку $\Delta$, обчислити найбільший інтервал $[\alpha_{k+1}^+;\beta_{k+1}^+]$, що належить відрізку $\Delta$**,** за допомогою наступних п'яти правил:

1) якщо $x, x' \in \Delta$ і $x \le z \le x'$, то $z \in \Delta$;

2) якщо $x \in \Delta$ і $f_0(x') \ge f_0(x)$, то $x' \in \Delta$;

3) якщо $x \notin \Delta$ і $f_0(x') \le f_0(x)$, то $x' \notin \Delta$;

4) якщо $x \in \Delta$ і $|x - x'| > \delta$, то $x' \notin \Delta$;

5) якщо $x \in [d-\delta; d]$ і $x' \in [d; d+\delta]$, то з $f_0(x) \ge f_0(x')$ і $|x - x'| \le \delta$ випливає, що $x \in \Delta$, а з $f_0(x) \ge f_0(x')$ і $|x - x'| > \delta$ випливає, що $x' \notin \Delta$.

IX. Якщо $k < n-2$, то покласти $k = k+1$ і перейти на крок V; інакше припинити обчислення.

***Теорема 1.*** *Якщо функція $f_0(x)$ належить класу $L_\delta$ на інтервалі $[d-\delta; d+\delta]$, то алгоритм 1 за $(n-1)$ ітерацію знаходить інтервал $[\alpha_{n-1}^+;\beta_{n-1}^+]$, що належить відрізку $\Delta$, такий, що $\beta_{n-1}^+ - \alpha_{n-1}^+ \ge \delta - \varepsilon$, причому алгоритм 1 є оптимальним алгоритмом пошуку області $\Delta$ з похибкою $\varepsilon$, тобто знаходить область $\Delta$ з похибкою $\varepsilon$, за найменше можливе число ітерацій.*

*Зауваження 1.* Якщо в алгоритмі 1 на кроці VII точки $x_k'$ і $x_k''$ обчислювати за формулами:

$$x_k' = \alpha_k + \nu_k F_{n-k-2} / F_{n-k}; \quad x_k'' = \beta_k - \nu_k F_{n-k-2} / F_{n-k},$$

то для цього алгоритму теорема 1 залишається в силі.

## 2.7. Методи пошуку глобального мінімуму, які використовують стохастичні автомати

З а д а ч а 0. Знайти $\arg\min\limits_{x\in[a_0;b_0]} f_0(x)$ для заданої функції $f_0:\ R^1 \to R^1$ і заданого інтервалу $[a_0;b_0]$.

*Припущення 0.* Функція $f_0$ неперервна на $[a_0;b_0]$ і має єдиний глобальний мінімум на $[a_0;b_0]$.

У початковій стадії методу інтервал $[a_0;b_0]$ розбивається на $m$ (досить велике число) підінтервалів $l_j$, $j=\overline{1,m}$, рівної довжини $w=(b_0-a_0)/m$ і кожному підінтервалу $l_j$, $j=\overline{1,m}$, ставиться у відповідність стан $S_j$, $j=\overline{1,m}$, стохастичного автомата. Стохастичний автомат задається вектором ймовірностей $p(k)=(p_1(k),...,p_m(k))$, кожна $j$-та компонента якого характеризує ймовірність $p_j(k)$ переходу автомата в стан $S_j$. На $k$-ій ітерації по заданому вектору $p(k)$ генерується стан автомата $S_{i_k}$, $i_k\in[1:m]$. Вихід стохастичного автомата належить інтервалу $[w(i_k-1);\ wi_k]$ з рівномірним розподілом ймовірності. Вектор ймовірностей $p(k)$ при переході від однієї ітерації до іншої перетворюється таким чином, що ймовірність $p_i(k)$, що відповідає інтервалу $l_i$, який містить точку глобального мінімуму, зростає, а всі ймовірності $p_j(k)$, $j=\overline{1,m}$, $j\neq i$, зменшуються. Приведений нижче алгоритм знаходить інтервал $l_{i_k}$, $i_k\in[1:m]$, у якому з ймовірністю досить близькою до одиниці знаходиться точка глобального мінімуму функції $f_0$ на $[a_0;b_0]$.

Методи пошуку глобального мінімуму, що використовують стохастичні автомати, доцільно застосовувати тоді, коли обчислення функції $f_0(x)$ супроводжується сильними збуреннями.

#### 1. Метод, який використовує модель Буша–Мостеллера

В алгоритмі 1 вектор ймовірностей $p(k)$ на кожній ітерації перетворюється за формулами Буша–Мостеллера.

***Алгоритм 1***

П о ч а т о к. І. Розбити інтервал $[a_0;b_0]$ на $m$ підінтервалів $l_i$ однакової довжини $w=(b_0-a_0)/m$ і поставити у відповідність кожному $l_i$, $i=\overline{1,m}$, єдиний стан автомата $S_i$, $i=\overline{1,m}$.

II. Покласти $p_i(0) = 1/m, \quad i = \overline{1,m}$.

III. Покласти $Q_i^* = \alpha_\infty$, $i = \overline{1,m}$, де $\alpha_\infty$ – досить велике додатне число.

IV. Вибрати константу $\varepsilon > 0$ ($\varepsilon$ — задана точність виконання умови оптимальності автомата у випадковому середовищі).

V. Покласти $k = 0$.

Основний цикл. VI. За допомогою вектора ймовірностей $p(k) = (p_1(k),..., p_m(k))$ згенерувати випадковий стан автомата $S(k) = S_{i_k}$, $i_k \in [1:m]$. Вибрати вихід автомата $x_{i_k}(k)$, що належить інтервалу $[w(i_k - 1), wi_k]$ (вважається, що випадкова величина $x_{i_k}(k)$ рівномірно розподілена на цьому інтервалі).

VII. Обчислити значення $f_0(x_{i_k})$.

VIII. Обчислити

$$Q^*_{\min} = \min_{i \in [1;m]} Q_i^* .$$

IX. Визначити вхід автомата $y(k+1)$ за правилами:

$$y(k+1) = \begin{cases} 0, & \text{якщо} \quad Q^*_{\min} < f_0\left(x_{i_k}\right); \\ 1, & \text{якщо} \quad Q^*_{\min} \geq f_0\left(x_{i_k}\right). \end{cases}$$



X. Покласти $Q_{i_k}^* = f_0\left(x_{i_k}\right)$ (решта значень $Q_i^*$, $i = \overline{1,m}$, $i \neq i_k$, не змінюються).

XI. Обчислити вектор $a(k) = (a_1(k),..., a_m(k))$, що визначає структуру автомата на $k$-ій ітерації, за правилами:

якщо $y(k+1) = 1$, то $a_{i_k}(k) = 1$ і $a_j(k) = 0$ для $j = \overline{1,m}$, $j \neq i_k$;

якщо $y(k+1) = 0$, то $a_j(k) = p_j(k)$, $j = \overline{1,m}$.

XII. Перетворити вектор ймовірностей $p(k)$ за формулою Буша–Мостеллера:

$$p(k+1) = (\lambda I + (1-\lambda) A(k)) p(k),$$

де $I$ – одинична матриця розміру $m \times m$; $\lambda$ – постійна величина, що належить інтервалу [0;1]; $A(k)$ – $m \times m$-матриця, що складається з $m$ однакових стовпців $a(k)$.

XIII. Якщо з заданою точністю $\varepsilon$ виконуються наступні умови оптимальності поведінки автомата у випадковому стаціонарному середовищі:

$$\left|p_{i_{k+1}}(k+1)-1\right| \le \varepsilon, \text{ якщо } Q^*_{i_{k+1}} < Q^*_j \text{ для всіх } j \ne i_{k+1}; \tag{1.7}$$

$$p_j(k+1) \le \varepsilon \text{ для всіх } j=\overline{1,m}, \quad j \ne i_{k+1}, \tag{1.8}$$

то зупинити обчислення (в цьому випадку точка глобального мінімуму з ймовірністю близькою до одиниці належить інтервалу $l_{i_{k+1}}$); інакше покласти $k=k+1$ і перейти на крок VI.

При вдалому розбитті інтервалу $[a_0;b_0]$ на підінтервали $l_j$, $j=\overline{1,m}$, алгоритм 1 приводить до інтервалу $[a_0+(i_{k+1}-1)w;\ a_0+i_{k+1}w)$, в якому з ймовірністю досить близькою до одиниці (за рахунок вибору $\varepsilon$), знаходиться точка глобального мінімуму функції $f_0(x)$ на інтервалі $[a_0;b_0]$.

### 2. Метод, який використовує усереднені значення функції



В цьому пункті наводиться алгоритм, в якому вектор ймовірностей $p(k)$ перетворюється за допомогою інформації про середні значення функції $f_0$ на відповідних підінтервалах.

***Алгоритм 2***

П о ч а т о к . I. Розбити інтервал $[a_0;b_0]$ на $m$ підінтервалів $l_i$ однакової довжини $w=(b_0-a_0)/m$ і поставити у відповідність кожному $l_i$, $i=\overline{1,m}$, єдиний стан стохастичного автомату $S_i$, $i=\overline{1,m}$.

II. Покласти $p_i(0)=1/m$, $i=\overline{1,m}$.

III. Вибрати константу $\varepsilon>0$ ($\varepsilon$ – задана точність виконання умови оптимальності стохастичного автомата у випадковому стаціонарному середовищі).

IV. Вибрати параметр алгоритму $\gamma>0$.

V. Обчислити масив середніх значень:

$$\bar{z}_i(0)=(f_0(w(i-1/2)))^{-\gamma}, \ i=\overline{1,m}.$$

VI. Покласти $k=0$.

О с н о в н и й   ц и к л. VII. За допомогою вектора ймовірностей $p(k)=(p_1(k),...,p_m(k))$ згенерувати випадковий стан автомата $S(k)=S_{i_k}$, $i_k\in[1:m]$. Вибрати вихід автомата $x_{i_k}(k)$, який належить інтервалу $[w(i_k-1);wi_k]$ (вважається, що випадкова величина $x_{i_k}(k)$ рівномірно розподілена на інтервалі $[w(i_k-1);wi_k]$).

VIII. Обчислити значення $f_0(x_{i_k})$.

IX. Обчислити значення

$$z_{i_k}(k+1) = (f_0(x_{i_k}))^{-\gamma}.$$

X. Обчислити масив середніх значень $\overline{z}_i(k+1)$, $i = 1,...,m$ за правилами:

$$\overline{z}_{i_k}(k+1) = \lambda \overline{z}_{i_k}(k) + (1-\lambda) z_{i_k}(k+1), \quad 0 < \lambda < 1;$$

$$\overline{z}_j(k+1) = \lambda \overline{z}_j(k), \quad j = \overline{1,m}; \quad j \neq i_k.$$

XI. Обчислити вектор ймовірностей $p(k+1) = (p_1(k+1),..., p_m(k+1))$ за формулами:

$$p_j(k+1) = \overline{z}_j(k+1) / \sum_{i=1}^{m} \overline{z}_i(k+1), \quad j = \overline{1,m}.$$

XII. Якщо з заданою точністю $\varepsilon$ виконуються умови оптимальності (1.7), (1.8) поведінки автомата у випадковому стаціонарному середовищі, то зупинити обчислення; інакше покласти $k = k+1$ і перейти на крок VII.

Алгоритм 2 забезпечує виконання умов (1.7), (1.8), якщо зроблено вдале розбиття інтервалу $[a_0; b_0]$ на підінтервали $l_j$, $j = \overline{1,m}$, і вибраний відповідний параметр алгоритму $\gamma$.



## 2.8. Адаптивні методи

### 1. Метод Кіфера–Вольфовиця

З а д а ч а 1. Знайти $\arg\max_{x \in R^1} f_0(x)$ для заданої функції $f_0 : R^1 \to R^1$.

*Припущення 1.* Функція $f_0(x)$ унімодальна.

Метод Кіфера–Вольфовиця застосовується для мінімізації унімодальних функцій, обчислення яких проводиться з випадковими похибками.

***Алгоритм 1***

П о ч а т о к. І. Вибрати довільне початкове наближення $x^0 \in R^1$.

II. Покласти $k = 0$.

О с н о в н и й ц и к л. III. Обчислити значення крокового множника $\rho_k$ і зміщення $\delta_k$, які задовольняють умовам теореми 1.

IV. Знайти величину $z(x^k + \delta_k)-$ результат обчислення з випадковими похибками значення функції $f_0(x)$ в точці $x^k + \delta_k$.

V. Знайти величину $z(x^k - \delta_k)-$ результат обчислення з випадковими похибками значення функції $f_0(x)$ в точці $x^k - \delta_k$.

VI. Обчислити наступне наближення

$$x^{k+1} = x^k + (\rho_k / \delta_k)(x^k + \delta_k) - z(x^k - \delta_k)).$$

VII. Покласти $k = k+1$ і перейти на крок III.

***Теорема 1.*** *Нехай функція* $f_0$ *є унімодальною і задовольняє нерівностям:*

$$|f_0(x) - f_0(y)| < \alpha_1 |x - x^*| + \alpha_2 < \infty,$$

*де* $x^*$ *– розв'язок задачі 1;* $\alpha_1$, $\alpha_2$ *– деякі константи.*

*Тоді, якщо в алгоритмі 1 похибка обчислень функції* $f_0$ *рівномірно обмежена і має нульове математичне сподівання, тобто:*

$$E(z(x^k \pm \delta_k) - f_0(x^k \pm \delta_k)) = 0;$$

$$\sum_{k=0}^{\infty} E(z(x^k \pm \delta_k) - f_0(x^k \pm \delta_k))^2 < \infty,$$

*і якщо крокові множники* $\rho_k$ *і зміщення* $\delta_k$ *такі, що:*

$$\rho_k > 0;\ \lim_{k\to\infty} \rho_k = 0;\ \lim_{k\to\infty} \delta_k = 0;$$

$$\sum_{k=0}^{\infty} \rho_k = \infty;\ \sum_{k=0}^{\infty} (\rho_k / \delta_k)^2 < \infty,$$

*то послідовність* $\left\{x^k\right\}_{k=0}^{\infty}$, *яка породжена алгоритмом 1, збігається до точки максимуму* $x^*$ *функції* $f_0$ *в середньоквадратичному і з ймовірністю 1, тобто:*

$$\lim_{k\to\infty} E(x^k - x^*)^2 = 0;\ \ P\{\lim_{k\to\infty} x^k = x^*\} = 1.$$

*Зауваження 1.* Якщо наближення $x^{k+1}$ на кроці VI алгоритму 1 обчислювати за формулою

$$x^{k+1} = x^k + \rho_k \operatorname{sign}\left\{\left(z(x^k + \delta_k) - z(x^k - \delta_k)\right)/2\delta_k\right\},$$

то отримаємо нормалізований варіант методу Кіфера–Вольфовиця.

**2. Простий перебір**

З а д а ч а 2. Знайти $\arg \min\limits_{x\in[a_0;b_0]} f_0(x)$ для заданої функції $f_0 : R^1 \to R^1$ і заданого відрізку $[a_0;b_0]$.

*Припущення 2.* Функція $f_0$ така, що на відрізку $[a_0;b_0]$ точка її локального мінімуму $x^*$ є точкою абсолютного мінімуму $f_0$ на відрізку $[a_0;b_0]$.

***Алгоритм 2***

П о ч а т о к. І. Задати: число $\varepsilon > 0$ – точність обчислення точки мінімуму функції $f_0$ на відрізку $[a_0;b_0]$; натуральне число $N$ (рекомендується вибирати число $N$ із відрізку $[10^1;10^3]$); покласти $k = 0$.

О с н о в н и й ц и к л. ІІ. Покласти $j = 0$.

III. Покласти $x_{k,0} = a_k$, обчислити $f_0(a_k)$ і покласти $f_0(x_{k,0}) = f_0(a_k)$.

IV. Обчислити зміщення

$$h_k = (b_k - a_k)/N.$$

V. Обчислити точку

$$x_{k,j+1} = x_{k,j} + h_k$$

та обчислити $f_0(x_{k,j+1})$.

VI. Якщо $f_0(x_{k,j+1}) > f_0(x_{k,j})$, то перейти на крок VII; інакше перейти на крок VIII.

VII. Якщо $j = 0$, то покласти $a_{k+1} = a_k$, $b_{k+1} = x_{k,1}$ і перейти на крок IX; інакше покласти $a_{k+1} = x_{k,j-1}$, $b_{k+1} = x_{k,j+1}$ і перейти на крок IX.

VIII. Якщо $j < N-1$, то покласти $j = j+1$ і перейти на крок V; інакше покласти $a_{k+1} = x_{k,N-1}$, $b_{k+1} = b_k$ і перейти на крок IX.

IX. Якщо $b_{k+1} - a_{k+1} > \varepsilon$, то перейти на крок X; інакше покласти $\overline{x}^* = (a_{k+1} + b_{k+1})/2$ і зупинити обчислення.

X. Покласти $k = k+1$ і перейти на крок II.

***Теорема 2.*** *Якщо виконано пропущення 2, то для довільного $\varepsilon > 0$ алгоритм 2 за скінченне число ітерацій приводить в точку $\overline{x}^*$ таку, що*

$$|\overline{x}^* - x^*| \le \varepsilon.$$

*Зауваження 2.* Недоліком методу простого перебору є те, що в багатьох «зайвих» точках доводиться обчислювати значення функції $f_0$, що небажано у випадку, коли $f_0$ визначається в результаті експерименту або коли для обчислення значення функції $f_0$ в точці потрібно значний машинний час.

**Приклад 1.** Знайти мінімум функції

$$f_0(x) = \max\left\{\frac{1}{3}x^3 - 2x^2 + 1, -2x^2 + 31\right\}$$

на проміжку [4;5] з точністю $\varepsilon = 0,05$.

*Розв'язування.*

**Алгоритм 2**

I. Задаємо $\varepsilon = 0,05$, $N = 10$, покладемо $k = 0$.

*1-а ітерація*:

II. Покладемо $j = 0$.

III. Покладемо $x_{0,0} = a = 4$, обчислюємо $f_0(4) = -1$ і покладемо $f_0(x_{0,0}) = -1$.

IV. Обчислюємо зміщення $h_0 = (b_0 - a_0) / N = 0{,}1$.

V. Обчислюємо точку $x_{0,1} = x_{0,0} + h_0 = 4 + 0{,}1 = 4{,}1$ та обчислюємо $f_0(x_{0,1}) = -2{,}62$.

VI. Оскільки $f_0(x_{0,1}) < f_0(x_{0,0})$,то переходимо на крок VIII.

VIII. Оскільки $j = 0 < N - 1 = 9$, то покладемо $j = 0 + 1 = 1$ і переходимо на крок V.

V. Обчислюємо точку $x_{0,2} = x_{0,1} + h_0 = 4{,}1 + 0{,}1 = 4{,}2$ та обчислюємо $f_0(x_{0,2}) = -4{,}28$.

VI. Оскільки $f_0(x_{0,2}) < f_0(x_{0,1})$, то переходимо на крок VIII.

VIII. Оскільки $j = 1 < N - 1 = 9$, то покладемо $j = 1 + 1 = 2$ і переходимо на крок V.

V. Обчислюємо точку $x_{0,3} = x_{0,2} + h_0 = 4{,}2 + 0{,}1 = 4{,}3$ та обчислюємо $f_0(x_{0,3}) = -5{,}98$.

VI. Оскільки $f_0(x_{0,3}) < f_0(x_{0,2})$, то переходимо на крок VIII.

VIII. Оскільки $j = 2 < N - 1 = 9$, то покладемо $j = 2 + 1 = 3$ і переходимо на крок V.

V. Обчислюємо точку $x_{0,4} = x_{0,3} + h_0 = 4{,}3 + 0{,}1 = 4{,}4$ та обчислюємо $f_0(x_{0,4}) = -7{,}72$.

VI. Оскільки $f_0(x_{0,4}) < f_0(x_{0,3})$, то переходимо на крок VIII.

VIII. Оскільки $j = 3 < N - 1 = 9$, то покладемо $j = 3 + 1 = 4$ і переходимо на крок V.

V. Обчислюємо точку $x_{0,5} = x_{0,4} + h_0 = 4{,}4 + 0{,}1 = 4{,}5$ та обчислюємо $f_0(x_{0,5}) = -9{,}125$.

VI. Оскільки $f_0(x_{0,5}) < f_0(x_{0,4})$, то переходимо на крок VIII.

VIII. Оскільки $j = 4 < N - 1 = 9$, то покладемо $j = 4 + 1 = 5$ і переходимо на крок V.

V. Обчислюємо точку $x_{0,6} = x_{0,5} + h_0 = 4{,}5 + 0{,}1 = 4{,}6$ та обчислюємо $f_0(x_{0,6}) = -8{,}8747$ .

VI. Оскільки $f_0(x_{0,6}) > f_0(x_{0,5})$, то переходимо на крок VII.

VII. Оскільки $j \neq 0$, то покладемо $a_1 = x_{0,4} = 4{,}4$, $b_1 = x_{0,6} = 4{,}6$ і переходимо на крок IX.

IX. Оскільки $b_1 - a_1 = 0{,}2 > \varepsilon = 0{,}05$, то переходимо на крок X.

X. Покладемо $k = 0 + 1 = 1$ і переходимо на крок II.

*2-а ітерація*:

II. Покладемо $j=0$.

III. Покладемо $x_{1,0}=a_1=4,4$, обчислюємо $f_0(4,4)=-7,72$ і покладемо $f_0(x_{1,0})=-7,72$.

IV. Обчислюємо зміщення $h_1=(b_1-a_1)/10=0,2/10= =0,02$ .

V. Обчислюємо точку $x_{1,1}=x_{1,0}+h_1=4,4+0,02=4,42$ та обчислюємо $f_0(x_{1,1})=-8,073$.

VI. Оскільки $f_0(x_{1,1})<f_0(x_{1,0})$, то переходимо на крок VIII.

VIII. Оскільки $j=0<N-1=9$, то покладемо $j=0+1=1$ і переходимо на крок V.

V. Обчислюємо точку $x_{1,2}=x_{1,1}+h_1=4,42+0,02=4,44$ та обчислюємо $f_0(x_{1,2})=-8,427$.

VI. Оскільки $f_0(x_{1,2})<f_0(x_{1,1})$,то переходимо на крок VIII.

VIII. Оскільки $j=1<N-1=9$, то покладемо $j=1+1=2$ і переходимо на крок V.

V. Обчислюємо точку $x_{1,3}=x_{1,2}+h_1=4,44+0,02=4,46$ та обчислюємо $f_0(x_{1,3})=-8,783$.

VI. Оскільки $f_0(x_{1,3})<f_0(x_{1,2})$, то переходимо на крок VIII.

VIII. Оскільки $j=2<N-1=9$, то покладемо $j=2+1=3$ і переходимо на крок V.

V. Обчислюємо точку $x_{1,4}=x_{1,3}+h_1=4,46+0,02=4,48$ та обчислюємо $f_0(x_{1,4})=-9,1408$.

VI. Оскільки $f_0(x_{1,4})<f_0(x_{1,3})$, то переходимо на крок VIII.

VIII. Оскільки $j=3<N-1=9$, то покладемо $j=3+1=4$ і переходимо на крок V.

V. Обчислюємо точку $x_{1,5}=x_{1,4}+h_1=4,48+0,02=4,5$ та обчислюємо $f_0(x_{1,5})=-9,125$.

VI. Оскільки $f_0(x_{1,5})>f_0(x_{1,4})$, то переходимо на крок VII.

VII. Оскільки $j\neq 0$, то покладемо $a_2=x_{1,3}=4,46$, $b_2=x_{1,5}=4,5$ і переходимо на крок IX.

IX. Оскільки $b_2-a_2=0,04<0,05$, то покладемо $\overline{x}^*=(a_2+b_2)/2=(4,46+4,5)/2=4,48$ і зупиняємо обчислення.

Таким чином, за дві ітерації методу простого перебору обчислили наближений розв’язок з точністю 0,05.

## Запитання та вправи для самостійної роботи до розділу 2.

1. Яку функцію називаєть унімодальною?
2. Запишіть правило обчислення членів послідовності Фібоначчі.
3. Для якого класу функцій можна застосовувати методи Фібоначчі?
4. В чому полягає відмінність модифікації методу Фібоначчі від основного алгоритму?
5. Для якого класу функцій методи Фібоначчі є оптимальними по кількості обчислень функції, яка мінімізується при заданій точності обчислення точки мінімуму $x^*$?
6. Чи використовуються в методі золотого перерізу числа Фібоначчі?
7. Для якого класу функцій можна застосовувати метод золотого перерізу?
8. Чи можна використовувати методи Фібоначчі і золотого перерізу для мінімізації строго опуклих донизу функцій?
9. Якою швидкістю збіжності володіють методи типу Ньютона?
10. Сформулюйте умови, при яких метод дотичних сходиться до точки мінімуму $x^*$.
11. Наведіть недоліки методу простого перебору.
12. Для яких функцій рекомендується використовувати методи пошуку глобального мінімуму, що використовують стохастичні автомати?
13. В чому полягає суть нормалізованого варіанту методу Кіфера–Вольфовиця?
14. Який алгоритм рекомендується використовувати для мінімізації функцій, обчислення яких проводиться із випадковими похибками?
15. Знайдіть найменше $N$, щоб $\left|F_n - 1/\sqrt{5}\left(\left(1+\sqrt{5}\right)/2\right)^n\right| < \varepsilon$ для усіх $n \geq N$. Розгляньте випадки $\varepsilon = 10^{-3}$, $\varepsilon = 10^{-4}$.
16. Знайдіть усі точки локального екстремуму функції $f_0(x) = \left|\left|\left|x^2 - 1\right| - 1\right| - 1\right|$ на відрізку $[a;b]$ при різних $a$ та $b$. При яких $a$ та $b$ ця функція буде унімодальною на відрізку $[a;b]$?
17. Знайдіть мінімум або максимум функції $f_0(x)$ на заданому відрізку $[a_0;b_0]$ з точністю $\varepsilon = 10^{-1}$, $\varepsilon = 10^{-2}$, $\varepsilon = 10^{-4}$:

| № | $f_0(x)$ | $[a_0;b_0]$ | № | $f_0(x)$ | $[a_0;b_0]$ |
|---|---|---|---|---|---|
| 1 | $f_0=\frac{x}{x^2+1}\to\min$ | $\left[-\frac{3}{2};0\right]$ | 14 | $f_0=\frac{\ln(x)-1}{x}\to\min$ | $[e;e^3]$ |
| 2 | $f_0=x+\frac{2}{x}\to\min$ | $\left[\frac{2}{3};3\right]$ | 15 | $f_0=x+\frac{4}{x^2}\to\min$ | $[1;4]$ |
| 3 | $f_0=x^4-2x^2+3\to\max$ | $[-1;1]$ | 16 | $f_0=\frac{x^2+1}{x}\to\max$ | $\left[-2;-\frac{1}{2}\right]$ |
| 4 | $f_0=\frac{x^4}{(x+1)^3}\to\min$ | $\left[-\frac{1}{2};1\right]$ | 17 | $f_0=x^2\exp(-x)\to\max$ | $[1;3]$ |
| 5 | $f_0=\frac{\ln(x)}{\sqrt{x}}\to\max$ | $[e;e^3]$ | 18 | $f_0=\exp(2x-x^2)\to\max$ | $[0;2]$ |
| 6 | $f_0=2x^3-6x^2-18x\to\max$ | $[0;2]$ | 19 | $f_0=x+\sqrt{3-x}\to\max$ | $[0;3]$ |
| 7 | $f_0=x^4-4x^3+6x^2-4x\to\min$ | $[-2;2]$ | 20 | $f_0=x^3-\frac{3}{4}x\to\min$ | $[-2;0]$ |
| 8 | $f_0=\frac{x^2-1}{x^2-4}\to\max$ | $[-1;1]$ | 21 | $f_0=\lvert x\rvert+(x-2)^2\to\min$ | $[-1;2]$ |
| 9 | $f_0=x+\frac{1}{x}\to\min$ | $\left[\frac{1}{2};2\right]$ | 22 | $f_0=2x^3-3x^2\to\max$ | $\left[-1;\frac{1}{2}\right]$ |
| 10 | $f_0=x^2+\exp(x)\to\min$ | $[-1;0]$ | 23 | $f_0=\frac{2x^3-32}{x^2}\to\min$ | $[3;5]$ |
| 11 | $f_0=\lvert x^2-3x+2\rvert+\lvert 2x-3\rvert\to\max$ | $[1;2]$ | 24 | $f_0=x+\frac{2}{x}-3\to\max$ | $\left[-5;-\frac{1}{2}\right]$ |
| 12 | $f_0=3x^4-4x^3+12x^2+1\to\min$ | $[-1;1]$ | 25 | $f_0=x-2\sqrt{x}\to\min$ | $[2;5]$ |
| 13 | $f_0=\frac{2x+1}{(x+3)(x-1)}\to\max$ | $\left[\frac{1}{2};2\right]$ | 26 | $f_0=x^4-2x^2+5\to\min$ | $\left[-2;-\frac{1}{2}\right]$ |

Розділ 3

# МЕТОДИ ОПТИМІЗАЦІЇ ДИФЕРЕНЦІЙОВАНИХ ФУНКЦІЙ В ЗАДАЧАХ БЕЗ ОБМЕЖЕНЬ

## 3.1. Аналітичні методи

Спочатку розглянемо аналітичний метод розв'язування задачі на безумовний екстремум, який базується на необхідних і достатніх умовах локального екстремуму.

Задача 1. Знайти $x^* = \arg\min_{x \in R^n} f_0(x)$ $\left(x^* = \arg\max_{x \in R^n} f_0(x)\right)$ для заданої неперервно диференційованої функції $f_0 : R^n \to R^1$.

### 1. Необхідна умова локальної оптимальності

Якщо $x^* \in R^n -$ точка локального екстремуму неперервно диференційованої в $R^n$ функції $f_0(x)$, то

$$\nabla f_0(x^*) = 0. \tag{3.1}$$



Точки $x^*$, які задовольняють рівність (3.1) називають ***стаціонарними*** точками функції $f_0(x)$. Таким чином, знаходження стаціонарних точок зводиться до розв'язування системи алгебраїчних рівнянь з $n$ змінними:

$$\begin{cases} \partial f_0(x)/\partial x_1 = 0, \\ \dots\dots\dots\dots\dots\dots \\ \partial f_0(x)/\partial x_n = 0. \end{cases}$$

### 2. Достатня умова локальної оптимальності

Нехай $f_0(x)$ двічі неперервно диференційована функція в точці $x^*$, де $x^* -$ стаціонарна точка. Тоді, якщо матриця $\nabla^2_{xx} f_0(x^*)$ є додатньо (від'ємно) визначеною, то $x^* -$ точка локального мінімуму (максимуму); якщо матриця $\nabla^2_{xx} f_0(x^*)$ є невизначеною, то $x^* -$ сідлова точка. Якщо матриця $\nabla^2_{xx} f_0(x^*)$ невід'ємно (недодатньо) визначена, то для визначення характеру стаціонарної точки $x^*$ потрібно досліджувати похідні більш високих порядків. Нагадаємо, що:

а) матриця $A$ називається ***додатньо визначеною*** (***невід'ємно визначеною***), якщо

$$(Ax, x) > 0 \qquad \left((Ax, x) \geq 0\right)$$

при всіх ненульових $x \in R^n$;

б) матриця $A$ називається ***від'ємно визначеною*** (*недодатньо визначеною*), якщо

$$(Ax, x) < 0 \qquad \left((Ax, x) \le 0\right)$$

при всіх ненульових $x \in R^n$;

в) матриця $A$ називається ***невизначеною***, якщо $(Ax, x) > 0$ для деяких ненульових $x \in R^n$ і $(A\overline{x}, \overline{x}) < 0$ для решту ненульових $\overline{x} \in R^n$.

Для перевірки знаковизначеності матриці $A$, як правило, використовують критерій Сильвестра: симетрична матриця $A$ є додатньо визначеною тоді і тільки тоді, коли всі її головні кутові мінори додатні. Щоб перевірити симетричну матрицю $A$ на від'ємну визначеність, необхідно перевірити матрицю $(-A)$ на додатну визначеність, або для того щоб матриця $A$ була від'ємно визначеною необхідно і достатньо, щоб головні мінори задовольняли умови: $M_1 < 0,\ M_2 > 0, \ldots, (-1)^n M_n > 0.$ Тепер запишемо аналітичний метод розв'язування задачі на безумовний екстремум.

***Алгоритм 1***

П о ч а т о к . I. Визначити градієнт $\nabla f_0(x)$.

II. Розв'язати систему алгебраїчних рівнянь з $n$ змінними:

$$\partial f_0(x) \big/ \partial x_j = 0, \quad j = \overline{1, n},$$

і отримати стаціонарні точки $x^{(k)}$, $k = \overline{1, M}$.

III. Визначити гесіан $\nabla^2_{xx} f_0(x)$.

IV. Покласти $k = 1$.

О с н о в н и й ц и к л .

V. Обчислити матрицю Гессе $\nabla^2_{xx} f_0(x^{(k)})$.

VI. Обчислити кутові мінори матриці $\nabla^2_{xx} f_0(x^{(k)})$. Якщо не всі кутові мінори ненульові, то для визначення характеру стаціонарної точки $x^{(k)}$ потрібно досліджувати похідні більш високих порядків; при цьому переходимо на крок IX. Інакше перейти на крок VII.

VII. Якщо $\nabla^2_{xx} f_0(x^{(k)})$ є додатньо визначеною, то стаціонарна точка $x^{(k)}$ є точкою локального мінімуму; перейти на крок IX. Інакше перейти на крок VIII.

VIII. Обчислити кутові мінори матриці $\left(-\nabla^2_{xx} f_0(x^{(k)})\right)$. Якщо матриця $\left(-\nabla^2_{xx} f_0(x^{(k)})\right)$ є додатньо визначеною, то $\nabla^2_{xx} f_0(x^{(k)})$ є від'ємно визначеною і стаціонарна точка $x^{(k)}$ є точкою локального максимуму. Інакше, матриця $\nabla^2_{xx} f_0(x^{(k)})$ є невизначеною і $x^{(k)}$ є сідловою точкою.

IX. Якщо $k < M$, то покласти $k = k + 1$ і перейти на крок V; інакше зупинити обчислення.

**Приклад 1.** Визначити точки локальних екстремумів функції $f_0(x) = x_1^3 - 2x_1x_2 + x_2^2 - 3x_1 - 2x_2$.

*Розв'язування.*

I. Визначаємо градієнт

$$\nabla f_0(x) = \begin{pmatrix} 3x_1^2 - 2x_2 - 3 \\ -2x_1 + 2x_2 - 2 \end{pmatrix}.$$

II. Розв'язуємо систему двох рівнянь з двома невідомими $x_1$ і $x_2$:

$$\begin{cases} 3x_1^2 - 2x_2 - 3 = 0, \\ -2x_1 + 2x_2 - 2 = 0 \end{cases} \Leftrightarrow \begin{cases} 3x_1^2 - 2x_2 - 3 = 0, \\ x_1 = x_2 - 1 \end{cases} \Leftrightarrow \begin{cases} 3(x_2-1)^2 - 2x_2 - 3 = 0, \\ x_1 = x_2 - 1 \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} 3x_2^2 - 6x_2 + 3 - 2x_2 - 3 = 0, \\ x_1 = x_2 - 1 \end{cases} \Leftrightarrow \begin{cases} x_2(3x_2 - 8) = 0, \\ x_1 = x_2 - 1 \end{cases} \Leftrightarrow \left[ \begin{array}{l} \begin{cases} x_2 = 0, \\ x_1 = -1 \end{cases} \\ \begin{cases} x_2 = 8/3, \\ x_1 = 5/3. \end{cases} \end{array} \right.$$

Таким чином, отримали дві стаціонарні точки $x^{(1)} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$ і $x^{(2)} = \begin{pmatrix} 5/3 \\ 8/3 \end{pmatrix}$, тому $M = 2$.

III. Визначаємо гесіан

$$\nabla_{xx}^2 f_0(x) = \begin{pmatrix} 6x_1 & -2 \\ -2 & 2 \end{pmatrix}.$$

IV. Покласти $k = 1$.

*1-а ітерація:*

V. Обчислюємо $\nabla_{xx}^2 f_0(x^{(1)}) = \begin{pmatrix} -6 & -2 \\ -2 & 2 \end{pmatrix}$.

VI. Обчислюємо кутові мінори матриці $\nabla_{xx}^2 f_0(x^{(1)})$:

$$M_1 = |-6| = -6 < 0;\ M_2 = \begin{vmatrix} -6 & -2 \\ -2 & 2 \end{vmatrix} = -6 \cdot 2 - (-2)(-2) = -16 < 0.$$

Всі кутові мінори ненульові, тому переходимо на крок VII.

VII. Матриця $\nabla_{xx}^2 f_0(x^{(1)})$ не є додатньо визначеною, тому переходимо на крок VIII.

VIII. Обчислюємо кутові мінори матриці $-\nabla_{xx}^2 f_0(x^{(1)}) = \begin{pmatrix} 6 & 2 \\ 2 & -2 \end{pmatrix}$:

$$M_1 = |6| = 6 > 0;\ M_2 = \begin{vmatrix} 6 & 2 \\ 2 & -2 \end{vmatrix} = 6 \cdot (-2) - 2 \cdot 2 = -16 < 0.$$

Тому матриця $\nabla^2_{xx} f_0(x^{(1)})$ є невизначеною, а стаціонарна точка $x^{(1)}$ є сідловою точкою.

IX. $k < 2,$ покласти $k = 1 + 1 = 2$ і перейти на крок V.

*2-а ітерація:*

V. Обчислюємо $\nabla^2_{xx} f_0(x^{(2)}) = \begin{pmatrix} -6 \cdot 5/3 & -2 \\ -2 & 2 \end{pmatrix} = \begin{pmatrix} 10 & -2 \\ -2 & 2 \end{pmatrix}.$

VI. Обчислюємо кутові мінори матриці $\nabla^2_{xx} f_0(x^{(2)})$:

$$M_1 = |10| = 10 > 0; \; M_2 = \begin{vmatrix} 10 & -2 \\ -2 & 2 \end{vmatrix} = 10 \cdot 2 - (-2)(-2) = 16 > 0.$$

Всі кутові мінори ненульові, тому переходимо на крок VII.

VII. Матриця $\nabla^2_{xx} f_0(x^{(2)})$ є додатньо визначеною, тому $x^{(2)}$ є точкою локального мінімуму, переходимо на крок IX.

IX. Зупиняємо обчислення.

Таким чином, знайшли точку локального мінімуму $x^{(2)} = \left(5/3; 8/3\right)^T$ і сідлову точку $x^{(1)} = (-1;0)^T$. Наступна таблиця описує поведінку функції $f_0$ в околі сідлової точки $x^{(1)}(-1;0)$.



| $x$ | $(-1;0)$ | $(-1,1;0)$ | $(-0,9;0)$ | $(-1;0,1)$ | $(-1;-0,1)$ | $(-1,1;-0,1)$ | $(-1,1;-0,1)$ |
|---|---|---|---|---|---|---|---|
| $f_0$ | 2 | 1,969 | 1,971 | 2,01 | 2,01 | 1,999 | 1,959 |

**Приклад 2.** Дослідити на екстремум функцію

$$f_0(x) = x_1^4 + x_2^4 - (x_1 + x_2)^2.$$

*Розв'язування.*

I. Визначаємо градієнт

$$\nabla f_0(x) = \begin{pmatrix} 4x_1^3 - 2(x_1 + x_2) \\ 4x_2 - 2(x_1 + x_2) \end{pmatrix}.$$

II. Розв'язуємо систему:

$$\begin{cases} 4x_1^3 - 2(x_1 + x_2) = 0 \\ 4x_2^3 - 2(x_1 + x_2) = 0 \end{cases} \Leftrightarrow \begin{cases} 2x_1^3 = x_1 + x_2 \\ 2x_2^3 = x_1 + x_2 \end{cases} \Leftrightarrow \begin{cases} x_1^3 = x_2^3 \\ 2x_2^3 = x_1 + x_2 \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} x_1 = x_2 \\ 2x_2^3 = 2x_2 \end{cases} \Leftrightarrow \begin{cases} x_1 = x_2 \\ x_2(x_2^2 - 1) = 0 \end{cases} \Leftrightarrow \left[ \begin{array}{l} \begin{cases} x_1 = 0 \\ x_2 = 0 \end{cases} \\ \begin{cases} x_1 = 1 \\ x_2 = 1 \end{cases} \\ \begin{cases} x_1 = -1 \\ x_2 = -1. \end{cases} \end{array} \right.$$

Таким чином, отримали три стаціонарні точки $x^{(1)} = (0;0)^T$, $x^{(2)} = (1;1)^T$, $x^{(3)} = (-1;-1)^T$, тому $M = 3$.

III. Визначаємо гесіан

$$\nabla_{xx}^2 f_0(x) = \begin{pmatrix} 12x_1^2 - 2 & -2 \\ -2 & 12x_2^2 - 2 \end{pmatrix}.$$

IV. Покладемо $k = 1$.

*1-а ітерація:*

V. Обчислюємо гесіан $\nabla_{xx}^2 f_0(x^{(1)}) = \begin{pmatrix} -2 & -2 \\ -2 & -2 \end{pmatrix}$.

VI. Обчислюємо кутові мінори:

$$M_1 = |-2| = -2 < 0;\ M_2 = \begin{vmatrix} -2 & -2 \\ -2 & -2 \end{vmatrix} = 0.$$

Безпосередня перевірка поведінки функції $f_0$ в околі точки $(0;0)^T$ показує, що точка $(0;0)^T$ є точкою локального максимуму. Відмітимо також, що $\sup\limits_{x \in R^2} f_0(x) = +\infty$. Переходимо на крок IX.

IX. $k = 1 < 3$, покладемо $k = 1 + 1 = 2$ і переходимо на крок V.

*2-а ітерація:*

V. Обчислюємо гесіан $\nabla_{xx}^2 f_0(x^{(2)}) = \begin{pmatrix} 10 & -2 \\ -2 & 10 \end{pmatrix}$.

VI. Обчислюємо кутові мінори:

$$M_1 = |10| = 10 > 0;\ M_2 = \begin{vmatrix} 10 & -2 \\ -2 & 10 \end{vmatrix} = 100 - 4 = 96 > 0.$$

Переходимо на крок VII.

VII. Матриця $\nabla_{xx}^2 f_0(x^{(2)})$ є додатньо визначеною, тому точка $x^{(2)} = (1;1)^T$ є точкою локального мінімуму. Переходимо на крок V.

*3-я ітерація:*

V. Обчислюємо гесіан $\nabla^2_{xx} f_0(x^{(3)}) = \begin{pmatrix} 10 & -2 \\ -2 & 10 \end{pmatrix}$, який дорівнює гесіану обчисленому на 2-й ітерації. Переходимо на крок VII.

VII. Точка $x^{(3)} = (-1;-1)^T$ є точкою локального мінімуму. Переходимо на крок IX.

IX. $k = 3 = M$, тому зупиняємо обчислення.

Таким чином, знайшли точку $x^{(1)} = (0;0)^T$ – точку локального максимуму і дві точки $x^{(2)} = (1;1)^T$, $x^{(3)} = (-1;-1)^T$ – точки локального мінімуму. Враховуючи поведінку функції $f_0$ на нескінченності, робимо висновок, що точки $x^{(2)} = (1;1)^T$ і $x^{(3)} = (-1;-1)^T$ є точками глобального мінімуму.

## 3.2. Градієнтні методи

З а д а ч а 1. Знайти $\arg\min_{x \in R^n} f_0(x)$ для заданої функції $f_0 : R^n \to R^1$.

*Припущення 1.* Функція $f_0$ – неперервно диференційована в $R^n$.

Надалі будемо розглядати числові методи розв'язування задачі. Відомо, що градієнт $\nabla f_0(x^k)$ є вектором нормалі до гіперповерхні рівня функції $f_0$ в точці $x^k$ (у двовимірному випадку – вектором нормалі до лінії рівня), який направлений у бік зростання функції $f_0$ (рис. 3.1)

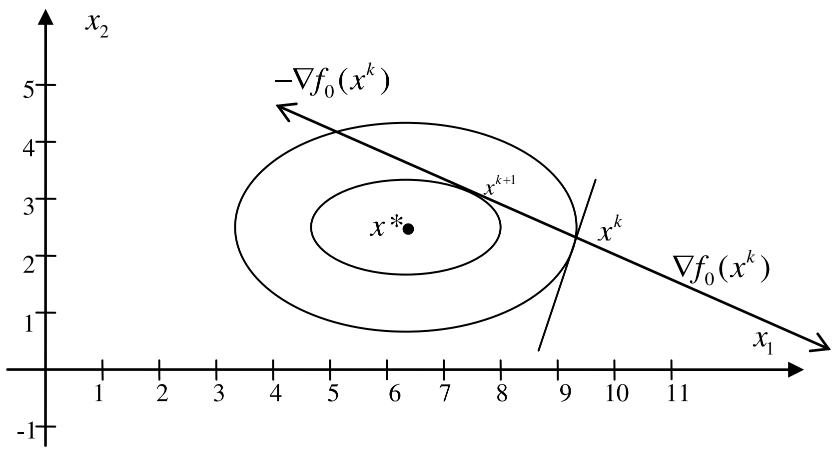



**Рис. 3.1**

Швидкість спадання функції $f_0$ в напрямі антиградієнта $-\nabla f_0(x^k)$ є найбільшою. В методі найшвидшого спуску шукають точку $x^{k+1}$, яка в напрямі антиградієнта дає найменше значення функції $f_0$. В різних варіантах градієнтного (антиградієнтного) спуску рухаються в напрямі

$-\nabla f_0(x^k)$ з певним кроком $\rho_k > 0$, або ж рухаються в напрямку $h^k$ "близькому" до $-\nabla f_0(x^k)$.

В градієнтних методах мінімізації за напрямок руху в $k$-й ітерації обирається вектор протилежний градієнту функції $f_0$ в точці $x^k$. Різні варіанти градієнтного методу відрізняються один від одного способом вибору крокового множника в $k$-й ітерації, а також тими чи іншими способами (різницевої) апроксимації градієнтів.

**1. Метод найшвидшого спуску**

В методі найшвидшого спуску кроковий множник $\rho_k$ $(k = 0,1,...)$ обчислюються із умови мінімуму функції $f_0$ в напрямку антиградієнта $-\nabla f_0\left(x^k\right)$, тобто

$$\rho_k = \arg\min_{\rho \ge 0} f_0\left(x^k - \rho \nabla f_0\left(x^k\right)\right).$$

Зауважимо, що для знаходження $\rho_k$ можна використовувати класичний аналітичний метод: $\partial f_0(x^k - \rho \nabla f_0(x^k)) / \partial \rho = 0$ або числові алгоритми одновимірної оптимізації розділу 2.

***Алгоритм 1***

Початок. I. Вибрати довільну початкову точку $x^0 \in R^n$ і покласти $k = 0$.

Основний цикл. II. Обчислити $\nabla f_0\left(x^k\right)$.

III. Якщо $\nabla f_0\left(x^k\right) = 0$, то покласти $x^* = x^k$ і закінчити обчислення; інакше перейти на крок IV.

IV. Обчислити кроковий множник $\rho_k$ із умови

$$f_0\left(x^k - \rho_k \nabla f_0\left(x^k\right)\right) = \min_{\rho \ge 0} f_0\left(x^k - \rho \nabla f_0\left(x^k\right)\right).$$

V. Обчислити наступне наближення

$$x^{k+1} = x^k - \rho_k \nabla f_0\left(x^k\right).$$

VI. Покласти $k = k + 1$ і перейти на крок II.

***Теорема 1.*** *Якщо виконано припущення 1 та* $(\boldsymbol{i})-$ *функція* $f_0$ *обмежена знизу в* $R^n$; $(\boldsymbol{ii})-$ *градієнт функції* $f_0$ *задовольняє умову Ліпшиця:*

$$\left\|\nabla f_0\left(x\right) - \nabla f_0\left(y\right)\right\| \le \alpha \left\|x - y\right\| \quad \forall x,\ \ y \in R^n,\ \alpha < \infty,$$

*то для нескінченної послідовності* $\left\{x^k\right\}_{k=0}^{\infty}$, *яка породжена алгоритмом 1, виконується співвідношення* $\left\|\nabla f_0\left(x^k\right)\right\| \to 0$ *при* $k \to \infty$.

***Теорема 1′.*** *Якщо виконані умови:* (***i***)*−* *функція* $f_0$ *двічі неперервно диференційована в* $R^n$; (***ii***)*−* *матриця других похідних* $\nabla^2_{xx} f_0(x)$ *функції* $f_0$ *задовольняє умовам:*

$$\beta\|y\|^2 \le \left(\nabla^2_{xx} f_0(x)y, y\right) \le \gamma\|y\|^2, \quad \gamma \ge \beta > 0,$$

*при будь-яких* $x, y \in R^n$, *то нескінченна послідовність* $\{x^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 1, збігається до точки мінімуму* $x^*$ *зі швидкістю геометричної прогресії із знаменником* $q = (\gamma - \beta)/(\gamma + \beta)$, *тобто*

$$\|x^{k+1} - x^*\| \le (\beta/\gamma)^{1/2} \|x^0 - x^*\| q^{k+1}.$$

Алгоритм 1 на практиці використовують рідко. Це пов’язано із неможливістю реалізувати крок IV в більшості практичних випадків, тому на практиці частіше використовують модифіковані методи найшвидшого спуску.

**Приклад 1.** За допомогою методу найшвидшого спуску знайти мінімум функції $f_0(x) = -4x_1 - 2x_2 + x_1^2 + x_2^2$, почавши ітераційний процес з точки $x^0 = (4;5)^T$.

*Розв’язування.* Знаходимо градієнт функції $f_0$

$$\nabla f_0(x) = \begin{pmatrix} -4 + 2x_1 \\ -2 + 2x_2 \end{pmatrix}.$$

I. $x^0 = (4;5)^T$, $k = 0$.

*1-а ітерація:*

II. $\nabla f_0(x^0) = \begin{pmatrix} -4 + 2\cdot 4 \\ -2 + 2\cdot 5 \end{pmatrix} = \begin{pmatrix} 4 \\ 8 \end{pmatrix}.$

III. $\nabla f_0(x^0) \ne 0,$ тому переходимо на крок IV.

IV. Обчислюємо кроковий множник $\rho_0$ із умови:

$$\rho_0 = \arg\min_{\rho \ge 0} f_0\left(x^0 - \rho\cdot\nabla f_0(x^0)\right),$$

де $x^0 - \rho\cdot\nabla f_0(x^0) = \begin{pmatrix} 4 \\ 5 \end{pmatrix} - \rho\begin{pmatrix} 4 \\ 8 \end{pmatrix} = \begin{pmatrix} 4 - 4\rho \\ 5 - 8\rho \end{pmatrix}.$

$$f_0\left(x^0 - \rho\cdot\nabla f_0(x^0)\right) = -4(4-4\rho) - 2(5-8\rho) + (4-4\rho)^2 + (5-8\rho)^2 =$$

$$= -16 + 16\rho - 10 + 16\rho + 16 - 32\rho + 16\rho^2 + 25 - 80\rho + 64\rho^2 = 80\rho^2 - 80\rho + 15.$$

Функція $f_0$ досягає мінімум по змінній $\rho$ у вершині параболи

$$\rho_{вер.} = -b/2a = -\left(-80/160\right) = 1/2 > 0.$$

Отже, $\rho_0 = 1/2$.

V. Обчислюємо наступне наближення

$$x^1 = \begin{pmatrix} 4 \\ 5 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 4 \\ 8 \end{pmatrix} = \begin{pmatrix} 4 & -2 \\ 5 & -4 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

VI. $k = 1$ і переходимо на крок II.

*2-а ітерація:*

II. $\nabla f_0(x^1) = \begin{pmatrix} -4 + 2\cdot 2 \\ -2 + 2\cdot 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0.$

III. $x^* = x^1 = \left(2;1\right)^T$.

Таким чином, знайшли точку мінімуму $x^* = (2;1)^T$, а $f_0(x^*) = -4\cdot 2 - 2\cdot 1 + 2^2 + 1^2 = -5.$



**2. Модифікований метод найшвидшого спуску**

***Алгоритм 2***

Кроки I–III такі, як в алгоритмі 1.

IV. Визначити число $\beta_k$

$$\beta_k = \arg\min_{\beta \ge 0} f_0\left(x^k - \beta \nabla f_0\left(x^k\right)\right).$$

V. Обчислити параметр $\lambda_k$, який задовольняє умову $(\boldsymbol{iii})$ теореми 2.

VI. Обчислити кроковий множник $\rho_k$ із умови

$$f_0\left(x^k - \rho_k \nabla f_0\left(x^k\right)\right) \le \left(1 - \lambda_k\right) f_0\left(x^k\right) + \lambda_k \beta_k . \qquad (3.2)$$

VII. Обчислити наступне наближення

$$x^{k+1} = x^k - \rho_k \nabla f_0\left(x^k\right).$$

VIII. Покласти $k = k+1$ і перейти на крок II.

***Теорема 2.*** *Якщо виконано припущення 1 та* $(\boldsymbol{i})-$ *градієнт функції* $f_0$ *задовольняє умову Ліпшиця:*

$$\left\|\nabla f_0\left(x\right) - \nabla f_0\left(y\right)\right\| \le \alpha \left\|x - y\right\| \quad \forall x,\ \ y \in R^n,\ \alpha < \infty;$$

$(\boldsymbol{ii})-$ *початкове наближення* $x^0$ *таке, що:*

$$\sup_{x', x'' \in X_0} \left\|x' - x''\right\| = \operatorname{diam} X_0 = \eta < \infty,$$

*де* $X_0 = \left\{ x \mid f_0(x) \le f_0(x^0),\ x \in R^n \right\}$;

$(\boldsymbol{iii})-$ *параметри* $\lambda_k$, $k = 0,1,...$, *задовольняють нерівностям* $\lambda \le \lambda_k \le 1$, *де* $\lambda -$ *довільна константа з напівінтервалу* $(0;1]$;

$(\boldsymbol{iv})-$ *функція* $f_0$ *опукла в* $R^n$, *то для нескінченної послідовності* $\left\{x^k\right\}_{k=0}^{\infty}$, *яка породжена алгоритмом 2, справедлива оцінка швидкості збіжності по функціоналу:*

$$f_0\left(x^k\right) - f_0^* \le 2\alpha\eta^2 \left( \sum_{i=0}^{k-1} \lambda_i \right)^{-1} \le 2\alpha\eta^2 / \lambda k,\ k = 1,2,....;\ \ f_0^* \triangleq \min_{x \in R^n} f_0(x).$$

***Теорема 2'.*** *Якщо виконані всі умови теореми 2 і функція* $f_0$ *є сильноопуклою донизу з параметром сильної опуклості* $v > 0$, *тобто*

$$f_0(\theta x + (1-\theta) y) \le \theta f_0(x) + (1-\theta) f_0(y) - \theta(1-\theta) v \|x - y\|^2$$

$$\forall x, y \in R^n,\ 0 \le \theta \le 1,$$

*то для нескінченної послідовності* $\left\{x^k\right\}_{k=0}^{\infty}$, *яка породжена алгоритмом 2, справедливі оцінки збіжності:*

$$f_0\left(x^k\right) - f_0^* \le \left(f_0\left(x^0\right) - f_0^*\right) \exp(-v\lambda k / 2\alpha),\ k = 1,2,... ;$$

$$\left\| x^k - x^* \right\|^2 \le (2/v)\left(f_0\left(x^0\right) - f_0^*\right) \exp(-v\lambda k / 2\alpha),\ k = 1,2,....$$

*Зауваження 2.* Реалізація алгоритму 2 на ЕОМ не призводить до шуканого розв'язку через наявність похибок обчислень. Тому алгоритм 2 використовують до тих пір, поки виконується нерівність:

$$\left(f_0\left(x^k\right) - f_0\left(x^{k+1}\right)\right) \Big/ \left\| \nabla f_0\left(x^k\right) \right\|^2 \ge \varepsilon, \tag{3.3}$$

де $\varepsilon > 0$ вибрана на даному етапі точність обчислень.

Якщо ж нерівність (3.3) порушується, то слід або підвищити точність обчислень, або перейти до іншого більш ефективного в даному разі методу. Така ситуація зазвичай трапляється в околі точки мінімуму $x^*$ або при попаданні в яр.

*Зауваження 2'.* Нерівність (3.2) еквівалентна нерівності

$$f_0\left(x^k\right) - f_0\left(x^{k+1}\right) \ge \lambda_k \left(f_0\left(x^k\right) - \beta_k\right).$$

Звідси бачимо, що при малих $\lambda_k$ точність обчислень $\beta_k$ (як мінімуму функції $f_0$ за напрямком $-\nabla f_0\left(x^k\right)$) може бути невисокою. Однак зі зменшенням $\lambda_k$ швидкість збіжності зменшується і тому важливо вдало вибирати $\lambda_k$ для кожного конкретного випадку.

### 3. Основний варіант градієнтного методу

***Алгоритм 3***

Початок. I. Вибрати довільну початкову точку $x^0 \in R^n$, довільну константу $\rho > 0$, довільний множник $\beta \in [1/2; 1)$, довільну константу $\varepsilon > 0$ $(\varepsilon \approx 1/2)$; покласти $k = 0$.

Основний цикл. II. Обчислити $\nabla f_0(x^k)$.

III. Якщо $\nabla f_0(x^k) = 0$, то покласти $x^* = x^k$ і закінчити обчислення; інакше перейти на крок IV.

IV. Покласти $\alpha = \rho$.

V. Обчислити точку

$$x = x^k - \alpha \nabla f_0(x^k).$$

VI. Якщо виконується нерівність

$$f_0(x) - f_0(x^k) \le \varepsilon\alpha(\nabla f_0(x^k), \nabla f_0(x^k)),$$

то покласти $\rho_k = \alpha$ і перейти на крок VII; інакше покласти $\alpha = \alpha\beta$ і перейти на крок V.

VII. Обчислити наступне наближення

$$x^{k+1} = x^k - \rho_k \nabla f_0(x^k).$$

VIII. Покласти $k = k + 1$ і перейти на крок II.

***Теорема 3.*** *Якщо виконано припущення 1 та* $(\boldsymbol{i})-$ *функція* $f_0$ *обмежена знизу*

$$f_0(x) \ge \delta_0 > -\infty, \ \forall x \in R^n;$$

$(\boldsymbol{ii})-$ *градієнт функція* $f_0$ *задовольняє умову Ліпшиця*

$$\|\nabla f_0(x) - \nabla f_0(y)\| \le \alpha_1 \|x - y\| \quad \forall x, \ y \in R^n, \ \alpha_1 < \infty,$$

*то для нескінченної послідовності* $\{x^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 3, виконується співвідношення*

$$\|\nabla f_0(x^k)\| \to 0 \ \textit{при} \ k \to \infty.$$

***Теорема 3'.*** *Якщо функція* $f_0$ *двічі неперервно диференційована в* $R^n$ *і її матриця других похідних* $\nabla_{xx}^2 f_0(x)$ *задовольняє нерівностям*

$$\gamma_1 \|y\|^2 \le (\nabla_{xx}^2 f_0(x) y, y) \le \gamma_2 \|y\|^2; \ \gamma_2 \ge \gamma_1 > 0$$

*при будь-яких* $x, y \in R^n$, *то для нескінченної послідовності* $\{x^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 3, справедливі співвідношення:*

$$x^k \to x^* \ \text{при} \ k \to \infty;$$

$$f_0(x^k) \to f_0(x^*) \ \text{при} \ k \to \infty,$$

*де* $x^*$ – *єдина точка мінімуму функції* $f_0$, *і наступні оцінки швидкості збіжності:*

$$\left\|x^k - x^*\right\| \le \delta_1 q^{k/2}, \ \ \delta_1 < \infty;$$

$$f_0\left(x^k\right) - f_0\left(x^*\right) \le \left(f_0\left(x^0\right) - f_0\left(x^*\right)\right) q^k,$$

*де* $q = 1 - 2\varepsilon(1-\varepsilon)\gamma_1(1+\gamma_1/\gamma_2)/\gamma_2$, *причому мінімальне значення* $q = 1 - (\gamma_1/2\gamma_2)(1+\gamma_1/\gamma_2)$ *досягається при* $\varepsilon = 1/2$.

***Теорема 3''.*** *Якщо виконані всі умови теореми 3 та існує таке число* $\delta_2 > 0$, *що*

$$\left\|\nabla f_0(x)\right\|^2 \ge \delta_2\left(f_0(x) - f_0^*\right) \ \ \forall x \in R^n; \quad f_0^* = \inf_{x \in R^n} f_0(x),$$

*то для послідовності* $\left\{x^k\right\}_{k=0}^{\infty}$, *яка породжена алгоритмом 3, справедлива оцінка*

$$f_0\left(x^k\right) - f_0^* \le q_1^k\left(f_0\left(x^0\right) - f_0^*\right), \ \ 0 < q_1 < 1.$$

**4. Модифікація основного варіанту градієнтного методу (метод з подрібненням кроку)**

***Алгоритм 4***

Початок.

I. Вибрати довільну початкову точку $x^1 \in R^n$, довільну константу $\rho_0 > 0$, довільний множник $\beta \in [1/2; 1)$, покласти $k = 1$.

Основний цикл.

II. Обчислити $\nabla f_0(x^k)$ і $\|\nabla f_0(x^k)\|$. Якщо $\nabla f_0(x^k) = 0$, то покласти $x^* = x^k$ і зупинитися; інакше перейти на крок III.

III. Обчислити одиничний вектор $h^k = \nabla f_0(x^k) / \|\nabla f_0(x^k)\|$.

IV. Покласти $\rho_k = \rho_{k-1}$.

V. Обчислити точку

$$x^{k+1} = x^k - \rho_k h^k.$$

VI. Якщо $f_0\left(x^{k+1}\right) < f_0\left(x^k\right)$, то покласти $k = k+1$ і перейти на крок II; інакше покласти $\rho_k = \beta\rho_k$ і перейти на крок V.

Для алгоритму 4 мають місце теореми аналогічні теореми 3 і $3'$.

**Приклад 2.** Використовуючи три ітерації алгоритму 4, розв'язати задачу безумовної мінімізації:

$$f_0(x) = x_1^2 + 2x_2^2 - 4x_1 + 2x_2 \to \min_{x \in R^2}.$$

*Розв'язування.* Спочатку знайдемо градієнт функції $f_0$:

$$\nabla f_0(x) = \begin{pmatrix} 2x_1 - 4 \\ 4x_2 + 2 \end{pmatrix}.$$

*1-а ітерація:*

I. $x^1 = (1;0)^T$, $\rho_0 = 1$, $\beta = 1/2$, $k = 1$.

II. $\nabla f_0(x^1) = \begin{pmatrix} 2\cdot 1 - 4 \\ 4\cdot 0 + 2 \end{pmatrix} = \begin{pmatrix} -2 \\ 2 \end{pmatrix} \neq 0.$ $\| \nabla f_0(x^1) \| = \sqrt{2^2 + 2^2} = 2\sqrt{2}.$

Переходимо на крок III.

III. $h^1 = \begin{pmatrix} -2 \\ 2 \end{pmatrix} \Big/ 2\sqrt{2} = \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \approx \begin{pmatrix} -0,7071 \\ 0,7071 \end{pmatrix}.$

IV. $\rho_1 = \rho_0 = 1.$

V. $x^2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} - 1 \cdot \begin{pmatrix} -0,7071 \\ 0,7071 \end{pmatrix} \cong \begin{pmatrix} -1,7071 \\ -0,7071 \end{pmatrix}.$

VI. $f_0(x^2) = -4\cdot 32;$ $f_0(x^1) = -3 \Rightarrow f_0(x^2) < f_0(x^1),$ тому $k = 2$ і переходимо на крок II.

*2-а ітерація:*

II. $\nabla f_0(x^2) = \nabla f_0 \begin{pmatrix} 1,7071 \\ -0,7071 \end{pmatrix} = \begin{pmatrix} -0,5858 \\ -0,8284 \end{pmatrix} \neq 0.$

$\| \nabla f_0(x^2) \| = \sqrt{0,5858^2 + 0,8284^2} = 1,0146.$ Переходимо на крок III.

III. $h^2 = \begin{pmatrix} -0,5858 \\ -0,8284 \end{pmatrix} \Big/ 1,0146 = \begin{pmatrix} -0,5774 \\ -0,8165 \end{pmatrix}.$

IV. $\rho_2 = 1.$

V. $x^3 = \begin{pmatrix} 1,7071 \\ -0,7071 \end{pmatrix} - 1 \cdot \begin{pmatrix} -0,5774 \\ -0,8165 \end{pmatrix} = \begin{pmatrix} 2,2845 \\ 0,1094 \end{pmatrix}.$

VI. $f_0(x^3) = -3,6763 \Rightarrow f_0(x^3) > f_0(x^2),$ тому $\rho_2 = (1/2)\cdot 1 = 0,5$ і переходимо на крок V.

V. $x^3 = \begin{pmatrix} 1,7071 \\ -0,7071 \end{pmatrix} - 0,5 \cdot \begin{pmatrix} -0,5774 \\ -0,8165 \end{pmatrix} = \begin{pmatrix} 1,9958 \\ -0,2989 \end{pmatrix}.$

VI. $f_0(x^3) = -4,4191 \Rightarrow f_0(x^3) < f_0(x^2),$ тому $k = 3$ і переходимо на крок II.

*3-я ітерація:*

II. $\nabla f_0(x^3)=\nabla f_0\begin{pmatrix}1,9958\\-0,2989\end{pmatrix}=\begin{pmatrix}-0,0084\\0,8044\end{pmatrix}\neq 0.$

$\|\nabla f_0(x^3)\|=\sqrt{0,0084^2+0,8044^2}=0,8044.$

III. $h^3=\begin{pmatrix}-0,0084\\0,8044\end{pmatrix}\Big/0,8044=\begin{pmatrix}-0,0104\\0,9999\end{pmatrix}.$

IV. $\rho_3=0,5.$

V. $x^4=\begin{pmatrix}1,9958\\-0,2989\end{pmatrix}-0,5\cdot\begin{pmatrix}-0,0104\\0,9999\end{pmatrix}=\begin{pmatrix}2,001\\-0,7989\end{pmatrix}.$

VI. $f_0(x^4)=-4,3213\Rightarrow f_0(x^4)>f_0(x^3),$ тому $\rho_3=0,5\cdot 0,5=0,25$ і переходимо на крок V.

V. $x^4=\begin{pmatrix}1,9958\\-0,2989\end{pmatrix}-0,25\begin{pmatrix}-0,0104\\0,9999\end{pmatrix}=\begin{pmatrix}1,9984\\-0,5489\end{pmatrix}.$

VI. $f_0(x^4)=-4,5012\Rightarrow f_0(x^4)<f_0(x^3),$ тому $k=4$ і зупиняємося.

Таким чином, виконавши три ітерації алгоритму 4, ми отримали наближений розв’язок задачі: $x^*\approx x^4=(1,9984;-0,5489)^T$, $f_0(x^*)=-4,5012,$ причому норма градієнта $\|\nabla f_0(x^*)\|=0,1956.$



**5. Градієнтний метод з постійним кроковим множником**

*Припущення 4.* (***i***)− функція $f_0$ неперервно диференційована; (***ii***)− градієнт функції $f_0$ задовольняє умову Ліпшиця з відомою константою $\alpha<\infty$.

***Алгоритм 5***

П о ч а т о к. I. Вибрати довільне початкове наближення $x^0\in R^n$ і довільну константу $\varepsilon\in(0;1)$ (доцільно вибирати $\varepsilon=1/2$); покласти $k=0$.

II. Вибрати довільне значення крокового множника $\rho$ із напіввідкритого інтервалу $\left(0;(1-\varepsilon)/\alpha\right]$.

О с н о в н и й ц и к л. III. Обчислити $\nabla f_0(x^k)$.

IV. Якщо $\nabla f_0(x^k)=0$, то покласти $x^*=x^k$ і закінчити обчислення; інакше перейти на крок V.

V. Обчислити наступне наближення

$$x^{k+1}=x^k-\rho\nabla f_0(x^k).$$

VI. Покласти $k=k+1$ і перейти на крок III.

***Теорема 5.*** *Якщо виконано припущення 5 і функція $f_0$ обмежена знизу, то для нескінченної послідовності $\{x^k\}_{k=0}^{\infty}$, яка породжена алгоритмом 5, виконується співвідношення $\left\|\nabla f_0\left(x^k\right)\right\| \to 0$ при $k \to \infty$.*

*Припущення 5′.* $(\boldsymbol{i})-$ функція $f_0$ двічі неперервно диференційована; $(\boldsymbol{ii})-$ матриця других похідних $\nabla_{xx}^2 f_0(x)$ функції $f_0$ задовольняє нерівностям:

$$\gamma_1\|y\|^2 \le \left(\nabla_{xx}^2 f_0(x)y, y\right) \le \gamma_2\|y\|^2; \ \gamma_2 \ge \gamma_1 > 0$$

при будь-яких $x, y \in R^n$ , причому константи $\gamma_2$ і $\gamma_1$ відомі.

***Алгоритм 5′***

Кроки I, III, IV, V, VI такі самі, як і в алгоритмі 5. Крок II алгоритму 5 слід замінити кроком II′, тобто вибрати довільне значення крокового множника $\rho$ із півінтервалу $\left(0; 2(1-\varepsilon)/\gamma_2\right]$.

***Теорема 5′.*** *Якщо виконані припущення 5′, то для нескінченної послідовності $\{x^k\}_{k=0}^{\infty}$, яка породжена алгоритмом 5′, справедливі граничні співвідношення:*

$$\lim_{k\to\infty} x^k = x^*; \ \lim_{k\to\infty} f_0\left(x^k\right) = f_0\left(x^*\right),$$

*де $x^*-$ єдина точка мінімуму функції $f_0$,*

*і наступні оцінки швидкості збіжності:*

$$\left\|x^k - x^*\right\| \le \delta_1 q^{k/2}, \ \delta_1 < \infty;$$

$$f_0\left(x^k\right) - f_0\left(x^*\right) \le \left(f_0\left(x^0\right) - f_0\left(x^*\right)\right)q^k,$$

*де* $q = 1 - 2\varepsilon(1-\varepsilon)\gamma_1\left(1+\gamma_1/\gamma_2\right)/\gamma_2$, *причому мінімальне значення* $q = 1 - \left(\gamma_1/2\gamma_2\right)\left(1+\gamma_1/\gamma_2\right)$ *досягається при* $\varepsilon = 1/2$.

***Алгоритм 5″***

Кроки I, III, IV, V, VI такі самі, як і в алгоритмі 5. Крок II алгоритму 5 слід замінити кроком II″, тобто вибрати довільне значення крокового множника $\rho$ із відкритого інтервалу $\left(0; 2/\gamma_2\right)$.

***Теорема 5″.*** *Якщо виконані припущення 5′, то для нескінченної послідовності $\{x^k\}_{k=0}^{\infty}$, яка породжена алгоритмом 5″, справедлива оцінка швидкості збіжності:*

$$\left\|x^k - x^*\right\| \le q^k \left\|x^0 - x^*\right\|,$$

*де* $q = \max\left\{\left|1-\rho\gamma_1\right|, \left|1-\rho\gamma_2\right|\right\}$, *причому мінімальне значення* $q = \left(\gamma_2 - \gamma_1\right)/\left(\gamma_2 + \gamma_1\right)$ *досягається при* $\rho = 2/\left(\gamma_2 + \gamma_1\right)$.

*Якщо ж в алгоритмі 5'' вибрати кроковий множник* $\rho = 2/(\gamma_2 + \gamma_1)$*, то, крім того, справедлива нерівність*

$$f_0\left(x^{k+1}\right) - f_0\left(x^*\right) \le \left((\gamma_2 - \gamma_1)/(\gamma_2 + \gamma_1)\right)^2 \left(f_0\left(x^k\right) - f_0\left(x^*\right)\right).$$

**6. Варіант градієнтного методу із матрицею прискорення збіжності**

*Припущення 6.* $(\boldsymbol{i})-$ функція $f_0$ неперервно диференційована; $(\boldsymbol{ii})-$ задана матриця прискорення збіжності $H(x)$ розмірності $n \times n$, елементами якої є неперервні функції від $x$.

***Алгоритм 6***

Початок. I. Вибрати довільне початкове наближення $x^0 \in R^n$ і довільні константи $\varepsilon \in (0;1)$, $\beta \in (0;1)$, $\rho > 0$ (доцільно вибирати $\varepsilon = 1/2$, $\beta \in (1/2; 4/5)$, $\rho = 1$); покласти $k = 0$.

Основний цикл. II. Обчислити вектор руху $h(x^k)$ до наступного наближення $x^{k+1}$ за формулою

$$h\left(x^k\right) = -H\left(x^k\right)\nabla f_0\left(x^k\right).$$

III. Якщо $h(x^k) = 0$, то покласти $x^* = x^k$ і завершити обчислення; інакше перейти на крок IV.

IV. Покласти $\mu = \rho$.

V. Обчислити значення

$$\theta\left(\mu, x^k\right) = f_0\left(x^k + \mu h\left(x^k\right)\right) - f_0\left(x^k\right) - \mu\varepsilon\left(\nabla f_0\left(x^k\right), h\left(x^k\right)\right).$$

VI. Якщо $\theta(\mu, x^k) \le 0$, то покласти $\rho_k = \mu$ і перейти на крок VII; інакше покласти $\mu = \beta\mu$ і перейти на крок V.

VII. Обчислити наступне наближення

$$x^{k+1} = x^k + \rho_k h\left(x^k\right).$$

VIII. Покласти $k = k + 1$ і перейти на крок II.

***Теорема 6.*** *Якщо виконані припущення 6 і умови:* $(\boldsymbol{i})-$ *функція* $f_0$ *двічі неперервно диференційована;* $(\boldsymbol{ii})-$ *матриця других похідних* $\nabla^2_{xx} f_0(x)$ *функції* $f_0$ *задовольняє нерівностям:*

$$\gamma_1 \|y\|^2 \le \left(\nabla^2_{xx} f_0(x) y, y\right) \le \gamma_2 \|y\|^2; \quad \gamma_2 \ge \gamma_1 > 0,$$

*при будь-яких* $x, y \in R^n$*;* $(\boldsymbol{iii})-$ $H(x)$ *– симетрична матриця, яка задовольняє нерівностям*

$$\gamma_3 \|y\|^2 \le \left(H(x) y, y\right) \le \gamma_4 \|y\|^2; \quad \gamma_4 \ge \gamma_3 > 0,$$

*при будь-яких* $x, y \in R^n$*, то для нескінченної послідовності* $\left\{x^k\right\}_{k=0}^{\infty}$*, яка породжена алгоритмом 6, справедливі оцінки швидкості збіжності*

$$\left\|x^k - x^*\right\| \le \delta_1 q^{k/2}, \ \ \delta_1 < \infty;$$

$$f_0\left(x^k\right) - f_0\left(x^*\right) \le \left(f_0\left(x^0\right) - f_0\left(x^*\right)\right) q^k,$$

*де* $q = 1 - 2\varepsilon(1-\varepsilon)\gamma_1\gamma_3(1+\gamma_1/\gamma_2)/(\gamma_2\gamma_4)$; $x^*$ *– єдина точка мінімуму функції* $f_0$, *причому мінімальне значення* $q = 1 - \gamma_1\gamma_3(1+\gamma_1/\gamma_2)/(2\gamma_2\gamma_4)$ *досягається при* $\varepsilon = 1/2$.

***Теорема 6′.*** *Якщо виконані припущення 6 і умови:* $(\boldsymbol{i})$ *–* $H(x)$ *– додатньо визначена матриця;* ***(ii)*** *– початкова точка* $x^0$ *така, що множина* $X_0 = \left\{x \mid f_0(x) \le f_0\left(x^0\right), x \in R^n\right\}$ *обмежена, то кожна гранична точка* $x'$ *нескінченної послідовності* $\left\{x^k\right\}_{k=0}^{\infty}$, *яка породжена алгоритмом 6, задовольняє умові* $\nabla f_0(x') = 0$.

*Зауваження 2.* Вибір матриці $H(x)$ істотно впливає на швидкість збіжності алгоритму 6 (прикладом цього є методи типу Ньютона і методи із розтягуванням простору).

### 7. Модифікований градієнтний метод, який не потребує обчислення похідних

Цей метод використовується в тих випадках, коли обчислення градієнту $\nabla_x f_0$ потребує значно більших витрат ніж обчислення значень функції $f_0$, яка мінімізується, або коли обчислення градієнту взагалі неможливе (наприклад, коли функція $f_0$ задана таблично або обчислюється за допомогою експериментів).

***Алгоритм 7***

Початок. I. Вибрати довільне початкове наближення $x^0 \in R^n$, яке задовольняє умові теореми 6, і константи $\varepsilon > 0$, $\lambda' > 0$, $\beta > 0$, $\lambda \in (0;1/2)$ (доцільно обирати $\varepsilon_0 \in \left[10^{-3};10^{-2}\right]$, $\lambda' \in \left[10^{-3};10^{-2}\right]$, $\beta \in [5;10]$, $\lambda = 0{,}4$).

II. Покласти $k = 0$, $\varepsilon = \varepsilon_0$.

Основний цикл. III. Обчислити вектор $h^k$, $j$-а компонента $h_j^k$ якого:

$$h_j^k = -1/\varepsilon\left(f_0\left(x^k + \varepsilon e^j\right) - f_0\left(x^k\right)\right)\big/\varepsilon, \quad j = \overline{1,n},$$

де $e^j$ – $j$-й орт.

IV. Обчислити

$$\Delta_k = f_0\left(x^k + \varepsilon\beta h^k\right) - f_0\left(x^k\right).$$

V. Якщо $\Delta_k < 0$, то, використовуючи алгоритм 7А, обчислити таке число $\bar{\rho}$, що

$$\bar{\rho}(1-\lambda)\Delta_k/(\beta\varepsilon) \le f_0\left(x^k+\bar{\rho}h^k\right)-f_0\left(x^k\right) \le \bar{\rho}\lambda\Delta_k/(\beta\varepsilon)$$

і перейти на крок VI; інакше покласти $\varepsilon = \varepsilon/2$ і перейти на крок III.

VI. Якщо $f_0\left(x^k+\bar{\rho}h^k\right)-f_0\left(x^k\right) \le -\lambda'\varepsilon$, то перейти на крок VII; інакше покласти $\varepsilon = \varepsilon/2$ і перейти на крок III.

VII. Обчислити наступне наближення

$$x^{k+1} = x^k + \bar{\rho}h^k .$$

VIII. Покласти $k = k+1$ і перейти на крок III.

***Алгоритм 7А*** **(алгоритм обчислення крокового множника $\bar{\rho}$ для алгоритму 7)**

I. Вибрати константу $\rho > 0$ (доцільно вибирати $\rho = 1$).

II. Визначити функції:

$$\psi_k(\alpha) = f_0\left(x^k+\alpha h^k\right)-f_0\left(x^k\right)-\alpha(1-\lambda)\Delta_k/(\beta\varepsilon);$$

$$\phi_k(\alpha) = f_0\left(x^k+\alpha h^k\right)-f_0\left(x^k\right)-\alpha\lambda\Delta_k/(\beta\varepsilon).$$

III. Покласти $\mu = \rho$.

IV. Обчислити значення $\psi_k(\mu)$.

V. Якщо $\psi_k(\mu) = 0$, то покласти $\rho = \mu$ і завершити обчислення; якщо $\psi_k(\mu) < 0$, то покласти $\mu = \rho + 1$ і перейти на крок IV; якщо $\psi_k(\mu) > 0$, то перейти на крок VI.

VI. Обчислити значення $\varphi_k(\mu)$.

VII. Якщо $\varphi_k(\mu) \le 0$, то покласти $\bar{\rho} = \mu$ і завершити обчислення; інакше покласти $a_0 = \mu - \rho$, $b_0 = \mu$ і перейти на крок VIII.

VIII. Покласти $i = 0$.

IX. Обчислити $v_i = (a_i + b_i)/2$.

X. Обчислити значення $\psi_k(v_i)$ і $\varphi_k(v_i)$.

XI. Якщо $\psi_k(v_i) \ge 0$ і $\varphi_k(v_i) \le 0$, то покласти $\bar{\rho} = v_i$ і завершити обчислення; інакше перейти на крок XII.

XII. Якщо $\psi_k(v_i) > 0$, то покласти $a_{i+1} = a_i$, $b_{i+1} = v_i$, $i = i+1$ і перейти на крок IX; інакше покласти $a_{i+1} = v_i$, $b_{i+1} = b_i$, $i = i+1$ і перейти на крок IX.

***Теорема 7.*** *Якщо виконані умови:* (***i***)*− функція* $f_0$ *неперервно диференційована;* (***ii***)*− початкове наближення* $x^0$ *в алгоритмі 6 таке, що множина* $X_0 = \left\{x \mid f_0(x) \le f_0\left(x^0\right), x \in R^n\right\}$ *обмежена, то послідовність* $x^0, x^1, \ldots,$ *яка породжена алгоритмом 6, або скінченна (тобто після скінченного числа ітерацій алгоритм 6 зациклюється в точці* $x^k$ *між* *кроками III і V або III і VI, повторюючи ділення* $\varepsilon$ *на 2), причому* $\nabla f_0(x^k)=0$, *або нескінченна і кожна її гранична точка* $x'$ *задовольняє умові* $\nabla f_0(x')=0$.

***Алгоритм 7′***

Усі кроки алгоритму 7, за винятком кроку V, залишаються без змін. Крок V алгоритму 7 слід замінити кроком V′, тобто якщо $\Delta_k<0$, то, використовуючи алгоритм 7Б, обчислити таке число $\bar{\rho}$, що

$$f_0\left(x^k+\bar{\rho}h^k\right)-f_0\left(x^k\right)\le\bar{\rho}\lambda\Delta_k/(\beta\varepsilon)$$

і перейти на крок VI; інакше покласти $\varepsilon=\varepsilon/2$ і перейти на крок III.

***Алгоритм 7Б* (алгоритм обчислення крокового множника $\bar{\rho}$ для алгоритму 7′)**

I. Вибрати константи $\gamma'\in(0,1)$, $\rho>0$ (доцільно вибирати $\gamma'\in(1/2,4/5)$, $\rho=1$).

II. Визначити функцію

$$\phi_k(\alpha)=f_0\left(x^k+\alpha h^k\right)-f_0\left(x^k\right)-\alpha\lambda\Delta_k/(\beta\varepsilon).$$

III. Покласти $\mu=\rho$.

IV. Обчислити значення $\varphi_k(\mu)$.

V. Якщо $\varphi_k(\mu)\le 0$, то покласти $\bar{\rho}=\mu$ і завершити обчислення; інакше покласти $\mu=\mu\gamma'$ і перейти на крок IV.

Для алгоритму 7′ справедлива теорема, аналогічна теоремі 7.

Позитивною стороною градієнтних методів є їх простота і те, що на кожній ітерації значення функції $f_0$ зменшується. До недоліків градієнтних методів можна віднести:

- збіжність (в загальному випадку) до стаціонарних точок чи точок локального мінімуму;
- неможливість безпосередньо застосовувати градієнтні методи в задачах недиференційованої оптимізації, та задачах умовної оптимізації;
- повільна збіжність для погано обумовлених функцій, гіперповерхні рівня яких відносяться до типу "яр". Наступний рисунок (рис. 3.2) ілюструє "зигзагоподібний" рух для методу найшвидшого спуску.

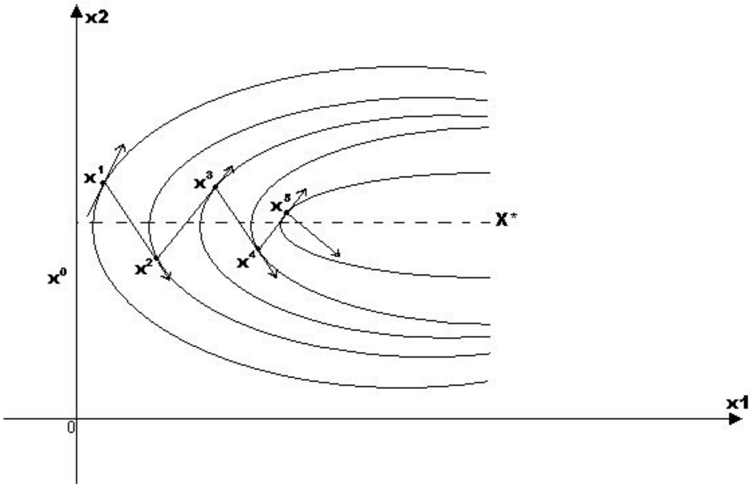



**Рис. 3.2**

## 3.3. Методи типу Ньютона

З а д а ч а  1. Знайти $\arg\min_{x\in R^n} f_0(x)$ для заданої неперервно диференційованої функції $f_0 : R^n \to R^1$.

На відміну від градієнтних методів, які використовують лінійну апроксимацію цільової функції $f_0$, методи типу Ньютона основані на квадратичній апроксимації функції $f_0$, використовуючи матрицю Гессе $\nabla_{xx}{}^2 f_0(x^k)$:

$$f_0(x) = f_0(x^k) + (\nabla f_0(x^k), x - x^k) +$$
$$+ (1/2)\cdot(\nabla_{xx}{}^2 f_0(x^k)(x - x^k), x - x^k).$$

Методи типу Ньютона володіють квадратичною (або надлінійною) швидкістю збіжності.

Для квадратичної функції вигляду $f_0(x) = (Hx, x)/2$ з додатньо визначеною матрицею $H$ точка мінімуму знаходиться за одну ітерацію. Збіжність до точки мінімуму функції $f_0$ гарантується у випадку коли $\nabla_{xx}{}^2 f_0(x^k)$ є додатньо визначеною матрицею. У більш загальному випадку методи типу Ньютона можуть визначати напрямки протилежні від напрямків до точки мінімуму.

Нагадаємо, що симетрична матриця додатньо визначена тоді і тільки тоді, коли всі її власні значення додатні. Якщо всі власні значення матриці $\nabla_{xx}{}^2 f_0(x^k)$ додатні, то локальна квадратична апроксимація $f_0(x)$ в точці

$x^k$ представляє собою еліпсоїд. Якщо ж, власні значення $\nabla_{xx}{}^2 f_0(x^k)$ мають протилежні знаки, то квадратична апроксимація являє собою сідло, а напрям руху в методах типу Ньютона буде в сторону сідла, а не точки мінімуму функції $f_0$.

#### 1. Метод Ньютона–Канторовича

*Припущення 1.* Функція $f_0$, яка мінімізується, є двічі неперервно диференційованою з оберненою матрицею Гессе $\nabla^2_{xx} f_0(x)$.

***Алгоритм 1***

Початок. I. Вибрати довільне початкове наближення $x^0 \in R^n$, покласти $k = 0$.

Основний цикл. II. Обчислити $\nabla f_0(x^k)$, $\nabla^2_{xx} f_0(x^k)$ та вектор руху $h^k$ до наступного наближення $x^{k+1}$ із системи рівнянь

$$-\nabla^2_{xx} f_0(x^k) h^k = \nabla f_0(x^k).$$

III. Якщо $h^k = 0$, то покласти $x^* = x^k$ і припинити обчислення; інакше перейти на крок IV.

IV. Обчислити наступне наближення

$$x^{k+1} = x^k + h^k.$$

V. Покласти $k = k + 1$ і перейти на крок II.

***Теорема 1.*** *Нехай виконується припущення 1. Тоді, якщо нескінченна послідовність* $\{x^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 1, збігається до точки* $x'$, *то* $\nabla f_0(x') = 0$.

**Приклад 1.** Методом Ньютона–Канторовича розв'язати задачу безумовної оптимізації

$$f_0(x) = -0,5x_2^3 + x_1 x_2 + 0,5x_1^2 - x_1 + 3x_2 + 4 \to \min_{x \in R^2},$$

починаючи ітераційний процес з точки $x^0 = (4; -1)^T$.

*Розв'язування.* Спочатку знаходимо градієнт і матрицю Гессе $H_1$ функції $f_0(x)$:

$$\nabla f_0(x) = \begin{pmatrix} x_2 + x_1 - 1 \\ -1,5x_2^2 + x_1 + 3 \end{pmatrix}, \qquad \nabla^2_{xx} f_0(x) = \begin{pmatrix} 1 & 1 \\ 1 & -3x_2 \end{pmatrix}.$$

Початок.

I. $x^0 = \begin{pmatrix} 4 \\ -1 \end{pmatrix}$, $k = 0$.

*1-а ітерація:*

II. $\nabla f_0\begin{pmatrix}4\\-1\end{pmatrix}=\begin{pmatrix}-1+4-1\\-1,5\cdot1+4+3\end{pmatrix}=\begin{pmatrix}2\\5,5\end{pmatrix},$

$\nabla^2_{xx} f_0\begin{pmatrix}4\\1\end{pmatrix}=\begin{pmatrix}1&1\\1&3\end{pmatrix}=G_0;\; G_0^{-1}=\frac{1}{2}\begin{pmatrix}3&-1\\-1&1\end{pmatrix}=\begin{pmatrix}1,5&0,5\\-0,5&0,5\end{pmatrix},$

$h^0=-G_0^{-1}\cdot\nabla f_0\begin{pmatrix}4\\1\end{pmatrix}=-\begin{pmatrix}1,5&-0,5\\-0,5&0.5\end{pmatrix}\begin{pmatrix}2\\5,5\end{pmatrix}=\begin{pmatrix}-3+2,75\\1-2,75\end{pmatrix}=$

$=\begin{pmatrix}-0,25\\-1,75\end{pmatrix}.$

III. $h^0\neq0,$ тому переходимо на крок IV.

IV. $x^1=x^0+h^0=\begin{pmatrix}4&-0,25\\-1&-1,75\end{pmatrix}=\begin{pmatrix}3,75\\-2,75\end{pmatrix}.$

V. $k=1$ і переходимо на крок II.

*2-а ітерація:*



II. $\nabla f_0\begin{pmatrix}3,75\\-2,75\end{pmatrix}=\begin{pmatrix}-2,75+3,75-1\\-1,5\cdot2,75^2+3,75+3\end{pmatrix}=\begin{pmatrix}0\\-4,59375\end{pmatrix},$

$\nabla^2_{xx} f_0\begin{pmatrix}3,75\\-2,75\end{pmatrix}=\begin{pmatrix}1&1\\1&8,25\end{pmatrix}=G_1,\; G_1^{-1}=\frac{1}{7,25}\begin{pmatrix}8,25&-1\\-1&1\end{pmatrix},$

$h^1=-G_1^{-1}\cdot\nabla f_0\begin{pmatrix}3,75\\-2,75\end{pmatrix}=-\frac{1}{7,25}\begin{pmatrix}8,25&-1\\-1&1\end{pmatrix}\begin{pmatrix}0\\-4,59375\end{pmatrix}=$

$=-\frac{1}{7,25}\begin{pmatrix}4,59375\\-4,59375\end{pmatrix}=\begin{pmatrix}-0,6336\\0,6336\end{pmatrix}.$

III. $h^1\neq0,$ тому переходимо на крок IV.

IV. $x^2=x^1+h^1=\begin{pmatrix}3,75-0,6336\\-2,75+0,6336\end{pmatrix}=\begin{pmatrix}3,1164\\-2,1164\end{pmatrix}.$

V. $k=2$ і переходимо на крок II.

*3-я ітерація:*

II. $\nabla f_0\begin{pmatrix}3,1164\\-2,1164\end{pmatrix}=\begin{pmatrix}-2,1164+3,1164-1\\-1,5\cdot2,1164^2+3,1164+3\end{pmatrix}=\begin{pmatrix}0\\-0,6023\end{pmatrix},$

$$\nabla^2_{xx} f_0\left(x^2\right) = \begin{pmatrix} 1 & 1 \\ 1 & 6,3492 \end{pmatrix} = G_2,\ G_2^{-1} = \frac{1}{5,3492}\begin{pmatrix} 6,3492 & -1 \\ -1 & 1 \end{pmatrix},$$

$$h^2 = -G_2^{-1}\cdot\nabla f_0\left(x^2\right) = -\frac{1}{5,3492}\begin{pmatrix} 6,3492 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ -0,6023 \end{pmatrix} =$$

$$= -\frac{1}{5,3492}\begin{pmatrix} 0,6023 \\ -0,6023 \end{pmatrix} = \begin{pmatrix} -0,1126 \\ 0,1126 \end{pmatrix}.$$

III. $h^2 \neq 0,$ тому переходимо на крок IV.

IV. $x^3 = x^2 + h^2 = \begin{pmatrix} 3,1164 - 0,1126 \\ -2,1164 + 0,1126 \end{pmatrix} = \begin{pmatrix} 3,0038 \\ -2,0038 \end{pmatrix}.$

V. $k = 3$ і переходимо на крок II.

*4-а ітерація:*

II. $\nabla f_0\left(x^3\right) = \begin{pmatrix} -2,0038 + 3,0038 - 1 \\ -1,5\cdot 2,0038^2 + 3,0038 + 3 \end{pmatrix} = \begin{pmatrix} 0 \\ -0,019 \end{pmatrix},$

$$\nabla^2_{xx} f_0\left(x^3\right) = \begin{pmatrix} 1 & 1 \\ 1 & 6,0114 \end{pmatrix} = G_3,\ G_3^{-1} = \frac{1}{5,0114}\begin{pmatrix} 6,0114 & -1 \\ -1 & 1 \end{pmatrix},$$

$$h^3 = -G_3^{-1}\cdot\nabla f_0\left(x^3\right) = -\frac{1}{5,0114}\begin{pmatrix} 6,0114 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ -0,019 \end{pmatrix} =$$

$$= -\frac{1}{5,0114}\begin{pmatrix} 0,019 \\ -0,019 \end{pmatrix} = \begin{pmatrix} -0,0038 \\ 0,0038 \end{pmatrix}.$$

III. $h^3 \neq 0,$ тому переходимо на крок IV.

IV. $x^4 = x^3 + h^3 = \begin{pmatrix} 3,0038 - 0,0038 \\ -2,0038 + 0,0038 \end{pmatrix} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}.$

V. $k = 4$ і переходимо на крок II.

*5-а ітерація:*

II. $\nabla f_0\left(x^4\right) = \begin{pmatrix} -2 + 3 - 1 \\ -1,5\cdot 4 + 3 + 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$

$$\nabla^2_{xx} f_0\left(x^4\right) = \begin{pmatrix} 1 & 1 \\ 1 & 6 \end{pmatrix} = G_4,\ G_4^{-1} = \frac{1}{5}\begin{pmatrix} 6 & -1 \\ -1 & 1 \end{pmatrix},$$

$$h^4 = -\frac{1}{5}\begin{pmatrix} 6 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

III. $h^4 = 0,\ x^* = x^4 = \begin{pmatrix} 3 \\ -2 \end{pmatrix}.$

Таким чином, за чотири ітерації алгоритму Ньютона–Канторовича знайшли оптимальну точку $x^* = (3;-2)^T$ і оптимальне значення функції $f_0$ :

$$f_0(3;-2) = -0,5\cdot(-8) - 6 + 0,5\cdot 9 - 3 + 3\cdot(-2) + 4 =$$
$$= 4 - 6 + 4,5 - 9 + 4 = -15 + 12,5 = -2,5.$$

**2. Узагальнений метод Ньютона–Канторовича**

*Припущення 2.* $(\boldsymbol{i})-$ функція $f_0$ двічі неперервно диференційована з оберненою матрицею Гессе $\nabla^2_{xx} f_0(x),$ яка задовольняє умову

$$\gamma_1 \|y\|^2 \le \left(\nabla^2_{xx} f_0(x) y, y\right) \le \gamma_2 \|y\|^2, \quad \gamma_2 \ge \gamma_1 > 0.$$

(Відмітимо, що у таких функцій існує єдина точка мінімуму $x^*$).

***Алгоритм 2***

Кроки I – III такі ж, як і в алгоритмі 1.

IV. Обчислити кроковий множник $\rho_k$ , який задовольняє умову

$$f_0(x^k + \rho_k h^k) = \min_{\rho \ge 0} f_0(x^k + \rho h^k).$$

V. Обчислити наступне наближення

$$x^{k+1} = x^k + \rho_k h^k.$$

VI. Покласти $k = k+1$ і перейти на крок II.

***Теорема 2.*** *Якщо виконано припущення 2, то нескінченна послідовність $\{x^k\}_{k=0}^{\infty}$, яка породжена алгоритмом 2, збігається до розв'язку $x^*$ із надлінійною швидкістю, тобто:*

$$\left\|x^{k+1} - x^*\right\| \le \nu_k \left\|x^k - x^*\right\|,$$

*де* $\nu_k = \left(\gamma_2 / \gamma_1^{\,3}\right)^{1/2} \left\|\nabla^2_{xx} f_0(x^k) - \nabla^2_{xx} f_0(x^k - \theta(x^k - x^*))\right\|, \;\; \theta \in [0;1].$

*Якщо, крім того, матриця других похідних $\nabla^2_{xx} f_0(x)$ функції $f_0$ задовольняє умову:*

$$\left\|\nabla^2_{xx} f_0(x) - \nabla^2_{xx} f_0(y)\right\| \le \alpha \|x - y\|, \;\; \alpha < \infty,$$

*при будь-яких $x, y \in R^n$, то нескінченна послідовність $\{x^k\}_{k=0}^{\infty}$ збігається до розв'язку $x^*$ з квадратичною швидкістю, тобто*

$$\left\|x^{k+1} - x^*\right\| \le (\gamma_2 / \gamma_1)^{1/2} (\alpha / \gamma_1) \left\|x^k - x^*\right\|^2.$$

Такий варіант узагальненого методу Ньютона–Канторовича в загальному випадку реалізувати на цифровій ЕОМ неможливо через вимогу одновимірної оптимізації на кроці IV алгоритму 2. Наведений нижче варіант позбавлений від цього недоліку.

***Алгоритм 2′***

Початок. I. Вибрати довільне початкове наближення $x^0 \in R^n$ і довільні константи $\varepsilon \in (0;1/2)$, $\beta \in (0;1)$, покласти $\rho = 1$ і $k = 0$.

Основний цикл. II. Обчислити вектор руху $h^k$ до наступного наближення $x^{k+1}$

$$h^k = -\left(\nabla^2_{xx} f_0(x^k)\right)^{-1} \nabla f_0(x^k).$$

III. Якщо $h^k = 0$, то покласти $x^* = x^k$ і припинити обчислення; інакше перейти на крок IV.

IV. Покласти $\mu = \rho$.

V. Обчислити значення

$$\delta = f_0(x^k + \mu h^k) - f_0(x^k) - \varepsilon\mu(\nabla f_0(x^k), h^k).$$

VI. Якщо $\delta \le 0$, то покласти $\rho_k = \mu$ або обчислити $\rho_k$ із нерівностей:

$$\rho_k(1-\varepsilon)(\nabla f_0(x^k), h^k) \le f_0(x^k + \rho_k h^k) - f_0(x^k) \le \rho_k \varepsilon(\nabla f_0(x^k), h^k)$$

і перейти на крок VII; інакше покласти $\mu = \mu\beta$ і перейти на крок V.

VII. Обчислити наступне наближення

$$x^{k+1} = x^k + \rho_k h^k.$$

VIII. Покласти $k = k+1$ і перейти на крок II.

***Теорема 2′.*** *Якщо виконано припущення 2, то нескінченна послідовність $\{x^k\}_{k=0}^{\infty}$, яка породжена алгоритмом 2′, збігається до точки мінімуму $x^*$ із надлінійною швидкістю, тобто:*

$$\left\|x^{N+l} - x^*\right\| \le \eta\lambda_N \ldots \lambda_{N+l},$$

*де $N$, $\eta < \infty$; $\lambda_{N+l} < 1$ при будь-якому $l \ge 0$; $\lambda_i \to 0$ при $i \to \infty$.*

*Якщо, крім того, матриця других похідних $\nabla^2_{xx} f_0(x)$ функції $f_0$ задовольняє умову:*

$$\left\|\nabla^2_{xx} f_0(x) - \nabla^2_{xx} f_0(y)\right\| \le \alpha \left\|x - y\right\|, \ \alpha < \infty,$$

*при будь-яких $x, y \in R^n$, то нескінченна послідовність $\{x^k\}_{k=0}^{\infty}$ збігається до розв'язку $x^*$ із квадратичною швидкістю, тобто*

$$\left\|x^{k+1} - x^*\right\| \le (\alpha / \gamma_1) \left\|x^k - x^*\right\|^2.$$

*Зауваження 2.* Спосіб вибору значень крокового множника $\rho_k$, $k = 0,1,\ldots,$ який використовується в алгоритмі 2′, при виконанні умов теореми 2′ гарантує, що, починаючи з деякої ітерації, алгоритм 2′ буде здійснюватися з одиничним кроковим множником $\rho_k = 1$, тобто алгоритм 2′ вироджується у звичайний метод Ньютона–Канторовича (алгоритм 1).

**Приклад 2.** Узагальненим методом Ньютона–Канторовича розв'язати задачу безумовної оптимізації квадратичної функції:

$$f_0(x) = \frac{1}{2}\left(\begin{pmatrix} 1 & 1 \\ 2 & 4 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) \to \min.$$

*Розв'язування.* Спочатку знайдемо градієнт і матрицю Гессе функції $f_0(x)$:

$$f_0(x) = \frac{1}{2}\left(\begin{pmatrix} x_1 + x_2 \\ 2x_1 + 4x_2 \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \frac{1}{2}\left(x_1^2 + x_1 x_2 + 2x_1 x_2 + 4x_2^2\right) =$$

$$= 0,5x_1^2 + 1,5x_1 x_2 + 2x_2^2,$$

$$\nabla f_0(x) = \begin{pmatrix} x_1 + 1,5x_2 \\ 1,5x_1 + 4x_2 \end{pmatrix}, \quad \nabla_{xx}^2 f_0(x) = \begin{pmatrix} 1 & 1,5 \\ 1,5 & 4 \end{pmatrix} = G_0.$$

**Алгоритм 2**

I. Вибираємо $x^0 = (5;5)^T$, покладемо $k = 0.$

*1-а ітерація:*

II. Обчислюємо градієнт, матрицю обернену до матриці Гессе та вектор руху $h^0$:

$$\nabla f_0(x^0) = \begin{pmatrix} 5 + 1,5 \cdot 5 \\ 1,5 \cdot 5 + 4 \cdot 5 \end{pmatrix} = \begin{pmatrix} 12,5 \\ 27,5 \end{pmatrix}, \quad G_0 = \begin{pmatrix} 1 & 1,5 \\ 1,5 & 4 \end{pmatrix},$$

$$G_0^{-1} = \frac{1}{4 - 2,25}\begin{pmatrix} 4 & -1,5 \\ -1,5 & 1 \end{pmatrix} = \frac{1}{1,75}\begin{pmatrix} 4 & -1,5 \\ -1,5 & 1 \end{pmatrix}.$$

Тепер обчислюємо вектор

$$h^0 = -G^{-1} \cdot \nabla f_0(x^0) = -\frac{1}{1,75} \cdot \begin{pmatrix} 4 & -1,5 \\ -1,5 & 1 \end{pmatrix} \cdot \begin{pmatrix} 12,5 \\ 27,5 \end{pmatrix} =$$

$$= -\frac{1}{1,75} \cdot \begin{pmatrix} 50 - 41,25 \\ -18,75 + 27,5 \end{pmatrix} = -\frac{4}{7} \cdot \begin{pmatrix} 8,75 \\ 8,75 \end{pmatrix} = -\begin{pmatrix} 5 \\ 5 \end{pmatrix} \neq 0.$$

III. Оскільки $h^0 \neq 0$, то переходимо на крок IV.

IV. Обчислюємо кроковий множник $\rho_0$:

$$f_0(x^0 + \rho h^0) = 0,5(x_1^0 + \rho h_1^0)^2 + 1,5(x_1^0 + \rho h_1^0) \cdot (x_2^0 + \rho h_2^0) \ + 2(x_2^0 + \rho h_2^0)^2 =$$

$$= 0,5(x_1^0)^2 + \rho x_1^0 h_1^0 + 0,5\rho^2 (h_1^0)^2 + 1,5x_1^0 x_2^0 + 1,5x_1^0 h_2^0 \rho + 1,5x_2^0 h_1^0 \rho +$$

$$+ 1,5h_2^0 h_1^0 \rho^2 + 2(x_2^0)^2 + 4x_2^0 h_2^0 \rho + 2\rho^2 (h_2^0)^2 = \rho^2[0,5(h_1^0)^2 + 1,5h_1^0 h_2^0 + 2(h_2^0)^2] +$$

$$+ \rho[x_1^0 h_1^0 + 1,5x_1^0 h_2^0 + 1,5x_2^0 h_1^0 + 4x_2^0 h_2^0] + 0,5(x_1^0)^2 + 1,5x_2^0 x_1^0 + 2(x_2^0)^2,$$

$$\rho_0 = -\frac{x_1{}^0 h_1{}^0 + 1{,}5 x_1{}^0 h_2{}^0 + 1{,}5 x_2{}^0 h_1{}^0 + 4 x_2{}^0 h_2{}^0}{(h_1{}^0)^2 + 3 h_1{}^0 h_2{}^0 + 4 (h_2{}^0)^2} =$$

$$= -\frac{5\cdot(-5) + 1{,}5\cdot 5\cdot(-5) + 1{,}5\cdot 5\cdot(-5) + 4\cdot 5\cdot(-5)}{25+75+100} = -\frac{-25-75-100}{200} = 1.$$

V. Обчислюємо наступне наближення

$$x^1 = x^0 + \rho_0 h^0 = \begin{pmatrix}5\\5\end{pmatrix} + \begin{pmatrix}-5\\-5\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}.$$

*2-а ітерація:*

II. Обчислюємо градієнт та вектор руху $h^1$ :

$$\nabla f_0(x^1) = \begin{pmatrix}0\\0\end{pmatrix},\ \ h^1 = \begin{pmatrix}0\\0\end{pmatrix}.$$

III. Оскільки $h^1 = 0$, то покладемо $x^* = x^1$ і зупиняємо обчислення.

Таким чином, за одну ітерацію узагальненого методу Ньютона–Канторовича знайшли оптимальний розв'язок $x^* = (0;0)^T$ задачі квадратичного програмування з додатньо визначеною і несиметричною матрицею $H = \begin{pmatrix}1 & 1\\ 2 & 4\end{pmatrix}$. Причому оптимальне значення цільової функції – $f_0(x^*) = 0$.



**3. Модифікація узагальненого методу Ньютона–Канторовича**

***Алгоритм 3***

Алгоритм 3 відрізняється від алгоритму 2′ тільки більш простим способом обчислення вектора $h^k$ на кроці II із системи рівнянь

$$-\left(\nabla^2_{xx} f_0(x^0)\right) h^k = \nabla f_0(x^k).$$

***Теорема 3.*** *Якщо виконується припущення 2 і матриця* $H = \left(\nabla^2_{xx} f_0(x^0)\right)^{-1}$ *задовольняє нерівності:*

$$\gamma_3 \| y \|^2 \le (Hy, y) \le \gamma \| y \|^2,\ \ \gamma_4 \ge \gamma_3 > 0,\ \ \forall y \in R^n,$$

*то для нескінченної послідовності* $\{x^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 3, справедливі оцінки швидкості збіжності:*

$$\left\| x^{k+1} - x^* \right\| \le \delta_1 q^{k/2},\ \ \delta_1 < \infty;$$

$$f_0(x^k) - f_0(x^*) \le \left( f_0(x^0) - f_0(x^*) \right) q^k,$$

*де* $q = 1 - 2\varepsilon(1-\varepsilon)\gamma_1\gamma_3(1+\gamma_1/\gamma_2)(\gamma_2\gamma_4)$.

*Якщо, крім того, початкове наближення $x^0$ в алгоритмі 3 вибране достатньо близько до точки $x^*$, то $\rho_k = 1,\ k = 0,1,\ldots,$ і послідовність $\{x^k\}_{k=0}^{\infty}$ збігається до $x^*$ із надлінійною швидкістю збіжності:*

$$\left\| x^{k+1} - x^* \right\| \le q_1 \left\| x^k - x^* \right\|,$$

*де*

$$q_1 = (1/\gamma_1) \max_{x \in X_0} \left\| \nabla^2_{xx} f_0(x^0) - \nabla^2_{xx} f_0(x) \right\|,$$

$$X_0 = \left\{ x \middle| f_0(x) \le f_0(x^0),\ x \in R^n \right\}.$$

В алгоритмі 3 на кожній ітерації для побудови вектора руху $h^k$ використовується одна і та ж матриця $-\nabla^2_{xx} f_0(x^0)$. Нижче наводиться модифікація узагальненого методу Ньютона–Канторовича, в якій відновлення матриці відбувається через $\tau$ $(\tau \ge 1)$ кроків. Такий алгоритм займає проміжне значення між алгоритмом 3 і узагальненим методом Ньютона–Канторовича (алгоритм 2).

***Алгоритм 3'***

Початок. I. Вибрати довільне початкове наближення $x^0 \in R^n$, константи $\varepsilon \in (0;1/2)$, $\beta \in (0;1)$ і натуральне число $\tau \ge 1$.

II. Покласти $j = 0,\ \rho = 1$.

Основний цикл. III. Покласти $i = 0$.

IV. Обчислити індекс $k = j\tau + i$.

V. Обчислити вектор руху $h^k$ до наступного наближення $x^{k+1}$ із системи

$$-\nabla^2_{xx} f_0(x^{j\tau}) h^k = \nabla f_0(x^k).$$

VI. Якщо $h^k = 0$, то покласти $x^* = x^k$ і припинити обчислення; інакше перейти на крок VII.

VII. Покласти $\mu = \rho$.

VIII. Обчислити значення

$$\delta = f_0(x^k + \mu h^k) - f_0(x^k) - \varepsilon\mu(\nabla f_0(x^k), h^k).$$

IX. Якщо $\delta \le 0$, то покласти $\rho_k = \mu$ і перейти на крок X; інакше покласти $\mu = \mu\beta$ і перейти на крок VIII.

X. Обчислити наступне наближення

$$x^{k+1} = x^k + \rho_k h^k.$$

XI. Якщо $i < \tau - 1$, то покласти $i = i + 1$ і перейти на крок IV; інакше перейти на крок XII.

XII. Покласти $j = j + 1$ і перейти на крок III.

***Теорема 3'.*** *Якщо виконані припущення 2 і:*

$$\left\|\nabla^2_{xx} f_0(x) - \nabla^2_{xx} f_0(y)\right\| \le \alpha \left\|x - y\right\|, \quad \alpha < \infty, \quad \forall x, y \in R^n,$$

*то нескінченна послідовність* $\{x^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 3′, збігається до точки мінімуму* $x^*$ *із надлінійною швидкістю, тобто*

$$\left\|x^{(j+1)\tau} - x^*\right\| \le \nu \left\|x^{j\tau} - x^*\right\|^{\tau+1}.$$

*Зауваження 3*. Збіжність з будь-якого початкового наближення є суттєвою перевагою узагальненого методу Ньютона–Канторовича і його модифікацій в порівнянні зі звичайним методом Ньютона–Канторовича, в якому збіжність гарантується лише при наявності достатньо вдалого початкового наближення.

### 4. Модифікований узагальнений метод Ньютона–Канторовича, який не вимагає обчислення матриці других похідних

***Алгоритм 4***

Початок. I. Вибрати початкове наближення $x^0 \in R^n$, яке задовольняє умовам теореми 4 і константи $\varepsilon_0 > 0,\ \beta > 0$, $\rho > 0,\ \alpha \in (0;1/2)$ (рекомендується вибирати: $\varepsilon_0 \in \left[10^{-3};10^{-2}\right]$, $\beta \in \left[10^{-3};10^{-2}\right],\ \alpha = 0,4;\ \rho = 1$).

II. Покласти $k = 0$.

Основний цикл. III. Обчислити $\nabla f_0(x^k)$.

IV. Якщо $\nabla f_0(x^k) = 0$, то покласти $x^* = x^k$ і припинити обчислення; інакше перейти на крок V.

V. Знайти $\varepsilon$ з умови $\varepsilon = \min\left\{\varepsilon_0, \left\|\nabla f_0(x^k)\right\|\right\}$.

VI. Знайти матрицю $H_k(\varepsilon)$ розмірності $n \times n$, $j$-й стовпчик якої обчислюється за формулою:

$$\frac{1}{\varepsilon}(\nabla f_0(x^k + \varepsilon e^j) - \nabla f_0(x^k)), \quad j = \overline{1,n},$$

де $e^j$ – $j$-й орт.

VII. Якщо існує обернена матриця $H_k^{-1}(\varepsilon)$ і виконується нерівність $(\nabla f_0(x^k), H_k^{-1}(\varepsilon)\nabla f_0(x^k)) > 0$, то обчислити вектор

$$h^k(\varepsilon) = -H_k^{-1}(\varepsilon)\nabla f_0(x^k)$$

і перейти на крок VIII; інакше покласти $\varepsilon = \varepsilon / 2$ і перейти на крок VI.

(Кроки VIII – XVIII призначені для обчислення такого $\bar{\rho}$, що

$$\bar{\rho}(1-\alpha)(\nabla f_0(x^k), h^k(\varepsilon)) \le f_0(x^k + \bar{\rho} h^k(\varepsilon)) - f_0(x^k) \le$$
$$\le \bar{\rho}\alpha(\nabla f_0(x^k), h^k(\varepsilon))).$$

VIII. Визначити функції:

$$\theta_k(\lambda) = f_0(x^k + \lambda h^k(\varepsilon)) - f_0(x^k) - \lambda(1-\alpha)(\nabla f_0(x^k), h^k(\varepsilon));$$

$$\varphi_k(\lambda) = f_0(x^k + \lambda h^k(\varepsilon)) - f_0(x^k) - \lambda\alpha(\nabla f_0(x^k), h^k(\varepsilon)).$$

IX. Покласти $\mu = \rho$.

X. Обчислити $\theta_k(\mu)$.

XI. Якщо $\theta_k(\mu) = 0,$ то покласти $\bar{\rho} = \mu$ і перейти на крок XIX; якщо $\theta_k(\mu) < 0,$ то покласти $\mu = \mu + \rho$ і перейти на крок X; якщо $\theta_k(\mu) > 0,$ то перейти на крок XII.

XII. Обчислити $\varphi_k(\mu)$.

XIII. Якщо $\varphi_k(\mu) \le 0,$ то покласти $\bar{\rho} = \mu$ і перейти на крок XIX; інакше покласти $a_0 = \mu - \rho,\ \ b_0 = \mu$, і перейти на крок XIV.

XIV. Покласти $i = 0$.

XV. Обчислити значення $\upsilon_i = (a_i + b_i)/2$.

XVI. Обчислити $\theta_k(\upsilon_i),\ \varphi_k(\upsilon_i)$.

XVII. Якщо $\theta_k(\upsilon_i) \ge 0$ і $\varphi_k(\upsilon_i) \le 0,$ то покласти $\bar{\rho} = \upsilon_i$ і перейти на крок XIX; інакше перейти на крок XVIII.

XVIII. Якщо $\theta_k(\upsilon_i) > 0,$ то покласти $a_{i+1} = a_i,\ b_{i+1} = \upsilon_i,\ i = i+1$ і перейти на крок XV; інакше покласти $a_{i+1} = \upsilon_i,\ b_{i+1} = b_i,\ i = i+1$ і перейти на крок XV.

XIX. Якщо виконується нерівність

$$f_0(x^k + \bar{\rho}h^k(\varepsilon)) - f_0(x^k) \le -\beta\varepsilon,$$

то перейти на крок XX; інакше покласти $\varepsilon = \varepsilon/2$ і перейти на крок VI.

XX. Обчислити наступне наближення

$$x^{k+1} = x^k + \bar{\rho}h^k(\varepsilon).$$

XXI. Покласти $k = k+1$ і перейти на крок III.

***Теорема 4.*** *Якщо виконані умови:* $(\boldsymbol{i})-$ *функція* $f_0-$ *двічі неперервно диференційована і її матриця Гессе невироджена при всіх x із достатньо великої області;* $(\boldsymbol{ii})-$ *функція* $f_0-$ *строго опукла;* $(\boldsymbol{iii})-$ *початкове наближення* $x^0$ *в алгоритмі 4 таке, що множина* $X_0 = \left\{x \middle| f_0(x) \le f_0(x^0),\ x \in R^n\right\}$ *обмежена, то нескінченна послідовність* $\left\{x^k\right\}_{k=0}^{\infty}$, *яка породжена алгоритмом 4, збігається до точки мінімуму* $x^*$ *функції* $f_0$.

## 3.4. Методи спряжених градієнтів

З а д а ч а 1. Знайти $\arg\min_{x \in R^n} f_0(x)$ для заданої неперервно диференційованої функції $f_0 : R^n \to R^1$.

#### 1. Загальна схема методу спряжених градієнтів

***Алгоритм 1***

П о ч а т о к . І. Вибрати початкове наближення $x^0 \in R^n$, яке задовольняє умови теореми 1; покласти $k = 0$.

О с н о в н и й ц и к л . ІІ. Обчислити $\nabla f_0(x^k)$.

ІІІ. Якщо $\nabla f_0(x^k) = 0$, то покласти $x^* = x^k$ і завершити обчислення; інакше перейти на крок IV.

IV. Обчислити вектор руху $h^k \in Q(x^k)$, де множина $Q(x)$ задається формулою

$$Q(x) = \left\{ h \middle| (\nabla f_0(x), h) < 0,\ h \in R^n \right\}.$$

V. Обчислити кроковий множник $\rho_k$, який задовольняє умову

$$f_0(x^k + \rho_k h^k) = \min_{\rho \ge 0} f_0(x^k + \rho h^k).$$

VI. Обчислити наступне наближення $x^{k+1} = x^k + \rho_k h^k$.

VII. Покласти $k = k + 1$ і перейти на крок II.

***Теорема 1.*** *Нехай $f_0$ – строго опукла донизу і двічі неперервно диференційована функція, причому для всіх $x \in X_0$,*

$$X_0 \triangleq \left\{ x \mid f_0(x) \le f_0(x^0),\ x \in R^n \right\}$$

*і для всіх $y \in R^n$ виконуються нерівності:*

$$\gamma_1 \|y\|^2 \le (\nabla^2_{xx} f_0(x) y, y) \le \gamma_2 \|y\|^2, \quad \gamma_1 > 0.$$

*Тоді якщо на кроці IV алгоритму 1 вектор $h^k$ вибирають так, що для деякого фіксованого $\alpha > 0$ виконується нерівність*

$$-(\nabla f_0(x^k), h^k) \ge \alpha \left\| \nabla f_0(x^k) \right\| \cdot \|h_k\|,$$

*то алгоритм 1 породжує або скінченну послідовність $\left\{x^k\right\}_{k=0}^{N}$, останній елемент якої мінімізує функцію $f_0$ в $R^n$, або нескінченну послідовність $\left\{x^k\right\}_{k=0}^{\infty}$, яка збігається до точки мінімуму $x^*$.*

Всі варіанти спряжених градієнтів відрізняються тільки різними способами обчислення векторів $h^k$ на кроці IV алгоритму 1 та способами обчислення крокового множника $\rho_k$ на кроці V алгоритму 1.

## 2. Метод спряжених градієнтів з відновленням

***Алгоритм 2***

Початок. I. Вибрати довільне початкове наближення $x^0 \in R^n$, натуральне число $\tau \ge n$ ($\tau$ – момент відновлення); покласти $k = 0$.

II. Обчислити $\nabla f_0(x^0)$ і покласти $g^0 = -\nabla f_0(x^0),\ h^0 = -\nabla f_0(x^0)$.

Основний цикл. III. Обчислити кроковий множник $\rho_k$, який задовольняє умову

$$f_0(x^k + \rho_k h^k) = \min_{\rho \ge 0} f_0(x^k + \rho h^k).$$

VI. Обчислити наступне наближення $x^{k+1} = x^k + \rho_k h^k$.

V. Обчислити $\nabla f_0(x^{k+1})$ і покласти $g^{k+1} = -\nabla f_0(x^{k+1})$.

VI. Якщо $g^{k+1} = 0$, тоді покласти $x^* = x^{k+1}$ і завершити обчислення; інакше перейти на крок VII.

VII. Обчислити коефіцієнт:

$$\beta_k = w\left(\frac{k+1}{\tau}\right)\frac{(g^{k+1} - g^k, g^{k+1})}{(g^k, g^k)},$$

де $w(t) = \begin{cases} 0, & \text{якщо } t - \text{ ціле число;} \\ 1, & \text{якщо } t - \text{ неціле число.} \end{cases}$

VIII. Обчислити вектор

$$h^{k+1} = g^{k+1} + \beta_k h^k.$$

IX. Покласти $k = k + 1$ і перейти на крок III.

***Теорема 2.*** *Якщо функція $f_0$ тричі неперервно диференційована та її матриця других похідних $\nabla^2_{xx} f_0(x)$ задовольняє нерівності:*

$$\gamma_1 \|y\|^2 \le (\nabla^2_{xx} f_0(x) y, y) \le \gamma_2 \|y\|^2, \quad 0 < \gamma_1 \le \gamma_2 < \infty$$

*для всіх $x, y \in R^n$, то нескінченна послідовність $\{x^k\}_{k=0}^{\infty}$, яка породжена алгоритмом 2 з довільним початковим наближенням $x^0$, збігається до розв'язку $x^*$ та існує таке ціле число $\nu > 0$ і константа $\delta_1 \in (0;\infty)$, що*

$$\|x^{k+n} - x^*\| \le \delta_1 \|x^k - x^*\|^2$$

*для всіх $k \ge \nu$, що належать множині $\{0, \tau, 2\tau, ...\}$.*

***Алгоритм 2'***

Початок. I. Вибрати довільне початкове наближення $x^0 \in R^n$, натуральне число $\tau \ge n$ ($\tau$ – момент відновлення); покласти $k = 0$.

II. Обчислити $\nabla f_0(x^0)$ і покласти $g^0 = -\nabla f_0(x^0),\ h^0 = -\nabla f_0(x^0)$.

Основний цикл. III. Обчислити кроковий множник $\rho_k$, який задовольняє умову

$$f_0(x^k+\rho_k h^k)=\min_{\rho\geq 0} f_0(x^k+\rho h^k).$$

IV. Обчислити наступне наближення

$$x^{k+1}=x^k+\rho_k h^k.$$

V. Обчислити $\nabla f_0(x^{k+1})$. Якщо $\nabla f_0(x^{k+1})=0$, тоді покласти $x^*=x^{k+1}$ і припинити обчислення; інакше перейти на крок VI.

VI. Якщо $k$ кратне $\tau$, то перейти на крок VII; інакше перейти на крок VIII.

VII. Покласти $g^{k+1}=-\nabla f_0(x^{k+1})$, $h^{k+1}=-\nabla f_0(x^{k+1})$ і перейти на крок XI.

VIII. Покласти $g^{k+1}=-\nabla f_0(x^{k+1})$.

IX. Обчислити коефіцієнт

$$\beta_k=\left\|g^{k+1}\right\|^2\Big/\left\|g^k\right\|^2.$$

X. Обчислити вектор $h^{k+1}=g^{k+1}+\beta_k h^k$.

XI. Покласти $k=k+1$ і перейти на крок III.

Для алгоритму $2'$ має місце теорема, аналогічна теоремі 2.

**Приклад 1.** Методом спряжених градієнтів розв'язати задачу безумовної оптимізації:

$$f_0(x)=-4x_1-2x_2+2x_1{}^2-x_1x_2+2x_2{}^2\to\min.$$



*Розв'язування.* Використаємо алгоритм $2'$, в якому кроковий множник $\rho_k$ знаходиться аналітично. Знайдемо $f_0(x^k+\rho h^k)$ як функцію від параметра $\rho$:

$$f_0(x^k+\rho h^k)=-4x_1{}^k-4\rho h_1{}^k-2x_2{}^k-2\rho h_2{}^k+2(x_1{}^k)^2+4x_1{}^k\rho h_1{}^k+2\rho^2(h_1{}^k)^2-$$
$$-(x_1{}^k+\rho h_1{}^k)(x_2{}^k+\rho h_2^k)+2(x_2{}^k)^2+4\rho x_2{}^k h_2{}^k+2\rho^2(h_2{}^k)^2=2\rho^2[(h_1{}^k)^2+(h_2{}^k)^2]-$$
$$-2\rho[2h_1{}^k+h_2{}^k-2x_1{}^k h_1{}^k-2x_2{}^k h_2{}^k]-4x_1{}^k-2x_2{}^k+2(x_1{}^k)^2+2(x_2{}^k)^2-x_1{}^k x_2{}^k-$$
$$-\rho x_1{}^k h_2^k-\rho h_1{}^k x_2{}^k-\rho^2 h_1{}^k h_2{}^k=\rho^2[2(h_1{}^k)^2+2(h_2{}^k)^2-h_1{}^k h_2{}^k]-\rho[x_1{}^k h_2{}^k+x_2{}^k h_1{}^k+$$
$$+4h_1{}^k+2h_2{}^k-4x_1{}^k h_1{}^k-4x_2{}^k h_2{}^k]-4x_1{}^k-2x_2{}^k+2(x_1{}^k)^2+2(x_2{}^k)^2-x_1{}^k x_2{}^k.$$

Мінімум цієї функції по $\rho$ досягається у вершині параболи

$$\rho_k=\frac{x_1{}^k h_2{}^k+x_2{}^k h_1{}^k+4h_1{}^k+2h_2{}^k-4x_1{}^k h_1{}^k-4x_2{}^k h_2{}^k}{2(2(h_1{}^k)^2+2(h_2{}^k)^2-h_1{}^k h_2{}^k)}.$$

Знайдемо градієнт функції $f_0$ : $\nabla f_0(x)=\begin{pmatrix}-4+4x_1-x_2\\-2-x_1+4x_2\end{pmatrix}$.

**Алгоритм $2'$**

I. Вибираємо початкове наближення $x^0=(0;0)^T,\tau=3,$ покладемо $k=0.$

II. Обчислюємо градієнт $\nabla f_0(x) = \begin{pmatrix} -4+0-0 \\ -2-0+0 \end{pmatrix} = \begin{pmatrix} -4 \\ -2 \end{pmatrix}$, покладемо $g^0 = (4;2)^T$; $h^0 = (4;2)^T$.

*1-а ітерація:*

III. Обчислюємо кроковий множник $\rho_0$:

$$\rho_0 = \frac{4\cdot(+4)+2\cdot(+2)}{2(2\cdot16+2\cdot4-8)} = \frac{+20}{2\cdot(32)} = \frac{5}{16}.$$

IV. Обчислюємо наступне наближення

$$x^1 = x^0 + \rho_0 h^0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \frac{5}{16}\cdot\begin{pmatrix} 4 \\ 2 \end{pmatrix} = \begin{pmatrix} 5/4 \\ 5/8 \end{pmatrix}.$$

V. Обчислюємо градієнт

$$\nabla f_0(x^1) = \begin{pmatrix} -4+4\cdot(5/4)-5/8 \\ -2-5/4+4\cdot(5/8) \end{pmatrix} = \begin{pmatrix} 3/8 \\ -3/4 \end{pmatrix} \neq 0,$$

тому переходимо на крок VI.

VI. 0 кратне 3, тому переходимо на крок VII.

VII. Покладемо $g^1 = (-3/8;3/4)^T$; $h^1 = (-3/8;3/4)^T$ і перейдемо на крок XI.

XI. Покладемо $k = 0+1 = 1$ і переходимо на крок III.

*2-а ітерація:*

III. Обчислюємо кроковий множник $\rho_1$:

$$\rho_1 = \frac{(5/4)\cdot(3/4)+(5/8)\cdot(-3/8)+4\cdot(-3/8)}{2\cdot(2\cdot 9/64+2\cdot 9/16+3/8\cdot 3/4)} +$$

$$+\frac{2\cdot(3/4)-3\cdot(5/4)\cdot(-3/8)-4\cdot(5/4)\cdot(3/4)}{2\cdot(2\cdot 9/64+2\cdot 9/16+3/8\cdot 3/4)} =$$

$$= \frac{15/16-15/64-12/8+6/4+45/32-60/32}{2\cdot(9/32+18/16+9/32)} = 5/72.$$

IV. Обчислюємо наступне наближення

$$x^2 = x^1 + \rho_1 h^1 = \begin{pmatrix} 5/4 \\ 5/8 \end{pmatrix} + \frac{5}{72}\cdot\begin{pmatrix} -3/8 \\ 3/4 \end{pmatrix} = \begin{pmatrix} 1{,}224 \\ 0{,}677 \end{pmatrix}.$$

V. Обчислюємо градієнт

$$\nabla f_0(x^2) = \begin{pmatrix} -4+4\cdot725/576-195/288 \\ -2-725/576+4\cdot(195/288) \end{pmatrix} = \begin{pmatrix} 0{,}3576 \\ -0{,}5503 \end{pmatrix} \neq 0.$$

VI. Індекс $k = 1$ не кратний 3, тому переходимо на крок VIII.

VIII. Покладемо $g^2 = \left(-0,3576;\ 0,5503\right)^T$.

IX. Обчислюємо коефіцієнт

$$\beta_1 = \| g^2 \|^2 / \| g^1 \|^2 = 0,430708 / 0,703125 = 0,612562.$$

X. Обчислюємо вектор

$$h^2 = g^2 + \beta_1 h^1 =$$

$$= \begin{pmatrix} -0,3576 \\ 0,5503 \end{pmatrix} + 0,612562 \cdot \begin{pmatrix} -0,375 \\ 0,75 \end{pmatrix} = \begin{pmatrix} -0,5873 \\ 1,0097 \end{pmatrix}.$$

XI. Покладемо $k = 1 + 1 = 2$ і переходимо на крок III.

*3-я ітерація:*

III. Обчислюємо кроковий множник $\rho_2$:

$$\rho_2 = \frac{(1,224 \cdot 1,0097 + 0,677 \cdot (-0,5873) + 4(-0,5873)}{2 \cdot (2 \cdot 0,5873^2 + 2 \cdot 1,0097^2 + 0,5873 \cdot 1,0097)} +$$

$$+ \frac{2 \cdot 1,0097 - 4 \cdot 1,224 \cdot (-0,5873) - 4 \cdot 0,677 \cdot 1,0097}{2 \cdot (2 \cdot 0,5873^2 + 2 \cdot 1,0097^2 + 0,5873 \cdot 1,0097)} =$$

$$= \frac{0,649624}{6,64365514} = 0,09778.$$

IV. Обчислюємо наступне наближення

$$x^3 = x^2 + \rho_2 h^2 = \begin{pmatrix} 1,224 \\ 0,677 \end{pmatrix} + 0,09778 \cdot \begin{pmatrix} -0,5873 \\ 1,0097 \end{pmatrix} = \begin{pmatrix} 1,167 \\ 0,776 \end{pmatrix}.$$

V. Обчислюємо градієнт

$$\nabla f_0(x^3) = \begin{pmatrix} -4 + 4 \cdot 1,167 - 0,776 \\ -2 - 1,167 + 4 \cdot 0,776 \end{pmatrix} = \begin{pmatrix} -0,108 \\ -0,063 \end{pmatrix} \neq 0.$$

VI. Індекс $k = 2$ не кратний 3, тому переходимо на крок VIII.

VIII. Покладемо $g^3 = \left(0,108; 0,063\right)^T$.

IX. Обчислюємо коефіцієнт

$$\beta_2 = \| g^3 \|^2 / \| g^2 \|^2 = 0,015633 / 0,430708 = 0,036296.$$

X. Обчислюємо вектор

$$h^3 = g^3 + \beta_2 h^2 =$$

$$= \begin{pmatrix} 0,108 \\ 0,063 \end{pmatrix} + 0,036296 \cdot \begin{pmatrix} -0,5873 \\ 1,0097 \end{pmatrix} = \begin{pmatrix} 0,08668 \\ 0,09965 \end{pmatrix}.$$

XI. Покладемо $k = 2 + 1 = 3$ і переходимо на крок III.

*4-а ітерація:*

III. Обчислюємо кроковий множник $\rho_3$:

$$\rho_3 = \frac{1,167\cdot 0,09965 + 0,776\cdot 0,08668 + 4\cdot 0,08668}{2\cdot(2\cdot 0,08668^2 + 2\cdot 0,09965^2 - 0,08668\cdot 0,09965)} +$$

$$+\frac{2\cdot 0,09965 - 4\cdot 1,167\cdot 0,08668 - 4\cdot 0,776\cdot 0,09965}{2\cdot(2\cdot 0,08668^2 + 2\cdot 0,09965^2 - 0,08668\cdot 0,09965)} =$$

$$= \frac{0,01563939}{0,052498855} = 0,297900.$$

IV. Обчислюємо наступне наближення

$$x^4 = x^3 + \rho_3 h^3 = \begin{pmatrix} 1,167 \\ 0,776 \end{pmatrix} + 0,2979\cdot\begin{pmatrix} 0,08668 \\ 0,09965 \end{pmatrix} = \begin{pmatrix} 1,1928 \\ 0,8057 \end{pmatrix}.$$

V. Обчислюємо градієнт

$$\nabla f_0(x^4) = \begin{pmatrix} -4 + 4\cdot 1,1928 - 0,8057 \\ -2 - 1,1928 + 4\cdot 0,8057 \end{pmatrix} = \begin{pmatrix} -0,0345 \\ 0,03 \end{pmatrix} \neq 0.$$

VI. Індекс $k = 3$ кратний 3, тому переходимо на крок VII.

VII. Покладемо $g^4 = (0,0345; -0,03)^T$; $h^4 = (0,0345; -0,03)^T$ і перейдемо на крок XI.

XI. Покладемо $k = 3 + 1 = 4$ і переходимо на крок III.

*5-а ітерація:*

III. Обчислюємо кроковий множник $\rho_4$:

$$\rho_4 = \frac{1,1928\cdot(-0,03) + 0,8057\cdot 0,0345 + 4\cdot 0,0345}{2\cdot(2\cdot(0,0345)^2 + 2\cdot(0,03)^2 + 0,0345\cdot 0,03)} +$$

$$+\frac{2\cdot(-0,03) - 4\cdot 1,1928\cdot 0,0345 - 4\cdot 0,8057\cdot(-0,03)}{2\cdot(2\cdot(0,0345)^2 + 2\cdot(0,03)^2 + 0,0345\cdot 0,03)} =$$

$$= \frac{0,00209025}{0,010431} = 0,200388.$$

IV. Обчислюємо наступне наближення

$$x^5 = x^4 + \rho_4 h^4 = \begin{pmatrix} 1,1928 \\ 0,8057 \end{pmatrix} + 0,200388\cdot\begin{pmatrix} 0,0345 \\ -0,03 \end{pmatrix} = \begin{pmatrix} 1,1997 \\ 0,7997 \end{pmatrix}.$$

V. Обчислюємо градієнт

$$\nabla f_0(x^5) = \begin{pmatrix} -4 + 4\cdot 1,1997 - 0,7997 \\ -2 - 1,1997 + 4\cdot 0,7997 \end{pmatrix} = \begin{pmatrix} -0,0009 \\ -0,0009 \end{pmatrix} \neq 0$$

і зупиняємо обчислення.

Таким чином, за 5 ітерацій отримали наближений розв'язок $x^* = (1,1997;\ 0,7997)^T$ і оптимальне значення функції

$$f_0(x^*) = -4 \cdot 1{,}1997 - 2 \cdot 0{,}7997 + 2 \cdot (1{,}1997)^2 -$$
$$-1{,}1997 \cdot 0{,}7997 + 2 \cdot (0{,}7997)^2 = -3{,}1999997.$$

**3. Модифікації методів спряжених градієнтів**

***Алгоритм 3***

Початок. I. Вибрати довільне початкове наближення $x^0 \in R^n$, константи $\varepsilon_0 > 0,\ \gamma_0 > 0,\ \beta \in (0;1),\ \beta' \in (0;1),\ \beta'' \in (0;1)$ (рекомендується $\varepsilon_0 = \cos 85^\circ,\ \gamma_0 = \cos 5^\circ,\ \beta = 0{,}6,\ \beta' = \beta'' = 0{,}8)$, натуральне число $\tau \geq n$ ($\tau$ – момент відновлення); покласти $k = 0$.

II. Обчислити $\nabla f_0(x^0)$.

III. Якщо $\nabla f_0(x^0) = 0$, то покласти $x^* = x^0$ і припинити обчислення; інакше покласти $g^0 = -\nabla f_0(x^0),\ h^0 = -\nabla f_0(x^0),\ \varepsilon = \varepsilon_0,\ \gamma = \gamma_0$ і перейти на крок IV.

Основний цикл. IV. Покласти $x = x^k$.

V. Обчислити вектор $h = h^k / \left\| h^k \right\|$.

VI. Визначити функцію $\theta : R^1 \to R^1$

$$\theta(\mu) = f_0(x + \mu h) - f_0(x).$$

VII. Покласти $\rho = 0$.

VIII. Обчислити скалярний добуток $(\nabla f_0(x + \rho h), h)$ і покласти $\varphi(\rho) = (\nabla f_0(x + \rho h), h)$.

IX. Якщо $\varphi(\rho) = 0$, то перейти на крок XVIII, інакше перейти на крок X.

X. Покласти $\lambda = 1$.

XI. Обчислити

$$\Delta = \theta(\rho - \lambda\varphi(\rho)) - \theta(\rho) + \left(1/2\right)\lambda\varphi^2(\rho).$$

XII. Якщо $\Delta \leq 0$, тоді покласти $\rho = \rho - \lambda\varphi(\rho)$ і перейти на крок XIII; інакше покласти $\lambda = \beta\lambda$ і перейти на крок XI.

XIII. Обчислити $\nabla f_0(x + \rho h)$.

XIV. Якщо $\nabla f_0(x + \rho h) = 0$, то покласти $x^* = x + \rho h$ і припинити обчислення; інакше перейти на крок XV.

XV. Обчислити

$$\varphi(\rho) = (\nabla f_0(x + \rho h), h).$$

XVI. Обчислити

$$\delta = \varphi(\rho) / \left\| \nabla f_0(x + \rho h) \right\|.$$

XVII. Якщо $\left| \delta \right| \leq \varepsilon$, то перейти на крок XXIII; інакше перейти на крок X.

XVIII. Обчислити наступне наближення $x^{k+1} = x + \rho h.$

XIX. Покласти $g^{k+1} = -\nabla f_0(x + \rho h)$.

XX. Обчислити коефіцієнт:

$$\beta_k = w\left(\frac{k+1}{\tau}\right)\frac{(g^{k+1} - g^k, g^k)}{\left\|g^k\right\|^2},$$

де

$$w(t) = \begin{cases} 0, \text{ якщо } \ t - \text{ціле число}, \\ 1, \text{ якщо } \ t - \text{неціле число}. \end{cases}$$

XXI. Обчислити вектор $h^{k+1} = g^{k+1} + \beta_k h^k$.

XXII. Покласти $k = k + 1$.

XXIII. Якщо виконується нерівність

$$(g^k, h^k) \geq \gamma \left\|g^k\right\| \left\|h^k\right\|,$$

то перейти на крок IV; інакше покласти $\varepsilon = \beta'\varepsilon$, $\gamma = \beta''\gamma$ і перейти на крок IV.

Для алгоритму 3 має місце теорема, аналогічна теоремі 2.

***Алгоритм 3′* (модифікація алгоритму 2, яка реалізується)**

Всі кроки алгоритму 2, окрім III, залишаються без змін. Крок III алгоритму 2 необхідно змінити кроком III′.

III′. Обчислити кроковий множник $\rho_k$, використовуючи для цього алгоритм 1Б чи алгоритм 2Б.

***Алгоритм 1Б* (алгоритм обчислення крокового множника $\rho_k$)**

I. Вибрати константи $\rho > 0$ і $\lambda \in (0;1/2)$.

II. Визначити функції:

$$\theta(\alpha, x^k) = f_0(x^k + \alpha h^k) - f_0(x^k) - \alpha(1-\lambda)(\nabla f_0(x^k), h^k);$$

$$\varphi(\alpha, x^k) = f_0(x^k + \alpha h^k) - f_0(x^k) - \alpha\lambda(\nabla f_0(x^k), h^k).$$

III. Покласти $\mu = \rho$.

IV. Обчислити значення $\theta(\mu, x^k)$.

V. Якщо $\theta(\mu, x^k) = 0$, то покласти $\rho_k = \mu$ і припинити обчислення; якщо $\theta(\mu, x^k) < 0$, то покласти $\mu = \mu + \rho$ і перейти на крок IV; якщо $\theta(\mu, x^k) > 0$, то перейти на крок VI.

VI. Обчислити значення $\varphi(\mu, x^k)$.

VII. Якщо $\varphi(\mu, x^k) \leq 0$, то покласти $\rho_k = \mu$ і припинити обчислення; інакше покласти $a_0 = \mu - \rho$, $b_0 = \mu$ і перейти на крок VIII.

VIII. Покласти $j = 0$.

IX. Обчислити $v_j = (a_j + b_j)/2$.

X. Обчислити значення $\theta(v_j, x^k)$, $\varphi(v_j, x^k)$.

XI. Якщо виконуються нерівності $\theta(v_j, x^k) \geq 0$ і $\varphi(v_j, x^k) \leq 0$, то покласти $\rho_k = v_j$ і припинити обчислення; інакше перейти на крок XII.

XII. Якщо $\theta(v_j, x^k) > 0$, то покласти $a_{j+1} = a_j, b_{j+1} = v_j$ і перейти на крок XIII; інакше покласти $a_{j+1} = v_j$, $b_{j+1} = b_j$ і перейти на крок XIII.

XIII. Покласти $j = j+1$ і перейти на крок IX.

***Алгоритм 2Б*** **(алгоритм обчислення крокового множника $\boldsymbol{\rho_к}$ за методом золотого перерізу)**

I. Вибрати довільні константи $\varepsilon_0 > 0,\ \alpha > 0,\ \rho > 0,\ \beta \in (0,1)$.

II. Обчислити дроби Фібоначі $F_1 = (3-\sqrt{5})/2$ і $F_2 = (\sqrt{5}-1)/2$.

III. Покласти $\varepsilon = \varepsilon_0$.

IV. Визначити функцію $\theta : R^1 \to R^1$

$$\theta(\mu) = f_0(x^k + \mu h^k) - f_0(x^k).$$

V. Обчислити значення $\theta(\rho)$ і $\theta(0)$.

VI. Якщо $\theta(\rho) > \theta(0)$, то покласти $a_0 = 0,\ b_0 = \rho$ і перейти на крок X; інакше покласти $i = 0,\ \mu_0 = 0$ і перейти на крок VII.

VII. Обчислити $\mu_{i+1} = \mu_i + \rho.$

VIII. Обчислити $\theta(\mu_{i+1})$.

IX. Якщо $\theta(\mu_{i+1}) \geq \theta(\mu_i)$, тоді покласти $a_0 = \mu_{i-1},\ b_0 = \mu_{i+1}$ і перейти на крок X; інакше покласти $i = i+1$ і перейти на крок VII.

X. Покласти $j = 0$.

XI. Обчислити $l_j = b_j - a_j$.

XII. Якщо $l_j \leq \varepsilon$, то перейти на крок XVII; інакше перейти на крок XIII.

XIII. Обчислити

$$v_j = a_j + F_1 l_j;\quad w_j = a_j + F_2 l_j.$$

XIV. Обчислити $\theta(v_j),\ \theta(w_j)$.

XV. Якщо $\theta(v_j) < \theta(w_j)$, то покласти $a_{j+1} = a_j,\ b_{j+1} = w_j$ і перейти на крок XVI; інакше покласти $a_{j+1} = v_j,\ b_{j+1} = b_j$ і перейти на крок XVI.

XVI. Покласти $j = j+1$ і перейти на крок XI.

XVI I. Покласти $\bar{\mu} = (a_j + b_j)/2.$

XVIII. Якщо $\theta(\bar{\mu}) \leq -\alpha\varepsilon$, тоді покласти $\rho_k = \bar{\mu}$ і припинити обчислення; інакше покласти $\varepsilon = \beta\varepsilon$ і перейти на крок V.

***Алгоритм 3′′* (реалізовна модифікація алгоритму 2′)**

Всі кроки алгоритму 2′, окрім кроку III, залишаються без змін. Крок III в алгоритмі 2′ необхідно замінити на крок III′′.

III′′. Обчислити множник $\rho_k$, використовуючи для цього алгоритм 1Б або алгоритм 2Б.

**4. Метод спряжених градієнтів для мінімізації квадратичних функцій**

З а д а ч а  4. Знайти $\arg\min\limits_{x \in R^n} f_0(x)$ для заданої квадратичної функції $f_0(x) \triangleq (Ax, x)/2 + (b, x) + \alpha$.

***Алгоритм 4***

П о ч а т о к. I. Вибрати довільне початкове наближення $x^0 \in R^n$ і покласти $k = 0$.

II. Обчислити $Ax^0 + b$ та покласти $h^0 = -(Ax^0 + b)$, $z^0 = -h^0$.

III. Якщо $h^0 = 0$, то покласти $x^* = x^0$ і припинити обчислення; інакше перейти на крок IV.

О с н о в н и й  ц и к л. IV. Якщо $k = 0$, то перейти на крок VII; інакше обчислити вектор $g^{k-1}$, $g^{k-1} = A(x^k - x^{k-1}) = \rho_{k-1} A h^{k-1}$ і перейти на крок V.

V. Обчислити коефіцієнт $\beta_k$ за однією з еквівалентних формул:

$$\beta_k = -\frac{(z^k, g^{k-1})}{(h^{k-1}, z^{k-1})}; \ \ \beta_k = -\frac{(z^k, z^k)}{(h^{k-1}, z^{k-1})}; \ \ \beta_k = \frac{(z^k, z^k)}{(z^{k-1}, z^{k-1})}.$$

VI. Обчислити вектор $h^k = -z^k + \beta_k h^{k-1}$.

VII. Обчислити кроковий множник $\rho_k$ за однією з еквівалентних формул:

$$\rho_k = -\frac{(z^0, h^k)}{(Ah^k, h^k)}; \ \ \rho_k = -\frac{(z^k, h^k)}{(Ah^k, h^k)}.$$

VIII. Обчислити наступне наближення

$$x^{k+1} = x^k + \rho_k h^k.$$

IX. Обчислити $Ax^{k+1} + b$ і покласти $z^{k+1} = Ax^{k+1} + b$.

X. Якщо $z^{k+1} = 0$, тоді покласти $x^* = x^{k+1}$ і припинити обчислення; інакше перейти на крок XI.

XI. Покласти $k = k + 1$ і перейти на крок IV.

***Теорема 4.*** *Якщо A – строго додатньо визначена симетрична матриця, то алгоритм 4 розв'язує задачу 4 за число ітерацій, що не перевищує n та* (***i***)*– для всіх i,* $0 \le i \le n-1$, *виконується нерівність* $(Ax^i + b, h^i) \ne 0$, *якщо* $Ax^i + b \ne 0$; (***ii***)*– точка* $x^i$ $(i = 1, \ldots, n-1)$ *є точкою* *мінімуму квадратичної функції* $f_0$ *на підпросторі, який утворений векторами* $h^0,...,h^{i-1}$ *і проходить через точку* $x^0$.

**5. Реалізовна модифікація методу зі змінною метрикою**

***Алгоритм 5***

П о ч а т о к. I. Вибрати довільне початкове наближення $x^0 \in R^n$, ціле число $\alpha$ $(1 \le \alpha \le 10)$, константу $\beta \in (1/2;4/5)$ і покласти $k=0,\ \ j=0$.

II. Обчислити $\nabla f_0(x^0)$. Якщо $\nabla f_0(x^0)=0$, то покласти $x^*=x^0$ і припинити обчислення; інакше перейти на крок III.

III. Покласти $H_0 = I$ ($I$ – одинична $n \times n$-матриця), $g_0 = \nabla f_0(x^0)$.

О с н о в н и й ц и к л. IV. Обчислити вектор $h^k = H_k g^k$.

V. Якщо число $k$ кратне $\alpha$, тоді покласти $j = j+1,\ \ \alpha = 3\alpha$ і перейти на крок VI; інакше перейти на крок VI.

VI. Покласти $x = x^k, h = h^k, i = 0$.

VII. Визначити функцію $\theta : R^1 \to R^1$

$$\theta(\alpha) = f_0(x+\alpha h) - f_0(x).$$

VIII. Покласти $\rho = 0$.

IX. Обчислити $\varphi(\rho) = (\nabla f_0(x+\rho h), h)$.

X. Якщо $\varphi(\rho)=0$, то покласти $\rho_k = \rho$ і перейти на крок XV; інакше перейти на крок XI.

XI. Покласти $\lambda = 1$.

XII. Обчислити $\Delta = \theta(\rho - \lambda\varphi(\rho)) - \theta(\rho) + \lambda\varphi^2(\rho)/2$.

XIII. Якщо $\Delta \le 0$, то покласти $i = i+1$ і перейти на крок XVI; інакше покласти $\lambda = \beta\lambda$ і перейти на крок XII.

XIV. Якщо $i < j$, то покласти $\rho = \rho - \lambda\varphi(\rho)$ і перейти на крок IX; інакше покласти $\rho_k = \rho - \lambda\varphi(\rho)$ і перейти на крок XV.

XV. Обчислити наступне наближення $x^{k+1} = x + \rho_k h$.

XVI. Обчислити $\nabla f_0(x^{k+1})$ і покласти $g^{k+1} = \nabla f_0(x^{k+1})$.

XVII. Якщо $g^{k+1} = 0$, то покласти $x^* = x^{k+1}$ і припинити обчислення; інакше перейти на крок XVIII.

XVIII. Обчислити вектори:

$$r^k = g^{k+1} - g^k;\ \ z^k = x^{k+1} - x^k.$$

XIX. Обчислити матрицю

$$H_{k+1} = H_k - \frac{H_k r^k (H_k r^k)^T}{(H_k r^k, r^k)} + \frac{z^k (z^k)^T}{(r^k, z^k)}.$$

XX. Покласти $k = k+1$ і перейти на крок IV.

## 3.5. Методи спряжених напрямків

З а д а ч а  1. Знайти $\arg\min_{x \in R^n} f_0(x)$ для заданої неперервно диференційованої функції $f_0 : R^n \to R^1$.

*Припущення 1.* Функція $f_0 -$ двічі неперервно диференційована та її матриця других похідних $\nabla^2_{xx} f_0(x)$ задовольняє нерівності:

$$\gamma_1 \|y\|^2 \le \left(\nabla^2_{xx} f_0(x) y, y\right) \le \gamma_2 \|y\|^2; \quad \gamma_2 \ge \gamma_1 > 0,$$

для будь-яких $x, y \in R^n$.

*Означення 1.* Вектори $p^0, p^1, ..., p^{n-1}$ називаються ***спряженими*** або $\boldsymbol{A}$ – ***ортогональними***, якщо $\left(p^i, Ap^j\right) = 0$ при $i \ne j$, де $A -$ строго додатньо визначена матриця.

#### 1. Метод спряжених напрямків із відновленням матриці

***Алгоритм 1***

П о ч а т о к . I. Вибрати довільне початкове наближення $x^0 \in R^n$, довільну симетричну строго додатньо визначену матрицю $H_0$, яка задовольняє умову:

$$\gamma_3 \|y\|^2 \le (H_0 y, y) \le \gamma_4 \|y\|^2, \quad \gamma_4 \ge \gamma_3 > 0, \ \forall y \in R^n,$$

зокрема, можна вибрати $H_0 = I$, де $I -$ одинична $n \times n$-матриця); покласти $k = 0$.

О с н о в н и й  ц и к л . II. Обчислити $\nabla f_0\left(x^k\right)$.

III. Якщо $\nabla f_0\left(x^k\right) = 0$, то покласти $x^* = x^k$ і закінчити обчислення; інакше перейти на крок IV.

IV. Якщо $k = 0$, то перейти на крок VIII; інакше перейти на крок V.

V. Якщо $k$ ділиться без остачі на $n$, то покласти $H_k = H_0$ і перейти на крок VIII; інакше перейти на крок VI.

VI. Обчислити вектори:

$$r^{k-1} = x^k - x^{k-1} = \rho_{k-1} h^{k-1};$$

$$g^{k-1} = \nabla f_0\left(x^k\right) - \nabla f_0\left(x^{k-1}\right).$$

VII. Обчислити матрицю $H_k$ за однією із наведених нижче формул:

$$H_k = H_{k-1} + \frac{r^{k-1}\left(r^{k-1}\right)^T}{\left(r^{k-1}, g^{k-1}\right)} - \frac{H_{k-1} g^{k-1}\left(g^{k-1}\right)^T H_{k-1}}{\left(H^T_{k-1} g^{k-1}, g^{k-1}\right)}; \tag{3.4}$$

$$H_k = H_{k-1} + \left(r^{k-1} - H_{k-1} g^{k-1}\right) \frac{\left(r^{k-1}\right)^T}{\left(r^{k-1}, g^{k-1}\right)}; \tag{3.5}$$

$$H_k = H_{k-1} + \left(r^{k-1} - H_0 g^{k-1}\right) \frac{\left(r^{k-1}\right)^T}{\left(r^{k-1}, g^{k-1}\right)}; \tag{3.6}$$

$$H_k = H_{k-1} - \frac{H_{k-1} g^{k-1} \left(r^{k-1}\right)^T}{\left(r^{k-1}, g^{k-1}\right)}; \tag{3.7}$$

$$H_k = H_{k-1} - \frac{H_0 g^{k-1} \left(r^{k-1}\right)^T}{\left(r^{k-1}, g^{k-1}\right)}; \tag{3.8}$$

$$H_k = H_0 - \frac{H_0 \nabla f_0(x^k) \left(H_0 \nabla f_0(x^k) + h^{k-1}\right)^T}{\left(H_0 \nabla f_0(x^k), \nabla f_0(x^k)\right) - \left(h^{k-1}, \nabla f_0(x^{k-1})\right)}; \tag{3.9}$$

$$H_k = H_0 + \frac{H_0 g^{k-1} \left(h^{k-1}\right)^T}{\left(h^{k-1}, \nabla f_0\left(x^{k-1}\right)\right)}; \tag{3.10}$$

$$H_k = H_0 + \frac{H_0 \nabla f_0\left(x^k\right) \left(h^{k-1}\right)^T}{\left(h^{k-1}, \nabla f_0\left(x^{k-1}\right)\right)}. \tag{3.11}$$

VIII. Обчислити вектор руху $h^k$ до наступного наближення $x^{k+1}$

$$h^k = -H_k^T \nabla f_0\left(x^k\right).$$

IX. Обчислити кроковий множник $\rho_k$ із умови

$$f_0\left(x^k + \rho_k h^k\right) = \min_{\rho \geq 0} f_0\left(x^k + \rho h^k\right).$$

X. Обчислити наступне наближення

$$x^{k+1} = x^k + \rho_k h^k.$$

XI. Покласти $k = k + 1$ і перейти на крок II.

***Теорема 1.*** *Якщо виконані припущення 1, то нескінченна послідовність $\left\{x^k\right\}_{k=0}^{\infty}$, яка породжена алгоритмом 1 (де матриця $H_k$ обчислюється за однією із формул (3.4) –(3.11)), збігається до розв'язку $x^*$ із надлінійною швидкістю:*

$$\left\|x^{(i+1)n} - x^*\right\| \leq \lambda_{in} \left\|x^{in} - x^*\right\|, \quad i = 0, 1, \ldots,$$

*де $\lambda_{in} \to 0$ при $i \to \infty$.*

*Зауваження 1.* Якщо на якійсь ітерації початкової стадії процесу мінімізації за алгоритмом 1 вектор $h^k = 0$, то необхідно почати процес спочатку, відновивши матрицю $H_0$.

**2. Метод спряжених напрямків без відновлення матриці**

***Алгоритм 2***

П о ч а т о к . I. Вибрати довільне початкове наближення $x^0 \in R^n$, довільну симетричну строго додатньо визначену матрицю $H_0$, яка задовольняє умову:

$$\gamma_3 \|y\|^2 \le (H_0 y, y) \le \gamma_4 \|y\|^2, \quad \gamma_4 \ge \gamma_3 > 0, \ \forall y \in R^n,$$

зокрема, можна вибрати $H_0 = I$, де $I -$ одинична $n \times n$-матриця); покласти $k = 0$.

О с н о в н и й ц и к л. II. Обчислити $\nabla f_0(x^k)$; якщо $\nabla f_0(x^k) = 0$, то покласти $x^* = x^k$ і закінчити обчислення; інакше перейти на крок III.

III. Якщо $k = 0$, то перейти на крок VI; інакше перейти на крок IV.

IV. Обчислити вектори:

$$r^{k-1} = x^k - x^{k-1} = \rho_{k-1} h^{k-1};$$

$$g^{k-1} = \nabla f_0\left(x^k\right) - \nabla f_0\left(x^{k-1}\right).$$

V. Обчислити матрицю $H_k$ за однією із наведених нижче формул:

$$H_k = H_{k-1} + \frac{r^{k-1}\left(r^{k-1}\right)^T}{\left(r^{k-1}, g^{k-1}\right)} - \frac{H_{k-1} g^{k-1}\left(g^{k-1}\right)^T H_{k-1}}{\left(H_{k-1}^T g^{k-1}, g^{k-1}\right)}; \tag{3.12}$$

$$H_k = H_{k-1} + \left(r^{k-1} - H_{k-1} g^{k-1}\right) \frac{\left(g^{k-1}\right)^T}{\left(r^{k-1}, g^{k-1}\right)}; \tag{3.13}$$

$$H_k = H_{k-1} + \left(r^{k-1} - H_0 g^{k-1}\right) \frac{\left(g^{k-1}\right)^T}{\left(r^{k-1}, g^{k-1}\right)}; \tag{3.14}$$

$$H_k = H_0 + \frac{H_0 g^{k-1}\left(h^{k-1}\right)^T}{\left(h^{k-1}, \nabla f_0\left(x^{k-1}\right)\right)}. \tag{3.15}$$

VI. Обчислити вектор руху $h^k$ до наступного наближення $x^{k+1}$

$$h^k = -H_k^T \nabla f_0\left(x^k\right).$$

VII. Обчислити кроковий множник $\rho_k$ із умови

$$f_0\left(x^k + \rho_k h^k\right) = \min_{\rho \ge 0} f_0\left(x^k + \rho h^k\right).$$

VIII. Обчислити наступне наближення

$$x^{k+1} = x^k + \rho_k h^k.$$

IX. Покласти $k = k + 1$ і перейти на крок II.

*Теорема 2. Нехай виконується припущення 1. Тоді:*

1) *якщо перші* $m$ *точок* ($m<\infty$ *і залежить від точки* $x^0$) *нескінченної послідовності* $\{x^k\}_{k=0}^{\infty}$ *обчислені за допомогою методу спряжених напрямків із відновленням матриці* $H_k$ *(тобто за допомогою алгоритму 1), а інші точки цієї послідовності – за допомогою методу спряжених напрямків без відновлення матриці (тобто за допомогою алгоритму 2), то послідовність* $\{x^k\}_{k=0}^{\infty}$ *збігається до розв'язку* $x^*$ *при довільному виборі початкового наближення* $x^0$;

2) *якщо початкове наближення* $x^0$ *вибрано в достатньо малому околі точки мінімуму* $x^*$, *то нескінченна послідовність* $\{x^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 2, збігається до розв'язку* $x^*$;

3) *нескінченна послідовність* $\{x^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 2, де матриця* $H_k$ *обчислюється за формулою (3.15), незалежно від вибору початкового наближення* $x^0$ *збігається до розв'язку* $x^*$ *із надлінійною швидкістю:*

$$\left\|x^{(i+1)n}-x^*\right\|\le\lambda_{in}\left\|x^{in}-x^*\right\|,\quad i=0,1,\ldots,$$

*де* $\lambda_{in}\to 0$ *при* $i\to\infty$.

*Зауваження 2.* Із оцінок швидкості збіжності методів спряжених напрямків випливає, що $n$ ітерацій методу спряжених напрямків еквівалентні (в розумінні швидкості збіжності) одній ітерації методів типу Ньютона–Канторовича. Однак вони значно переважають по швидкості збіжності градієнтні методи.

З теоретичної точки зору методи спряжених напрямків без відновлення матриці краще використовувати в тому випадку, коли $H_{in}\to\left(\nabla_{xx}^2 f_0\left(x^{in}\right)\right)^{-1}$ при $i\to\infty$.

Рекомендується використовувати алгоритм 1 із відновленням матриці, якщо матриця $H_k$ обчислюється за формулами (3.7)-(3.11) , а алгоритм 2 без відновлення матриці, коли $H_k$ обчислюється за формулами (3.4)-(3.6).

### 3. Мінімізація квадратичних функцій за допомогою методу спряжених напрямків

Задача 3. Знайти $\arg\min_{x\in R^n} f_0(x)$ для заданої квадратичної функції

$$f_0(x)\triangleq(Ax,x)/2+(b,x)+\alpha .$$

***Алгоритм 3***

Початок. I. Вибрати довільне початкове наближення $x^0 \in R^n$, довільну симетричну строго додатньо визначену матрицю $H_0$ (зокрема можна вибрати $H_0 = I$, де $I-$ одинична $n \times n$-матриця); покласти $k = 0$.

Основний цикл. II. Обчислити $Ax^k + b$ і покласти

$$z^k = Ax^k + b.$$

Якщо $z^k = 0$, то покласти $x^* = x^k$ і завершити обчислення; інакше перейти на крок III.

III. Якщо $k = 0$, то перейти на крок VI; інакше перейти на крок IV.

IV. Обчислити вектори:

$$r^{k-1} = x^k - x^{k-1} = \rho_{k-1} h^{k-1}; \;\; g^{k-1} = Ar^{k-1}.$$

V. Обчислити матрицю $H_k$ за будь-якою із формул (3.4)-(3.11).

VI. Обчислити вектор руху $h^k$ до наступного наближення $x^{k+1}$

$$h^k = -H_k^T z^k.$$

VII. Обчислити кроковий множник $\rho_k$ за однією із формул:

$$\rho_k = -\frac{\left(z^0, h^k\right)}{\left(Ah^k, h^k\right)}; \qquad \rho_k = -\frac{\left(z^k, h^k\right)}{\left(Ah^k, h^k\right)}.$$

VIII. Обчислити наступне наближення

$$x^{k+1} = x^k + \rho_k h^k.$$

IX. Покласти $k = k + 1$ і перейти на крок II.

***Теорема 3.*** *Якщо A – додатньо визначена симетрична матриця із постійними членами (тобто* $\left(Ax, x\right) > 0$ *при* $\forall x \neq 0$*), то алгоритм 3 розв'язує задачу 3 за число ітерацій, яке не перевищує* n*, та*

$(\boldsymbol{i})-$ *точки* $x^0, x^1, ..., x^n$*, які породжені алгоритмом 3 при різних способах обчислення матриць* $H_k$*, одні й ті ж;* $(\boldsymbol{ii})-$ *для всіх* $i$, $0 \le i \le n-1$ *виконується умова* $\left(Ax^i + b, h^i\right) \neq 0$, *якщо* $Ax^i + b \neq 0$; $(\boldsymbol{iii})-$ *точка* $x^i$, $(i = \overline{1, n-1})$ *є точкою мінімуму квадратичної функції* $f_0$ *на підпросторі, який утворений векторами* $h^0, h^1, ..., h^{i-1}$ *і який проходить через точку* $x^0$; $(\boldsymbol{iv})-$ *якщо в алгоритмі 3 матрицю* $H_k$ *обчислювати:*

*за формулами (3.4)-(3.6), то* $H_n = A^{-1}$;

*за формулами (3.9), (3.11), то* $H_n = H_0$;

*за формулою (3.7), то* $H_n = 0$;

*за формулою (3.10), то* $H_n \neq H_0$.

Нижче наведений метод спряжених напрямків, в якому безпосередньо знаходять вектор руху $h^k$ без обчислення матриці $H_k$.

***Алгоритм 3'***

Початок. I. Вибрати довільне початкове наближення $x^0 \in R^n$, довільну симетричну додатньо визначену матрицю $H_0$ (зокрема, можна вибрати $H_0 = I$, де $I-$ одинична $n \times n$-матриця); покласти $k = 0$.

Основний цикл. II. Обчислити $Ax^k + b$ і покласти

$$z^k = Ax^k + b.$$

III. Якщо $z^k = 0$, то покласти $x^* = x^k$ і завершити обчислення; інакше перейти на крок IV.

IV. Якщо $k = 0$, то обчислити вектор $h^0$ за формулою $h^0 = -H_0 z^0$ і перейти на крок V; інакше обчислити вектор $h^k$ за будь-яким із алгоритмів 1A, 2A, 3A, 4A і перейти на крок V.

V. Обчислити кроковий множник $\rho_k$ за однією із формул:

$$\rho_k = -\frac{\left(z^0, h^k\right)}{\left(Ah^k, h^k\right)}; \qquad \rho_k = -\frac{\left(z^k, h^k\right)}{\left(Ah^k, h^k\right)}.$$

VI. Обчислити наступне наближення

$$x^{k+1} = x^k + \rho_k h^k.$$

IX. Покласти $k = k + 1$ і перейти на крок II.

***Алгоритм 1A*** (алгоритм обчислення вектора $h^k$)

I. Обчислити вектор $g^{k-1} = z^k - z^{k-1}$.

II. Обчислити коефіцієнт $\beta_k$ за однією із еквівалентних формул:

$$\beta_k = 1 - \frac{\left(H_0 z^k, g^{k-1}\right)}{\left(H_0 z^k, z^k\right) - \left(h^{k-1}, z^{k-1}\right)};$$

$$\beta_k = -\frac{\left(h^{k-1}, z^{k-1}\right)}{\left(H_0 z^k, z^k\right) - \left(h^{k-1}, z^{k-1}\right)}.$$

III. Обчислити вектор

$$h^k = -\beta_k H_0 z^k + \left(1 - \beta_k\right) h^{k-1}.$$

***Алгоритм 2A*** (алгоритм обчислення вектора $h^k$)

I. Обчислити коефіцієнт

$$\beta_k = \frac{\left(H_0 z^k, z^k\right)}{\left(H_0 z^k, z^k\right) - \left(h^{k-1}, z^{k-1}\right)}.$$

II. Обчислити вектор

$$h^k = -H_0 z^k + \beta_k \left(H_0 z^k + h^{k-1}\right).$$

***Алгоритм 3А*** (алгоритм обчислення вектора $h^k$)

I. Обчислити вектор $g^{k-1} = z^k - z^{k-1}$.

II. Обчислити коефіцієнт $\beta_k$ за однією із формул:

$$\beta_k = \frac{\left(H_0 z^k, g^{k-1}\right)}{\left(h^{k-1}, g^{k-1}\right)}; \quad \beta_k = -\frac{\left(H_0 z^k, g^{k-1}\right)}{\left(h^{k-1}, z^{k-1}\right)}.$$

III. Обчислити вектор

$$h^k = -H_0 z^k + \beta_k h^{k-1}.$$

***Алгоритм 4А* (алгоритм обчислення вектора $h^k$)**

I. Обчислити коефіцієнт $\beta_k$ за однією із еквівалентних формул:

$$\beta_k = -\frac{\left(H_0 z^k, z^k\right)}{\left(h^{k-1}, z^{k-1}\right)}; \quad \beta_k = \frac{\left(H_0 z^k, z^k\right)}{\left(H_0 z^{k-1}, z^{k-1}\right)}.$$

II. Обчислити вектор

$$h^k = -H_0 z^k + \beta_k h^{k-1}.$$

***Теорема 3′.*** *Якщо А – додатньо визначена симетрична матриця із постійними членами (тобто $(Ax, x) > 0$ при $\forall x \neq 0$), то алгоритм 3′ розв'язує задачу 3 за число ітерацій, яке не перевищує $n$, і для послідовності $x^0, x^1, ..., x^n$, яка породжена алгоритмом 3′ справедливі твердження (**i**), (**ii**), (**iii**) теореми 3.*

Методи спряжених напрямків можуть бути використані і для мінімізації опуклої квадратичної функції:

$$f_0(x) \underline{\underline{\Delta}} (Ax, x)/2 + (b, x) + \alpha,$$

де $(Ax, x) \geq 0$ при $x \neq 0$. При цьому слід враховувати два випадки.

1. Мінімум квадратичної функції $f_0$ існує. Тоді алгоритм 3 і алгоритм 3′ розв'язує задачу мінімізації $f_0$ за число ітерацій, яке не перевищує $n$.

2. Функція $f_0$ не досягає мінімуму. Тоді при деякому $k$, $0 \leq k \leq n-1$ буде виконуватися умова $\left(Ah^k, h^k\right) = 0$, що тягне $\rho_k = \infty$.

**Приклад 1.** Методом спряжених напрямків розв'язати задачу безумовної оптимізації

$$f_0(x) = x_1^{\ 2} + 2x_2^{\ 2} - 2x_1 x_2 - 2x_1 - 4x_2 \to \min.$$

*Розв'язування.* Використаємо алгоритм 2, в якому матриця $H_k$ на кроці V обчислюється за формулою (3.15).

Кроковий множник $\rho_k$ на кроці VII знайдемо аналітично:

$$f_0(x^k + \rho h^k) = (x_1^{\ k} + \rho h_1^{\ k})^2 + 2(x_2^{\ k} + \rho h_2^{\ k})^2 - 2(x_1^{\ k} + \rho h_1^{\ k}) \cdot (x_2^{\ k} + \rho h_2^{\ k}) -$$
$$-2(x_1^{\ k} + \rho h_1^{\ k}) - 4(x_2^{\ k} + \rho h_2^{k}) = (x_1^{\ k})^2 + \rho^2 (h_1^{\ k})^2 + 2\rho x_1^{\ k} h_1^{\ k} + 2(x_2^{\ k})^2 +$$
$$+4\rho x_2^{\ k} h_2^{\ k} + 2\rho^2 (h_2^{\ k})^2 - 2x_1^{\ k} x_2^{\ k} - 2\rho x_1^{\ k} h_2^{\ k} - 2\rho x_2^{\ k} h_1^{\ k} - 2\rho^2 h_1^{\ k} h_2^{\ k} -$$

$$-2x_1^k - 2\rho h_1^k - 4x_2^k - 4\rho h_2^k = \rho^2[(h_1^k)^2 + 2(h_2^k)^2 - 2h_1^k h_2^k] -$$
$$-2\rho[2h_2^k + h_1^k + x_1^k h_2^k + x_2^k h_1^k - 2x_2^k h_2^k - x_1^k h_1^k] +$$
$$+(x_1^k)^2 + 2(x_2^k)^2 - 2x_1^k x_2^k - 2x_1^k - 4x_2^k.$$

Мінімум цієї функції по $\rho$ досягається у вершині параболи

$$\rho_k = \frac{2h_2^k + h_1^k + x_2^k h_1^k + x_1^k h_2^k - 2x_2^k h_2^k - x_1^k h_1^k}{(h_1^k)^2 + 2(h_2^k)^2 - 2h_1^k h_2^k}.$$

Знайдемо градієнт функції $f_0$ : $\nabla f_0(x) = \begin{pmatrix} 2x_1 - 2x_2 - 2 \\ 4x_2 - 2x_1 - 4 \end{pmatrix}$.

**Алгоритм 2**

I. Вибираємо початкове наближення $x^0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $H_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$;

покладемо $k = 0$.

*1-а ітерація:*

II. Обчислюємо градієнт $\nabla f_0(x^0) = \begin{pmatrix} -2 \\ -4 \end{pmatrix} \neq 0.$

III. Оскільки $k = 0$, то переходимо на крок VI.

VI. Обчислюємо вектор $h^0 = -I \cdot \nabla f_0(x^0) = (2;4)^T$.

VII. Обчислюємо кроковий множник

$$\rho_0 = \frac{2\cdot 4 + 2 + 0 + 0 - 0 - 0}{4 + 2\cdot 16 - 2\cdot 2\cdot 4} = \frac{10}{20} = 0{,}5.$$

VIII. Обчислюємо наступне наближення

$$x^1 = x^0 + \rho_0 h^0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + 0{,}5\cdot\begin{pmatrix} 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

IX. Покладемо $k = 0 + 1 = 1$ і ми переходимо на крок II.

*2-а ітерація:*

II. Обчислюємо градієнт

$$\nabla f_0(x^1) = \begin{pmatrix} 2\cdot 1 - 2\cdot 2 - 2 \\ 4\cdot 2 - 2\cdot 1 - 4 \end{pmatrix} = \begin{pmatrix} -4 \\ 2 \end{pmatrix} \neq 0.$$

III. Оскільки $k \neq 0$, то переходимо на крок IV.

IV. Обчислюємо вектори:

$$r^0 = x^1 - x^0 = \rho_0 h^0 = 0{,}5\cdot\begin{pmatrix} 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix},$$

$$g^0 = \nabla f_0(x^1) - \nabla f_0(x^0) = \begin{pmatrix} -4 \\ 2 \end{pmatrix} - \begin{pmatrix} -2 \\ -4 \end{pmatrix} = \begin{pmatrix} -2 \\ 6 \end{pmatrix}.$$

V. Обчислюємо матрицю $H_1$ за формулою (3.15):

$$H_1 = H_0 + \frac{H_0 g^0 (h^0)^T}{(h^0, \nabla f_0(x^0))} =$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{\begin{pmatrix} -2 \\ 6 \end{pmatrix} \cdot (2;4)}{\left( \begin{pmatrix} 2 \\ 4 \end{pmatrix}, \begin{pmatrix} -2 \\ -4 \end{pmatrix} \right)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{1}{-20} \begin{pmatrix} -4 & -8 \\ 12 & 24 \end{pmatrix} =$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1/5 & 2/5 \\ -3/5 & -6/5 \end{pmatrix} = \begin{pmatrix} 6/5 & 2/5 \\ -3/5 & -1/5 \end{pmatrix}.$$

VI. Обчислюємо вектор

$$h^1 = -H_1^T \cdot \nabla f_0(x^1) = \begin{pmatrix} 6/5 & 2/5 \\ -3/5 & -1/5 \end{pmatrix} \cdot \begin{pmatrix} -4 \\ 2 \end{pmatrix} = \begin{pmatrix} 6 \\ 2 \end{pmatrix}.$$

VII. Обчислюємо кроковий множник

$$\rho_1 = \frac{2 \cdot 2 + 6 + 2 \cdot 6 + 2 \cdot 1 - 2 \cdot 2 \cdot 2 - 1 \cdot 6}{36 + 2 \cdot 4 - 2 \cdot 6 \cdot 2} = \frac{10 + 12 + 2 - 8 - 6}{36 + 8 - 24} = 0,5.$$

VIII. Обчислюємо наступне наближення

$$x^2 = x^1 + \rho_1 h^1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix} + 0,5 \cdot \begin{pmatrix} 6 \\ 2 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \end{pmatrix}.$$

IX. Покладемо $k = 1 + 1 = 2$ і переходимо на крок II.

*3-я ітерація:*

II. Обчислюємо градієнт

$$\nabla f_0(x^2) = \begin{pmatrix} 2 \cdot 4 - 2 \cdot 3 - 2 \\ 4 \cdot 3 - 2 \cdot 4 - 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0,$$

тому покладемо $x^* = x^2 = (4;3)^T$ і зупиняємо обчислення.

Таким чином, за дві ітерації методу спряжених напрямків отримали оптимальний розв'язок $x^* = x^2 = (4;3)^T$ і оптимальне значення цільової функції

$$f_0(x^*) = 4^2 + 2 \cdot 3^2 - 2 \cdot 4 \cdot 3 - 2 \cdot 4 - 4 \cdot 3 = -10.$$

#### 4. Модифікований метод спряжених напрямків, який не потребує обчислення похідних

***Алгоритм 4***

Початок. I. Вибрати: початкове наближення $x^{1,0} \in R^n$, яке задовольняє умовам теореми 4; ортонормований координатний базис

$h^{1,1}, h^{1,2}, ..., h^{1,n}$; довільну константу $\varepsilon \in (0;1)$ (доцільно вибрати $\varepsilon \in \left[10^{-10}; 10^{-3}\right]$).

II. Покласти $k = 1$.

О с н о в н и й ц и к л. III. Покласти $i = 1$.

IV. Обчислити кроковий множник $\rho_{k,i}$ із умови

$$f_0\left(x^{k,i-1} + \rho_{k,i} h^{k,i}\right) = \min_{\rho} f_0\left(x^{k,i-1} + \rho h^{k,i}\right).$$

V. Обчислити наступне наближення

$$x^{k,i} = x^{k,i-1} + \rho_{k,i} h^{k,i}.$$

VI. Якщо $i < n$, то покласти $i = i + 1$ і перейти на крок IV; інакше перейти на крок VII.

VII. Обчислити $\gamma_k = \left\| x^{k,n} - x^{k,0} \right\|$.

VIII. Обчислити вектор

$$h^{k,n+1} = \left(x^{k,n} - x^{k,0}\right) \big/ \gamma_k .$$

IX. Обчислити кроковий множник $\rho_{k,n+1}$ із умови

$$f_0\left(x^{k,n} + \rho_{k,n+1} h^{k,n+1}\right) = \min_{\rho} f_0\left(x^{k,n} + \rho h^{k,n+1}\right).$$

X. Обчислити наступне наближення

$$x^{k,n+1} = x^{k,n} + \rho_{k,n+1} h^{k,n+1}.$$

XI. Знайти індекс $s \in [1:n]$, який задовольняє умові

$$\rho_{k,s} = \max\left\{\rho_{k,1}, \rho_{k,2}, ..., \rho_{k,n}\right\}.$$

XII. Якщо $k = 1$, то обчислити визначник $\Delta_k$, стовпцями якого є вектори $h^{k,1}, h^{k,2}, ..., h^{k,n}$; інакше перейти на крок XIII.

XIII. Якщо $(\rho_{k,s} \Delta_k / \gamma_k) \geq \varepsilon$, то перейти на крок XIV; інакше перейти на крок XVIII.

XIV. Покласти $j = 1$.

XV. Якщо $j = s$, то покласти $h^{k+1,j} = h^{k,n+1}$ і перейти на крок XVI; інакше покласти $h^{k+1,j} = h^{k,j}$ і перейти на крок XVI.

XVI. Якщо $j < n$, то покласти $j = j + 1$ і перейти на крок XV; інакше перейти на крок XVII.

XVII. Обчислити $\Delta_{k+1} = \rho_{k,s} \Delta_k \big/ \gamma_k$ і перейти на крок XXII.

XVIII. Покласти $j = 1$.

XIX. Покласти $h^{k+1,j} = h^{k,j}$.

XX. Якщо $j < n$, то покласти $j = j + 1$ і перейти на крок XIX; інакше перейти на крок XXI.

XXI. Покласти $\Delta_{k+1} = \Delta_k$ і перейти на крок XXII.

XXII. Покласти $x^{k+1,0} = x^{k,n+1}$.

XXIII. Покласти $k = k+1$ і перейти на крок III.

***Теорема 4.*** *Якщо виконані умови:* $(\boldsymbol{i})-$ *функція* $f_0$ *– неперервно диференційована;* $(\boldsymbol{ii})-$ *функція* $f_0-$ *строго опукла;* $(\boldsymbol{iii})-$ *початкове наближення* $x^{1,0}$ *в алгоритмі 4 таке, що множина*

$$X_0 = \left\{ x \mid f_0\left(x\right) \le f_0\left(x^{1,0}\right),\ x \in R^n \right\}$$

*обмежена, то послідовність* $x^{k,i}$, $i = \overline{0,n}$; $k = 1,2,...,$ *яка породжена алгоритмом 4, збігається до єдиної точки мінімуму* $x^*$ *функції* $f_0$.

## 3.6. Методи псевдообернених операторів

З а д а ч а 0. Знайти $\arg\min\limits_{x \in R^n} f_0\left(x\right)$ для заданої функції $f_0 : R^n \to R^1$.

*Припущення 0.* $(\boldsymbol{i})-$ функція $f_0\left(x\right)-$ двічі неперервно диференційована по Фреше; $(\boldsymbol{ii})-$ градієнт $\nabla f_0(x)$ функції $f_0(x)$ задовольняє в $R^n$ умові Ліпшиця з константою $\gamma_1 \ge 0$; $(\boldsymbol{iii})-$ матриця других похідних $\Delta^2_{xx} f_0(x)$ (гесіан) функції $f_0(x)$ задовольняє в $R^n$ умові Ліпшиця з константою $\gamma_2 \ge 0$.

Методи псевдообернених операторів по суті є квазіньютоновськими. Швидкість збіжності цих методів – надлінійна. На $k$-й ітерації за напрямок руху до наступного наближення $x^{k+1}$ обираємо вектор $\left(-B_k \nabla f_0\left(x^k\right)\right)$, де $B_k$ – матриця, що апроксимує обернений гесіан функції $f_0(x)$ в точці $x^k$. Обчислення матриці $B_k$ основане на алгоритмах стійкого псевдообернення прямокутних матриць. Перевагою методів псевдообернених операторів є те, що вони не потребують обчислення матриці других похідних функції $f_0(x)$, стійкі до збурень, які виникають при похибках апроксимації і дискретизації обчислень, володіють надлінійною швидкістю збіжності і тоді, коли гесіан вироджений.

#### 1. Основний метод

***Алгоритм 1***

П о ч а т о к . I. Вибрати довільне початкове наближення $x^0 \in R^n$.

II. Вибрати значення параметра $m$, $m \ge n$.

III. Покласти $k = 0$.

О с н о в н и й ц и к л . IV. Обчислити $\nabla f_0\left(x^k\right)$.

V. Якщо $k \geq m$, то перейти на крок VI; інакше покласти $x^{k+1} = x^k - \nabla f_0\left(x^k\right)$ і перейти на крок XI.

VI. Сформувати додаткову $n \times m$-матрицю $A_{k,m}$, $j$-м стовпцем якої є вектор:

$$a^j = x^{k-(m-j)} - x^{k-(m-j+1)}, \quad j = \overline{1,m}.$$

VII. Сформувати допоміжну $n \times m$-матрицю $H_{k,m}$, $j$-м стовпцем якої є вектор:

$$h^j = \nabla f_0\left(x^{k-(m-j)}\right) - \nabla f_0\left(x^{k-(m-j+1)}\right), \quad j = \overline{1,m}.$$

VIII. Використовуючи алгоритм $1'$ стійкого псевдообернення скінченновимірних прямокутних матриць, обчислити псевдообернені матриці $A_{k,m}^{+}$ і $\left(H_{k,m} A_{k,m}^{+}\right)^{+}$, де символ $C^{+}$ позначає матрицю, псевдообернену до $C$.

IX. Обчислити матрицю $B_k$, яка апроксимує обернений гесіан

$$B_k = \left(H_{k,m} A_{k,m}^{+}\right)^{+} + I - H_{k,m} A_{k,m}^{+} \left(H_{k,m} A_{k,m}^{+}\right)^{+}.$$

X. Обчислити наступне наближення

$$x^{k+1} = x^k - B_k \nabla f_0\left(x^k\right)$$

XI. Покласти $k = k+1$ і перейти на крок IV.

Нижче наводиться алгоритм $\Pi_{\varepsilon}^{(0)}$ для псевдообернення $n \times m$-матриці $C_n$, яка складається із рядків $a^i$, $i = \overline{1,n}$, $\left(a^i = \left(a_1^i, a_2^i, \ldots, a_m^i\right)\right)$.

***Алгоритм 1′*** **(регуляризований рекурентний алгоритм $\Pi_{\varepsilon}^{(0)}$ стійкого псевдообернення скінченновимірних прямокутних матриць)**

Початок. I. Вибрати параметр регуляризації $\varepsilon \geq 0$.

II. Обчислити $\left\|a^1\right\|_1 = \sum_{j=1}^{m} \left|a_j^1\right|$.

III. Якщо $\left\|a^1\right\|_1 > \varepsilon$, то обчислити матрицю розмірності $n \times 1$

$$C_1^{\natural} = \left(a^1\right)^T \left(a^1 \left(a^1\right)^T\right)^{-1}$$

і перейти на крок IV; якщо $\left\|a^1\right\|_1 \leq \varepsilon$, то покласти $C_1^{\natural} = 0$ і перейти на крок IV.

IV. Покласти $i = 1$.

Основний цикл. V. Якщо $i < n$, то перейти на крок VI; інакше зупинити обчислення.

VI. Обчислити вектор-рядок $d^i$ розмірності $i$

$$d^i = a^{i+1} C_i^{\natural}.$$

VII. Обчислити вектор-рядок $b^i$ розмірності $m$:

$$b^i = a^{i+1} - d^i C_i,$$

де $C_i$ – матриця розмірності $i \times m$, яка складається з перших $i$ рядків початкової матриці $C_n$.

VIII. Обчислити $\left\| b^i \right\|_1 = \sum_{j=1}^{m} \left| b_j^i \right|$, якщо $\left\| b^i \right\|_1 > \varepsilon$, то перейти на крок IX; інакше перейти на крок XI.

IX. Обчислити вектор-стовпець $g^i$ розмірності $m$

$$g^i = \left( I - C_i^{\natural} C_i \right) \left( a^{i+1} \right)^T.$$

X. Обчислити вектор-стовпець $r^{i+1}$ розмірності $m$

$$r^{i+1} = g^i \left( b^i g^i \right)^{-1}$$

і перейти на крок XII.

XI. Обчислити вектор-стовпець $r^{i+1}$ розмірності $m$

$$r^{i+1} = C_i^{\natural} \left( d^i \right)^T \left( 1 + d^i \left( d^i \right)^T \right)^{-1}$$

і перейти на крок XII.

XII. Обчислити матрицю $S_i$ розмірності $m \times i$

$$S_i = C_i^{\natural} - r^{i+1} d^i.$$

XIII. Скласти матрицю $C_{i+1}^{\natural}$ розмірності $m \times (i+1)$, яка складається з матриці $S_i$ і вектор-стовпця $r^{i+1}$

$$C_{i+1}^{\natural} = \left( S_i \vdots r^{i+1} \right).$$

XIV. Покласти $i = i + 1$ і перейти на крок V.

При $\varepsilon = 0$ матриця $C_n^{\natural}$, яка породжується алгоритмом $1'$, завжди співпадає із псевдооберненою до $C_n$, тобто $C_n^{\natural} = C_n^+$, а алгоритм $1'$ є методом Гревілля псевдообернення прямокутних матриць.

При деякому (малому) $\varepsilon > 0$ алгоритм $\Pi_\varepsilon^{(0)}$ стійкий до малих збурень обчислень, а також до збурень елементів псевдооберненої матриці $C_n$ (метод Гревілля дуже чутливий до помилок обчислень). Ця стійкість має місце для широкого класу матриць навіть неповного рангу.

*Означення 1.* Послідовність $\left\{ x^k \right\}_{k=0}^{\infty}$, яка породжена алгоритмом 1, збігається надлінійно до стаціонарної точки $x^*$ функції $f_0(x)$, якщо при деякому $k_1 < \infty$ досягається $Px^{k_1} = Px^*$ або при всіх $k \geq 0$ – $Px^k \neq Px^*$ і $\lim\limits_{k \to \infty} \left\| P\left( x^{k+1} - x^* \right) \right\| / \left\| P\left( x^k - x^* \right) \right\| = 0$, де $P$ – ортопроектор на образ $R\left( \nabla_{xx}^2 f_0(x^*) \right)$ гесіана $\nabla_{xx}^2 f_0(x^*)$ в точці $x^*$.

***Теорема 1.*** *Нехай виконується припущення 0 і наступні умови:* $(\boldsymbol{iv})-$ *послідовність* $\{x^k\}_{k=0}^{\infty}$*, яка породжена алгоритмом 1, збігається до стаціонарної точки* $x^*$ *функції* $f_0(x)$; $(\boldsymbol{v})-$ *при кожному* $k \geq m$ *матриця* $A_{k,m}$ *має образ* $R(A_{k,m})$*, який містить образ* $R\left(\nabla_{xx}^2 f_0\left(x^k\right)\right)$ *матриці других похідних, знайденої в точці* $x^k$; $(\boldsymbol{vi})-$ *при кожному* $k \geq 0$ *виконується*

$$Px^k \neq Px^*,$$

*де* $P$ *– ортопроектор на образ* $R\left(\nabla_{xx}^2 f_0(x^*)\right)$ *гесіана* $\nabla_{xx}^2 f_0(x^*)$ *в точці* $x^*$;

$(\boldsymbol{vii})-$ *при кожному* $k \geq 0$ *виконується*

$$R\left(B_k^T\right) \supset R(P).$$

*Тоді достатніми умовами надлінійної швидкості збіжності є:*

$$\lim_{k\to\infty} \left\| P\left(B_k^+ - \nabla_{xx}^2 f_0(x^*)\right)\left(x^{k+1} - x^k\right)\right\| / \left\|P\left(x^{k+1} - x^k\right)\right\| = 0 \tag{3.16}$$

*і виконання умов в деякому околі точки* $x^*$:

$$\begin{gathered}\left\|\nabla_{xx}^2 f_0\left(x^k + \tau\left(x^{k+1} - x^k\right)\right)P - P\nabla_{xx}^2 f_0\left(x^k + \tau\left(x^{k+1} - x^k\right)\right)\right\| \leq \\ \leq \beta_1 \left\|P\left(x^{k+1} - x^k\right)\right\|, \quad 0 < \beta_1 < \infty, \quad \forall \tau \in [0,1].\end{gathered} \tag{3.17}$$

*Якщо для кожного* $x^{k+1}$ *має місце нерівність* (3.17) *і*

$$\left\|\nabla_{xx}^2 f_0\left(x^* + \tau\left(x^{k+1} - x^*\right)\right)P - P\nabla_{xx}^2 f_0\left(x^* + \tau\left(x^{k+1} - x^*\right)\right)\right\| \leq$$

$$\leq \beta_1 \left\|P\left(x^{k+1} - x^*\right)\right\|, \quad 0 < \beta_1 < \infty, \quad \forall \tau \in [0,1],$$

*то* (3.16) *є і необхідною умовою.*

***Теорема 1'.*** *Нехай виконується припущення 0, умови* $(\boldsymbol{iv})$, $(\boldsymbol{vi})$ *теореми 1, а в деякому околі точки* $x^*$ *виконується нерівність* (3.17) , $R(A_{k,m}) \supset R\left(\nabla_{\mathrm{xx}}^2 f_0\left(x^k\right)\right)$ *і нехай матриця* $B_k$ *на кроці IX алгоритму 1 обчислюється за формулою:*

$$B_k = \left(H_{k,m}A_{k,m}^+\right)^{\natural} + I - H_{k,m}A_{k,m}^+\left(H_{k,m}A_{k,m}^+\right)^{\natural},$$

*де псевдообернення* $G^{\natural}$ *означає, що воно ведеться алгоритмом 1' стійкого псевдообернення* $\Pi_{\varepsilon_k}^{(0)}$ *з* $\varepsilon_k$*, що задовольняють умовам:*

$$\max_{j\in[k-m+1,k]} \left\{\left\|x^k - x^j\right\|, \left\|x^j - x^{j-1}\right\|\right\} \leq \beta_2 \varepsilon_k^{2+\beta};$$

$$\lim_{k\to\infty} \varepsilon_k = 0,$$

*(тут* $\beta_2$, $\beta$ *– деякі додатні сталі), тоді* $\{x^k\}_{k=0}^{\infty}$ *збігається надлінійно.*

## 2. Стійке псевдообернення погано обумовлених матриць

*Означення 2.* Для прямокутної матриці $C$ ***псевдовласними числами*** називають числа $\lambda_i$, такі, що $\lambda_i^2$ є власними числами матриці $C^T C$. Відношення $\nu = \lambda_{\max} / \lambda_{\min}$, де $\lambda_{\max}$, $\lambda_{\min}$ – максимальне і мінімальне з ненульових псевдовласних чисел матриці $C$. Якщо $\nu$ – велике, то кажуть, що матриця $C$ ***погано обумовлена***.

Алгоритм $1'$ не застосовується для псевдообернення погано обумовлених матриць. Нижче наводиться модифікація алгоритму $\Pi_{\varepsilon}^{(0)}$, основана на послідовному уточнені псевдоoберненої матриці, що дає можливість псевдообертати і погано обумовлені, навіть вироджені матриці. Алгоритм 2 слугує для псевдообернення матриці $C_n$ розмірності $n \times m$, яка складається з рядків $a^i$, $i = 1,...,n$. На виході алгоритму 2 одержується матриця $C_n^{+}$, яка приймається за псевдообернену до $C_n$.

***Алгоритм 2*** **(алгоритм** $\Pi^{(0)}(\varepsilon, \delta, \mu, \rho)$ **стійкого псевдообернення погано обумовлених матриць)**

П о ч а т о к . І. Вибрати довільні константи (параметри алгоритму) $\varepsilon > 0,\ \delta > 0,\ \mu > 0,\ \rho > 0$ (рекомендується $\mu << \rho,\ \delta$ – «мале число»).

ІІ. Обчислити $\left\| a^1 \right\|_1$.

(Норма $\left\| b \right\|_1$ вектора $b = (b_1,...,b_m)$ обчислюється за формулою

$$\left\| b \right\|_1 = \sum_{j=1}^{m} \left| b_j \right| ).$$

ІІІ. Якщо $\left\| a^1 \right\|_1 > \varepsilon$, то обчислити матрицю $C_1^{+}$ розмірності $m \times 1$

$$C_1^{+} = \left(a^1\right)^T \left(a^1 \left(a^1\right)^T\right)^{-1}$$

і перейти на крок IV; якщо $\left\| a^1 \right\|_1 \le \varepsilon$, то покласти $C_1^{+} = 0$, де 0 – нуль-матриця розмірності $m \times 1$, і перейти на крок IV.

IV. Обчислити матрицю $G_1$ розмірності $m \times m$

$$G_1 = I - C_1^{+} C_1 .$$

V. Обчислити матрицю $Q_1$ розмірності $m \times m$

$$Q_1 = C_1^{+} (C_1^{+})^T .$$

VI. Покласти $i = 1$.

О с н о в н и й ц и к л . VII. Якщо $i < n$, то перейти на крок VIII; інакше зупинити обчислення.

VIII. Обчислити $\left\| a^{i+1} G_i \right\|_1$; якщо $\left\| a^{i+1} G_i \right\|_1 > \varepsilon$, то перейти на крок XV; інакше перейти на крок IX.

IX. Обчислити вектор-стовпець $r^{i+1}$ розмірності $m$

$$r^{i+1} = Q_i \left(a^{i+1}\right)^T \left(1 + a^{i+1} Q_i \left(a^{i+1}\right)^T\right)^{-1}.$$

X. Обчислити матрицю $S_i$ розмірності $m \times i$

$$S_i = \tilde{C}_i - r^{i+1} a^{i+1} \tilde{C}_i.$$

XI. Скласти матрицю $\tilde{C}_{i+1}$ розмірності $m \times (i+1)$, яка складається з матриці $S_i$ і вектор-стовпця $r^{i+1}$

$$\tilde{C}_{i+1} = \left(S_i \vdots r^{i+1}\right).$$

XII. Покласти $G_{i+1} = G_i$.

XIII. Обчислити матрицю $Q_{i+1}$ розмірності $m \times m$ за формулою

$$Q_{i+1} = \left(I - r^{i+1} a^{i+1}\right) Q_i.$$

XIV. Покласти $i = i + 1$ і перейти на крок VII.

XV. Обчислити вектор-стовпець $r^{i+1}$ розмірності $m$

$$r^{i+1} = G_i \left(a^{i+1}\right)^T \left(a^{i+1} G_i \left(a^{i+1}\right)^T\right)^{-1}.$$

XVI. Обчислити матрицю $S_i$ розмірності $m \times i$

$$S_i = \left(I - r^{i+1} a^{i+1}\right) \tilde{C}_i.$$

XVII. Скласти матрицю $\tilde{C}_{i+1}$ розмірності $m \times (i+1)$, яка складається з матриці $S_i$ і вектор-стовпця $r^{i+1}$

$$\tilde{C}_{i+1} = \left(S_i \vdots r^{i+1}\right).$$

XVIII. Обчислити матрицю $G_{i+1}$ розмірності $m \times m$

$$G_{i+1} = \left(I - r^{i+1} a^{i+1}\right) G_i.$$

XIX. Обчислити матрицю $Q_{i+1}$ розмірності $m \times m$ за формулою

$$Q_{i+1} = \left(I - r^{i+1} a^{i+1}\right) Q_i \left(I - r^{i+1} a^{i+1}\right)^T + r^{i+1} \left(r^{i+1}\right)^T.$$

XX. Обчислити величину:

$$\delta' = \left| \operatorname{tr} G_{i+1} - \operatorname{rang} G_{i+1} \right|,$$

де $\operatorname{tr} G_{i+1}$ – слід матриці $G_{i+1}$; $\operatorname{rang} G_{i+1}$ – ранг матриці $G_{i+1}$.

XXI. Якщо $\delta' \le \delta$, то покласти $i = i + 1$ і перейти на крок VII; інакше перейти на крок XXII.

(Кроки XXII-XXXVII алгоритму уточнюють одержані вище матриці $\tilde{C}_{i+1}$, $G_{i+1}$, $Q_{i+1}$).

XXII. Покласти $\overline{Q}_0^1 = Q_{i+1}$.

XXIII. Покласти $j = 1$.

XXIV. Обчислити матрицю $\left(\overline{C}_1^j\right)^{+}$ розмірності $m \times 1$

$$\left(\overline{C}_1^j\right)^{+} = \overline{Q}_0^j \left(a^1\right)^T \left(\rho^{-2}\mu^2 + a^1 \overline{Q}_0^j \left(a^1\right)^T\right)^{-1}.$$

XXV. Покласти $l = 0$.

XXVI. Обчислити вектор-стовпець $\overline{r}^{\,l+1}$ розмірності $m$

$$\overline{r}^{\,l+1} = \overline{Q}_l^j \left(a^{l+1}\right)^T \left(\rho^{-2}\mu^2 + a^{l+1} \overline{Q}_l^j \left(a^{l+1}\right)^T\right)^{-1}.$$

XXVII. Обчислити матрицю $\overline{Q}_{l+1}^j$ розмірності $m \times m$

$$\overline{Q}_{l+1}^j = \left(I - \overline{r}^{\,l+1} a_{l+1}^j\right) \overline{Q}_l^j.$$

XXVIII. Якщо $l = 0$, то перейти на крок XXXI; інакше перейти на крок XXIX.

XXIX. Обчислити матрицю $S_l^j$ розмірності $m \times l$

$$S_l^j = \left(I - \overline{r}^{\,l+1} a^{l+1}\right) \left(\overline{C}_l^j\right)^{+}.$$

Національний університет
водного господарства
та природокористування

XXX. Скласти матрицю $\left(\overline{C}_{l+1}^j\right)^{+}$ розмірності $m \times (l+1)$, яка складається з матриці $S_l^j$ і вектор-стовпця $\overline{r}^{\,l+1}$

$$\left(\overline{C}_{l+1}^j\right)^{+} = \left(S_l^j \vdots \overline{r}^{\,l+1}\right).$$

XXXI. Якщо $l < i$, то покласти $l = l + 1$ і перейти на крок XXXVI; інакше перейти на крок XXXII.

XXXII. Обчислити матрицю $\left(C_{i+1}^j\right)^{+}$ розмірності $m \times (i+1)$

$$\left(C_{i+1}^j\right)^{+} = \left(\rho + \rho^{-2}\mu^{-2}\right) \left(\overline{C}_{i+1}^j\right)^{+}.$$

XXXIII. Обчислити матрицю $Q_{i+1}^j$ розмірності $m \times m$

$$Q_{i+1}^j = \left(C_{i+1}^j\right)^{+} \left(\left(C_{i+1}^j\right)^{+}\right)^T.$$

XXXIV. Обчислити матрицю $G_{i+1}^j$ розмірності $m \times m$

$$G_{i+1}^j = I - \left(C_{i+1}^j\right)^{+} \left(C_{i+1}\right).$$

XXXV. Обчислити величину

$$\delta'' = \left| \operatorname{tr} G_{i+1}^j - \operatorname{rang} G_{i+1}^j \right|.$$

Якщо $\delta'' \le \delta$, то перейти на крок XXXVI; інакше перейти на крок XXXVII.

XXXVI. Покласти $C_{i+1}^{\mathbb{N}} = \left(C_{i+1}^{j}\right)^{\mathbb{N}}$, $G_{i+1} = G_{i+1}^{j}$, $Q_{i+1} = Q_{i+1}^{j}$, $i = i+1$ і перейти на крок VII.

XXXVII. Покласти $\bar{Q}_0^{j+1} = Q_{i+1}^{j}$, $j = j+1$ і перейти на крок XXIV.

## 3.7. Стохастичні квазіградієнтні методи

### 1. Загальний стохастичний квазіградієнтний метод для детермінованих задач

З а д а ч а 1. Знайти $\arg\min\limits_{x} f_0\left(x\right)$ для заданої функції $f_0 : R^n \to R^1$.

*Припущення 1.* (***i***) – функція $f_0$ – неперервно диференційована; (***ii***) – градієнт функції $f_0$ задовольняє умову Ліпшиця, тобто для будь-яких точок $x, y$

$$\left\|\nabla f_0(x) - \nabla f_0(y)\right\| \le \gamma \left\|x - y\right\|, \qquad 0 < \gamma < \infty.$$

У стохастичних квазіградієнтних методах наступні наближення до шуканого розв'язку знаходять у напрямку стохастичного квазіградієнту – випадкового вектора $\xi^k$, умовне математичне сподівання якого у деякому змісті близьке до градієнту (або узагальненого градієнту) даної функції, тобто:

Національний університет
водного господарства
та природокористування

$$E\left(\xi^k / x^0, \ldots, x^k\right) = \alpha_k \nabla f_0\left(x^k\right) + b^k, \;\; b^k \to 0, \;\; \alpha_k \to \text{const} > 0.$$

***Алгоритм 1***

П о ч а т о к . I. Вибрати довільну початкову точку $x^0 \in R^n$ і числа $\rho_o, \gamma_0$.

II. Покласти $k = 0$.

О с н о в н и й ц и к л . III. Обчислити реалізацію $\xi^k$ випадкового вектора $\xi$, умовне математичне сподівання якого задовольняє рівності:

$$E\left(\xi / x^0, \ldots, x^k\right) = \alpha_k \nabla f_0\left(x^k\right) + b^k,$$

де $\alpha_k$ – невід'ємна випадкова величина і $b^k$ – випадковий вектор, вимірні відносно $\sigma$-підалгебри $\mathcal{B}_k$ індукованої сімейством випадкових величин $\left(x^0, \ldots, x^k\right)$; $\nabla f_0\left(x^k\right)$ – градієнт функції $f_0$ в точці $x^k$.

IV. Обчислити вектор $x^{k+1} = x^k - \rho_k \gamma_k \xi^k$.

V. Обчислити кроковий $\rho_{k+1}$ та нормуючий $\gamma_{k+1}$ множники.

VI. Покласти $k = k+1$ і перейти на крок III.

***Теорема 1.*** *Нехай виконуються припущення 1 і, крім того, для кожного числа $\delta < \infty$ існує таке число $C_\delta < \infty$, що $\left\|\nabla f_0\left(x\right)\right\| \le C_\delta$ при $f_0\left(x\right) \le \delta$; $\eta_k\left(\omega\right)$ – випадкова величина, вимірна відносно $\sigma$-підалгебри $\mathcal{B}_k$,* *індукованої величинами* $x^0,...,x^k$, *є такою, що для будь-якого* $\delta < \infty$ *та деякого числа* $c_\delta$

$$E\left(\left\|\xi^k\right\|^2 / x^0,...,x^k\right) \le \eta_k^2 \le c_\delta$$

*при* $f_0\left(x^s\right) \le \delta$, $s = \overline{0,k}$; *нормуючий множник* $\gamma_k$ *задовольняє умові* $0 < \gamma \le \gamma_k \eta_k \le r_k < \infty$; *крім того, величини* $\rho_k$, $r_k$, *вимірні відносно* $\sigma$-*підалгебри* $\mathcal{B}_k$, *є такими, що* $\rho_k \ge 0$, $\alpha_k \ge 0$, $\left\|b^k\right\| / \alpha_k \to 0$ *рівномірно з ймовірністю* 1 *і*

$$\sum_{k=0}^{\infty} E\left(\rho_k \left\|b^k\right\| + \rho_k^2 r_k^2\right) < \infty; \quad \sum_{k=0}^{\infty} \rho_k \alpha_k \overset{\text{м.н.}}{=} \infty.$$

*Тоді нескінченна послідовність* $\left\{x^k\left(\omega\right)\right\}_{k=0}^{\infty}$, *яка породжена алгоритмом 1, є такою, що майже для кожного* $\omega$ *послідовність* $\left\{f_0\left(x^k\left(\omega\right)\right)\right\}_{k=0}^{\infty}$ *збігається, і* $\left\|\nabla f_0\left(x^{k_s}\left(\omega\right)\right)\right\| \to 0$ *майже напевно при* $s \to \infty$ *для деякої підпослідовності* $\left\{x^{k_s}\left(\omega\right)\right\}_{s=0}^{\infty}$.

***Теорема 1'.*** *Нехай виконуються припущення 1 і, крім того, для кожного числа* $\delta < \infty$ *існує число* $C_\delta < \infty$ *таке, що* $\left\|\nabla f_0(x)\right\| \le C_\delta$ *при* $f_0(x) \le \delta$; $\alpha_k \equiv 1; d_k(w)$ *– така випадкова величина, вимірна відносно* $\sigma$-*підалгебри* $\mathcal{B}_k$, *індукованої величинами* $x^0,...,x^k$, *що для будь-якого* $\delta$ *та деякого числа* $c_\delta < \infty$:

$$\sum_{j=1}^{n} D\left(\xi_j^k / x^0,...,x^k\right) \le d_k^2 \le c_\delta$$

*при* $f_0\left(x^s\right) \le \delta$, $s = \overline{0,k}$; *нормуючий множник* $\gamma_k$ *задовольняє умові* $0 < \gamma \le \gamma_k d_k \le r_k < \infty$; *крім того, величини* $\rho_k, r_k$, *вимірні відносно* $\sigma$-*підалгебри* $\mathcal{B}_k$, *є такими, що* $\rho_k \ge 0$, $\rho_k \to 0$, $\left\|b^k\right\| \to 0$ *рівномірно з ймовірністю 1 і*

$$\sum_{k=0}^{\infty} E\left(\rho_k \left\|b^k\right\| + \rho_k^2 r_k^2\right) < \infty; \quad \sum_{k=0}^{\infty} \rho_k = \infty$$

*з ймовірністю 1. Тоді нескінченна послідовність точок* $\left\{x^k\left(\omega\right)\right\}_{k=0}^{\infty}$, *яка породжена алгоритмом 1, є такою, що майже для кожного* $\omega$ *послідовність* $\left\{f_0\left(x^k\left(\omega\right)\right)\right\}_{k=0}^{\infty}$ *збігається, і* $\left\|\nabla f_0\left(x^{k_s}\left(\omega\right)\right)\right\| \to 0$ *майже напевно при* $s \to \infty$.

Н а с л і д о к. Якщо $f_0(x)$ унімодальна, тобто має тільки глобальний екстремум, і $\left\|\nabla f_0(x)\right\| \neq 0$ в усіх точках *x*, за винятком оптимальної, то з теорем 1 та 1′ випливає, що послідовність $\left\{f_0\left(x^k(\omega)\right)\right\}_{k=0}^{\infty}$ майже напевно збігається до глобального екстремуму.

**2. Загальний стохастичний квазіградієнтний метод для стохастичних задач**

З а д а ч а 2. Знайти $\arg\min_x E\varphi_0(x,\omega)$ для заданої функції $\varphi_0 : R^n \times \Omega \to R^1$.

*Припущення 2.* Функція $f_0(x) \triangleq E\varphi_0(x,\omega)$ двічі неперервно диференційована.

***Алгоритм 2***

П о ч а т о к. I. Вибрати: довільне початкове наближення $x^0 \in R^n$; початкові значення крокового та нормуючого множників $\rho_0$ та $\gamma_0$, початкове значення величини зміщення $\Delta_0$ (числа $\rho_0$, $\gamma_0$ і $\Delta_0$ вибираємо у відповідності з умовами теореми 2).

II. Покласти $k = 0$.

О с н о в н и й ц и к л. III. Вибрати натуральне число $s_k \geq 1$.

IV. Знайти $\left\{\xi^{k,s}\right\}_{s=1}^{s_k}$ – серію незалежних спостережень вектора $\xi = \left(\xi_1,\ldots,\xi_n\right)^T$ з незалежними та рівномірно розподіленими на $[-1;1]$ компонентами в *k*-й ітерації.

V. Знайти $\left\{\omega^{k,s}\right\}_{s=1}^{s_k}$ – серію незалежних по $k = 0,1,\ldots,$ спостережень станів природи $\omega$.

VI. Обчислити вектор

$$x^{k+1} = x^k - \rho_k\gamma_k \sum_{s=1}^{s_k} \frac{\varphi_0\left(x^k + \Delta_k\xi^{k,s}, \omega^{k,s}\right) - \varphi_0\left(x^k, \omega^{k,s}\right)}{\Delta_k}\xi^{k,s}.$$

VII. Обчислити кроковий множник $\rho_{k+1}$, нормуючий множник $\gamma_{k+1}$ та зміщення $\Delta_{k+1}$, які задовольняють умовам теореми 2.

VII. Покласти $k = k+1$ і перейти на крок III.

***Теорема 2.*** *Нехай виконуються припущення 2 і умови:* $(\boldsymbol{i})-$ $\mu_k(\omega)$ – *така випадкова величина, вимірна відносно* $\sigma$*-підалгебри* $\mathcal{B}_k$, *індукованої величинами* $(x^0,\ldots,x^k)$, *що для будь-якого* $\alpha$ *та деякого числа* $\tau(\alpha) < \infty$*:*

$$\sum_{j=1}^{n} D\left(\frac{\varphi_0\left(x^k+\Delta_k \xi^{k,s}, \omega^{k,s}\right)-\varphi_0\left(x^k, \omega^{k,s}\right)}{\Delta_k} \xi^{k,s} / x^k\right) \le \mu_k^2 \le \tau(\alpha)$$

*при* $f_0(x^s) \le \alpha,\ s = 0,1,...,k;$ $(\boldsymbol{ii})-$ *нормуючий множник* $\gamma_k$ *задовольняє умові* $0 < \gamma \le \gamma_k \mu_k \le r_k < \infty;$ $(\boldsymbol{iii})-$ *такі величини* $\rho_k$, $r_k$, $\Delta_k$, *вимірні відносно* $\sigma$ *-підалгебри* $\mathcal{B}_k$, *що* $\rho_k \ge 0$, $\rho_k \to 0$, $\Delta_k \to 0$ *рівномірно з ймовірністю 1 і:*

$$\sum_{k=0}^{\infty} E\left(\rho_k \Delta_k + \rho_k^2 r_k^2\right) < \infty; \quad \sum_{k=0}^{\infty} \rho_k = \infty$$

*з ймовірністю 1.*

*Тоді нескінченна послідовність точок* $\left\{x^k(\omega)\right\}_{k=0}^{\infty}$, *яка породжена алгоритмом 2, є такою, що майже для кожного* $\omega$ *послідовність* $\left\{f_0\left(x^k(\omega)\right)\right\}_{k=0}^{\infty}$ *збігається, і* $\left\|\nabla f_0\left(x^{k_s}(\omega)\right)\right\| \to 0$ *майже напевно при* $s \to \infty$ *для деякої підпослідовності* $k_s$.

**3. Стохастичний квазіградієнтний метод з процедурою переривання**

Задача 3. Знайти $\arg\min_x f_0(x)$ для неперервно диференційованої функції $f_0 : R^n \to R^1$.

Стохастичний квазіградієнтний метод з процедурою переривання оснований на побудові у $k$-й ітерації вектора стохастичного квазіградієнту $\xi^k$, умовне математичне сподівання якого:

$$E\left(\xi^k / x^0,...,x^k\right) = \nabla f_0\left(x^k\right) + b^k,$$

де $b^k$ – вектор, вимірний відносно $\sigma$-підалгебри $\mathcal{B}_k$, індукованої випадковими величинами ($x^0,..., x^k$). Якщо у деякій ітерації рух по стохастичному квазіградієнту виходить за межі сфери визначеного радіусу, то виникає переривання, і алгоритм починає працювати з довільної точки, яка належить наперед заданій обмеженій замкненій множині.

Алгоритми такого роду відображають ті ситуації, які часто зустрічаються на практиці, коли в процесі обчислень доводиться змінювати або параметри алгоритму, або сам алгоритм.

***Алгоритм 3***

Початок. I. Вибрати: довільну константу $\alpha$ $(0 < \alpha < \infty)$ і задати довільну замкнену множину $B$, яка міститься у сфері радіусу $\alpha$, тобто

$$B \subset S \triangleq \{x \mid \|x\| \le \alpha\}.$$

Вибрати довільну початкову точку $x^0 \in B$.

II. Задати правило формування послідовності крокових множників $\{\rho_k\}_{k=0}^{\infty}$.

III. Покласти $k=0$; знайти $\rho_0$.

Основний цикл. IV. Обчислити випадковий вектор $\xi^k$, який задовольняє умову:

$$E\left(\xi^k / x^0,...,x^k\right) = \nabla f_0\left(x^k\right) + b^k,$$

де $b^k$ – вектор, вимірний відносно $\sigma$-підалгебри $\mathcal{B}_k$, індукованої ($x^0,...,x^k$).

V. Обчислити вектор:

$$x^{k+1} = \begin{cases} x^k - \rho_k \xi^k, & \text{якщо} \quad \left\|x^k(\omega)\right\| \le \alpha; \\ z^{k+1} \in B, & \text{якщо} \quad \left\|x^k(\omega)\right\| > \alpha. \end{cases}$$

де $z^{k+1}$ – довільна точка множини $B$.

VI. Знайти $\rho_{k+1.}$

VII. Покласти $k=k+1$ і перейти на крок IV.

***Теорема 3.*** *Нехай $f_0$ неперервно диференційована функція і виконані умови: $(\boldsymbol{i})-$ існує така постійна $\delta$, що*

$$\left\|\xi^k\right\| + \left\|\nabla f_0\left(x^k\right)\right\| + \left\|b^k\right\| \le \delta;$$

$(\boldsymbol{ii})-$

$$\max_{x\in B} f_0(x) \le \inf_{\|x\|>\alpha} f_0(x);$$

$(\boldsymbol{iii})-$ *величини $\rho_k$, вимірні відносно $\sigma$-підалгебри $\mathcal{B}_k$, і задовольняють умовам:*

$$\rho_k \ge 0, \ \sum_{k=0}^{\infty}\rho_k = \infty, \ \sum_{k=0}^{\infty}\rho_k\left\|b^k\right\| < \infty \ \text{ м.н.}; \ \sum_{k=0}^{\infty} E\rho_k^2 < \infty.$$

*Тоді послідовність $\left\{x^k(\omega)\right\}_{k=0}^{\infty}$, яка породжена алгоритмом 3, є такою, що майже для кожного $\omega$ послідовність $\left\{f_0\left(x^k(\omega)\right)\right\}_{k=0}^{\infty}$ збігається, і будь-яка гранична точка послідовності $\left\{x^k(\omega)\right\}_{k=0}^{\infty}$ належить $X^* \triangleq \left\{x \mid \nabla f_0(x) = 0\right\}$ майже при кожному $\omega$.*

*Зауваження 3.* В даній теоремі, на відміну від теореми 1, не вимагається глобальної умови Ліпшиця. Умови рівномірної по $k$

обмеженості (умова (*i*)), які вимагаються в теоремі 3, легко послабити, використовуючи нормуючий множник $\gamma_k$. Ці умови обмеженості, однак, виконуються, якщо тільки точки $x^k$ змінюється в обмеженій області, а випадкові збурення мають усічені закони розподілу, тобто належить також обмеженій області.

#### 4. Стохастичний квазіградієнтний метод з постійним кроковим множником

Задача 4. Знайти $\arg\min\limits_x E_\omega f_0(x,\omega)$ для заданої функції $f_0 : R^n \times \Omega \to R^1$.

*Припущення 4.* Майже при всіх $\omega$ виконуються нерівності:

$$\|\nabla f_0(x,\omega)\| \le T < \infty, \quad \forall x \in R^n;$$

$$\|\nabla f_0(x,\omega) - \nabla f_0(y,\omega)\| \le L\|x - y\|, \quad \forall x, y \in R^n;$$

$$\min_{x\in R^n} E_\omega f_0(x,\omega) \triangleq F^* > -\infty.$$

***Алгоритм 4***

Початок. I. Вибрати довільне початкове наближення $x^0 \in R^n$, кроковий множник $\lambda > 0$.

II. Покласти $k = 0$.

Основний цикл. III. Обчислити реалізацію $\omega^k$ випадкового параметра $\omega$.

IV. Обчислити $\nabla f_0(x^k,\omega^k)$.

V. Обчислити наступне наближення

$$x^{k+1} = x^k - \lambda \nabla f_0(x^k,\omega^k).$$

VI. Покласти $k = k+1$ і перейти на крок III.

***Теорема 4.*** *Нехай виконується припущення 4, тоді для будь-яких чисел $\varepsilon \in (0;1)$ і $\delta > 0$ алгоритм 4 з постійним кроковим множником $\lambda$,*

$$0 < \lambda \le \gamma\delta^2 \Bigg/ \left( 2LT^2 \left( \left[ \left( 1 + \frac{2LT^2(F(x^0) - F^*)}{\gamma(1-\gamma)\delta^4} \right) \times \frac{1}{\varepsilon} \left( \frac{2LT^2}{\gamma\delta^2} \right)^2 \right] + 1 \right) \right),$$

$$0 < \gamma < 1, \qquad \left( F(x) \triangleq E_\omega f_0(x,\omega) \right)$$

*приводить в область $\delta$-екстремальних розв'язків*

$$X_\delta = \left\{ x \,\middle|\, \delta \ge \|\nabla E_\omega f_0(x,\omega)\| \right\}$$

*за $s$ $(s \le \varepsilon(\gamma^2\delta^4/(8n\lambda LT^4))^2)$ ітерацій з ймовірністю $P \ge 1 - \varepsilon$.*

## 3.8. Метод локальних варіацій

Задача 1. Знайти $\arg\min_{x\in R^n} f_0(x)$ для заданої неперервно диференційованої функції $f_0 : R^n \to R^1$.

Метод локальних варіацій не вимагає обчислення похідних функції, яка мінімізується і не є очевидною модифікацією алгоритму, яка вимагає таких обчислень.

***Алгоритм 1***

Початок. I. Вибрати початкове наближення $x^0 \in R^n$, яке задовольняє умовам теореми 1; константу $\rho_0 > 0$ (рекомендується $\rho_0 = 1$).

II. Обчислити $n$-вимірні вектори $h^i$, $i = \overline{1,2n}$:

$$h^{2j-1} = e^j, \quad h^{2j} = -e^j, \quad j = \overline{1,n};$$

де $e^j$, $j = \overline{1,n - j}$-й стовпчик одиничної $n\times n$-матриці.

III. Покласти $k = 0,\ x = x^0$.

IV. Обчислити $f_0(x)$.

Основний цикл. V. Покласти $\rho = \rho_k$.

VI. Покласти $i = 1$.

VII. Обчислити $f_0(x + \rho h^i)$.

VIII. Якщо $f_0(x + \rho h^i) < f_0(x)$, то покласти $x = x + \rho h^i$ і перейти на крок VI; інакше перейти на крок IX.

IX. Якщо $i < 2n$, то покласти $i = i + 1$ і перейти на крок VII; інакше перейти на крок X.

X. Покласти $x^{k+1} = x$, $\rho_{k+1} = \rho / 2$, $k = k + 1$ і перейти на крок V.

***Теорема 1.*** *Якщо функція $f_0$ неперервно диференційована і початкове наближення $x^0$ в алгоритмі 1 є таким, що множина*

$$X_0 \triangleq \left\{ x \middle| f_0(x) \le f_0(x^0),\ \ x \in R^n \right\}$$

*обмежена, то кожна гранична точка $x'$ послідовності $\{x^k\}_{k=0}^{\infty}$, яка породжена алгоритмом 1, задовольняє умову $\nabla f_0(x') = 0$. Якщо, крім того, множина*

$$X_* \triangleq \left\{ x \middle| \nabla f_0(x) = 0,\ \ x \in R^n \right\}$$

*складається з скінченної кількості точок і кожна її точка є або локальним мінімумом, або локальним максимумом, то послідовність $\{x^k\}_{k=0}^{\infty}$ збігається до такої точки $\bar{x}$, що $\nabla f_0(\bar{x}) = 0$.*

Метод локальних варіацій особливо ефективний, коли функція $f_0$ має вигляд:

$$f_0(x) = \sum_{i=1}^{n} f_i(x_i), \quad x = (x_1, x_2, \ldots, x_n)^T.$$

## 3.9. Методи випадкового пошуку

### 1. Метод випадкового пошуку в опуклих задачах мінімізації

З а д а ч а 1. Знайти $\arg\min_{x \in R^n} f_0(x)$ для заданої функції $f_0 : R^n \to R^1$.

*Припущення 1.* $(\boldsymbol{i})-$ функція $f_0 -$ опукла донизу в $R^n$; $(\boldsymbol{ii})-$ функція $f_0$ неперервно диференційована в $R^n$ та її градієнт задовольняє умову Ліпшиця в $R^n$, тобто

$$\| \nabla f_0(x) - \nabla f_0(y) \| \le \gamma \| x - y \|, \quad \forall x, y \in R^n.$$

У методі випадкового пошуку на $k$-й ітерації за відомим наближенням $x^k$ обчислюється наступне наближення $x^{k+1}$ як точка, у деякому змісті близька до точки мінімуму функції $f_0$ на прямій $x^k - \rho h^k$, $\rho \in (-\infty; \infty)$, де $h^k$ незалежна реалізація одиничного випадкового вектора рівномірно розподіленого на одиничній сфері із центром в початку координат. Для опуклої донизу функції алгоритм генерує послідовність $\{x^k\}_{k=0}^{\infty}$, для якої:

$$f_0(x^k) - f_0(x^*) = O(1/k),$$

де $x^*-$ розв'язок задачі 1, $O(t)-$ величина порядку $t$, а для сильноопуклої функції $f_0$ виконується:

$$x^k \to x^* \text{ і } f_0(x^k) \to f_0(x^*) \text{ при } k \to \infty$$

зі швидкістю збіжності геометричної прогресії.

***Алгоритм 1***

П о ч а т о к . I. Вибрати довільне початкове наближення $x^0 \in R^n$.

II. Покласти $k = 0$.

О с н о в н и й ц и к л . III. Якщо $\nabla f_0(x^k) = 0$, то покласти $x^* = x^k$ і припинити обчислення; інакше перейти на крок IV.

IV. Знайти незалежну реалізацію $h^k$ одиничного випадкового вектора $\xi$, рівномірно розподіленого на одиничній $n$-вимірній сфері із центром в початку координат.

V. Обчислити кроковий множник $\rho_k$, який задовольняє нерівність:

$$f_0(x^k - \rho_k h^k) \le (1 - \lambda_k) f_0(x^k) + \lambda_k w_k, \tag{3.18}$$

де $\lambda_k -$ довільна точка фіксованого відрізку $\lambda \le \lambda_k \le 1$ ($\lambda -$ довільна константа з напівінтервалу $(0;1]$, $w_k = \min_{\rho \in (-\infty;\infty)} f_0(x^k - \rho h^k)$.

VI. Обчислити наступне наближення

$$x^{k+1} = x^k - \rho_k h^k.$$

VII. Покласти $k = k+1$ і перейти на крок III.

***Теорема 1.*** *Нехай виконується припущення 1 і нехай множина* $X_0 \triangleq \left\{ x \middle| f_0(x) \le f_0(x^0),\ \ x \in R^n \right\}$ *обмежена. Тоді, якщо послідовність* $\{x^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 1, є такою, що* $\nabla f_0(x^k) \ne 0,\ k = 0,1,\ldots,$ *та існує номер* $m_0$, *для якого* $\alpha_{m_0} \ne 0$,

*де*

$$\alpha_k \triangleq \left( \nabla f_0(x^k), h^k \right) / \left\| \nabla f_0(x^k) \right\|,$$

*то для всіх* $m \ge m_0 + 1$ *виконується нерівність:*

$$f_0(x^m) - f_0(x^*) \le c_1 \left[ \sum_{k=0}^{m-1} (\alpha_k)^2 \right]^{-1},$$

*де* $c_1 -$ *додатна константа.*

***Теорема 1'.*** *Нехай виконуються всі припущення теореми 1 і функція* $f_0(x)$ *сильноопукла. Тоді для послідовності* $\{x^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 1, виконуються нерівності:*

$$f_0(x^m) - f_0(x^*) \le \left( f_0(x^0) - f_0(x^*) \right) \exp\left( -\frac{\nu\lambda}{2\gamma} \sum_{i=0}^{m-1} (\alpha_i)^2 \right),\ m = 1,2,\ldots;$$

$$\| x^m - x^* \| \le \frac{2}{\nu} \left( f_0(x^0) - f_0(x^*) \right) \exp\left( -\frac{\nu\lambda}{2\gamma} \sum_{i=0}^{m-1} (\alpha_i)^2 \right),\ \ m = 1,2,\ldots,$$

*тут* $\nu -$ *параметр сильноопуклої донизу функції.*

*Зауваження 1.* При виконанні припущень теореми 1 алгоритм 1 породжує послідовність $\{x^k\}_{k=0}^{\infty}$, для якої з ймовірністю рівною одиниці здійснюється подія «для довільного $\varepsilon$, яке задовольняє нерівностям $0 < \varepsilon < 1/n$, знайдеться таке число $l(\varepsilon)$, що нерівності:

$$f_0(x^k) - f_0(x^*) \le 2\gamma\eta^2 n / (k\lambda(1 - n\varepsilon))$$

справедливі для всіх цілих $k \ge l(\varepsilon)$».

Звідси, зокрема, випливає, що з ймовірністю рівною одиниці для будь-якого числа $\beta > 2\gamma\eta^2 n / \lambda$ знайдеться число $\overline{l}(\beta)$ таке, що нерівності:

$$f_0(x^k) - f_0(x^*) \le \beta / k$$

справедливі для всіх цілих $k \ge \overline{l}(\beta)$.

*Зауваження 1'.* При виконанні припущень теореми $1'$ алгоритм 1 породжує послідовність $\{x^k\}_{k=0}^{\infty}$, для якої з ймовірністю рівною одиниці виконується: $x^k \to x^*$ і $f_0(x^k) \to f_0(x^*)$ при $k \to \infty$ зі швидкістю збіжності геометричної прогресії.

**2. Адаптивний метод випадкового пошуку**

З а д а ч а 2. Знайти $\arg\min_{x \in R^n} f_0(x)$ для заданої функції $f_0 : R^n \to R^1$.

У даному адаптивному алгоритмі випадкового пошуку величина крокового множника змінюється залежно від результату, отриманого на попередній ітерації. Вектор $\xi^k$, який визначає напрямок руху в $k$-й ітерації до наступного наближення $x^{k+1}$, є незалежною реалізацією випадкового вектора $\xi$, розподіленого на одиничній сфері із щільністю $p_\xi(y)$

$$\int\limits_{\substack{\|y\|=1 \\ (g,y)>0}} p_\xi(y)ds = 1/2, \quad \forall g \neq 0. \tag{3.19}$$

***Алгоритм 2***

П о ч а т о к . I. Вибрати довільне початкове наближення $x^0 \in R^n$, довільне початкове значення крокового множника $\rho_0$ і константи $\gamma_1 > 1$, $0 < \gamma_2 < 1$; покласти $k = 0$.

О с н о в н и й ц и к л . II. Знайти незалежну реалізацію $\xi^k$ випадкового вектора $\xi$, розподіленого на одиничній сфері із щільністю $p_\xi(y)$, яка задовольняє умові (3.19) (цій вимозі задовольняє випадковий вектор $\xi$, рівномірно розподілений на сфері, а також проекція на одиничну сферу вектора, рівномірно розподіленого в кубі або на його поверхні).

III. Якщо $f_0(x^k + \rho_k \xi^k) < f_0(x^k)$, то покласти $h^k = \rho_k \xi^k$; інакше покласти $h^k = 0$.

IV. Обчислити наступне наближення

$$x^{k+1} = x^k + h^k.$$

V. Обчислити наступне значення крокового множника згідно умов:

$$\rho_{k+1} = \begin{cases} \gamma_1 \rho_k, & \text{якщо } f_0(x^{k+1}) < f_0(x^k), \\ \gamma_2 \rho_k, & \text{якщо } f_0(x^{k+1}) \geq f_0(x^k). \end{cases}$$

VI. Покласти $k = k+1$ і перейти на крок II.

***Теорема 2.*** *Якщо виконані умови:* $(\boldsymbol{i})-$ *функція* $f_0-$ *обмежена знизу;* $(\boldsymbol{ii})-$ *функція* $f_0-$ *неперервно диференційована;* $(\boldsymbol{iii})-$ $x^*-$ *єдина точка мінімуму функції* $f_0$; $(\boldsymbol{iv})- \nabla f_0(x) \neq 0$ *при* $x \neq x^*$; $(\boldsymbol{v})-$ *множина* $\{x | f_0(x) \leq \alpha\}$ *обмежена для всіх* $\alpha \leq f_0(x^0)$; $(\boldsymbol{vi})- \gamma_1 \gamma_2 > 1$, *то послідовність* $\{x^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 2, збігається з ймовірністю 1 до точки мінімуму* $x^*$.

## Запитання та вправи для самостійної роботи до розділу 3.

1. Сформулюйте означення гіперповерхні рівня функції $f_0(x),\ x \in R^n$ .
2. В якому напрямку здійснюється рух на кожній ітерації градієнтних методів?
3. В чому полягає різниця між методом найшвидшого спуску і його модифікацією?
4. Яка швидкість збіжності методу найшвидшого спуску і модифікації методу найшвидшого спуску?
5. Яка швидкість збіжності основного варіанту градієнтного методу?
6. Сформулюйте умови, при яких градієнтний метод з постійним кроковим множником є збіжним.
7. Які приклади матриць прискорення збіжності градієнтних методів Вам відомі?
8. В якому напрямку здійснюється рух до наступного наближення $x^{k+1}$ в модифікованому градієнтному методі, що не вимагає обчислення похідних?
9. Які вимоги ставляться до функції $f_0$ в методі Ньютона–Канторовича і узагальненому методі Ньютона–Канторовича?
10. Яка суттєва перевага узагальненого методу Ньютона–Канторовича порівняно із звичайним методом Ньютона–Канторовича?
11. Як обчислюється вектор руху до наступного наближення $x^{k+1}$ в методах типу Ньютона?
12. В якому напрямку здійснюється рух до наступного наближення $x^{k+1}$ в методах спряжених градієнтів?
13. Які переваги застосування методу спряжених градієнтів для мінімізації квадратичних функцій із строго додатньо визначеною симетричною матрицею?
14. Сформулюйте означення системи із n спряжених векторів (А – ортогональних векторів).
15. Дайте оцінку швидкості збіжності методів спряжених напрямків, методів типу Ньютона і градієнтних методів.
16. Які переваги застосування методу спряжених напрямків для мінімізації квадратичних функцій із строго додатньо визначеною симетричною матрицею?
17. В якому напрямку здійснюється рух до наступного наближення $x^{k+1}$ в методах псевдообернених операторів?

18. Назвіть переваги методів псевдообернених операторів.
19. Сформулюйте означення стохастичного квазіградієнту.
20. Сформулюйте умови, при яких стохастичний квазіградієнтний метод майже напевно збігається до глобального екстремуму.
21. В якому напрямку здійснюється рух до наступного наближення в методі локальних варіацій?
22. Для якого класу функцій особливо ефективний метод локальних варіацій?
23. З яким кроковим множником і в якому напрямку здійснюється рух до наступного наближення $x^{k+1}$ в методах випадкового пошуку?
24. Знайдіть в явному вигляді кроковий множник $\rho_k$ у методі найшвидшого спуску для квадратичної функції.
25. Знайдіть точку мінімуму функції $f(x) = (x_1^2 - x_2^2)/2 + (1 - x_1^2)/2$. Виконайте одну ітерацію класичного методу Ньютона з початкової точки $x^0 = (2;2)^T$. Обчислити значення $f(x^0), f(x^1)$.
26. Розв'яжіть наступні задачі безумовної оптимізації для квадратичної функції
$$f(x) = ax_1^2 + bx_1x_2 + cx_2^2 + dx_1 + ex_2 \to \min,$$
запропонованими методами безумовної оптимізації. Коефіцієнти $a,b,c,d,e$ задані в таблиці:

| №з/п | a | b | c | d | e | | №з/п | a | b | c | d | e |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. | 2 | 2 | -2 | 2 | -2 | | 12. | 9 | 3 | 3 | -1 | 1 |
| 2. | 2 | 2 | -3 | 2 | -2 | | 13. | 5 | -5 | 3 | 1 | -3 |
| 3. | 5 | -2 | 5 | -2 | -2 | | 14. | 2 | 2 | 3 | -1 | 1 |
| 4. | 5 | 4 | 1 | -16 | -11 | | 15. | 3 | -1 | 1 | 2 | -4 |
| 5. | 4 | -2 | 3 | -2 | -2 | | 16. | 3 | 2 | 2 | 2 | -2 |
| 6. | 3 | 1 | 6 | -5 | -12 | | 17. | 6 | -2 | 1 | 2 | -9 |
| 7. | 8 | -5 | 1 | 6 | 2 | | 18. | 6 | 4 | 6 | 2 | -2 |
| 8. | 2 | 2 | 4 | -2 | -2 | | 19. | 7 | -4 | 5 | 2 | -2 |
| 9. | 4 | -2 | 1 | 12 | -5 | | 20. | 3 | 3 | 3 | -3 | 5 |
| 10. | 6 | 2 | 1 | -2 | 1 | | 21. | 7 | -4 | 3 | 2 | 3 |
| 11. | 7 | -3 | 2 | -9 | -1 | | 22. | 9 | 5 | 1 | -2 | 2 |

**Р о з д і л 4**

# МЕТОДИ ОПТИМІЗАЦІЇ НЕДИФЕРЕНЦІЙОВАНИХ ФУНКЦІЙ І ВІДШУКАННЯ СІДЛОВИХ ТОЧОК В ЗАДАЧАХ БЕЗ ОБМЕЖЕНЬ

## 4.1. Означення узагальненого градієнта (субградієнта). Приклади обчислення субградієнтів

В розділі 3 припускалось, що цільова функція $f_0(x)$ неперервно диференційована. Проте на практиці часто зустрічаються оптимізаційні задачі, в яких цільова функція не має неперервних частинних похідних (тобто $f_0$ є негладкою функцією). Для прикладу розглянемо динамічну задачу планування і управління запасами. Виробниче підприємство планує свою роботу на $m$ кварталів. Нехай $p_i$, $i=\overline{1,m}$ – план виробництва продукту в $i$-му кварталі; $y_i$, $i=\overline{1,m}$ – фактичний обсяг виробництва продукту на кінець $i$-го кварталу. Вважається, що плани виробництва кожного кварталу різні, тому підприємство може виготовляти частину продукції передчасно (для забезпечення надійності виконання плану наступного кварталу) і зберігати її у вигляді запасів.

Позначимо через $\alpha_i$, $i=\overline{1,m}$ – витрати на зберігання одиниці продукції в $i$-му кварталі, а через $q_i(z_{i-1},z_i)$, $i=\overline{1,m}$ – функцію виробничих витрат, пов'язаних із зменшенням або збільшенням виробництва продукції:

$$q_i(z_{i-1},z_i)=\begin{cases} s_i(z_i-z_{i-1}), z_i\ge z_{i-1}, \\ r_i(z_{i-1}-z_i), z_i<z_{i-1}. \end{cases}$$

Задача полягає у знаходженні таких обсягів виробництва $z_i$ і запасів $y_i$ в кожному місяці $i=\overline{1,m}$, для яких загальні витрати мінімальні, тобто

$$f_0(x)=f_0(y,z)=\sum_{i=1}^{m}\alpha_i y_i+\sum_{i=1}^{m}q_i(z_{i-1},z_i)\to\min$$

при умовах:

$$y_i=y_{i-1}+z_i-p_i,\ \ i=\overline{1,m};$$
$$z_i\ge 0,\ \ y_i\ge 0,\ \ i=\overline{1,m}.$$

Сформульовану задачу можна представити як частинний випадок мінімаксної задачі наступним чином. Спочатку цільову функцію запишемо у вигляді

$$f_0(x)=\sum_{i=1}^{m}\alpha_i y_i+\sum_{i=1}^{m}\max\left\{s_i(z_i-z_{i-1}),r_i(z_{i-1}-z_i)\right\}.$$

Потім, вводячи допоміжні змінні $0 \le \lambda_i \le 1,\ i = \overline{1,m}$, цільову функцію представимо так

$$f_0(x) = \sum_{i=1}^{m} \alpha_i y_i + \max_{\lambda_i, i\in[1;m]} \sum_{i=1}^{m} \left[ \lambda_i s_i (z_i - z_{i-1}) + (1-\lambda_i) r_i (z_{i-1} - z_i) \right].$$

В оптимізаційних задачах з негладкою цільовою функцією (або негладкими функціями-обмеженнями) замість градієнта використовують поняття субградієнт або узагальнений градієнт.

*Означення 1.* Вектор $\hat{\nabla} f_0(\bar{x})$ називається ***узагальненим градієнтом*** (або ***субградієнтом***) функції $f_0$ в точці $\bar{x}$, якщо для будь-якого $x \in R^n$ виконується нерівність

$$f_0(x) - f_0(\bar{x}) \ge \hat{\nabla} f_0(\bar{x})(x - \bar{x}).$$

Відмітимо, що якщо $f_0(x)$ диференційована в точці $\bar{x}$, то узагальнений градієнт співпадає з градієнтом, тобто $\hat{\nabla} f_0(\bar{x}) = \nabla f_0(\bar{x})$.

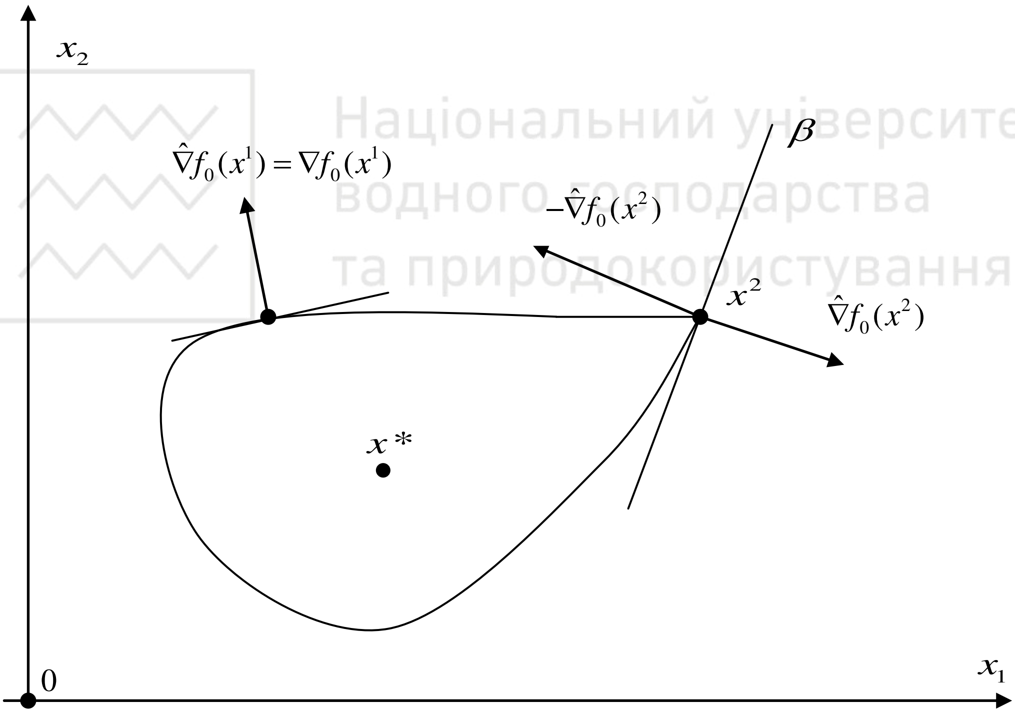



**Рис. 4.1**

На рис. 4.1 зображено субградієнт функції $f_0$ в точці $x^1$, який дорівнює градієнту і субградієнт в точці $x^2$, в якій не існує похідної.

Субградієнт функції в точці $x^2$ напрямлений вздовж вектора нормалі до опорної гіперплощини $\beta$, проведеної до лінії рівня в точці $x^2$. Зрозуміло, що в точці $x^2$ існує безліч субградієтів (так як існує безліч опорних гіперплощин). Відмітимо одну особливість методів узагальнених градієнтів.

Бачимо, що якщо для мінімізації функції $f_0$ в точці $x^2$ використовувати зображений на рис. 4.1 антисубградієнт, то ні при якому значенні крокового множника значення цільової функції в цьому напрямі не буде зменшуватись.

Вкажемо також на те, що множина субградієнтів $G(\bar{x})$ функції $f_0$ в будь-якій точці $\bar{x}$ є замкнутою, обмеженою і опуклою. Необхідною умовою мінімуму функції $f_0$ в точці $x^*$ являється $0 \in G(x^*)$.

Наведемо приклади обчислення субградієнтів деяких типів функцій.

**Приклад 1.** Для опуклої функції $f_0$ скалярного аргументу $x$ множиною субградієнтів $G(\bar{x})$ в точці $\bar{x}$ є опукла комбінація правосторонньої та лівосторонньої похідних в цій точці:

$$G(\bar{x}) = \left\{ \hat{\nabla} f_0(\bar{x}) \right\} = \lambda f_0^-(\bar{x}) + (1-\lambda) f_0^+(\bar{x}),$$

де $0 \le \lambda \le 1, f_0^-(\bar{x}), f_0^+(\bar{x}) -$ відповідно, похідна зліва та справа від точки $\bar{x}$, тобто:

$$f_0^-(\bar{x}) = \lim_{\varepsilon \to +0} \nabla f_0(\bar{x} - \varepsilon), \qquad f_0^+(\bar{x}) = \lim_{\varepsilon \to +0} \nabla f_0(\bar{x} + \varepsilon).$$

**Приклад 2.** Для сепарабельної функції $f_0(x) = \sum_{j=1}^{n} f_j(x_j)$, де $f_j$ – опуклі по $x_j$, $j = \overline{1,n}$, множина субградієнтів $G(\bar{x})$ в точці $\bar{x}$ є такою:

$$G(\bar{x}) = \left\{ \lambda_j f_j^-(\bar{x}_j) + (1-\lambda_j) f_j^+(\bar{x}_j),\ j = \overline{1,n} \right\},$$

де $0 \le \lambda \le 1,\ f_j^-(\bar{x}_j),\ f_j^+(\bar{x}_j) -$ відповідно лівостороння та правостороння похідні функції $f_j$ в точці $\bar{x}_j$.

**Приклад 3.** Для мінімаксної задачі $f_0(x) = \max_{y \in Y} f(x,y) \to \min_{x \in X}$ при виконанні умов:

а) $f(x,y) -$ опукла по $x$ для будь-якого $y \in Y$;

б) для кожного $x \in X$ існує $y(x)$ таке, що $f_0(x) = f(x, y(x)) = \max_{y \in Y} f(x,y)$;

в) множина $X -$ опукла;

г) для кожного $y \in Y$ існує субградієнт $\hat{\nabla}_x f(x,y)$,

субградієнт функції $f_0$ в точці $\bar{x} \in X$ обчислюється за формулою

$$\hat{\nabla} f_0(\bar{x}) = \hat{\nabla}_x f(\bar{x}, y) \Big|_{y = y(\bar{x})}.$$

**Приклад 4.** Для задачі про найкраще чебишевське наближення

$$f_0(x) = \max_{i \in [1;m]} \left| \sum_{j=1}^{n} \alpha_{ij} x_j - b_i \right| \to \min_x$$

субградієнт $\hat{\nabla} f_0(x)$ такий:

$$\hat{\nabla} f_0(x) = (g_1(x), g_2(x), ..., g_n(x))^T,$$

де

$$g_j(x) = \begin{cases} \alpha_{i(x)j}, & \text{якщо } \sum_{j=1}^{n} \alpha_{i(x)j} x_j - b_{i(x)} \geq 0, \\ -\alpha_{i(x)j}, & \text{якщо } \sum_{j=1}^{n} \alpha_{i(x)j} x_j - b_{i(x)} < 0. \end{cases}$$

Тут $i(x)$ задовольняє умову

$$\left| \sum_{j=1}^{n} \alpha_{i(x)j} x_j - b_{i(x)} \right| = \max_{i \in [1;m]} \left| \sum_{j=1}^{n} \alpha_{ij} x_j - b_i \right|.$$

**Приклад 5.** Для сформульованої вище в цьому підрозділі задачі планування і управління запасами з цільовою функцією вигляду:

$$f_0(x) = \sum_{i=1}^{m} \alpha_i y_i + \max_{\lambda_i, i \in [1;m]} \sum_{i=1}^{m} \left[ \lambda_i s_i (z_i - z_{i-1}) + (1 - \lambda_i) r_i (z_{i-1} - z_i) \right]$$

(тут $x = (y, z)^T = (y_1, y_2, ..., y_m, z_1, z_2, ..., z_m)^T, 0 \leq \lambda_i \leq 1$)

субградієнт такий:

$$\hat{\nabla} f_0(x) = (\alpha_1, \alpha_2, ..., \alpha_m, \hat{\nabla}_{z_1} f_0(x), \hat{\nabla}_{z_2} f_0(x), ..., \hat{\nabla}_{z_m} f_0(x))^T,$$

де $\hat{\nabla}_{z_i} f_0(x) = s_i \lambda_i(z) - r_i(1 - \lambda_i(z)) - s_{i+1} \lambda_{i+1}(z) - r_{i+1}(1 - \lambda_{i+1}(z)),$

$$\lambda_i(z) = \begin{cases} 1, & z_i - z_{i-1} \geq 0, \\ 0, & z_i - z_{i-1} < 0. \end{cases}$$

**Приклад 6.** Для задачі

$$f_0(x) = \left| \sum_{j=1}^{n} \alpha_j x_j - b \right|$$

узагальнений градієнт такий:

$$\hat{\nabla} f_0(x) = (\hat{\nabla}_1 f_0(x), \hat{\nabla}_2 f_0(x), ..., \hat{\nabla}_n f_0(x))^T,$$

$$\text{де } \hat{\nabla}_j f_0(x) = \begin{cases} \alpha_j, & \text{якщо } \sum_{j=1}^{n} \alpha_j x_j - b \geq 0, \\ -\alpha_{j,} & \text{якщо } \sum_{j=1}^{n} \alpha_j x_j - b < 0. \end{cases}$$

## 4.2. Опуклі функції та їх властивості

*Означення 2.* Функцію $f_0 : R^n \to R^1$, називають ***опуклою донизу*** (***опуклою***) в $R^n$, якщо для будь-яких $x^1, x^2 \in R^n$ і будь-якого $\lambda \in [0;1]$ виконується нерівність

$$f_0(\lambda x^1 + (1-\lambda)x^2) \le \lambda f_0(x^1) + (1-\lambda) f_0(x^2).$$

*Означення 3.* Функцію $f_0 : R^n \to R^1$, називають ***опуклою доверху*** (***догори, вгнутою, увігнутою***) в $R^n$, якщо $-f_0$ є опуклою донизу в $R^n$.

На рис. 4.2. а) зображено опуклу донизу функцію $f_0$, а на рис. 4.2. б) – опуклу доверху функцію $f_0$.

Якщо в означеннях 2,3 замість $R^n$ поставити $X$, де $X \subset R^n$ – опукла множина, то будемо мати означення опуклої донизу чи доверху функції $f_0$ на множині $X$.

Замінивши в означеннях 2 (3) знаки $\le (\ge)$ на знаки строгої нерівності <(>), будемо мати означення строго опуклої донизу (доверху) функції $f_0$.



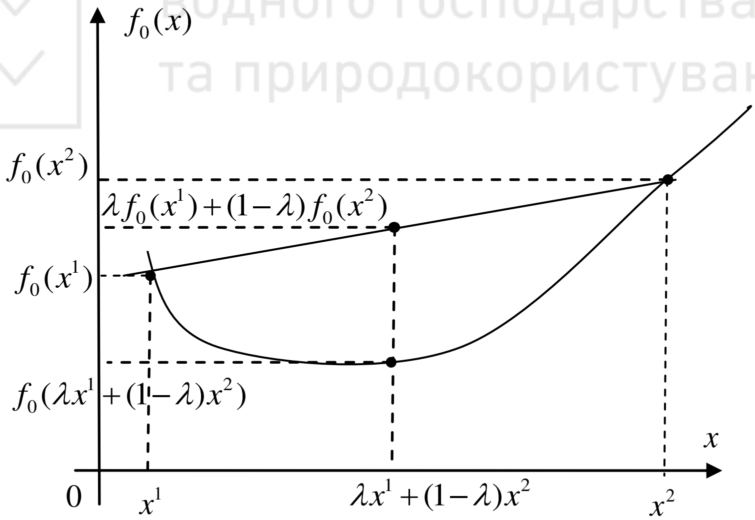



**Рис. 4.2, а**

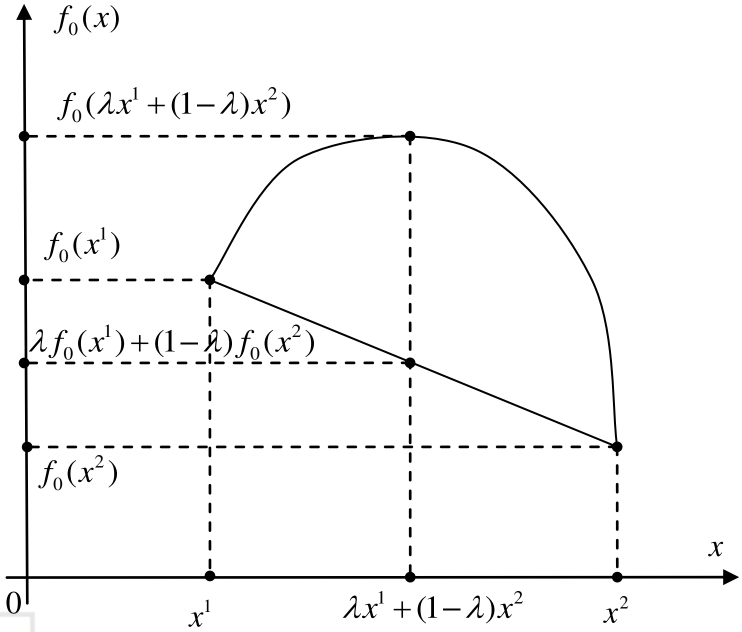



**Рис. 4.2, б**

Наведемо приклади опуклих функцій в $R^n$ або на опуклій множині $X$:

а) лінійна функція $f_0(x)=(c,x)$ є опуклою донизу і опуклою доверху;

б) квадратична функція $f_0(x)=(Ax,x)+(c,x)+d$ є опуклою донизу, якщо матриця $A$ невід'ємно визначена і опуклою доверху, якщо матриця $A$ недодатньо визначена;

в) сума опуклих донизу (доверху) функцій $f_0(x)=\sum_{j=1}^{m} f_j(x)$ є опуклою донизу (доверху) функцією;

г) якщо $g_0(x)$ опукла донизу (доверху) функція, то функція $f_0(x)=\max(g_0(x);0)$ є також опуклою донизу (доверху);

д) якщо функція $g_0(x)$ опукла донизу (доверху) і невід'ємна, то функція $f_0(x)=g_0^2(x)$ також опукла донизу (доверху);

г) якщо функції $g_j(x)$, $j=\overline{1,m}$ опуклі донизу, то функція $f_0(x)=\max\limits_{j\in[1;m]} g_j(x)$ є також опукла донизу.

***Основні властивості опуклих функцій:***

а) для довільної опуклої донизу функції $h(x)$, визначеної на опуклій множині $X \subset R^n$, множина $G=\left\{x \middle| h(x)\le 0,\ x\in X\right\}$ є опуклою множиною;

б) (нерівність Ієнсена) якщо $f_0(x)$ опукла донизу на опуклій множині $X \subset R^n$; $x^i \in X$, $i = \overline{1,m}$, $\lambda_i \geq 0$, $i = \overline{1,m}$ та $\sum_{i=1}^{m} \lambda_i = 1$, то

$$f_0\left(\sum_{i=1}^{m} \lambda_i x^i\right) \leq \sum_{i=1}^{m} \lambda_i f(x^i);$$

в) якщо функція $f_0(x)$ опукла донизу на опуклій множині $X \subset R^n$, то в кожній точці $\overline{x} \in X$ вона має похідну по довільному напрямку $h\left(\|h\| = 1\right)$

$$D_h f_0(\overline{x}) = \lim_{\alpha \to +0} \frac{f_0(\overline{x} + \alpha h) - f_0(x)}{\alpha};$$

г) якщо функція $f_0(x)$ опукла донизу і диференційована на опуклій множині $X \subset R^n$, то в кожній точці $\overline{x} \in X$ виконується нерівність

$$\nabla f_0(\overline{x}), x - \overline{x}) \leq f_0(x) - f_0(\overline{x}) \quad \forall x \in X;$$

д) (обернена до властивості г)) якщо для диференційованої на опуклій множині $X \subset R^n$ функції $f_0(x)$ в кожній внутрішній точці $\overline{x} \in X$ має місце нерівність

$$\left(\nabla f_0(\overline{x}), x - \overline{x}\right) \leq f_0(x) - f_0(\overline{x}) \quad \forall x \in X;$$

то функція $f_0(x)$ опукла донизу на множині $X$;

е) якщо функція $f_0(x)$ опукла донизу на опуклій множині $X \subset R^n$, то в кожній внутрішній точці $\overline{x} \in X$ існує субградієнт $\hat{\nabla} f_0(\overline{x})$;

є) якщо функція $f_0(x)$ опукла донизу на опуклій множині $X \subset R^n$, то множина її субградієнтів $G(\overline{x})$ для кожної внутрішньої точки $\overline{x} \in X$ є непорожньою, обмеженою, опуклою і замкненою;

ж) якщо функція $f_0(x)$ опукла донизу на опуклій множині $X \subset R^n$, то похідна цієї функції по довільному напрямку $h\left(\|h\| = 1\right)$ обчислюється за формулою:

$$D_h f_0(\overline{x}) = \max_{\hat{\nabla} f_0(\overline{x}) \in G(\overline{x})} (\hat{\nabla} f_0(\overline{x}), h);$$

з) якщо функція $f_0(x)$ опукла донизу на опуклій замкненій множині $X \subset R^n$, то локальний мінімум, що досягається в деякій точці множини $X$, є глобальним(абсолютним) мінімумом функції $f_0(x)$ на множині $X$;

і) (необхідні і достатні умови мінімуму опуклої донизу функції) якщо функція $f_0(x)$ опукла донизу на опуклій множині $X \subset R^n$, то внутрішня точка $x^* \in X$ є точкою абсолютного мінімуму функції $f_0(x)$ на множині

$X$ тоді і тільки тоді, коли нуль-вектор належить множині субградієнтів функції $f_0(x)$ в точці $x^*$, тобто $0 \in G(x^*)$.

## 4.3. Методи узагальненого градієнтного спуску

З а д а ч а 0. Знайти $\arg\min f_0\left(x\right)$ для заданої опуклої донизу функції $f_0 : R^n \to R^1$.

*Припущення 0.* Множина розв'язків $X^*$ задачі 0 непорожня.

**1. Метод зі сталим кроковим множником**

В алгоритмі 1 на $k$-й ітерації в якості вектору руху $h^k$ до наступного наближення $x^{k+1}$ обираємо протилежний до одиничного вектора узагальненого градієнта функції $f_0$ в точці $x^k$. Кроковий множник $\rho$ є сталою величиною. При заданому $\delta > 0$ можна вказати таке $\rho = \rho(\delta)$, що породжена алгоритмом 1 послідовність $\left\{x^k\right\}_{k=0}^{\infty}$ попадає в область, де функція $f_0$ відрізняється від свого мінімуму на величину $\delta$.

***Алгоритм 1***

П о ч а т о к. І. Вибрати довільне початкове наближення $x^0 \in R^n$, постійний кроковий множник $\rho > 0$; покласти $k = 0$.

О с н о в н и й ц и к л. ІІ. Обчислити узагальнений градієнт $g\left(x^k\right)$ функції $f_0$ в точці $x^k$. Якщо $g\left(x^k\right)$=0, то $x^k \in X^*$ і припинити обчислення; інакше перейти на крок ІІІ.

ІІІ. Обчислити вектор $h^k$ (який визначає напрямок руху до наступного наближення $x^{k+1}$)

$$h^k = g\left(x^k\right) / \left\| g\left(x^k\right) \right\|.$$

IV. Обчислити наступне наближення

$$x^{k+1} = x^k - \rho h^k.$$

V. Покласти $k = k + 1$ і перейти на крок ІІ.

***Теорема 1.*** *Якщо $f_0$ опукла донизу функція, то для довільного $\delta > 0$ можна знайти таке $\rho(\delta)$, що для послідовності $x^0, x^1, ..., x^k, ...$, породженої алгоритмом 1, при $\rho = \rho(\delta)$ знайдеться таке $k = k^*$, що $x^{k^*} \in X^*$, або така підпослідовність $x^{k_0}, x^{k_1}, ..., x^{k_i}, ...$, що*

$$\lim_{i \to \infty} f_0\left(x^{k_i}\right) - \min_{x \in R^n} f_0\left(x\right) < \delta.$$

***Теорема 1'.*** *Якщо опукла донизу функція $f_0$ має область мінімумів $X^*$ таку, що містить сферу радіуса $r > \rho/2$, то для послідовності*

$x^0, x^1, ..., x^k, ...,$ *породженої алгоритмом 1, знайдеться таке* $k = k*$, *що* $x^{k*} \in X*$.

**2. Основний метод**

В алгоритмі 2 на $k$-й ітерації в якості вектору, який визначає напрям руху до наступного наближення $x^{k+1}$, обирається одиничний вектор узагальненого градієнта функції $f_0$ в точці $x^k$. Кроковий множник $\rho_k$ задовольняє класичним умовам:

$$\rho_k \geq 0, \ \beta = 0,1,..., \ \sum_{k=0}^{\infty} \rho_k = \infty; \ \lim_{k \to \infty} \rho_k = 0. \tag{4.1}$$

***Алгоритм 2***

Початок. I. Вибрати довільне початкове наближення $x^0 \in R^n$ і покласти $k = 0$.

Основний цикл. II. Обчислити узагальнений градієнт $g(x^k)$ функції $f_0$ в точці $x^k$. Якщо $g(x^k)$=0, то покласти $x* = x^k$ і припинити обчислення; інакше перейти на крок III.

III. Обчислити вектор

$$h^k = g(x^k) / \| g(x^k) \|.$$

IV. Обчислити значення крокового множника $\rho_k$, яке задовольняє умовам теореми 2.

V. Обчислити наступне наближення

$$x^{k+1} = x^k - \rho_k h^k.$$

V. Покласти $k = k + 1$ і перейти на крок II.

***Теорема 2.*** *Нехай* $f_0(x)$ *– опукла донизу функція, область мінімумів* $X*$ *якої обмежена. Тоді, якщо крокові множники* $\rho_k$, $k = 0,1,...,$ *є такими, що:*

$$\rho_k \geq 0, \ \ \beta = 0,1,..., \ \ \lim_{k \to \infty} \rho_k = 0; \ \ \sum_{k=0}^{\infty} \rho_k = \infty,$$

*то нескінченна послідовність* $\{x^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 2, задовольняє граничним співвідношенням:*

$$\lim_{k \to \infty} \min_{x \in X*} \| x^k - x \| = 0; \quad \lim_{k \to \infty} f_0(x^k) = \min_{x \in R^n} f_0(x).$$

***Теорема 2'.*** *Нехай множина мінімумів*

$$X* \triangleq \left\{ x \mid f_0(x) = \inf_{x \in R^n} f_0(x) \right\}$$

*опуклої донизу функції* $f_0$ *непорожня і крокові множники* $\rho_k$, $k = 0,1,...,$ *задовольняють умовам:*

$$\rho_k > 0,\ k = 0,1,...;\ \lim_{k\to\infty}\rho_k = 0;\ \sum_{k=0}^{\infty}\rho_k = \infty.$$

*Тоді, якщо виконується одна з наступних п'яти умов:* $(\boldsymbol{iv})-$ *множина* $X^*$ *обмежена;* $(\boldsymbol{v})-\ \sum_{k=0}^{\infty}\rho_k^2 < \infty$; $(\boldsymbol{vi})-\ \operatorname{int} X^* \neq \varnothing$; $(\boldsymbol{vii})-$ *множина* $X^*$ *є лінійним многовидом в* $R^n$; $(\boldsymbol{viii})-n=2$, *то граничні точки нескінченної послідовності* $\{x^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 2, належать множині* $X^*$. *При цьому умови* $(\boldsymbol{v})$ *і* $(\boldsymbol{vii})$ *забезпечують збіжність послідовності* $\{x^k\}_{k=0}^{\infty}$ *до точки* $\bar{x} \in X^*$, *а умова* $(\boldsymbol{vi})$ *забезпечує скінченність алгоритму 2.*

**Приклад 1.** Використовуючи алгоритм 2, розв'язати перевизначену систему алгебраїчних рівнянь

$$\begin{cases} x_1 - 2x_2 + x_3 = 4, \\ 2x_1 + x_2 + x_3 = 1, \\ x_1 - x_2 - x_3 = 2, \\ 2x_1 + 2x_2 + x_3 = 4, \\ -x_1 + x_2 - 2x_3 = 2. \end{cases}$$

*Розв'язування.* Розв'язок даної системи у відповідності із прикладом 4 підрозділу 4.1 будемо шукати як розв'язок наступної задачі про найкраще чебишевське наближення

$$f_0(x) = \max_{i\in[1;5]}\left|\sum_{j=1}^{3}\alpha_{ij}x_j - b_i\right| \to \min_x .$$

Виконаємо п'ять ітерацій алгоритму 2.

**Алгоритм 2**

I . Вибираємо початкове наближення $x^0 = (0;0;0)^T$ і покладемо $k = 0$.

*1-а ітерація:*

II. Обчислюємо узагальнений градієнт $\hat{\nabla} f_0(x^0)$:

а) спочатку обчислюємо нев'язки $\delta_i = \left|\sum_{j=1}^{3}\alpha_{ij}x_j - b_i\right|$ для $i = \overline{1,5}$:

$i=1: \delta_1 = |0-4| = |-4| = 4,$

$i=2: \delta_2 = |0-1| = |-1| = 1,$

$i=3: \delta_3 = |0-2| = |-2| = 2,$

$i=4: \delta_4 = |0-4| = |-4| = 4,$

$i=5: \delta_5 = |0-2| = |-2| = 2;$

б) знаходимо $i(x^0)=1$;

в) оскільки $\sum_{j=1}^{3}\alpha_{1j}x^0{}_j - b_1 < 0$, то компоненти узагальненого градієнта такі: $g_1(x^0)=-1$; $g_2(x^0)=2$; $g_3(x^0)=-1$.

Отже, узагальнений градієнт в точці $x^0$ знайдений $\hat{\nabla} f_0(x^0)=(-1;2;-1)^T \neq 0$; переходимо на крок III.

III. Обчислюємо вектор

$$h^0=\hat{\nabla} f_0(x^0)/\left\|\hat{\nabla} f_0(x^0)\right\|=(-1;2;-1)^T/\sqrt{1+4+1}=$$

$$=(1/\sqrt{6})(-1;2;-1)^T \cong (-0,408;0,816;-0,408)^T.$$

IV. Покладемо $\rho_0=1$.

V. Обчислюємо наступне наближення

$$x^1=x^0-\rho_0 h^0=\begin{pmatrix}0\\0\\0\end{pmatrix}-\begin{pmatrix}-0,408\\0,816\\-0,408\end{pmatrix}=\begin{pmatrix}0,408\\-0,816\\0,408\end{pmatrix}.$$

VI. Покладемо $k=0+1=1$ і переходимо на крок II.

*2-а ітерація:*

II. Обчислюємо узагальнений градієнт $\hat{\nabla} f_0(x^1)$:

а) спочатку обчислюємо нев'язки:

$i=1:\delta_1=|0,408-2(-0,816)+0,408-4|=|-1,552|=1,552,$

$i=2:\delta_2=|2\cdot 0,408-0,816+0,408-1|=|-0,592|=0,592,$

$i=3:\delta_3=|0,408+0,816-0,408-2|=|-1,184|=1,184,$

$i=4:\delta_4=|2\cdot 0,408-2\cdot 0,816+0,408-4|=|-4,408|=4,408,$

$i=5:\delta_5=|-0,408-0,816-2\cdot 0,408-2|=|-4,04|=4,04;$

б) знаходимо $i(x^1)=4$;

в) оскільки $\sum_{j=1}^{3}\alpha_{4j}x_j^1 - b_4 < 0$, то компоненти узагальненого градієнта такі: $g_1(x^1)=-2$; $g_2(x^1)=-2$; $g_3(x^1)=-1$.

Отже, узагальнений градієнт в точці $x^1$ знайдений: $\hat{\nabla} f_0(x^1)=(-2;-2;-1)^T \neq 0$; переходимо на крок III.

III. Обчислюємо вектор

$$h^1=\hat{\nabla} f_0(x^1)/\left\|\hat{\Delta} f_0(x^1)\right\|=(-2;-2;-1)^T/\sqrt{4+4+1}=$$

$$=(1/3)(-2;-2;-1)^T.$$

IV. Покладемо $\rho_1 = 0,8.$

V. Обчислюємо наступне наближення

$$x^2 = x^1 - \rho_1 h^1 = \begin{pmatrix} 0,408 \\ -0,816 \\ 0,408 \end{pmatrix} - 0,8 \begin{pmatrix} -2/3 \\ -2/3 \\ -1/3 \end{pmatrix} = \begin{pmatrix} 0,941 \\ -0,283 \\ 0,674 \end{pmatrix}.$$

VI. Покладемо $k = 1+1 = 2$ і переходимо на крок II.

*3-я ітерація:*

II. Обчислюємо узагальнений градієнт $\widehat{\nabla} f_0(x^2)$:

а) спочатку обчислюємо нев'язки:

$i = 1: \delta_1 = |0,941 - 2(-0,283) + 0,674 - 4| = |-1,819| = 1,819,$

$i = 2: \delta_2 = |2 \cdot 0,941 - 0,283 + 0,674 - 1| = |1,273| = 1,273,$

$i = 3: \delta_3 = |0,941 + 0,283 - 0,674 - 2| = |-1,45| = 1,45,$

$i = 4: \delta_4 = |2 \cdot 0,941 - 2 \cdot 0,283 + 0,674 - 4| = |-2,01| = 2,01,$

$i = 5: \delta_5 = |-0,941 - 0,283 - 2 \cdot 0,674 - 2| = |-4,572| = 4,572;$

б) знаходимо $i(x^2) = 5$;

в) оскільки $\sum_{j=1}^{3} \alpha_{5j} x_j^2 - b_5 < 0$, то компоненти узагальненого градієнта такі:

$$g_1(x^2) = 1;\ g_2(x^2) = -1;\ g_3(x^2) = 2.$$

Отже, узагальнений градієнт в точці $x^2$ знайдений: $\widehat{\nabla} f_0(x^2) = (1; -1; 2)^T \neq 0$; переходимо на крок III.

III. Обчислюємо вектор

$$h^2 = \widehat{\nabla} f_0(x^2) / \left\| \hat{\Delta} f_0(x^2) \right\| = (1; -1; 2)^T / \sqrt{1+4+1} =$$

$$= (1/\sqrt{6})(1; -1; 2)^T \cong (0,408; -0,408; 0,816)^T.$$

IV. Покладемо $\rho_2 = 0,7.$

V. Обчислюємо наступне наближення

$$x^3 = x^2 - \rho_2 h^2 = \begin{pmatrix} 0,941 \\ -0,283 \\ 0,674 \end{pmatrix} - 0,7 \begin{pmatrix} 0,408 \\ -0,408 \\ 0,816 \end{pmatrix} = \begin{pmatrix} 0,655 \\ 0,003 \\ 0,103 \end{pmatrix}.$$

VI. Покладемо $k = 2+1 = 3$ і переходимо на крок II.

*4-а ітерація:*

II. Обчислюємо узагальнений градієнт $\widehat{\nabla} f_0(x^3)$:

а) спочатку обчислюємо нев'язки:

$i=1: \delta_1 = |0,655 - 2\cdot 0,003 + 0,103 - 4| = |-3,248| = 3,248,$

$i=2: \delta_2 = |2\cdot 0,655 - 0,003 + 0,103 - 1| = |0,416| = 0,416,$

$i=3: \delta_3 = |0,655 - 0,003 - 0,103 - 2| = |-1,451| = 1,451,$

$i=4: \delta_4 = |2\cdot 0,655 + 2\cdot 0,003 + 0,103 - 4| = |-2,581| = 2,581,$

$i=5: \delta_5 = |-0,655 + 0,003 - 2\cdot 0,103 - 2| = |-2,858| = 2,858;$

б) знаходимо $i(x^3) = 1$;

в) оскільки $\sum_{j=1}^{3} \alpha_{1j} x_j^3 - b_1 < 0$, то компоненти узагальненого градієнта

такі: $g_1(x^3) = -1$; $g_2(x^3) = 2$; $g_3(x^3) = -1$;

отже, узагальнений градієнт в точці $x^3$ знайдений:
$\widehat{\nabla} f_0(x^3) = (-1; 2; -1)^T \neq 0$; переходимо на крок III.



III. Обчислюємо вектор

$$h^3 = \widehat{\nabla} f_0(x^3) / \left\| \widehat{\nabla} f_0(x^3) \right\| = (-1; 2; -1)^T / \sqrt{1+4+1} =$$
$$= (1/\sqrt{6})(-1; 2; -1)^T \cong (-0,408; 0,816; -0,408)^T.$$

IV. Покладемо $\rho_3 = 0,5$.

V. Обчислюємо наступне наближення

$$x^4 = x^3 - \rho_3 h^3 = \begin{pmatrix} 0,655 \\ 0,003 \\ 0,103 \end{pmatrix} - 0,5 \begin{pmatrix} -0,408 \\ 0,816 \\ -0,408 \end{pmatrix} = \begin{pmatrix} 0,859 \\ -0,405 \\ 0,307 \end{pmatrix}.$$

VI. Покладемо $k = 3+1 = 4$ і переходимо на крок II.

*5-а ітерація*:

II. Обчислюємо узагальнений градієнт $\widehat{\nabla} f_0(x^4)$:

а) спочатку обчислюємо нев'язки:

$i=1: \delta_1 = |0,859 + 2\cdot 0,405 + 0,307 - 4| = |-2,024| = 2,024,$

$i=2: \delta_2 = |2\cdot 0,859 - 0,405 + 0,307 - 1| = |-0,62| = 0,62,$

$i=3: \delta_3 = |0,859 + 0,405 - 0,307 - 2| = |-1,043| = 1,043,$

$i=4: \delta_4 = |2\cdot 0,859 - 2\cdot 0,405 + 0,307 - 4| = |-2,785| = 2,785,$

$i=5: \delta_5 = |-0,859 - 0,405 - 2\cdot 0,307 - 2| = |-3,878| = 3,878;$

б) знаходимо $i(x^4)=5$;

в) оскільки $\sum_{j=1}^{3}\alpha_{5j}x_j^4-b_5<0$, то компоненти узагальненого градієнта

такі: $g_1(x^4)=1$; $g_2(x^4)=-1$; $g_3(x^4)=2$;

отже, узагальнений градієнт в точці $x^4$ знайдений:
$\widehat{\nabla} f_0(x^4)=(1;-1;2)^T\neq 0$; переходимо на крок III.

III. Обчислюємо вектор

$$h^4=\widehat{\nabla} f_0(x^4)/\left\|\widehat{\nabla} f_0(x^4)\right\|=(1;-1;2)^T/\sqrt{1+1+4}\cong$$

$$\cong(0,408;-0,408;0,816)^T.$$

IV. Покладемо $\rho_4=0,4$.

V. Обчислюємо наступне наближення

$$x^5=x^4-\rho_4h^4=\begin{pmatrix}0,859\\-0,405\\0,307\end{pmatrix}-0,4\begin{pmatrix}0,408\\-0,408\\0,816\end{pmatrix}=\begin{pmatrix}0,695\\-0,242\\-0,019\end{pmatrix}.$$

VI. Покладемо $k=4+1=5$ і зупиняємось.

Таким чином, за п'ять ітерацій алгоритму 2 знайшли наближений розв'язок перевизначеної системи рівнянь $x^5\approx(0,695;-0,242;-0,019)^T$, причому $f_0(x^5)=3,13$.



**3. Модифікація основного методу**

В алгоритмі 3 на $k$-й ітерації в якості вектора, який визначає напрям руху до наступного наближення $x^{k+1}$, обирається вектор протилежний до узагальненого градієнта функції $f_0$ в точці $x^k$. Кроковий множник $\rho_k$ задовольняє класичним умовам (4.1).

***Алгоритм 3***

Початок. I. Вибрати довільне початкове наближення $x^0\in R^n$ і покласти $k=0$.

Основний цикл. II. Обчислити узагальнений градієнт $g\left(x^k\right)$ функції $f_0$ в точці $x^k$.

III. Обчислити кроковий множник $\rho_k$, який задовольняє умовам теореми 3.

IV. Обчислити наступне наближення

$$x^{k+1}=x^k-\rho_k g(x^k).$$

V. Покласти $k=k+1$ і перейти на крок II.

***Теорема 3.*** *Нехай* $f_0(x)-$ *опукла донизу функція, область мінімумів* $X^*$ *якої обмежена. Тоді, якщо:* $(\boldsymbol{i})-$ *крокові множники* $\rho_k$, $k=0,1,...,$ *такі, що* $\rho_k \geq 0$; $\lim\limits_{k\to\infty}\rho_k=0$; $\sum\limits_{k=0}^{\infty}\rho_k=\infty$; $(\boldsymbol{ii})-$ *послідовність узагальнених градієнтів* $\{g(x^k)\}_{k=0}^{\infty}$, *яка породжена алгоритмом 3, обмежена, то нескінченна послідовність* $\{x^k\}_{k=0}^{\infty}$ *задовольняє граничним співвідношенням:*

$$\lim_{k\to\infty}\min_{x\in X^*}\left\|x^k-x\right\|=0;\quad \lim_{k\to\infty}f_0\left(x^k\right)=\min_{x\in R^n}f_0\left(x\right).$$

**4. Перший метод із спеціальним вибором крокового множника**

Алгоритм 4 може бути застосований для мінімізації функції $f_0(x)$, яка задовольняє припущенню 4.

*Припущення 4.* Нехай $f_0(x)$ – така опукла донизу функція, що при деякому $0\leq\varphi<\dfrac{\pi}{2}$, для всіх $x\in R^n$ виконується нерівність:

$$\left(g\left(x\right),x-x^*\left(x\right)\right)\geq\cos\varphi\left\|g\left(x\right)\right\|\cdot\left\|x-x^*\left(x\right)\right\|,\qquad(4.2)$$

де $x^*(x)-$ точка, що належить множині мінімумів функції $f_0(x)$ і лежить на найкоротшій відстані від $x$.

В алгоритмі 4 на $k$-й ітерації в якості вектору руху $h^k$ до наступного наближення $x^{k+1}$ обираємо протилежний до одиничного вектора узагальненого градієнта функції $f_0$ в точці $x^k$. Крокові множники $\rho_k$, $k=0,1,...,$ обчислюються у відповідності з рекурентною формулою:

$$\rho_{k+1}=\rho_k\beta(\varphi),\ \ \rho_0=\rho_0\left(x^0,\varphi\right),$$

де

$$\beta\left(\varphi\right)=\begin{cases}\sin\varphi, & \pi/4\leq\varphi<\pi/2,\\ 1/\left(2\cos\varphi\right), & 0\leq\varphi<\pi/4.\end{cases}\qquad(4.3)$$

Послідовність $\{x^k\}_{k=0}^{\infty}$, яка породжена алгоритмом 4, збігається до множини мінімумів із швидкістю геометричної прогресії , знаменник якої $q=\beta(\varphi)$.

***Алгоритм 4***

П о ч а т о к . І. Вибрати довільну початкову точку $x^0\in R^n$.

II. Знайти кут $0\leq\varphi<\pi/2$, який задовольняє нерівності (4.2) для всіх $x\in R^n$.

III. Обчислити значення $\beta\left(\varphi\right)$ за формулою (4.3).

IV. Обчислити початкове значення крокового множника $\rho_0$, яке задовольняє нерівність:

$$\rho_0 \geq \begin{cases} \left\| x^*\left(x^0\right) - x^0 \right\| \cos\varphi, & \pi/4 \leq \varphi < \pi/2, \\ \left\| x^*\left(x^0\right) - x^0 \right\| \big/ \left(2\cos\varphi\right), & 0 \leq \varphi < \pi/4, \end{cases}$$

де $x^*\left(x^0\right)-$ точка, що належить множині мінімумів функції $f_0(x)$ і лежить на найкоротшій відстані від точки $x^0$.

V. Покласти $k = 0$.

Основний цикл. VI. Обчислити вектор напрямку руху $h^k$ до наступного наближення $x^{k+1}$

$$h^k = -g\left(x^k\right) \big/ \left\| g\left(x^k\right) \right\|.$$

VII. Обчислити наступне наближення

$$x^{k+1} = x^k + \rho_k h^k .$$

VIII. Обчислити значення крокового множника:

$$\rho_{k+1} = \rho_k \beta(\varphi),$$

де $\beta(\varphi)$ визначається за формулою (4.3).

IX. Покласти $k = k+1$ і перейти на крок VI.

***Теорема 4.*** *Нехай виконується припущення 4. Тоді послідовність* $x^k$, $k = 0,1,...$, *яка породжена алгоритмом 4, є такою, що або при деякому* $\bar{k}$ – $g\left(x^{\bar{k}}\right) = 0$ *і* $x^{\bar{k}}$ *належить області мінімумів, або при всіх* $k = 0,1,...$, *буде виконуватися нерівність*

$$\left\| x^k - x^*\left(x^k\right) \right\| \leq \begin{cases} \rho_k / \cos\varphi, & \pi/4 \leq \varphi < \pi/2, \\ 2\cos\varphi \cdot \rho_k, & 0 \leq \varphi < \pi/4. \end{cases}$$

*Зауваження 4.* В нерівності (4.2) $\cos\varphi$ показує ступінь витягнутості гіперповерхонь рівня функції $f_0$. Якщо в деякому околі мінімуму функції $f_0$ не існує кута $\varphi < \pi/2$, який задовольняє нерівності (4.2), то така функція називається суттєво витягнутою і для її мінімізації алгоритм 4 незастосовний. У цьому випадку слід використовувати універсальний метод вибору крокових множників, як у алгоритмі 3.

### 5. Другий метод із спеціальним вибором крокового множника

Алгоритм 5 може бути застосований для мінімізації функції $f_0(x)$, яка задовольняє припущенням 5.

*Припущення 5.* $(\boldsymbol{i})-$ функція $f_0$ опукла донизу в $R^n$; $(\boldsymbol{ii})-$ функція $f_0$ має єдину точку мінімуму $x^*$; $(\boldsymbol{iii})-$ для будь-якого числа $\alpha > 0$ існує таке скінченне число $\tau > 0$, що для будь-якої пари точок $x,\ z \in Y$

$$Y \triangleq \left\{ y \mid \left\| y - x^* \right\| \le \alpha \right\},$$

такої, що $f_0(x) = f_0(z) \ne f_0(x^*)$, виконується умова

$$\left( \left\| x - x^* \right\| / \left\| z - x^* \right\| \right) \le \tau .$$

В алгоритмі 5 на $k$-й ітерації в якості вектору руху $h^k$ до наступного наближення $x^{k+1}$ обираємо протилежний до одиничного вектора узагальненого градієнта функції $f_0$ в точці $x^k$. Крокові множники $\rho_k$, $k = 0,1,...,$ обчислюються у відповідності з рекурентними формулами:

$$\rho_{k+1} = \rho_k \frac{\sqrt{\sigma^2 - 1}}{\sigma}; \; \rho_0 = \rho_0\left(x^0, \sigma\right),$$

де $\sigma \ge \sqrt{2}$ – характеризує ступінь «витягнутості» гіперповерхонь рівня функції $f_0$.

Послідовність $\left\{x^k\right\}_{k=0}^{\infty}$, яка породжена алгоритмом 5, збігається до точки мінімуму функції $f_0$ зі швидкістю геометричної прогресії із знаменником $q = \sqrt{\sigma^2 - 1} / \sigma$.

***Алгоритм 5***

Початок. I. Вибрати довільне початкове наближення $x^0 \in R^n$.

II. Знайти число $\sigma \ge \sqrt{2}$ і значення крокового множника $\rho_0$, які задовольняють умовам теореми 5.

III. Покласти $k = 0$.

Основний цикл. IV. Обчислити вектор напрямку руху $h^k$ до наступного наближення $x^{k+1}$

$$h^k = -g\left(x^k\right) \Big/ \left\| g\left(x^k\right) \right\|.$$

V. Обчислити наступне наближення

$$x^{k+1} = x^k + \rho_k h^k .$$

VI. Обчислити значення крокового множника

$$\rho_{k+1} = \rho_k \sqrt{\sigma^2 - 1} / \sigma .$$

VII. Прокласти $k = k + 1$ і перейти на крок IV.

***Теорема 5.*** *Нехай виконуються припущення 5, і нехай числа $\sigma$ і $\rho_0$ задовольняють умовам:*

(***i***)– $\sigma \ge \sqrt{2}$;

(***ii***)– $\rho_0 \ge \left( \left\| x^0 - x^* \right\| / \sigma \right)$;

(***iii***) – *для будь-якої пари точок* $x, z \in Y$, $Y \triangleq \{x \mid \|x - x^*\| \le \sigma\rho_0\}$, *таких, що* $f_0(x) = f_0(z) \ne f_0(x^*)$ *виконується умова*

$$\left(\|x - x^*\| / \|z - x^*\|\right) \le \sigma .$$

*Тоді послідовність* $\{x^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 5, збігається до точки* $x^*$ *зі швидкістю геометричної прогресії*

$$\|x^k - x^*\| \le \rho_k \sigma, \quad k = 0,1,...,$$

*знаменник якої* $q = \sqrt{\sigma^2 - 1} / \sigma$ *(за виключенням випадку, коли для деякого* $k = \bar{k}$, $g\left(x^{\bar{k}}\right) = 0$, *тобто* $x^{\bar{k}} = x^*$*).*

#### 6. Метод, що використовує апріорне значення мінімуму функції

Алгоритм 6 оснований на апріорному знанні значення $f_0^*$ функції $f_0$ в точці мінімуму. На $k$-й ітерації за вектор руху $h^k$ до наступного наближення $x^{k+1}$ обирається узагальнений градієнт функції $f_0$ в точці $x^k$. Кроковий множник $\rho_k$ $(k = 0,1,...)$, обчислюється за формулою:

$$\rho_k = \gamma\left(f_0\left(x^k\right) - f_0^*\right) / \left\|g\left(x^k\right)\right\|^2 ,$$

де $\gamma$ – константа, яка задовольняє нерівності $0 < \gamma < 2$. Послідовність $\{x^k\}_{k=0}^{\infty}$, яка породжена алгоритмом 6, збігається до точки мінімуму $x^*$ із швидкістю геометричної прогресії із знаменником $q = \left(1 - \gamma(2-\gamma)\alpha^2 / \lambda^2\right)^{1/2} < 1$, де $\lambda$ та $\alpha$ – деякі константи.

***Алгоритм 6***

Початок. I. Вибрати довільне початкове наближення $x^0 \in R^n$ і константу $\gamma$, $0 < \gamma < 2$; покласти $k = 0$.

Основний цикл. II. Обчислити $f_0\left(x^k\right)$ та узагальнений градієнт $g\left(x^k\right)$ та покласти $h^k = -g\left(x^k\right)$.

III. Обчислити значення крокового множника

$$\rho_k = \gamma\left(f_0\left(x^k\right) - f_0^*\right) \Big/ \left\|h^k\right\|^2 .$$

IV. Обчислити наступне наближення

$$x^{k+1} = x^k + \rho_k h^k .$$

V. Покласти $k = k + 1$ і перейти на крок II.

***Теорема 6.*** *Нехай опукла функція* $f_0$ *має єдину точку мінімуму* $x^*$, *причому відоме значення* $f_0^*$ *функції* $f_0$ *в точці* $x^*$. *Тоді при* $0 < \gamma < 2$

*послідовність* $\{x^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 6, збігається до* $x^*$ *із будь-якого початкового наближення* $x^0 \in R^n$.

***Теорема 6'.*** *Нехай* $(\boldsymbol{i})-$ $f_0(x)$ *– сильноопукла функція, причому,*

$$f_0(x) - f_0(x^*) \geq \alpha \|x - x^*\|^2, \quad \alpha > 0;$$

$(\boldsymbol{ii})-$ *функція* $f_0(x)$ *задовольняє умову Ліпшиця в області*

$$Y \triangleq \left\{ x \middle| \left(\|x - x^*\|\right) \leq \left(\|x^0 - x^*\|\right) \right\},$$

*тобто для всіх* $x, y \in Y$ *виконується* $f_0(x) - f_0(y) \leq \lambda \|x - y\|$.

*Тоді при* $0 < \gamma < 2$ *послідовність* $\{x^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 6, збігається до точки мінімуму* $x^*$ *із швидкістю геометричної прогресії*

$$\|x^k - x^*\| \leq q^k \|x^0 - x^*\|$$

*із знаменником* $q = \left(1 - \gamma(2 - \gamma)\alpha^2 / \lambda^2\right)^{1/2} < 1$.

**7. Метод, що стійкий до похибок обчислень**

Алгоритм 7 вказує на стійкість методу узагальненого градієнтного спуску щодо малих похибок в обчисленні точок $x^k$, $k = 0,1,...$ і узагальнених градієнтів $g(x^k)$ в цих точках. На $k$-й ітерації за вектор руху $h^k$ до наступного наближення $x^{k+1}$ обирається одиничний вектор протилежний до узагальненого градієнта функції $f_0$ в точці $y^k$, який лежить в $\delta_k$-околі точки $x^k$. Послідовність $\{x^k\}_{k=0}^{\infty}$, яка породжена алгоритмом 7, збігається до точки $x^* \in X^*$,
де

$$X^* = \left\{ x \middle| f_0(x) = f_0^* \right\}; \; f_0^* = \min_{x \in R^n} f_0(x).$$

***Алгоритм 7***

П о ч а т о к . I. Вибрати довільне початкове наближення $x^0 \in R^n$ і покласти $k = 0$.

О с н о в н и й ц и к л . II. Обчислити кроковий множник $\rho_k$ і величину змішення $\delta_k$, які задовольняють умовам теореми 7.

III. Обчислити узагальнений градієнт $g(y^k)$ функції $f_0$ в будь-якій точці $y^k$, яка задовольняє нерівність

$$\|y^k - x^k\| \leq \delta_k.$$

IV. Обчислити вектор руху $h^k$ до наступного наближення $x^{k+1}$

$$h^k = -g\left(y^k\right) \Big/ \left\| g\left(y^k\right) \right\|.$$

V. Обчислити наступне наближення

$$x^{k+1} = x^k + \rho_k h^k.$$

VI. Покласти $k = k+1$ і перейти на крок II.

***Теорема 7.*** *Нехай* $f_0$ *– опукла донизу функція, яка має непорожню множину мінімумів* $X^*$. *Тоді, якщо числа* $\delta_k$, $\rho_k$, $k = 0,1,\ldots$, *вибрати таким чином, що:*

$$\rho_k > 0; \ \ \delta_k > 0; \ \ \lim_{k\to\infty} \delta_k = 0;$$

$$\sum_{k=0}^{\infty} \delta_k \rho_k < \infty; \ \ \sum_{k=0}^{\infty} \rho_k^2 < \infty; \ \ \sum_{k=0}^{\infty} \rho_k = \infty,$$

*то послідовність* $\left\{x^k\right\}_{k=0}^{\infty}$, *яка породжена алгоритмом 7, задовольняє граничному співвідношенню*

$$\lim_{k\to\infty} x^k = x^* \in X^*.$$

#### 8. Багатокроковий метод узагальненого градієнтного спуску



*Припущення 8.* Функція $f_0$ – опукла донизу.

В нижче наведеному алгоритмі напрямок спуску вибирається з використанням узагальнених градієнтів і значень функції $f_0$ на попередніх ітераціях. На кожній ітерації необхідно розв'язувати спеціальну задачу мінімізації, яка відповідним нормуванням зводиться до задачі лінійного програмування. Крокові множники $\rho_k$ задовольняють класичним умовам.

***Алгоритм 8***

Початок. I. Вибрати довільне натуральне число $m \geq 1$.

II. Вибрати довільний набір точок $\left\{x^{-m+1},\ldots,x^0\right\}$.

III. Вибрати константу $\alpha$, $\alpha > f_0\left(x^0\right)$.

IV. Покласти $k = 0$.

Основний цикл. V. Обчислити $\hat{\nabla} f_0\left(x^k\right)$ – узагальнений градієнт функції $f_0$ в точці $x^k$.

VI. Обчислити кроковий множник $\rho_k$, який задовольняє умовам теореми 8.

VII. Якщо $\hat{\nabla} f_0\left(x^k\right) = 0$, то покласти $x^* = x^k$ і припинити обчислення; інакше перейти на крок VIII.

VIII. Якщо $f_0\left(x^k\right) > \alpha$, то покласти $\Im_k = \{k\}$ і перейти на крок IX, якщо $f_0\left(x^k\right) \le \alpha$, то покласти $\Im_k = \overline{k-m+1,k}$ і перейти на крок IX.

IX. Обчислити вектор $h^k$ з умови:

$$\min_{\|h\| \le \rho_k} \varphi_k(h) = \varphi_k\left(h^k\right),$$

де функція

$$\varphi_k(h) \triangleq \max_{j \in \Im_k} \left[ f_0\left(x^j\right) + \left(x^k - x^j, \hat{\nabla} f_0\left(x^j\right)\right) + \left(h, \hat{\nabla} f_0\left(x^j\right)\right) \right].$$

X. Обчислити наступне наближення

$$x^{k+1} = x^k + h^k.$$

XI. Покласти $k = k+1$ і перейти на крок V.

***Теорема 8.*** *Нехай виконуються умови:* $(\boldsymbol{i})-$ *функція* $f_0$ *опукла донизу;* $(\boldsymbol{ii})-$ *множина*

$$X^* \triangleq \left\{ \bar{x} \,\middle|\, f_0(\bar{x}) = \inf_{x \in R^n} f_0(x) \right\}$$

*непорожня і обмежена;* $(\boldsymbol{iii})-$ *крокові множники* $\rho_k$, $k = 0,1,\ldots$ *задовольняють умовам:*

$$\rho_k \to +0 \text{ при } k \to \infty, \ \sum_{k=0}^{\infty} \rho_k = \infty.$$

*Тоді нескінченна послідовність* $\left\{x^k\right\}_{k=0}^{\infty}$, *яка породжена алгоритмом 8, задовольняє граничним співвідношенням:*

$$\min_{x \in X^*} \left\| x - x^k \right\| \to 0 \text{ при } k \to \infty;$$

$$f_0\left(x^k\right) \to \inf_{x \in R^n} f_0(x) \text{ при } k \to \infty.$$

*Зауваження 8.* Якщо множина $X^*$ містить внутрішні точки, то послідовність $\left\{x^k\right\}_{k=0}^{\infty}$, яка породжена алгоритмом 8, скінченна.

### 9. $\varepsilon$-субградієнтний метод

З а д а ч а 9. Знайти $\arg\min_{x \in R^n} f_0(x)$ для заданої функції $f_0 : R^n \to R^1$.

*Припущення 9.* $(\boldsymbol{i})-$ функція $f_0$ – опукла напівнеперервна знизу; $(\boldsymbol{ii})-$ $\inf_{x \in R^n} f_0(x) > -\infty$; $(\boldsymbol{iii})-$ принаймні в одній точці $\bar{x} \in R^n$ виконується умова $f_0(\bar{x}) < \infty$.

*Означення 9.* Для довільного $\varepsilon > 0$ вектор $g_\varepsilon(x) \in R^n$ називається $\varepsilon$-***субградієнтом функції*** $f_0$ в точці $x$, якщо $f_0(z) \geq f_0(x) - \varepsilon + (z - x, g_\varepsilon(x))$ для всіх $z \in R^n$. Множина $\varepsilon$-субградієнтів в точці $x$ позначається через $G_\varepsilon(x)$.

***Алгоритм 9***

Початок. I. Вибрати вектор $x^0 \in R^n$ такий, що $f_0(x^0) < \infty$.

II. Вибрати константи $\varepsilon_0 > 0$ та $0 < \alpha < 1$.

III. Покласти $k = 0$.

Основний цикл. IV. Обчислити:

$$\varepsilon_{k+1} = \alpha^i \varepsilon_k ,$$

де $i$ – найменше невід'ємне ціле число, при якому $0 \notin G_{\varepsilon_{k+1}}(x^k)$.

(Якщо $x^k$ не є точкою мінімуму функції $f_0$, то завжди існує невід'ємне ціле число $i$, при якому виконується включення $0 \notin G_{\alpha^i \varepsilon_k}(x^k)$).

V. Знайти вектор $h^k$, для якого виконується нерівність

$$\sup_{g \in G_{\varepsilon_{k+1}}(x^k)} (h^k, g) < 0 .$$

VI. Покласти $x^{k+1} = x^k + \rho_k h^k$, де $\rho_k > 0$ таке, що виконується нерівність $f_0(x^k) - f_0(x^{k+1}) > \varepsilon_{k+1}$.

VII. Покласти $k = k + 1$ і перейти на крок IV.

***Теорема 9.*** *Нехай виконуються припущення 9. Тоді або нескінченна послідовність* $\{x^k\}_{k=0}^{\infty}$ *задовольняє граничному співвідношенню*

$(\boldsymbol{iv})$ –

$$\lim_{k \to \infty} f_0(x^k) = \min_{x \in R^n} f_0(x),$$

*або* $f_0(x^m) = \min_{x \in R^n} f_0(x)$ *при деякому* $m > 0$. *Якщо, крім того, множина*

$$X^* \triangleq \left\{ x^* \middle| f_0(x^*) = \min_{x \in R^n} f_0(x) \right\}$$

*непорожня і обмежена, то:* $(\boldsymbol{v})$ – *кожна збіжна підпослідовність послідовності* $\{x^k\}_{k=0}^{\infty}$ *має границю в* $X^*$ *і хоча б одна така підпослідовність існує;* $(\boldsymbol{vi})$ – *при кожному* $\varepsilon > 0$ *існує* $\overline{m} \geq 0$ *таке, що* $x^k \in X^* + \varepsilon Y$ *при всіх* $k \geq \overline{m}$, *де* $Y \triangleq \{ x \,|\, \| x \| \leq 1 \}$; $(\boldsymbol{vii})$ – *якщо мінімум функції* $f_0$ *досягається в єдиній точці* $x^*$, *то* $\{x^k\}_{k=0}^{\infty}$ *прямує до* $x^*$.

## 4.4. Методи градієнтного типу з розтягненням простору

Задача 0. Знайти $\arg\min_{x \in R^n} f_0(x)$ для заданої майже диференційованої функції $f_0 : R^n \to R^1$.

*Означення 0.* Функція $f_0(x)$ називається ***майже диференційованою,*** якщо вона задовольняє наступним умовам:

а) в будь-якій обмеженій області задовольняє умову Ліпшиця;

б) майже всюди диференційована;

в) її градієнт неперервний на тій множині, де він визначений.

Відмітимо, що опукла функція майже всюди диференційована (це означає, що майже всюди, тобто за виключенням точок множини міри нуль, узагальнений градієнт співпадає з градієнтом).

*Означення 1.* ***Майже градієнтом*** функції $f_0(x)$ в точці $\bar{x}$ називається вектор, який являється граничною точкою деякої послідовності градієнтів $\nabla f_0(x^1),\ \nabla f_0(x^2),\ldots,$ де $\left\{x^k\right\}_{k=1}^{\infty}$ – така послідовність точок, що збігаються до точки $\bar{x}$, що у всіх точках цієї послідовності $f_0(x)$ диференційована.

*Означення 2.* Назвемо ***узагальненим майже градієнтом*** функції $f_0(x)$ в точці $\bar{x}$ довільний вектор $g$, що належить випуклому замиканню множини майже градієнтів.



Суть методів градієнтного типу з розтягненням простору полягає у побудові в процесі послідовних наближень лінійних операторів, що змінюють метрику простору, і виборі напрямку спуску, який відповідає антиградієнту в просторі з новою метрикою.

*Означення 3.* ***Оператором розтягнення*** простору $R^n$ у напрямку $\xi \in R^n$ $(\|\xi\| = 1)$ з коефіцієнтом $\alpha$ $(\alpha \ge 0)$, називається оператор $G_\alpha(\xi)$, що діє на вектор $x$, представлений у формі

$$x = \gamma_\xi(x)\xi + d_\xi(x); \quad (\xi, d_\xi(x)) = 0,$$

(тут $\gamma_\xi(x) = (x,\xi);$ $d_\xi(x) = x - (x,\xi)\xi$ ) наступним чином:

$$G_\alpha(\xi)x = \alpha\gamma_\xi(x)\xi + d_\xi(x) = x + (\alpha - 1)(x,\xi)\xi,$$

де $G_\alpha(\xi)$ – лінійний симетричний оператор.

Оператор $G_\beta(\xi) = G_{1/\alpha}(\xi)$ називається ***оператором «стиснення».***

Нижче наводяться алгоритми з розтягненням простору у напрямках: спочатку майже градієнта, а потім різниці двох послідовних майже градієнтів.

**1. Метод градієнтного типу з розтягненням простору в напрямку майже градієнта**

В алгоритмі 1 в $k$-й ітерації наступне наближення $x^{k+1}$ знаходять за формулою:

$$x^{k+1} = x^k - B_k \rho_k \xi^k,$$

де $\xi^k$ – одиничний вектор майже градієнта функції $\varphi_k(y) \triangleq f_0(B_k y)$, яка отримується з $f_0(x)$ при використанні лінійного перетворення простору $y = A_k x$; $B_k$ – оператор, обернений до результуючого оператора $A_k$ перетворення простору ( $A_k$ отримується в результаті послідовного застосування операторів розтягнення простору у напрямку нормованих майже градієнтів $\xi^0, \xi^1, ..., \xi^{k-1}$ з коефіцієнтами $\alpha_1, \alpha_2, ..., \alpha_k$: $A_k = G_{\alpha_k}(\xi^{k-1}) A_{k-1}$); $\rho_k$ – кроковий множник.

Оператори $B_{k+1}$ визначаються рекурентними співвідношеннями:

$$B_{k+1} = B_k G_{\beta_{k+1}}(\xi^k); \quad B_0 = I,$$

де $\beta_{k+1} \triangleq 1/\alpha_{k+1}$ – коефіцієнти «стиснення» простору.

***Алгоритм 1***

Початок. I. Вибрати початкове наближення $x^0 \in R^n$ і неособливу матрицю $B_0$ (можна вибрати $B_0 = I$, де $I$ – одинична $n \times n$-матриця); покласти $k = 0$.

Основний цикл. II. Обчислити майже градієнт $g(x^k)$ функції $f_0$ у точці $x^k$.

III. Якщо $g(x^k) = 0$, то покласти $x^* = x^k$ і закінчити обчислення, інакше перейти на крок IV.

IV. Обчислити оператор $B_k^T$, спряжений до оператора $B_k$.

V. Обчислити майже градієнт $\tilde{g}(y^k)$ функції $\varphi_k(y) \triangleq f_0(B_k y)$ у точці $y^k = B_k^{-1} x^k$

$$\tilde{g}(y^k) = B_k^T \, g(x^k).$$

VI. Обчислити напрямок розтягнення простору

$$\xi^k = \tilde{g}(y^k) / \| \tilde{g}(y^k) \| .$$

VII. Обчислити значення крокового множника $\rho_k$.

VIII. Обчислити наступне наближення

$$x^{k+1} = x^k - \rho_k B_k \xi^k.$$

IX. Знайти коефіцієнт розтягнення простору $\alpha_{k+1}$.

X. Обчислити коефіцієнт «стиснення» простору

$$\beta_{k+1} = 1/a_{k+1} \ .$$

XI. Обчислити оператор $B_{k+1}$, обернений до результуючого оператора $A_{k+1}$ перетворення простору, за формулою:

$$B_{k+1} = B_k G_{\beta_{k+1}}(\xi^k),$$

де оператор "стиснення" $G_{\beta_{k+1}}(\xi^k)$ обчислюється згідно означенню 3.

XII. Покласти $k = k+1$ і перейти на крок II.

***Теорема 1.*** *Нехай* $f_0(x)$ *– майже диференційована функція і* $x^*$ *– точка її локального мінімуму. Тоді, якщо в процесі реалізації алгоритму 1 виконуються умови:*

$(\boldsymbol{i})-\ \left\|x^k - x^*\right\| \le \sigma,\quad \sigma > 0,\quad k = 1,2,...;$

$(\boldsymbol{ii})-\ 1+\delta \le \alpha_k \le \overline{\alpha},\qquad \delta > 0,$

*то існує така підпослідовність* $\{x^{k_s}\}_{s=1}^{\infty}$ *та число* $\gamma > 0$, *що*

$$\|\ \tilde{g}(y^{k_s})\ \| \le \gamma \left( \prod_{j=1}^{k_s} \alpha_j \right)^{-\frac{1}{n}}, \qquad s = 1,2,...$$

***Теорема 1'.*** *Нехай* $f_0(x)$ *– майже диференційована функція, визначена в деякому околі* $S_\sigma$ *точки* $x^*$ *локального мінімуму, і в тих точках, де функція диференційована, її похідна* $f'_{0,\mu(x)}(x)$ *по напрямку* $\mu(x) = x - x^*$ *задовольняє нерівності:*

$$\lambda \le \frac{f'_{0,\mu(x)}(x)}{f_0(x) - f_0(x^*)} \le \lambda',\ \ \lambda' > \lambda > 0, \qquad x \ne x^*.$$

*Тоді, якщо в алгоритмі 1 прийняти:*

$(\boldsymbol{i})-\ x^0 \in S_\sigma;$

$(\boldsymbol{ii})-\ \rho_k = \dfrac{2\lambda\lambda'}{\lambda+\lambda'} \dfrac{f_0(x^k) - f_0(x^*)}{\left\|\tilde{g}(y^k)\right\|};\ \ \alpha_{k+1} \equiv \alpha = \dfrac{\lambda' + \lambda}{\lambda' - \lambda},$

*то знайдуться константа* $\gamma'$ *і підпослідовність індексів* $\{k_s\}_{s=1}^{\infty}$, $k_s < k_{s+1}$ *такі, що*

$$f_0(x^{k_s}) - f_0(x^*) \le \gamma' \alpha^{-\frac{k_s}{n}}, \qquad s = 1,2,...$$

*Зауваження 1.* Описаний в теоремі $1'$ спосіб вибору коефіцієнта розтягнення $\alpha_k$, $k = 1,2,...$ та крокового множника $\rho_k$, $k = 0,1,2,...$ знаходить безпосереднє застосування при розв'язуванні систем нелінійних рівнянь:

$$f_i(x) = 0, \qquad i = \overline{1,n},$$

шляхом зведення до задачі $\min\limits_{x} \max\limits_{i \in [1;n]} \left| f_i(x) \right|$.

В регулярному випадку при достатньо доброму початковому наближені $x^0$ константи $\lambda$ і $\lambda'$ можна вибирати близькими до одиниці, що забезпечує швидку збіжність.

При розв'язанні опуклих задач, в загальному випадку $f_0(x^*)$ невідоме, тому виникає питання про вибір $f_0(x^*)$ під час обчислень. Справедлива наступна теорема.

***Теорема 1''.*** *Нехай опукла донизу функція* $f_0(x)$ *має наступні властивості:* $(\boldsymbol{i})-$ *існує така стала* $\bar{\gamma} > 1$, *що якщо функція* $\varphi(\tau)$,

$$\varphi(\tau) \triangleq f_0((1-\tau)x^1 + \tau x^2), \quad 0 \le \tau \le 1,$$

*строго спадаюча по* $\tau$, *то виконується нерівність:*

$$f'_{0,(x^1-x^2)}(x^1) \le \bar{\gamma}(f_0(x^1) - f_0(x^2)),$$

*де* $f'_{0,(x^1-x^2)}(x^1)$ *– похідна функції* $f_0(x)$ *по напрямку* $(x^1 - x^2)$ *в точці* $x^1$;

$(\boldsymbol{ii})-$ $\lim\limits_{\|x\| \to \infty} f_0(x) = +\infty.$

*Тоді, якщо в алгоритмі 1 прийняти:*

$$\alpha_{k+1} = (\bar{\gamma}+1)/(\bar{\gamma}-1); \quad \rho_k = 2\bar{\gamma}(f_0(x^k) - \bar{f})/(\bar{\gamma}+1)\left\|\tilde{g}(y^k)\right\|,$$

*де* $\bar{f}$ *– більше або рівне* $f^* \triangleq \min\limits_{x \in R^n} f_0(x)$, *то послідовність* $\{\rho_k\}_{k=0}^{\infty}$ *є обмеженою і для будь-якого* $\varepsilon > 0$ *знайдеться таке* $\bar{k}$, *що*

$$f_0(x^{\bar{k}}) \le \bar{f} + \varepsilon$$

*(обчислення перериваються, якщо на деякому кроці* $f_0(x^{\bar{k}}) \le \bar{f}$*). Якщо* $\bar{f}$ *вибране меншим за* $f^*$, *то послідовність* $\{\rho_k\}_{k=0}^{\infty}$ *є необмеженою.*

*Зауваження 1'.* Теорема 1″ дозволяє будувати алгоритм мінімізації функції $f_0(x)$, що задовольняє умовам теореми, при невідомому $f^*$. Цей алгоритм пов'язаний з підбором $f^*$ з використанням наступних ознак: якщо при деякому $\bar{f}$ кроковий множник $\rho_k$ перевищує наперед задане достатньо велике число, то $\bar{f}$ збільшується; при наближені $f_0(x^k)$ до $\bar{f}$ – значення $\bar{f}$ зменшується.

***Теорема 1'''.*** *Нехай функція* $f_0$ *опукла донизу в* $R^n$ *і в процесі застосування алгоритму 1 виконуються наступні умови:*

$(\boldsymbol{i})-$ $\left\|g(x^k)\right\| \le \sigma, \qquad k = 0,1,...;$

$(\boldsymbol{ii})-$ $1 < \alpha \equiv \alpha_k; \qquad \beta_k = 1/\alpha_k.$

*Тоді для послідовності* $\{x^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 1, справедливі наступні нерівності:*

$$\min_{r\in[1;k]} \left\| \tilde{g}(y^r) \right\| \le \sigma \sqrt{k(\alpha^2 - 1)} \,/\, \sqrt{\alpha^{2k/n} - 1}\,,\ k = 1,2,\dots .$$

**2. Методи градієнтного типу з розтягненням простору в напрямку різниці двох послідовних майже градієнтів ($r(\alpha)$ - алгоритм)**

***Алгоритм 2***

Початок. I. Вибрати довільне початкове наближення $x^0 \in R^n$, коефіцієнт розтягнення простору $\alpha > 1$ і покласти $\tilde{g}^0 = 0$ ($0 -$ $n$-вимірний нуль-вектор), $B_0 = I$ ($I -$ одинична матриця), $k = 0$.

Основний цикл. II. Обчислити майже градієнт $g(x^k)$ функції $f_0$ в точці $x^k$; якщо він визначений неоднозначно, то взяти такий майже градієнт, що $(B_k \tilde{g}^k, g(x^k)) \le 0$.

III. Обчислити вектор

$$\bar{g}^k = B_k^T g(x^k).$$

Зауваження: вектор $\bar{g}^k$ є майже градієнтом функції $\varphi_k(y) \triangleq f_0(B_k y)$ в точці $y = A_k x^k$, де $A_k \triangleq B_k^{-1}$ – оператор розтягнення простору після $k$-ї ітерації.

IV. Обчислити різницю двох майже градієнтів від функції $\varphi_k(y)$, обчислених в точках $y^k = A_k x^k$, $\tilde{y}^k = A_k x^{k-1}$, за формулою

$$r^k = \bar{g}^k - \tilde{g}^k .$$

V. Обчислити напрямок розтягнення простору

$$\xi^k = r^k / \left\| r^k \right\|.$$

VI. Обчислити оператор $B_{k+1}$, обернений результуючому оператору $A_{k+1}$ перетворення простору після $k+1$-ї ітерації:

$$B_{k+1} = B_k G_\beta(\xi^k),$$

де $\beta = 1/\alpha$ – коефіцієнт «стиснення» простору.

VII. Обчислити майже градієнт $\tilde{g}^{k+1}$ функції $\varphi_{k+1}(y) \triangleq f_0(B_{k+1} y)$ в точці $y = B_{k+1}^{-1} x^k$:

$$\tilde{g}^{k+1} = G_\beta(\xi^k) \cdot \bar{g}^k = B_{k+1}^T g(x^k).$$

VIII. Обчислити значення крокового множника $\rho_k$ з умови

$$\rho_k = \arg \overline{\min_{\rho \ge 0}}\, f_0(x^k - \rho B_{k+1} \tilde{g}^{k+1}) \qquad (4.4)$$

(операція $\min$ означає визначення найближчого до нуля локального мінімуму).

IX. Обчислити наступне наближення

$$x^{k+1} = x^k - \rho_k B_{k+1} \widetilde{g}^{k+1}.$$

*Зауваження 2.* Крок IX фактично реалізує крок найшвидшого спуску для функції $\varphi_{k+1}(y) \triangleq f_0(B_{k+1} y)$:

$$y^{k+1} = \widetilde{y}^k - \rho_k \widetilde{g}^{k+1}.$$

X. Покласти $k = k+1$ і перейти на крок II.

***Теорема 2.*** *Якщо виконуються умови: $(\boldsymbol{i})-$ функція $f_0$ – неперервна і кусково-гладка; $(\boldsymbol{ii})-$ $\lim\limits_{\|x\|\to\infty} f_0(x) = +\infty$; $(\boldsymbol{iii})-$ функція $f_0$ володіє властивістю: в довільній скінченій області $X$ для довільного $\delta > 0$ знайдеться $\varepsilon > 0$, що якщо $x, y \in X$, $f_0(x) - f_0(y) < \varepsilon$ і на відрізку між $x$ та $y$ функція $f_0$ спадає, то $\|x - y\| < \delta$ (для опуклих донизу функцій ця властивість є аналогом сильної опуклості донизу), то граничною точкою послідовності $\{x^k\}_{k=0}^{\infty}$, яка породжена алгоритмом 2 при довільному $\alpha > 1$, є деяка точка $\bar{x}$, множина майже градієнтів якої утворює лінійно залежне сімейство векторів. При цьому, в силу монотонності процесу спуску $\left(f_0(x^0) \geq f_0(x^1) \geq \ldots \geq f_0(x^k) \geq \ldots\right)$, послідовність $\{f_0(x^k)\}_{k=0}^{\infty}$ збігається до $f_0(\bar{x})$.*

***Теорема 2′.*** *Якщо виконуються всі умови теореми 1 і $x^*-$ ізольована точка локального мінімуму, $x^0$ – така точка, що зв'язна компонента множини $\left\{x \mid f_0(x^*) \leq f_0(x) \leq f_0(x^0)\right\}$, що містить точки $x^*$, $x^0$, не має крім $x^*$ інших точок $z$, у яких сімейство майже градієнтів $\hat{G}(z)$ лінійно залежне, то послідовність $\{x^k\}_{k=0}^{\infty}$, яка породжена алгоритмом 1 (з початковою точкою $x^0$, яка задовольняє умовам теореми), збігається до $x^*$.*

*Зауваження 2′.* На відміну від інших градієнтних методів мінімізації негладких функцій (узагальненого градієнтного спуску з розтягненням простору в напрямку майже градієнтів) $r(\alpha)$ – алгоритм забезпечує монотонність спуску за визначеним по формулі (4.4) способом вибору крокового множника: $f_0(x^0) \geq f_0(x^1) \geq \ldots \geq f_0(x^k)$. Він суттєво відрізняється від звичайних релаксаційних методів наступним: якщо $\rho_k = 0$, то це не значить, що процес спуску закінчується. При виконанні серії ітерацій з нульовим кроком точка $x^k$ стоїть на місці, але змінюються $B_k$ і $\widetilde{g}^k$. В цей час, ніби у прихованій формі, здійснюється пошук необхідного напрямку спуску.

## 4.5. Методи локального випадкового пошуку

Задача 0. Знайти $\arg\min_{x \in R^n} f_0(x)$ для заданої неперервної функції $f_0$: $R^n \to R^1$.

### 1. Метод локального випадкового пошуку з парною пробою

***Алгоритм 1***

Початок. I. Вибрати довільне початкове наближення $x^0 \in R^n$, пробний кроковий множник $\gamma > 0$ та робочий кроковий множник $\rho > 0$; покласти $k = 0$.

Основний цикл. II. Обчислити незалежну реалізацію $\xi^k$ випадкового одиничного вектора $\xi$, що рівномірно розподілений за всіма напрямками простору параметрів $x$.

III. Обчислити вектор руху $h^k$ до наступного наближення $x^{k+1}$

$$h^k = \xi^k sign[f_0(x^k - \gamma\xi^k) - f_0(x^k + \gamma\xi^k)].$$

IV. Обчислити наступне наближення

$$x^{k+1} = x^k + \rho h^k .$$

V. Покласти $k = k + 1$ і перейти на крок II.

*Зауваження 1*. Характерною особливістю даного алгоритму є його підвищена тенденція до «блукання», навіть у тому випадку, якщо розв'язок задачі 0 знайдений. Проте, алгоритм 1 приводить з великою ймовірностю у $\varepsilon$-окіл точок локального мінімуму, що залежить від кроку $\rho$, розмірності простору $R^n$ та вигляду функції $f_0$. (Наприклад, у центральному полі, тобто коли $f_0(x) = x - x^*$, для того, щоб попасти у $\varepsilon$-окіл точки $x^*$ з великою ймовірністю, необхідно кроковий множник $\rho$ вибирати з умови $\rho \le \varepsilon / \alpha_n^*$, де при $n = 2$ $\alpha_n^* = 0{,}7$; при $n = 3$ $\alpha_n^* = 0{,}9$; при $n = 4$ $\alpha_n^* = 1{,}0$).

### 2. Метод локального випадкового пошуку з поверненням при невдалому кроці

***Алгоритм 2***

Початок. I. Вибрати довільне початкове наближення $x^0 \in R^n$ , кроковий множник $\rho > 0$; покласти $k = 0$.

Основний цикл. II. Обчислити незалежну реалізацію $\xi^k$ випадкового одиничного вектора $\xi$, що рівномірно розподілений за всіма напрямками простору $R^n$.

III. Обчислити вектор $z$ за формулою $z = x^k + \rho\xi^k$.

IV. Якщо $f_0(z) < f_0(x^k)$, то покласти $x^{k+1} = z$ та перейти на крок V; інакше покласти $x^{k+1} = x^k$ та перейти на крок V.

V. Покласти $k = k+1$ та перейти на крок II.

*Зауваження 2*. Алгоритм 2 (як і алгоритм 1) може бути рекомендований для оптимізації систем, функція якості котрих змінюється з часом порівняно з великою швидкістю.

#### 3. Метод локального випадкового пошуку з лінійною екстраполяцією

Суть такого методу пошуку зводиться до наступного. Після невдалого випадкового кроку виконується подвійний крок в протилежному напрямку, але функція цілі $f_0$ у цьому стані не обчислюється, а екстраполюється в припущенні про лінійний характер функції цілі $f_0$.

***Алгоритм 3***

Початок. I. Вибрати довільне початкове наближення $x^0 \in R^n$, кроковий множник $\rho > 0$; обчислити $f_0(x^0)$; покласти $k = 0$.

Основний цикл. II. Обчислити незалежну реалізацію $\xi^k$ випадкового одиничного вектора $\xi$, що рівномірно розподілений за всіма напрямками простору $R^n$.

III. Обчислити вектор $z = x^k + \rho\xi^k$.

IV. Обчислити $f_0(z)$ – значення функції $f_0$ в точці $z$; якщо $f_0(z) < f_0(x^k)$, то покласти $x^{k+1} = z$, $f_0(x^{k+1}) = f_0(z)$ та перейти на крок VII; інакше перейти на крок V.

V. Обчислити вектор $x^{k+1} = x^k - \rho\xi^k$.

VI. Обчислити «значення» функції $f_0$ в точці $x^{k+1}$

$$f_0(x^{k+1}) = 2f_0(x^k) - f_0(z).$$

VII. Покласти $k = k+1$ та перейти на крок II.

*Зауваження 3*. Алгоритм 3 володіє підвищеною швидкодією порівняно з алгоритмами 1 та 2, тому що він використовує несприятливі кроки, проте він погано працює у нелінійній задачі.

#### 4. Метод випадкового пошуку по найкращій пробі з накопиченням

Суть методу полягає у наступному. У $k$-й ітерації з точки $x^k$ робиться $m$ $(m \geq 1)$ пробних кроків за незалежними випадковими напрямками $\xi^{k,j}$, $j = \overline{1,m}$, та визначається найкращий напрямок $\xi^{k,j^*}$. Робочий крок робиться саме в цьому напрямку.

***Алгоритм 4***

Початок. I. Вибрати довільне початкове наближення $x^0 \in R^n$, пробний кроковий множник $\gamma > 0$ та робочий кроковий множник $\rho > 0$, число пробних кроків $m \geq 1$; покласти $k = 0$.

Основний цикл. II. Обчислити $m$ незалежних реалізацій $\xi^{k,i}$, $i = \overline{1,m}$, випадкового одиничного вектора $\xi$, що рівномірно розподілений за всіма напрямками простору $R^n$.

III. Обчислити індекс $i^*$, що задовольняє умову

$$f_0(x^k + \gamma\xi^{k,i^*}) = \min_{i \in [1;m]} f_0(x^k + \gamma\xi^{k,i}).$$

IV. Обчислити наступне наближення

$$x^{k+1} = x^k + \rho\xi^{k,i^*}.$$

V. Покласти $k = k + 1$ і перейти на крок II.

*Зауваження 4*. Із збільшенням числа пробних кроків напрямок вектора руху $\xi^{k,j^*}$ наближається до напрямку, протилежного до градієнта $\nabla f_0(x^k)$, та у границі при $m \to \infty$, $\gamma \to 0$ збігається з ним.

**5. Метод статистичного градієнта**

***Алгоритм 5***

Початок. I. Вибрати довільну початкову точку $x^0 \in R^n$, пробний кроковий множник $\gamma > 0$, робочий кроковий множник $\rho > 0$, число пробних кроків $m \geq 1$; покласти $k = 0$.

Основний цикл. II. Обчислити $m$ незалежних реалізацій $\xi^{k,i}$, $i = \overline{1,m}$, випадкового одиничного вектора $\xi$, що рівномірно розподілений за всіма напрямками простору $R^n$.

III. Обчислити прирости функції $f_0$

$$\Delta_i = f_0(x^k + \gamma\xi^{k,i}) - f_0(x^k), \quad i = \overline{1,m}.$$

IV. Обчислити вектор $g^k = \sum_{i=1}^{m} (\xi^{k,i}\Delta_i)$.

V. Обчислити вектор руху $h^k$ до наступного наближення $x^{k+1}$ за формулою

$$h^k = -g^k / \|g^k\|.$$

VI. Обчислити наступне наближення $x^{k+1} = x^k + \rho h^k$.

VII. Покласти $k = k + 1$ та перейти на крок II.

*Зауваження 5.* Вектор $g^k$ з’являється статистичною оцінкою градієнтного напрямку функції $f_0$ в точці $x^k$ (при $m \to \infty$, $\gamma \to 0$ вектор $g^k$ наближається до напрямку градієнта $\nabla f_0(x^k)$).

## 4.6. Квазіградієнтні методи

Задача 0. Знайти $\arg\min_x f_0(x)$ для слабоопуклої донизу функції $f_0 : R^n \to R^1$.

*Означення 1.* Функція $f_0$ називається ***слабоопуклою донизу***, якщо для кожного $x$ з довільної замкнутої обмеженої множини існує непорожня множина $G(x)$ векторів $\tilde{\nabla} f_0(x)$ таких, що для всіх $z$ і $\tilde{\nabla} f_0(x) \in G(x)$:

$$f_0(z) - f_0(x) \geq \left(\tilde{\nabla} f_0(x), z - x\right) + r(z, x) \qquad (4.5)$$

та $r(x, y)\|x - y\|^{-1} \to 0$ рівномірно по $x$ при $y \to x$.

*Означення 2.* Вектор $\tilde{\nabla} f_0(x)$, що задовольняє нерівності (4.5) називається ***квазіградієнтом*** слабоопуклої функції $f_0$ в точці $x$.

Слабоопуклими функціями є диференційовані, а також опуклі функції (не обов'язково диференційовані). У першому випадку квазіградієнтом є звичайний градієнт, а в другому – узагальнений градієнт. Клас слабоопуклих функцій являється замкнутим щодо операції взяття максимуму, тобто якщо $f(x, y)$ – слабоопуклі при кожному значенні $y$ функції, то

$$f_0(x) \triangleq \max_{y \in Y} f(x, y) = f(x, y(x))$$

є слабоопуклою функцією та

$$\tilde{\nabla} f_0(x) = \tilde{\nabla}_x f(x, y)\Big|_{y = \arg\max_{y \in Y} f(x, y)}.$$

#### 1. Квазіградієнтний метод мінімізації слабоопуклої донизу функції

Задача 1. Знайти $\arg\min_x f_0(x)$ для слабоопуклої донизу функції $f_0 : R^n \to R^1$.

Множину розв’язків $X^*$ для цієї задачі визначимо рівністю $X^* = \{x | 0 \in G(x)\}$, де $G(x)$ – множина квазіградієнтів функції $f_0$ в точці $x$.

Сутність даного методу полягає в побудові на $k$-й ітерації квазіградієнту слабоопуклої функції в точці $x^k$. Якщо рух у напрямку квазіградієнту виводить за межі спеціальним чином побудованої множини

$S$, то процес побудови такої послідовності переривається, і алгоритм починає працювати з довільної точки деякої підмножини $A \subset S$.

***Алгоритм 1***

Початок. I. Вибрати довільне початкове наближення $x^0$, константу $\delta > 0$, кроковий множник $\rho_0$.

II. Побудувати множину $S$ за правилом:

$$S = Q\left(f_0\left(x^0\right) + \delta\right),$$

де $Q(\alpha) \triangleq \{x | f_0(x) \le \alpha\}$.

III. Задати довільну компактну підмножину $A$ множини $S$.

IV. Покласти $k = 0$.

Основний цикл. V. Обчислити квазіградієнт $\tilde{\nabla} f_0\left(x^k\right)$ функції $f_0$ в точці $x^k$.

VI. Обчислити вектор:

$$x^{k+1} = \begin{cases} x^k - \rho_k \tilde{\nabla} f_0\left(x^k\right), & \text{якщо } x^k + \rho_k \tilde{\nabla} f_0(x^k) \in S, \\ y \in A, & \text{якщо } x^k + \rho_k \tilde{\nabla} f_0(x^k) \notin S, \end{cases}$$

де $y$ – довільна точка множини $A$.

VII. Обчислити кроковий множник $\rho_{k+1}$.

VIII. Покласти $k = k + 1$ і перейти на крок V.

***Теорема 1.*** *Якщо виконані умови:* $(\boldsymbol{i})-$ $X^* \triangleq \{x | 0 \in G(x)\}$ *компактна;* $(\boldsymbol{ii})-$ $Q(\alpha) \triangleq \{x | f_0(x) \le \alpha\}$ *компактна для кожного* $\alpha$; $(\boldsymbol{iii})-$ *функція* $f_0$ *приймає на* $X^*$ *скінчене число значень;* $(\boldsymbol{iv})-$ *крокові множники задовольняють умовам:*

$$\rho_k > 0;\ \rho_k \to 0;\ \rho_{k+1}/\rho_k \to 1;\ \sum_{k=0}^{\infty} \rho_k = \infty,$$

*то будь-яка гранична точка послідовності* $\{x^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 1, належить множині розв'язків* $X^*$ *задачі 1.*

#### 2. Стохастичний квазіградієнтний метод мінімізації слабоопуклої функції

***Алгоритм 2***

Початок. I. Вибрати довільну константу $\alpha \in (0;\infty)$ і довільну замкнуту множину $B$, яка знаходиться у сфері радіуса $\alpha$, тобто $B \subset S \triangleq \{x | \|x\| \le \alpha\}$, вибрати довільну початкову точку $x^0 \in R^n$.

II. Задати правило формування послідовності крокових множників $\{\rho_k\}_{k=0}^{\infty}$.

III. Покласти $k = 0$, знайти $\rho_0$.

Основний цикл. IV. Обчислити випадковий вектор $\xi^k$, умовне математичне сподівання якого:

$$E\left(\xi^k / x^0,...,x^k\right) = \tilde{\nabla} f_0\left(x^k\right) + b^k ,$$

де $b^k$ – вектор, вимірний відносно $\sigma$-підалгебри $\mathcal{B}_k$, індукованої випадковими величинами $\left(x^0, x^1, \ldots, x^k\right)$; $\tilde{\nabla} f_0\left(x^k\right)$ – квазіградієнт слабоопуклої функції.

V. Обчислити вектор:

$$x^{k+1} = \begin{cases} x^k - \rho_k \xi^k, & \text{якщо } \| x^k \| \le \alpha, \\ z^{k+1} \in B, & \text{якщо } \| x^k \| > \alpha, \end{cases}$$

де $z^{k+1}$ – довільна точка множини $B$.

VI. Знайти $\rho_{k+1}$.

VII. Покласти $k = k + 1$ і перейти на крок IV.

***Теорема 2.*** *Нехай $f_0$ слабоопукла донизу функція і, крім того, виконуються умови:*

$(\boldsymbol{i})$ – $\max\limits_{x \in B} f_0(x) < \inf\limits_{\|x\| > \alpha} f_0(x)$;

$(\boldsymbol{ii})$ – $\left\|\xi^k\right\| + \left\|\tilde{\nabla} f_0\left(x^k\right)\right\| + \left\|b^k\right\| \le \beta$;

$(\boldsymbol{iii})$ – *крокові множники $\rho_k$, вимірні відносно $\sigma$-підалгебри $\mathcal{B}_k$, і такі, що*

$$\rho_k > 0; \;\; \rho_{k+1} / \rho_k \to 1; \;\; \sum_{k=0}^{\infty} \rho_k = \infty;$$

$$\sum_{k=0}^{\infty} \rho_k \left\|b^k\right\| < \infty \text{ м. н., } \sum_{k=0}^{\infty} E \rho_k^2 < \infty .$$

*Тоді майже при кожному $\omega$ граничні точки послідовності $\left\{x^k(\omega)\right\}_{k=0}^{\infty}$, породженої алгоритмом 2, належать множині $X^* \triangleq \left\{x \middle| 0 \in G(x)\right\}$, і послідовність $\left\{f_0\left(x^k(\omega)\right)\right\}_{k=0}^{\infty}$ збігається ($G(x)$ – множина квазіградієнтів функції $f_0$ в точці $x$).*

## 4.7. $\varepsilon$ – квазіградієнтні методи

### 1. $\varepsilon$ – квазіградієнтний метод мінімізації опуклих функцій

Задача 1. Знайти arg $\min\limits_{x \in R^n} f_0(x)$, $f_0 : R^n \to R^1$ – опукла донизу функція.

За напрямок руху до наступного наближення $x^{k+1}$ в даному методі обирають $\varepsilon-$квазіградієнт опуклої функції $f_0$ у точці $x^k$.

***Алгоритм 1***

Початок. I. Вибрати довільне початкове наближення $x^0 \in R^n$, кроковий множник $\rho_0$ та число $\varepsilon_0$, що задовольняють умовам теореми 1, і деяку константу $\alpha \geq 1$.

II. Покласти $k = 0$.

Основний цикл. III. Покласти $\varepsilon = \varepsilon_k$.

IV. Обчислити $\varepsilon-$квазіградієнт $g_\varepsilon(x^k)$ функції $f_0$ у точці $x^k$.

V. Обчислити нормуючий множник

$$\gamma_k = \begin{cases} 1, & \text{якщо } \left\| g_\varepsilon(x^k) \right\| \leq \alpha, \\ \left\| g_\varepsilon(x^k) \right\|^{-2}, & \text{якщо } \left\| g_\varepsilon(x^k) \right\| > \alpha. \end{cases}$$

VI. Обчислити наступне наближення

$$x^{k+1} = x^k - \rho_k \gamma_k g_\varepsilon(x^k).$$

VII. Обчислити кроковий множник $\rho_{k+1}$ та число $\varepsilon_{k+1}$, що задовольняють умовам теореми 1.

VIII. Покласти $k = k+1$ та перейти на крок III.

***Теорема 1.*** *Нехай функція $f_0$ опукла донизу і множина $X^* \triangleq \{ x | 0 \in G(x,0) \}$ (тут $G(x,\varepsilon)$ $(\varepsilon \geq 0)$) – множина $\varepsilon-$квазіградієнтів функції $f_0$ у точці $x$) компактна. Тоді, якщо числові послідовності $\{\rho_k\}_{k=0}^{\infty}$ і $\{\varepsilon_k\}_{k=0}^{\infty}$ є такими, що:*

$$\rho_k \to +0, \ \ \varepsilon_k \to +0 \ \ \textit{при} \ \ k \to \infty \ \ \textit{і} \ \ \sum_{k=0}^{\infty} \rho_k = \infty,$$

*то границя будь-якої збіжної підпослідовності послідовності $\{x^k\}_{k=0}^{\infty}$, яка породжена алгоритмом 1, належить множині X*.*

***Теорема 1'.*** *Нехай функція $f_0$ задовольняє умовам теореми 1 і $\varepsilon_k = const > 0$, $k = 0,1,\ldots$. Тоді, якщо числова послідовність $\{\rho_k\}_{k=0}^{\infty}$ є такими, що:*

$$\rho_k \to +0 \ \ \textit{при} \ \ k \to \infty, \ \sum_{k=0}^{\infty} \rho_k = \infty,$$

*то існує підпослідовність $\{x^{k_s}\}_{s=0}^{\infty}$ послідовності $\{x^k\}_{k=0}^{\infty}$, яка породжена алгоритмом 1, яка збігається до множини $X_\varepsilon^* \triangleq \left\{ x \middle| f_0(x) \leq \min_{x'} f_0(x') + \varepsilon \right\}$.*

**2. $\varepsilon -$квазіградієнтний метод мінімізації слабоопуклих функцій**

Задача 2. Знайти $\arg\min\limits_{x \in R^n} f_0(x)$, для заданої слабоопуклої функції $f_0 : R^n \to R^1$.

В даному ітераційному методі наступне $(k+1)$-е наближення $x^{k+1}$ вибирають у напрямку $\varepsilon -$квазіградієнта $g_\varepsilon(x^k)$ для функції $f_0$ у точці $x^k$ ($k$-му наближенні).

*Означення 2.* Вектор $g_\varepsilon(x)$ називають ***ε-квазіградієнтом слабоопуклої*** функції $f_0$ у точці $x\ (\varepsilon > 0)$, якщо для всіх $z \in R^n$ виконується нерівність

$$f_0(z) - f_0(x) \ge (g_\varepsilon(x), z - x) + r(x, z) - \varepsilon$$

і відношення $r(x,z)/\|z - x\|$ прямує до нуля рівномірно по $x$ при $z \to x$ (у кожній компактній підмножині з $R^n$).

Якщо функція $f_0$ визначається рівністю

$$f_0(x) = \max_{y \in Y} \varphi(x, y)\ ,$$

де множина $Y -$ компактна, а $\varphi(x, y) -$ неперервна за сукупністю змінних та рівномірно слабоопукла по $x$ для кожного $y$, то $\varepsilon -$квазіградієнтом функції $f_0$ у точці $\bar{x}$ є квазіградієнт функції $\varphi(x, \bar{y})$ по $x$ в точці $x = \bar{x}$, де $\bar{y}$ – довільна точка з множини

$$Y_\varepsilon^*(\bar{x}) \triangleq \left\{ y^* \middle| \phi(\bar{x}, y^*) \ge \phi(\bar{x}, y) - \varepsilon, \forall y \in Y \right\}.$$

***Алгоритм 2***

Початок. I. Вибрати довільне початкове наближення $x^0 \in R^n$, деяку константу $\delta > 0$, кроковий множник $\rho_0$ та величину $\varepsilon_0$, які задовольняють умовам теореми 2.

II. Покласти $k = 0$.

Основний цикл. III. Якщо $f_0(x^k) > f_0(x^0) + \delta$, то покласти $x^{k+1} = x^0$ і перейти на крок VI; інакше перейти на крок IV.

IV. Обчислити $\varepsilon_k$–квазіградієнт $g_{\varepsilon_k}(x^k)$ функції $f_0$ у точці $x^k$.

V. Обчислити наступне наближення

$$x^{k+1} = x^k - \rho_k g_{\varepsilon_k}(x^k).$$

VI. Обчислити кроковий множник $\rho_{k+1}$ та величину $\varepsilon_{k+1}$, що задовольняють умовам теореми 2.

VII. Покласти $k = k+1$ і перейти на крок III.

***Теорема 2.*** *Нехай функція $f_0$ слабоопукла і виконані умови: (**i**)$-$ множини $\{x | f_0(x) \le \alpha\}$ компактні при всіх $\alpha$; (**ii**)$-$ множина*

$X^* \triangleq \{x \,|\, 0 \in G(x,0)\}$ *компактна;* $(\boldsymbol{iii})-$ *числові послідовності* $\{\rho_k\}_{k=0}^{\infty}$, $\{\varepsilon_k\}_{k=0}^{\infty}$ *такі, що:*

$$\rho_k \to +0,\ \rho_{k+1}/\rho_\beta \to 1,\ \rho_{k+1}/\rho_k \to 1,\ \varepsilon_k \to +0,$$

$$\text{при } k \to \infty;\ \sum_{k=0}^{\infty}\rho_k = \infty;\ \sum_{k=0}^{\infty}\varepsilon_k < \infty.$$

*Тоді границя будь-якої збіжної підпослідовності послідовності* $\{x^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 2, належить множині X*.*

*Зауваження 2.* Для того, щоб уникнути обчислення на кожній ітерації значення функції $f_0(x^k)$, крок III алгоритму замінюють на III′.

III′. Якщо $\|x^k\| > \beta$, то покласти $x^{k+1} = x^0$ та перейти на крок VI; інакше перейти на крок IV (тут константу $\beta$ обирають з умови $f_0(x) > f_0(x^0) + \delta$ при $\|x\| > \beta$).

## 4.8. Методи узагальнених майже градієнтів для мінімізації функцій, які задовольняють локальну умову Ліпшиця



**1. Детермінований випадок**

З а д а ч а 1. Знайти $\arg\min\limits_{x\in R^n} f_0(x)$ для функції $f_0 : R^n \to R^1$, яка задовольняє локальну умову Ліпшиця.

Якщо існує практично ефективний спосіб обчислення узагальненого майже градієнта $\hat{g}(x)$ функції $f_0(x)$ у будь-якій точці $x \in R^n$, то за напрямок руху до наступного наближення $x^{k+1}$ в нижче наведених алгоритмах вибирається вектор $\xi^k$, який дорівнює узагальненому майже градієнтові функції $f_0$ у точці $\tilde{x}^k$, рівномірно розподіленій в $n$-вимірному кубі з центром у точці $x^k$ зі стороною $\alpha_k > 0$, тобто

$$\xi^k = \hat{g}(\tilde{x}^k). \tag{4.6}$$

В інших випадках за напрямок руху $\xi^k$ вибирається скінченно-різницевий аналог узагальненого майже градієнта $\hat{g}(\tilde{x}^k)$:

$$\xi^k = \frac{1}{\alpha_k}\sum_{i=1}^{n}\Big(f_0(\tilde{x}_1^k,\ldots,x_i^k + \alpha_k/2,\ldots,\tilde{x}_n^k) - \\ - f_0(\tilde{x}_1^k,\ldots,x_i^k - \alpha_k/2,\ldots,\tilde{x}_n^k)\Big)e^i, \tag{4.7}$$

де $e^i$, $i = \overline{1,n}$ $-$ $i$-й орт.

***Алгоритм 1***

П о ч а т о к . I. Вибрати довільне початкове наближення $x^0 \in R^n$, кроковий множник $\rho_0$ і величину зміщення $\alpha_0$, які задовольняють умовам теореми 1.

II. Покласти $k = 0$.

О с н о в н и й ц и к л . III. Обчислити реалізацію $\tilde{x}^k$ випадкової точки, рівномірно розподіленої в $n$-вимірному кубі з центром у точці $x^k$ зі стороною $\alpha_k$.

IV. Обчислити вектор $\xi^k$ за формулою (4.6) або (4.7).

V. Обчислити наступне наближення

$$x^{k+1} = x^k - \rho_k \xi^k .$$

VI. Обчислити значення крокового множника $\rho_{k+1}$ і величину зміщення $\alpha_{k+1}$, які задовольняють умовам теореми 1.

VII. Покласти $k = k+1$ і перейти на крок III.

***Теорема 1.*** *Нехай функція $f_0$ в будь-якій обмеженій області задовольняє умову Ліпшиця і, крім того, виконуються умови*

$$\rho_k > 0,\ \ \alpha_k > 0,\ \ k = 0,1,\ldots;\ \ \sum_{k=0}^{\infty} \rho_k = \infty;\ \ \sum_{k=0}^{\infty} \rho_k^2 < \infty;$$

$$\rho_k / \alpha_k \to 0,\ \ (\alpha_k - \alpha_{k+1}) / \rho_k \to 0,\ \ \alpha_k \to 0 \ \ \text{при}\ k \to \infty .$$

*Тоді, якщо послідовність $\left\{x^k\right\}_{k=0}^{\infty}$, яка породжена алгоритмом 1, належить обмеженій множині простору $R^n$, то з ймовірністю 1 граничні точки послідовності $\left\{x^k\right\}_{k=0}^{\infty}$ належать множині $X^* \triangleq \left\{x^* \mid 0 \in G(x^*)\right\}$ і послідовність $\left\{f_0(x^k)\right\}_{k=0}^{\infty}$ майже напевно збігається.*

Тут і далі $G(x)-$ множина узагальнених майже градієнтів функції $f_0$ в точці $x$.

#### 2. Стохастичний випадок

З а д а ч а 2 . Знайти $\arg\min\limits_{x} Ef_0(x,\omega)$ для заданої функції $f_0 : R^n \times \Omega \to R^1$.

*Припущення 2.* Функція $f_0$ задовольняє локальну умову Ліпшиця

$$| f_0(x,\omega) - f_0(y,\omega) | \le \gamma(\omega) \| x - y \|,\ \ E\gamma^2(\omega) < \infty .$$

Якщо існує досить ефективний спосіб побудови узагальненого майже градієнта $\hat{g}(x,\omega)$ функції $f_0$ в довільній точці $(x,\omega)$, то в наведеному нижче алгоритмі за напрямок руху $\xi^k$ до наступного наближення $x^{k+1}$ вибирається вектор:

$$\xi^k = \hat{g}(\tilde{x}^k, \omega^k), \tag{4.8}$$

де $\tilde{x}^k$ – реалізація випадкового вектора рівномірно розподіленого в $n$-вимірному кубі з центром в точці $x^k$ і стороною $\alpha_k > 0$; $\omega^k$ – незалежні спостереження $\omega$. Якщо обчислювати $\hat{g}(x,\omega)$ важко, то за напрямок руху $\xi^k$ вибирається скінчено-різницевий аналог узагальненого майже градієнта $\hat{g}(\tilde{x}^k, \omega^k)$ у вигляді:

$$\begin{aligned}\xi^k = \frac{1}{2\alpha_k}\sum_{i=1}^{n}\Big( f_0(\tilde{x}_1^k, ..., x_i^k + \alpha_k, ..., \tilde{x}_n^k, \omega^k) - \\ - f_0(\tilde{x}_1^k, ..., x_i^k - \alpha_k, ..., \tilde{x}_n^k, \omega^k)\Big) e^i,\end{aligned} \tag{4.9}$$

де $e^i$, $i = \overline{1,n}$ – $i$-й орт.

***Алгоритм 2***

П о ч а т о к . І. Вибрати довільне початкове наближення $x^0 \in R^n$, значення крокового множника $\rho_0$ і зміщення $\alpha_0$, які задовольняють умовам теореми 2; покласти $k = 0$.

О с н о в н и й  ц и к л . ІІ. Обчислити реалізації $\tilde{x}_i^k$, $i = \overline{1,n}$, випадкових величин, рівномірно розподілених на відрізках:

$$\left[x_i^k - \alpha_k, x_i^k + \alpha_k\right], \quad i = \overline{1,n}.$$

ІІІ. Обчислити реалізацію $\omega^k$ випадкового параметра $\omega$.

IV. Обчислити вектор $\xi^k$ напрямку руху до наступного наближення $x^{k+1}$ за формулою (4.8) (або (4.9)).

V. Обчислити наступне наближення

$$x^{k+1} = x^k - \rho_k \xi^k.$$

VI. Обчислити значення крокового множника $\rho_{k+1}$ і зміщення $\alpha_{k+1}$, які задовольняють умовам теореми 2.

VII. Покласти $k = k+1$ і перейти на крок ІІ.

***Теорема 2.*** *Нехай виконані припущення 2 і умови:*

$$\rho_k > 0, \ \ \alpha_k > 0 \ \ \text{при} \ \ k = 0,1,...;$$

$$\sum_{k=0}^{\infty}\rho_k = \infty, \ \ \sum_{k=0}^{\infty}\rho_k^2 < \infty, \ \ \sum_{k=0}^{\infty}(\rho_k/\alpha_k)^2 < \infty;$$

$$\alpha_k \to 0 \ \ і \ \ |\alpha_k - \alpha_{k+1}|/\rho_k \to 0 \ \ \text{при} \ \ k \to \infty.$$

*Тоді існує підпослідовність* $\left\{x^{k_s}\right\}_{s=0}^{\infty}$ *послідовності* $\left\{x^k\right\}_{k=0}^{\infty}$, *яка породжена алгоритмом 2, яка збігається до точки* $x^* \in X^* \triangleq \left\{x^* | 0 \in G(x^*)\right\}$.

## 4.9. Метод усереднення напрямків спуску для мінімізації функцій, що задовольняють умову Ліпшиця

Задача 1. Знайти $\arg\min_{x\in R^n} f_0(x)$ для заданої функції $f_0 : R^n \to R^1$.

*Припущення 1.* Функція $f_0$ задовольняє локальну умову Ліпшиця.

***Алгоритм 1***

Початок. I. Вибрати будь-яке початкове наближення $x^0 \in R^n$.

II. Вибрати будь-яке натуральне число *l*.

III. Покласти $k = 0$.

Основний цикл. IV. Знайти кроковий множник $\rho_k$, параметри $\alpha_k$ та $\delta_k$, які задовольняють умови теореми 1.

V. Обчислити реалізацію $\tilde{x}^k$ випадкової точки, рівномірно розподіленої в $n$-вимірному кубі з центром в точці $x^k$ і стороною $\alpha_k$.

VI. Обчислити вектор $\theta^k$ за формулою:

$$\theta^k = \sum_{i=1}^{n} \frac{f_0(\tilde{x}^k + \delta_k e^i) - f_0(\tilde{x}^k)}{\delta_k} e^i,$$

де $e^i$, $i = \overline{1,n}$ – $i$-й орт.

VII. Якщо $k \geq l$, то обчислити наступне наближення

$$x^{k+1} = x^k - \rho_k \sum_{j=k-l}^{k} \theta^j$$

і перейти на крок VIII; інакше обчислити наступне наближення за формулою

$$x^{k+1} = x^k - \rho_k \sum_{j=0}^{k} \theta^j$$

і перейти на крок VIII.

VIII. Покласти $k = k+1$ і перейти на крок IV.

***Теорема 1.*** *Якщо функція $f_0$ задовольняє локальну умову Ліпшиця і виконуються умови:*

$$\rho_k > 0,\ \ \alpha_k > 0,\ \ k = 0,1,...;$$

$$\sum_{k=0}^{\infty} \rho_k = \infty,\ \sum_{k=0}^{\infty} \rho_k^2 < \infty,\ \lim_{k\to\infty} (\rho_k / a_k) = 0;$$

$$\lim_{k\to\infty} a_k = 0,\ \lim_{k\to\infty} (\delta_k / a_k) = 0,\ \overline{\lim_{k\to\infty}} (\rho_k / \rho_{k+1}) < \infty;$$

$$\lim_{k\to\infty} \left( |a_k - a_{k+1}| / \rho_k \right) = 0,\ \lim_{k\to\infty} (a_{k+1} / a_k) = 1,$$

*то усі граничні точки послідовності $\{x^k\}_{k=0}^{\infty}$, яка породжена алгоритмом 1, з ймовірністю 1 належать множині $X^* \triangleq \{x^* | 0 \in G(x^*)\}$, де $G(x)$ –*

*множина узагальнених майже градієнтів функції $f_0$ в точці $x$, і, крім того, послідовність $\left\{f_0(x^k)\right\}_{k=0}^{\infty}$ збігається майже напевно.*

## 4.10. Методи послідовних наближень для розв'язування дискретних мінімаксних задач

З а д а ч а  0. Знайти $\arg\min\limits_{x\in R^n}\max\limits_{i\in\Im}\varphi_i(x)$ для заданих функцій $\varphi_i : R^n \to R^1,\ i\in\Im$ та заданої множини $\Im$.

*Припущення 0.* $(\boldsymbol{i})-$ функції $\varphi_i,\ i\in\Im$ – неперервно диференційовані в $R^n$.

У даному підрозділі описуються методи, в яких на $k$-й ітерації за напрямок руху до наступного наближення $x^{k+1}$ вибирається вектор $h^k(\varepsilon)$, що задовольняє умову:

$$\max_{i\in\mathrm{K}_\varepsilon(x^k)}\left(\nabla\varphi_i(x^k), h^k(\varepsilon)\right)=\min_{\|h\|=1}\max_{i\in\mathrm{K}_\varepsilon(x^k)}\left(\nabla\varphi_i(x^k), h\right),$$

де

$$\mathrm{K}_\varepsilon(x^k)\triangleq\left\{i\,|\max_{i\in\Im}\varphi_i\left(x^k\right)-\varphi_i(x^k)\le\varepsilon,\quad i\in\Im\right\}.$$

Вектор $h^k(\varepsilon)$ при $\varepsilon\ge 0$ називають ***напрямком ε-найшвидшого спуску***.



#### 1. Перший метод послідовних наближень

### *Алгоритм 1*

П о ч а т о к . I. Вибрати довільне початкове наближення $x^0\in R^n$.

II. Вибрати константи $\varepsilon_0>0$ і $\alpha_0>0$.

III. Покласти $k=0$.

О с н о в н и й  ц и к л . IV. Покласти $\varepsilon=0$ і перейти на крок VII.

V. Покласти $j=0$.

VI. Покласти $\varepsilon=\varepsilon_j$.

VII. Знайти множину індексів

$$K_\varepsilon(x^k)=\left\{i\,|\max_{i\in\Im}\varphi_i(x^k)-\varphi_i(x^k)\le\varepsilon,\quad i\in\Im\right\}.\qquad(4.10)$$

VIII. Знайти багатогранник $L_\varepsilon(x^k)$, що є опуклою оболонкою, натягнутою на множину точок

$$\left\{\nabla\varphi_i(x^k),\ i\in\mathrm{K}_\varepsilon(x^k)\right\}.$$

IX. Якщо $\varepsilon=0$, то перейти на крок X; інакше перейти на крок XI.

X. Використовуючи алгоритм $1'$, визначити, чи належить початок координат багатограннику $L_0(x^k)$. Якщо початок координат належить $L_0(x^k)$, то покласти $x^* = x^k$ і припинити обчислення; інакше перейти на крок XI.

XI. Використовуючи алгоритм $1''$, знайти точку $z_\varepsilon$, що є найближчою до початку координат точкою багатогранника $L_\varepsilon(x^k)$.

XII. Обчислити $\psi_\varepsilon(x^k)$ за формулою $\psi_\varepsilon(x^k) = -\| z_\varepsilon \|$.

XIII. Якщо виконується нерівність $\psi_\varepsilon(x^k) \le -\varepsilon(\alpha_0 / \varepsilon_0)$, то перейти на крок XIV; інакше покласти $\varepsilon_{j+1} = \varepsilon_0 / 2^{j+1}$, $j = j+1$ і перейти на крок VI.

XIV. Обчислити вектор $h^k(\varepsilon)$ – напрямок $\varepsilon$-найшвидшого спуску функції $\max\limits_{i \in \Im} \varphi_i(x)$ у точці $x^k$ за формулою $h^k(\varepsilon) = -(1/\|z_\varepsilon\|) z_\varepsilon$.

XV. Обчислити кроковий множник $\rho_k$ з умови

$$\max_{i \in \Im} \varphi_i(x^k + \rho_k h^k(\varepsilon)) = \min_{\rho \ge 0} \max_{i \in \Im} \varphi_i(x^k + \rho h^k(\varepsilon)).$$

XVI. Обчислити наступне наближення

$$x^{k+1} = x^k + \rho_k h^k(\varepsilon).$$

XVII. Покласти $k$=$k$+1 і перейти на крок IV.

***Теорема 1.*** *Якщо виконане припущення 0 і початкове наближення* $x^0$ *в алгоритмі 1 таке, що множина*

$$X_0 \triangleq \left\{ x \mid \max_{i \in \Im} \varphi_i(x) \le \max_{i \in \Im} \varphi_i(x^0), \quad x \in R^n \right\}$$

*обмежена, то будь-яка гранична точка нескінченної послідовності* $\{x^k\}_{k=o}^{\infty}$, *яка породжена алгоритмом 1, є стаціонарною точкою функції* $\max\limits_{i \in \Im} \varphi_i(x)$.

(Точка $x^* \in R^n$, для якої виконується нерівність

$$\inf_{\|g\|=1} \max_{i \in \mathrm{K}_0(x^*)} \left( \frac{\partial \varphi_i(x^*)}{\partial x}, g \right) \ge 0,$$

називається ***стаціонарною точкою*** функції $\max\limits_{i \in \Im} \varphi_i(x)$ на $R^n$. Якщо функція $\max\limits_{i \in \Im} \varphi_i(x)$ опукла донизу, то стаціонарна точка $x^*$ є точкою мінімуму).

*Зауваження 1.* Щоб на $k$-й ітерації зменшити кількість порівнянь $\varphi_\varepsilon(x^k)$ із $-\ \varepsilon(\alpha_0 / \varepsilon_0)$, рекомендується починати порівняння не з $\varepsilon = \varepsilon_0$, а зі значення $\varepsilon$, отриманого на попередній $(k-1)$-й ітерації, тобто при такому $\varepsilon$, яке отримується при виході з циклу, утвореного кроками VI-XIII на $(k-1)$-й ітерації.

*Алгоритм 1′* **(алгоритм визначення належності початку координат багатограннику $L_\varepsilon(x^k)$, який є опуклою оболонкою множини векторів $\left\{\nabla\varphi_i(x^k),\ i\in \mathrm{K}_\varepsilon(x^k)\right\}$).**

I. Обчислити параметр $l$ – кількість елементів множини $\mathrm{K}_\varepsilon(x^k)$.

II. Обчислити $i_0$ – найменший елемент множини $\mathrm{K}_\varepsilon(x^k)$.

III. Покласти $j=1$.

IV. Покласти $m=0$.

V. Обчислити індекс $i=i_0+m$.

VI. Якщо індекс $i$ належить множині $\mathrm{K}_\varepsilon(x^k)$, то покласти $g^i=\nabla\varphi_i(x^k)$ і перейти на крок VII; інакше покласти $m=m+1$ і перейти на крок V.

VII. Якщо $j<l$, то покласти $j=j+1$ і перейти на крок IV; інакше перейти на крок VIII.

VIII. Розв'язати задачу лінійного програмування в $(l+1)$- вимірному просторі векторів ($\gamma_1,\gamma_2,...,\gamma_l,\alpha$):

знайти $\arg\min\alpha$ при умовах:

$$\sum_{j=1}^{l}\gamma_j g_s^j-\alpha\le 0,\quad s=\overline{1,n};\qquad -\sum_{j=1}^{l}\gamma_j g_s^j-\alpha\le 0,\quad s=\overline{1,n};$$

$$\sum_{j=1}^{l}\gamma_j=1;\qquad \gamma_j\ge 0,\quad j=\overline{1,l};\quad \alpha\ge 0.$$

Позначити через $(\gamma_1^0,\gamma_2^0,...,\gamma_l^0,\alpha^0)$ розв'язок цієї задачі.

IX. Якщо $\alpha^0=0$, то початок координат належить багатограннику $L_\varepsilon(x^k)$; якщо $\alpha^0>0$, то початок координат не належить багатограннику $L_\varepsilon(x^k)$.

*Алгоритм 1″* **(алгоритм обчислення найближчої до початку координат точки багатогранника $L_\varepsilon(x^k)$, що є опуклою оболонкою множини векторів $Q=\left\{\nabla\varphi_i(x^k),\ i\in \mathrm{K}_\varepsilon(x^k)\right\}$).**

I. Як початкове наближення $y^0$ вибрати довільну точку багатогранника $L_\varepsilon(x^k)$ (зокрема у якості $y^0$ можна брати вектор $\nabla\varphi_{i_0}(x^k)$, для якого

$$(\nabla\varphi_{i_0}(x^k),\nabla\varphi_{i_0}(x^k))=\min_{i\in \mathrm{K}_\varepsilon(x^k)}(\nabla\varphi_i(x^k),\nabla\varphi_i(x^k)),$$

чи вектор, рівний $\sum_{j=1}^{l}\gamma_j^0 g^j$, де $(\gamma_1^0,\gamma_2^0,...,\gamma_l^0)$ – вектор, отриманий при розв'язуванні задачі лінійного програмування на кроці VIII алгоритму 1′, а вектори $g^j$, $j=\overline{1,l}$, визначаються кроками I-VII алгоритму 1′).

II. Покласти $m=0$.

III. Знайти точку $\overline{y}^m$ множини $Q$, для якої

$$(\overline{y}^m, y^m)=\min_{i\in K_\varepsilon(x^k)}(\nabla\varphi_i(x^k), y^m).$$

IV. Якщо $(y^m-\overline{y}^m, y^m)=0$, то покласти $z_\varepsilon = y^m$ і припинити обчислення; інакше перейти на крок V.

V. Знайти параметр $\tau_m\in[0;1]$, що задовольняє умову

$$\left\|y^m+\tau_m(\overline{y}^m-y^m)\right\|=\min_{t\in[0;1]}\left\|y^m+t(\overline{y}^m-y^m)\right\|.$$

VI. Обчислити наступне наближення

$$y^{m+1}=y^m+\tau_m(\overline{y}^m-y^m).$$

VII. Покласти $m=m+1$ і перейти на крок III.

***Теорема 1′.*** *Якщо виконані припущення теореми 1, то нескінченна послідовність* $\{y^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 1″, збігається до найближчої до початку координат точки* $z_\varepsilon$ *багатогранника* $L_\varepsilon(x^k)$.

**2. Модифікація першого методу послідовних наближень**

***Алгоритм 2***

Усі кроки алгоритму 1, за винятком кроку XV, залишаються без змін. Крок XV записати у вигляді:

XV. Використовуючи алгоритм 2′, обчислити кроковий множник $\rho_k$, що задовольняє умову:

$$\left(\max_{i\in\Im}\varphi_i(x^k)-\max_{i\in\Im}\varphi_i(x^k+\rho_k h^k(\varepsilon))\right)\Big/\Big(\max_{i\in\Im}\varphi_i(x^k)- \\ -\min_{\rho\ge 0}\max_{i\in\Im}\varphi_i(x^k+\rho h^k(\varepsilon))\Big)\ge\beta, \tag{4.11}$$

де $\beta\in(0;1)-$ довільний параметр, фіксований для усіх значень $k$.

***Алгоритм 2′*** **(алгоритм обчислення за скінчене число ітерацій крокового множника $\rho_k$, який задовольняє умові (4.11))**

I. Вибрати довільний параметр $\beta\in(0;1)$ і константу $\delta_0>0$.

II. Обчислити $\mu=(\sqrt{5}-1)/2$.

III. Визначити функцію $f_k : R^1\to R^1$ за правилом

$$f_k(t) = \max_{i \in \Im} \varphi_i(x^k + th^k(\varepsilon)). \tag{4.12}$$

IV. Обчислити значення $f_k(\delta_0)$ та $f_k(0)$. Якщо $f_k(\delta_0) < f_k(0)$, то перейти на крок V; інакше перейти на крок X.

V. Покласти $s = 1$.

VI. Обчислити $\delta_s = \delta_{s-1}\mu$.

VII. Якщо $f_k(\delta_s) \geq f_k(\delta_{s-1})$, то перейти на крок IX; інакше перейти на крок VIII.

VIII. Покласти $s = s+1$ і перейти на крок VI.

IX. Покласти $\delta_0^{(1)} = 0$, $\delta_0^{(2)} = \delta_{s-1}$, $\delta_0^{(3)} = \delta_s$ і перейти на крок XV.

X. Покласти $s = 1$.

XI. Обчислити $\delta_s = \delta_{s-1} / \mu$.

XII. Якщо $f_k(\delta_s) \leq f_k(0)$, то перейти на крок XIV; інакше перейти на крок XIII.

XIII. Покласти $s = s+1$ і перейти на крок XI.

XIV. Покласти $\delta_0^{(1)} = 0$, $\delta_0^{(2)} = \delta_s$, $\delta_0^{(3)} = \delta_{s-1}$.

XV. Покласти $j = 0$.

XVI. Знайти множину індексів $K \subset \Im$

$$K \triangleq \left\{ i \mid \max_{i \in \Im} \varphi_i(x^k + \delta_j^{(2)} h^k(\varepsilon)) - \varphi_i(x^k + \delta_j^{(2)} h^k(\varepsilon)) \leq 0 \right\}.$$

Обчислити похідну справа $f_k'(\delta_j^{(2)} + 0)$ і похідну зліва $f_k'(\delta_j^{(2)} - 0)$ функції $f_k$ у точці $\delta_j^{(2)}$:

$$f_k'(\delta_j^{(2)} + 0) = \max_{i \in \mathrm{K}} (\nabla \varphi_i(x^k + \delta_j^{(2)} h^k(\varepsilon)), h^k(\varepsilon));$$

$$f_k'(\delta_j^{(2)} - 0) = \min_{i \in \mathrm{K}} (\nabla \varphi_i(x^k + \delta_j^{(2)} h^k(\varepsilon)), h^k(\varepsilon)).$$

XVII. Обчислити

$$\Delta_j = \max\left\{ f_k'(\delta_j^{(2)} + 0)(\delta_j^{(2)} - \delta_j^{(3)}), f_k'(\delta_j^{(2)} - 0)(\delta_j^{(2)} - \delta_j^{(1)}), 0 \right\}.$$

XVIII. Якщо виконується нерівність

$$(f_k(0) - f_k(\delta_j^{(2)})) \cdot (1 - \beta) \geq \beta \Delta_j,$$

то покласти $\rho_k = \delta_j^{(2)}$ і припинити обчислення; інакше перейти на крок XIX.

XIX. Обчислити $\gamma_j$ – середину відрізка $[\delta_j^{(1)}; \delta_j^{(3)}]$,

$$\gamma_j = (\delta_j^{(1)} + \delta_j^{(3)}) / 2.$$

XX. Обчислити точку $\delta_j^{(4)}$, симетричну $\delta_j^{(2)}$ відносно точки $\gamma_j$.

XXI. Якщо $f_k(\delta_j^{(4)}) < f_k(\delta_j^{(2)})$, то покласти:

$$\delta_{j+1}^{(1)} = \delta_j^{(2)}; \quad \delta_{j+1}^{(2)} = \delta_j^{(4)}; \quad \delta_{j+1}^{(3)} = \delta_j^{(3)}$$

і перейти на крок XXII; інакше покласти:

$$\delta_{j+1}^{(1)} = \delta_j^{(1)}; \quad \delta_{j+1}^{(2)} = \delta_j^{(2)}; \quad \delta_{j+1}^{(3)} = \delta_j^{(4)}$$

і перейти на крок XXII.

XXII. Покласти $j = j+1$ та перейти на крок XVI.

***Теорема 2.*** *Якщо виконані умови теореми 1 і функції* $f_k(t)$, $k = 0,1, \ldots,$ *опуклі на множині* $[0;\infty)$, *то будь-яка гранична точка нескінченної послідовності* $\{x^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 2, є стаціонарною точкою функції* $\max\limits_{i\in\Im} \varphi_i(x)$.

#### 3. Другий метод послідовних наближень

Алгоритм 3 застосовується для знаходження $\varepsilon$- стаціонарних точок функції $\max\limits_{i\in\Im} \varphi_i(x)$. Якщо $x^* - \varepsilon$ - стаціонарна точка та $\varphi_i(x)$, $i \in \mathrm{K}_\varepsilon(x^*)$ – опуклі донизу функції, то $\max\limits_{i\in\Im} \varphi_i(x^*)$ є наближене значення для мінімуму функції $\max\limits_{i\in\Im} \varphi_i(x)$ з абсолютною похибкою, що не перевищує $\varepsilon$.

Алгоритм 3' оснований на використанні алгоритму 3 і дає можливість обчислювати стаціонарні точки функції $\max\limits_{i\in\Im} \varphi_i(x)$.

Точка $x^* \in R^n$ називається ***ε-стаціонарною точкою*** функції $\max\limits_{i\in\Im} \varphi_i(x)$, якщо

$$\min_{\|g\|=1} \max_{i\in \mathrm{K}_\varepsilon(x^*)} \left( \frac{\partial \varphi_i(x^*)}{\partial x}, g \right) \geq 0$$

(тут $g \in R^n$, $\mathrm{K}_\varepsilon(x)$ визначається згідно (4.10)).

***Алгоритм 3***

Початок. I. Вибрати довільне початкове наближення $x^0 \in R^n$, константу $\varepsilon > 0$.

II. Покласти $k = 0$.

Основний цикл. III. Знайти множину індексів

$$\mathrm{K}_\varepsilon(x^k) \triangleq \left\{ i \mid \max_{i\in\Im} \varphi_i(x^k) - \varphi_i(x^k) \leq \varepsilon, \;\; i \in \Im \right\}.$$

IV. Знайти багатогранник $L_\varepsilon(x^k)$, що є опуклою оболонкою, натягнутою на множину точок

$$\left\{ \nabla \varphi_i(x^k), \; i \in \mathrm{K}_\varepsilon(x^k) \right\}.$$

V. Використовуючи алгоритм $1'$, визначити, чи належить початок координат багатограннику $L_\varepsilon(x^k)$. Якщо початок координат належить

$L_\varepsilon(x^k)$, то покласти $x^* = x^k$ і припинити обчислення; інакше перейти на крок VI.

VI. Використовуючи алгоритм 1″, знайти точку $z_\varepsilon$, яка є найближчою до початку координат точкою багатогранника $L_\varepsilon(x^k)$.

VII. Обчислити вектор $h^k(\varepsilon)$ – напрямок $\varepsilon-$ найшвидшого спуску функції $\max\limits_{i\in\Im}\varphi_i(x)$ у точці $x^k$

$$h^k(\varepsilon) = -\left(1/\|z_\varepsilon\|\right) z_\varepsilon.$$

VIII. Обчислити кроковий множник $\rho_k$ з умови

$$\max_{i\in\Im}\varphi_i(x^k + \rho_k h^k(\varepsilon)) = \min_{\rho\ge 0}\max_{i\in\Im}\varphi_i(x^k + \rho h^k(\varepsilon)).$$

IX. Обчислити наступне наближення

$$x^{k+1} = x^k + \rho_k h^k(\varepsilon).$$

X. Покласти $k = k+1$ і перейти на крок III.

***Теорема 3.*** *Якщо виконані умови теореми 1, то будь-яка гранична точка нескінченної послідовності $\left\{x^k\right\}_{k=0}^{\infty}$, яка породжена алгоритмом 3, є $\varepsilon$-стаціонарною точкою функції $\max\limits_{i\in\Im}\varphi_i(x)$.*

*Зауваження 3.* Обчислення крокового множника $\rho_k$ на кроці VIII алгоритму 3 з умови мінімуму практично нездійснене. Рекомендується використовувати алгоритм 2′ для обчислення за скінченне число ітерацій крокового множника $\rho_k$, що задовольняє умові (4.11). У випадку, коли $f_k(t),\ k = 0,1,...,$ − опуклі на $[0;\infty)$ функції, теорема 3 залишається в силі.

Алгоритм 3 може бути також використаний для знаходження стаціонарних точок функції $\max\limits_{i\in\Im}\varphi_i(x)$, тільки на кроці VI алгоритму 3 необхідно додатково обчислювати значення

$$\varphi_\varepsilon(x^k) = -\|z_\varepsilon\|.$$

***Алгоритм 3′***

Початок. I. Вибрати довільне початкове наближення $x^{0,0} \in R^n$, константи $\varepsilon_0 > 0,\ \rho_0 > 0$.

II. Покласти $l = 0,\ \ k_0 = 0$.

Основний цикл. III. Покласти $x^0 = x^{l,k_l}$, $\varepsilon = \varepsilon_l$, $\rho = \rho_l$.

IV. Використовуючи алгоритм 3, обчислити таку точку $x^{k_l}$, що

$$\varphi_\varepsilon(x^{k_l}) \ge -\rho$$

(така точка отримується за скінченне число ітерацій алгоритму 3).

V. Покласти $x^{l+1,k_{l+1}} = x^{k_l}$, $\varepsilon_{l+1} = \varepsilon_l/2,\ \ \rho_{l+1} = \rho_l/2$ і перейти на крок III.

***Теорема 3'.*** *Якщо виконане припущення 0 і* $x^{0,0}$ *таке, що множина*

$$\left\{x \mid \max_{i\in\Im}\varphi_i(x) \le \max_{i\in\Im}\varphi_i(x^{0,0}), \quad x\in R^n\right\}$$

*обмежена, то будь-яка гранична точка нескінченної послідовності* $\left\{x^{l,k_l}\right\}_{l=0}^{\infty}$, *яка породжена алгоритмом 3', є стаціонарною точкою функції* $\max_{i\in\Im}\varphi_i(x)$.

#### 4. Модифікація другого методу послідовних наближень

*Припущення 4*. Функції $\varphi_i$, $i\in\Im$ – двічі неперервно диференційовані в просторі $R^n$.

***Алгоритм 4***

Початок. I. Вибрати довільне початкове наближення $x^0\in R^n$, константи $\alpha>0$ і $\varepsilon>0$.

II. Покласти $k=0$.

III. Знайти множину

$$X_0 \triangleq \left\{x \mid \max_{i\in\Im}\varphi_i(x) \le \max_{i\in\Im}\varphi_i(x^0), \quad x\in R^n\right\}.$$

IV. Обчислити константу

$$\beta_1 = \max_{x\in X_0}\max_{i\in\Im}\left\|\nabla\varphi_i(x)\right\|.$$

V. Знайти множину

$$Y_0 = \left\{y \mid \|y\| \le \alpha\beta_1, \quad y\in R^n\right\}.$$

VI. Обчислити константу:

$$\beta_2 = \max_{x\in X_0,\, y\in Y_0}\max_{i\in\Im}\left\|\frac{\partial^2\varphi_i(x+y)}{\partial x^2}\right\|,$$

де

$$\left\|\frac{\partial^2\varphi_i(x+y)}{\partial x^2}\right\| = \sqrt{\sum_{j_1,j_2=1}^{n}\left(\frac{\partial^2\varphi_i(x+y)}{\partial x_{j_1}\partial x_{j_2}}\right)^2}.$$

VII. Якщо $\beta_2=0$, то обчислити константу $\bar{\alpha}=\min\left\{\alpha,\ \varepsilon/(2\beta_1^2)\right\}$ і перейти на крок IX; інакше перейти на крок VIII.

VIII. Обчислити константу

$$\bar{\alpha} = \min\left\{\alpha,\ 2/\beta_2,\ \frac{-\beta_1+\sqrt{\beta_1^2+\beta_2\varepsilon}}{\beta_1\beta_2}\right\}.$$

IX. Вибрати довільне $\alpha_0\in(0,\bar{\alpha})$.

Основний цикл. X. Знайти множину індексів

$$\mathrm{K}_{\varepsilon}(x^k) \underline{\underline{\Delta}} \left\{ i \mid \max_{i \in \Im} \varphi_i(x^k) - \varphi_i(x^k) \le \varepsilon,\ i \in \Im \right\}.$$

XI. Використовуючи алгоритм 1', визначити, чи належить початок координат багатограннику $L_{\varepsilon}(x^k)$, який є опуклою оболонкою, натягнутою на множину точок $\left\{ \nabla\varphi_i(x^k),\ \ i \in K_{\varepsilon}(x^k) \right\}$. Якщо початок координат належить $L_{\varepsilon}(x^k)$, то покласти $x^* = x^k$ і припинити обчислення; інакше перейти на крок XII.

XII. Використовуючи алгоритм 1'', знайти точку $z_{\varepsilon}$, що є найближчою до початку координат точкою багатогранника $L_{\varepsilon}(x^k)$.

XIII. Обчислити $\rho_k = \|z_{\varepsilon}\|$.

XIV. Обчислити вектор

$$h^k(\varepsilon) = -(1/\|z_{\varepsilon}\|) z_{\varepsilon}.$$

XV. Обчислити наступне наближення

$$x^{k+1} = x^k + \alpha_0 \rho_k h^k(\varepsilon).$$

XVI. Покласти $k = k+1$ і перейти на крок X.



***Теорема 4.*** *Якщо виконане припущення 4 і початкове наближення* $x^0$ *таке, що множина* $X_0$ *обмежена, то будь-яка гранична точка нескінченної послідовності* $\{x^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 4, є* $\varepsilon$*-стаціонарною точкою функції* $\max_{i \in \Im} \varphi_i(x)$.

## 4.11. Методи Ерроу–Гурвиця розв'язування неперервних мінімаксних задач

### 1. Детермінований метод Ерроу–Гурвиця

З а д а ч а  1. Знайти $\arg\min_{x \in R^n} \max_{y \in R^m} \varphi(x, y)$ для заданої функції $\varphi : R^n \times R^m \to R^1$.

*Припущення 1.* $(\boldsymbol{i})$ – функція $\varphi(x, y)$ неперервно диференційована по $x$ та $y$; $(\boldsymbol{ii})$ – функція $\varphi(x, y)$ опукла донизу по $x$ при будь-якому $y$ і опукла доверху по $y$ при будь-якому $x$.

В методі Ерроу–Гурвиця на $k$-й ітерації обчислюють $(k+1)$-е наближення $(x^{k+1}, y^{k+1})$ в напрямку антиградієнта по $x$ і градієнта по $y$ функції $\varphi(x, y)$, що обчислюються в точці $(x^k, y^k)$.

Крокові множники по змінним $x$ та $y$ задовольняють класичним умовам і можуть бути різними.

***Алгоритм 1***

Початок. I. Вибрати довільне початкове наближення $x^0 \in R^n$, $y^0 \in R^m$.

II. Покласти $k = 0$.

Основний цикл. III. Обчислити вектори $\nabla_x \varphi(x^k, y^k)$ і $\nabla_y \varphi(x^k, y^k)$ – градієнти функції $\varphi(x, y)$ відповідно по змінним $x$ та $y$, обчислені в точці $(x^k, y^k)$.

IV. Обчислити крокові множники $\rho_k$ і $\rho'_k$, що задовольняють умовам теореми 1.

V. Обчислити наступні наближення:

$$x^{k+1} = x^k - \rho_k \nabla_x \varphi(x^k, y^k),$$
$$y^{k+1} = y^k + \rho'_k \nabla_y \varphi(x^k, y^k).$$

VI. Покласти $k = k + 1$ і перейти на крок III.

***Теорема 1.*** *Нехай виконуються припущення 1 і нехай (**i**) – у функції $\varphi(x, y)$ існує принаймні одна сідлова точка; (**ii**) – крокові множники $\rho_k$ і $\rho'_k$ алгоритму 1 задовольняють умовам:*

$$\rho_k > 0 \ \ i \ \ \rho'_k > 0 \ \ при \ \ k = 0, 1, \ldots, \ \sum_{k=0}^{\infty} \rho_k = \infty;$$
$$\rho'_k \to 0, \ \ \rho_k / \rho'_k \to 1 \ \ i \ \ \rho_{k+1} / \rho_k \to 1 \ \ при \ \ k \to \infty.$$

*Тоді, якщо послідовності $\{x^k\}_{k=0}^{\infty}$ і $\{y^k\}_{k=0}^{\infty}$, які породжені алгоритмом 1, обмежені, то будь-яка гранична точка $\bar{x}$ послідовності $\{x^k\}_{k=0}^{\infty}$ є першою компонентою $\bar{x}$ деякої сідлової точки $(\bar{x}, \bar{y})$ функції $\varphi(x, y)$.*

*Зауваження 1.* При зроблених в теоремі припущенях щодо послідовності $\{y^k\}_{k=0}^{\infty}$ можна стверджувати лише її збіжність до множини:

$$\tilde{Y} = \{\tilde{y} \mid \varphi(x^*, \tilde{y}) = \max_{y \in R^m} \varphi(x^*, y), \ \ \tilde{y} \in R^m\},$$

де $x^*$ – вектор, що належить множині перших компонент сідлових точок функції $\varphi(x, y)$.

В деяких випадках $\tilde{Y}$ співпадає з множиною других компонент сідлових точок функції $\varphi(x, y)$, але в загальному випадку ці множини різні (прикладом цього є функція $\varphi(x, y) = xy$).

**Приклад 1.** Розв'язати неперервну мінімаксну задачу:

$$\text{знайти } \arg \min_{x \in R} \max_{y \in R} \left( 2x^2 - 4x - 4y^2 + y - 3xy + 1 \right),$$

використовуючи чотири ітерації детермінованого методу Ерроу–Гурвиця.

*Розв'язування.* Обчислимо градієнти функції $\varphi(x, y)$ по змінних $x$ і $y$:

$$\nabla_x \varphi(x, y) = 4x - 4 - 3y\,;\ \nabla_y \varphi(x, y) = -8y + 1 - 3x\,.$$

**Алгоритм 1**

I. Вибираємо початкове наближення $x^0 = 2$; $y^0 = 0,8$.

II. Покладемо $k = 0$.

*1-а ітерація:*

III. Обчислюємо градієнти:

$$\nabla_x \varphi(x^0, y^0) = 4 \cdot 2 - 4 - 3 \cdot 0,8 = 1,6\,;$$

$$\nabla_y \varphi(x^0, y^0) = -8 \cdot 0,8 + 1 - 3 \cdot 2 = -11,4\,.$$

IV. Знаходимо крокові множники $\rho_0 = 0,25$, $\rho_0' = 0,4$.

V. Обчислюємо наступні наближення:

$$x^1 = x^0 - \rho_0 \cdot \nabla_x \varphi(x^0, y^0) = 2 - 0,25 \cdot 1,6 = 1,6\,;$$

$$y^1 = y^0 + \rho_0' \cdot \nabla_y \varphi(x^0, y^0) = 0,8 + 0,4 \cdot (-11,4) = -3,76\,.$$

VI. Покладемо $k = 0 + 1 = 1$ і переходимо на крок III.

*2-а ітерація:*

III. Обчислюємо градієнти:

$$\nabla_x \varphi(x^1, y^1) = 4 \cdot 1,6 - 4 - 3 \cdot (-3,76) = 13,68\,;$$

$$\nabla_y \varphi(x^1, y^1) = -8 \cdot (-3,76) + 1 - 3 \cdot 1,6 = 26,28\,.$$

IV. Знаходимо крокові множники $\rho_1 = 0,055$, $\rho_1' = 0,14$.

V. Обчислюємо наступні наближення:

$$x^2 = x^1 - \rho_1 \cdot \nabla_x \varphi(x^1, y^1) = 1,6 - 0,055 \cdot 13,68 = 0,8476\,;$$

$$y^2 = y^1 + \rho_1' \cdot \nabla_y \varphi(x^1, y^1) = -3,76 + 0,14 \cdot 26,28 = -0,0808\,.$$

VI. Покладемо $k = 1 + 1 = 2$ і переходимо на крок III.

*3-я ітерація:*

III. Обчислюємо градієнти:

$$\nabla_x \varphi(x^2, y^2) = 4 \cdot 0,8476 - 4 - 3 \cdot (-0,0808) = -0,3672\,;$$

$$\nabla_y \varphi(x^2, y^2) = -8 \cdot (-0,0808) + 1 - 3 \cdot 0,8476 = -0,8964\,.$$

IV. Знаходимо крокові множники $\rho_2 = 0,015$, $\rho_2' = 0,12$.

V. Обчислюємо наступні наближення:

$$x^3 = x^2 - \rho_2 \cdot \nabla_x \varphi(x^2, y^2) = 0,8476 + 0,015 \cdot 0,3672 = 0,8531\,;$$

$$y^3 = y^2 + \rho_2' \cdot \nabla_y \varphi(x^2, y^2) = -0,0808 - 0,12 \cdot 0,8964 = -0,1884\,.$$

VI. Покладемо $k=2+1=3$ і переходимо на крок III.

*4-а ітерація:*

III. Обчислюємо градієнти:

$$\nabla_x\varphi\left(x^3,y^3\right)=4\cdot0,8531-4-3\cdot\left(-0,1884\right)=-0,0224;$$

$$\nabla_y\varphi\left(x^3,y^3\right)=-8\cdot\left(-0,1884\right)+1-3\cdot0,8531=-0,0521.$$

IV. Знаходимо крокові множники $\rho_3=0,013$, $\rho_3'=0,11$.

V. Обчислюємо наступні наближення:

$$x^4=x^3-\rho_3\cdot\nabla_x\varphi\left(x^3,y^3\right)=0,8531+0,013\cdot0,0224=0,8534;$$

$$y^4=y^3+\rho_3'\cdot\nabla_y\varphi\left(x^3,y^3\right)=-0,1884-0,11\cdot0,0521=-0,1941.$$

VI. Покладемо $k=3+1=4$ і зупиняємо обчислення.

Таким чином, за чотири ітерації алгоритму 1 обчислили точку $\left(x^4;y^4\right)=\left(0,8534;-0,1941\right)^T$, в якій функція $\varphi$ приймає значення

$$\varphi\left(x^4,y^4\right)=2\cdot0,8534^2-4\cdot0,8534-4\cdot0,1941^2-0,1941-$$

$$-3\cdot0,8534\cdot\left(-0,1941\right)+1=-0,804881,$$

а градієнти такі:

$$\nabla_x\varphi\left(x^4,y^4\right)=4\cdot0,8534-4-3\cdot\left(-0,1941\right)=-0,0041;$$

$$\nabla_y\varphi\left(x^4,y^4\right)=-8\cdot\left(-0,1941\right)+1-3\cdot0,8534=-0,0074.$$



**2. Стохастичний метод Ерроу–Гурвиця**

З а д а ч а 2. Знайти $\arg\min\limits_{x\in R^n}\max\limits_{y\in R^m}E_\omega\varphi(x,y,\omega)$ для заданої функції $\varphi:R^n\times R^m\times\Omega\to R^1$.

*Припущення 2.* (***i***) – функція $f_0(x,y)\triangleq E_\omega\varphi(x,y,\omega)$ неперервно диференційована по *x* та *y*; (***ii***) – функція $f_0(x,y)$ опукла донизу по *x* при будь-якому *y* і опукла доверху по *y* при будь-якому *x*.

В стохастичному методі Ерроу–Гурвиця на *k*-й ітерації обчислюють (*k*+1)-е наближення в напрямку статистичних оцінок антиградієнта по *x* і градієнта по *y* функції $f_0(x,y)$ в точці $(x^k, y^k)$.

Крокові множники по змінних *x* та *y* задовольняють класичним умовам і можть бути різними.

***Алгоритм 2***

П о ч а т о к. I. Вибрати довільне початкове наближення $x^0\in R^n$, $y^0\in R^m$.

II. Покласти $k=0$.

О с н о в н и й   ц и к л. III. Обчислити крокові множники $\rho_k$ і $\rho'_k$, що задовольняють умовам теореми 2.

IV. Обчислити реалізації $\xi^k$, $\zeta^k$ випадкових векторів $\tilde{\xi}^k$, $\tilde{\zeta}^k$, що задовольняють умовам:

$$E(\tilde{\xi}^k / \Phi_k) = \nabla_x f_0(x^k, y^k), \quad E(\|\tilde{\xi}^k\|^2 / \Phi_k) \le \alpha,$$

$$E(\tilde{\zeta}^k / \Phi_k) = \nabla_y f_0(x^k, y^k), \quad E(\|\tilde{\zeta}^k\|^2 / \Phi_k) \le \alpha,$$

де $\Phi_k$ – $\sigma$-алгебра, що породжується випадковими величинами $x^0, y^0, x^1, y^1, \ldots, x^k, y^k$; $\alpha < \infty$.

V. Обчислити наступні наближення:

$$x^{k+1} = x^k - \rho_k \xi^k, \quad y^{k+1} = y^k + \rho'_k \zeta^k.$$

VI. Покласти $k = k+1$ і перейти на крок III.

***Теорема 2.*** *Нехай виконуються припущення 2 і крокові множники* $\rho_k$, $\rho'_k$ *задовольняють умовам:*

$$\rho_k > 0,\ \rho'_k > 0 \ \textit{при}\ k = 0, 1, \ldots;$$

$$\sum_{k=0}^{\infty} (\rho'_k)^2 < \infty,\ \rho_k / \rho'_k \to 0 \ \textit{при}\ k \to \infty,\ \sum_{k=0}^{\infty} \rho_k = \infty.$$

*Тоді, якщо послідовності* $\{x^k\}_{k=0}^{\infty}$ *і* $\{y^k\}_{k=0}^{\infty}$, *які породжені алгоритмом 2, обмежені майже напевно, то будь-яка гранична точка послідовності* $\{x^k\}_{k=0}^{\infty}$ *мінімізує функцію* $\max_{y \in R^m} f_0(x, y)$, *тобто є розв'язком задачі 2.*

## 4.12. Квазіградієнтні методи розв'язування дискретних мінімаксних задач стохастичного програмування

### 1. Мінімізація функції $E_\omega \max_{i \in \Im} \varphi_i(\boldsymbol{x}, \omega)$

З а д а ч а   1. Знайти $\arg\min_{x \in R^n} E_\omega \max_{i \in \Im} \varphi_i(x, \omega)$ для заданих функцій $\varphi_i : R^n \times \Omega \to R^1$, $i \in \Im$ і заданої множини індексів $\Im$.

*Припущення 1.* $(\boldsymbol{i})$ – функції $\varphi_i$, $i \in \Im$ неперервно диференційовані по *x*, причому градієнти по $x$ функції $\varphi_i$, $i \in \Im$ задовольняють при кожному $\omega$ локальну умову Ліпшиця

$$\|\nabla_x \varphi_i(x, \omega) - \nabla_x \varphi_i(y, \omega)\| \le \gamma_Z(\omega) \|x - y\|, \quad i \in \Im,$$

із інтегрованою константою Ліпшиця $\gamma_Z(\omega)$, яка задовольняє умові $E\gamma_Z(\omega) < \infty$ для $x, y$, які належать замкненій обмеженій множині $Z$; $(\boldsymbol{ii})$ –

для кожного $\omega$ є можливість обчислити значення функцій $\varphi_i(x,\omega)$, $i \in \Im$ і їх похідних по $x$: $\nabla_x \varphi_i(x,\omega)$, $i \in \Im$.

У алгоритмі 1 на *k*-й ітерації за вектор, який визначає напрямок руху до наступного наближення $x^{k+1}(\omega)$, вибирається $\nabla_x \varphi_{i_k}(x^k(\omega),\omega^k)$, де індекс $i_k \in \Im$ задовольняє умові

$$\varphi_{i_k}(x^k,\omega^k) = \max_{i\in\Im} \varphi_i(x^k,\omega^k) .$$

***Алгоритм 1***

Початок. I. Вибрати довільне початкове наближення $x^0 \in R^n$.

II. Вибрати таку константу $\delta > 0$, щоб виконувалась нерівність

$$\inf_{\|x\|\ge\delta} E \max_{i\in\Im} \varphi_i(x,\omega) > E \max_{i\in\Im} \varphi_i(x,\omega) \text{ при } \|x\| < \delta.$$

III. Покласти $k = 0$.

Основний цикл. IV. Обчислити $\omega^k$ – реалізацію випадкового елемента $\omega$.

V. Обчислити індекс $i_k$, який задовольняє умову

$$\varphi_{i_k}(x^k(\omega),\omega^k) = \max_{i\in\Im} \varphi_i(x^k(\omega),\omega^k) .$$

VI. Обчислити $\nabla_x \varphi_{i_k}(x^k(\omega),\omega^k)$.

VII. Знайти кроковий множник $\rho_k$, який задовольняє умовам теореми 1.

VIII. Якщо $\|x^k(\omega)\| > \delta$, то покласти $x^{k+1}(\omega) = x^k(\omega)$ і перейти на крок X; інакше перейти на крок IX.

IX. Обчислити

$$x^{k+1}(\omega) = x^k(\omega) - \rho_k \nabla_x \varphi_{i_k}(x^k(\omega),\omega^k).$$

X. Покласти $k = k+1$ і перейти на крок IV.

***Теорема 1.*** *Якщо виконані припущення 1 та* $(\boldsymbol{iii})-$ $E \max_{i\in\Im} \varphi_i(x,\omega) \to \infty$ *при* $\|x\| \to \infty$; $(\boldsymbol{iv})-$ *функція* $E \max_{i\in\Im} \varphi_i(x,\omega)$ *приймає на множині розв'язків X* задачі 1 не більше ніж зліченне число значень;* $(\boldsymbol{v})-$ *послідовність крокових множників* $\{\rho_k\}_{k=0}^{\infty}$ *задовольняє умовам:*

$$\rho_k > 0,\ k = 0, 1, \ldots,\ \sum_{k=0}^{\infty} \rho_k = \infty,\ \sum_{k=0}^{\infty} \rho_k^2 < \infty;$$

$$\rho_{k+1} / \rho_k \to 1 \text{ при } k \to \infty,$$

*то майже для всіх* $\omega$ *граничні точки послідовності* $\{x^k(\omega)\}_{k=0}^{\infty}$, *яка породжена алгоритмом 1, належать множині розв'язків X* задачі 1.*

**2. Мінімізація функції $\max_{i\in\Im} E_\omega \varphi_i(x,\omega)$**

З а д а ч а 2. Знайти $\arg\min_{x\in R^n} \max_{i\in\Im} E_\omega \varphi_i(x,\omega)$ для заданих функцій $\varphi_i : R^n \times \Omega \to R^1$, $i\in\Im$ і заданої множини індексів $\Im$.

*Припущення 2.* Функції $\varphi_i(x,\omega)$, $i\in\Im$ і розподіл $\omega$ такі, що виконуються наступні умови:

$(\boldsymbol{i})$ – існує константа $\delta<\infty$ така, що $(\nabla f_i(x),x)>0$ для $\|x\|>\delta$, де $f_i(x) \triangleq E_\omega \varphi_i(x,\omega)$;

$(\boldsymbol{ii})$ – градієнти функцій $f_i(x)$, $i\in\Im$, задовольняють локальну умову Ліпшиця

$$\|\nabla f_i(x)-\nabla f_i(y)\| \le \gamma_Z \|x-y\|,\ \gamma_Z<\infty$$

для всіх $x, y$, які належать довільній замкненій обмеженій множині Z;

$(\boldsymbol{iii})$ – при кожному $\omega$ є можливість обчислювати значення функцій $\varphi_i(x,\omega)$, $i\in\Im$ та їх похідних по $x$: $\nabla_x \varphi_i(x,\omega)$, $i\in\Im$.

В алгоритмі 2 на *k*-й ітерації за вектор, який визначає напрямок руху до наступного наближення $x^{k+1}(\omega)$, обирається $\nabla_x \varphi_{i_k}(x^k(\omega),\omega^k)$, де індекс $i_k\in\Im$ задовольняє умові $z_{i_k}^k = \max_{i\in\Im} z_i^k$ (тут при кожному $i\in\Im$ допоміжна послідовність чисел $\{z_i^k\}_{k=0}^\infty$ має властивість $z_i^k - f_i(x^k(\omega))\to 0$ при $k\to\infty$).

***Алгоритм 2***

П о ч а т о к. І. Вибрати константу $\delta>0$, яка задовольняє умові ($\boldsymbol{i}$) припущення 2.

ІІ. Вибрати довільне початкове наближення $x^0\in R^n$, яке задовольняє умові $\|x^0\|\le\delta$.

ІІІ. Вибрати довільні числа $z_i^0$, $i\in\Im$.

IV. Покласти $k=0$.

О с н о в н и й ц и к л. V. Обчислити індекс $i_k$, який задовольняє умову

$$z_{i_k}^k(\omega) = \max_{i\in\Im} z_i^k(\omega).$$

VI. Обчислити $\omega^k$ – реалізацію випадкового елемента $\omega$.

VII. Обчислити $\nabla_x \varphi_{i_k}(x^k(\omega),\omega^k)$.

VIII. Знайти кроковий множник $\rho_k$ і параметр $\sigma_k$, які задовольняють умовам теореми 2.

IX. Обчислити наступне наближення

$$x^{k+1}(\omega) = \begin{cases} x^o, & \text{якщо } \left\| x^k(\omega) \right\| > 2\delta; \\ x^k(\omega) - \rho_k \nabla_x \varphi_{i_k}(x^k(\omega), \omega^k), & \text{якщо } \left\| x^k(\omega) \right\| \le 2\delta. \end{cases}$$

X. Обчислити значення величин:

$$z_i^{k+1}(\omega) = z_i^k(\omega) + \sigma_k(\varphi_i(x^k(\omega), \omega^k) - z_i^k(\omega)), \quad i \in \Im.$$

XI. Покласти $k = k + 1$ і перейти на крок V.

***Теорема 2.*** *Якщо виконані припущення 2 і* $(\boldsymbol{iv})-$ $E_\omega | \varphi_i(x,\omega)|^2 < \infty$, $E_\omega \| \nabla_x \varphi_i(x,\omega) \|^2 < \infty$, $i \in \Im$; $(\boldsymbol{v})-$ *послідовність крокових множників* $\{\rho_k\}_{k=0}^{\infty}$ *задовольняє умовам:*

$$\rho_k > 0, \ \ k = 0, 1, ...; \ 0 \le \sigma_k \le 1, \ \ k = 0, 1, \ ...; \ \sum_{k=0}^{\infty} \rho_k = \infty, \ \sum_{k=0}^{\infty} \sigma_k^2 < \infty;$$

$$\rho_k / \sigma_k \to 0 \ \ при \ \ k \to \infty; \ \ \rho_{k+1} / \rho_k \to 0 \ \ при \ \ k \to \infty;$$

$(\boldsymbol{vi})-$ *функція* $\max_{i \in \Im} E_\omega \varphi_i(x,\omega)$ *приймає на множині розв'язків X* задачі 2 не більш ніж злічене число значень, то майже для всіх* $\omega$ *граничні точки послідовності* $\{x^k(\omega)\}_{k=0}^{\infty}$, *яка породжена алгоритмом 2, належать множині розв'язків X* задачі 2.*

## Запитання та вправи для самостійної роботи до розділу 4.

1. Сформулюйте означення узагальненого градієнта (субградієнта) функції $f_0(x)$ в точці $\bar{x} \in R^n$.
2. Сформулюйте означення опуклої донизу (опуклої) функції $f_0(x)$, $x \in R^n$.
3. Сформулюйте означення опуклої догори функції $f_0(x)$, $x \in R^n$.
4. Які функції називають строго опуклими донизу та строго опуклими доверху?
5. Перерахуйте основні властивості опуклих функцій.
6. Сформулюйте необхідні і достатні умови оптимальності опуклих функцій.
7. В якому напрямку здійснюється рух до наступного наближення $x^{k+1}$ в алгоритмі узагальненого градієнтного спуску з постійним кроковим множником?
8. Запишіть класичні умови, яким задовільняють крокові множники $\rho_k$ основного і модифікованого алгоритмів узагальненого градієнтного спуску.

9. Якими властивостями володіє послідовність $x^k$, $k = 0,1,2,...,$ що породжена основним і модифікованим алгоритмом узагальненого градієнтного спуску?
10. Яка швидкість збіжности послідовності $x^k$, $k = 0,1,2,...,$ що породжена другим алгоритмом зі спеціальним вибором кроку?
11. Запишіть формулу обчислення крокового множника $\rho_k$ в алгоритмі узагальненого градієнта, що використовує апріорне значення мінімуму функції.
12. В чому полягає сутність методів градієнтного типу з розтягненням простору?
13. Сформулюйте означення оператора розтягнення простору $R^n$ в напрямку вектора $\xi$ $(\|\xi\|=1)$ з коефіцієнтом розтягу $\alpha$ $(\alpha \geq 0)$.
14. Який оператор називають оператором «стиснення» простору $R^n$ ?
15. Яка особливість $r(\alpha)$-алгоритму в порівнянні з іншими методами узагальненого градієнтного спуску?
16. Для мінімізації яких функцій рекомендується використовувати методи локального випадкового пошуку?
17. Які із п’яти наведених алгоритмів локального випадкового пошуку на кожній ітерації забезпечують зменшення значення цільової функції $f_0$ ?
18. Сформулюйте означення слабоопуклої функції та її квазіградієнта.
19. Приведіть приклади слабоопуклих функцій і вкажіть як обчислюються їх квазіградієнти.
20. Сформулюйте умови, яким задовольняють крокові множники $\rho_k$ в квазіградієнтних методах мінімізації слабоопуклих донизу функцій.
21. Дайте означення $\varepsilon$-квазіградієнта слабоопуклої функції.
22. В якому напрямку здійснюється рух до наступного наближення $x^{k+1}$ в методах узагальнених майже градієнтів для мінімізації функцій, що задовольняють локальну умову Ліпшиця?
23. Сформулюйте детерміновану дискретну мінімаксну (максимінну) задачу.
24. Сформулюйте умови, з яких визначається кроковий множник $\rho_k$ в першому методі послідовних наближень і його модифікації.
25. Які точки знаходять другим методом послідовних наближень при розв’язуванні дискретних мінімаксних задач?

26. Сформулюйте детерміновану неперервну мінімаксну (максимінну) задачу.
27. В яких напрямках здійснюється рух по змінних $x$ та $y$ в детермінованому методі Ерроу–Гурвиця розв'язування неперервних мінімаксних задач?
28. Сформулюйте стохастичну неперервну мінімаксну (максимінну) задачу.
29. В яких напрямках здійснюється рух по змінних $x$ та $y$ в стохастичному методі Ерроу–Гурвиця розв'язування неперервних мінімаксних задач?
30. Дайте означення сідлової точки $(x^*; y^*)$.
31. Сформулюйте умови, яким задовольняють крокові множники в узагальненому градієнтному методі відшукання сідлових точок.
32. Сформулюйте дискретну мінімаксну задачу стохастичного програмування.
33. Використовуючи означення опуклої множини, доведіть, що прямокутник $X = \left\{ x \in R^2 : 0 \le x_1 \le a,\ 0 \le x_2 \le b \right\}$ є опуклою множиною.
34. Чи є опуклими множини:
    – $X = \left\{ (x_1, x_2) \in R^2 : x_1 - x_2{}^2 \le 0,\ x_2 - x_1{}^2 \le 0 \right\}$;
    – $X = \left\{ (x_1, x_2) \in R^2 : x_1{}^2 + x_2{}^2 \le 1,\ x_1 - x_2{}^2 \ge 0 \right\}$?
35. Знайдіть мінімальне значення $\beta$, при якому множина $X$ є опуклою:
$$X = \left\{ (x_1, x_2) \in R^2 : (x_1{}^2 + 1) x_2 \le 5,\ x_2 \ge \beta \right\}.$$
36. Дослідіть опуклість функцій:
    – $f(x) = \dfrac{3}{x},\ x > 0$;
    – $f(x) = \exp(3x_1 - x_2),\ (x_1, x_2) \in R^2$;
    – $f(x) = 3 - x_1{}^2 - x_2{}^2,\ (x_1, x_2) \in R^2$;
    – $f(x) = 3x_1 - 2x_1 x_2,\ (x_1, x_2) \in R^2$;
    – $f(x) = -\dfrac{2}{x_1} - \dfrac{2}{x_2},\ x_1 > 0,\ x_2 > 0$.
37. Знайдіть субградієнти функцій:
    – $f(x) = |2x + 2|,\ x \in R^1$;
    – $f(x) = |x + 3| + |x - 3|,\ x \in R^1$;
    – $f(x) = |2x_1 + 2x_2| + |2x_1 - 2x_2|,\ (x_1, x_2) \in R^2$;
    – $f(x) = \max\{3x^2 - 3, (x - 3)\},\ x \in R^1$.

**Розділ 5**

# МЕТОДИ РОЗВ'ЯЗУВАННЯ ЗАДАЧ ЛІНІЙНОГО ПРОГРАМУВАННЯ

## 5.1. Приклади задач лінійного програмування (ЛП)

**Приклад 1.** Фірма виготовляє лак для паркету і для меблів із сировини двох типів $C_1$ і $C_2$. В таблиці наведені дані, які характеризують виробничий процес.

| | Витрати сировини в тоннах на 1 тонну лаку для: | | Максимально можливі щоденні витрати сировини (тонн) |
|---|---|---|---|
| | паркету | меблів | |
| Сировина $C_1$ | 4 | 3 | 20 |
| Сировина $C_2$ | 2 | 2,5 | 12,5 |
| Прибуток (в тис. грн) на тонну лаку | 6 | 5 | |



Відділ реалізації обмежив щоденне виробництво лаку для меблів не більше 3-х тонн і встановив вимогу на щоденне виробництво лаку для меблів не більше, ніж на півтори тонни від щоденного виробництва лаку для паркету. Комерційний директор поставив задачу: знайти такі оптимальні об'єми щоденного виготовлення обох видів, які дають найбільший прибуток фірмі.

*Розв'язування.* Для розв'язування цієї задачі введемо такі змінні:
$x_1$ – щоденний об'єм (в тоннах) випуску лаку для паркету;
$x_2$ – щоденний об'єм (в тоннах) випуску лаку для меблів.
Побудуємо цільову функцію (функцію цілі) – щоденний прибуток фірми

$$L(x_1, x_2) = 6x_1 + 5x_2.$$

Запишемо обмеження задачі:

а) щоденні витрати сировини $C_1$

$$4x_1 + 3x_2 \le 20;$$

б) щоденні витрати сировини $C_2$

$$2x_1 + 2,5x_2 \le 12,5;$$

в) щоденне виробництво лаку для меблів не більше 3-х тонн

$$x_2 \le 3;$$

г) щоденне виробництво лаку для меблів не більше, ніж на півтори тонни від виробництва для паркету

$$x_2 - x_1 \le 1,5;$$

д) випуск продукції невід'ємний:

$$x_1 \ge 0;\ \ x_2 \ge 0.$$

Тепер сформулюємо математичну модель задачі лінійного програмування (ЛП):

$$\text{максимізувати}\ \ L(x_1, x_2) = 6x_1 + 5x_2 \qquad (5.1)$$

при обмеженнях:

$$\begin{gathered} 4x_1 + 3x_2 \le\ 20; \\ 2x_1 + 2,5x_2 \le\ 12,5; \\ -x_1 + x_2 \le\ 1,5; \\ x_2 \le 3;\ \ x_1 \ge 0;\ \ x_2 \ge 0. \end{gathered} \qquad (5.2)$$

***Допустимим розв'язком задачі ЛП*** називається будь-яка пара $(\bar{x}_1; \bar{x}_2)^T$, яка задовольняє всі нерівності (5.2); наприклад, точка $(1;2)^T$ є допустимим розв'язком, причому прибуток $L(1;2) = 6{\cdot}1 + 5{\cdot}2 = 16$ тис. грн; точка (3;2) також є допустимим розв'язком, а прибуток $L(3;2) = 6{\cdot}3 + 5{\cdot}2 = 28$ тис. грн. Зрозуміло, що таких допустимих розв'язків існує безліч. Тому для пошуку найкращого (оптимального розв'язку) потрібні ефективні алгоритми.

***Вправи 1.***

1. Серед заданих точок знайти допустимі розв'язки:

а) $x_1 = 1,5;\ \ x_2 = 3,5;$ б) $x_1 = 1,5;\ \ x_2 = 5;$

в) $x_1 = 2,5;\ \ x_2 = 2,5;$ г) $x_1 = 2,5;\ \ x_2 = 3,5;$

д) $x_1 = 2;\ \ x_2 = 3,5;$ е) $x_1 = 4;\ \ x_2 = -0,5;$

2. Побудувати нову область обмежень, враховуючи наступні зміни в постановці задачі:

а) максимально можливі щоденні витрати сировини $C_1$ дорівнюють 21 т., а мінімально можливі щоденні витрати сировини $C_1$ – 16 т.;

б) мінімальне щоденне сумарне виробництво лаків обох типів складає 5 т.;

в) щоденне виробництво лаку для меблів не менше щоденного виробництва лаку для паркету;

г) відношення щоденного виробництва лаку для меблів до загального об'єму виробництва лаків обох типів не повинно перевищувати 2/3.

3. Серед точок із заданої множини допустимих розв'язків $\bar{X} = \{(0,5;\ 0,5)^T, (1;\ 2,5)^T,\ (2;2)^T,\ (0,5;\ 4)^T,\ (1,5;\ 3,5)^T,\ (2,5;\ 2,5)^T,$ $(3;\ 2,25)^T,\ (3,25;2)^T\}$

знайти «найкращу» точку, тобто точку, в якій цільова функція $L(x_1, x_2)$ приймає максимальне значення.

4. Для допустимого розв'язку (4;1) визначити необхідні об'єми сировини $C_1$ і $C_2$.

## 5.2. Графічний спосіб розв'язування задач лінійного програмування

Графічним способом практично завжди можна розв'язувати тільки задачі лінійного програмування розмірності два, хоча відповідні ідеї покладені в основу побудови загальних методів розв'язування задач ЛП (симплекс-методів).

Графічний спосіб розв'язування задач ЛП складається з двох етапів.

<u>Етап I.</u> Будується область допустимих розв'язків, які задовольняють усім обмеженням задачі ЛП.

<u>Етап II.</u> В області допустимих розв'язків знаходиться оптимальний розв'язок (оптимальні розв'язки, чи встановлюється, що задача ЛП оптимального розв'язку не має).



#### 1. Задача ЛП з максимізацією цільової функції

Розв'яжемо графічним способом задачу ЛП із підрозділу 5.1.

$$\text{максимізувати } L(x_1, x_2) = 6x_1 + 5x_2 \tag{5.3}$$

при обмеженнях:

$$\begin{aligned} &4x_1 + 3x_2 \le 20; && \text{(а)} \\ &2x_1 + 2{,}5x_2 \le 12{,}5; && \text{(б)} \\ &-x_1 + x_2 \le 1{,}5; && \text{(в)} \\ &x_2 \le 3; && \text{(г)} \\ &x_1 \ge 0;\ x_2 \ge 0. && \text{(д)} \end{aligned} \tag{5.4}$$

<u>Етап I.</u> ***Побудова області допустимих розв'язків.***

Спочатку будуємо систему координат $(x_1Ox_2)$.

Далі проводимо пряму $4x_1 + 3x_2 = 20$, яка отримується із нерівності (а) заміною знака "$\le$" на знак "=". Для цього достатньо скласти таблицю:

| | | |
|---|---|---|
| $x_1$ | 0 | 5 |
| $x_2$ | 20/3 | 0 |

і через дві точки $(0;\ 20/3)^T$, $(5;0)^T$ провести пряму (а). Пряма (а) ділить площину $(x_1Ox_2)$ на дві півплощини. Підставляємо «тестову» точку $(0;0)^T$

в нерівність (а) і пересвідчуємося, що вона задовольняє цю нерівність, а тому нерівність (а) визначає півплощину, яка містить точку $(0;0)^T$ , (на рис. 5.1 ця півплощина вказана стрілочкою від прямої (а)). Аналогічним чином будуються півплощини, які визначаються нерівностями (б), (в), (г).

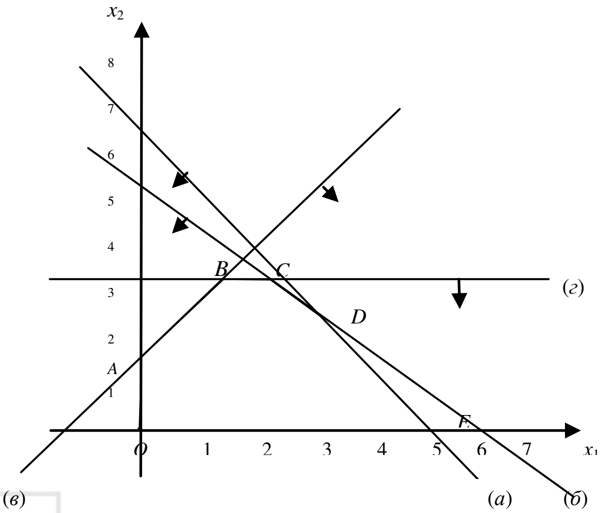



**Рис. 5.1**



Нерівності (д) вказують на те, що область допустимих розв'язків лежить в першому квадранті (вище осі $Ox_1$ і правіше осі $Ox_2$). Таким чином, шестикутник $OABCDE$ визначає область допустимих розв'язків задачі ЛП (5.3), (5.4) (рис. 5.1).

Е т а п  I I. ***Знаходження оптимального розв'язку.***

Знайдемо координати вершин шестикутника $OABCDE$, розв'язавши відповідні системи рівнянь:

$$A:\begin{cases}-x_1+x_2=1,5\\ x_1=0\end{cases}\Leftrightarrow\begin{cases}x_2=1,5\\ x_1=0\end{cases}\Leftrightarrow A(0;1,5);$$

$$B:\begin{cases}-x_1+x_2=1,5\\ x_1=3\end{cases}\Leftrightarrow\begin{cases}x_1=x_2-1,5\\ x_2=3\end{cases}\Leftrightarrow\begin{cases}x_1=1,5\\ x_2=3\end{cases}\Leftrightarrow B(1,5;3);$$

$$C:\begin{cases}x_2=3\\ 2x_1+2,5x_2=12,5\end{cases}\Leftrightarrow\begin{cases}x_2=3\\ x_1=\dfrac{12,5-2,5\cdot x_2}{2}\end{cases}\Leftrightarrow\begin{cases}x_2=3\\ x_1=\dfrac{5}{2}\end{cases}\Leftrightarrow C(2,5;3);$$

$$D:\begin{cases}4x_1+3x_2=20\\2x_1+2,5x_2=12,5\end{cases}\Leftrightarrow\begin{cases}4x_1+3x_2=20\\-2x_2=-5\end{cases}\Leftrightarrow\begin{cases}x_1=\dfrac{20-3x_2}{4}\\x_2=2,5\end{cases}\Leftrightarrow$$

$$\Leftrightarrow\begin{cases}x_1=\dfrac{20-7,5}{4}\\x_2=2,5\end{cases}\Leftrightarrow\begin{cases}x_1=3,125\\x_2=2,5\end{cases}\Leftrightarrow D(3,125;2,5);$$

$$E:\begin{cases}4x_1+3x_2=20\\x_2=0\end{cases}\Leftrightarrow\begin{cases}x_1=5\\x_2=0\end{cases}\Leftrightarrow E(5;0).$$

Побудуємо спочатку пряму $6x_1+5x_2=15$ (ця пряма отримується із (5.3) підстановкою замість $L(x_1,x_2)$ числа 15), склавши таблицю:

| | | |
|---|---|---|
| $x_1$ | 0 | 2,5 |
| $x_2$ | 3 | 0 |

Тепер побудуємо пряму $6x_1+5x_2=25$ (ця пряма отримується із (5.3) підстановкою замість $L(x_1,x_2)$ числа 25), склавши таблицю:

| | | |
|---|---|---|
| $x_1$ | 0 | 25/6 |
| $x_2$ | 5 | 0 |

Очевидно, що прямі $6x_1+5x_2=15$ і $6x_1+5x_2=25$ є паралельними, причому подвійна стрілка вказує напрям зростання цільової функції $L(x_1,x_2)$ (рис. 5.2). Із рис. 5.2 бачимо, що максимального значення цільова функція $L(x_1,x_2)$ набуває в точці $D$ (бо при подальшому паралельному перенесенні прямої $6x_1+5x_2=\text{const}$ в напрямі $\Rightarrow$ перетину цієї прямої з областю допустимих розв'язків не буде). Отже, оптимальний розв'язок задачі ЛП (5.3), (5.4) $-x^*_1=3,125;\ x^*_2=2,5,$ при цьому значення цільової функції $L(x^*_1,x^*_2)=6\cdot3,125+5\cdot2,5=31,25$. Звідси можна зробити висновок, що для фірми найвигідніше щоденно випускати 3,125 т. лаку для паркету та 2,5 т. лаку для меблів і отримувати щоденний прибуток 31,25 тис. грн.

Із рис. 5.2 випливає, що якщо змінювати коефіцієнти цільової функції $L(x_1,x_2)$ (тобто змінювати кут нахилу прямих $L(x_1,x_2)=ax_1+bx_2$), то оптимальний розв'язок буде розміщений або у вершинах шестикутника *OABCDE*, або оптимальними розв'язками будуть всі точки відрізків *AB*, чи *BC*, чи *CD* тощо. У цьому і полягає основна ідея побудови загального симплексного методу розв'язування задач ЛП, який буде розглянуто в наступних підрозділах.

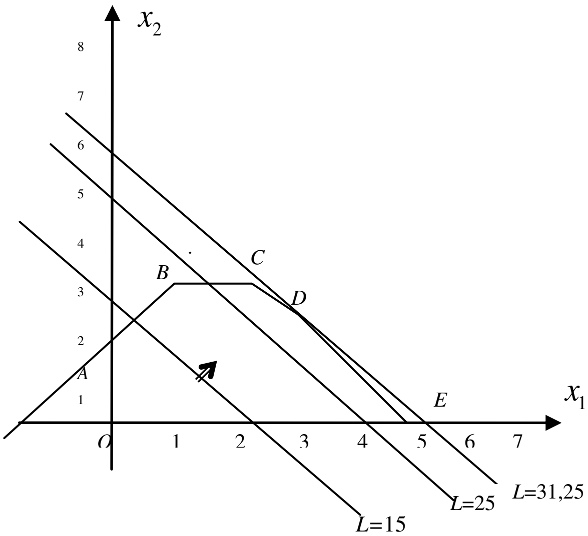



**Рис. 5.2**

***Вправи 1.***

1. Для кожної нерівності визначити допустиму півплощину:

а) $1,5x_1 + 2x_2 \geq 4$;

б) $-4x_1 + 7x_2 \geq 3$;

в) $3x_1 + 3,5x_2 \leq -2$.

Національний університет
водного господарства
та природокористування

2. Визначити напрям зростання цільової функції $L(x_1, x_2)$:

а) максимізувати $L(x_1, x_2) = 2x_1 + 1,5x_2$;

б) максимізувати $L(x_1, x_2) = -5x_1 - 0,5x_2$;

в) максимізувати $L(x_1, x_2) = 1,5x_1 - 0,3x_2$.

3. Для моделі (5.3), (5.4) побудувати область допустимих розв'язків і знайти оптимальний розв'язок при наступних незалежних змінах умов задачі:

а) щоденний об'єм виробництва лаку для меблів не повинен перевищувати 2,7 т. і бути не меншим 1,2 т.;

б) щоденний об'єм виробництва лаку для паркету повинен бути не меншим 1 т. і не перевищувати 4,5 т.;

в) щоденні витрати сировини $C_2$ повинні бути не менші 8 т.

4. Для області допустимих розв'язків (5.4) знайти оптимальний розв'язок для наступних цільових функцій:

а) $L(x_1, x_2) = x_1 + 2,5x_2$;

б) $L(x_1, x_2) = 10x_1 + 5x_2$;

в) $L(x_1, x_2) = 4x_1 + x_2$.

**2. Задача ЛП з мінімізацією цільової функції**

**Приклад 1 (задача «дієти»).** Комбікормовий завод щоденно виготовляє не менше 1200 т. комбікормів, які складаються із суміші вівса і пшениці, склад якої задається в таблиці.

| | Процентний вміст (%) | | Вартість (тис. грн. за 1 т.) |
|---|---|---|---|
| | Білок | Клітковина | |
| Овес | 7 | 1 | 0,5 |
| Пшениця | 50 | 7 | 0,8 |

Замовники продукції вимагають, щоб у комбікормах було не менше 20% білків і не більше 6% клітковини. Головний технолог заводу хоче визначити склад суміші з урахуванням вимог замовників і виробництва, мінімізуючи вартість комбікормів.

Вводимо змінні задачі, які необхідно визначити:

$x_1$ – кількість вівса у тоннах, який використовують у щоденному виробництві комбікормів;

$x_2$ – кількість пшениці у тоннах, яка використовується в щоденному виробництві комбікормів.

Отже, цільова функція, яку потрібно мінімізувати має вигляд:

$$L(x_1, x_2) = 0,5x_1 + 0,8\,x_2 .$$

Вимога щоденного виробництва комбікормів – не менше 1200 т. записується таким чином:

$$x_1 + x_2 \geq 1200.$$

Кількість білка в $x_1$ тоннах вівса дорівнює $0,07x_1$ (тонн), а кількість білка в $x_2$ тоннах пшениці дорівнює $0,5x_2$ (тонн). Сумарна кількість білка повинна бути не менша 20% від маси суміші $x_1 + x_2$. Таким чином, отримуємо нерівність

$$0,07x_1 + 0,5x_2 \geq 0,2\ (x_1 + x_2).$$

Аналогічно міркуючи, отримуємо нерівність для клітковини

$$0,01x_1 + 0,07x_2 \leq 0,06(x_1 + x_2).$$

Після зведення подібних членів у двох останніх нерівностях можемо сформулювати задачу ЛП з мінімізацією цільової функції:

$$\text{мінімізувати } L(x_1, x_2) = 0,5x_1 + 0,8\,x_2$$

при обмеженнях:

$$x_1 + x_2 \geq 1200; \qquad (а)$$
$$0,13x_1 - 0,3x_2 \leq 0; \qquad (б)$$
$$0,05x_1 - 0,01x_2 \geq 0; \qquad (в)$$
$$x_1 \geq 0,\ \ x_2 \geq 0.$$

На рис. 5.3 зображена допустима область розв'язків *ABCD* (до речі, ця область є необмеженою). Зауважимо, що прямі $0,13x_1 - 0,3x_2 = 0$ і $0,05x_1 - 0,01x_2 = 0$ проходять через початок координат (0;0), тому для визначення півплощин, які визначаються нерівностями (б) і (в) , в якості тестової точки потрібно брати не $(0;0)^T$, а якусь іншу точку , наприклад, $(0;\ 500)^T$. Побудуємо також дві прямі , які відповідають цільовій функції при $L(x_1, x_2) = 200$ і $L(x_1, x_2) = 1000$, і таким чином, отримаємо напрям спадання цільової функції, який позначимо стрілкою $\Rightarrow$. Найменшого значення цільова функція $L(x_1, x_2)$ приймає в точці *B*, яка є точкою перетину прямих (а) і (б) .



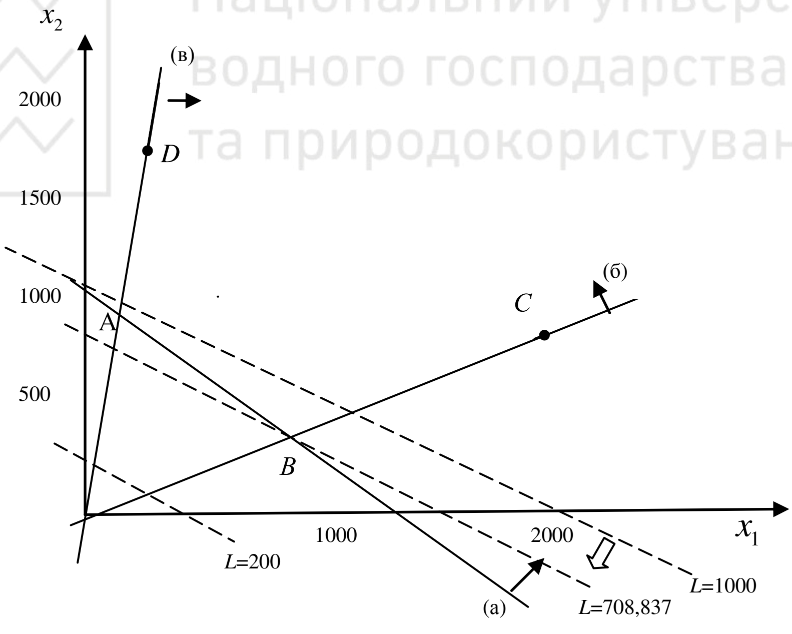



**Рис. 5.3**

Розв'язавши систему рівнянь

$$\begin{cases} x_1 + x_2 = 1200, \\ 0,13x_1 - 0,3x_2 = 0, \end{cases}$$

знайдемо координати точки $B$: $B\left(36000/43;15600/43\right)$. Отже, оптимальна кількість вівса, яка використовується в щоденному виробництві комбікормів складає $36000/43 \approx 837,21$ т., а пшениці – $15600/43 \approx 362,79$ т.; при цьому мінімальна вартість комбікормів складає $0,5 \cdot 36000/43 + 0,8 \cdot 15600/43 \approx 708,837$ тис. грн.

***Вправи 2.***

1. Необхідно побудувати нову допустиму область розв'язків і знайти оптимальний розв'язок при наступних змінах:

а) кількість вівса, що використовується в щоденному виробництві комбікормів не перевищує 500 т.;

б) кількість вівса, що використовується в щоденному виробництві комбікормів не перевищує 900 т., а пшениці – 1100 т.;

в) щоденне виробництво комбікормів не більше 1200 т.

2. Визначити напрям спадання таких цільових функцій:

а) мінімізувати $L\left(x_1, x_2\right) = 1,7x_1 + 1,5x_2$;

б) мінімізувати $L\left(x_1, x_2\right) = -4x_1 + 5x_2$;

в) мінімізувати $L\left(x_1, x_2\right) = 0,6x_1 - 2,2x_2$.



### 3. Додаткові змінні в задачах лінійного програмування

У даному пункті вводяться додаткові невід'ємні змінні: залишкова та надлишкова, які зв'язані з нерівностями «≤», «≥», відповідно; і вільна змінна, яка може приймати будь-яке значення із $R$.

А. Залишкова змінна. Нерівності «≤» в обмеженнях задачі ЛП в основному інтерпретують, як обмеження на використання ресурсів, а залишкова змінна показує кількість невикористаного ресурсу. Для прикладу 1 із підрозділу 5.1 розглянемо перше обмеження пов'язане із щоденними витратами сировини $C_1$

$$4x_1 + 3x_2 \le 20.$$

Ця нерівність еквівалентна рівності:

$$4x_1 + 3x_2 + s_1 = 20,$$

де залишкова змінна $s_1$ $(s_1 \ge 0)$ дорівнює невикористаній кількості сировини $C_1$

$$s_1 = 20 - \left(4x_1 + 3x_2\right).$$

Б. Надлишкова змінна. Нерівність «≥» вказує на те, «що щось» повинно бути не менше від певної величини. *Надлишкова* змінна визначає перевищення значення лівої частини нерівності над цією величиною. В прикладі 2 нерівність

$$x_1 + x_2 \geq 1200$$

показує, що щоденне виробництво комбікормів не менше 1200 т. Математично ця нерівність еквівалентна рівності

$$x_1 + x_2 - s_2 = 1200,$$

де $s_2 \geq 0$. Додатне значення надлишкової змінної $s_2$ показує перевищення щоденного виробництва комбікормів над мінімальним значенням в 1200 т.

В. Вільна змінна. В наступному прикладі описується ситуація, коли змінні (вільні) можуть приймати будь-які дійсні значення (додатні, від'ємні і нуль).

**Приклад 2.** Кафе виготовляє і торгує м'ясними котлетами і гуляшем. На порцію котлет йде 0,15 кг. м'яса, а на гуляш – 0,10 кг. м'яса. На початок робочого дня в кафе в наявності є 100 кг. м'яса, причому м'ясо можна ще докупити впродовж дня, але уже з націнкою в 15 коп. за 1 кг. М'ясо, яке залишається, в кінці робочого дня віддається (безкоштовно) в притулок для бездомних. Дохід кафе від одної порції котлет 25 коп., а від одної порції гуляшу 20 коп. За день кафе не може реалізувати більше 700 порцій страв. Яка повинна бути доля кожного страви (тобто, яка кількість котлет і яка кількість гуляшу) в щоденному виробництві кафе, щоб його прибуток був максимальний?



Позначимо кількість порцій котлет – $x_1$, а кількість порцій гуляшу – $x_2$, які виготовляє кафе щоденно. Якщо кафе не докуповує м'ясо протягом дня, то маємо таке обмеження

$$0,15x_1 + 0,10x_2 \leq 100,$$

а якщо докуповує, то маємо наступне обмеження

$$0,15x_1 + 0,10x_2 \geq 100.$$

Оскільки ми не знаємо, яка із цих нерівностей буде мати місце, то логічно використати рівність

$$0,15x_1 + 0,10x_2 + x_3 = 100,$$

де $x_3$ – вільна змінна, яка відіграє роль як залишкової, так і надлишкової змінних. Скористаємось стандартним математичним прийомом для представлення вільної змінної $x_3$:

$$x_3 = x_3^+ - x_3^-,$$

де $x_3^+ \geq 0$ і $x_3^- \geq 0$.

Якщо $x_3^+ > 0$ і $x_3^- = 0$, то змінна $x_3$ відіграє роль залишкової. Якщо $x_3^- > 0$ і $x_3^+ = 0$, то змінна $x_3$ відіграє роль надлишкової. Таким чином, наше обмеження можна записати у вигляді рівності

$$0,15x_1 + 0,10x_2 + x_3^+ - x_3^- = 100,$$

а цільова функція приймає вигляд:

$$\text{максимізувати } L(x_1, x_2, x_3^-) = 0,25x_1 + 0,20x_2 - 0,15x_3^-.$$

***Вправи 3.***

1. Для допустимого розв'язку $x_1 = 3,\ x_2 = 2$ (приклад 1, п.5.1) знайти кількість недовикористаної сировини $C_1$ і $C_2$.
2. Знайти перевищення над мінімальним допустимим об'ємом виробництва комбікормів (приклад 1, п.5.2), на які витрачається 700 т. вівса і 800 т. пшениці.

#### 4. Особливі випадки розв'язування задач лінійного програмування

I. Приклад задачі ЛП, яка не має розв'язку в допустимій області:

$$L(x) = 4x_1 + 2x_2 \to \max$$

при обмеженнях:

$$-x_1 + 2x_2 \le 4, \qquad (а)$$

$$2x_1 + 5x_2 \ge 5, \qquad (б)$$

$$x_1 \ge 0, x_2 \ge 0.$$



*Розв'язування.* Спочатку побудуємо допустиму область $X$ :

$-x_1 + 2x_2 = 4$ :

| | | |
|---|---|---|
| $x_1$ | 0 | -4 |
| $x_2$ | 3 | 0 |

$2x_1 + 5x_2 = 5$ :

| | | |
|---|---|---|
| $x_1$ | 0 | 2,5 |
| $x_2$ | 1 | 0 |

З рис. 5.4 бачимо, що допустима область $X$ є необмеженою. Тепер побудуємо лінії рівня цільової функції $L(x) = c_1$. Для цього знайдемо градієнт $L(x)$ : $\nabla L(x) = \begin{pmatrix} \partial L/\partial x_1 \\ \partial L/\partial x_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$ і побудуємо цей градієнт як вектор $OM$.

Проведемо пряму перпендикулярну до $OM$ і подвійною стрілкою вкажемо напрям зростання цільової функції $L(x)$. Зрозуміло, що задача ЛП розв'язку немає (тобто $\sup\limits_{x \in X} L(x) = +\infty$ ).

За аналогією не важко побудувати приклад задачі ЛП, яка також не буде мати розв'язку, причому $\inf\limits_{x \in X} L(x) = -\infty$.

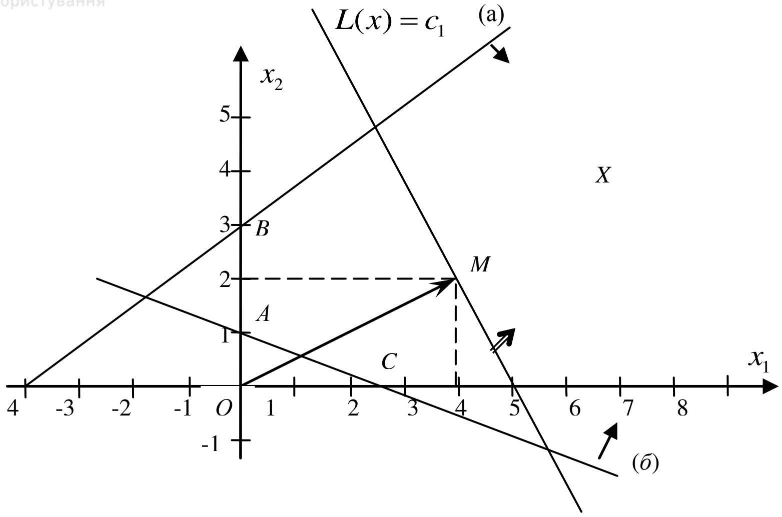



**Рис. 5.4**

II. Приклад задачі ЛП, яка має безліч розв'язків в допустимій області:

$$L(x) = 9x_1 + 4x_2 \to \max$$

при обмеженнях:

$$-x_1 + 2x_2 \le 5, \qquad (а)$$
$$9x_1 + 4x_2 \le 40, \qquad (б)$$
$$3x_1 + 5x_2 \ge 4, \qquad (в)$$
$$x_1 \ge 0, \quad x_2 \ge 0.$$



*Розв'язування.* Спочатку побудуємо допустиму область $X$ :

$-x_1 + 2x_2 = 5$:

| | | |
|---|---|---|
| $x_1$ | 0 | 3 |
| $x_2$ | 2,5 | 4 |

$9x_1 + 4x_2 = 40$:

| | | |
|---|---|---|
| $x_1$ | 0 | 40/9 |
| $x_2$ | 10 | 0 |

$3x_1 + 5x_2 = 4$:

| | | |
|---|---|---|
| $x_1$ | 0 | 4/3 |
| $x_2$ | 4/5 | 0 |

З рис. 5.5 бачимо, що допустимою областю розв'язків $X$ є п'ятикутник *ABCDE*. Тепер будуємо лінії рівня цільової функції $L(x) = c_2$. Знаходимо градієнт $\nabla L(x) = \begin{pmatrix} 9 \\ 4 \end{pmatrix}$ і наносимо його на графік. Проводимо пряму перпендикулярно до $\nabla L(x)$ і подвійною стрілкою вказуємо напрям зростання цільової функції $L(x)$.

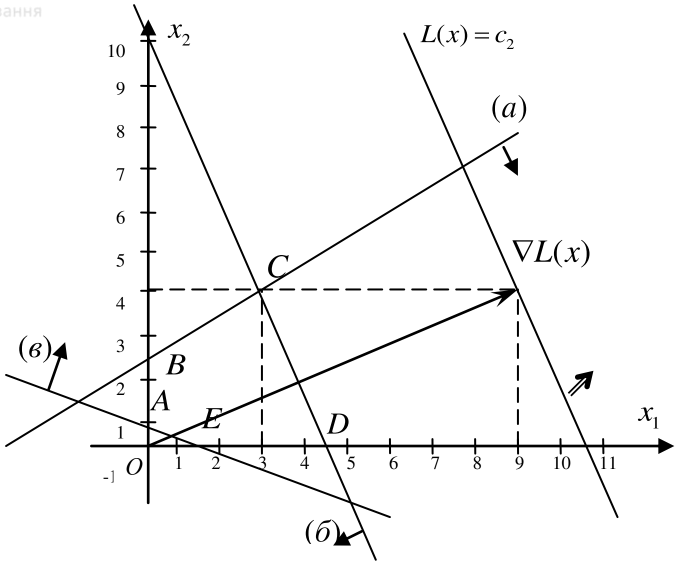



**Рис. 5.5**

Приймемо до уваги, що кутові коефіцієнти лінії рівня цільової функції та прямої, яка відповідає нерівності (б), однакові $k = -9/4$ і робимо висновок, що всі точки відрізка $CD$ є розв'язками задачі ЛП.

## 5.3. Графічний аналіз чутливості

Важливою є проблема дослідження розв'язку задачі ЛП в залежності від змін коефіцієнтів цільової функції і коефіцієнтів в нерівностях, які задають обмеження. Такі дослідження називають ***аналіз чутливості***. В цьому підрозділі ми проведемо аналіз чутливості графічним способом змінюючи:

- коефіцієнти цільової функції;
- значення констант в правій частині нерівностей, які задають обмеження.

#### 1. Аналіз чутливості коефіцієнтів цільової функції

Цільова функція задачі ЛП у випадку двох змінних має вигляд

$$L = L(x_1, x_2) = c_1 x_1 + c_2 x_2 ; \tag{5.5}$$

Представимо цільову функцію (при $c_1 \neq 0$) наступним чином

$$x_2 = -\frac{c_1}{c_2} x_1 + \frac{L}{c_2} . \tag{5.6}$$

Зміна значень коефіцієнтів $c_1$ і $c_2$ цільової функції (5.5) призводить до зміни кута нахилу прямої (5.6). Із графічного способу розв'язання задачі ЛП витікає, що при суттєвій зміні кута нахилу прямої (5.6) оптимальний розв'язок може досягатися в іншій кутовій точці області допустимих розв'язків. Також очевидно, що існує інтервал зміни відношень $c_1 / c_2$, при яких оптимальний розв'язок не змінюється. Задача аналізу чутливості коефіцієнтів цільової функції полягає у визначенні *інтервалу оптимальності* для відношення $c_1 / c_2$ тобто: якщо значення $c_1 / c_2$ належить цьому інтервалу, то оптимальний розв'язок задачі ЛП залишається незмінним. Продемонструємо аналіз чутливості коефіцієнтів цільової функції на прикладі задачі ЛП підрозділу 5.2.

**Приклад 1.** Зобразимо в системі координат $x_1Ox_2$ область допустимих розв'язків (5.4) і цільову функцію (5.3) при $L(x_1, x_2) = 31.25$ для задачі ЛП із підрозділу 5.2.

З рис. 5.6 бачимо, що цільова функція (5.3) досягає максимального значення в кутовій точці *D*. При зміні коефіцієнтів цільової функції $L(x_1, x_2) = c_1x_1 + c_2x_2$ точка *D* залишається оптимальним розв'язком поки кут нахилу лінії $L$ буде лежати між кутами нахилу прямих $4x_1 + 3x_2 = 20$ (що еквівалентно $x_2 = -\left(4/3\right)x_1 + 20/3$) і $2x_1 + 2{,}5x_2 = 12{,}5$ (що еквівалентно $x_2 = -\left(2/2{,}5\right)x_1 + 12{,}5/2{,}5$) (нагадуємо, що перетин цих прямих визначає точку *D*).

Алгебраїчно це записується у вигляді таких нерівностей:

$$-\frac{4}{3} \le -\frac{c_1}{c_2} \le -\frac{2}{2{,}5}.$$

Помноживши нерівності на (-1) , отримаємо:

$$0{,}8 \le \frac{c_1}{c_2} \le \frac{4}{3}. \tag{5.7}$$

Отже, якщо коефіцієнти цільової функції (5.5) задовольнять нерівності (5.7), то оптимальний розв'язок задачі ЛП буде досягатися в точці *D*. Також відмітимо, що якщо пряма $z = c_1x_1 + c_2x_2$ співпадає із прямою $4x_1 + 3x_2 = 20$, то оптимальним розв'язком буде будь-яка точка відрізка *DE*; аналогічно: якщо пряма $z = c_1x_1 + c_2x_2$ співпадає із прямою $2x_1 + 2{,}5x_2 = 12{,}5$, то оптимальним розв'язком буде будь-яка точка відрізка *CD*.

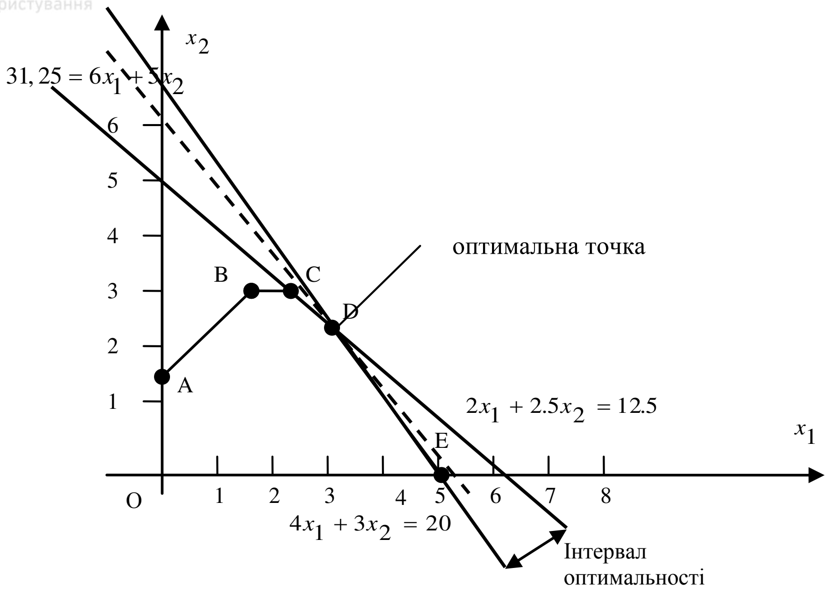



**Рис. 5.6**

Нерівність (5.7) можна також використовувати для визначення інтервалу оптимальності для одного із коефіцієнтів цільової функції (5.5) при умові, що інший коефіцієнт залишається постійним, наприклад, якщо в нашій моделі ми зафіксуємо значення коефіцієнта $c_2$ ($c_2 = 5$), тоді інтервал оптимальності для коефіцієнта $c_1$ отримуємо із нерівності (5.7) при $c_2 = 5$:

$$0,8 \le \frac{c_1}{5} \le \frac{4}{3} \quad \Leftrightarrow \quad 4 \le c_1 \le \frac{20}{3}.$$

***Вправи 1.***

1. Для прикладу 1 п 5.3 визначити інтервал оптимальності коефіцієнта $c_2$ при умові, що коефіцієнт $c_1$ є незмінним ($c_1 = 6$).

2. Для задачі ЛП максимізувати $L(x_1, x_2) = 4x_1 + x_2$ при обмеженнях:

$$2x_1 + 3x_2 \le 12,$$
$$2x_1 - x_2 \le 4,$$
$$x_1 \ge 0, x_2 \ge 0,$$

графічно визначити:

а) інтервал оптимальності для відношення коефіцієнтів $c_1 / c_2$ цільової функції;

б) інтервал оптимальності для коефіцієнта $c_1$ при фіксованому значенні коефіцієнта $c_2$ ($c_2 = 1$);

в) інтервал оптимальності для коефіцієнта $c_2$ при фіксованому значенні коефіцієнта $c_1$ ($c_1 = 4$).

3. Для задачі “дієти” із прикладу 1 п. 5.2:

а) визначити інтервал оптимальності для відношення вартості тонни вівса до вартості тонни пшениці;

б) чи залишиться оптимальним раніше знайдений розв’язок (точка В), якщо вартість тонни вівса залишиться постійною (0,5 тис. грн.), а вартість тони пшениці зменшиться до 0,7 тис. грн.?

в) чи залишиться оптимальним раніше знайдений розв’язок (точка В), якщо вартість тонни вівса зменшиться на 20%, а вартість тонни пшениці збільшиться на 20%?

#### 2. Аналіз чутливості коефіцієнтів функцій обмежень. Вартість ресурсів

В багатьох задачах ЛП права частина нерівностей в обмеженнях являє собою верхню границю доступних ресурсів. В цьому пункті вивчається чутливість оптимального розв’язку задачі ЛП до змін обмежень, які накладаються на ресурси. Такий аналіз задачі ЛП пропонує просту міру чутливості розв’язку, яку називають ***вартістю одиниці ресурсу*** – при зміні кількості доступних ресурсів (на одиницю) значення цільової функції в оптимальному розв’язку зміниться на вартість одиниці ресурсу. Проілюструємо цей вид аналізу на прикладі.



**Приклад 2.**

В моделі прикладу 1 п.5.1 перші дві нерівності представляють собою обмеження на використання сировини $C_1$ і $C_2$ відповідно. Визначимо вартість одиниць цих ресурсів. Почнемо з обмеження для сировини $C_1$. Нагадаємо, що в даній задачі ЛП оптимальний розв’язок досягається в точці *D*, яка є точкою перетину прямих, що відповідають обмеженням на сировину $C_1$ і $C_2$ (рис. 5.7).

При зменшенні рівня доступності сировини $C_1$ оптимальний розв’язок задачі ЛП буде прямувати по відрізку *DC* до точки *C*, а при збільшенні рівня доступності сировини $C_1$ оптимальний розв’язок задачі ЛП буде прямувати по відрізку *DF* до точки *F*. Зрозуміло, що зміна рівня доступності сировини $C_1$, яка призводить до виходу точки перетину *D* із відрізка *CF*, веде до нездійсненності оптимального розв’язку в точці *D*.

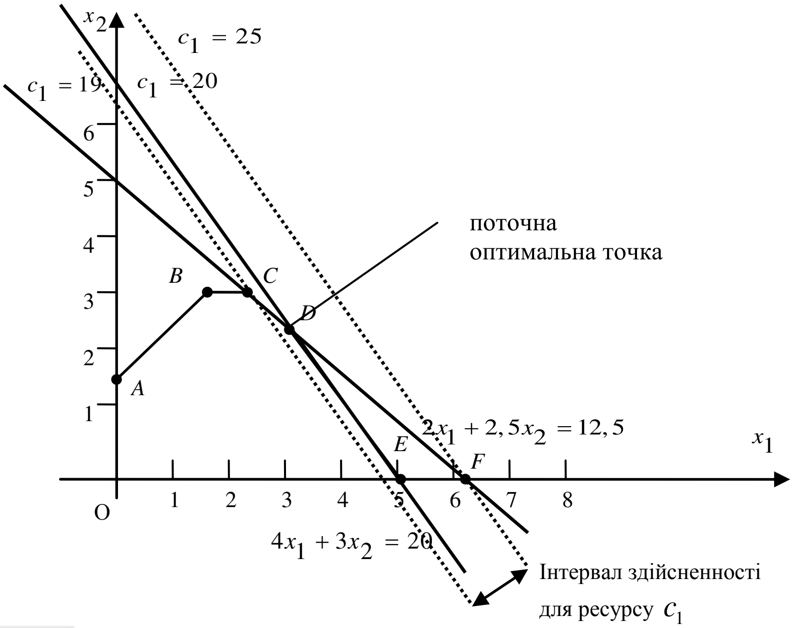



**Рис. 5.7**



Тому кінцеві точки *C*(2,5;3) і *F*(6,25;0) відрізка *CF* визначають ***інтервал здійсненності*** для ресурсу $C_1$. Кількість сировини $C_1$, яка відповідає точці *C*(2,5;3), дорівнює:

$$4x_1 + 3x_2 = 4*2,5 + 3*3 = 10 + 9 = 19 \text{ т.}$$

Аналогічно кількість сировини $C_1$, яка відповідає точці *F*(6,25;0) дорівнює:

$$4x_1 + 3x_2 = 4*6,25 + 3*0 = 25 \text{ т.}$$

Таким чином, інтервал здійсненності для ресурсу $C_1$ наступний

$$19 \le c_1 \le 25 .$$

Можемо стверджувати, що існуючий рівень ресурсу $C_1 = 20$ т. може бути зменшений на 1 тонну або збільшений на 5 тонн, при цьому гарантується , що оптимальний розв'язок буде досягатися в точці *D* – точці перетину прямих, які відповідають обмеженням на ресурси $C_1$ і $C_2$.

Обчислимо вартість одиниці сировини $C_1$. При зміні кількості сировини $C_1$ від 19 до 25 тонн значення цільової функції *L* будуть відповідати положенню точки *D* на відрізку *CF*. Позначимо через $y_1$ вартість одиниці ресурсу $C_1$. Справедлива формула

$$y_1 = \frac{\textit{Зміна} \text{ значення } L \text{ при переміщенні т. } D \text{ від } C \text{ до } F}{\textit{Зміна} \text{ кількості } \mathrm{C}_1 \text{ при переміщенні т. } D \textit{ від } C \text{ до } F}.$$

Якщо точка $D$ співпадає з точкою $C(2{,}5;3)$, то $L = 6*2{,}5 + 5*3 = 30$ (тис. грн.); якщо ж точка $D$ співпадає з точкою $F(6{,}25;0)$, то $L = 6*6{,}25 + 5*0 = 37{,}25$ (тис. грн.). Тепер можемо обчислити вартість одиниці ресурсу $C_1$:

$$y_1 = \frac{37{,}25 - 30}{25 - 19} = \frac{7{,}25}{6} = 1{,}25 \text{ (тис. грн. на тонну ресурсу } C_1\text{)}.$$

Цей результат показує, що зміна кількості ресурсу $C_1$ на одну тонну (якщо загальна кількість ресурсу не менше 19 тонн і не більше 25 тонн) призводить до зміни значення цільової функції в оптимальному розв'язку на 1,25 тис. грн.

Аналогічно проводиться аналіз щодо ресурсу $C_2$. На рис. 5.8 бачимо, що інтервал здійсненності для ресурсу $C_2$ визначається кінцевими точками *E* і *G* відрізка *EG*, де $E(5;0)$ і $G(11/4;3)$.

Точка *G* знаходиться на перетині прямих *BC* і *ED*. Знаходимо кількість сировини $C_2$, яка відповідає точці *E*:

$$2x_1 + 2{,}5x_2 = 2*5 + 2{,}5*0 = 10 \text{ т.}$$

і кількість сировини $C_2$, яка відповідає точці *G*:

$$2x_1 + 2{,}5x_2 = 2*11/4 + 2{,}5*3 = 5{,}5 + 7{,}5 = 13 \text{ т.}$$

Значення цільової функції в точці *E* дорівнює:

$$6x_1 + 5x_2 = 6*5 + 5*0 = 30 \text{ (тис. грн.)},$$

а в точці *G*:

$$6x_1 + 5x_2 = 6*11/4 + 5*3 = 33/2 + 15 = 63/2 = 31{,}5 \text{ (тис. грн.)}.$$

Обчислимо вартість ресурсу $C_2$:

$$y_2 = \frac{31{,}5 - 30}{13 - 10} = 0{,}5 \text{ (тис. грн. на тонну сировини } C_2\text{)}.$$

Отже, зміна кількості сировини $C_2$ на одну тонну призводить до зміни значення цільової функції в оптимальній точці на 0,5 тис. грн.

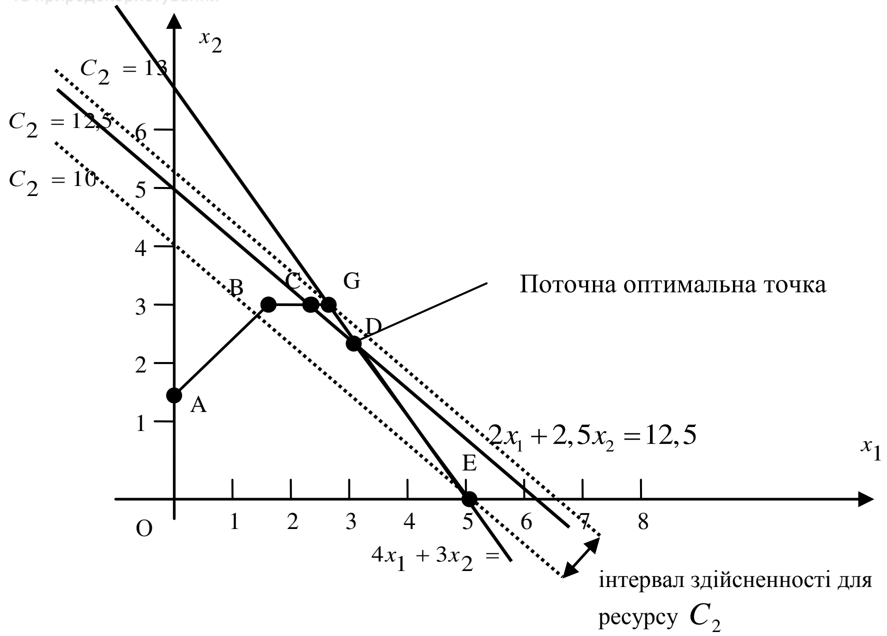



**Рис. 5.8**

***Вправи 2.***

1. Завод виготовляє продукцію двох типів П1 і П2. Об’єм реалізації продукту П1 за день складає не менше 70% від загальної реалізації продуктів П1 і П2. Технологічно виробництво продукції П1 не може перевищувати 70 одиниць в день. У виробництві обох типів продукції використовується сировина $C_1$, щоденне використання якої не повинно перевищувати 400 т. При виробництві одиниці продукції П1 використовується 4 т. сировини $C_1$, а одиниці продукції П2 – 8 т. сировини $C_1$. Вартість одиниці продукції П1 складає 30 тис. грн., а одиниці продукції П2 – 120 тис. грн. Визначити:

а) оптимальну структуру виробництва заводу (максимізувати сумарну вартість виготовленої за день продукції обох типів);

б) вартість одиниці сировини $C_1$;

в) інтервал зміни сировини $C_1$, при якому справедлива дана вартість одиниці сировини $C_1$;

г) (за допомогою графічного аналізу чутливості) як зміниться значення цільової функції при зміні максимального рівня виробництва продукції П1 на $\pm 10\%$, $\pm 20\%$, $\pm 30\%$.

## 5.4. Загальна задача лінійного програмування

***Загальна задача ЛП (ЗЗЛП)*** формулюється так:
знайти вектор $x = (x_1, x_2, ..., x_n)^T \in R^n$, який мінімізує (максимізує) лінійну цільову функцію

$$L(x) = c_1 x_1 + c_2 x_2 + ... + c_n x_n = \sum_{j=1}^{n} c_j x_j \tag{5.8}$$

і задовольняє систему лінійних обмежень:

$$\sum_{j=1}^{n} a_{ij} x_j \theta_i b_i, \qquad i = \overline{1,m}; \tag{5.9}$$

$$x_j \geq 0, \quad j = \overline{1,k}, \quad k \leq n, \tag{5.10}$$

де $c_j,\ j = \overline{1,n};\ a_{ij},\ i = \overline{1,m},\ j = \overline{1,n},\ b_i,\ i = \overline{1,m}$ – задані дійсні числа; $\theta_i,\ i = \overline{1,m},$ – один із знаків $\geq,\ \leq,\ =$.

Обмеження (5.10) мають назву ***прямих***, обмеження (5.9) – ***непрямих*** обмежень. Вектор $x$, який задовольняє обмеженням (5.9) і (5.10) називається ***допустимим розв'язком*** (***допустимим вектором, точкою, планом***) задачі ЛП. Множина всіх допустимих розв'язків ЗЗЛП називається ***допустимою областю*** (***допустимою множиною***) ЗЗЛП. Позначимо цю множину буквою $X$:

$$X = \left\{ x = (x_1, ..., x_n)^T \in R^n : \sum_{j=1}^{n} a_{ij} x_j \theta_i b_i,\ i = \overline{1,m};\ x_j \geq 0,\ j = \overline{1,k},\ k \leq n \right\}.$$

Вектор $x^* \in X$, який доставляє цільовій функції $L(x)$ мінімум (максимум) в допустимій області, називається ***оптимальним розв'язком*** (***оптимальним вектором, точкою, планом***) ЗЗЛП

$$x^* = \arg\min_{x \in D} L(x) \qquad \left( x^* = \arg\max_{x \in D} L(x) \right).$$

Зауважимо, що задача максимізації цільової функції $L(x)$ еквівалентна задачі мінімізації цільової функції $L_1(x) = -L(x)$ на допустимій області $X$. Значення цільової функції в оптимальній точці $L(x^*)$ називається ***оптимальним значенням цільової функції***.

Множина точок $x = (x_1, x_2, ..., x_n)^T \in R^n$, яка задовольняє умову

$$\sum_{j=1}^{n} a_{ij} x_j \geq b_i \qquad \left(\sum_{j=1}^{n} a_{ij} x_j \leq b_i\right),\ \ i = \overline{1,m}$$

називається ***півпростором*** в $R^n$, а множина точок $x = (x_1, x_2, ..., x_n)^T \in R^n$, яка задовольняє умову

$$\sum_{j=1}^{n} a_{ij} x_j = b_i, \quad i = \overline{1,m}$$

називається ***гіперплощиною*** в $R^n$.

*Означення 1.* ***Многогранною множиною*** називається перетин скінченого числа півпросторів.

*Означення 2.* Обмежена многогранна множина називається ***многогранником***.

Відомо, що допустима область $X$ (5.9) є опуклою многогранною множиною. Отже, ЗЗЛП полягає у знаходженні мінімуму (максимуму) лінійної цільової функції $L(x)$ на опуклій многогранній множині.

*Означення 3.* Точка $x$ опуклої множини $X$ називається ***кутовою*** (***крайньою***), якщо її не можна представити у вигляді:

$$x = \alpha x^1 + (1-\alpha) x^2,$$

де $\alpha \in (0;1)$; $x^1$, $x^2 \in X$, $x^1 \neq x^2$.

*Означення 4.* Кутові точки опуклої многогранної множини називають її ***вершинами***.

***Теорема 1.*** *Нехай в ЗЗЛП допустима область $X$ є многогранником. Тоді цільова функція $L(x)$ досягає оптимального значення у вершині многогранника. Якщо ж цільова функція досягає оптимального значення у двох чи більше точках, то вона досягає оптимального значення у будь-якій точці, що є їх опуклою лінійною комбінацією.*

З наведеної теореми випливає, що розв'язки ЗЗЛП потрібно шукати серед вершин її допустимої області. Перебрати всі вершини допустимої області в реальних задачах майже нереально, тому для розв'язування ЗЗЛП будуть запропоновані ефективні алгоритми, які здійснюють цілеспрямований перебір незначної кількості вершин допустимої області.

## 5.5. Стандартна задача лінійного програмування. Базис, базисний розв'язок

Спочатку запишемо *стандартну задачу лінійного програмування* (*СЗЛП*) *у* ***координатній формі***:

$$L(x) = \sum_{j=1}^{n} c_j x_j \to \min \qquad (5.11)$$

при обмеженнях:

$$\sum_{j=1}^{n} a_{ij} x_j = b_i, \; i = \overline{1,m}, \qquad (5.12)$$

$$x_j \geq 0, \;\; j = \overline{1,n}. \qquad (5.13)$$

Якщо ввести матрицю $A = \left(a_{ij}\right)_{i=\overline{1,m}}^{j=\overline{1,n}}$, вектори $c = (c_1, c_2, ..., c_n)^T$, $b = (b_1, b_2, ..., b_m)^T$, то прийдемо до ***матричної форми*** СЗЛП:

$$L(x) = (c, x) \to \min \tag{5.14}$$

при обмеженнях:

$$Ax = b; \tag{5.15}$$

$$x \ge 0. \tag{5.16}$$

Позначивши стовпці матриці $A$ через $A_j$, $j = \overline{1,n}$, $A_j = (a_{1j}, ..., a_{mj})^T$, отримаємо ***векторну форму*** СЗЛП:

$$L(x) = (c, x) \to \min \tag{5.17}$$

при обмеженнях:

$$\sum_{j=1}^{n} A_j x_j = b; \tag{5.18}$$

$$x \ge 0. \tag{5.19}$$

Будь-яка ЗЗЛП зводиться до стандартної задачі ЛП наступним чином:

а) нерівності обмеження перетворюються у рівності за допомогою введення в ліву частину нерівності (5.9) додаткових змінних – залишкових і надлишкових (ці змінні введені в підрозділі 5.2).

*Приклад нерівності типу* "$\le$".

Нерівність $2x_1 + 3x_2 \le 7$, еквівалентна рівності $2x_1 + 3x_2 + s_1 = 7$, де $S_1$ – залишкова змінна, причому $s_1 \ge 0$.

*Приклад нерівності типу* "$\ge$". Нерівність $1{,}5x_1 - 4x_2 \ge 3$, еквівалентна рівності $1{,}5x_1 - 4x_2 - s_2 = 3$, де $s_2$ – надлишкова змінна, причому $s_2 \ge 0$;

б) вільну змінну $x_j$ (тобто змінну, яка може приймати і від'ємні і додатні значення, а також нуль) можна представити як різницю двох невід'ємних змінних:

$$x_j = x_j^{\ +} - x_j^{\ -}; \qquad x_j^+ \ge 0, x_j^- \ge 0. \tag{5.20}$$

Наприклад, для $x_j = 10$ покладемо $x_j^+ = 10, x_j^- = 0$; для $x_j = -8$ покладемо $x_j^+ = 0,\ x_j^- = 8$. В обох випадках змінні $x_j^+, x_j^-$ невід'ємні;

в) якщо цільова функція $L(x)$ максимізується, то поклавши $L_1(x) = -L(x)$ прийдемо до задачі мінімізації цільової функції $L_1(x)$.

**Приклад 1.** Загальну задачу ЛП:

$$L_1(x) = 3x_1 - 1{,}5x_2 - 2x_3 \to \max$$

при обмеженнях:

$$2x_1 - x_2 + x_3 \le 3,$$
$$-x_1 + 2x_2 + x_3 \ge 2,$$
$$2x_1 + 3x_2 + 1,5x_3 = 15,$$
$$x_1 \ge 0,\ x_2 \ge 0,\ x_3 \in R$$

звести до стандартної форми.

Для зведення цієї задачі до стандартної задачі ЛП виконаємо наступне:

а) введемо нову цільову функцію $L_1(x) = -L(x) = = -3x_1 + 1,5x_2 + 2x_3$;

б) додамо до лівої частини першої нерівності додаткову(залишкову) змінну $x_4$ і отримаємо рівність;

в) віднімемо додаткову (надлишкову) змінну $x_5$ від лівої частини другої нерівності;

г) вільну змінну $x_3$ представимо у вигляді $x_3 = x_3^+ - x_3^-$, де $x_3^+$ і $x_3^-$ – невід'ємні змінні; потім цю заміну підставимо в цільову функцію $L_1(x)$ і обмеження рівності.

В результаті приходимо до наступної стандартної задачі ЛП:

$$L_1(x) = -3x_1 + 1,5x_2 + 2x_3^+ - 2x_3^- \to \min$$

при обмеженнях:

$$2x_1 - x_2 + 3x_3^+ - 3x_3^- + x_4 = 3,$$
$$-x_1 + 2x_2 + x_3^+ - x_3^- - x_5 = 2,$$
$$2x_1 + 3x_2 + 1,5x_3^+ - 1,5x_3^- = 15,$$
$$x_1 \ge 0,\ x_2 \ge 0,\ x_3^+ \ge 0;\ x_3^- \ge 0,\ x_4 \ge 0,\ x_5 \ge 0.$$



Відмітимо, що перехід від ЗЗЛП до СЗЛП призводить до збільшення розмірності простору змінних задачі, а отже і до збільшення складності задачі. У наведеному вище прикладі 1 ЗЗЛП сформульована в просторі $R^3$, а СЗЛП – в просторі $R^6$.

Надалі, не обмежуючи загальності, припускається, що для СЗЛП (5.14)-(5.16) виконується умова

$$rangA = m < n, \qquad (5.21)$$

а отже система рівнянь $Ax = b$ є сумісною і має безліч розв'язків. З умови (5.21) випливає, що максимальне число лінійно незалежних векторів $A_j,\ j = \overline{1,n}$ дорівнює числу *m*.

Якщо серед розв'язків системи $Ax = b$ немає невід'ємних, то допустима область $X$ є порожньою.

*Означення 1.* Ненульовий допустимий розв'язок $x$ стандартної задачі ЛП називається ***базисним***, якщо система векторів $A_j$, які відповідають додатнім компонентам $x_j$ цього розв'язку, є лінійно незалежною.

Нульовий допустимий розв'язок також вважають базисним.

Максимальне число додатних компонент базисного розв'язку СЗЛП дорівнює $m$.

*Означення 2.* Базисний розв'язок називається ***невиродженим***, якщо числа його додатних компонент дорівнює $m$.

*Означення 3.* Базисний розв'язок називається ***виродженим***, якщо числа його додатних компонент менше $m$.

Надалі будемо припускати (не обмежуючи загальності), що:

а) у випадку невиродженого базисного розв'язку перші його $m$ компонент є додатними:

$$x = (x_1, x_2, ..., x_m, 0, ..., 0),\ x_j > 0 \text{ при } j = \overline{1, m}, \tag{5.22}$$

причому вектори $A_1, ..., A_m$ – лінійно незалежні; ці вектори називають ***базисом***, що породжує базисний розв'язок $x$, а утворену ними матрицю $B = (A_1, ..., A_m)$ – ***базисною матрицею*** (зрозуміло, що у невиродженому випадку базис єдиний);

б) у випадку виродженого базисного розв'язку:

$$x = (x_1, x_2, ..., x_r, 0, ..., 0)^T,\quad x_j > 0 \text{ при } j = \overline{1, r},\ r < m, \tag{5.23}$$

причому вектори $A_1, ..., A_r$ – лінійно незалежні, у цьому випадку базисом, що породжує розв'язок (5.23), є будь-який набір $m$ лінійно незалежних векторів, який включає вектори $A_1, ..., A_r$. Ця система векторів формує базисну матрицю (зрозуміло, що у виродженому випадку базис і базисна матриця не єдині).

*Означення 4.* Змінні $x_j$, які відповідають базисним векторам називають ***базисними***, а решту змінних називають ***небазисними***.

Відмітимо, що у невиродженому випадку всі базисні змінні додатні, а небазисні – нульові, у виродженому випадку і серед базисних змінних є рівні нулю.

***Теорема 1.*** *Допустимий розв'язок $x$ СЗЛП є вершиною її допустимої множини $D$ тоді і тільки тоді, коли $x$ – базисний розв'язок.*

**Приклад 2.** Для СЗЛП:

$$L = 4x_1 + 5x_2 - 3x_3 - 2x_4 \to \min$$

при обмеженнях:

$$x_1 + 2{,}5x_2 - 3x_3 + x_4 = 4,$$
$$x_1 + 2x_2 + x_3 + 2x_4 = 4,$$
$$x_i \ge 0,\ i = \overline{1, 4}$$

знайти:

а) невироджений базисний розв'язок;

б) вироджений базисний розв'язок;

в) відповідні їм базиси, базисні матриці, базисні та небазисні змінні.

*Розв'язування.* а) Виберемо базисними змінними $x_2$ та $x_3$, а небазисними змінними $x_1$ та $x_4$. Покладемо $x_1 = 0$, $x_4 = 0$ та із системи обмежень отримаємо:

$$\begin{cases} 2,5x_2 - 3x_3 = 4, \\ 2x_2 + x_3 = 4 \mid 3 \end{cases} \Leftrightarrow \begin{cases} 8,5x_2 = 16, \\ x_3 = 4 - 2x_2, \end{cases} \Leftrightarrow \begin{cases} x_2 = 16/8,5 = 32/17, \\ x_3 = 4 - 64/17 = 4/17. \end{cases}$$

Отже, $(0;32/17;4/17;0)^T$ – невироджений базисний розв'язок; вектори $A_2 = \begin{pmatrix} 2,5 \\ 2 \end{pmatrix}$ та $A_3 = \begin{pmatrix} -3 \\ 1 \end{pmatrix}$ утворюють базис; базисна матриця $B = \begin{pmatrix} 2,5 & -3 \\ 2 & 1 \end{pmatrix}$;

б) Виберемо базисними змінними $x_1$ та $x_2$, а небазисними змінними $x_3$ та $x_4$. Покладемо $x_3 = 0$, $x_4 = 0$ та із системи обмежень отримаємо:

$$\begin{cases} x_1 + 2,5x_2 = 4, \\ x_1 + 2x_2 = 4, \end{cases} \Leftrightarrow \begin{cases} x_2 = 0, \\ x_1 = 4. \end{cases}$$

Отже, $(4;0;0;0)^T$ – вироджений базисний розв'язок; вектори $A_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ та $A_2 = \begin{pmatrix} 2,5 \\ 2 \end{pmatrix}$ утворюють базис; базисна матриця $B = \begin{pmatrix} 1 & 2,5 \\ 1 & 2 \end{pmatrix}$.



## 5.6. Канонічна задача лінійного програмування. Метод Жордана–Гауса перебору вершин допустимої області

### 1. Канонічна задача лінійного програмування

***Канонічною задачею ЛП (КЗЛП)*** називається стандартна задача ЛП:

$$L(x) = (c, x) \to \min$$

при обмеженнях:

$$Ax = b, \quad x \ge 0,$$

в якій матриця $A$ містить одиничну підматрицю розмірності $m \times m$ і вектор $b$ – невід'ємний. Надалі будемо вважати (не обмежуючи загальності), що перші $m$ стовпців матриці $A$ утворюють одиничну підматрицю. Отже, в координатній формі КЗЛП має вигляд:

$$L(x) = \sum_{j=1}^{n} c_j x_j \to \min \tag{5.24}$$

при обмеженнях:

$$x_i + \sum_{j=m+1}^{n} \alpha_{ij} x_j = b_i,\ i = \overline{1,m}; \qquad (5.25)$$

$$x_j \geq 0,\ \ j = \overline{1,n}; \qquad (5.26)$$

$$b_i \geq 0,\ \ i = \overline{1,m}. \qquad (5.27)$$

Дуже просто для КЗЛП (5.24)-(5.27) відшукується базисний розв’язок: $x = (b_1, ...b_m, 0, ..., 0)^T$;

базис: $A_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ ... \\ 0 \end{pmatrix},\ A_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ ... \\ 0 \end{pmatrix},\ \ldots,\ A_m = \begin{pmatrix} 0 \\ 0 \\ 0 \\ ... \\ 1 \end{pmatrix}$;

базисна матриця: $B = \begin{pmatrix} 1\ 0...0 \\ 0\ 1...0 \\ 0\ 0...1 \end{pmatrix}$;

базисні змінні: $x_1, ..., x_m$;

небазисні змінні: $x_{m+1}, ..., x_n$.

Запишемо КЗЛП (5.24)-(5.27) у матричній формі (у вигляді стандартної задачі ЛП):

$$L(x) = (c, x) \to \min$$

при обмеженнях:

$$\hat{A}x = b; \qquad (5.28)$$

$$x \geq 0; b \geq 0, \qquad (5.29)$$

де $\hat{A} = (I, A_{m+1}, ..., A_n)$, $I -$ одинична матриця розмірності $m \times m$, $A_j = (\alpha_{1j}, ..., \alpha_{mj})^T$, $j = \overline{m+1, n}$; $b = (b_1, ..., b_m)^T$.

Справедливе обернене твердження: кожну СЗЛП (5.14)-(5.16) при виконанні умови (5.21), можна звести до канонічної задачі ЛП, попередньо забезпечивши виконання умов: перші $m$ векторів-стовпців $A_1, ..., A_m$ матриці $A$ утворюють базис, а базисна матриця $B = (A_1, ..., A_m)$.

Помножимо зліва рівність (5.15) на $B^{-1}$:

$$B^{-1}Ax = B^{-1}b.$$

Отримаємо:

$\hat{A}x = \hat{b}$, де $\hat{A} = B^{-1}A$, $\hat{b} = B^{-1}b$, матриця $\hat{A}$ дорівнює $(I, A_{m+1}, ..., A_n)$.

Ми вказали теоретичний шлях переходу від СЗЛП до КЗЛП. Як діють на практиці буде описано в наступних підрозділах.

**2. Метод Жордана–Гауса перебору вершин допустимої області**

Основною операцією в симплекс-методі розв'язування КЗЛП є перехід від одного базисного розв'язку до іншого.

Перейдемо від базисного розв'язку

$$x = (b_1, ..., b_{r-1}, b_r, b_{r+1}, ..., b_m, 0, ..., 0)^T$$

КЗЛП (5.24)–(5.27) до нового базисного розв'язку
$\bar{x} = (\bar{b}_1, ..., \bar{b}_{r-1}, 0, \bar{b}_{r+1}, ..., \bar{b}_m, 0, ..., 0, \bar{b}_s, 0, ..., 0)^T$, якому відповідає нова КЗЛП, в якій з числа базисних виключена змінна $x_r$ $(r \in \overline{1,m})$, а змінна $x_s$ $(s \in \overline{m+1,n})$ введена до числа базисних. Цей перехід здійснюється ***методом виключення Жордана–Гауса*** (із забезпеченням канонічності нової задачі ЛП).

Спочатку із всіх рівнянь (окрім $r$-го) виключимо змінну $x_s$: для цього до кожного $i$-го рівняння ($i \neq r$) додаємо $r$-те рівняння, яке множиться на $\left(-\alpha_{is} / \alpha_{rs}\right)$, а $r$-те рівняння множиться на $1/\alpha_{rs}$. Від системи рівнянь (5.25) приходимо до наступної системи:

$$\begin{cases} x_i + \bar{\alpha}_{ir} x_r + \sum\limits_{j=m+1}^{n} \bar{\alpha}_{ij} x_j = \bar{b}_i, & i = 1, ..., r-1, r+1, ..., m, \\ \bar{\alpha}_{ir} x_r + \sum\limits_{j=m+1}^{n} \bar{\alpha}_{ij} x_j = \bar{b}_i, & i = r, \end{cases} \tag{5.30}$$

де

$$\bar{\alpha}_{ij} = \begin{cases} \alpha_{ij} - \alpha_{rj}\alpha_{is} / \alpha_{rs}, & i = 1, ..., r-1, r+1, ..., m, \\ \alpha_{ij} / \alpha_{rs}, & i = r; \quad j = \overline{1,n}; \end{cases} \tag{5.31}$$

$$\bar{b}_i = \begin{cases} b_i - b_r\alpha_{is} / \alpha_{rs}, & i = 1, ..., r-1, r+1, ..., m, \\ b_i / \alpha_{rs}, & i = r. \end{cases} \tag{5.32}$$

З (5.31) для $s$-го стовпця маємо:

$$\bar{\alpha}_{is} = \begin{cases} 1, & i = r, \\ 0, & i \neq r. \end{cases}$$

Для забезпечення канонічності нової задачі ЛП, необхідно також виконання умов $\bar{b}_i > 0$, $i = \overline{1,m}$. Для цього індекси $r$ та $s$ потрібно вибирати такими, щоб справджувались нерівності

$$b_r / \alpha_{rs} > 0; \ \ b_i - b_r\alpha_{is} / \alpha_{rs} > 0, \quad i \neq r. \tag{5.33}$$

Для цього індекс $r$ повинен вибиратись із умови

$$b_r / \alpha_{rs} = \min_{\alpha_{is} > 0} \left(b_i / \alpha\right). \tag{5.34}$$

Позначимо $\theta_s = b_r / \alpha_{rs}$.

Новий базисний розв'язок має вигляд:

$$\bar{x} = (b_1 - \theta_s \alpha_{1s}, ..., b_{r-1} - \theta_s \alpha_{r-1,s}, 0, b_{r+1} - \theta_s \alpha_{r+1,s}, ..., b_m - \theta_s \alpha_{ms}, 0, ..., \theta_s, 0, ..., 0)^T.$$

Враховуючи, що $b_r - \theta_s \alpha_{rs} = 0$, для значення цільової функції $L$ в точці $\bar{x}$ можна записати такі рівності:

$$L(\bar{x}) = \sum_{i=1}^{n} c_i \bar{x}_i = \sum_{i=1}^{m} c_i (b_i - \theta_s \alpha_{is}) + c_s \theta_s = \sum_{i=1}^{m} c_i b_i + c_s \theta_s - \theta_s \sum_{i=1}^{m} c_i a_{is} =$$
$$= \sum_{i=1}^{m} c_i b_i + \theta_s (c_s - \sum_{i=1}^{m} c_i \alpha_{is}) = L(x) + \theta_s \Delta_s, \qquad (5.35)$$

де ***симплекс-різниця*** $\Delta_s$ обчислюється за формулою

$$\Delta_s = c_s - \sum_{i=1}^{m} c_i \alpha_{is} = c_s - z_s. \qquad (5.36)$$

Легко переконатись, що симплекс-різниці базисних змінних дорівнюють нулю.

Із (5.35) випливає, що якщо $\theta_s > 0$ і $\Delta_s < 0$, то при переході від базису $x$ до базису $\bar{x}$ значення цільової функції $L$ зменшується на величину $|\theta_s \Delta_s|$; якщо $\theta_s = 0$, то значення цільової функції не змінюється (випадок виродженого базисного розв'язку).

***Теорема 1.* (*Критерій оптимальності базисного розв'язку ЗЛП*).** *Якщо для деякого базисного розв'язку $x^*$ виконуються нерівності $\Delta_j \geq 0$, $j = \overline{1,n}$, то $x^*$ – оптимальний розв'язок задачі лінійного програмування.*



***Теорема 2.* (*Критерій необмеженості цільової функції ЗЛП на допустимій області*).** *Якщо для деякого базисного розв'язку $\bar{x}$ існує хоча б одне $j$, таке що $\Delta_j < 0$, і вектор обмежень $A_j \leq 0$, то цільова функція ЗЛП необмежена на допустимій множині, тобто*

$$\min_{x \in X} L(x) = -\infty.$$

### 3. Табличний симплекс-метод

Наведемо алгоритм табличного симплекс-методу розв'язування канонічної задачі лінійного програмування:

$$L(x) = \sum_{j=m+1}^{n} c_j x_j \to \min$$

при обмеженнях:

$$x_i + \sum_{j=m+1}^{n} \alpha_{ij} x_j = a_i^0, \quad i = \overline{1,m};$$

$$x_j \geq 0, \ j = \overline{1,n};$$

$$a_i^0 \geq 0, \ i = \overline{1,m}.$$

***Алгоритм 3***

I. Складається і послідовно заповнюється симплекс-таблиця:

| $c$ | | $a$ | $c_1$ | $c_2$ | … | $c_s$ | … | $c_n$ | |
|---|---|---|---|---|---|---|---|---|---|
| | $x_б$ | $a^0$ | $a^1$ | $a^2$ | … | $a^s$ | … | $a^n$ | $\Theta$ |
| $c_1$ | $x_1$ | $a_1^0$ | $\alpha_{11}$ | $\alpha_{12}$ | … | $\alpha_{1s}$ | … | $\alpha_{1n}$ | |
| $c_2$ | $x_2$ | $a_2^0$ | $\alpha_{21}$ | $\alpha_{22}$ | … | $\alpha_{2s}$ | … | $\alpha_{2n}$ | |
| … | … | … | … | … | … | | … | … | |
| $c_r$ | $x_r$ | $a_r^0$ | $\alpha_{r1}$ | $\alpha_{r2}$ | … | $\alpha_{rs}$ | … | $\alpha_{rn}$ | |
| … | … | … | … | … | … | | … | … | |
| $c_m$ | $x_m$ | $a_m^0$ | $\alpha_{m1}$ | $\alpha_{m2}$ | … | $\alpha_{ms}$ | … | $\alpha_{mn}$ | |
| | $\Delta$ | $L(x)$ | $\Delta_1$ | $\Delta_2$ | … | $\Delta_s$ | … | $\Delta_n$ | |

II. Обчислення проводяться починаючи з останнього рядка: обчислюється значення цільової функції $L(x)=\sum_{i=1}^{m} c_i a_i^0$ та обчислюється симплекс-різниці (оцінки векторів або відносні оцінки змінних $x_j,\ j=\overline{1,n}$):

$$\Delta_j = c_j - \sum_{i=1}^{m} c_i \alpha_{ij}, \qquad j=\overline{1,n}$$

(тут мається на увазі, що базисними змінними є $m$ перших змінних $x_1, x_2, ..., x_m$, у випадку, коли базисними змінними є $x_{i_1}, x_{i_2}, ..., x_{i_m}$, то

$$\Delta_j = c_j - \sum_{k=1}^{m} c_{i_k} \alpha_{kj},\ \ j=\overline{1,n};\ \ L(x)=\sum_{k=1}^{m} c_{i_k} a_k^0,$$

причому, потрібно використовувати той факт, що базисним змінним відповідають нульові оцінки, тобто $\Delta_j = 0,\quad j \in \{i_1, i_2, ..., i_m\}$ ).

Якщо $\Delta_j \geq 0,\ j=\overline{1,n}$ то зупинити обчислення. В цьому випадку знайшли оптимальний розв'язок КЗЛП:

$$x^* = \left(a_1^0, a_2^0, ..., a_m^0, 0, ..., 0\right)^T \text{ і } L\left(x^*\right) = \sum_{i=1}^{m} c_i a_i^0;$$

інакше перейти на крок III .

III. Якщо існує принаймні один такий індекс $j$, що $\Delta_j < 0$ і $\alpha_{ij} \leq 0$ для всіх $i=\overline{1,m}$, то зупинити обчислення. В цьому випадку цільова функція необмежена знизу на допустимій множині: $\min_{x\in X} L(x) = -\infty$. Інакше перейти на крок IV.

IV. Знайти індекси $s$ та $r$ із умов:

$$\Delta_s = \min\left\{\Delta_j \mid \Delta_j < 0,\ \ j = \overline{1,n}\right\},$$

$$a_r^0 / \alpha_{rs} = \min\left\{a_i^0 / \alpha_{is} \mid \alpha_{is} > 0,\ i = \overline{1,m}\right\}.$$

V. Використовуючи формули (5.30)–(5.32) (в яких замість векторв $b$ використовувати $a^0$), перераховуємо матрицю $A$ та стовпець $a^0$; змінну $x_r$ виключаємо з числа базисних змінних, а змінну $x_s$ вводимо до базисних; коефіцієнт $c_r$ замінюємо на $c_s$; переходимо на крок I.

Відмітимо, що $r$-й рядок, $s$-й стовпець та елемент $\alpha_{rs}$ називають ***ведучими*** (***розв'язуючими***).

**Приклад 1.** Розв'язати табличним симплекс-методом ЗЗЛП:

$$\bar{L}(x) = 3x_1 + 4x_2 \to \max$$

при обмеженнях:

$$2x_1 + 2x_2 \le 5,$$

$$2x_1 - x_2 \le 2,$$

$$x_1 \ge 0, x_2 \ge 0.$$

*Розв'язування.*

Спочатку, вводячи додаткові змінні $x_3$ і $x_4$, переходимо до СЗЛП (вона є і канонічною) на мінімум:

$$L(x) = -3x_1 - 4x_2 \to \min$$

при обмеженнях:

$$2x_1 + 2x_2 + x_3 = 5,$$

$$2x_1 - x_2 + x_4 = 2,$$

$$x_j \ge 0, \quad j = \overline{1,4}.$$

*1-а ітерація:*

I. Складаємо і послідовно заповнюємо симплекс-таблицю 1.

Таблиця 1.

| c | | | -3 | -4 | 0 | 0 | |
|---|---|---|---|---|---|---|---|
| | $x_б$ | $a^0$ | $a^1$ | $a^2$ | $a^3$ | $a^4$ | $\Theta$ |
| 0 | $x_3$ | 5 | 2 | 2 | 1 | 0 | 2,5 |
| 0 | $x_4$ | 2 | 2 | -1 | 0 | 1 | |
| | $\Delta$ | 0 | -3 | -4 | 0 | 0 | |

II. Обчислюємо:

$$L(x) = 0 \cdot 5 + 0 \cdot 2 = 0;$$

$$\Delta_1 = c_1 - (c_3\alpha_{11} + c_4\alpha_{21}) = -3 - (0+0) = -3,$$

$$\Delta_2 = c_2 - (c_3\alpha_{12} + c_4\alpha_{22}) = -4 - (0+0) = -4.$$

III. Переходимо на крок IV.

IV. Знаходимо індекс $s = 2$ (бо оцінка $\Delta_2$ мінімальна).

Обчислюємо відношення $\theta_1 = a_1^0 / \alpha_{1s} = 2/2 = 1$, бо $\alpha_{12} > 0$, і покладемо $r = 1$.

V. Перераховуємо матрицю $A$ і вільний стовпець $a^0$:

$$\overline{\alpha}_{11} = \alpha_{11}/\alpha_{12} = 2/2 = 1; \quad \overline{\alpha}_{12} = 1;$$

$$\overline{\alpha}_{13} = \alpha_{13}/\alpha_{12} = 1/2 = 0,5; \quad \overline{\alpha}_{14} = \alpha_{14}/2 = 0;$$

$$\overline{\alpha}_{21} = \alpha_{21} - \frac{\alpha_{11}\alpha_{22}}{\alpha_{12}} = 2 - 2(-1)/2 = 3; \quad \overline{\alpha}_{22} = 0;$$

$$\overline{\alpha}_{23} = \alpha_{23} - \alpha_{13}\alpha_{22}/\alpha_{12} = 0 - 1(-1)/2 = 0,5;$$

$$\overline{\alpha}_{24} = \alpha_{24} - \alpha_{14}\alpha_{22}/\alpha_{12} = 1 - 0 = 1;$$

$$\overline{a}_1^0 = a_1^0/\alpha_{12} = 5/2 = 2,5;$$

$$\overline{a}_2^0 = a_2^0 - a_1^0\alpha_{22}/a_{12} = 2 - 5(-1)/2 = 4,5.$$

Змінну $x_2$ введемо до базисних, а змінну $x_3$ виводимо; переходимо на крок I.

*2-а ітерація*:

I. Послідовно заповнимо симплекс-таблицю 2.

Таблиця 2.

| $c$ | | | -3 | -4 | 0 | 0 | |
|---|---|---|---|---|---|---|---|
| | $x_б$ | $a^0$ | $a^1$ | $a^2$ | $a^3$ | $a^4$ | $\Theta$ |
| -4 | $x_2$ | 2,5 | 1 | 1 | 0,5 | 0 | |
| 0 | $x_4$ | 4,5 | 3 | 0 | 0,5 | 1 | |
| | $\Delta$ | -10 | 1 | 0 | 2 | 0 | |

II. Обчислюємо:

$$L(x) = (-4) \cdot 2,5 + 0 \cdot 4,5 = -10,$$

$$\Delta_1 = -3 - ((-4)1 + 0) = -3 + 4 = 1,$$

$$\Delta_3 = 0 - ((-4) \cdot 0,5 + 0) = 2.$$

Оскільки $\Delta_j \geq 0,\ j = \overline{1,4}$, то зупинимо обчислення. Таким чином, знайдений оптимальний розв'язок КЗЛП: $x^* = (0; 2,5; 0; 4,5)^T$ і оптимальне значення цільової функції $L(x^*) = -10$, а значить відомий оптимальний розв'язок початкової задачі $x^* = (0; 2,5)^T$ і оптимальне значення цільової функції $\overline{L}(x^*) = -(-10) = 10$.

## 5.7. Симплекс-метод та його варіанти

### 1. Розв'язування СЗЛП з $n$ змінними графічним способом в просторі $R^2$

Спочатку розглянемо специфічний випадок розв'язування стандартної задачі лінійного програмування:

$$\sum_{i=1}^{n} c_i x_i \to \min$$

при обмеженнях:

$$\sum_{j=1}^{n} \alpha_{ij} x_j = a_i^0, \quad i = \overline{1,m}; \; x_j \geq 0, \quad j = \overline{1,n}$$

за умови, що ранг матриці $A$ дорівнює $n - 2 = m$.

Таку задачу можна розв'язати графічним методом, виразивши, наприклад, всі базисні змінні $x_1, x_2, ..., x_{n-2}$ через дві небазисні змінні $x_{n-1}, x_n$:

$$x_1 = \varphi_1(x_{n-1}, x_n) \geq 0,$$
$$x_2 = \varphi_2(x_{n-1}, x_n) \geq 0,$$
$$........$$
$$x_{n-2} = \varphi_{n-2}(x_{n-1}, x_n) \geq 0.$$

Підставляючи в цільову функцію замість $x_i, i = \overline{1, n-2}$ функції $\varphi_i$, приходимо до наступної задачі в просторі $R^2$:

$$\sum_{i=1}^{n-2} c_i \varphi_i(x_{n-1}, x_n) + c_{n-1} x_{n-1} + c_n x_n \to \min$$

при обмеженнях:

$$\varphi_1(x_{n-1}, x_n) \geq 0;$$
$$\varphi_2(x_{n-1}, x_n) \geq 0;$$
$$.....$$
$$\varphi_{n-2}(x_{n-1}, x_n) \geq 0;$$
$$x_{n-1} \geq 0, \quad x_n \geq 0,$$

яка розв'язується графічно.

**Приклад 1.** Розв'язати стандартну задачу ЛП:

$$L(x) = 5x_5 + 2x_4 \to \max$$

при обмеженнях:

$$\begin{cases} 2x_1 - 4x_2 + 2x_3 + x_4 - 7x_5 = 18; \\ x_1 - x_2 - 2x_3 - 5x_4 + 6x_5 = -19; \\ 2x_2 - 4x_3 - 7x_4 + 13x_5 = -4; \\ x_i \geq 0,\ i = \overline{1,5}. \end{cases}$$

Отже, $n = 5$, $m = 3$. Спочатку знайдемо ранг матриці обмежень $A$. Для цього обчислимо визначник складений із трьох перших стовпців матриці $A$:

$$\det A' = \begin{vmatrix} 2 & -4 & 2 \\ 1 & -1 & -2 \\ 0 & 2 & -4 \end{vmatrix} = 8 + 4 + 0 - 0 + 8 - 16 = 4 \neq 0.$$

Звідси робимо висновок, що ранг матриці $A$ дорівнює 3 (тобто $rang\, A = 5 - 2 = 3 = m$).

Визначимо змінні $x_1, x_2, x_3$ через $x_4, x_5$, виконуючи наступні еквівалентні перетворення систем:

$$\begin{cases} x_1 - x_2 - 2x_3 - 5x_4 + 6x_5 = -19; \\ 2x_2 - 4x_3 - 7x_4 + 13x_5 = -4; \\ 2x_1 - 4x_2 + 2x_3 + x_4 - 7x_5 = -18; \end{cases} \Leftrightarrow \begin{cases} x_1 - x_2 - 2x_3 - 5x_4 + 6x_5 = -19; \\ x_2 - 2x_3 - \frac{7}{2}x_4 + \frac{13}{2}x_5 = -2; \\ -2x_2 + 6x_3 + 11x_4 - 19x_5 = 20; \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} x_1 - x_2 - 2x_3 - 5x_4 + 6x_5 = -19; \\ x_2 - 2x_3 - \frac{7}{2}x_4 + \frac{13}{2}x_5 = -2; \\ 2x_3 + 4x_4 - 6x_5 = 16; \end{cases} \Leftrightarrow \begin{cases} x_1 - x_2 - 2x_3 - 5x_4 + 6x_5 = -19; \\ x_2 - 2x_3 - \frac{7}{2}x_4 + \frac{13}{2}x_5 = -2; \\ x_3 + 2x_4 - 3x_5 = 8; \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} x_1 - \quad 4x_3 - \frac{17}{2}x_4 + \frac{25}{2}x_5 = -21; \\ x_2 + \qquad \frac{1}{2}x_4 + \frac{1}{2}x_5 = 14; \\ x_3 + 2x_4 - 3x_5 = 8; \end{cases} \Leftrightarrow$$

$$\Leftrightarrow\begin{cases} x_1 - \qquad \frac{1}{2}x_4+\frac{1}{2}x_5=11; \\ \quad x_2 + \qquad \frac{1}{2}x_4+\frac{1}{2}x_5=14; \\ \qquad x_3+2x_4-3x_5=8; \end{cases} \Leftrightarrow \begin{cases} x_1=\frac{1}{2}x_4-\frac{1}{2}x_5+11; \\ x_2=-\frac{1}{2}x_4-\frac{1}{2}x_5+14; \\ x_3=-2x_4+3x_5+8. \end{cases}$$

Враховуючи умову невід'ємності змінних $x_1,\ x_2,\ x_3$ отримаємо наступну систему нерівностей у просторі $R^2$:

$$\begin{cases} (1/2)x_4-(1/2)x_5+11\ge 0; & (1) \\ (-1/2)x_4-(1/2)x_5+14\ge 0; & (2) \\ -2x_4+3x_5+8\ge 0; & (3) \end{cases}$$

$$x_4\ge 0,\ \ x_5\ge 0.$$

Зобразимо допустиму область в системі координат $(x_4Ox_5)$ – п'ятикутник $ABCDO$.

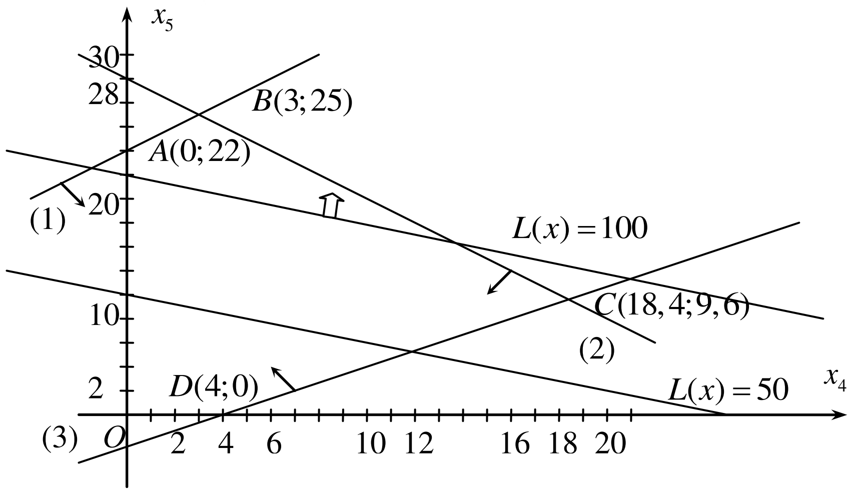
Побудуємо дві лінії рівня для функції цілі

$$5x_5+2x_4=50 \text{ і } 5x_5+2x_4=100.$$

Встановлюємо напрямок зростання цільової функції.

Робимо висновок, що максимальне значення цільової функції досягається в точці $B(3;25)$:

$$x_4^*=3,\ \ x_5^*=25.$$




Тепер можемо знайти оптимальні значення змінних $x_1, x_2, x_3$:

$$\begin{cases} x_1^* = (1/2)\cdot 3 - (1/2)\cdot 25 + 11 = 0; \\ x_2^* = (-1/2)\cdot 3 - (1/2)\cdot 25 + 14 = 0; \\ x_3^* = -2\cdot 3 + 3\cdot 25 + 8 = 77 \end{cases}$$

і оптимальне значення цільової функції

$$L(x^*) = 5\cdot 25 + 2\cdot 3 = 131.$$

Тепер переходимо до загального випадку розв'язування СЗЛП.

### 2. Симплекс-метод розв'язування невиродженої стандартної задачі лінійного програмування

З а д а ч а  2. Знайти $\arg\max\limits_{x\in X}(c,x)$ для заданого вектора $c=(c_1,...,c_n)^T$ та заданої множини

$$X \triangleq \left\{ x \middle| Ax = a^0,\quad x \ge 0,\ x \in R^n \right\},$$

де $A$ – матриця розмірності $m\times n$:

$$A = \begin{pmatrix} \alpha_{11}\alpha_{12}...\alpha_{1n} \\ \alpha_{21}\alpha_{22}...\alpha_{2n} \\ .................. \\ \alpha_{m1}\alpha_{m2}...\alpha_{mn} \end{pmatrix},\quad a^0 = \left(a_1^0, a_2^0, ..., a_m^0\right)^T.$$

Стовпці матриці $A$ позначаються через $a^1$, $a^2$, ..., $a^n$.

*Припущення 2*. (***i***) – ранг матриці $A$ дорівнює $m$; (***ii***) – $n > m$; (***iii***) – множина $X$ непорожня.

*Означення 1*. Допустимий розв'язок $x$ (тобто вектор $x\in X$) називається ***опорним*** (***базисним***) розв'язком задачі 2, якщо система векторів $a^j$, що відповідає його додатнім компонентам $(x_j > 0)$, лінійно-незалежна.

*Означення 2*. ***Базисом опорного розв'язку*** $x$ називають систему $m$ лінійно-незалежних векторів, яка включає всі вектори $a^j$, що відповідають додатнім складовим опорного розв'язку $x$.

*Означення 3*. Опорний розв'язок $x$ називається ***невиродженим***, якщо число його додатних компонент дорівнює $m$ (якщо воно менше $m$, то опорний розв'язок називається ***виродженим***).

*Означення 4*. Стандартна задача лінійного програмування називається ***невиродженою***, якщо всі її опорні розв'язки невироджені.

*Означення 5*. Компоненти опорного розв'язку, що відповідають векторам його базису, називаються ***базисними***, а інші – ***небазисними***.

В подальшому базисні вектори будемо позначати через

$$a^{i_1}, a^{i_2}, ..., a^{i_m};\ \ i_k \in [1:n],\quad k=\overline{1,m}.$$

Базисні вектори $a^{i_k}$ характеризуються номером $i_k$ та позицією $k$, яку він займає в базисі. Матриця $B$, що складена з векторів $a^{i_1}, a^{i_2}, ..., a^{i_m}$, називається ***базисною***

$$B = (a^{i_1}, ..., a^{i_m}).$$

В симплекс-методі розв'язування задачі 2 починається з відомого опорного розв'язку та його базису. На кожній ітерації алгоритму проводиться перевірка опорного розв'язку на оптимальність. Якщо опорний розв'язок не оптимальний, то вказується спосіб, що дозволяє по даному опорному побудувати другий опорний розв'язок, більш близький до оптимального. Через скінченне число ітерацій або знаходиться розв'язок задачі 2, або встановлюється необмеженість цільової функції задачі 2 на множині $X$.

***Алгоритм 2***

Початок. I. Знайти базисний (опорний) розв'язок $x^1 = (x_1^1, ... x_n^1)^T$ задачі 2, базис $a^{i_1}, a^{i_2}, ..., a^{i_m}$ цього опорного розв'язку та обчислити матрицю $B^{-1}$, обернену до початкової базисної матриці $B$, що складається із стовпців $a^{i_1}, a^{i_2}, ..., a^{i_m}$.

(В основному на практиці знаходять опорний розв'язок, якому відповідає одиничний базис $a^{i_1} = (1, 0, ..., 0)^T, ...,$ $a^{i_m} = (0, 0, ..., 1)^T$. Для знаходження початкового базису і опорного розв'язку використовують алгоритм 3).

II. Розкласти по початковому базису $a^{i_1}, a^{i_2}, ..., a^{i_m}$ вектори $a^0, a^1, ..., a^n$ (тобто, знайти числа $z_{kj}$, $k = \overline{1, m}$; $j = \overline{0, n}$, такі, що:

$$a^0 = z_{10}a^{i_1} + z_{20}a^{i_2} + ... + z_{m0}a^{i_m};$$

$$a^1 = z_{11}a^{i_1} + z_{21}a^{i_2} + ... + z_{m1}a^{i_m};$$

$$.....................$$

$$a^n = z_{1n}a^{i_1} + z_{2n}a^{i_2} + ... + z_{mn}a^{i_m};$$

за формулами:

$$z_{10} = x_{i_1}^1, \ z_{20} = x_{i_2}^1, \ ..., \ z_{mo} = x_{i_m}^1;$$

$$(z_{1j}, z_{2j}, ..., z_{mj})^T = B^{-1}(\alpha_{1j}, \alpha_{2j}, ... \alpha_{mj})^T, \quad j = \overline{1, n}.$$

Відмітимо, що завчасно відомий розклад по початковому базису базисних векторів $a^{i_k}$, $k = \overline{1, m}$:

$$z_{li_k} = 0 \text{ при } k \neq l; \ \ z_{li_k} = 1 \text{ при } k = l,$$

$$k = \overline{1, m}; \quad l = \overline{1, m}.$$

Основний цикл. III. Для кожного $j \in [1:n]$ обчислити оцінку

$$\Delta_j = \sum_{k=1}^{m} c_{i_k} z_{kj} - c_j. \tag{5.37}$$

IV. Якщо всі $\Delta_j \geq 0$ $(j = \overline{1,n})$, то покласти $x^* = x^1$ та припинити обчислення (в цьому випадку опорний розв'язок оптимальний і оптимум цільової функції дорівнює $\sum_{k=1}^{m} c_{i_k} z_{k0}$);

якщо при деякому $j \in [1:n]$ виконується $\Delta_j < 0$, то перейти на крок V.

V. Покладемо $j = 1$.

VI. Якщо $\Delta_j < 0$, то перейти на крок VII; інакше перейти на крок VIII.

VII. Якщо при всіх $k \in [1:m]$ виконується нерівність $z_{kj} \leq 0$, то припинити обчислення (в цьому випадку цільова функція $(c, x)$ не обмежена (зверху) на допустимій множині $X$); інакше перейти на крок VIII.

VIII. Якщо $j < n$, то покласти $j = j + 1$ і перейти на крок VI; інакше перейти на крок IX.

IX. Вибрати згідно завчасно обумовленого правила такий індекс $s \in [1:n]$, щоб $\Delta_s < 0$. Звичайно на практиці за $\Delta_s$ приймають або найменшу із від'ємних оцінок $\Delta_j$, або ж від'ємну оцінку $\Delta_j$ з найменшим індексом $j$.

X. Обчислити відношення $z_{k0} / z_{ks}$ для тих $k$, для яких $z_{ks} > 0$, та позначити мінімальне з цих відношень через $\theta_0$.

Знайти такий індекс $r \in [1:m]$, що $z_{rs} > 0$ та $z_{r0} / z_{rs} = \theta_0$.

XI. При кожному $j \in [1:n]$ обчислити

$$\theta_j = z_{rj} / z_{rs}.$$

XII. Покласти $\bar{z}_{kj} = z_{kj}$, $k = \overline{1,m}$; $j = \overline{1,n}$.

XIII. Перейти до нового базису, який отримуємо заміною вектора $a^{i_r}$ в попередньому базисі вектором $a^s$ (тобто покласти $i_r = s$).

XIV. Обчислити координати всіх векторів $a^0, a^1, ..., a^n$ в новому базисі за основними формулами:

при $k \neq r$

$$z_{kj} = \bar{z}_{kj} - \theta_j \bar{z}_{ks}, \ j = \overline{0,n}; \ k = 1, ..., r-1, r+1, ... m;$$

при $k = r$

$$z_{kj} = \theta_j, \ j = \overline{0,n}.$$

XV. Перейти до нового опорного розв'язку $x^1 = (x_1^1, ..., x_n^1)^T$:

$$x_{i_k}^1 = z_{k0}, \quad k = \overline{1,m},$$

інші координати вектора $x^1$ дорівнюють нулю.

XVI. Перейти на крок III.

***Теорема 2.*** *Якщо виконані припущення 2 і задача 2 невироджена, то через скінченне число ітерацій процес розв'язування задачі 2 алгоритмом 2* *закінчиться або на кроці IV( в цьому випадку знаходиться оптимальний розв'язок задачі 2), або на кроці VII (в цьому випадку встановлюється, що оптимального розв'язку задачі 2 немає).*

*Зауваження 2.* Оцінки $\Delta_j$, $j=\overline{1,n}$, доцільно обчислювати за формулою (5.37) тільки на першій ітерації. На наступних ітераціях нову оцінку $\Delta_j$ потрібно обчислювати через стару оцінку $\bar{\Delta}_j$, що отримана на попередній ітерації, за формулою

$$\Delta_j = \bar{\Delta}_j - \theta_j \bar{\Delta}_s, \qquad j=\overline{1,n}.$$

**3. Методи пошуку початкового базису**

А. Перший метод. Нехай початкова задача лінійного програмування має вигляд.

Задача 3. Знайти $\arg\max\limits_{x\in X}(c,x)$ для заданого вектора $c=(c_1,c_2,...,c_n)$ та множини $X$, що задається співвідношеннями:

$$\begin{aligned}
&\alpha_{11}x_1 + ... + \alpha_{1s}x_s + x_{s+1} = a_1^0;\\
&\alpha_{21}x_1 + ... + \alpha_{2s}x_s + x_{s+2} = a_2^0;\\
&\dots\dots\dots\dots\dots\dots\dots\\
&\alpha_{k1}x_1 + ... + \alpha_{ks}x_s + x_n = a_k^0;\\
&\alpha_{k+1,1}x_1 + ... + \alpha_{k+1,s}x_s = a_{k+1}^0;\\
&\dots\dots\dots\dots\dots\dots\dots\\
&\alpha_{m1}x_1 + ... + \alpha_{ms}x_s = a_m^0;\\
&x_j \ge 0, \qquad j=\overline{1,n}.
\end{aligned}$$

В обмеженнях задачі 3 виділені змінні $x_{s+1},...,x_n$ із різними одиничними векторами $a^{s+1},...,\ a^n$. Якщо таких змінних в задачі лінійного програмування немає, то в задачі 3 потрібно покласти $s=n$.

*Припущення 3.* Обмеження задачі 3 такі, що $a_i^0 \ge 0$, $i=\overline{1,m}$.

Для відшукування початкового базису та опорного розв'язку задачі 3 в обмеження цієї задачі штучно вводять $m-k$ нових змінних $y_1, y_2,..., y_{m-k}$ з таким розрахунком, щоб нова система рівнянь-обмежень мала повний одиничний базис. Потім з допомогою алгоритму 2 розв'язують задачу лінійного програмування з новими обмеженнями і з спеціально побудованою цільовою функцією. В результаті розв'язування „нової" задачі приходять або до опорного розв'язку початкової задачі 3, або безпосередньо до оптимального розв'язку задачі 3. Крім того, якщо початкова задача 2 не розв'язується, то це виявляється в результаті розв'язування «нової» задачі.

«Нова» задача є задачею лінійного програмування у просторі $R^{n+m-k}$ (або $R^{s+m}$).

Задача 3′. Знайти $\arg \max\limits_{(x_1,\ldots,x_n,y_1,\ldots,y_{m-k})\in Y} (-y_1 - y_2 - \ldots - y_{m-k})$ для множини $Y$, що задається співвідношеннями:

$$\begin{aligned}
&\alpha_{11}x_1 + \ldots + \alpha_{1s}x_s + x_{s+1} &&= a_1^0;\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&\alpha_{k1}x_1 + \ldots + \alpha_{ks}x_s + x_n &&= a_k^0;\\
&\alpha_{k+1,1}x_1 + \ldots + \alpha_{k+1,s}x_s + y_1 &&= a_{k+1}^0;\\
&\alpha_{k+2,1}x_1 + \ldots + \alpha_{k+2,s}x_s + y_2 &&= a_{k+2}^0;\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&\alpha_{m1}x_1 + \ldots + \alpha_{ms}x_s + \ldots + y_{m-k} &&= a_m^0;\\
&x_j \ge 0, \quad j = \overline{1,n}; \quad y_i \ge 0, \quad i = \overline{1,m-k}.
\end{aligned}$$

Задача 3′ має опорний розв'язок $\left(0,\ldots,0,a_1^0,a_2^0,\ldots,a_m^0\right)^T$ з одиничним базисом $a^{s+1},a^{s+2},\ldots,a^n,a^{n+1},\ldots,a^{n+m-k}$, тому для неї використовується алгоритм 1 – симплекс-метод розв'язування стандартної задачі лінійного програмування. При цьому задача 3′ явно має оптимальний розв'язок, оскільки її цільова функція обмежена зверху числом нуль. Змінні $y_1,\ldots,y_{m-k}$ називають ***штучними змінними***, а вектори $a^{n+1},\ldots,a^{n+m-k}$ – ***штучними векторами.***

***Алгоритм 3***

I. Використовуючи алгоритм 2, обчислити вектор $\left(x_1^0,x_2^0,\ldots,x_n^0,y_1^0,\ldots,y_{m-k}^0\right)^T$ – оптимальний розв'язок задачі 3′ і оптимальний базис задачі 3′ – $a^{i_1},a^{i_2},\ldots,a^{i_m}$.

II. Якщо хоча б при одному $\overline{j}\in[1:m-k]$ виконується нерівність $y_{\overline{j}} > 0$, то припинити обчислення (в цьому випадку початкова задача 3 недопустима, тобто немає допустимих розв'язків); якщо $y_1^0 = y_2^0 = \ldots = y_{m-k}^0 = 0$, то перейти на крок III.

III. Якщо оптимальний базис задачі 3′ не містить штучних векторів $a^j$, $j = n+1,\ldots,n+m-k$, то припинити обчислення, оскільки оптимальний базис задачі 3′ є початковим базисом задачі 3, а відповідне йому значення основних змінних $x_1^0,x_2^0,\ldots,x_n^0$ складають початковий опорний розв'язок задачі 3; якщо оптимальний базис задачі 3′ має хоча б один штучний вектор $a^j$, $j\in[n+1:n+m-k]$, то перейти на крок IV.

IV. Вектор $(x_1^0, x_2^0, ..., x_n^0)^T$ є опорним розв'язком задачі 3. Для відшукування початкового базису задачі 3 перейти на крок V.

V. Якщо в розкладі векторів $a^1, a^2, ..., a^n$ по базису $a^{i_1}, a^{i_2}, ..., a^{i_m}$ всі коефіцієнти $z_{kj}$ при штучних векторах дорівнюють нулю, то, видаливши із системи векторів $a^{i_1}, a^{i_2}, ..., a^{i_m}$ штучні вектори, отримаємо початковий базис задачі 3, який відповідає опорному розв'язку $(x_1^0, ... x_n^0)^T$; інакше перейти на крок VI.

VI. Знайти вектор $a^s, \left(s \in [1:n]\right)$, в розкладі якого по базису $a^{i_1}, a^{i_2}, ..., a^{i_m}$ є число $z_{rs}$ при штучному векторі $a^{i_r}$, відмінне від нуля. Ввести в базис вектор $a^s$ замість штучного вектора $a^{i_r}$ (тобто покласти $i_r = s$).

VII. Якщо в новому базисі немає штучних векторів, то припинити обчислення (в цьому випадку знаходиться опорний розв'язок задачі 3 і його базис); інакше перейти на крок VIII.

VIII. Розкласти по новому базису вектори $a^1, a^2, ..., a^n$ (як на кроці XIV алгоритму 1).

IX. Якщо в розкладі векторів $a^1, a^2, ..., a^n$ по базису $a^{i_1}, a^{i_2}, ..., a^{i_m}$ всі коефіцієнти $z_{kj}$ при штучних векторах дорівнюють нулю, то перейти на крок V; інакше перейти на крок VI.

***Теорема 3.*** *При виконанні припущень 3 алгоритм 3 за скінченне число ітерацій або знаходить опорний розв'язок задачі 3 і відповідний йому базис, що складається з $l$ векторів ( $l$ – ранг матриці A задачі 3), або встановлює, що задача 3 не має допустимих розв'язків.*

Б. Д р у г и й  м е т о д. Нижче приводиться алгоритм відшукування опорного розв'язку задачі 2 і його базису, який зводиться до розв'язування *M*-задачі – спеціально побудованої задачі лінійного програмування в просторі $R^{n+m}$, яка містить параметр $M$ (тут $M$ – достатньо велике число). В процесі розв'язування $M$-задачі або встановлюється необмеженість цільової функції задачі 2, або знаходиться оптимальний розв'язок задачі 2, або встановлюється, що задача 2 не має допустимих розв'язків.

*Припущення 3'*. Обмеження задачі 2 такі, що $a_i^0 \geq 0$, $i = \overline{1,m}$.

***Алгоритм 3'***

I. Вибрати достатньо велике число $M > 0$.

II. Використовуючи алгоритм 2, обчислити вектор $(x_1^0, ..., x_n^0, y_1^0 ..., y_m^0)^T$ – оптимальний розв'язок наступної задачі лінійного програмування.

***M - задача.*** Знайти

$$\arg \max_{(x_1, ... x_n, y_1, ... y_m)} \left[(c, x) - My_1 - My_2 - ... - My_m\right]$$

при обмеженнях:

$$\begin{aligned}
&\alpha_{11}x_1+\alpha_{12}x_2+\ldots+\alpha_{1n}x_n+y_1 &&=a_1^0;\\
&\alpha_{21}x_1+\alpha_{22}x_2+\ldots+\alpha_{2n}x_n+y_2 &&=a_2^0;\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&\alpha_{m1}x_1+\alpha_{m2}x_2+\ldots+\alpha_{mn}x_n+y_m &&=a_m^0;\\
&x_j\ge 0,\quad j=\overline{1,n};\quad y_i\ge 0,\quad i=\overline{1,m}.
\end{aligned}$$

$M$-задача має початковий опорний розв'язок $\left(0,...,0,a_1^0,...,a_m^0\right)$ з одиничним базисом $a^{n+1},...,a^{n+m}$, тому для неї можна застосувати алгоритм 2. Якщо при розв'язуванні $M$-задачі алгоритмом 2 виявиться, що її цільова функція необмежена зверху на допустимій множині, то цільова функція задачі 2 також необмежена зверху на допустимій множині $X$.

III. Якщо при деякому $i_0\in[1:m]$ виконується $y_{i_0}^0>0$, то припинити обчислення ( в цьому випадку задача 0 не має допустимих розв'язків); якщо при всіх $i\in[1:m]$ виконується $y_i^0=0$, то покласти $x^*=\left(x_1^0,...,x_n^0\right)^T$ і припинити обчислення (тобто в цьому випадку вектор $(x_1^0,x_2^0,...,x_n^0)^T$ є оптимальним розв'язком задачі 2).

***Теорема 3'.*** *Нехай виконується припущення 3'. Якщо для всіх достатньо великих значень $M$ $M$-задача має оптимальний розв'язок, то алгоритм 3' або встановлює недопустимість задачі 2, або знаходить оптимальний її розв'язок, а якщо при тих же умовах $M$-задача не розв'язується, то не розв'язується і задача 2.*

**Приклад 2.** Розв'язати симплекс-методом загальну задачу ЛП:

$$L(x)=-2x_1-x_2-7x_3\to\min$$

при обмеженнях:

$$\begin{gathered}
x_1+2x_2+3x_3\le 4,\\
-x_1-4x_2+10x_3\le 7,\\
x_j\ge 0,\ \ j=\overline{1,3}.
\end{gathered}$$

*Розв'язування*. Вводячи залишкові змінні $x_4$ і $x_5$ зведемо цю задачу до стандартної задачі ЛП:

$$\overline{L}(x)=2x_1+x_2+7x_3\to\max$$

при обмеженнях:

$$\begin{gathered}
x_1+2x_2+3x_3+x_4=4,\\
-x_1-4x_2+10x_3+x_5=7,\\
x_j\ge 0,\quad j=\overline{1,5}.
\end{gathered}$$

В СЗЛП маємо $n=5,\ m=2;\quad c=(2;1;7;0;0)^T$,

$$a^0=\begin{pmatrix}4\\7\end{pmatrix}, a^1=\begin{pmatrix}1\\-1\end{pmatrix}, a^2=\begin{pmatrix}2\\-4\end{pmatrix},\ a^3=\begin{pmatrix}3\\10\end{pmatrix}, a^4=\begin{pmatrix}1\\0\end{pmatrix}, a^5=\begin{pmatrix}0\\1\end{pmatrix}.$$

**Алгоритм 2**

I. Опорний (базисний) розв'язок $x^1=(0;0;0;4;7)^T$; базис $a^{i_1}=\begin{pmatrix}1\\0\end{pmatrix}$, $a^{i_2}=\begin{pmatrix}0\\1\end{pmatrix}$ (тобто, $i_1=4,\ i_2=5$). Базисна матриця $B=\begin{pmatrix}1&0\\0&1\end{pmatrix}$, $B^{-1}=\begin{pmatrix}1&0\\0&1\end{pmatrix}$.

II. Знаходимо коефіцієнти розкладу по початковому базису векторів $a^j$, $j=\overline{0,n}$:

$j=0:(z_{10};z_{20})=(4;7)$; $\quad j=1:(z_{11};z_{21})=(1;-1)$;

$j=2:(z_{12};z_{22})=(2;-4)$; $\quad j=3:(z_{13};z_{23})=(3;10)$;

$j=4:(z_{14};z_{24})=(1;0)$; $\quad j=5:(z_{15};z_{25})=(0;1)$.

*1-а ітерація:*

III. Обчислюємо оцінки $\Delta_j$, $j=\overline{1,5}$:

$j=1:\Delta_1=c_{i_1}z_{11}+c_{i_2}z_{21}-c_1=0+0-2=-2;$

$j=2:\Delta_2=c_{i_1}z_{12}+c_{i_2}z_{22}-c_2=0-1=-1;$

$j=3:\Delta_3=c_{i_1}z_{13}+c_{i_2}z_{23}-c_3=0-7=-7;$

$j=4:\Delta_4=c_{i_1}z_{14}+c_{i_2}z_{24}-c_4=0-0=0;$

$j=5:\Delta_5=c_{i_1}z_{15}+c_{i_2}z_{25}-c_5=0-0=0.$

IV. Оскільки $\Delta_1<0,\Delta_2<0,\Delta_3<0,$ тому переходимо на крок V.

V-VIII. Оскільки при всіх $j=\overline{1,5}$ існує $z_{1j}>0$, то переходимо на крок IX.

IX. Вибираємо $s=3$, оскільки $\Delta_3=-7$ .

X. Обчислюємо відношення $z_{k0}/z_{ks}$, $k=\overline{1,2}$, $\theta_0$ і $r$:

$k=1:z_{10}/z_{13}=4/3;\quad k=2:z_{20}/z_{23}=7/10=\theta_0;\ r=2.$

XI. Обчислюємо $\theta_j=z_{2j}/10$, $j=\overline{1,5}$:

$j=1:\theta_1=z_{21}/10=-1/10=-0,1;$

$j=2:\theta_2=z_{22}/10=-4/10=-0,4;$

$j=3:\theta_3=z_{23}/10=10/10=1;$

$j=4:\theta_4=z_{24}/10=0/10=0;$

$j=5:\theta_5=z_{25}/10=1/10=0,1.$

XII. Запам'ятовуємо коефіцієнти розкладу і оцінки:

$$\left(\overline{z}_{10};\overline{z}_{20}\right)=\left(4;7\right);\ \left(\overline{z}_{11};\overline{z}_{21}\right)=\left(1;-1\right);\ \left(\overline{z}_{12};\overline{z}_{22}\right)=\left(2;-4\right);$$

$$\left(\overline{z}_{13};\overline{z}_{23}\right)=\left(3;10\right);\ \left(\overline{z}_{14};\overline{z}_{24}\right)=\left(1;0\right);\ \left(\overline{z}_{15};\overline{z}_{25}\right)=\left(0;1\right);$$

$$\overline{\Delta}_1=-2,\ \overline{\Delta}_2=-1,\ \overline{\Delta}_3=-7,\ \overline{\Delta}_4=0,\ \overline{\Delta}_5=0.$$

XIII. Переходимо до нового базису:

$$a^{i_2}=a^3=\begin{pmatrix}3\\10\end{pmatrix};\ i_1=4,\ i_2=3,\ B=\begin{pmatrix}1&3\\0&10\end{pmatrix}.$$

XIV. Обчислюємо координати векторів $a^j$, $j=\overline{0,n}$ в новому базисі:

$$k=1:\quad j=0: z_{10}=\overline{z_{10}}-\theta_0\overline{z_{13}}=4-0,7\cdot 3=1,9;$$

$$j=1: z_{11}=\overline{z_{11}}-\theta_1\overline{z_{13}}=1-(-0,1)\cdot 3=1,3;$$

$$j=2: z_{12}=\overline{z_{12}}-\theta_2\overline{z_{13}}=2-(-0,4)\cdot 3=3,2;$$

$$j=3: z_{13}=\overline{z_{13}}-\theta_3\overline{z_{13}}=3-1\cdot 3=0;$$

$$j=4: z_{14}=\overline{z_{14}}-\theta_4\overline{z_{13}}=1-0=1;$$

$$j=5: z_{15}=\overline{z_{15}}-\theta_5\overline{z_{13}}=0-0,1\cdot 3=-0,3;$$

$$k=2:\quad j=0: z_{20}=\theta_0=0,7;\qquad j=1: z_{21}=\theta_1=-0,1;$$

$$j=2: z_{22}=\theta_2=-0,4;\quad j=3: z_{23}=\theta_3=1;$$

$$j=4: z_{24}=\theta_4=0;\qquad j=5: z_{25}=\theta_5=0,1.$$

XV. Переходимо до нового опорного розв'язку:

$$k=1: x_{i_1}^1=x_4^1=z_{10}=1,9;$$

$$k=2: x_{i_2}^1=x_3^1=z_{20}=0,7;$$

$$x^1=(0;0;0,7;1,9;0)^T.$$

XVI. Переходимо на крок III.

*2-а ітерація:*

X. Обчислюємо для $j=\overline{1,n}$ оцінки (враховуючи зауваження 2):

$$j=1: \Delta_1=\overline{\Delta_1}-\theta_1\Delta_3=-2-(-0,1)\cdot(-7)=-2,7;$$

$$j=2: \Delta_2=\overline{\Delta_2}-\theta_2\Delta_3=-1-(-0,4)\cdot(-7)=-3,8;$$

$$j=3: \Delta_3=\overline{\Delta_3}-\theta_3\Delta_3=-7-1\cdot(-7)=0;$$

$$j=4: \Delta_4=\overline{\Delta_4}-\theta_4\Delta_3=0-0=0;$$

$$j=5: \Delta_5=\overline{\Delta_5}-\theta_5\Delta_3=0-0,1\cdot(-7)=0,7.$$

V. Оскільки $\Delta_1<0$, $\Delta_2<0$, тому переходимо на крок V.

V–VIII. Оскільки при всіх $j=\overline{1,5}$ існує $z_{mj}>0$, то переходимо на крок IX.

IX. Вибираємо $s=2$, оскільки $\Delta_2=-3,8$.

X. Обчислюємо відношення $z_{k0}/z_{ks}$, $k=\overline{1,2}$, $\theta_0$ і $r$:

$k=1:\dfrac{z_{10}}{z_{12}}=\dfrac{1,9}{3,2}=\dfrac{19}{32}$, $z_{22}<0$, тому $\theta_0=\dfrac{19}{32}$, $r=1$.

XI. Обчислимо $\theta_j$, $j=\overline{1,5}$:

$j=1:\theta_1=z_{11}/z_{12}=1,3/3,2=13/32$; $j=2:\theta_2=z_{12}/z_{12}=1$;
$j=3:\theta_3=z_{13}/z_{12}=0$; $j=4:\theta_4=z_{14}/z_{12}=1/3,2=5/16$;
$j=5:\theta_5=z_{15}/z_{12}=-0,3/3,2=-3/32$.

XII. Запам'ятовуємо коефіцієнти розкладу і оцінки:

$$\left(\overline{z_{10}};\overline{z_{20}}\right)=(1,9;0,7);\ \left(\overline{z_{11}};\overline{z_{21}}\right)=(1,3;-0,1);\ \left(\overline{z_{12}};\overline{z_{22}}\right)=(3,2;-0,4);$$
$$\left(\overline{z_{13}};\overline{z_{23}}\right)=(0;1);\ \left(\overline{z_{14}};\overline{z_{24}}\right)=(1;0);\ \left(\overline{z_{15}};\overline{z_{25}}\right)=(-0,3;0,1);$$
$$\overline{\Delta}_1=-2,7;\ \overline{\Delta}_2=-3,8;\ \overline{\Delta}_3=0;\ \overline{\Delta}_4=0;\ \overline{\Delta}_5=0,7.$$

XIII. Переходимо до нового базису:

$$a^{i_1}=a^2=\begin{pmatrix}2\\-4\end{pmatrix};\ \ i_1=2,\ i_2=3,\ B=\begin{pmatrix}2&3\\-4&10\end{pmatrix}.$$

XIV. Обчислюємо координати векторів $a^j$, $j=\overline{0,n}$ в новому базисі:

$k=1$: $j=0:z_{10}=\theta_0=19/32$; $j=1:z_{11}=\theta_1=13/32$;
$j=2:z_{12}=\theta_2=1$; $j=3:z_{13}=\theta_3=0$;
$j=4:z_{14}=\theta_4=5/16$; $j=5:z_{15}=\theta_5=-3/32$.

$k=2$: $j=0:z_{20}=0,7-(19/32)\cdot(-0,4)=15/16$;
$j=1:z_{21}=-0,1-(13/32)\cdot(-0,4)=1/16$;
$j=2:z_{22}=-0,4-1\cdot(-0,4)=0$;
$j=3:z_{23}=1-0=1$;
$j=4:z_{24}=0-(5/16)\cdot(-0,4)=1/8$;
$j=5:z_{25}=0,1-(-3/32)\cdot(-0,4)=1/16$.

XV. Переходимо до нового опорного розв'язку:

$$k=1:x^1_{i_1}=x^1_2=z_{10}=19/32;$$
$$k=2:x^1_{i_2}=x^1_3=z_{20}=15/16;$$
$$x^1=(0;19/32;15/16;0;0)^T.$$

XVI. Переходимо на крок III.

*3-я ітерація:*

XI. Обчислюємо для $j=\overline{1,n}$ оцінки:

$j=1:\Delta_1=-2,7-\dfrac{13}{32}(-3,8)=-\dfrac{37}{32}$; $\quad j=2:\Delta_2=-3,8-1\cdot(-3,8)=0$;

$j=3:\Delta_3=0-0=0$; $\quad j=4:\Delta_4=0-(5/16)(-3,8)=19/16$;

$j=5:\Delta_5=0,7-(3/32)(-3,8)=11/32$.

VI. Оскільки $\Delta_1<0$, тому переходимо на крок V.

V-VIII. Оскільки при всіх $j=\overline{1,5}$ існує $z_{mj}>0$, то переходимо на крок IX.

IX. Вибираємо $s=1$, оскільки $\Delta_1=-37/32$.

X. Обчислюємо відношення $z_{k0}/z_{ks}$, $k=\overline{1,2}$, $\theta_0$ і $r$:

$k=1: z_{10}/z_{11}=19/13$; $\quad k=2: z_{20}/z_{21}=15$,

тому $\theta_0=19/13$, $r=1$.

XI. Обчислюємо $\theta_j$, $j=\overline{1,5}$:

$j=1:\theta_1=z_{11}/z_{11}=1$; $\quad j=2:\theta_2=z_{12}/z_{11}=32/13$;

$j=3:\Theta_3=\dfrac{z_{13}}{z_{11}}=0$; $\quad j=4:\Theta_4=\dfrac{z_{14}}{z_{11}}=10/13$;

$j=5:\Theta_5=\dfrac{z_{15}}{z_{11}}=-3/13$.

XII. Запам'ятовуємо коефіцієнти розкладу і оцінки:

$$\left(\overline{z_{10}};\overline{z_{20}}\right)=(19/32;15/16);\ \left(\overline{z_{11}};\overline{z_{21}}\right)=(13/32;1/16);$$

$$\left(\overline{z_{12}};\overline{z_{22}}\right)=(1;0);\ \left(\overline{z_{13}};\overline{z_{3}}\right)=(0;1);\ \left(\overline{z_{14}};\overline{z_{24}}\right)=(5/16;1/8);$$

$$\left(\overline{z_{15}};\overline{z_{25}}\right)=(3/32;1/16);$$

$$\overline{\Delta}_1=-37/32,\ \overline{\Delta}_2=0,\ \overline{\Delta}_3=0,\ \overline{\Delta}_4=19/16,\ \overline{\Delta}_5=11/32.$$

XIII. Переходимо до нового базису:

$$a^{i_1}=a^1=\begin{pmatrix}1\\-1\end{pmatrix};\ i_1=1,\ i_2=3,\ B=\begin{pmatrix}1&3\\-1&10\end{pmatrix}.$$

XIV. Обчислюємо координати векторів $a^j$, $j=\overline{0,n}$ в новому базисі:

$k=1:\quad j=0: z_{10}=\theta_0=19/13;\quad j=1: z_{11}=\theta_1=1;$

$j=2: z_{12}=\theta_2=32/13;\quad j=3: z_{13}=\theta_3=0;$

$j=4: z_{14}=\theta_4=10/13;\quad j=5: z_{15}=\theta_5=-3/13;$

$$k=2:\quad j=0: z_{20}=\frac{15}{16}-\frac{19}{13}\cdot\frac{1}{16}=\frac{11}{13};\qquad j=1: z_{21}=\frac{1}{16}-1\cdot\frac{1}{16}=0;$$

$$j=2: z_{22}=0-\frac{32}{13}\cdot\frac{1}{16}=-\frac{2}{13};\qquad j=3: z_{23}=1-0=1;$$

$$j=4: z_{24}=\frac{1}{8}-\frac{10}{13}\cdot\frac{1}{16}=\frac{1}{13};\qquad j=5: z_{25}=\frac{1}{16}-\left(-\frac{3}{13}\right)\cdot\frac{1}{16}=\frac{1}{13}.$$

XV. Переходимо до нового опорного розв’язку:

$$k=1: x_{i_1}^1=x_1^1=z_{10}=\frac{19}{13};\quad k=2: x_{i_2}^1=x_3^1=z_{20}=\frac{11}{13};\quad x^1=\left(\frac{19}{13};0;\frac{11}{13};0;0\right)^T.$$

XVI. Переходимо на крок III.

*4-а ітерація:*

III. Обчислюємо для $j=\overline{1,n}$ оцінки:

$$j=1: \Delta_1=-\frac{37}{32}-1\cdot\left(-\frac{37}{32}\right)=0;\qquad j=2: \Delta_2=0-\frac{32}{13}\cdot\left(-\frac{37}{32}\right)=\frac{37}{13};$$

$$j=3: \Delta_3=0-0=0;\qquad j=4: \Delta_4=\frac{19}{16}-\frac{10}{13}\cdot\left(-\frac{37}{32}\right)=\frac{27}{13};$$

$$j=5: \Delta_5=\frac{11}{32}-\frac{3}{13}\cdot\left(-\frac{37}{32}\right)=\frac{1}{13}.$$

IV. Оскільки всі $\Delta_j\ge 0$, $j=\overline{1,5}$, то $x^*=x^1=\left(\frac{19}{13};0;\frac{11}{13};0;0\right)^T$ і завершуємо обчислення.

Отже, оптимальний розв’язок $x^*=\left(\frac{19}{13};0;\frac{11}{13};0;0\right)^T$ СЗЛП отримали на четвертій ітерації, а оптимальний розв’язок початкової ЗЗЛП – $x^*=\left(\frac{19}{13};0;\frac{11}{13}\right)^T$, при цьому, оптимальне значення цільової функції $L(x^*)=-2\cdot\left(\frac{19}{13}\right)-0-7\cdot\left(\frac{11}{13}\right)=-\frac{115}{13}$.

**4. Симплекс-метод розв'язування виродженої стандартної задачі лінійного програмування**

В цьому пункті приводиться симплекс-метод розв'язування виродженої задачі 2, для якої виконуються умови (***i***), (***ii***), (***iii***) припущення 2.

Якщо застосувати алгоритм 2 для розв'язування виродженої задачі 2, то можливе зациклювання процесу розв'язування задачі (тобто повернення до уже пройденого базису нескінченне число разів), оскільки на кроці X алгоритму 2 у виродженому випадку існує більше ніж один номер $r$, для якого досягається мінімальне відношення $\theta_0 = z_{r0} / z_{rs}$. Симплекс-метод для розв'язування виродженої задачі лінійного програмування включає в себе спеціальне правило вибору вектора, що виводиться з базису, яке гарантує від зациклювання процесу розв'язуваня задачі.

***Алгоритм 4***

I–IX (кроки I–IX такі, як і в алгоритмі 2).

X. Обчислити відношення $z_{k0} / z_{ks}$ для тих $k$, для яких $z_{ks} > 0$ та позначити мінімальне з цих відношень через $\theta_0$.

XI. Покласти $j = 0$.

XII. Покласти $\theta_0^0 = \theta_0$.

XIII. Обчислити індекси $r_1, r_2, ... r_{l_j}$ такі, що:

$$z_{r_1 0} / z_{r_1 s} = z_{r_2 0} / z_{r_2 s} = ... = z_{r_{l_j} 0} / z_{r_{l_j} s} = \theta_0^0 .$$

XIV. Якщо $l_j = 1$, то замість вектора $a^{i_{r_1}}$ ввести в базис вектор $a^s$ (тобто покласти $i_{r_1} = s$) і перейти на крок XIX; якщо $l_j > 1$, то перейти на крок XV.

XV. Обчислити

$$\theta_0^{j+1} = \min \left\{ z_{r_1 (j+1)} / z_{r_1 s}, z_{r_2 (j+1)} / z_{r_2 s}, ..., z_{r_{l_j} (j+1)} / z_{r_{l_j} s} \right\} .$$

XVI. Покласти $\overline{r}_1 = r_1$, $\overline{r}_2 = r_2, ...,$ $\overline{r}_{l_j} = r_{l_j}$.

XVII. Серед множини індексів $\{\overline{r}_1, \overline{r}_2, ..., \overline{r}_{l_j}\}$ знайти індекси $r_1, r_2, ..., r_{l_{j+1}}$, для яких:

$$z_{r_1 (j+1)} / z_{r_1 s} = z_{r_2 (j+1)} / z_{r_2 s} = ... = z_{r_{l_{j+1}} (j+1)} / z_{r_{l_{j+1}} s} = \theta_0^{j+1} .$$

XVIII. Покласти $j = j + 1$ і перейти на крок XIV.

XIX. При кожному $j \in [1 : n]$ обчислити

$$\theta_j = z_{r_1 j} / z_{r_1 s} .$$

XX. Покласти $\overline{z}_{kj} = z_{kj}$, $k = \overline{1, m}$; $j = \overline{1, n}$.

XXI. Обчислити координати векторів $a^0, a^1, ..., a^n$ в новому базисі:

при $k \neq r_1$

$$z_{kj} = \overline{z}_{kj} - \theta_j \overline{z}_{ks}, \quad j = \overline{0, n};$$
$$k = 1, ..., r_1 - 1, r_1 + 1, ..., m;$$

при $k = r_1$

$$z_{kj} = \theta_j, \ \ j = \overline{0, n}.$$

XXII. Перейти до нового опорного розв'язку $x^1 = (x_1^1, ..., x_n^1)^T$ : $x_{i_k}^1 = z_{k0}$, $k = \overline{1, m}$, решту координат вектора $x^1$ дорівнюють нулю.

XXIII. Перейти на крок III.

***Теорема 4.*** *Якщо виконані умови* (***i***), (***ii***), (***iii***) *припущення 2, то через скінченне число ітерацій процес розв'язування задачі 2 алгоритмом 4 закінчиться або на кроці IV* (*знаходиться оптимальний розв'язок задачі 2, який рівний опорному* $x^1 = (x_1^1, ..., x_n^1)$ ), *або на кроці VII* (*встановлюється, що оптимального розв'язку задачі 2 немає*).

*Зауваження 4.* На практиці алгоритм 4 використовується рідко, оскільки він потребує значно більше часу для розв'язування задачі лінійного програмування в порівнянні з алгоритмом 2, а зациклювання процесу розв'язування виродженої задачі алгоритмом 2 малоймовірно. Якщо ж при розв'язуванні задачі 2 алгоритмом 2 відбулося зациклювання на деякому опорному розв'язку, то потрібно використовувати алгоритм 4 для отримання нового опорного розв'язку і далі продовжувати розв'язування задачі 2 з допомогою алгоритму 2.

### 5. Модифікований симплекс-метод

В багатьох випадках модифікований симплекс-метод (або метод оберненої матриці) більш економний в обчисленнях, ніж звичайний (особливо це проявляється, якщо число рівнянь-обмежень $m$ істотно менше розмірності простору $n$). На кожній ітерації в модифікованому алгоритмі обчислюється матриця, обернена до базисної, що зводиться до обчислення за основними формулами $m$ векторів розмірності $m$ (нагадаємо, що в основному симплекс-методі на кожній ітерації за основними формулами обчислюється $n - m + 1$ вектор). Модифікований симплекс-метод застосовується для розв'язування невироджених і вироджених стандартних задач лінійного програмування.

***Алгоритм 5***

Початок. I. Знайти початковий опорний розв'язок $x^1 = (x_1^1, x_2^1, ..., x_n^1)^T$ задачі 2 і базис $a^{i_1}, a^{i_2}, ..., a^{i_m}$ цього опорного розв'язку.

(В загальному випадку для знаходження початкового опорного розв'язку і відповідного йому базису використовують алгоритм 3 або алгоритм 3′).

II. Обчислити матрицю $B^{-1} = \left(\beta_{kj}\right)_{k=\overline{1,m}}^{j=\overline{1,m}}$, обернену до початкової базисної матриці $B$ .

III. Покласти $z_{k0} = x_{i_k}^1$ , $k = \overline{1,m}$ .

Основний цикл. IV. Покласти $c_{баз} = (c_{i_1}, c_{i_2}, ..., c_{i_m})$ .

V. Обчислити вектор

$$b = c_{баз} B^{-1} .$$

VI. Для кожного $j \in [1:n]$ обчислити оцінку

$$\Delta_j = ba^j - c_j , \quad j = \overline{1,n} .$$

VII. Якщо всі $\Delta_j \ge 0$ $(j = \overline{1,n})$, то покласти $x^* = x^1$ і припинити обчислення (тобто в цьому випадку опорний розв'язок $x^1$ оптимальний); якщо при деякому $j \in [1:n]$ виконується $\Delta_j < 0$, то перейти на крок VIII.

VIII. Вибрати згідно завчасно обумовленого правила такий індекс $s \in [1:n]$, щоб $\Delta_s < 0$.

IX. Обчислити вектор-стовпець

$$\left(z_{1s}, z_{2s}, ..., z_{ms}\right)^T = B^{-1} a^s .$$

X. Якщо при всіх $k \in [1:m]$ виконується $z_{ks} \le 0$, то припинити обчислення (в цьому випадку цільова функція $(c, x)$ не обмежена (зверху) на допустимій множині $X$); інакше перейти на крок XI.

XI. Обчислити відношення $z_{k0} / z_{ks}$ для всіх $k \in [1:m]$, для яких $z_{ks} > 0$ та позначити мінімальне з цих відношень через $\theta_0$ .

XII. Знайти такий індекс $r \in [1:m]$, що $z_{rs} > 0$ і $z_{r0} / z_{rs} = \theta_0$ .

XIII. Покласти $\bar{\beta}_{kj} = \beta_{kj}$ , $k = 1, ..., m;$ $j = \overline{1,m}$ , і $\bar{z}_{k0} = z_{k0}$, $k = \overline{1,m}$ .

XIV. Перейти до нового базису $a^{i_1}, ..., a^{i_m}$ , який отримуємо заміною вектора $a^{i_r}$ в попередньому базисі вектором $a^s$ (тобто покласти $i_r = s$ ).

XV. Обчислити за основними формулами матрицю $B^{-1} = \left(\beta_{kj}\right)_{k=\overline{1,m}}^{j=\overline{1,m}}$, обернену до нової базисної матриці:

$$\beta_{rj} = \bar{\beta}_{rj} / z_{rs} , \quad j = \overline{1,m};$$

$$\beta_{kj} = \bar{\beta}_{kj} - (\bar{\beta}_{rj} / z_{rs}) z_{ks} , \quad k = \overline{1,m}; \quad j = \overline{1,m}; \quad k \ne r.$$

XVI. Обчислити за основними формулами нові значення базисних координат опорного розв'язку:

$$z_{r0} = z_{r0} / z_{rs} ;$$

$$z_{k0} = \overline{z}_{k0} - (\overline{z}_{r0} / z_{rs}) z_{ks}, \quad k = 1, 2, ..., r-1, r+1, ..., m.$$

XVII. Перейти до нового опорного розв'язку $x^1 = (x_1^1, ..., x_n^1)^T$ :

$$x_{i_k}^1 = z_{k0}, \quad k = \overline{1, m},$$

інші координати вектора $x^1$ дорівнюють нулю.

XVIII. Перейти на крок IV.

***Теорема 5.*** *Якщо виконані припущення 2 і задача 2 невироджена, то процес розв'язування задачі 2 алгоритмом 5 через скінченне число ітерацій закінчиться або на кроці VII (знаходиться оптимальний розв'язок задачі 2), або на кроці X (встановлюється, що оптимального розв'язку задачі 2 немає).*

*Зауваження 5.* На практиці алгоритм 5 застосовується також для розв'язування виродженої задачі лінійного програмування, для якої виконуються умови $(\boldsymbol{i}), (\boldsymbol{ii}), (\boldsymbol{iii})$ припущення 2. В цьому випадку теоретично можливе зациклювання процесу розв'язування задачі 2 алгоритмом 5. Якщо при розв'язуванні задачі 2 алгоритмом 5 на деякому опорному розв'язку все-таки відбудеться зациклювання, то (аналогічно пункту 4) необхідно включити в алгоритм 5 спеціальне правило вибору вектора, що виводиться з базису, яке гарантує незациклювання.



*Зауваження 5'.* Якщо на кроці XV алгоритму 5 матрицю $B^{-1}$ обчислити за формулою $B^{-1} = D_r \overline{B}^{-1}$, де $\overline{B}^{-1}$ – матриця, обернена до базисної на попередній ітерації, а $D_r$ має вигляд:

$$D_r = \begin{Bmatrix} 1 \ldots 0 & -z_{1s} / z_{rs} & 0 \ldots 0 \\ & \ldots\ldots\ldots\ldots\ldots\ldots & \\ 0 \ldots 1 & -z_{r-1,s} / z_{rs} & 0 \ldots 0 \\ 0 \ldots 0 & 1 / z_{rs} & 0 \ldots 0 \\ 0 \ldots 0 & -z_{r+1,s} / z_{rs} & 1 \ldots 0 \\ & \ldots\ldots\ldots\ldots\ldots\ldots & \\ 0 \ldots 0 & -z_{ms} / z_{rs} & 0 \ldots 1 \end{Bmatrix},$$

то такий модифікований симплекс-метод називається ***мультиплікативним***. Для запису матриці $D_r$ достатньо знати $m+1$ число: число $r$ та числа $-z_{1s} / z_{rs}, ..., 1/ z_{rs}, ..., -z_{ms} / z_{rs}$. Якщо позначити через $D_{r_k}$ матрицю $D_r$ в $k$-й ітерації мультиплікативного алгоритму і за початкову базисну вибрати одиничну матрицю $I$, то обернена до базисної в $k$-й ітерації буде обчислюватися за формулою

$$B^{-1} = D_{r_k} D_{r_{k-1}} ... D_{r_1} I .$$

З цієї формули бачимо, що для обчислення матриці $B^{-1}$ в $k$-й ітерації мультиплікативного алгоритму потрібно знати $(m+1)k$ чисел, що на початковій стадії обчислення приводить до економії оперативної пам'яті ЕОМ.

## 5.8. Двоїстий симплекс-метод

***Двоїстою*** (***спряженою***) до задачі 2 п. 5.7 називають наступну задачу.

З а д а ч а 0. Знайти $\arg\min\limits_{y \in Y}(a^0, y)$ для заданого вектора $a^0$ і множини $Y$, заданої співвідношеннями

$$Y = \{ y \mid A^T y \geq c^T, \; y \in R^m \}.$$

*Припущення 0*. (***i***) – ранг матриці $A$ дорівнює $m$; (***ii***) – $n > m$; (***iii***) – задача 0 п. 5.7 – розв'язна.

#### 1. Основний метод

*Означення 1*. ***Майже допустимим опорним розв'язком задачі*** 2 п. 5.7 називається вектор $\tilde{x} = (\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_n)^T$, який задовольняє рівнянню $Ax = a^0$, причому вектори $a^j$, що відповідають ненульовим координатам вектора $\tilde{x}$, лінійно незалежні.

*Означення 2*. ***Базисом*** майже допустимого опорного розв'язку $\tilde{x}$ задачі 2 п. 5.7 називається будь-який упорядкований набір з $\tilde{m}$ лінійно незалежних векторів $a^j$, який містить всі вектори $a^j$, що відповідають ненульовим координатам вектора $\tilde{x}$.

Практичний інтерес представляє майже допустимий опорний розв'язок і його базис, для яких виконується $\Delta_j \geq 0, \; j = \overline{1,n}$ (часто такий розв'язок називається ***псевдопланом*** задачі 2 п. 5.7, а базис – ***псевдобазисом*** задачі 2 п. 5.7). В алгоритмі 1 на кожній ітерації здійснюється перехід від одного базису задачі 0 до іншого (або, що теж саме, від одного псевдобазису задачі 2 п. 5.7 до іншого).

*Означення 3*. Допустимий розв'язок $y$ двоїстої задачі 0 називається ***опорним***, якщо серед умов $A^T y \geq c^T$, які він перетворює в рівності, є $m$ лінійно незалежних умов.

*Означення 4*. ***Базисом*** опорного розв'язку двоїстої задачі 0 називається довільна система з $m$ лінійно незалежних векторів $a^{i_1}, ..., a^{i_m}$ прямої задачі, для яких

$$(y, a^{i_s}) = c_{i_s}, \; s = \overline{1,m}.$$

*Означення 5.* Опорний розв'язок $y$ двоїстої задачі 0 називається ***невиродженим***, якщо для будь-якого вектора $a^j$, що не входить у його базис, $(y, a^j) > c_j$.

*Означення 6.* Двоїста задача 0, всі опорні розв'язки якої є невиродженими, називається ***невиродженою***.

Двоїстий симплекс-метод (або метод послідовного уточнення оцінок) розв'язування задачі 2 п. 5.7 по своїй суті – це застосування звичайного симплекс-методу до двоїстої задачі 0, доповнене побудовою на кожній ітерації $n$-вимірного вектора $\tilde{x}$, який є майже допустимим опорним розв'язком задачі 2 п. 5.7 (звідси походить назва методу). Двоїстий симплекс-метод полягає в такому послідовному переході від одного майже допустимого розв'язку $\tilde{x}$ до іншого, що через скінченне число переходів або одержимо оптимальний розв'язок задачі 2 п. 5.7, або встановимо, що задача 2 п. 5.7 не має допустимих розв'язків. Оптимальному розв'язку $x^*$ задачі 2 п. 5.7 відповідає оптимальний розв'язок $y^*$ задачі 0, причому оптимальні значення прямої і двоїстої задач співпадають.

Оптимальний розв'язок двоїстої задачі визначається за формулою $y^* = (c_{i_1}, c_{i_2}, ..., c_{i_m})B^{-1}$, де $B-$ базисна матриця оптимального опорного розв'язку задачі 2 п. 5.7; $i_1, ..., i_m -$ індекси базисних векторів.

Для початку роботи алгоритму необхідно мати опорний розв'язок $\bar{y} = (\bar{y}_1, ..., \bar{y}_m)^T$ спряженої (двоїстої) задачі і його базис $a^{i_1}, ..., a^{i_m}$, на основі яких обчислюється початковий майже допустимий опорний розв'язок з невідємними оцінками за правилом:

спочатку знаходять коефіцієнти розкладу $z_{l0}$, $l = \overline{1, m}$ вектора $a^0$ по базису $a^{i_1}, ..., a^{i_m}$; далі складають вектор $\tilde{x} = (\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_n)$ $l$-а компонента, $\tilde{x}_{i_l}$ $(l = \overline{1, m})$ якого дорівнює $z_{l0}$, інші компоненти вектора $\tilde{x}$ дорівнюють нулю. Алгоритми пошуку опорного розв'язку спряженої задачі і його базису наведені в пункті 2.

Алгоритм 1 застосовують для виродженої і невиродженої задачі 2 п. 5.7. Якщо при використанні алгоритму для розв'язування задачі 2 п. 5.7 у виродженому випадку виникає зациклювання , то в алгоритмі 1 необхідно застосовувати процедуру, що гарантує відсутність зациклювання .

***Алгоритм 1***

Початок. I. Знайти майже допустимий опорний розв'язок $\tilde{x} = (\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_n)^T$ задачі 2 п. 5.7 (з невід'ємними оцінками $\Delta_j$, $j = \overline{1, n}$) і базис $a^{i_1}, a^{i_2}, ..., a^{i_m}$ цього майже допустимого опорного розв'язку.

II. Обчислити матрицю $B^{-1}$ обернену до початкової базисної матриці $B$, яка складається зі стовпців $a^{i_1}, a^{i_2}, ..., a^{i_m}$.

III. Розкласти по початковому базису $a^{i_1}, a^{i_2}, ..., a^{i_m}$ вектори $a^0, a^1, a^2, ..., a^n$ (тобто знайти числа $z_{kj}$, $k = \overline{1,m}$; $j = \overline{0,n}$ такі, що $a^j = \sum_{k=1}^{m} z_{kj} a^{i_k}$, $j = \overline{0,n}$) за формулами:

$$z_{10} = \tilde{x}_{i_1},\ z_{20} = \tilde{x}_{i_2}, ...,\ z_{m0} = \tilde{x}_{im};$$

$$(z_{1j}, z_{2j}, ..., z_{mj})^T = B^{-1}(\alpha_{1j}, \alpha_{2j}, ..., \alpha_{mj})^T,\ j = \overline{1,n}.$$

Відзначимо, що заздалегідь відомий розклад по початковому базису базисних векторів $a^{i_k}$, $k = \overline{1,m}$:

$$z_{li_k} = 0 \text{ при } k \neq l,\ z_{li_k} = 1 \text{ при } k = l,\ k = \overline{1,m};\ l = \overline{1,m}.$$

IV. Обчислити для кожного $j \in [1:n]$ оцінку

$$\Delta_j = \sum_{k=1}^{m} c_{i_k} z_{kj} - c_j.$$

Основний цикл. V. Якщо для всіх $k \in [1:m]$ виконується нерівність $z_{k0} \geq 0$, то покласти $x^* = \tilde{x}$ і припинити обчислення (тобто в цьому випадку майже допустимий опорний розв'язок $\tilde{x}$ оптимальний); інакше перейти на крок VI.

VI. Якщо існує таке $k \in [1:m]$, що $z_{k0} < 0$ і при всіх $j \in [1:n]$ виконується $z_{kj} \geq 0$, то припинити обчислення (у цьому випадку задача 2 п. 5.7 не має допустимих розв'язків); інакше перейти на крок VII.

VII. Вибрати (по заздалегідь вибраному правилу) індекс $r \in [1:m]$ такий, щоб $z_{r0} < 0$.

VIII. Обчислити відношення $\Delta_j / z_{rj}$ для всіх тих $j \in [1:n]$, для яких $z_{rj} < 0$ і позначити мінімальне по абсолютній величині із цих відношень через $\theta_0$ (нагадаємо, що $\Delta_j \geq 0,\ j \in [1:n]$).

IX. Знайти індекс $s \in [1:n]$ такий, що $z_{rs} < 0$ і $-\Delta_s / z_{rs} = \theta_0$.

Для невиродженої задачі 2 п. 5.7 індекс $s$ визначається однозначно. У виродженому випадку мінімальне відношення $\theta_0$ може досягатися для декількох індексів $j \in [1:n]$. Тому на практиці в якості $s$ вибирається найменший індекс $j$, для якого досягається мінімальне відношення. Теоретично у виродженому випадку необхідно застосовувати процедуру вибору $s$, що гарантує відсутність зациклювання (див. пункт 3).

X. Покласти $\overline{z}_{kj} = z_{kj}$, $k = \overline{1,m}$; $j = \overline{0,n}$ і $\overline{\Delta}_j = \Delta_j$, $j = \overline{1,n}$.

XI. Перейти до нового базису $a^{i_1}, a^{i_2}, ..., a^{i_m}$, який отримується заміною вектора $a^{i_r}$ у попередньому базисі вектором $a^s$ (тобто покласти $i_r = s$).

XII. Обчислити координати векторів $a^0, a^1, ..., a^n$ у новому базисі:

$$z_{kj} = \overline{z}_{kj} - (\overline{z}_{rj} / \overline{z}_{rs})\overline{z}_{ks}, \quad j = \overline{0,n};$$

$$k = 1, ..., \ r-1, \ r+1, ..., m;$$

$$z_{rj} = \overline{z}_{rj} / \overline{z}_{rs}, \quad j = \overline{0,n}.$$

XIII. Обчислити за основними формулами оцінки $\Delta_j, \quad j \in [1:n]$, що відповідають новому базису:

$$\Delta_j = \overline{\Delta}_j - (\overline{z}_{rj} / \overline{z}_{rs})\overline{\Delta}_s, \quad j = \overline{1,n}.$$

XIV. Перейти до нового майже допустимого опорного розв’язку:

$$\tilde{x} = (\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_n)^T;$$

$$\tilde{x}_{i_k} = z_{k0}, \quad k = \overline{1,m};$$

інші координати вектора $\tilde{x}$ дорівнюють нулю.

XV. Перейти на крок V.

***Теорема 1.*** *Якщо виконані припущення 0 і задача 0 невироджена, то процес розв’язування задачі 2 п. 5.7 за допомогою алгоритму 1 закінчиться через скінченне число ітерацій або на кроці V (в цьому випадку знаходиться оптимальний розв’язок задачі 2 п. 5.7), або на кроці VI (у цьому випадку встановлюється, що задача 2 п. 5.7 не має допустимих розв’язків).*

*Зауваження 1*. Оскільки число ітерацій, необхідних для розв’зування задачі лінійного програмування, визначається в основному кількістю непрямих обмежень, то двоїстий симплекс-метод доцільно застосовувати, коли число непрямих обмежень істотно перевищує число змінних.

### 2. Методи пошуку початкового опорного розв’язку спряженої задачі

У підпункті 2.1 наводиться метод пошуку опорного розв’язку спряженої задачі 0 за умови, що відомий довільний допустимий розв’язок задачі 0. В підпункті 2.2 на основі методу наведеного в 2.1 розглядається алгоритм, який дозволяє знайти оптимальний розв’язок спряженої задачі (а значить, і прямої) без попереднього обчислення допустимого розв’язку спряженої задачі.

***2.1. Метод пошуку опорного розв’язку спряженої задачі по відомому допустимому розв’язку.***

Попередньо вкажемо клас задач, у яких визначення допустимого розв’язку є простою операцією. Нехай у задачі 2 п. 5.7 одна і та ж компонента $\alpha_{tj}$, $j=\overline{1,n}$ всіх векторів $a^j$, $j=\overline{1,n}$, додатня. Побудувати вектор:

$$y' = (\underbrace{\overbrace{0,...,0,y_t}^{t},0,...,0}_{m})^T,$$

де $y_t \geq \max\limits_{j\in[1;n]}(c_j/\alpha_{tj})$.

Вектор $y'$ задовольняє обмеженням задачі 0 і тому є допустимим розв'язком задачі 0.

*Припущення 2*. $(\boldsymbol{i})-$ ранг матриці $A$ дорівнює $m$; $(\boldsymbol{ii})-$ $n>m$; $(\boldsymbol{iii})-$ допустима множина спряженої задачі 0 непорожня.

***Алгоритм 2***

I. Знайти вектор $y'=(y_1',y_2',...,y_2')^T$ – такий допустимий розв'язок спряженої задачі 0, що:

$$\sum_{i=1}^{m}\alpha_{ij}y_i' > c_j,\quad j=\overline{1,n_1};\tag{5.38}$$

$$\sum_{i=1}^{m}\alpha_{ij}y_i' = c_j,\quad j=\overline{n_1+1,n}.\tag{5.39}$$



II. Знайти число $r-$ кількість лінійно незалежних умов в (5.39). Якщо $r=m$, то припинити обчислення (тобто $y'-$ опорний розв'язок задачі 0, а вектори $a^{i_1},...,a^{i_m}$, які визначаються лінійно незалежними умовами, утворюють базис цього опорного розв'язку); інакше перейти на крок III.

III. Виразити із системи рівностей:

$$\sum_{i=1}^{m}\alpha_{ij}y_i = c_j,\quad j=\overline{n_1+1,n},\tag{5.40}$$

$r$ змінних $y_1,...,y_r$ через інші змінні:

$$y_j=\sum_{i=r+1}^{m}d_{ij}y_i+d_j,\quad j=\overline{1,r}.\tag{5.41}$$

IV. Замінити змінні $y_i$, $i=\overline{1,r}$ у нерівностях:

$$\sum_{i=1}^{m}\alpha_{ij}y_i > c_j,\quad j=\overline{1,n_1},$$

їх виразами з (5.41) і одержати систему нерівностей відносно $y_{r+1},...,y_m$

$$\sum_{i=r+1}^{m}\alpha_{ij}'y_i > c_j',\quad j=\overline{1,n_1}.\tag{5.42}$$

V. Підставити в лінійну форму $(a^0,y)$ задачі 0 значення змінних $y_1,...,y_r$ з (5.41) і одержати лінійну функцію $g_1$ змінних $y_{r+1},...,y_m$

$$g_1(y_{r+1},...,y_m)=\sum_{i=r+1}^{m}b_iy_i+b_0\ .$$

VI. Якщо $b_i = 0,\ i = \overline{r+1, m},$ то перейти на крок XIII; інакше перейти на крок VII.

VII. Обчислити

$$\mu_j = \sum_{i=r+1}^{m} \alpha'_{ij} b_i, \quad j = \overline{1, n_1}.$$

VIII. Якщо $\mu_j \le 0, \quad j = \overline{1, n_1}$, то припинити обчислення (у цьому випадку лінійна форма задачі 0 не обмежена знизу на допустимій множині); інакше перейти на крок IX.

IX. Обчислити

$$\Delta_j = \sum_{i=r+1}^{m} \alpha'_{ij} y'_i - c'_j, \quad j = \overline{1, n_1}. \tag{5.43}$$

X. Обчислити

$$\theta_0 = \min_{\mu_j > 0} (\Delta_j / \mu_j).$$

XI. Перейти до нового допустимого розв'язку $y'' = (y''_1, ..., y''_m)^T$ за правилами:

$$y''_i = y'_i - \theta_0 b_i, \quad i = \overline{r+1, m};$$

$y''_1, ..., y''_r$ знаходять через $y''_{r+1}, ..., y''_m$ згідно (5.41).

XII. Обчислити індекс $k \in [1 : n_1]$ такий, що

$$\theta_0 = (\Delta_k / \mu_k), \quad \mu_k > 0$$

і припинити обчислення.

У цьому випадку побудований новий допустимий розв'язок $y''$, який перетворює в рівності $r+1$ лінійно незалежних умов з системи обмежень задачі 0, тобто $r$ лінійно незалежних умов із системи (5.39)

$$\sum_{i=1}^{m} \alpha_{ij} y''_i = c_j, \quad j = \overline{n_1 + 1, n},$$

і лінійно незалежну від системи (5.39) умову з (5.38)

$$\sum_{i=1}^{m} \alpha_{ik} y''_i = c_k, \quad k \in [1 : n_1].$$

XIII. Обчислити величини $\Delta_j, \quad j = \overline{1, n_1}$ згідно (5.43).

XIV. Обчислити величину $\theta_0$ за правилом:

$$\theta_0 = \begin{cases} \min\limits_{\alpha'_{r+1,j} > 0} (\Delta_j / \alpha'_{r+1,j}), & \text{якщо серед } \alpha'_{r+1,j},\ j = \overline{1, n_1}, \text{ є додатні величини;} \\ -\min\limits_{\alpha'_{r+1,j} < 0} (-\Delta_j / \alpha'_{r+1,j}), & \text{якщо } \alpha'_{r+1,j} \le 0 \text{ для } j = 1, ..., n_1. \end{cases}$$

XV. Перейти до нового допустимого розв'язку $y'' = (y_1'',..., y_m'')^T$ за правилами:

$$y_i'' = y_i', \quad i = \overline{r+2, m};$$

$$y_{r+1}'' = y_{r+1}' - \theta_0,$$

$y_1'',..., y_r''$ знаходять через $y_{r+1}'',..., y_m''$ за формулою (5.41).

XVI. Обчислити такий індекс $k \in [1 : n_1]$, що

$$\theta_0 = \Delta_k / \alpha_{r+1,k}' ,$$

і припинити обчислення (див. крок XII).

Тепер можна застосувати алгоритм 2 до допустимого розв'язку $y''$ (тобто зробити ще одну ітерацію) і одержати допустимий розв'язок $y'''$, який перетворює в рівності $r+2$ лінійно незалежні обмеження задачі 0 і т.д. Не більш ніж за $m - r'$ ітерацій буде отриманий опорний розв'язок спряженої задачі 0 (тут $r'$ – число лінійно незалежних умов-рівностей для $y'$).

***2.2. Метод пошуку оптимального розв'язку спряженої задачі без попереднього обчислення допустимого розв'язку.***

***Алгоритм 2'***

I. Вибрати достатньо велике число $\beta > 0$ ($\beta$ більше будь-якого числа, з яким його доводиться по ходу розв'язування порівнювати).

II. Скласти розширену спряжену задачу:

$$\text{знайти } \arg\min(\beta y_0 + \sum_{i=1}^{m} a_i^0 y_i)$$

при обмеженнях:

$$y_0 + \sum_{i=1}^{m} \alpha_{ij} y_i \geq c_j, \quad j = \overline{1, n}; \quad y_0 \geq 0 .$$

(Прямою розширеною задачею є задача 2 п. 5.7 (зі змінними $x_0, x_1,..., x_n$), у якій до обмежень $Ax = a^0$, $x \geq 0$ додані обмеження $x_0 + x_1 + ... + x_n = \beta$; $x_0 \geq 0$).

III. Обчислити

$$y_0' = \max\{c_1,..., c_n, 0\} .$$

IV. Скласти вектор

$$y' = (\underbrace{y_0', 0,..., 0}_{m+1})^T ,$$

який є допустимим розв'язком розширеної спряженої задачі.

V. Відштовхуючись від допустимого розв'язку $y'$ за допомогою методу, наведеного в підпункті 2.1, обчислити опорний розв'язок розширеної спряженої задачі, його базис і майже допустимий опорний розв'язок розширеної прямої задачі.

VI. Розв'язати пряму розширену задачу двоїстим симплекс-методом і одержати вектори: $\overline{x}^* = (x_0^*, x_1^*, ..., x_n^*)^T$ – оптимальний розв'язок прямої розширеної задачі; $\overline{y}^* = (y_0^*, y_1^*, ..., y_m^*)^T$ – оптимальний розв'язок розширеної спряженої задачі.

VII. Якщо $y_0^* = 0$, то $y^* = (y_1^*, ..., y_m^*)^T$ – оптимальний розв'язок спряженої задачі 0, а $x^* = (x_1^*, ..., x_m^*)^T$ – оптимальний розв'язок задачі 2 п. 5.7.

Якщо $y_0^* > 0$, то спряжена задача 0 і пряма задача 2 п. 5.7 – нерозв'язні.

Алгоритм 2′ за скінченне число ітерацій знаходить оптимальний розв'язок прямої задачі 2 п. 5.7 і двоїстої задачі 0 без попереднього обчислення допустимого розв'язок спряженої задачі.

Недоліком алгоритму 2′ є проблема вибору необхідного числа $\beta > 0$.

### 3. Правило вибору вектора $a^s$, який вводиться в базис, що гарантує від зациклювання у виродженому випадку



Якщо при використанні алгоритму 1 для розв'язування задачі лінійного програмування у виродженому випадку виникло зациклювання (тобто повернення до вже пройденого базису), то на кроці IX алгоритму 1 варто застосувати наступну процедуру вибору індексу $s \in [1:n]$, що гарантує від зациклювання.

***Алгоритм 3***

Початок. I. Позначити через $\Im_1$ сукупність індексів $j$, на яких досягається

$$\theta_0 = \min_{z_{rj}<0}(-\Delta_j / z_{rj}).$$

II. Покласти $v = 1$.

Основний цикл. III. Якщо $\Im_v$ містить один елемент, то цей елемент прийняти як індекс $s$ і припинити обчислення; інакше перейти на крок IV.

IV. Якщо всі індекси $j \in \Im_v$ більші ніж $i_{m-v+1}$, то покласти $s = \min_{j\in\Im_v} j$; інакше перейти на крок V.

V. Побудувати множину $\Im_{v+1}$, включаючи в неї ті індекси $j$, на яких досягається

$$\min_{\substack{j\in\Im_v \\ j<i_{m-v+1}}} (z_{m-v+1,\,j} / z_{rj}).$$

VI. Якщо $v < m$, то покласти $v = v + 1$ і перейти на крок III; інакше перейти на крок VII.

VII. Покласти $s = \min\limits_{j \in \Im_{m+1}} j$ .

**Приклад 1.** (*Задача А*) За допомогою алгоритму 1 знайти мінімум функції:

$$L(x) = x_1 + x_2 + 2x_3 + 8x_4 \to \min_x$$

при обмеженнях:

$$2x_1 - x_2 + 3x_3 - 2x_4 = 3,$$
$$-x_1 + 3x_2 - 4x_3 + 4x_4 = 1,$$
$$x_j \geq 0, \ \ j = \overline{1,4}.$$

*Розв'язування* (*Задача Б*). Спочатку перейдемо до задачі максимізації функції:

$$\overline{L}(x) = -x_1 - x_2 - 2x_3 - 8x_4 \to \max_x$$

при обмеженнях:

$$2x_1 - x_2 + 3x_3 - 2x_4 = 3,$$
$$-x_1 + 3x_2 - 4x_3 + 4x_4 = 1,$$
$$x_j \geq 0, j = \overline{1,4}.$$

Отже, в прямій задачі маємо: $c = (-1;-1;-2;-8)^T$ ,

$$A = \begin{pmatrix} 2 & -1 & 3 & -2 \\ -1 & 3 & -4 & 4 \end{pmatrix}, \ a^0 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \ a^1 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \ a^2 = \begin{pmatrix} -1 \\ 3 \end{pmatrix}, \ a^3 = \begin{pmatrix} 3 \\ -4 \end{pmatrix},$$

$$a^4 = \begin{pmatrix} -2 \\ 4 \end{pmatrix}, \ n = 4, \ m = 2.$$

Двоїстою до цієї задачі є наступна:

$$\overline{L}^{\partial в}(y) = (a^0, y) = 3y_1 + y_2 \to \min_y$$

при обмеженнях:

$$\begin{cases} 2y_1 - y_2 \geq -1, \\ -y_1 + 3y_2 \geq -1, \\ 3y_1 - 4y_2 \geq -2, \\ -2y_1 + 4y_2 \geq -8, \end{cases} \qquad y \in R^2.$$

Складаємо матрицю В із першого і третього стовпців матриці А і перевіряємо чи така матриця $B = \begin{pmatrix} 2 & 3 \\ -1 & -4 \end{pmatrix}$ є невиродженим базисом двоїстої задачі (тобто покладемо $i_1 = 1, \ i_2 = 3$). Знаходимо опорний розв'язок двоїстої задачі розв'язуючи систему:

$$\begin{cases}(a^{i_1}, y) = c_1 \\ (a^{i_2}, y) = c_3\end{cases} \Leftrightarrow \begin{cases}\left(\begin{pmatrix}2\\-1\end{pmatrix},\begin{pmatrix}y_1\\y_2\end{pmatrix}\right) = -1 \\ \left(\begin{pmatrix}3\\-4\end{pmatrix},\begin{pmatrix}y_1\\y_2\end{pmatrix}\right) = -2\end{cases} \Leftrightarrow \begin{cases}2y_1 - y_2 = -1 \\ 3y_1 - 4y_2 = -2\end{cases} \Leftrightarrow \begin{cases}\overline{y_1} = -2/5 \\ \overline{y_2} = 1/5.\end{cases}$$

Перевіряємо нерівності для небазисних векторів:

$$\begin{cases}(a^2, \overline{y}) > c_2 \\ (a^4, \overline{y}) > c_4\end{cases} \Leftrightarrow \begin{cases}\left(\begin{pmatrix}-1\\3\end{pmatrix},\begin{pmatrix}\overline{y_1}\\\overline{y_2}\end{pmatrix}\right) > -1 \\ \left(\begin{pmatrix}-2\\4\end{pmatrix},\begin{pmatrix}\overline{y_1}\\\overline{y_2}\end{pmatrix}\right) > -8\end{cases} \Leftrightarrow \begin{cases}-\overline{y_1} + 3\overline{y_2} > -1 \\ 2\overline{y_1} - 4\overline{y_2} > -8\end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases}2/5 + 3/5 > -1 \\ 4/5 + 4/5 > -8\end{cases} \Leftrightarrow \begin{cases}1 > -1 \\ 8/5 > -8.\end{cases}$$

Оскільки нерівності справджуються, то матриця $B = \begin{pmatrix}2 & 3\\-1 & -4\end{pmatrix}$ є невиродженим базисом двоїстої задачі, а $\overline{y} = (-2/5; 1/5)^T$ – опорний розв'язок двоїстої задачі.

**Алгоритм 1**

I. Виберемо $i_1 = 1$, $i_2 = 3$, тобто $a^{i_1} = \begin{pmatrix}2\\-1\end{pmatrix}$, $a^{i_2} = \begin{pmatrix}3\\-4\end{pmatrix}$,

$B = \begin{pmatrix}2 & 3\\-1 & -4\end{pmatrix}$, $\overline{y} = (-2/5; 1/5)^T$.

II. Обчислюємо матрицю обернену до $B$:

$$B^{-1} = \frac{1}{2\cdot(-4) - (-1)\cdot 3}\begin{pmatrix}-4 & -3\\1 & 2\end{pmatrix} = \begin{pmatrix}4/5 & 3/5\\-1/5 & -2/5\end{pmatrix}.$$

III. Розкладаємо по початковому базису $a^{i_1}$, $a^{i_2}$ вектори $a^0, a^1, ..., a^4$:

$$\begin{pmatrix}z_{10}\\z_{20}\end{pmatrix} = \begin{pmatrix}4/5 & 3/5\\-1/5 & -2/5\end{pmatrix}\begin{pmatrix}3\\1\end{pmatrix}, \quad \begin{matrix}z_{10} = 12/5 + 3/5 = 3, \\ z_{20} = -3/5 - 2/5 = -1,\end{matrix}$$

$z_{11} = 1$, $z_{21} = 0$ ($a^1 -$ базисний вектор),

$$\begin{pmatrix}z_{12}\\z_{22}\end{pmatrix} = \begin{pmatrix}4/5 & 3/5\\-1/5 & -2/5\end{pmatrix}\begin{pmatrix}-1\\3\end{pmatrix}, \quad \begin{matrix}z_{12} = -4/5 + 9/5 = 1, \\ z_{22} = 1/5 - 6/5 = -1;\end{matrix}$$

$z_{13} = 0, z_{23} = 1$, ($a^3 -$базисний вектор),

$$\begin{pmatrix}z_{14}\\z_{24}\end{pmatrix} = \begin{pmatrix}4/5 & 3/5\\-1/5 & -2/5\end{pmatrix}\begin{pmatrix}-2\\4\end{pmatrix}, \quad \begin{matrix}z_{14} = -8/5 + 12/5 = 4/5, \\ z_{24} = 2/5 - 8/5 = -6/5.\end{matrix}$$

IV. Знаходимо майже допустимий опорний розв’язок прямої задачі:

$$\tilde{x}_1^1 = z_{10} = 3,\ \tilde{x}_3^1 = z_{20} = -1,\ \tilde{x}^1 = (3;0;-1;0)^T,$$

V. Обчислимо оцінки:

$j=1: \Delta_1 = c_{i_1} z_{11} + c_{i_2} z_{21} - c_1 = -1\cdot 1 + (-2)\cdot 0 - (-1) = 0,$

$j=2: \Delta_2 = c_{i_1} z_{12} + c_{i_2} z_{22} - c_2 = -1\cdot 1 + (-2)\cdot(-1) - (-1) = 2,$

$j=3: \Delta_3 = c_{i_1} z_{13} + c_{i_2} z_{23} - c_3 = -1\cdot 0 + (-2)\cdot 1 - (-2) = 0,$

$j=4: \Delta_4 = c_{i_1} z_{14} + c_{i_2} z_{24} - c_4 = -1\cdot\left(4/5\right) + (-2)\cdot(-6/5) - (-8) = 48/5.$

*1-а ітерація:*

VI. Нерівності $z_{k0} \ge 0,\ k = 1,2$ не виконуються, тому переходимо на крок VII.

VII. При $k=2,\ z_{k0} < 0,$ проте нерівності $z_{2j} \ge 0,\ j = \overline{1,4}$ не виконуються, тому переходимо на крок VIII.

VIII. Знаходимо індекс $r = 2$.

IX. Обчислюємо відношення $\Delta_j / z_{2j}$ для $j = 2;4$ і знаходимо $\theta_0$:

$$\frac{\Delta_2}{z_{22}} = \frac{2}{-1} = -2;\ \frac{\Delta_4}{z_{24}} = \frac{48}{5}\Big/\frac{-6}{5} = -8;\ \theta_0 = |-2| = 2.$$

X. Знаходимо індекс $s = 2$.

XI. Запам’ятовуємо коефіцієнти розкладу і оцінки:

$\overline{z}_{kj} = z_{kj},\ k = \overline{1,2},\ k = \overline{1,2},\ j = \overline{0,4}$ і $\overline{\Delta}_j = \Delta_j,\ j = \overline{1,4}$:

$$\overline{z}_{10} = 3,\ \overline{z}_{20} = -1,\ \overline{z}_{11} = 1,\ \overline{z}_{21} = 0,\ \overline{z}_{12} = 1,\ \overline{z}_{22} = -1,$$

$$\overline{z}_{13} = 0,\ \overline{z}_{23} = 1,\ \overline{z}_{14} = 4/5,\ \overline{z}_{24} = -6/5,$$

$$\overline{\Delta_1} = 0,\ \overline{\Delta_2} = 2,\ \overline{\Delta_3} = 0,\ \overline{\Delta_4} = 48/5.$$

XII. Переходимо до нового базису $a^{i_1}$, $a^{i_2}$, замінюючи в старому базисі вектор $a^{i_2}$ на вектор $a^2$ (тобто покласти $i_1 = 1,\quad i_2 = 2$), $B = \begin{pmatrix} 2 & -1 \\ -1 & 3 \end{pmatrix}$.

XIII. Обчислюємо координати векторів $a^0, a^1, ..., a^4$ в новому базисі:

$k=1:\ j=0: z_{10} = \overline{z}_{10} - \left(\overline{z}_{20}/\overline{z}_{22}\right)\overline{z}_{12} = 3 - (-1/-1)\cdot 1 = 2,$

$j=1: z_{11} = 1$ ($a^1$ – базисний вектор),

$j=2: z_{12} = 0$ ($a^2$ – базисний вектор),

$$j=3: z_{13}=\bar z_{13}-\left(\bar z_{23}/\bar z_{22}\right)\bar z_{12}=0-\left(-1/-1\right)\cdot 1=1,$$

$$j=4: z_{14}=\bar z_{14}-\left(\bar z_{24}/\bar z_{22}\right)\bar z_{12}=4/5-\left(6/5\right)\cdot 1=-2/5;$$

$$k=2:\ j=0: z_{10}=\bar z_{20}-\left(\bar z_{20}/\bar z_{22}\right)=(-1/-1)=1,$$

$$j=1: z_{21}=0\ \ (a^1-\text{базисний вектор}),$$

$$j=2: z_{22}=1\ \ (a^2-\text{базисний вектор}),$$

$$j=3: z_{23}=\left(\bar z_{23}/\bar z_{22}\right)=\left(1/-1\right)=-1,$$

$$j=4: z_{24}=\left(\bar z_{24}/\bar z_{22}\right)=\left(-6/5\right)/(-1)=6/5.$$

XIV. Обчислюємо оцінки, які відповідають новому базису:

$$j=1: \Delta_1=\overline{\Delta_1}-0\cdot\overline{\Delta_2}=0, \qquad j=2: \Delta_2=\overline{\Delta_2}-1\cdot\overline{\Delta_2}=0,$$

$$j=3: \Delta_3=\overline{\Delta_3}+1\cdot\overline{\Delta_2}=0+2=2,$$

$$j=4: \Delta_4=\overline{\Delta_4}-\left(6/5\right)\cdot 2=48/5-12/5=36/5.$$

XV. Переходимо до нового опорного розв’язку:

$$\tilde x_1^1=z_{10}=2,\ \ \tilde x_2^1=z_{20}=1,\ \ \tilde x^1=\left(2;1;0;0\right)^T.$$

XVI. Переходимо на крок VI.

*2-а ітерація:*

VI. Нерівності $z_{k0}>0,\ k=1,2$ виконуються, тому $x^*=\left(2;1;0;0\right)^T$ і зупиняємо обчислення.

Таким чином, за одну ітерацію знайшли оптимальний розв’язок $x^*=\left(2;1;0;0\right)^T$ задачі Б, він також є оптимальним розв’язком задачі А.

Мінімальне значення цільової функції в задачі Б:

$$\bar L(x^*)=-2-1-0-0=-3,$$

а максимальне значення цільової функції в задачі А:

$$L(x^*)=-\bar L(x^*)=3.$$

Оптимальний розв’язок двоїстої задачі наступний:

$$y^*=\left(c_{i_1},c_{i_2}\right)B^{-1}=\left(-1,-1\right)\begin{pmatrix}2 & -1\\ -1 & 3\end{pmatrix}^{-1}=$$

$$=\left(-1,-1\right)\frac{1}{2\cdot 3-(-1)(-1)}\begin{pmatrix}3 & 1\\ 1 & 2\end{pmatrix}=(-1;-1)\frac{1}{5}\begin{pmatrix}3 & 1\\ 1 & 2\end{pmatrix}=$$

$$=\left(1/5\right)\left(-3-1-1-2\right)=\left(-4/5;-3/5\right).$$

Оптимальне значення цільової функції двоїстої задачі $L^{дв}(y^*)=3\left(-4/5\right)+\left(-3/5\right)=-3,$ тобто $\bar L(x^*)=L^{дв}(y^*)$.

## 5.9. Методи послідовного скорочення нев'язок

Процес обчислення оптимального розв'язку методом скорочення нев'язок не потребує попереднього визначення опорного розв'язку прямої чи двоїстої задачі. Для початку процесу достатньо знати лише довільний допустимий розв'язок двоїстої задачі. При наявності опорного розв'язку початкової задачі 2 п. 5.7 і довільного допустимого розв'язку двоїстої задачі 0 п. 5.8 модифікація методу скорочення нев'язок – метод двосторонніх оцінок дозволяє отримати наближений розв'язок з наперед заданою похибкою при суттєво меншій кількості ітерацій, ніж це вимагається для отримання оптимального розв'язку.

#### 1. Метод послідовного скорочення нев'язок

*Припущення 1.* $(\boldsymbol{i})-$ $a_i^0 \geq 0,\ i = \overline{1,m},$ $(\boldsymbol{ii})-$ ранг матриці $A$ дорівнює $m$; $(\boldsymbol{iii})-$ $n > m;$ $(\boldsymbol{iv})-$ двоїста задача 2 п. 5.8 має допустимі розв'язки.

Метод послідовного скорочення нев'язок використовує пряму і двоїсту задачі лінійного програмування. Правила переходу від однієї ітерації до іншої забезпечує скорочення нев'язок (різниць між лівими і правими частинами умов-рівностей задачі 2 п. 5.7). Через скінченне число ітерації або нев'язки зводяться до нуля (тобто знаходиться оптимальний розв'язок задачі 2 п. 5.7), або встановлюється нерозв'язність задачі 2 п. 5.7. На кожній ітерації алгоритму необхідно обчислювати оптимальні розв'язки деякої допоміжної задачі і двоїстої до неї. В методі необхідно попередньо обчислювати допустимий розв'язок двоїстої задачі 0 п. 5.8.

***Алгоритм 1***

Початок. I. Знайти довільний допустимий розв'язок $y = (y_1,\ y_2,\ \ldots,\ y_m)^T$ двоїстої задачі 0 п. 5.8.

Основний цикл. II. Знайти всі індекси $j \in [1:n],$ для яких

$$\sum_{i=1}^{m} \alpha_{ij} y_i = c_j\ ,$$

і позначити множину таких індексів через

$$\Im \triangleq \left\{ j_1, j_2, \ldots, j_l \right\}.$$

III. Використовуючи модифікований симплекс-метод, знайти оптимальний розв'язок $(\bar{x}_{j_1}^*, \bar{x}_{j_2}^*, \ldots, \bar{x}_{j_l}^*, \varepsilon_1^*, \varepsilon_2^*, \ldots, \varepsilon_m^*)$ допоміжної задачі 1 лінійного програмування в просторі $R^{l+m}$:

Задача 1. Знайти $\arg \min\limits_{(\bar{x}_{j_1}, \ldots, \bar{x}_{j_l}, \varepsilon_1, \ldots, \varepsilon_m)} \sum\limits_{i=1}^{m} \varepsilon_i$ при обмеженнях:

$$\sum_{j\in\Im}\alpha_{ij}\bar{x}_j+\varepsilon_i=a_i^0,\quad i=\overline{1,m};$$

$$\bar{x}_j\ge 0 \text{ при } j\in\Im;\quad \varepsilon_i\ge 0,\quad i=\overline{1,m}.$$

IV. Знайти оптимальний розв'язок $\mu^*=(\mu_1^*,\mu_2^*,...,\mu_m^*)$ задачі, двоїстої до приведеної вище допоміжної задачі 1, який можна обчислити за формулою:

$$\mu^*=(\bar{c}_{i_1},\bar{c}_{i_2},...,\bar{c}_{i_m})B_\Im^{-1},$$

де $B_\Im^{-1}$ – матриця, обернена до оптимальної базисної матриці задачі 1; $(\bar{c}_{i_1},\bar{c}_{i_2},...,\bar{c}_{i_m})$ – координати вектора $(\underbrace{0,...,0}_{l},\ \underbrace{1,1,\ ...,1}_{m})$, які відповідають оптимальному базису задачі 1. При розв'язуванні задачі 1 модифікованим симплекс-методом матриця $B_\Im^{-1}$ і вектор $(\bar{c}_{i_1},\bar{c}_{i_2},...,\bar{c}_{i_m})$ відомі.

Двоїста до задачі 1 має вигляд:

знайти $\arg\max\limits_{(\mu_1,...,\mu_m)}\sum\limits_{i=1}^{m}a_i^0\mu_i$ при обмеженнях:

$$\sum_{i=1}^{m}\alpha_{ij}\mu_i\le 0 \text{ при } j\in\Im;$$

$$\mu_i\le 1,\quad i=\overline{1,m}.$$

V. Якщо нев'язки $\varepsilon_i^*,\ i=\overline{1,m}$ задовольняють рівності

$$\varepsilon_i^*=0,\quad i=\overline{1,m},$$

то обчислити вектор $x^*=(x_1^*,x_2^*,...,x_n^*)^T$ за правилом:

$$\left.\begin{array}{lll} x_j^*=\bar{x}_j^* & \text{при} & j\in\Im \\ x_j^*=0 & \text{при} & j\notin\Im \end{array}\right\},\quad j=\overline{1,n},$$

і зупинити обчислення (в цьому випадку знаходиться оптимальний розв'язок $x^*$ задачі 2 п. 5.7);
якщо при деякому $i\in[1:m]$ $\varepsilon_i^*>0$, то перейти на крок VI.

VI. Обчислити величини:

$$\delta_j^*=\sum_{i=1}^{m}\alpha_{ij}\mu_i^*,\quad j=\overline{1,n}.$$

VII. Якщо при всіх $j\in[1:n]$ виконується нерівність $\delta_j^*\le 0$, то припинити обчислення (в цьому випадку задача 2 п. 5.7 не має допустимих розв'язків); інакше перейти на крок VIII.

VIII. Обчислити величини $\Delta_j$ для тих $j$, для яких $\delta_j^*>0$:

$$\Delta_j=\sum_{i=1}^{m}\alpha_{ij}y_i-c_j\ .$$

IX. Обчислити відношення $\Delta_j / \delta_j^*$ для тих $j$, для яких $\delta_j^* > 0$, і мінімальне з цих відношень позначити через $\theta_0$.

X. Покласти $\overline{y} = y$.

XI. Перейти до нового допустимого розв'язку $y = (y_1, \ldots, y_m)^T$ двоїстої задачі 0 п. 5.8 за правилом:

$$y_i = \overline{y}_i - \theta_0 \mu_i^*, \quad i = \overline{1,m}.$$

XII. Перейти на крок II.

***Теорема 1.*** *Якщо виконані припущення 1 і розширена задача 1 невироджена, то процес розв'язування задачі 2 п. 5.7 алгоритмом 1 через скінченне число ітерацій закінчиться або на кроці V (знаходиться оптимальний розв'язок задачі 2 п. 5.7), або на кроці VII (встановлюється, що задача 2 п. 5.7 не має допустимих розв'язків).*

*Зауваження 1.* Алгоритм 1 можна застосовувати для розв'язування задачі 2 п. 5.7 і в тому випадку, коли розширена задача 1 вироджена. Для цього необхідно при розв'язуванні допоміжних задач модифікованим симплекс-методом користуватися правилом, яке гарантує від зациклювання.

*Зауваження 1'.* Початковий опорний розв'язок кожної наступної допоміжної задачі 1 слід складати з відповідних додатних компонент оптимального розв'язку попередньої допоміжної задачі 1, що суттєво скоротить число ітерацій, необхідних для розв'язання чергової допоміжної задачі 1.



**2. Метод двосторонніх оцінок**

Для розв'язування задач лінійного програмування методом двосторонніх оцінок попередньо необхідно знаходити (окрім від початкового допустимого розв'язку двоїстої задачі 0 п. 5.8) початковий опорний розв'язок задачі 2 п. 5.7. На кожній ітерації алгоритму обчислюється деяка оцінка (нев'язка $\varepsilon$) відхилення одержаного опорного розв'язку від оптимального. У методі двосторонніх оцінок необхідно на кожній ітерації розв'язувати допоміжну задачу лінійного програмування з відомим початковим опорним розв'язком та його базисом.

***Алгоритм 2***

П о ч а т о к . І. Знайти опорний розв'язок $\overline{x} = \left(\overline{x}_1, ..., \overline{x}_n\right)^T$ задачі 2 п. 5.7 та допустимий розв'язок $\overline{y} = \left(\overline{y}_1, ...., \overline{y}_m\right)^T$ спряженої задачі 0 п. 5.8.

II. Обчислити величини:

$$\Delta_j = \sum_{i=1}^{m} \alpha_{ij} \overline{y}_i - c_j, \quad j \in [1:n].$$

Основний цикл. III. Знайти всі індекси $j \in [1:n]$, для яких виконується $\Delta_j = 0$, і позначити множину таких індексів через $\Im_Y$.

IV. Позначити через $\Im_X$ множину індексів векторів базису опорного розв'язку $\bar{x}$.

V. Знайти множину

$$\Im = \Im_Y \bigcup \Im_X .$$

VI. Використовуючи модифікований симплекс-метод, знайти оптимальний розв'язок $\bar{x}^*$ наступної допоміжної задачі лінійного програмування.

Задача 2. Знайти $\arg\min\limits_{x_j} \sum\limits_{j\in\Im} \Delta_j x_j$ при обмеженнях:

$$\sum_{j\in\Im} a^j x_j = a^0;$$

$$x_j \geq 0, \quad j \in \Im.$$

VII. Знайти оптимальний розв'язок $\mu^* = \left(\mu_1^*, ..., \mu_m^*\right)$ задачі, двоїстої до задачі 2:

$$\mu^* = \left(\Delta_{i_1}, ..., \Delta_{i_m}\right) B_\Im^{-1},$$

де $B_\Im^{-1}$ – матриця, обернена до оптимальної базисної матриці задачі 2; $i_1, \ldots, i_m$ – номери базисних векторів оптимального розв'язку задачі 2.

VIII. Обчислити нев'язку

$$\varepsilon = \sum_{j\in\Im} \Delta_j \bar{x}_j^* .$$

Нев'язка $\varepsilon$ є різницею між значеннями лінійних форм двоїстої та прямої задач при даних $\bar{x}$ і $\bar{y}$. На практиці обчислення можна припиняти, коли $\varepsilon$ не перевищує наперед заданої точності обчислення мінімуму цільової функції.

IX. Якщо $\varepsilon = 0$, то обчислити оптимальний розв'язок $x^* = \left(x_1^*, ..., x_n^*\right)^T$ задачі 2 п. 5.7:

$$\left.\begin{aligned} x_j^* &= \bar{x}_j^* \quad \text{при} \quad j \in \Im \\ x_j^* &= 0 \quad \text{при} \quad j \notin \Im \end{aligned}\right\}, \quad j = \overline{1, n},$$

та припинити обчислення; інакше перейти на крок X.

X. Обчислити параметри:

$$\delta_j^* = \sum_{i=1}^m \alpha_{ij} \mu_i^*, \quad j = \overline{1, n}.$$

Якщо при усіх $j \in [1:n]$ виконується $\delta_j^* \le 0$, то припинити обчислення (у цьому випадку задача 2 п. 5.7 не має допустимих розв'язків); інакше перейти на крок XI.

XI. Обчислити відношення $\Delta_j / \delta_j^*$ для тих $j$, для яких $\delta_j^* > 0$ і мінімальне з цих відношень позначити через $\theta_0$.

XII. Перейти до нового опорного розв'язку $\bar{x} = \left(\bar{x}_1, ..., \bar{x}_n\right)^T$ задачі 2 п. 5.7:

$$\left.\begin{array}{lll} x_j^* = \bar{x}_j^* & \text{при} & j \in \Im \\ x_j^* = 0 & \text{при} & j \notin \Im \end{array}\right\}, \quad j = \overline{1, n},$$

XIII. Покласти $y = \bar{y}$; $\bar{\Delta}_j = \Delta_j$, $j = \overline{1, n}$.

XIV. Перейти до нового допустимого розв'язку $\bar{y} = \left(\bar{y}_1, ..., \bar{y}_m\right)^T$ двоїстої задачі 0 п. 5.8

$$\bar{y} = y - \theta_0 \mu^T *.$$

XV. Обчислити величини $\Delta_j$, $j \in [1:n]$, які відповідають новому допустимому розв'язку $\bar{y}$:

$$\Delta_j = \bar{\Delta}_j - \theta_0 \delta_j^*, \; j \in [1:n].$$

XVI. Перейти на крок III.

***Теорема 2.*** *Якщо виконані припущення 1, то у невиродженому випадку процес розв'язування задачі 2 п. 5.7 алгоритмом 2 через скінченне число ітерацій закінчиться або на кроці IX (знаходиться оптимальний розв'язок задачі 2 п. 5.7), або на кроці X (встановлюється, що задача 2 п. 5.7 не має допустимих розв'язків).*

Алгоритм 2 може застосовуватися і для виродженого випадку. Тоді при розв'язуванні допоміжної задачі 2 теоретично може виникнути зациклення. При зацикленні треба використовувати спеціальне правило, яке гарантує від зациклювання.

## 5.10. Методи блочного програмування

Методи блочного програмування дозволяють отримати розв'язок екстремальної задачі з допомогою розв'язування ряду допоміжних задач, кожна з яких в певному сенсі є більш простою, ніж початкова. В лінійному програмуванні методи блочного програмування застосовуються для розв'язування задач великого розмірності і задач з блочною структурою матриці обмежень.

**1. Метод розкладання Данцига–Вульфа**

З а д а ч а 1. Знайти $\arg\max\limits_{x\in X}(c,x)$ для заданого вектора $c=(c_1,c_2,...,c_n)$ і множини $X$, що задається співвідношеннями:

$$A^0x=b^0; \tag{5.44}$$

$$A^1x=b^1; \tag{5.45}$$

$$x\geq 0, \tag{5.46}$$

де $A^0$ – матриця розмірності $m\times n$; $A^1$ – матриця розмірності $m_1\times n$; $b^0=(b_1,...,b_m)$; $b^1=(b_{m+1},...,b_{m+m_1})$.

***1.1. Випадок обмеженості множини $X^1$.***

*Припущення 1.* $(\boldsymbol{i})-$ множина $X^1$, що визначається умовами (5.45) – (5.46) є обмеженою; $(\boldsymbol{ii})-$ $\operatorname{rang}\begin{pmatrix}A^0\\A^1\end{pmatrix}=m+m_1$; $(\boldsymbol{iii})-$ $m+m_1<n.$

В методах розкладання початкова задача 1 розмірності $(m+m_1)\times n$ зводиться до розв'язування наступної задачі лінійного програмування розмірності $(m+1)\times l$.

***z*** ***-задача.*** Знайти $\arg\max\limits_{z}\sum\limits_{i=1}^{l}\sigma_i z_i$ при обмеженнях:

$$\sum_{i=1}^{l}p^i z_i=b^0; \tag{5.47}$$

$$\sum_{i=1}^{l}z_i=1; \tag{5.48}$$

$$z_i\geq 0,\quad i=\overline{1,l}, \tag{5.49}$$

де $x^1,x^2,...x^l$ – вершини многогранника $X^1$;

$$\sigma_i=(c,x^i),\ \ i=\overline{1,l};\quad p^i=A^0x^i,\quad i=\overline{1,l}. \tag{5.50}$$

Оптимальний розв'язок $(z_1^*,z_2^*,...,z_l^*)$ $z$-задачі однозначно визначає оптимальний розв'язок $x^*$ задачі 1:

$$x^*=\sum_{i=1}^{l}z_i^*x^i.$$

Для розв'язування $z$-задачі не обов'язково знати всі вершини множини $X^1$ (тобто вектори $x^1,x^2,...,x^l$). Необхідно лише знати вершини, що входять в початковий опорний базис $z$-задачі. В процесі розв'язування $z$-задачі модифікованим симплекс-методом перевірка на оптимальність опорного розв'язку $z$-задачі проводиться з допомогою розв'язування деякої допоміжної задачі лінійного програмування з обмеженнями (5.45) – (5.46).

***Алгоритм 1***

Початок. І. Знайти початковий опорний розв'язок $z^1 = (z_l^1, ..., z_l^1)$ $z$-задачі, базис $\begin{pmatrix} p^{i_1} \\ 1 \end{pmatrix}, ..., \begin{pmatrix} p^{i_{m+1}} \\ 1 \end{pmatrix}$ цього опорного розв'язку та коефіцієнти $\sigma_{i_k}$, $k = \overline{1, m+1}$ лінійної форми $z$-задачі, що відповідають початковому опорному розв'язку.

ІІ. Обчислити матрицю $B^{-1}$, обернену до матриці, що складена з базисних векторів $\begin{pmatrix} p^{i_1} \\ 1 \end{pmatrix}, ..., \begin{pmatrix} p^{i_{m+1}} \\ 1 \end{pmatrix}$:

$$B^{-1} \triangleq (\beta_{ij})_{i=\overline{1,m+1}}^{j=\overline{1,m+1}}.$$

Основний цикл. ІІІ. Покласти $\sigma_{баз} = (\sigma_{i_1}, \sigma_{i_2}, ..., \sigma_{i_{m+1}})$.

IV. Обчислити вектор $\overline{\Lambda} \triangleq (\Lambda, \lambda) = (\lambda_1, \lambda_2, ..., \lambda_m, \lambda)$ за формулою

$$\overline{\Lambda} = \sigma_{баз} B^{-1}.$$

V. Покласти $y_{k0} = z_{i_k}^1$, $k = \overline{1, m+1}$.

VI. Обчислити $n$-вимірний вектор $c^\Lambda = (c_1^\Lambda, c_2^\Lambda, ..., c_n^\Lambda)$ за формулою

$$c^\Lambda = c - \Lambda A^0$$

(в звичайній формі запису координати вектора $c^\Lambda$ обчислюються за формулою

$$c_j^\Lambda = c_j - \sum_{i=1}^{n} \lambda_i \alpha_{ij}^0, \qquad j = \overline{1, n}).$$

VII. Обчислити вектор $x^\nu$ – оптимальний розв'язок наступної задачі лінійного програмування в просторі $R^n$. Підзадача $X_\Lambda$:

знайти $\arg\max_x (c^\Lambda, x)$ при обмеженнях:

$$A^1 x = b^1;$$
$$x \geq 0.$$

Для розв'язування підзадачі $X_\Lambda$ зручно застосовувати симплекс-метод або модифікований симплекс-метод. Причому, на кожній наступній ітерації за початковий опорний розв'язок та його базис можна брати оптимальний розв'язок та його базис, що отримані на попередній ітерації.

VIII. Якщо виконується рівність $(c^\Lambda, x^\nu) = \lambda$, то покласти $z^* = z^1$ та перейти на крок IX (тобто, в цьому випадку опорний розв'язок $z^1 = (z_1^1, ..., z_l^1)$ є і оптимальним розв'язком $z$-задачі); інакше перейти на крок X.

IX. Обчислити вектор $x^*$ – оптимальний розв'язок задачі 1:

$$x^* = \sum_{k=1}^{m+1} z_{i_k}^* x^{i_k}$$

та припинити обчислення.

X. Обчислити вектор $\overline{p}^{\nu} = \begin{pmatrix} p^{\nu} \\ 1 \end{pmatrix}$, де $p^{\nu} = A^0 x^{\nu}$.

XI. Обчислити вектор (тобто, розкласти вектор $\overline{p}^{\nu}$ по базисних векторах)

$$(y_{1\nu}, y_{2\nu}, ..., y_{m+1,\nu})^{\mathrm{T}} = B^{-1}\overline{p}^{\nu}.$$

XII. Обчислити коефіцієнт лінійної форми $z$-задачі

$$\sigma_{\nu} = (c, x^{\nu}).$$

XIII. Обчислити відношення $y_{k0} / y_{k\nu}$ для тих $k \in [1:(m+1)]$, для яких $y_{k\nu} > 0$ та позначити мінімальне з цих відношень через $\theta_0$.

XIV. Знайти індекс $r \in [1:(m+1)]$ такий, що $y_{r\nu} > 0$ і $y_{r0} / y_{r\nu} = \theta_0$.

XV. Покласти $\overline{\beta}_{ij} = \beta_{ij}$, $i = \overline{1, m+1}$; $j = \overline{1, m+1}$ і $\overline{y}_{k0} = y_{k0}$, $k = \overline{1, m+1}$.

Перейти до нового базису $\begin{pmatrix} p^{i_1} \\ 1 \end{pmatrix}, \ldots, \begin{pmatrix} p^{i_r} \\ 1 \end{pmatrix}, \ldots, \begin{pmatrix} p^{i_{m+1}} \\ 1 \end{pmatrix}$, який отримуємо заміною вектора $\overline{p}^{i_r} = \begin{pmatrix} p^{i_r} \\ 1 \end{pmatrix}$ в попередньому базисі вектором $\overline{p}^{\nu} = \begin{pmatrix} p^{\nu} \\ 1 \end{pmatrix}$, (тобто покласти $i_r = \nu$).

XVI. Обчислити за основними формулами матрицю $B^{-1} \triangleq (\beta_{ij})_{i=\overline{1,m+1}}^{j=\overline{1,m+1}}$, що є оберненою до нової базисної матриці:

$$\beta_{rj} = \overline{\beta}_{rj} / y_{r\nu}, \quad j = \overline{1, m+1};$$

$$\beta_{ij} = \overline{\beta}_{ij} - \left(\overline{\beta}_{rj} / \overline{\beta}_{r\nu}\right) y_{i\nu}, \;\; i = \overline{1, m+1}, \;\; j = \overline{1, m+1}; \;\; i \neq r.$$

XVII. Обчислити за основними формулами нові значення базисних координат опорного розв'язку:

$$y_{r0} = \overline{y}_{r0} / y_{r\nu};$$

$$y_{k0} = \overline{y}_{k0} - \left(\overline{y}_{r0} / y_{r\nu}\right) y_{k\nu}, \quad k = 1, \ldots, r-1, r+1, \ldots, m+1.$$

XVIII. Перейти до нового опорного розв'язку $z$-задачі $z^1 = (z_1^1, ..., z_l^1)$:

$$z_{i_k}^1 = y_{k0}, \;\; k = \overline{1, m+1},$$

інші координати вектора $z^1$ дорівнюють нулю.

XIX. Перейти на крок III.

***Теорема 1.*** *Якщо допустима множина розв'язків X задачі 1 не порожня, то при виконанні припущень 1 алгоритм 1 за скінченне число* *ітерацій призводить до оптимального розв’язку задачі 1 (припускається, що якщо в виродженому випадку виникає зациклювання, то по аналогії з пунктом 4 п. 5.7 застосовується спеціальне правило виходу з циклу).*

***1.2. Випадок необмеженості множини $X^1$.***

*Припущення 1′.* $(\boldsymbol{i})-$ множина $X^1$, яка визначається умовами (5.45) – (5.46) є необмеженою; $(\boldsymbol{ii})-$ $\operatorname{rang}\begin{pmatrix} A^0 \\ A^1 \end{pmatrix} = m + m_1$; $(\boldsymbol{iii})-$ $m + m_1 < n$.

Позначимо через $x^i$, $i = \overline{1, l_1}$, вершини множини $X^1$, а направляючі вектори необмежених ребер $X^1$ через $x^{l_1+1}$, $x^{l_1+2}$, ..., $x^l$.

Якщо виконуються припущення 1′, то початкова задача 1 зводиться до розв’язання $z$-задачі, в якій умову (5.48) необхідно замінити на:

$$\sum_{i=1}^{l} \varepsilon_i z_i = 1,$$

де $\varepsilon_i = 1$ при $i = \overline{1, l_1}$; $\varepsilon_i = 0$ при $i = \overline{l_1 + 1, l}$. До зміненої $z$-задачі можна застосувати алгоритм 1 із наступними змінами: оскільки на кроці VII алгоритму 1 підзадача $X_\Lambda$ не завжди є розв’язувана (цільова функція може бути необмеженою на множині $X^1$), то можливі два випадки:

1°) оптимальний розв’язок $x^\nu$ під задачі $X_\Lambda$ існує;

2°) при розв’язуванні підзадачі $X_\Lambda$ є оцінка $\Delta_j < 0$, для якої всі $u_{kj} \le 0$, $k = \overline{1, m_1}$ (тобто цільова функція підзадачі $X_\Lambda$ необмежена зверху).

У випадку 1° кроки алгоритму 1 залишаться без змін. У випадку 2° кроки I–VI залишаться без змін, а на кроці VII алгоритму 1 в якості вектора $x^\nu$ вибирається $n$-вимірний вектор, у якого координати з номерами $j_1, j_2, ..., j_{m_1}$ дорівнюють, відповідно, $-u_{1j}, -u_{2j}, ..., -u_{m_1 j}$, $j$-та координата дорівнює одиниці, а інші координати нульові (тут $j_1, j_2, ..., j_{m_1}$ – номери базисних векторів підзадачі $X_\Lambda$) і здійснюється перехід на крок X, на якому обчислюється вектор $\overline{p}^\nu$ за зміненою формулою:

$$\overline{p}^\nu = \begin{pmatrix} p^\nu \\ 0 \end{pmatrix},$$

де $p^\nu = A^0 x^\nu$.

У випадку 2° вектор $x^\nu$ є направляючим вектором необмеженого ребра многогранника $X^1$ (тут $u_{kj}$, $k = \overline{1, m_1}$ – коефіцієнти розкладу $j$-го вектора матриці обмежень в підзадачі $X_\Lambda$).

Крок XI алгоритму 1 потрібно доповнити перевіркою на необмеженість цільової функції $z$-задачі.

XI. Обчислити вектор

$$(y_{1\nu},..., y_{m+1,\nu})^{\mathrm{T}} = \mathrm{B}^{-1}\overline{p}^{\nu}.$$

Якщо при всіх $k \in [1:(m+1)]$ виконується $y_{k\nu} \le 0$, то припинити обчислення (в цьому випадку цільова функція $z$-задачі (а значить, і задачі 1) необмежена на допустимій множині); інакше перейти на крок XII.

Інші кроки алгоритму 1 залишаються без змін.

*Зауваження 1.* Принцип розкладання, в свою чергу, можна застосувати і для розв'язування допоміжної задачі на кроці VII алгоритму 1 і т.д. Тому розв'язування задач лінійного програмування з будь-яким числом обмежень в принципі можна звести до розв'язування ряду задач лінійного програмування, кількість обмежень в кожній з яких не перевищує заданого числа $\overline{m}$.

***Теорема 1'.*** *Якщо допустима множина розв'язків X задачі 1 не порожня, то при виконанні припущення 1' процес розв'язування задачі 1 зміненим алгоритмом 1 через скінченне число ітерацій або закінчиться на кроці IX (знаходиться оптимальний розв'язок задачі 1), або на кроці XI (встановлюється, що цільова функція задачі 1 необмежена на допустимій множині).*



**2. Метод розкладання для розв'язування задач лінійного програмування з блочно-діагональною матрицею**

Метод розкладання, що описаний в попередньому пункті, особливо ефективний при спеціальній блочно-діагональній структурі матриці $A^1$ в задачі 1:

$$A^1 = \begin{pmatrix} A_1 & 0 & ... & 0 \\ 0 & A_2 & ... & 0 \\ . & . & . & . \\ 0 & 0 & ... & A_s \end{pmatrix},$$

де $A_t = (\alpha_{ij}^t)$, $t = \overline{1,s}$ – матриці розмірності $m_t \times n_t$. Початкова задача лінійного програмування записується у вигляді.

З а д а ч а 2. Знайти $\arg\max\limits_x \sum\limits_{t=1}^{s}(c^t, x^t)$ при обмеженнях:

$$\sum_{t=1}^{s} A_t^0 x^t = b^0, \tag{5.51}$$

$$A_t x^t = b^t, \qquad t = \overline{1,s}, \tag{5.52}$$

$$x^t \ge 0, \qquad t = \overline{1,s}, \tag{5.53}$$

де $x^t - n_t$-вимірний вектор-стовпець; $A_t^0 = (\alpha_{ij}^{t0})$ – матриця розмірності $m \times n_t$; $A_t = (\alpha_{ij}^t)$ – матриця розмірності $m_t \times n_t$; $b^0 - m$-вимірний вектор-стовпець; $b^t - m_t$-вимірний вектор-стовпець; $c^t - n_t$-вимірний вектор-рядок; $x = (x^1, x^2, ..., x^s)^T$ .

Розв’язування задачі 2 зводиться до розв’язування наступної задачі лінійного програмування розмірності $(m+s) \times \hat{l}$ $(\hat{l} = \sum_{t=1}^{s} l_t)$:

$$z_1\text{-задача. Знайти } \arg\max_{z^t} \sum_{t=1}^{s} \sum_{i=1}^{l_t} \sigma_i^t z_i^t$$

при обмеженнях:

$$\sum_{t=1}^{s} \sum_{i=1}^{l_t} p_{(i)}^t z_i^t = b^0 ; \tag{5.54}$$

$$\sum_{i=1}^{l_t} \varepsilon_i^t z_i^t = 1 , \quad t = \overline{1, s} ; \tag{5.55}$$

$$z_i^t \geq 0 , \ i = \overline{1, l_t} ; \qquad t = \overline{1, s} , \tag{5.56}$$

де $\sigma_i^t = (c^t, x_{(i)}^t)$; $p_{(i)}^t = A_t^0 x_{(i)}^t$; $x_{(i)}^t$, $i = \overline{1, l_t}$ при кожному $t \in [1:s]$ відповідає вершинам та направляючим векторам необмежених ребер багатогранної множини $G_t$, $G_t \triangleq \left\{ x^t \,\middle|\, A_t x^t = b^t, \ x^t \geq 0 \right\}$, (припускається, що при $i \in [1:l_t^1]$, $x_{(i)}^t$ відповідає вершині, а при $i \in [(l_t^1 + 1):l_t]$ – направляючим векторам необмежених ребер $G_t$);

$$\varepsilon_i^t = \begin{cases} 1, \text{ якщо } x_{(i)}^t \text{ відповідає вершині множини } G_t, \\ 0, \text{ якщо } x_{(i)}^t \text{ відповідає нескінченому ребру множини } G_t. \end{cases}$$

Для зручності запису обмеження (5.54) і (5.55) перепишемо у вигляді:

$$\sum_{t=1}^{s} \sum_{i=1}^{l_t} \overline{p}_{(i)}^t z_i^t = \overline{b}^0 ,$$

де

$$\overline{p}_{(i)}^t = \begin{pmatrix} p_{(i)}^t \\ \varepsilon_i^t e^t \end{pmatrix}, \ \overline{b}^0 = \begin{pmatrix} b_0 \\ I_s \end{pmatrix},$$

тут $e^t - s$-вимірний вектор з одиницею на $t$-му місці; $I_s -$ $s$-вимірний вектор з одиничними компонентами.

Оптимальний розв’язок $z_1$-задачі однозначно визначає оптимальний розв’язок задачі 2, тобто, якщо числа $z_i^{*t}$ визначають розв’язок $z_1$-задачі, то вектори:

$$x^{*t} = \sum_{i=1}^{l_t} z^{*t}_{\ i} x^t_{(i)}, \quad t = \overline{1, s}, \tag{5.57}$$

утворюють оптимальний розв’язок

$$x^* = (x^{*1}, x^{*2}, ..., x^{*s})^T \tag{5.58}$$

задачі 2.

Для розв’язування $z_1$-задачі немає необхідності знати завчасно всі вершини багатогранників $G_t$, $t = \overline{1, s}$ (необхідна лише інформація про вершини, що входять в початковий базис). $z_1$-задача розпадається на $s$ (по числу блоків) підзадач розмірності $m_t \times n_t$, $t = \overline{1, s}$, кожна з яких розв’язується модифікованим симплекс-методом. Оптимальні значення лінійних форм цих підзадач використовують для перевірки опорного розв’язку $z_1$-задачі на оптимальність і для відновлення базису $z_1$-задачі, якщо опорний розв’язок не оптимальний.

***Алгоритм 2***

Початок 1. Знайти початковий базис $z_1$-задачі:

$$\overline{p}^{t_1}_{(i_1)}, \overline{p}^{t_2}_{(i_2)}, ..., \overline{p}^{t_{m+s}}_{(i_{m+s})},$$

базисні змінні опорного розв’язку $z_1$-задачі, що відповідають цьому базису:

$$z^{t_1}_{i_1}, z^{t_2}_{i_2}, ..., z^{t_{m+s}}_{i_{m+s}},$$

та коефіцієнти лінійної форми $z_1$-задачі, що відповідають початковому базису:

$$\sigma^{t_1}_{i_1}, \sigma^{t_2}_{i_2}, ..., \sigma^{t_{m+s}}_{i_{m+s}}.$$

II. Обчислити матрицю $B^{-1}$, обернену до матриці, що складена з базисних векторів $\overline{p}^{t_k}_{(i_k)}$, $k = \overline{1, m+s}$:

$$B^{-1} \triangleq \left(\beta_{ij}\right)^{j=\overline{1,m+s}}_{i=\overline{1,m+s}}.$$

Основний цикл. III. Покласти $\sigma_{баз} = (\sigma^{t_1}_{i_1}, ..., \sigma^{t_{m+s}}_{i_{m+s}})$.

IV. Обчислити вектор $\overline{\Lambda} \triangleq (\Lambda^1, \Lambda^2) = (\lambda^1_1, ...\lambda^1_m, \lambda^2_1, ..., \lambda^2_s)$ за формулою

$$\overline{\Lambda} = \sigma_{баз} B^{-1}.$$

V. Покласти

$$y_{k0} = z^{t_k}_{i_k}, \quad k = \overline{1, m+s}.$$

VI. Покласти $t = 1$.

VII. Розв’язати наступну задачу лінійного програмування розмірності $m_t \times n_t$ ($X^{(t)}_{\Lambda}$-задачу):

знайти $\arg\max\limits_{x^t} (c^t - \Lambda^1 A^0_t, x^t)$ при обмеженнях:

$$A_t x^t = b^t; \qquad x^t \geq 0$$

(при цьому вважаються відомими вершини $x_{i_1}^{t_1}, x_{i_2}^{t_2}, ..., x_{i_{m+s}}^{t_{m+s}}$).

Процес розв'язування $X_{\Lambda}^{(t)}$-задачі може закінчитися одним з двох виходів: вихід $1^0$ – $X_{\Lambda}^{(t)}$-задача розв'язувана; вихід $2^0$ – лінійна форма $X_{\Lambda}^{(t)}$-задачі необмежена на множині $G_t$). Якщо $X_{\Lambda}^{(t)}$-задача розв'язується, то позначити її розв'язок через $x_{(\nu_t)}^t$ та перейти на крок VIII; якщо $X_{\Lambda}^{(t)}$-задача не розв'язується (тобто лінійна форма $X_{\Lambda}^{(t)}$-задачі необмежена на множині $G_t$), то позначити через $x_{(\nu_t)}^t$ направляючий вектор необмеженого ребра множини $G_t$, що породжує вектор $\overline{p}_{(\nu_t)}^t$ з від'ємною оцінкою $\Delta_{\nu_t}^t$ та перейти на крок IX.

(Правило вибору такого $x_{(\nu_t)}^t$ описано в пункті 1).

VIII. Обчислити оцінку

$$\Delta_{\nu_t}^t = -(c^t - \Lambda^1 A_t^0, x_{(\nu_t)}^t) + \lambda_t^2$$

та перейти на крок X.

IX. Обчислити оцінку $\Delta_{\nu_t}^t$ за формулою

$$\Delta_{\nu_t}^t = -(c^t - \Lambda^1 A_t^0, x_{(\nu_t)}^t)$$

та перейти на крок X.

X. Якщо $t < s$, то покласти $t = t+1$ та перейти на крок VII; інакше перейти на крок XI.

XI. Якщо виконується умова

$$\min_{t \in [1:s]} \Delta_{\nu_t}^t = 0, \tag{5.59}$$

то опорний розв'язок та базис $z_1$-задачі оптимальні; в цьому випадку знайти оптимальний розв'язок початкової задачі 2 по (5.57) – (5.58), де $z_i^{*t}$ дорівнюють відповідним базисним змінним $z_1$-задачі при $t = t_k$, $i = i_k$, $k = \overline{1, m+s}$ та $z_i^{*t} = 0$ у протилежному випадку (вектори $x_i^t$, що відповідають номерам базисних змінних, теж відомі) та припинити обчислення. Якщо умова (5.59) не виконується, то перейти на крок XII.

На наступних кроках алгоритму здійснюється перехід до нового базису $z_1$-задачі.

XII. Обчислити індекс $\mu \in [1:s]$, для якого виконується

$$\Delta_{\nu_\mu}^{\mu} = \min_{t \in [1:s]} \Delta_{\nu_t}^t .$$

XIII. Обчислити $(m+s)$-вимірний вектор $\overline{p}^{\mu}_{(\nu_\mu)}$ та відповідний коефіцієнт лінійної форми $z_1$-задачі:

$$\overline{p}^{\mu}_{(\nu_\mu)} = \begin{pmatrix} A^0_\mu \cdot x^\mu_{(\nu_\mu)} \\ \varepsilon^\mu_{\nu_\mu} \cdot e^\mu \end{pmatrix}; \quad \sigma^\mu_{\nu_\mu} = (c^\mu, x^\mu_{(\nu_\mu)}) .$$

XIV. Обчислити вектор (тобто обчислити коефіцієнти розкладу вектора $\overline{p}^{\mu}_{(\nu_\mu)}$ по базисних векторах)

$$(y_{1\mu}, ..., y_{m+s,\mu})^T = B^{-1}\overline{p}^{\mu}_{(\nu_\mu)} .$$

XV. Обчислити індекс $r \in [1:(m+s)]$, який задовольняє умовам:

$$y_{r0} / y_{r\mu} = \min_{y_{k\mu}>0}(y_{k0} / y_{k\mu}), \quad y_{r\mu} > 0 .$$

XVI. Покласти $\overline{\beta}_{ij} = \beta_{ij},\ i = \overline{1, m+s};\ \ j = \overline{1, m+s}$, та $\overline{y}_{k0} = y_{k0}$, $k = \overline{1, m+s}$.

XVII. Перейти до нового базису $\overline{p}^{t_1}_{(i_1)}, ..., \overline{p}^{t_{m+s}}_{(i_{m+s})}$, який отримується заміною вектора $\overline{p}^{t_r}_{(i_r)}$ в попередньому базисі вектором $\overline{p}^{\mu}_{(\nu_\mu)}$ (тобто покласти $t_r = \mu,\ i_r = \nu_\mu$).

Відмітимо, що при переході до нового базису змінюється також коефіцієнт лінійної форми $\sigma^{t_r}_{i_r}$ та точка $x^{t_r}_{i_r}$.

XVIII. Обчислити за основними формулами матрицю $B^{-1}$, обернену до нової базисної матриці:

$$\beta_{rj} = \overline{\beta}_{rj} / y_{r\mu}, \quad j = \overline{1, m+s} ;$$

$$\beta_{ij} = \overline{\beta}_{ij} - (\overline{\beta}_{rj} / \overline{\beta}_{r\mu}) y_{i\mu}, \quad i = \overline{1, m+s} ; \quad j = \overline{1, m+s} ; \quad i \neq r .$$

XIX. Обчислити по основних формулах нові значення базисних координат опорного розв'язку $z_1$-задачі:

$$y_{r0} = \overline{y}_{r0} / y_{r\mu} ;$$

$$y_{k0} = \overline{y}_{k0} - (\overline{y}_{r0} / y_{r\mu}) y_{k\mu}, \ \ k = 1, ..., r-1, r+1, ..., m+s .$$

XX. Покласти $z^{t_k}_{i_k} = y_{k0}$, $k = \overline{1, m+s}$, та перейти на крок III.

***Теорема 2.*** *Якщо допустима множина розв'язків задачі 2 непорожня, рівняння-обмеження задачі 2 лінійно незалежні і цільова функція задачі 2 обмежена на допустимій множині розв'язків, то за скінченне число ітерацій алгоритм 2 призводить до оптимального розв'язку задачі 2 (при цьому припускається, що якщо у виродженому випадку виникає зациклювання, то, по аналогії з пунктом 4 п. 5.7, застосовується спеціальне правило виходу з циклу).*

**3. Метод, який використовує узагальнений градієнтний спуск**

З а д а ч а 3. Знайти $\arg\max_x (c, x)$ для заданого вектора $c \in R^n$ при обмеженнях:

$$\sum_{j=1}^{n} \alpha_{ij}^1 x_j \le b_i^1, \; i = \overline{1, m_1}; \tag{5.60}$$

$$\sum_{j=1}^{n} \alpha_{kj}^2 x_j \le b_k^2, \quad k = \overline{1, m_2}; \tag{5.61}$$

$$x_j \ge 0, \quad j = \overline{1, n}. \tag{5.62}$$

Задача двоїста до задачі 3, має вигляд.

З а д а ч а 3 $'$. Знайти $\arg\min_{(y,z)} (\sum_{i=1}^{m_1} b_i^1 y_i + \sum_{k=1}^{m_2} b_k^2 z_k)$ при обмеженнях:

$$\sum_{i=1}^{m_1} \alpha_{ij}^1 y_i + \sum_{k=1}^{m_2} \alpha_{kj}^2 z_k \ge c_j, \quad j = \overline{1, n};$$

$$y_i \ge 0, \; i = \overline{1, m_1}; \quad z_k \ge 0, \; k = \overline{1, m_2},$$

тут $y = (y_1, ..., y_{m_1})$, $z = (z_1, ..., z_{m_2})$.

*Припущення 3.* $(\boldsymbol{i})-$ задача 3 має оптимальний розв'язок; $(\boldsymbol{ii})-$ система обмежень (5.61) та (5.62) вирізає у просторі обмежень багатогранник $X_1$.

Метод розв'язування задачі 3$'$, що приводиться нижче, зводиться до мінімізації (методом узагальненого градієнтного спуску) опуклої кусково-лінійної функції

$$\varphi^*(y) \triangleq \max_{x \in X_1} \left( \sum_{j=1}^{n} c_j x_j + \sum_{i=1}^{m_1} y_i \left( b_i^1 - \sum_{j=1}^{n} \alpha_{ij}^1 x_j \right) \right)$$

при простих обмеженнях $y_i \ge 0$, $i = \overline{1, m_1}$. На кожній ітерації алгоритму потрібно розв'язувати задачу лінійного програмування з обмеженнями (5.61) та (5.62).

Метод узагальненого градієнтного спуску зручно застосовувати до таких задач лінійного програмування, двоїсті до яких мають блочну структуру.

***Алгоритм 3***

П о ч а т о к. І. Покласти $y^1 = \underbrace{(0, 0, ..., 0)}_{m_1}$.

ІІ. Покласти $l = 1$.

О с н о в н и й ц и к л. ІІІ. Покласти $y = y^l$.

IV. Розв'язати задачу лінійного програмування:

$$\text{знайти } \arg\max_x \left( \sum_{j=1}^{n} c_j x_j + \sum_{i=1}^{m_1} y_i \left( b_i^1 - \sum_{j=1}^{n} \alpha_{ij}^1 x_j \right) \right)$$

при обмеженнях (5.61) та (5.62);

позначити її розв'язок через $x^*(y^l)$, а через $z(y^l)$ – вектор оптимального розв'язку двоїстої задачі.

V. Обчислити вектор $h^l = (h_1^l, ..., h_{m_1}^l)$,

$$h_i^l = b_i^1 - \sum_{j=1}^{n} \alpha_{ij}^1 x_j^*(y^l), \quad i = \overline{1, m_1}.$$

VI. Обчислити кроковий множник $\rho_l$.

VII. Обчислити вектор

$$y^{l+1} = \pi^+(y^l - \rho_l h^l \left\| h^l \right\|^{-1}),$$

де $\pi^+$ – оператор, що залишає невід'ємні координати без змін і перетворює від'ємні координати в нуль.

VIII. Покласти $l = l + 1$ та перейти на крок III.

***Теорема 3.*** *Якщо виконуються припущення 3 та крокові множники задовольняють умовам:*

$$\rho_l > 0, \quad l = 1,2,...; \quad \rho_l \underset{l \to \infty}{\to} 0; \quad \sum_{l=1}^{\infty} \rho_l = \infty,$$

*то послідовність $\{\varphi^*(y^l)\}_{l=1}^{\infty}$ збігається до оптимального значення функції $\varphi^*(y)$ на множині $y \geq 0$, при цьому пара векторів $(y^l, z(y^l))$ збігається до оптимального розв'язку двоїстої задачі 3'.*



## 5.11. Модифікований симплекс-метод для розв'язування задачі лінійного програмування з двосторонніми обмеженнями

З а д а ч а 0. Знайти $\arg\max_{x \in X} \sum_{j=1}^{n} c_j x_j$ для заданого вектора $c = (c_1, \ldots, c_n)$ і заданої множини обмежень:

$$X \triangleq \left\{ x \middle| Ax = b, \ \alpha_j \leq x_j \leq \beta_j, \ j = \overline{1,n}; \ x \in R^n \right\},$$

де $A$ – $m \times n$-матриця; $b$ – $n$-вимірний вектор; $\alpha_j$, $\beta_j$, $j \in [1:n]$ – деякі числа.

*Припущення 0.* $(\boldsymbol{i})$ – ранг матриці $A \triangleq (\alpha_{ij})_{i=\overline{1,m}}^{j=\overline{1,n}} = (a^1, \ldots, a^n)$ дорівнює $m$; $(\boldsymbol{ii})$ – $n > m$; $(\boldsymbol{iii})$ – множина обмежень $X$ – непорожня.

*Означення 1.* Допустимий розв'язок $\overline{x} = (\overline{x}_1, \ldots, \overline{x}_n)^T$ задачі 0 називається ***опорним***, якщо система векторів $a^j$, які відповідають компонентам $\overline{x}_j$, для яких

$$\alpha_j < \overline{x}_j < \beta_j, \tag{5.63}$$

лінійно незалежна.

*Означення 2.* Система $m$ лінійно незалежних векторів, яка містить сукупність всіх векторів, яким відповідають компоненти $\overline{x}_j$ опорного розв'язку $\overline{x}$, які задовольняють (5.63), називається ***базисом опорного розв'язку*** $\overline{x}$.

*Означення 3.* Опорний розв'язок $\overline{x}$ задачі 0 називається ***невиродженим***, якщо $m$ його компонент задовольняють нерівностям (5.63). Задача 0 називається ***невиродженою***, якщо всі її опорні розв'язки невироджені.

Нижче наводиться модифікований симплекс-метод розв'язування задачі лінійного програмування із двосторонніми обмеженнями (задачі 0). Для початку роботи алгоритму потрібно мати опорний розв'язок і його базис. Метод приводить до розв'язку задачі 0 (або встановлює необмеженість лінійної форми на множині $X$) за число ітерацій, яке не перевищує $N = C_n^m 2^{n-m}$ (ця оцінка справедлива для невиродженої задачі 0).

Якщо алгоритм 1 застосовувати для розв'язування виродженої задачі 0, то (теоретично) можливе зациклювання (тобто повернення до вже пройденого базису). Якщо при розв'язуванні задачі 0 алгоритмом 1 виникає зациклювання, варто застосовувати процедуру вибору вектора $a^{i_r}$, який виводиться з базису, що виключає зациклювання. У початковій стадії алгоритму потрібно обчислювати матрицю, обернену до початкової базисної.

#### 1. Основний метод

***Алгоритм 1***

Початок. I. Знайти опорний розв'язок $\overline{x} = (\overline{x}_1, \ldots, \overline{x}_n)^T$ задачі 0 і його базис $a^{i_1}, \ldots, a^{i_m}$ (точніше множину $\Im_{\overline{x}}$, що складається з індексів $i_1, \ldots, i_m$, де $i_k \in [1:n]$ при $k = \overline{1,m}$).

II. Покласти $z_{k0} = \overline{x}_{i_k}$, $k = \overline{1,m}$.

Числа $z_{k0}$, $k = \overline{1,m}$ є коефіцієнтами розкладу вектора $a^0 = b - \sum\limits_{j \in \Im_{\overline{x}}} \overline{x}_j a^j$ по початковому базису (тут $\Im_{\overline{x}}$ – множина, яка складається з базисних індексів $i_1, \ldots, i_m$).

III. Обчислити матрицю $B^{-1}$, обернену до початкової базисної матриці $B$, яка складається з векторів $a^{i_1},\dots,a^{i_m}$.

IV. Розкласти по базису $a^{i_1},\dots,a^{i_m}$ вектори $a^j$, $j=\overline{1,n}$ (тобто знайти числа $z_{kj}$ такі, що $a^j=\sum_{k=1}^{m} z_{kj}a^{i_k}$, $j=\overline{1,n}$) за правилом:

$$(z_{1j},\dots,z_{mj})^T=B^{-1}(\alpha_{1j},\dots,\alpha_{mj})^T,\ j=\overline{1,n}.$$

Заздалегідь відомий розклад базисних векторів $a^{i_k}$, $k=\overline{1,m}$: $z_{li_k}=0$ при $k\neq l$ і $z_{li_k}=1$ при $k=l$ $\left(k=\overline{1,m};\ l=\overline{1,m}\right)$. Заздалегідь також відомо, що для базисних векторів $\Delta_j=0$, тобто, при $j\in\Im_{\bar{x}}$, $\Delta_j=0$.

V. Обчислити оцінки:

$$\Delta_j=\begin{cases} y_j, & \text{якщо } \alpha_j\le x_j<\beta_j;\\ -y_j, & \text{якщо } x_j=\beta_j,\end{cases}$$

де $y_j=\sum_{k=1}^{m} c_{i_k}z_{kj}-c_j$, $j=\overline{1,n}$.

О с н о в н и й  ц и к л. VI. Якщо всі $\Delta_j\ge 0$ $(j=\overline{1,n})$, то покласти $x^*=\bar{x}$ і припинити обчислення (у цьому випадку опорний розв'язок $\bar{x}$ є оптимальним розв'язком задачі 0); інакше (тобто, якщо при якомусь $j\in[1:n]$ виконується $\Delta_j<0$) перейти на крок VII.

VII. Якщо існує індекс $j\in[1:n]\setminus\Im_{\bar{x}}$, для якого $\Delta_j<0$ і виконується одна з умов:

(***i***) – при $\bar{x}_j=\alpha_j$ для $z_{kj}>0$ – $\alpha_{i_k}=-\infty$, для $z_{kj}<0$ – $\beta_{i_k}=\infty$;

(***ii***) – при $\bar{x}_j=\beta_j$ для $z_{kj}>0$ – $\beta_{i_k}=\infty$, для $z_{kj}<0$ – $\alpha_{i_k}=-\infty$,

то припинити обчислення (у цьому випадку функція цілі задачі 0 необмежена на допустимій множині $X$); інакше перейти на крок VIII.

VIII. Знайти індекс $s\in[1:n]$ такий, що $\Delta_s=\min\limits_{j\in[1:n]}\Delta_j$ і покласти $a^{i_{m+1}}=a^s$ (тобто одержати розширений базис $a^{i_1},\dots,a^{i_m},a^{i_{m+1}}$).

Індекс $s$ належить множині $[1:n]\setminus\Im_{\bar{x}}$ і визначає вектор $a^s$, який може включатися в старий базис. В якості $s$ можна брати будь-який індекс, для якого $\Delta_s<0$.

IX. Покласти $z_{m+1,s}=-1$, $z_{m+1,0}=\bar{x}_s$.

X. Якщо $\bar{x}_s=\alpha_s$, то покласти $v=0$ і перейти на крок XI; якщо $\bar{x}_s=\beta_s$ то покласти $v=1$ і перейти на крок XIII.

XI. Обчислити величини $\gamma_k$ при таких $k\in\left[1:(m+1)\right]$, що $z_{ks}\neq 0$:

$$\gamma_k = \begin{cases} \alpha_{i_k}, & \text{якщо } z_{ks} > 0; \\ \beta_{i_k}, & \text{якщо } z_{ks} < 0. \end{cases}$$

XII. Обчислити параметр $\theta_0$ перетворення опорного розв'язку $\bar{x}$

$$\theta_0 = \min_{\substack{z_{ks} \neq 0 \\ k \in [1; m+1]}} \frac{z_{k0} - \gamma_k}{z_{ks}},$$

позначити через $r$ індекс такий, що $\theta_0 = \dfrac{z_{r0} - \gamma_r}{z_{rs}}$, і перейти на крок XV.

*Примітка.* У невиродженому випадку індекс $r$ визначається однозначно. У виродженому випадку в якості $r$ вибирається найменший індекс, на якому досягається мінімум. Якщо у виродженому випадку виникає зациклювання, то варто використати правило, що виводить із циклу.

XIII. Обчислити величини $\gamma_k$ при $k \in \left[1 : (m+1)\right]$, для яких $z_{ks} \neq 0$:

$$\gamma_k = \begin{cases} \alpha_{i_k}, & \text{якщо } z_{ks} < 0; \\ \beta_{i_k}, & \text{якщо } z_{ks} > 0. \end{cases}$$



XIV. Обчислити параметр $\theta_0$ перетворення опорного розв'язку $\bar{x}$ за правилом

$$\theta_0 = \min_{\substack{z_{ks} \neq 0 \\ k \in [1; m+1]}} \frac{\gamma_k - z_{k0}}{z_{ks}},$$

позначити через $r$ індекс, такий, що $\theta_0 = \dfrac{\gamma_r - z_{r0}}{z_{rs}}$, і перейти на крок XV (див. крок XII).

XV. Покласти $x' = \bar{x}$; $z'_{kj} = z_{kj}$, $k = \overline{1,m}$, $j = \overline{1,n}$; $y'_j = y_j$; $j = \overline{1,n}$; $\Delta'_j = \Delta_j$, $j = \overline{1,n}$ і перейти на крок XVI.

XVI. Обчислити новий опорний розв'язок $\bar{x} = \left(\bar{x}_1, \ldots, \bar{x}_n\right)^T$ за правилом:

$$\bar{x}_l = \begin{cases} x'_{i_k} - (-1)^v \theta_0 z'_{ks}, & \text{якщо } l = i_k,\ k = \overline{1,m}; \\ x'_s + (-1)^v \theta_0, & \text{якщо } l = s; \\ x'_l, & \text{якщо } l \notin \Im_{\bar{x}},\ l \neq s. \end{cases}$$

XVII. Покласти $i_r = s$, тобто, замінити в розширеному базисі $a^{i_1}, \ldots, a^{i_m}, a^{i_{m+1}}$ значення вектора $a^{i_r}$ значенням вектора $a^s$ і перейти до нового базису $a^{i_1}, \ldots, a^{i_m}$, який отримується із перших $m$ векторів розширеного базису (при переході до нового базису старий базис не змінюється, якщо $r = m+1$ і у старому базисі змінюється значення одного вектора $a^{i_r}$, якщо $r < m+1$).

XVIII. Знайти коефіцієнти $z_{kj},\ k = \overline{1,m};\ j = \overline{1,n}$ розкладу векторів $a^1, \ldots, a^n$ по векторах нового базису:

а) якщо $r = m+1$ (якщо базис не змінився), то покласти

$$z_{kj} = z'_{kj},\ \ k = \overline{1,m};\ \ j = \overline{1,n};$$

б) якщо $r < m+1$ (тобто якщо базис змінився), то обчислити $z_{kj}$ за основними формулами:

$$z_{rj} = z'_{rj} \big/ z'_{rs},\ \ j = \overline{1,n};$$

$$z_{kj} = z'_{kj} - (z'_{rj} \big/ z'_{rs}) z'_{ks},\ \ \ j = \overline{1,n};\ k = 1, \ldots, r-1, r+1, \ldots, m.$$

XIX. Покласти $z_{k0} = \bar{x}_{i_k},\ k = \overline{1,m}$.

XX. Обчислити оцінки $\Delta_j,\ \ j = \overline{1,n}$, які відповідають новому базису:

а) якщо $r = m+1$ (базис не змінився), то покласти

$$\Delta_j = \begin{cases} \Delta'_j, & \text{якщо} \quad j \neq s; \\ -\Delta'_j, & \text{якщо} \quad j = s; \end{cases}$$



б) якщо $r < m+1$ (базис змінився), то обчислити величини $y_j,\ j = \overline{1,n}$ за основними формулами:

$$y_j = y'_j - (z'_{rj} \big/ z'_{rs}) y'_s,\ \ j = \overline{1,n},$$

і покласти:

$$\Delta_j = (-1)^{v_j} y_j,\ \ \ j = \overline{1,n},$$

де $v_j = 0$, якщо $\bar{x}_j = \alpha_j$, і $v_j = 1$, якщо $\bar{x}_j = \beta_j$.

Заздалегідь відомо, що $\Delta_j = 0$, якщо $j \in \Im_{\bar{x}}$. Тому обчислення $\Delta_j$ варто проводити для $j \notin \Im_{\bar{x}}$, тобто, коли або $\bar{x}_j = \alpha_j$, або $\bar{x}_j = \beta_j$.

XXI. Перейти на крок VI.

***Теорема 1.*** *Якщо виконані припущення 0, то в невиродженому випадку процес розв'язування задачі 0 алгоритмом 1 через скінченне число ітерацій закінчиться або на кроці VI (знаходиться оптимальний розв'язок задачі 0), або на кроці VII (встановлюється, що функція цілі задачі 0 необмежена на допустимій множині X).*

Алгоритм 1 може застосовуватися також для розв'язування вироджених задач лінійного програмування із двосторонніми обмеженнями. Тільки при виникненні циклу варто використати спеціальне правило виводу вектора із базису, що виключає зациклювання.

**2. Правило вибору індексу *r* (який визначає вектор $a^{i_r}$, що виводиться із базису) для запобігання зациклювання**

Якщо при використанні алгоритму 1 для розв'язування задачі лінійного програмування у виродженому випадку виникло зациклювання, то на кроці XII або XIV алгоритму варто застосувати наступну процедуру вибору індексу *r*, яка гарантує від зациклювання.

***Алгоритм 2***

Початок. I. Обчислити лінійно незалежну систему векторів $w^k$, $k=\overline{1,m}$:

$$w^k=\begin{cases} a^{i_k}, & \text{якщо } \ z_{k0}<\beta_{i_k}; \\ -a^{i_k}, & \text{якщо } \ z_{k0}=\beta_{i_k}. \end{cases}$$

II. Обчислити $y_{kj}$, $k=\overline{1,m}$; $j=\overline{1,m}$ – коефіцієнти розкладу векторів $w^k$, $k=\overline{1,m}$ по базису $a^{i_1},\ldots,a^{i_m}$, тобто такі числа, що:

$$w^j=\sum_{k=1}^{m}y_{kj}a^{i_k}, \quad j=\overline{1,m}.$$

III. Покласти $y_{m+1,j}=0$, $j=\overline{1,m}$.

Основний цикл. IV. Якщо $\overline{x}_s=\alpha_s$, то перейти на крок V; якщо $\overline{x}_s=\beta_s$, то перейти на крок X.

V. Знайти множину $G_0$, яка складається з тих індексів $k$, на яких досягається мінімум

$$\theta_0=\min_{\substack{z_{ks}\neq 0\\ k\in[1;m+1]}}\frac{z_{k0}-\gamma_k}{z_{ks}}.$$

VI. Покласти $\mu=0$.

VII. Якщо $G_\mu$ складається з одного елемента, то прийняти *r* рівним цьому елементу і припинити обчислення, інакше (якщо $G_\mu$ складається більш ніж з одного елемента) перейти на крок VIII.

VIII. Знайти множину $G_{\mu+1}$, яка складається з індексів $k\in G_\mu$, на яких досягається мінімум

$$\min_{k\in G_\mu}\left(y_{k,\mu+1}/z_{ks}\right).$$

IX. Покласти $\mu=\mu+1$ і перейти на крок VII.

X. Знайти множину $G_0$, яка складається з індексів $k$, на яких досягається мінімум

$$\theta_0=\min_{\substack{z_{ks}\neq 0\\ k\in[1;m+1]}}\frac{\gamma_k-z_{k0}}{z_{ks}}.$$

XI. Покласти $\mu = 0$.

XII. Якщо $G_\mu$ складається з одного елемента, то прийняти $r$ рівним цьому елементу й припинити обчислення; інакше (якщо $G_\mu$ складається більш ніж з одного елемента) перейти на крок XIII.

XIII. Знайти множину $G_{\mu+1}$, яка складається з індексів $k \in G_\mu$, на яких досягається максимум

$$\max_{k \in G_\mu} (y_{k,\mu+1} / z_{ks}).$$

XIV. Покласти $\mu = \mu + 1$ і перейти на крок XII.

***Теорема 2.*** *Алгоритм 2 генерує скінченну послідовність множин* $G_0, G_1, \ldots, G_t$, $t \le m$, *таку, що* $G_t$ *містить єдиний елемент, який приймається за r.*

## 5.12. Модифікований симплекс-метод для розв'язування загальної задачі лінійного програмування



З а д а ч а 1. Знайти $\arg\max_x \sum_{j=1}^{n} c_j x_j$ для заданого вектора $c = (c_1, \ldots, c_n)$ при обмеженнях:

$$\sum_{j=1}^{n} \alpha_{ij} x_j = b_i, \quad i = 1, \ldots, p; \tag{5.64}$$

$$\sum_{j=1}^{n} \alpha_{ij} x_j \le b_i, \quad i = p+1, \ldots, m; \tag{5.65}$$

$$x_j \ge 0, \quad j = q+1, \ldots, n. \tag{5.66}$$

*Припущення 1.* $(\boldsymbol{i})-$ рівняння (5.64) лінійно незалежні; $(\boldsymbol{ii})-$ вектори $a^j = (\alpha_{1j}, \alpha_{2j}, \ldots, \alpha_{mj})^T$, $j = \overline{1,q}-$ лінійно незалежні, $(\boldsymbol{iii})-$ задача 1 має допустимий розв'язок.

*Означення 1.* Допустимий розв'язок $x = (x_1, \ldots, x_n)^T$ задачі 1 називається ***опорним***, якщо він породжує неособливу квадратну підматрицю

$$A_x \triangleq \begin{pmatrix} \alpha_{i_1 j_1} & \alpha_{i_1 j_2} & \ldots & \alpha_{i_1 j_s} \\ \alpha_{i_2 j_1} & \alpha_{i_2 j_2} & \ldots & \alpha_{i_2 j_s} \\ \ldots & \ldots & \ldots & \ldots \\ \alpha_{i_s j_1} & \alpha_{i_s j_2} & \ldots & \alpha_{i_s j_s} \end{pmatrix}$$

матриці $A \triangleq (\alpha_{ij})_{i=\overline{1,m}}^{j=\overline{1,n}}$ таку, що:

а) стовпці матриці $A_x$ містять всі стовпці матриці $A$, для яких компонента $x_j$ додатна в $x$ або не обмежена вимогою невід'ємності, $j_l = l$ для $l = \overline{1, q}$;

б) рядки матриці $A_x$ утворені умовами (5.64) і деякими з умов (5.65), причому всі такі обмеження-нерівності задовольняються допустимим розв'язком $x$ як строгі рівності, $i_k = k$ для $k = \overline{1, p}$.

Підматриця $A_x$, яка породжена опорним розв'язком $x$, називається його ***базисом***.

*Означення 2.* Опорний розв'язок $x$ задачі 1 з базисом $A_x$ називається ***невиродженим***, якщо:

$$x_{j_l} > 0 \quad \text{при} \quad l = \overline{q+1, s};$$

$$\sum_{j=1}^{n} \alpha_{ij} x_j < b_i \quad \text{при} \quad i \neq i_k \quad (k = \overline{1, s}).$$

З одного боку, задача 1 може бути розв'язана методами, наведеними вище, якщо попередньо звести її до стандартного вигляду з $n+m-(p+q)$ невід'ємними змінними і $m-q$ умовами рівностями, з іншого боку – існує ряд класів задач лінійного програмування (наприклад, двоїсті задачі при $m < n-m$), які вигідно розв'язувати в їх природному записі. Нижче наводиться модифікація симплекс-методу для розв'язування задачі 1 у природному записі. У початковій стадії алгоритму потрібно мати опорний розв'язок і його базис, а також необхідно обчислювати матрицю, обернену до початкової базисної. На кожній ітерації алгоритму опорний розв'язок перевіряється на оптимальність; якщо він не оптимальний, то здійснюється перехід до нового опорного розв'язку і нового базису (кращого від попереднього, тобто такого, що дає більше значення цільової функції). У невиродженому випадку за скінченне число ітерацій знаходиться оптимальний розв'язок $x^*$, або встановлюється необмеженість лінійної форми на допустимій множині розв'язків. У виродженому випадку теоретично може виникнути зациклювання, тоді необхідно застосовувати спеціальне правило, що гарантує від зациклювання.

***Алгоритм 1***

Початок. I. Знайти початковий опорний розв'язок $x = (x_1, \ldots, x_n)^T$ задачі 1 і його базис $A_x$, утворений елементами, які стоять на перетині рядків і стовпців матриці $A$ з номерами $i_1, \ldots, i_s$ та $j_1, \ldots, j_s$, відповідно.

II. Обчислити матрицю $A_x^{-1} = (\beta_{lk})_{l=\overline{1,s}}^{k=\overline{1,s}}$, обернену до матриці $A_x$.

Основний цикл. III. Обчислити числа:

$$\lambda_k = \sum_{l=1}^{s} c_{j_l} \beta_{lk}, \quad k = \overline{1,s}.$$

IV. Обчислити величини:

$$\Delta_j = \sum_{k=1}^{s} \alpha_{i_k j} \lambda_k - c_j, \quad j = \overline{1,n}.$$

V. Якщо виконуються нерівності:

$$\lambda_k \ge 0, \quad k = \overline{p+1,s};$$

$$\Delta_j \ge 0, \quad j = \overline{1,n},$$

то покласти $x^* = x$ і припинити обчислення (у цьому випадку опорний розв'язок $x$ є оптимальним розв'язком задачі 1); інакше покласти $\bar{A}_x = A_x$ (для цього потрібно покласти $\bar{j}_l = j_l$, $l = \overline{1,s}$; $\bar{i}_k = i_k$, $k = \overline{1,s}$); $\bar{x} = x$; $\bar{\beta}_{lk} = \beta_{lk}$; $l = \overline{1,s}$; $k = \overline{1,s}$ і перейти на крок VI.

VI. Серед величин $\lambda_k$, $k = \overline{p+1,s}$ і $\Delta_j$, $j = \overline{1,n}$ знайти найменшу. Якщо найменшою величиною виявилося $\Delta_r$, $r \in [1:n]$, то перейти на крок VII, а якщо $\lambda_t$, $t \in [(p+1):s]$, то перейти на крок XXI.

VII. Обчислити величини:

$$z_{lr} = \sum_{k=1}^{s} \beta_{lk} \alpha_{i_k r}, \quad l = \overline{1,s}.$$

VIII. Обчислити параметр

$$\theta_0' = \begin{cases} \min\limits_{\substack{z_{lr} > 0 \\ l > q}} (x_{j_l} / z_{lr}), & \text{якщо існує } z_{lr} > 0, \ l > q; \\ \infty, & \text{якщо } z_{lr} \le 0 \text{ для } l = \overline{q+1,s}. \end{cases}$$

IX. Обчислити параметр $\theta_0''$ за формулою:

$$\theta_0'' = \begin{cases} \min\limits_{\delta_i < 0} (-\Delta^{(i)} / \delta_i), & \text{якщо існує } \delta_i < 0, \ i \in [1:m]; \\ \infty, & \text{якщо } \delta_i \ge 0, \ i = \overline{1,m}, \end{cases} \tag{5.67}$$

де

$$\Delta^{(i)} = b_i - \sum_{j=1}^{n} \alpha_{ij} x_j, \quad i = \overline{1,m};$$

$$\delta_i = \sum_{l=1}^{s} \alpha_{ij_l} z_{lr} - \alpha_{ir}, \quad i = \overline{1,m}.$$

X. Обчислити параметр $\theta_0$ за формулою

$$\theta_0 = \min\{\theta_0', \theta_0''\}.$$

Якщо $\theta_0 = \infty$, то припинити обчислення (у цьому випадку цільова функція задачі 1 необмежена на допустимій множині); якщо $\theta_0 < \infty$, то перейти на крок XI.

XI. Обчислити новий опорний розв’язок $x = (x_1, \dots, x_n)^T$ за формулою:

$$x = \bar{x} - \theta_0 h,$$

де $h = (h_1, \dots, h_n)^T$ – вектор, який визначається за правилами:

$$h_j = \begin{cases} z_{lr}, \text{ якщо } j = j_l,\ l = \overline{1, s}; \\ -1, \text{ якщо } j = r; \\ 0, \text{ в інших випадках.} \end{cases}$$

XII. Якщо $\theta_0 = \theta_0'$, то перейти на крок XIII, якщо $\theta_0 = \theta_0''$, то перейти на крок XVII.

XIII. Обчислити індекс $v \in \left[(q+1) : s\right]$ такий, що

$$\theta_0' = x_{j_v} \big/ z_{vr}, \quad z_{vr} > 0. \tag{5.68}$$

XIV. Перейти до нового базису $A_x$, який отримується зі старого базису $\bar{A}_x$ шляхом заміни його $v$-стовпця стовпцем

$$(\alpha_{i_1 r}, \alpha_{i_2 r}, \dots, \alpha_{i_s r})^T \text{ (тобто покласти } j_v = r).$$

XV. Обчислити матрицю $A_x^{-1} = (\beta_{lk})_{l=\overline{1,s}}^{k=\overline{1,s}}$ обернену новій базисній матриці за правилами:

$$\beta_{lk} = \begin{cases} \bar{\beta}_{lk} - \bar{\beta}_{vk} z_{lr} \left(1/z_{vr}\right), & l \neq v; \\ \bar{\beta}_{vk} \left(1/z_{vr}\right), & l = v, \end{cases}$$

де $l = \overline{1, s}$ і $k = \overline{1, s}$.

XVI. Перейти на крок III.

XVII. Обчислити індекс $\mu \in \left[1 : m\right]$ такий, що $\theta_0'' = -\Delta^{(\mu)} \big/ \delta_\mu$, $\delta_\mu < 0$, $\mu \neq i_k$, $k = \overline{1, s}$.

XVIII. Перейти до нового базису $A_x$, який отримується шляхом обрамлення базису $\bar{A}_x$ рядком $(\alpha_{\mu j_1}, \alpha_{\mu j_2}, \dots, \alpha_{\mu j_s}, \alpha_{\mu r})$ і стовпцем $(\alpha_{i_1 r}, \alpha_{i_2 r}, \dots, \alpha_{i_s r}, \alpha_{\mu r})^T$; покласти $i_{s+1} = \mu$, $j_{s+1} = r$ і перейти на крок XIX.

XIX. Обчислити матрицю $A_x^{-1} = (\beta_{lk})_{l=\overline{1,s+1}}^{k=\overline{1,s+1}}$, обернену до нової базисної матриці за правилами:

$$\beta_{lk} = \begin{cases} \bar{\beta}_{lk} + \hat{\delta}_k z_{lr} / \alpha_0 & \text{для } l,\ k = \overline{1,s}; \\ -\hat{\delta}_k / \alpha_0 & \text{для } l = s+1,\ k = \overline{1,s}; \\ -z_{lr} / \alpha_0 & \text{для } l = \overline{1,s},\ k = s+1; \\ 1/\alpha_0 & \text{для } l = k = s+1, \end{cases}$$

де $\hat{\delta}_k$ та $\alpha_0$ обчислюються, відповідно, за формулами:

$$\hat{\delta}_k = \sum_{\gamma=1}^{s} \alpha_{\mu j_\gamma} \bar{\beta}_{\gamma k}, \quad k = \overline{1,s};$$

$$\alpha_0 = \alpha_{\mu r} - \sum_{\gamma=1}^{s} \alpha_{\mu j_\gamma} z_{\gamma r}.$$

XX. Покласти $s = s+1$ і перейти на крок III.

XXI. Знайти параметр $\theta_0''$ із умов:

$$\theta_0'' = \begin{cases} \infty, & \text{якщо } \hat{\delta}^{(i)} \ge 0, \quad i = \overline{1,m}; \\ \min\limits_{\substack{\hat{\delta}^{(i)} < 0 \\ i \in [1;m]}} (-\Delta^{(i)} / \hat{\delta}^{(i)}), & \text{якщо існує } \hat{\delta}^{(i)} < 0; \end{cases} \tag{5.69}$$

$$\Delta^{(i)} = b_i - \sum_{j=1}^{n} \alpha_{ij} x_j, \quad i = \overline{1,m}; \quad \hat{\delta}^{(i)} = \sum_{\gamma=1}^{s} \alpha_{ij_\gamma} \beta_{\gamma t}, \quad i = \overline{1,m}.$$

XXII. Знайти параметр $\theta_0'$ із умов:

$$\theta_0' = \begin{cases} \min\limits_{\substack{\beta_{\gamma t} > 0 \\ \gamma > q}} (x_{j_\gamma} / \beta_{\gamma t}), & \text{якщо існує } \beta_{\gamma t} > 0, \quad \gamma > q, \\ \infty, & \text{якщо } \beta_{\gamma t} \le 0, \quad \gamma = \overline{q+1,s}. \end{cases} \tag{5.70}$$

XXIII. Обчислити параметр $\theta_0$ за формулою

$$\theta_0 = \min\{\theta_0', \theta_0''\}.$$

Якщо $\theta_0 = \infty$, то припинити обчислення (у цьому випадку цільова функція задачі 1 необмежена на допустимій множині); якщо $\theta_0 < \infty$, то перейти на крок XXIV.

XXIV. Обчислити новий опорний розв'язок $x = (x_1, \ldots, x_n)^T$ за формулою:

$$x = \bar{x} - \theta_0 h,$$

де $h = (h_1, \ldots, h_n)^T$ – вектор, який визначається за правилом

$$h_j = \begin{cases} \beta_{lt}, & \text{якщо } j = j_l,\ l = \overline{1,s}, \\ 0, & \text{в інших випадках.} \end{cases}$$

XXV. Якщо $\theta_0 = \theta_0'$, то перейти на крок XXVI, якщо $\theta_0 = \theta_0''$ то перейти на крок XXX.

XXVI. Обчислити індекс $v \in [(q+1):s]$ такий, що

$$\theta_0' = (x_{j_v} / \beta_{vt}), \quad \beta_{vt} > 0.$$

Якщо задача 1 невироджена, то індекс $v$ визначається однозначно. У виродженому випадку можна вибрати будь-який індекс $v$, на якому досягається мінімум в (5.70).

XXVII. Перейти до нового базису $A_x$, який отримується зі старого базису $\overline{A}_x$ шляхом видалення його $t$-го рядка і $v$-го стовпця, (тобто покласти $j_1 = \overline{j}_1, \ldots, j_{v-1} = \overline{j}_{v-1}, j_v = \overline{j}_{v+1}, \ldots, \quad j_{s-1} = \overline{j}_s$; $\quad i_1 = \overline{i}_1, \ldots, i_{t-1} = \overline{i}_{t-1}$, $i_t = \overline{i}_{t+1}, \ldots, i_{s-1} = \overline{i}_s$, де $\overline{j}_l$, $\overline{i}_k$, $l = \overline{1,s}$, $k = \overline{1,s}$ визначають старий базис $\overline{A}_x$ ).

XXVIII. Обчислити матрицю $A_x^{-1} = (\beta_{lk})_{l=\overline{1,s-1}}^{k=\overline{1,s-1}}$, обернену до нової базисної матриці за правилами:

$$\beta_{lk} = \begin{cases} \overline{\beta}_{lk} - (\overline{\beta}_{vk}\overline{\beta}_{lt} / \overline{\beta}_{vt}), & \text{якщо } l < v,\ k < t; \\ \overline{\beta}_{l+1,k} - (\overline{\beta}_{vk}\overline{\beta}_{l+1,t} / \overline{\beta}_{vt}), & \text{якщо } l \ge v,\ k < t; \\ \overline{\beta}_{l,k+1} - (\overline{\beta}_{v,k+1}\overline{\beta}_{lt} / \overline{\beta}_{vt}), & \text{якщо } l < v,\ k \ge t; \\ \overline{\beta}_{l+1,k+1} - (\overline{\beta}_{v,k+1}\overline{\beta}_{l+1,t} / \overline{\beta}_{vt}), & \text{якщо } l \ge v,\ k \ge t, \\ \qquad l = \overline{1,s-1}, \quad k = \overline{1,s-1}. \end{cases}$$

XXIX. Покласти $s = s - 1$ і перейти на крок III.

XXX. Обчислити індекс $\mu \in [1:m]$ такий, що

$$\theta_0'' = -\Delta^{(\mu)} / \hat{\delta}^{(\mu)}, \quad \hat{\delta}^{(\mu)} < 0; \quad \mu \ne i_k, \quad k = \overline{1,s}.$$

Якщо задача 1 невироджена, то індекс $\mu$ визначається однозначно. У випадку виродженої задачі можна вибрати будь-який індекс $\mu$, на якому досягається мінімум в (5.69).

XXXI. Обчислити величини

$$\hat{\delta}_k = \sum_{l=1}^{s} \alpha_{\mu j_l} \beta_{lk}, \quad k = \overline{1,s}.$$

XXXII. Перейти до нового базису $A_x$, який отримується зі старого базису $\overline{A}_x$ шляхом заміни його $t$-го рядка рядком $(\alpha_{\mu j_1}, \alpha_{\mu j_2}, \ldots, \alpha_{\mu j_s})$ (тобто покласти $i_t = \mu$ ).

XXXIII. Обчислити матрицю $A_x^{-1} = (\beta_{lk})_{l=\overline{1,s}}^{k=\overline{1,s}}$, обернену новій базисній матриці за правилами

$$\beta_{lk} = \begin{cases} \overline{\beta}_{lk} - (\overline{\beta}_{lt}\hat{\delta}_k / \hat{\delta}^{(\mu)}), & k \neq t; \\ \overline{\beta}_{lt} / \hat{\delta}^{(\mu)}, & k = t, \end{cases} \quad l,\ k = \overline{1, s}.$$

XXXIV. Перейти на крок III.

***Теорема 1.*** *Якщо виконані припущення 1, то в невиродженому випадку процес розв'язування задачі 1 алгоритмом 1 через скінчення число ітерацій закінчиться або на кроці V (знаходиться оптимальний розв'язок $x^*$ задачі 1), або на одному із кроків X або XXIII (встановлюється, що функція цілі задачі 1 необмежена на допустимій множині розв'язків).*

## 5.13. Ітераційні методи

Ітераційні методи являють собою послідовність однорідних за процедурою виконання ітерацій, які призводять в границі до оптимального розв'язку задачі. Їх застосовують тоді, коли необхідно швидко знайти грубий розв'язок задачі, а також в задачах великої розмірності.

**1. Ітераційний метод Петшиковського**

Задача 1. Знайти $\arg\max\limits_x (c, x)$ для заданого вектора $c \in R^n$ при обмеженнях:

$$\varphi_i(x) \triangleq (\overline{a}^i, x) - a_i^0 \geq 0,\ i = \overline{1, m};$$

$$\varphi_{m+j}(x) \triangleq x_j = (e^j, x) \geq 0,\quad j = \overline{1, n_1},\ (n_1 \leq n), \tag{5.71}$$

де $\overline{a}^i$ – $i$-й вектор-рядок матриці $A$ задачі 2 п.5.7; $e^j$ – $n$-вимірний одиничний вектор з одиницею на $j$-му місці.

Якщо покласти $\overline{a}^{m+i} = e^i,\ a_{m+i}^0 = 0,\ i = \overline{1, n_1}$, то обмеження (5.71) задачі 1 можна переписати в наступній еквівалентній формі:

$$\varphi_i(x) \triangleq (\overline{a}^i, x) - a_i^0 \geq 0,\ i = \overline{1, m + n_1}. \tag{5.72}$$

На відміну від розглянутих раніше, метод Петшиковського не є скінченим. Розв'язок задачі 1 одержуємо лише в границі.

В методі Петшиковського задача 1 з обмеженнями (5.72) зводиться до задачі безумовної максимізації опуклої екстремальної функції:

$$g_\mu(x) \triangleq (c, x) + \frac{\mu}{2}\sum_{i=1}^{m+n_1} \delta_i(x)\left[\varphi_i(x)\right]^2,$$

де

$$\delta_i(x) = \begin{cases} 0, & \text{якщо} \quad \varphi_i(x) \geq 0, \\ -1, & \text{якщо} \quad \varphi_i(x) < 0. \end{cases}$$

Для розв'язання задачі безумовної оптимізації застосовується градієнтний метод, який при досить великих $\mu$ дає добре наближення розв'язку задачі 1.

***Алгоритм 1***

П о ч а т о к . I. Вибрати довільне початкове наближення $x^0 \in R^n$.

II. Вибрати параметр $\mu > 0$.

III. Покласти $k = 0$.

О с н о в н и й  ц и к л . IV. Обчислити величини:

$$\beta_i^k = \varphi_i\left(x^k\right),\ i = \overline{1, m+n_1}.$$

V. Обчислити градієнт функції $g_\mu$ в точці $x = x^k$

$$\nabla g_\mu\left(x^k\right) = c + \mu \sum_{i=1}^{m+n_1} \delta_i\left(x^k\right) \beta_i^k \overline{a}^i.$$

VI. Обчислити величини:

$$\xi_i^k = \left(\overline{a}^i, \nabla g_\mu\left(x^k\right)\right),\ i = \overline{1, m+n_1}.$$

VII. Для всіх $i \in \left[1 : \left(m+n_1\right)\right]$, для яких $\xi_i^k \neq 0$, обчислити відношення $\theta_i^k = -\beta_i^k / \xi_i^k$ і розташувати їх в порядку зростання

$$\theta_{i_1}^k \leq \theta_{i_2}^k \leq \ldots \leq \theta_{i_s}^k.$$

VIII. Визначити функцію

$$\zeta(\theta) \triangleq \frac{\partial}{\partial \theta} g_\mu\left(x^k + \theta \nabla g_\mu\left(x^k\right)\right) = \left(c, \nabla g_\mu\left(x^k\right)\right) +$$

$$+ \mu \sum_{i=1}^{m+n_1} \delta_i\left(x^k + \theta \nabla g_\mu\left(x^k\right)\right)\left(\beta_i^k + \theta \xi_i^k\right) \xi_i^k.$$

IX. Знайти індекс $l \in [1 : s]$, при якому виконуються нерівності:

$$\zeta\left(\theta_{i_l}^k\right) > 0; \quad \zeta\left(\theta_{i_{l+1}}^k\right) \leq 0.$$

X. Обчислити кроковий множник

$$\rho_k = \left[\zeta\left(\theta_{i_l}^k\right)\theta_{i_{l+1}}^k - \zeta\left(\theta_{i_{l+1}}^k\right)\theta_{i_l}^k\right] \Big/ \left[\zeta\left(\theta_{i_l}^k\right) - \zeta\left(\theta_{i_{l+1}}^k\right)\right].$$

XI. Обчислити наступне наближення

$$x^{k+1} = x^k + \rho_k \nabla g_\mu\left(x^k\right).$$

XII. Покласти $k = k+1$ і перейти на крок IV.

***Теорема 1.*** *Послідовність* $\left\{x^k\right\}_{k=0}^{\infty}$, *яка породжена алгоритмом 1, при довільному* $x^0 \in R^n$ *і* $\mu > 0$ *має властивість:*

$$\lim_{k \to \infty} d_1\left(x^k, X_\mu\right) = 0,$$

*де* $X_\mu$ *– множина, на якій досягається максимум функції* $g_\mu$; $d_1(x,Y)$ *– відстань від точки* $x$ *до множини* $Y$.

***Теорема 1′.*** *Для довільного* $\varepsilon > 0$ *існує таке* $\mu_0 > 0$, *що при довільному* $\mu > \mu_0$ *нерівність*

$$d_1(x, X^*) < \varepsilon$$

*виконується для довільного* $x \in X_\mu$, *де* $X^*$ *– множина розв'язків задачі 1.*

З теореми 1 і 1′ випливає, що при досить великих значеннях $k$ і $\mu$ точка $x^k$, яка отримана за допомогою алгоритму 1, досить добре наближує розв'язок задачі 1.

**2. Ітераційний метод, який використовує модифіковану функцію Лагранжа**

З а д а ч а 2. Знайти $\arg\max\limits_{x\in X}(c,x)$ для заданого вектора $c \in R^n$ і заданої множини:

$$X \triangleq \left\{x \mid Ax = a^0,\ x \ge 0,\ x \in R^n\right\},$$

де $a^0 \in R^m$; $A$ – матриця розмірності $m \times n$.

*Припущення 2.* Множина $X^*$ розв'язків задачі 2 непорожня.

Наведений нижче алгоритм оснований на побудові модифікованої функції Лагранжа і потребує на кожній ітерації розв'язувати задачу мінімізації квадратичної функції. Цей метод знаходить розв'язки задачі 2 за скінченну кількість ітерацій $k_1$, що залежить від величини коефіцієнтів штрафу і вибору початкового наближення двоїстої змінної (при цьому припускається, що задача мінімізації квадратичної функції розв'язується за скінченну кількість ітерацій).

***Алгоритм 2***

П о ч а т о к. I. Вибрати довільне початкове наближення двоїстої змінної $y^0 \in R^m$.

II. Вибрати довільну послідовність штрафних коефіцієнтів $\{\alpha_k\}_{k=0}^{\infty}$, $\alpha_k > 0,\ k = 0, 1, \ldots$.

III. Покласти $k = 0$.

О с н о в н и й ц и к л. IV. Визначити модифіковану функцію Лагранжа

$$\psi(x, y, \alpha_k) \triangleq (c,x) + \left(y, a^0 - Ax\right) + \frac{\alpha_k}{2}\left\|Ax - a^0\right\|^2.$$

V. Обчислити наближення $x^k \ge 0$ основної змінної з умови

$$\psi\left(x^k, y^k, \alpha_k\right) = \min_{x\ge 0}\psi\left(x, y^k, \alpha_k\right).$$

VI. Обчислити наступне наближення двоїстої змінної

$$y^{k+1} = y^k + \alpha_k \left(a^0 - Ax^k\right).$$

VII. Покласти $k = k+1$ і перейти на крок IV.

***Теорема 2.*** *Якщо виконуються припущення 2 і така послідовність штрафних коефіцієнтів* $\{\alpha_k\}_{k=0}^{\infty}$, *що* $\alpha_k \ge \alpha > 0,\ k = 0, 1, ...,$ *то алгоритм 2 скінченний, тобто знайдеться такий номер* $k_1$, *що* $x^{k_1} \in X^*,\ \ y^{k_1} \in Y^*$, *де* $X^*, Y^*$ *– множина сідлових точок функції Лагранжа*

$$\varphi(x, y) \triangleq (c, x) + \left(y, a^0 - Ax\right).$$

*Число* $k_1$ *тим менше, чим менше відношення* $\rho\left(y^0, Y^*\right)/\alpha$. *Зокрема для довільного* $y^0$ *знайдеться* $\bar{\alpha}$ *таке, що при* $\alpha_0 \ge \bar{\alpha}$ *алгоритм 2 закінчиться за одну ітерацію, тобто буде* $x^1 \in X^*,\ \ y^1 \in Y^*$ *(тут* $\rho(x, Y) = \inf\limits_{y \in Y} \|x - y\|$ *).*

*Зауваження 2.* Важливою проблемою в алгоритмі 2 є вибір послідовності штрафних коефіцієнтів $\alpha_k,\ k = 0, 1, \ldots$.

Збільшення $\alpha_k$ вигідне з точки зору зменшення числа ітерацій. Однак при великих значеннях $\alpha_k$ важко розв'язати задачу мінімізації функції $\psi\left(x, y^k, \alpha_k\right)$. Тому задовільне значення $\alpha_k$ слід вибирати в процесі обчислення.



**3. Ітераційний метод Федоренка**

Задача 3. Знайти $\arg\min\limits_{y \in Y} \sum\limits_{i=1}^{n} y_i h_i^0$ для заданого вектора $h^0 = \left(h_1^0, ..., h_n^0\right)$ і заданої множини

$$Y \triangleq \left\{ y \mid b_j + \sum_{i=1}^{n} y_i h_i^j = 0,\ \ j = \overline{1, m};\ \ \ \alpha_i^- \le y_i \le \alpha_i^+,\ \ i = \overline{1, n};\ \ \ y \in R^n \right\}.$$

Якщо відобразити точки $y = (y_1, ..., y_n)$ простору $R^n$ в точки $x = (x_1, ..., x_{m+1})$ простору $R^{m+1}$ за правилом

$$x = \begin{pmatrix} 0 \\ b_1 \\ \vdots \\ b_m \end{pmatrix} + y_1 \begin{pmatrix} h_1^0 \\ h_1^1 \\ \vdots \\ h_1^m \end{pmatrix} + \ldots + y_n \begin{pmatrix} h_n^0 \\ h_n^1 \\ \vdots \\ h_n^m \end{pmatrix}, \tag{5.73}$$

то образом $n$-вимірного «паралелепіпеда»

$$\sigma = \left\{ y \mid \alpha_i^- \le y_i \le \alpha_i^+,\ \ i = \overline{1, n} \right\}$$

буде опуклий многокутник $\tau$ в просторі $R^{m+1}$. Тоді метод розв'язування задачі 3 зводиться до відшукання мінімального числа $\lambda$, для якого $\lambda e \in \tau$, де $e = (1,0,...,0) \in R^{m+1}$ (мінімальне значення $\lambda$ позначимо через $\lambda *$).

Наведений нижче алгоритм визначає за скінченне число ітерацій наближений розв'язок $\tilde{x}$ допоміжної задачі і відповідний наближений розв'язок $\bar{y}$ задачі 3 (точність наближеного розв'язку $\tilde{x}$ задається константою $\delta_0$). Об'єм обчислень залежить від вибору параметрів алгоритму. На кожній основній ітерації алгоритму 3 використовується задане число «малих» ітерацій довільного вибраного алгоритму квадратичного програмування.

***Алгоритм 3***

Початок. I. Вибрати вектор $g \in R^{m+1}$ такий, що $(g,e) = 1$ (тут $e = (1,0,...,0)$).

II. Покласти:

$$a^i = \left(h_i^0, h_i^1, ..., h_i^m\right),\ i = \overline{1,n};\ c = \left(0, b_1, ..., b_m\right).$$

III. Задати число $l > 0$ – кількість «малих» ітерацій розв'язування допоміжної задачі квадратичного програмування (зазвичай $l \in [10;20]$).

IV. Задати константи $\beta \in (0;1)$, $\varepsilon > 0$, $d*$ $(0 < d* < 1)$, $\eta* > 0$, $\delta_0 > 0$, $\omega_0 > 0$.

Основний цикл. V. Обчислити величини:

$$\varphi_i = \left(a^i, g\right),\ i = \overline{1,n}.$$

VI. Обчислити величини:

$$\hat{\alpha}_i = \begin{cases} \alpha_i^-, & \text{якщо} \quad \varphi_i \ge 0, \\ \alpha_i^+, & \text{якщо} \quad \varphi_i < 0, \quad i = \overline{1,n}. \end{cases}$$

VII. Обчислити вектор

$$\hat{x} = c + \sum_{i=1}^{n} \hat{\alpha}_i a^i.$$

VIII. Обчислити значення

$$f_0(g) = (\hat{x}, g). \tag{5.74}$$

Відмітимо, що точка $\hat{x}$ – розв'язок задачі

$$\min_{x \in \tau} (x, g) = f_0(g),$$

при цьому $f_0(g)$ дає точну оцінку знизу для $\lambda*$: $f_0(g) \le \lambda *$.

IX. Знайти підмножину індексів $\Im_\varepsilon$ $\left(\Im_\varepsilon \subset \{1,...,n\}\right)$ за правилом:

$$i \in \Im_\varepsilon,\ \text{якщо}\ |\phi_i| \le \varepsilon \left\|a^i\right\|_G,$$

де норма $\|z\|_G$ визначається як відстань точки $z$ від прямої $\lambda e$ $(-\infty < \lambda < \infty)$, в площині $G$, що проходить через $z$ ортогонально вектору $g$, і обчислюється за правилом

$$\|z\|_G = \|z - (z, g)\varepsilon\|. \tag{5.75}$$

X. Обчислити точку

$$\bar{x} = c + \sum_{i \notin \Im\varepsilon} \hat{\alpha}_i a^i .$$

XI. Покласти $v = 1$.

XII. Вибрати алгоритм $\hat{A}$ для розв'язування наступної задачі квадратичного програмування:

$$\text{знайти } \arg \min_{\substack{y_i \in [\alpha_i^-; \alpha_i^+] \\ i \in \Im_\varepsilon}} \| \bar{x} + \sum_{i \in \Im_\varepsilon} y_i a^i \|_G , \tag{5.76}$$

де норма $\| \cdot \|_G$ визначається за правилом (5.75).

XIII. Знайти наближений розв'язок $y_i^v, i \in \Im_\varepsilon$, задачі (5.76), застосовуючи $l$ ітерацій алгоритму $\hat{A}$.

XIV. Обчислити вектор

$$\tilde{x} = \bar{x} + \sum_{i \in \Im_\varepsilon} y_i^v a^i .$$

XV. Обчислити величину

$$\eta = \left|(\tilde{x}, g) - f_0(g)\right| / \left|f_0(g)\right| .$$

Якщо $\eta > \eta^*$, то покласти $\varepsilon = \beta\varepsilon$ і перейти на крок IX; якщо $\eta \le \eta^*$, то перейти на крок XVI.

XVI. Якщо $\|\tilde{x} - (\tilde{x}, g)e\| \le \delta_0$, то зупинити обчислення (у цьому випадку точка $\tilde{x}$ приймається за розв'язок допоміжної задачі, а вектор $\bar{y} = (\bar{y}_1, ..., \bar{y}_n)$, де

$$\bar{y}_i = \begin{cases} y_i^v, & i \in \Im_\varepsilon; \\ \hat{\alpha}_i, & i \notin \Im_\varepsilon, \quad i = \overline{1, n}, \end{cases}$$

приймається за наближений розв'язок задачі 3); інакше перейти на крок XVII.

XVII. Обчислити вектор

$$\upsilon = (\tilde{x} - (\tilde{x}, g)e) / \|\tilde{x} - (\tilde{x}, g)e\| .$$

XVIII. Обчислити значення

$$\psi_i = (\upsilon, a^i) - \varphi_i(e, \upsilon), \; i = \overline{1, n} .$$

XIX. Обчислити значення

$$d = \frac{1}{\left\| \tilde{x} - \left( \tilde{x}, g \right) e \right\|} \sum_{i \in \Im_\varepsilon} \psi_i \times \begin{cases} y_i^\nu - \alpha_i^+, & \text{якщо} \quad \psi_i \le 0, \\ y_i^\nu - \alpha_i^-, & \text{якщо} \quad \psi_i > 0. \end{cases}$$

XX. Якщо $d > d*$ (в цьому випадку задача (5.76) не розв'язана з необхідною точністю), то покласти $\nu = \nu + 1$ і перейти на крок XIII, приймаючи за початкове наближення алгоритму $\hat{A}$ наближений розв'язок $y_i^{\nu-1}$ задачі (5.76), одержаний на $(\nu - 1)$-й «малій» ітерації; якщо $d \le d*$, то перейти на крок XXI для перерахунку вектора $g$.

(«Малі» ітерації породжуються кроками XI–XX алгоритму 3.

Новий вектор $\tilde{g}$ визначається як

$$\tilde{g} = \tilde{g}(\rho) = \left(1 - \rho(e, \upsilon)\right) g + \rho \upsilon,$$

де параметр $\rho = \arg\max_\rho \min_{x \in \tau} \left( x, \tilde{g}(\rho) \right)$. На кроках XXI–XXX алгоритму методом «ділення відрізка» обчислюється необхідне значення $\rho$ за скінченне число ітерацій.

Величина $(\tilde{x}, g)$ є верхньою оцінкою для $\lambda*$, точною лише при $\|\tilde{x}\|_G = 0$ ).

XXI. Обчислити значення

$$\zeta_0 = (c, \upsilon) - \xi_0(e, \upsilon), \quad \xi_0 = (c, g).$$

XXII. Покласти $\bar{w}^1 = (1;0)$, $\tilde{w}^1 = (0;1)$.

XXIII. Для вектора $w = (w_1, w_2)$ визначити функцію

$$\zeta(w) = \zeta_0 + \sum_{i=1}^{n} \psi_i \times \begin{cases} \alpha_i^+, & \text{якщо } w_1 \varphi_i + w_2 \psi_i < 0, \\ \alpha_i^-, & \text{в іншому випадку.} \end{cases}$$

XXIV. Якщо $\zeta\left(\tilde{w}^1\right) > 0$, то зупинити обчислення (в цьому випадку задача 3 не має розв'язку), якщо $\zeta\left(\tilde{w}^1\right) < 0$, то перейти на крок XXV.

XXV. Покласти $j = 1$.

XXVI. Обчислити вектор

$$\chi^j = \frac{1}{2}\left( \bar{w}^j + \tilde{w}^j \right).$$

XXVII. Якщо $\zeta\left(\chi^j\right) > 0$, то покласти

$$\bar{w}^{j+1} = \chi^j, \quad \tilde{w}^{j+1} = \tilde{w}^j, \quad j = j + 1,$$

і перейти на крок XXVIII; якщо $\zeta\left(\chi^j\right) < 0$, то покласти

$$\bar{w}^{j+1} = \bar{w}^j, \quad \tilde{w}^{j+1} = \chi^j, \quad j = j + 1,$$

і перейти на крок XXVIII.

XXVIII. Якщо $\| \bar{w}^j - \tilde{w}^j \| \le w_0$, то перейти на крок XXIX; інакше перейти на крок XXVI.

XXIX. Обчислити

$$\rho = \left(\bar{w}_2^j + \tilde{w}_2^j\right) \Big/ \left(\bar{w}_1^j + \tilde{w}_1^j\right),$$

де $\left(\bar{w}_1^j, \bar{w}_2^j\right) = \bar{w}^j$; $\left(\tilde{w}_1^j, \tilde{w}_2^j\right) = \tilde{w}^j$.

XXX. Обчислити вектор

$$\tilde{g} = \left(1 - \rho\left(e, \upsilon\right)\right) g + \rho \upsilon.$$

XXXI. Покласти $g = \tilde{g}$ і перейти на крок V.

***Теорема 3.*** *Нехай основний ітераційний цикл V–XXXI приводить від вектора $g^k$ до вектора $g^{k+1}$. Нехай цей перехід здійснюється в ситуації, коли наближене розв'язування задачі мінімізації (5.76) дало точку $\tilde{x}^k$ (при переході від кроку XX до XXI), що задовольняє умову $d \le d^*$. Тоді*

$$f_0\left(g^{k+1}\right) > f_0\left(g^k\right) + \varepsilon \left\|\tilde{x}^k\right\|_G \left(1 - d^*\right),$$

*де $f_0\left(g\right)$ визначається за (5.74), а норма $\| \cdot \|_G$ – за (5.75).*

Алгоритм 3 за скінченну кількість ітерацій знаходить наближений розв'язок $\tilde{x}$ допоміжної задачі і відповідний наближений розв'язок $\bar{y}$ задачі 3.

Збіжність $\left\|\bar{w}^j - \tilde{w}^j\right\|$ до нуля має порядок $\left(1/2\right)^j$, тому необхідна точність $w_0$ досягається за порівняно невелике число ітерацій по $j$.



#### 4. Алгоритм «Заєць» розв'язування прямої і двоїстої задач лінійного програмування

З а д а ч а 4 (п р я м а з а д а ч а). Знайти $\arg\max\limits_{x \in X}\left(c, x\right)$ для заданого вектора $c \in R^n$ і заданої множини

$$X \triangleq \left\{x \mid Ax \le b,\ x \ge 0,\ b \in R^m,\ x \in R^n\right\}.$$

З а д а ч а 4′ (д в о ї с т а з а д а ч а). Знайти $\arg\min\limits_{y \in Y}\left(y, b\right)$ для заданого вектора $b \in R^m$ і заданої множини

$$Y \triangleq \left\{y \mid yA \ge c,\ \ y \ge 0,\ \ c \in R^n,\ \ y \in R^m\right\}.$$

*Припущення 4.* Задачі 4 і 4′ мають розв'язки.

Розв'язування пари двоїстих задач 4 і 4′ зводиться до розв'язування наступної додаткової задачі нелінійного програмування:

$$\text{знайти } \arg\min_{\left(x, y, \xi, \eta\right)} g\left(x, y, \xi, \eta\right)$$

при обмеженнях:

$$x \ge 0,\ \ \xi \ge \sigma x,\ \ \xi \ge \tau e^n,\ \ y \ge 0,\ \ \eta \ge \delta y,\ \ \eta \ge \tau e^m, \qquad (5.77)$$

де

$$g(x, y, \xi, \eta) = (y, b) + (q(y), \xi) - (c, x) - (p(x), \eta);$$
$$q(y) = \max\{-yA + c,\ 0\};\ \ p(x) = \max\{Ax - b,\ 0\};$$

$e^n \in R^n$, $e^m \in R^m$ – вектори з одиничними компонентами; $\sigma, \tau -$ константи, що задовольняють умовам $1 < \sigma < 3,\ \tau > 0$.

Якщо вектор $(x^*, y^*, \xi^*, \eta^*)$ – розв'язок допоміжної задачі, то $x^*-$ розв'язок 4, а $y^*$ – розв'язок задачі $4'$.

***Алгоритм 4***

Початок. I. Вибрати довільне початкове наближення $x^0 \in R_+^n,\ y^0 \in R_+^m$.

II. Вибрати константи $\sigma, \tau$ з умов :

$$1 < \sigma < 3,\ \ \tau > 0.$$

III. Вибрати вектори $\xi^0 = (\xi_1^0, ..., \xi_n^0)^T$, $\eta^0 = (\eta_1^0, ..., \eta_m^0)$, що задовольняють обмеженням (5.77) при $x = x^0,\ y = y^0$.

IV. Вибрати константи $l$ і $\hat{\rho} < \dfrac{1}{2}$ $(l > \hat{\rho})$, які визначають крокові множники алгоритму.

V. Вибрати константу $\varepsilon > 0$, яка визначає точність обчислення оптимального значення задач 4 і $4'$.

VI. Покласти $x = x^0,\ y = y^0,\ \xi = \xi^0,\ \eta = \eta^0,\ \varphi_1 = -\infty,\ \psi_1 = +\infty,\ \bar{y} = y^0$, $d = 0,\ k = 1,\ \rho = 1/2$.

Основний цикл. VII. Обчислити вектори $\hat{x}(y) = (\hat{x}_1(y), ..., \hat{x}_n(y))^T$, $\hat{y}(x) = (\hat{y}_1(x), ..., \hat{y}_m(x))$ за формулами:

$$\hat{x}_j(y) = \begin{cases} 0, & \text{якщо} \quad q_j = 0, \\ \xi_j, & \text{якщо} \quad q_j > 0, \end{cases} \quad j = \overline{1, n};$$

$$\hat{y}_v(x) = \begin{cases} 0, & \text{якщо} \quad p_v = 0, \\ \eta_v, & \text{якщо} \quad p_v > 0, \end{cases} \quad v = \overline{1, m},$$

де

$$p = \max\{Ax - b, 0\};\ \ q = \max\{c - yA, 0\}.$$

VIII. Обчислити наступні наближення $x^{k+1}$ і $y^{k+1}$:

$$x^{k+1} = (1 - \rho)x^k + \rho\hat{x}(y);$$
$$y^{k+1} = (1 - \rho)y^k + \rho\hat{y}(x)$$

і покласти $x = x^{k+1},\ y = y^{k+1}$.

IX. Якщо $\rho \geq \hat{\rho}$, то перейти на крок X, якщо $\rho < \hat{\rho}$, то перейти на крок XII.

X. Покласти $k = k+1$, $d = d + \rho$ і перейти на крок XI.

XI. Якщо $d \geq l$, то покласти $d = 0$, $\rho = \dfrac{\rho}{2}$ і перейти на крок VII; якщо $d < l$, то перейти на крок VII.

XII. Обчислити значення $\varphi(x,\eta)$, де функція $\varphi(x,\eta)$ задається за правилом

$$\varphi(x,\eta) = (c,x) - (\eta,p),\, p = \max\{Ax - b, 0\} \tag{5.78}$$

і покласти $\varphi = \varphi(x,\eta)$.

XIII. Якщо $\varphi > \varphi_1$, то перейти на крок XIV, інакше перейти на крок XVII.

XIV. Покласти $\varphi_1 = \varphi$, $\bar{x} = x$.

XV. Обчислити вектор

$$\xi(x) = \max\{\sigma x, \tau e^n\}$$

і покласти $\xi = \xi(x)$.



XVI. Обчислити значення $\psi(\bar{y},\xi)$, де функція $\psi(y,\xi)$ задається за правилом

$$\psi(y,\xi) = (y,b) - (q,\xi),\; q = \max\{c - yA, 0\} \tag{5.79}$$

і покласти $\psi_1 = \psi(\bar{y},\xi)$.

XVII. Обчислити значення $\psi(y,\xi)$, де функція $\psi(y,\xi)$ задається за правилом (5.79), і покласти $\psi = \psi(y,\xi)$.

XVIII. Якщо $\psi < \psi_1$, то перейти на крок XIX, інакше перейти на крок XXII.

XIX. Покласти $\psi_1 = \psi, \bar{y} = y$.

XX. Обчислити вектор

$$\eta(y) = \max\{\sigma y, \tau e^m\}$$

і покласти $\eta = \eta(y)$.

XXI. Обчислити значення $\varphi(\bar{x},\eta)$, де функція $\varphi(x,\eta)$ задається за правилом (5.78) і покласти $\varphi_1 = \varphi(\bar{x},\eta)$.

XXII. Обчислити

$$f_k = (\varphi_1 + \psi_1)/2 .$$

XXIII. Обчислити

$$\delta = (\psi_1 - \varphi_1)/(|f_k| + 1).$$

XXIV. Якщо $\delta \le \varepsilon$, то покласти $\overline{x}^k = \overline{x}$, $\overline{y}^k = \overline{y}$ і зупинити обчислення; інакше перейти на крок X.

***Теорема 4.*** *Якщо виконується припущення 4, то точки* $\overline{x}^k$, $\overline{y}^k$ *і значення* $f_k$, *які породжені алгоритмом 4, задовольняють співвідношенням:*

$$\lim_{\varepsilon \to 0} \overline{x}^k = x^*; \;\; \lim_{\varepsilon \to 0} \overline{y}^k = y^*; \;\; \lim_{\varepsilon \to 0} f_k = (c, x^*) = (b, y^*) .$$

### 5. Ітераційний метод, який використовує модифіковану функцію Лагранжа для розв'язання двоїстої пари задач лінійного програмування

З а д а ч а 5. Знайти $\arg\max\limits_x \sum\limits_{j=1}^{n} c_j x_j$ при обмеженнях:

$$\sum_{j=1}^{n} \alpha_{ij} x_j \le b_i , \quad i = \overline{1,m} ; \quad x_j \ge 0 , \quad j = \overline{1,n} .$$

Двоїстою до задачі 5 є наступна.

З а д а ч а 5′. Знайти $\arg\max\limits_y \sum\limits_{i=1}^{m} b_i y_i$

при обмеженнях:

$$\sum_{i=1}^{m} \alpha_{ij} y_i \ge c_j , \quad j = \overline{1,n} ; \quad y_i \ge 0 , \quad i = \overline{1,m} .$$

*Припущення 5.* Множини оптимальних розв'язків $X^*$ і $Y^*$, відповідно, задач 5 і 5′ непорожні і обмежені.

В наведеному нижче алгоритмі знаходження оптимальних розв'язків $X^*$ і $Y^*$задач 5 і 5′ зводиться до обчислення множини сідлових точок модифікованої функції Лагранжа пари двоїстих задач 5 і 5′:

$$\varphi_\alpha(x, y) = \sum_{j=1}^{n} c_j x_j + \sum_{i=1}^{m} b_i y_i - \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{ij} x_j y_i -$$

$$-\sum_{i=1}^{m} \left[ \frac{\alpha}{2} \left( [-\varepsilon_i(x)]_+ \right)^2 + q_\alpha(\varepsilon_i(x)) y_i \right] + \sum_{j=1}^{n} \left[ \frac{\alpha}{2} \left( [-\Delta_j(y)]_+ \right)^2 + q_\alpha(\Delta_j(y)) x_j \right],$$

де

$$\Delta_j(y)=\sum_{i=1}^{m}\alpha_{ij}y_i-c_j\,,\quad j=\overline{1,n}\,;\quad x=(x_1,...,x_n),\;\; y=(y_1,...,y_m)\,;$$

$$\varepsilon_i(x)=-\sum_{j=1}^{n}\alpha_{ij}x_j+b_i\,,\quad i=\overline{1,m}\,;\quad [u]_+=\begin{cases}0, & u\le 0,\\ u, & u>0;\end{cases}$$

$$q_\alpha(t)=\begin{cases}0, & \text{якщо} \quad t\le 0,\\ \alpha t^2/2, & \text{якщо} \quad 0<t\le 1/\alpha,\\ t-1/2\alpha, & \text{якщо} \quad t>1/\alpha\,;\end{cases}$$

$\alpha>0$ – параметр.

При цьому множина сідлових точок функції $\varphi_\alpha(x,y)$ при довільному $\alpha>0$ складається з множини $X^*\times Y^*$. Сідлові точки функції $\varphi_\alpha(x,y)$ знаходяться градієнтним методом.

Відмітимо, що застосувати градієнтний метод для знаходження сідлових точок звичайної функції Лагранжа $\varphi(x,y)$ у вигляді

$$\varphi(x,y)=\sum_{j=1}^{n}c_jx_j+\sum_{i=1}^{m}b_iy_i-\sum_{i=1}^{m}\sum_{j=1}^{n}\alpha_{ij}x_jy_i$$

в загальному випадку неможливо через нестійкості множини сідлових точок цієї функції (означення стійкості множини сідлових точок наводиться в пункті 3 п.7.6).

***Алгоритм 5***

Початок. I. Вибрати довільне початкове наближення $(x^0,y^0)\in R^n\times R^m$.

II. Вибрати довільний параметр $\alpha>0$.

III. Покласти $k=0$.

IV. Визначити функції $\Delta_j(y)$, $j=\overline{1,n}$ і $\varepsilon_i(x)$, $i=\overline{1,m}$ відповідно за правилами:

$$\Delta_j(y)=\sum_{i=1}^{m}\alpha_{ij}y_i-c_j\,,\quad j=\overline{1,n}\,,\;\bigl(y=(y_1,...,y_m)\bigr);$$

$$\varepsilon_i(x)=-\sum_{j=1}^{n}\alpha_{ij}x_j+b_i\,,\quad i=\overline{1,m}\,,\;\bigl(x=(x_1,...,x_n)\bigr).$$

Основний цикл. V. Обчислити кроковий множник $\rho_k$, який задовольняє умовам теореми 5.

VI. Обчислити вектор $\tilde{y}^k=(\tilde{y}_1^k,...,\tilde{y}_m^k)$ за правилами:

$$\tilde{y}_i^k = \begin{cases} y_i^k - \alpha\varepsilon_i\left(x^k\right), & \text{якщо} \quad \varepsilon_i\left(x^k\right) \le 0; \\ y_i^k\left(1-\alpha\varepsilon_i\left(x^k\right)\right), & \text{якщо} \quad 0 < \varepsilon_i\left(x^k\right) \le 1/\alpha; \quad i = \overline{1,m}. \\ 0, & \text{якщо} \quad \varepsilon_i\left(x^k\right) > 1/\alpha, \end{cases}$$

VII. Обчислити наступне наближення $x^{k+1} = \left(x_1^{k+1},...,x_n^{k+1}\right)$ за правилами:

$$x_j^{k+1} = \begin{cases} \max\left\{x_j^k - \rho_k\Delta_j\left(\tilde{y}^k\right),0\right\}, & \text{якщо} \quad \Delta_j(y^k) \le 0; \\ \max\left\{x_j^k - \rho_k\left[\Delta_j\left(\tilde{y}^k\right) - \frac{\alpha}{2}\left(\Delta_j\left(\tilde{y}^k\right)\right)^2\right],0\right\}, & \text{якщо} \quad \Delta_j(y^k) \in (0;1/\alpha]; \\ \max\left\{x_j^k - \rho_k\left[\Delta_j\left(\tilde{y}^k\right) - \Delta_j\left(y^k\right) + \frac{1}{2\alpha}\right],0\right\}, & \text{якщо} \quad \Delta_j(y^k) > 1/\alpha, \\ j = \overline{1,n}. \end{cases}$$

VIII. Обчислити вектор $\tilde{x}^k = \left(\tilde{x}_1^k,...,\tilde{x}_n^k\right)$ за правилами:

$$\tilde{x}_j^k = \begin{cases} x_j^k - \alpha\Delta_j\left(y^k\right), & \text{якщо} \quad \Delta_j\left(y^k\right) \le 0; \\ x_j^k\left(1-\alpha\Delta_j\left(y^k\right)\right), & \text{якщо} \quad 0 < \Delta_j\left(y^k\right) \le 1/\alpha; \quad j = \overline{1,n}. \\ 0, & \text{якщо} \quad \Delta_j\left(y^k\right) > 1/\alpha, \end{cases}$$

IX. Обчислити наступне наближення $y^{k+1} = \left(y_1^{k+1},...,y_m^{k+1}\right)$ за правилами:

$$y_i^{k+1} = \begin{cases} \max\left\{y_i^k - \rho_k\varepsilon_i\left(\tilde{x}^k\right),0\right\}, & \text{якщо} \quad \varepsilon_i(x^k) \le 0; \\ \max\left\{y_i^k - \rho_k\left[\varepsilon_i\left(\tilde{x}^k\right) - \frac{\alpha}{2}\left(\varepsilon_i\left(x^k\right)\right)^2\right],0\right\}, & \text{якщо} \quad \varepsilon_i(x^k) \in (0;1/\alpha]; \\ \max\left\{y_i^k - \rho_k\left[\varepsilon_i\left(\tilde{x}^k\right) - \varepsilon_i\left(x^k\right) + \frac{1}{2\alpha}\right],0\right\}, & \text{якщо} \quad \varepsilon_i(x^k) > 1/\alpha, \\ i = \overline{1,m}. \end{cases}$$

X. Покласти $k = k+1$ і перейти на крок V.

***Теорема 5.*** *Нехай виконується припущення 5. Тоді існує така константа* $\rho > 0$*, що якщо крокові множники* $\rho_k$ *в алгоритмі 5 вибирати згідно умов:*

$$0 < \rho_k < \rho, \quad k = 0, 1, ...; \quad \lim_{k\to\infty}\rho_k = 0; \quad \lim_{k\to\infty}\sum_{t=0}^{k}\rho_t = \infty,$$

*то послідовності* $\left\{x^k\right\}_{k=0}^{\infty}$ *і* $\left\{y^k\right\}_{k=0}^{\infty}$, *які породженні алгоритмом 5, при довільному* $\alpha > 0$ *збігаються, відповідно, до* $X^*$ *і* $Y^*$.

## 5.14. Методи параметричного програмування

**1. Випадок наявності параметра в цільовій функції**

Задача 1. Знайти $\arg\max_x \sum_{j=1}^{n}(c_j^1 + \lambda c_j^2)x_j$ для заданих чисел $c_j^1$, $c_j^2$, $j = \overline{1,n}$ при обмеженнях:

$$Ax = a^0;\ \ x \geq 0,$$

де $\lambda -$ параметр.

*Припущення 1.* $(\boldsymbol{i})-$ множина обмежень задачі 1 – непорожня; $(\boldsymbol{ii})-$ ранг матриці $A$ дорівнює $m$; $(\boldsymbol{iii})-$ $n > m.$

*Означення 1. Базис опорного розв’язку задачі 1 оптимальний* для деякого значення $\lambda$, якщо оцінки $\Delta_j$, всіх векторів $a^j$, $j = \overline{1,n}$ відносно цього базису, обчислені при даному $\lambda$, невід’ємні. Вся сукупність значень параметра $\lambda$, при якому базис оптимальний, називається ***множиною оптимальності*** цього базису.



Наведений нижче алгоритм в невиродженому випадку вказує (за скінченне число ітерацій), при яких значеннях $\lambda \in (-\infty;\infty)$ задача 1 нерозв’язна і при яких – має розв'язок, причому, у випадку розв’язності обчислюється оптимальний розв’язок і його базис. В цьому алгоритмі потрібно розв’язувати задачу 1 симплекс-методом при одному або декількох значеннях параметра $\lambda$, проте початковий опорний розв’язок і його базис обчислюються один раз. У випадку виродженості задачі лінійного програмування в алгоритмі 1 теоретично можливе зациклювання (тоді потрібно застосовувати спеціальне правило, яке гарантує від зациклювання в симплекс-методі).

***Алгоритм 1***

Початок. I. Вибрати довільне число $\lambda_0$.

II. Покласти $t = 0$.

Основний цикл. III. Покласти $\lambda = \lambda_t$.

IV. Покласти $k = 0$. Застосувати для розв’язування задачі 1 симплекс-метод (якщо $t > 0$, то початковий опорний розв'язок і його базис для розв’язування задачі 1 при $\lambda = \lambda_t$ відомі з попередніх ітерацій).

Якщо задача 1 при $\lambda = \lambda_t$ має розв’язок, то перейти на крок V (при цьому вважаються відомими: оптимальний базис $a^{i_1},...,a^{i_m}$; оптимальний розв'язок $(x_1^*,...,x_n^*)^T$; коефіцієнти розкладу векторів $a^0,a^1,...,a^n$ по оптимальному базису, тобто числа $z_{lj}$, $j=\overline{0,n}$; $l=\overline{1,m}$).

Якщо задача 1 при $\lambda=\lambda_t$ немає розв'язку (тобто цільова функція задачі 1 необмежена), то перейти на крок XXVIII (у цьому випадку відомі деякий опорний розв'язок задачі і його базис $a^{i_1},...,a^{i_m}$; коефіцієнти розкладу $z_{lj}$, і, крім того, індекс $j_t\in[1:n]$ є таким, що $\Delta_{j_t}<0$ і при всіх $l\in[1:m]$ виконується $z_{lj_t}\le 0$).

V. Для кожного $j\in[1:n]$ обчислити складові оцінок $\Delta_j$ векторів $a^j$, $j=\overline{1,n}$, відносно оптимального базису $a^{i_1},...,a^{i_m}$ задачі 1 при $\lambda=\lambda_t$:

$$\Delta_j=\Delta_j^1+\lambda_t\Delta_j^2; \tag{5.80}$$

$$\Delta_j^1=\sum_{l=1}^m c_{i_l}^1 z_{lj}-c_j^1,\quad j=\overline{1,n}; \tag{5.81}$$

$$\Delta_j^2=\sum_{l=1}^m c_{i_l}^2 z_{lj}-c_j^2,\quad j=\overline{1,n}. \tag{5.82}$$

VI. Обчислити значення $\underline{\lambda}_k$ за правилом

$$\underline{\lambda}_k=\begin{cases}-\infty, & \text{якщо всі}\quad \Delta_j^2\le 0;\\ \max\limits_{\Delta_j^2>0}(-\Delta_j^1/\Delta_j^2), & \text{в іншому випадку.}\end{cases} \tag{5.83}$$

VII. Обчислити значення $\overline{\lambda}_k$ за правилом

$$\overline{\lambda}_k=\begin{cases}+\infty, & \text{якщо всі}\quad \Delta_j^2\ge 0;\\ \min\limits_{\Delta_j^2<0}(-\Delta_j^1/\Delta_j^2), & \text{в іншому випадку.}\end{cases} \tag{5.84}$$

VIII. Роздрукувати: «Множина оптимальності базису $a^{i_1},...,a^{i_m}$ складається із всіх значень $\lambda$, які задовольняють умові $\underline{\lambda}_k\le\lambda\le\overline{\lambda}_k$ ».

Якщо $\overline{\lambda}_k=\infty$ та $\underline{\lambda}_k=-\infty$, то припинити обчислення; інакше перейти на крок IX.

IX. Якщо $k=0$, то покласти:

$$\tilde{a}^{i_1}=a^{i_1},\ \tilde{a}^{i_2}=a^{i_2},\ \ldots,\ \tilde{a}^{i_m}=a^{i_m};\quad \tilde{\Delta}_j^1=\Delta_j^1,\quad j=\overline{1,n};$$

$$\tilde{\Delta}_j^2=\Delta_j^2,\quad j=\overline{1,n};\ \tilde{\underline{\lambda}}_0=\underline{\lambda}_0;\quad \tilde{z}_{lj}=z_{lj},\ \ l=\overline{1,m};\quad j=\overline{0,n},$$

і перейти на крок X; інакше перейти на крок X.

X. Якщо $\overline{\lambda}_k=\infty$, то перейти на крок XVIII; інакше обчислити такий індекс $s\in[1:n]$, що

$$-\Delta_s^1/\Delta_s^2=\overline{\lambda}_k;\quad \Delta_s^2<0 \tag{5.85}$$

і перейти на крок XI.

XI. Якщо при кожному $l \in [1:m]$ виконується $z_{ls} \le 0$, то видати на друк: «Лінійна форма задачі 1 для $\lambda > \overline{\lambda}_k$ необмежена в допустимій області» і перейти на крок XVIII; інакше перейти на крок XII.

XII. Обчислити індекс $r \in [1:m]$, який задовольняє умові

$$z_{r0} / z_{rs} = \min_{z_{ls}>0}(z_{l0} / z_{ls}), \quad z_{rs} > 0. \tag{5.86}$$

XIII. Перейти до нового базису, який отримується заміною вектора $a^{i_r}$ в попередньому базисі вектором $a^s$ (тобто покласти $i_r = s$).

XIV. Покласти $\overline{z}_{lj} = z_{lj}$, $l = \overline{1,m}$; $j = \overline{0,n}$; $\overline{\Delta}_j^1 = \Delta_j^1$, $\overline{\Delta}_j^2 = \Delta_j^2$, $j = \overline{1,n}$.

XV. Обчислити координати всіх векторів $a^0, a^1, ...., a^n$ у новому базисі за основними формулами:

при $l \ne r$

$$z_{lj} = \overline{z}_{lj} - (\overline{z}_{rj} / \overline{z}_{rs})\overline{z}_{ls}, \quad j = \overline{0,n}; \; l = 1, \,...,\, r-1,\, r+1, \,...,\, m, \tag{5.87}$$

при $l = r$

$$z_{lj} = \overline{z}_{rj} / \overline{z}_{rs}, \quad j = \overline{0,n}. \tag{5.88}$$

XVI. Обчислити складові оцінок $\Delta_j$ векторів $a^j$, $j = \overline{1,n}$ щодо нового базису $a^{i_1}, ..., a^{i_m}$:

$$\Delta_j^1 = \overline{\Delta}_j^1 - \left(\overline{z}_{rj} / \overline{z}_{rs}\right)\overline{\Delta}_s^1, \qquad j = \overline{1,n}; \tag{5.89}$$

$$\Delta_j^2 = \overline{\Delta}_j^2 - \left(\overline{z}_{rj} / \overline{z}_{rs}\right)\overline{\Delta}_s^2, \qquad j = \overline{1,n}. \tag{5.90}$$

XVII. Покласти $k = k+1$ і перейти на крок VI.

XVIII. Якщо $\tilde{\lambda}_0 = -\infty$, то припинити обчислення; інакше покласти $k = 0$ і перейти на крок XIX.

XIX. Перейти до початкового оптимального базису, початкових оцінок $\Delta_j^1$, $\Delta_j^2$ та початкових коефіцієнтів розкладу $z_{lj}$, тобто покласти

$$a^{i_1} = \tilde{a}^{i_1},\; a^{i_2} = \tilde{a}^{i_2},\, ...,\; a^{i_m} = \tilde{a}^{i_m};\; \Delta_j^1 = \tilde{\Delta}_j^1,\; \Delta_j^2 = \tilde{\Delta}_j^2, \quad j = \overline{1,n};$$

$$\underline{\lambda}_0 = \underline{\tilde{\lambda}}_0;\; z_{lj} = \tilde{z}_{lj},\; l = \overline{1,m}; \quad j = \overline{0,n}.$$

XX. Обчислити індекс $s \in [1:n]$ такий, що

$$-\Delta_s^1 / \Delta_s^2 = \underline{\lambda}_k; \qquad \Delta_s^2 > 0.$$

XXI. Якщо при кожному $l \in [1:m]$ виконується $z_{ls} \le 0$, то видати на друк «Лінійна форма задачі 1 для $\lambda < \underline{\lambda}_k$ необмежена у допустимій області», і припинити обчислення; інакше перейти на крок XXII.

XXII. Обчислити індекс $r \in [1:m]$, який задовольняє умові (5.86).

XXIII. Перейти до нового базису, який отримується заміною вектора $a^{i_r}$ в попередньому базисі вектором $a^s$ (тобто, покласти $i_r = s$).

XXIV. Покласти:

$$\overline{z}_{lj} = z_{lj}, \; l = \overline{1,m}; \quad j = \overline{0,n};$$

$$\overline{\Delta}_j^1 = \Delta_j^1, \; \overline{\Delta}_j^2 = \Delta_j^2, \quad j = \overline{1,n}.$$

XXV. Обчислити величини $z_{lj}, \; l = \overline{1,m}, \; j = \overline{0,n};$ $\Delta_j^1, \; j = \overline{1,n};$ $\Delta_j^2, \; j = \overline{1,n},$ відповідно за (5.87) - (5.90).

XXVI. Обчислити значення $\underline{\lambda}_{k+1}$ та $\overline{\lambda}_{k+1}$ за формулами (5.83) і (5.84).

Видати на друк: «Множина оптимальності базису $a^{i_1},...a^{i_m}$ складається із всіх значень $\lambda$, які задовольняють умові

$$\underline{\lambda}_{k+1} \le \lambda \le \overline{\lambda}_{k+1}$$ ».

Якщо $\underline{\lambda}_{k+1} = -\infty$, то припинити обчислення; інакше перейти на крок XXVII.

XXVII. Покласти $k = k + 1$ і перейти на крок XX.

XXVIII. Обчислити складові оцінки $\Delta_{j_t}$ :

$$\Delta_{j_t} = \Delta_{j_t}^1 + \lambda_t \Delta_{j_t}^2;$$

$$\Delta_{j_t}^1 = \sum_{l=1}^{m} c_{i_l}^1 z_{lj_t} - c_{j_t}^1; \qquad \Delta_{j_t}^2 = \sum_{l=1}^{m} c_{i_l}^2 z_{lj_t} - c_{j_t}^2$$

(індекси $i_l$, $l = \overline{1,m}$ і $j_t$ визначені на кроці IV ).

XXIX. Якщо $\Delta^2{}_{j_t} = 0$, то видати на друк: «Задача 1 немає розв’язку при всіх $\lambda \in (-\infty;\infty)$» і припинити обчислення; інакше перейти на крок XXX.

XXX. Обчислити

$$\lambda_{t+1} = -\left(\Delta_{j_t}^1 \big/ \Delta_{j_t}^2\right).$$

Якщо $\Delta^2{}_{j_t} > 0$, то видати на друк: «Задача 1 немає розв'язку при $\lambda < \lambda_{t+1}$» і перейти на крок XXXI для аналізу задачі при $\lambda \ge \lambda_{t+1}$.

Якщо $\Delta^2{}_{j_t} < 0$, то видати на друк: «Задача 1 немає розв'язку при $\lambda > \lambda_{t+1}$» і перейти на крок XXXII для аналізу задачі при $\lambda \le \lambda_{t+1}$.

XXXI. Якщо на попередніх ітераціях був проведений аналіз задачі 1 для $\lambda \ge \lambda_{t+1}$, то припинити обчислення; інакше покласти $t = t + 1$ і перейти на крок III.

XXXII. Якщо на попередніх ітераціях був проведений аналіз задачі 1 для $\lambda \le \lambda_{t+1}$, то припинити обчислення; інакше покласти $t = t + 1$ і перейти на крок III.

***Теорема 1.*** *Якщо виконані припущення 1, то у невиродженому випадку алгоритм 1 за скінченне число ітерацій розбиває числову вісь* $\lambda$ *на області, у яких задача 1 або немає розв'язку (внаслідок необмеженості цільової функції), або має оптимальний розв'язок (причому відомі оптимальний базис і оптимальний опорний розв'язок).*

**2. Випадок наявності параметра в правих частинах обмежень**

Задача 2. Знайти $\arg\max_x (c, x)$ для заданого вектора $c \in R^n$ при обмеженнях

$$\sum_{j=1}^{n} \alpha_{ij} x_j = b_i^1 + \lambda b_i^2, \quad i = \overline{1,m}; \tag{5.91}$$

$$x_j \geq 0, \quad j = \overline{1,n}, \tag{5.92}$$

де $\lambda$ – параметр.

*Припущення 2.* $(\boldsymbol{i})-$ цільова функція задачі 2 обмежена на допустимій множині при всіх $\lambda$; $(\boldsymbol{ii})-$ ранг матриці $A$ дорівнює $m$; $(\boldsymbol{iii})-$ $n > m$.

Якщо при деякому значенні параметра $\lambda$ коефіцієнти розкладу вектора $b = b^1 + \lambda b^2$ (тут $b^1 = \left(b_1^1, ..., b_m^1\right)^T$, $b^2 = \left(b_1^2, ..., b_m^2\right)^T$) по векторах базису спряженої задачі невід'ємні, то цей базис є оптимальним базисом задачі 2 для даного $\lambda$.

*Означення 2.* Вся сукупність значень параметра $\lambda$, при яких базис оптимальний, називається ***множиною оптимальності*** цього базису.

Наведений нижче алгоритм розбиває числову вісь $\lambda$ на області, в кожній з яких обчислюється оптимальний базис, що не змінюється зі зміною $\lambda$, а також виділяє область (якщо вона існує), в якій обмеження задачі 2 несумісні. Алгоритм базується на застосуванні двоїстого симплекс-методу для розв'язування задачі 2 при одному або декількох значеннях параметра $\lambda$. Алгоритм 2 застосовується у невиродженому та виродженому випадках. У виродженому випадку теоретично можливе зациклювання. Якщо зациклювання відбудеться, то необхідно застосовувати спеціальне правило, що гарантує від зациклювання у двоїстому симплекс-методі.

***Алгоритм 2***

Початок. I. Вибрати довільне число $\lambda_0$.

II. Покласти $t = 0$.

Основний цикл. III. Покласти $\lambda = \lambda_t$.

IV. Покласти $k = 0$. Застосувати до розв'язування задачі 2 двоїстий симплекс-метод.

Якщо в результаті знаходиться оптимальний розв'язок задачі 2 і його базис (позначити оптимальний розв'язок через

$\left(x_1\left(\lambda_t\right),\ x_2\left(\lambda_t\right), ..., x_n\left(\lambda_t\right)\right)^T$, його базис через $a^{i_1}, ..., a^{i_m}$, оцінки через $\Delta_j$, $j=\overline{1,n}$, а коефіцієнти розкладу векторів $a^j$, $j=\overline{1,n}$ по базисних векторах через $z_{lj}$, $l=\overline{1,m}$, $j=\overline{1,n}$), то перейти на крок V; якщо в результаті виявиться несумісність умов (5.91) і (5.92) при $\lambda=\lambda_t$, то перейти на крок XXXI.

У цьому випадку, відповідно до ознаки несумісності (див. крок VI алгоритму 1, п.5.8) існує таке $l_t \in [1:m]$, що $z_{l_t 0} < 0$ і при всіх $j \in [1:n]$ виконується $z_{l_t j} \geq 0$, де $z_{l0}$, $l=\overline{1,m}$ – коефіцієнти розкладу вектора $b^1 + \lambda_t b^2$ через базисні вектори $a^{i_1}, ..., a^{i_m}$ спряженої задачі.

V. Розв'язати систему рівнянь

$$b^1 = \sum_{l=1}^{m} a^{i_l} x_{i_l}^1\left(\lambda_t\right)$$

відносно $x_{i_l}^1(\lambda_t)$.

VI. Розв'язати систему рівнянь

$$b^2 = \sum_{l=1}^{m} a^{i_l} x_{i_l}^2\left(\lambda_t\right)$$

відносно $x_{i_l}^2\left(\lambda_t\right)$.

VII. Покласти $z_{l0}^1 = x_{i_l}^1\left(\lambda_t\right)$; $z_{l0}^2 = x_{i_l}^2\left(\lambda_t\right)$, $l=\overline{1,m}$.

VIII. Обчислити $\underline{\lambda}_k$ за правилами:

$$\underline{\lambda}_k = \begin{cases} -\infty, & \text{якщо всі } z_{l0}^2 \leq 0, \\ \max\limits_{z_{l0}^2>0}(-z_{l0}^1 / z_{l0}^2), & \text{якщо існує } z_{l0}^2 > 0. \end{cases} \quad l=\overline{1,m}, \qquad 5.93)$$

IX. Обчислити значення $\overline{\lambda}_k$ за правилами:

$$\overline{\lambda}_k = \begin{cases} +\infty, & \text{якщо всі } z_{l0}^2 \geq 0, \\ \min\limits_{z_{l0}^2<0}(-z_{l0}^1 / z_{l0}^2), & \text{якщо існує } z_{l0}^2 < 0, \end{cases} \quad l=\overline{1,m}. \qquad (5.94)$$

X. Видати на друк: «Множина оптимальності базису $a^{i_1}, a^{i_2}, ..., a^{i_m}$ складається із всіх значень $\lambda$, які задовольняють умові $\underline{\lambda}_k \leq \lambda \leq \overline{\lambda}_k$ » і перейти на крок XI.

XI. Якщо $\underline{\lambda}_k = -\infty$ та $\overline{\lambda}_k = +\infty$, то припинити обчислення; інакше перейти на крок XII.

XII. Якщо $k=0$, то покласти:

$$\tilde{a}^{i_1} = a^{i_1}, ..., \tilde{a}^{i_m} = a^{i_m};\ \tilde{\Delta}_j = \Delta_j,\ j=\overline{1,n};\ \underline{\tilde{\lambda}}_0 = \underline{\lambda}_0;$$

$$\tilde{z}_{lj} = z_{lj},\ l=\overline{1,m};\ \ j=\overline{1,n};\ \tilde{z}_{l0}^1 = z_{l0}^1,\ \ \tilde{z}_{l0}^2 = z_{l0}^2,\ \ l=\overline{1,m},$$

і перейти на крок XIII; інакше перейти на крок XIII.

XIII. Якщо $\bar{\lambda}_k = +\infty$, то перейти на крок XXI; інакше обчислити індекс $r \in [1:m]$ такий, що

$$-z_{r0}^1 / z_{r0}^2 = \bar{\lambda}_k; \quad z_{r0}^2 < 0.$$

XIV. Обчислити індекс $s \in [1:n]$ такий, що

$$-\Delta_s / z_{rs} = \min_{z_{rj}<0}(-\Delta_j / z_{rj}). \qquad (5.95)$$

XV. Покласти:

$$\bar{z}_{lj} = z_{lj}, \ l = \overline{1,m}; \ j = \overline{0,n}; \ \bar{\Delta}_j = \Delta_j, \ \ j = \overline{1,n};$$

$$\bar{z}_{l0}^1 = z_{l0}^1, \quad \bar{z}_{l0}^2 = z_{l0}^2, \ \ l = \overline{1,m}.$$

XVI. Перейти до нового базису $a^{i_1},...,a^{i_m}$, який отримується заміною вектора $a^{i_r}$ в попередньому базисі вектором $a^s$ (тобто покласти $i_r = s$).

XVII. Обчислити координати всіх векторів $a^1$, ..., $a^n$ у новому базисі за основними формулами:

при $l \neq r$

$$z_{lj} = \bar{z}_{lj} - \left(\bar{z}_{rj} / \bar{z}_{rs}\right)\bar{z}_{ls}, \ \ j = \overline{1,n}; \ l = 1, ..., r-1, r+1, ..., m; \qquad (5.96)$$

при $l = r$

$$z_{lj} = \bar{z}_{rj} / \bar{z}_{rs}, \ \ j = \overline{1,n}. \qquad (5.97)$$

XVIII. Обчислити оцінки

$$\Delta_j = \bar{\Delta}_j - \left(\bar{z}_{rj} / \bar{z}_{rs}\right)\bar{\Delta}_s, \quad j = \overline{1,n}. \qquad (5.98)$$

XIX. Обчислити величини $z_{l0}^1$ та $z_{l0}^2$, $l = \overline{1,m}$, за основними формулами:

$$z_{l0}^1 = \bar{z}_{l0}^1 - \left(\bar{z}_{r0}^1 / \bar{z}_{rs}\right)\bar{z}_{ls}, \ \ l = 1, ..., r-1, r+1, ..., m; \ z_{r0}^1 = \bar{z}_{r0}^1 / \bar{z}_{rs}; \qquad (5.99)$$

$$z_{l0}^2 = \bar{z}_{l0}^2 - \left(\bar{z}_{r0}^2 / \bar{z}_{rs}\right)\bar{z}_{ls}, \ \ l = 1, ..., r-1, r+1, ..., m; \ z_{r0}^2 = \bar{z}_{r0}^2 / \bar{z}_{rs}. \qquad (5.100)$$

XX. Покласти $k = k+1$ і перейти на крок VIII.

XXI. Якщо $\tilde{\underline{\lambda}}_0 = -\infty$, то припинити обчислення; інакше покласти $k = 0$ і перейти на крок XXII.

XXII. Перейти до початкового оптимального базису, початкових оцінок, початкових коефіцієнтів розкладу, тобто покласти:

$$a^{i_1} = \tilde{a}^{i_1}, \ \ a^{i_2} = \tilde{a}^{i_2}, ..., \ a^{i_m} = \tilde{a}^{i_m}; \ \Delta_j = \tilde{\Delta}_j, \ \ j = \overline{1,n}; \ \ \underline{\lambda}_0 = \tilde{\underline{\lambda}}_0;$$

$$z_{lj} = \tilde{z}_{lj}, \quad j = \overline{1,n}; \quad l = \overline{1,m}; \ \ z_{l0}^1 = \tilde{z}_{l0}^1, \quad z_{l0}^2 = \tilde{z}_{l0}^2, \quad l = \overline{1,m}.$$

XXIII. Обчислити індекс $r \in [1:m]$ такий, що

$$-z_{r0}^1 / z_{r0}^2 = \underline{\lambda}_k \ \text{ і } \ z_{r0}^2 > 0.$$

XXIV. Обчислити індекс $s \in [1:n]$, який задовольняє умові (5.95).

XXV. Покласти:

$$\bar{z}_{lj} = z_{lj},\ l = \overline{1,m},\quad j = \overline{1,n};\quad \bar{\Delta}_j = \Delta_j,\quad j = \overline{1,n};$$

$$\bar{z}_{l0}^1 = z_{l0}^1,\quad \bar{z}_{l0}^2 = z_{l0}^2,\quad l = \overline{1,m}.$$

XXVI. Перейти до нового базису, який отримується заміною вектора $a^{i_r}$ попереднього базису вектором $a^s$ (тобто покласти $i_r = s$).

XXVII. Обчислити величини:

$$z_{lj},\quad l = \overline{1,m};\ j = \overline{1,n};\quad \Delta_j,\quad j = \overline{1,n};\quad z_{l0}^1,\ l = \overline{1,m};\ z_{l0}^2,\ l = \overline{1,m},$$

відповідно за (5.96) - (5.100).

XXVIII. Обчислити значення $\underline{\lambda}_{k+1}$ та $\bar{\lambda}_{k+1}$ за (5.93) і (5.94), відповідно. Видати на друк: «Множина оптимальності базису $a^{i_1}, a^{i_2}, ..., a^{i_m}$ складається із всіх значень $\lambda$, які задовольняють умові $\underline{\lambda}_{k+1} \le \lambda \le \bar{\lambda}_{k+1}$ ».

XXIX. Якщо $\underline{\lambda}_{k+1} = -\infty$, то припинити обчислення; інакше перейти на крок XXX.

XXX. Покласти $k = k + 1$ і перейти на крок XXXIII.

XXXI. Обчислити величини $z_{l0}^1,\ z_{l0}^2,\ l = \overline{1,m},$ за тими правилами, що і на кроках V, VI, VII.

XXXII. Якщо $z_{l_t 0}^2 = 0$, то видати на друк: «Обмеження задачі 2 несумісні при всіх $\lambda \in (-\infty; \infty)$» і припинити обчислення; інакше перейти на крок XXXIII.

XXXIII. Обчислити

$$\lambda_{t+1} = -\left(z_{l_t 0}^1 / z_{l_t 0}^2\right).$$

Якщо $z_{l_t 0}^2 > 0$, то видати на друк: «Обмеження задачі 2 несумісні при всіх $\lambda < \lambda_{t+1}$» і перейти на крок XXXIV для аналізу задачі при $\lambda \ge \lambda_{t+1}$; якщо $z_{l_t 0}^2 < 0$, то видати на друк: «Обмеження задачі 2 несумісні при всіх $\lambda > \lambda_{t+1}$» і перейти на крок XXXV для аналізу задачі при $\lambda \le \lambda_{t+1}$.

XXXIV. Якщо на попередніх ітераціях був проведений аналіз задачі 2 для $\lambda \ge \lambda_{t+1}$, то припинити обчислення; інакше покласти $t = t + 1$ і перейти на крок III.

XXXV. Якщо на попередніх ітераціях був проведений аналіз задачі 2 для $\lambda \le \lambda_{t+1}$, то припинити обчислення; інакше покласти $t = t + 1$ і перейти на крок III.

***Теорема 2.*** *Якщо виконані припущення 2, то у невиродженому випадку алгоритм 2 за скінченне число ітерацій розбиває числову вісь $\lambda$ на області, у яких або задача 2 має оптимальний базис і оптимальний розв'язок, або обмеження задачі 2 несумісні.*

## Запитання та вправи для самостійної роботи до розділу 5.

1. Сформулюйте загальну, стандартну та канонічну задачі лінійного програмування.
2. Які Ви знаєте способи переходу від загальної до стандартної задачі лінійного програмування?
3. Які Ви знаєте способи переходу від стандартної до канонічної задачі лінійного програмування?
4. Сформулюйте означення допустимого розв’язку задачі лінійного програмування.
5. Сформулюйте означення півпростору та гіперплощини в $R^n$.
6. Сформулюйте означення многогранної множини та многогранника.
7. Яка точка $x \in X$ називається кутовою (крайньою)?
8. Які точки опуклої многогранної множини $X$ називаються її вершинами?
9. Де потрібно шукати розв’язки загальної задачі лінійного програмування?
10. Поясніть, які змінні в задачах лінійного програмування називають залишковими, надлишковими, вільними.
11. Який допустимий розв’язок називають базисним (опорним) розв’язком стандартної задачі лінійного програмування?
12. Що називають базисом опорного (базисного) розв’язку?
13. Яку матрицю називають базисною?
14. Який базисний розв’язок називають невиродженим?
15. Який базисний розв’язок називають виродженим?
16. Яка стандартна задача лінійного програмування називається невиродженою, а яка виродженою?
17. Які компоненти опорного (базисного) розв’язку називають базисними, а які небазисними?
18. Який вектор $x^0$ вибирається у якості початкового в стандартному симплекс-методі?
19. Запишіть формули для обчислення оцінок $\Delta_j$, $j = \overline{1,n}$ в симплекс-методі на першій ітерації і на наступних ітераціях.
20. При виконанні яких умов цільова функція $(c, x)$ необмежена на допустимій множині $X$ ?
21. Запишіть умову, при якій симплекс-метод завершує обчислення знаходженням оптимального розв’язку $x^*$ задачі лінійного програмування.
22. Чи можна використовувати симплекс-метод для розв’язування вироджених стандартних задач лінійного програмування?
23. Назвіть переваги модифікованого симплекс-методу (або методу базисної матриці).

24. Який симплекс-метод називають мультиплікативним?
25. Сформулюйте пряму стандартну задачу лінійного програмування і двоїсту (спряжену) до неї.
26. Який розв’язок двоїстої задачі називають опорним?
27. Що називають базисом опорного розв’язку двоїстої задачі?
28. Сформулюйте означення невиродженого опорного розв’язку і невиродженої двоїстої задачі.
29. Охарактеризуйте сутність двоїстого симплекс-методу.
30. При виконанні яких умов двоїстий симплекс-метод: а) знаходить оптимальний розв’язок $x^*$ прямої задачі; б) встановлює, що пряма задача не має допустимих розв’язків?
31. В якому випадку рекомендується використовувати двоїстий симплекс-метод?
32. Опишіть сутність методів послідовного скорочення нев’язок і двосторонніх оцінок.
33. Назвіть переваги методів послідовного скорочення нев’язок.
34. В чому полягає специфіка методів блочного програмування?
35. В чому полягає сутність ітераційних методів розв’язувння задач лінійного програмування?
36. В яких випадках рекомендується використовувати ітераційні методи?
37. Яку допоміжну задачу потрібно розв’язувати на кожній ітерації ітераційного методу, який використовує модифіковану функцію Лагранжа?
38. Сформулюйте задачі параметричного програмування з параметром в цільовій функції та з параметром в правій частині обмежень.
39. Складіть математичну модель задачі оптимального планування виробництва. На підприємстві виготовляється $n$ типів продукції з $m$ типів сировини. Відомо, що на виготовлення одиниці продукції $j$-го типу потрібно $\alpha_{ij}$ одиниць сировини $i$-го типу. Підприємство має в своєму розпорядженні $b_i$, $i=\overline{1,m}$ одиниць сировини $i$-го типу. Відомо також, що з кожної одиниці продукції $j$-го типу $(j=\overline{1,n})$ підприємство отримує $c_j$ одиниць прибутку.

    Потрібно дізнатися скільки одиниць $x_1, x_2, ..., x_n$ кожного типу продукції повинно виготовити підприємство, щоб забезпечити собі максимальний прибуток.
40. Складіть математичну модель задачі про оптимальне використання посівної площі. Нехай під посів $p$ культур відведено $r$ земельних ділянок площею, відповідно, $b_1, b_2, ..., b_r$ гектарів. Відомо, що середня урожайність $i$-ї культури $(i=\overline{1,p})$ на $j$-й ділянці $(j=\overline{1,r})$

складає $\alpha_{ij}$ центнерів з гектара, а прибуток за один центнер $i$-ї культури складає $c_i$ гривень. Потрібно визначити, яку площу на кожній ділянці потрібно відвести під кожну культуру, щоб одержати максимальний прибуток, якщо в сукупності треба зібрати не менше $d_i$ центнерів $i$-ї культури.

41. Спираючись на геометричну інтерпретацію задач лінійного програмування, розв'яжіть наступні задачі:

а) $L(x_1, x_2) = x_1 + 2x_2 \to \max(\min)$

при обмеженнях

$x_1 - x_2 \le 1,\ \ 2x_1 + 3x_2 \le 6,\ \ x_1 \ge 0,\ \ x_2 \ge 0;$

б) $L(x_1, x_2) = 2x_1 - x_2 \to \max(\min)$

при обмеженнях

$3x_1 + 4x_2 \le 9,\ \ 2x_1 - x_2 \le 4,\ x_1 \ge 0,\ x_2 \ge 0;$

в) $L(x_1, x_2) = -3x_1 - 4x_2 \to \max(\min)$

при обмеженнях

$1{,}5x_1 - 2{,}5x_2 \le 5,\ \ 2x_1 + 4x_2 \le 7,\ \ x_1 \le 3,\ x_1 \ge 0,\ x_2 \ge 0.$

42. Розв'яжіть графічно наступні задачі лінійного програмування попередньо звівши їх до розмірності два:

а) $L(x_1, x_2, x_3) = 2x_1 + x_2 + x_3 \to \max(\min)$

при обмеженнях:

$x_1 + x_2 + x_3 = 4,\ \ x_1 - x_2 + x_3 \le 2,\ \ x_1 \ge 0,\ \ x_2 \ge 0,\ \ x_3 \ge 0;$

б) $L(x_1, x_2, x_3) = -3x_2 - 3x_3 \to \max(\min)$

при обмеженнях:

$x_1 - 3x_2 + x_3 = 5,\ \ x_2 - x_3 = 2,\ x_1 \ge 0,\ x_2 \ge 0,\ \ x_3 \ge 0;$

в) $L(x_1, x_2) = -x_1 - 2x_2 - 3{,}5x_3 - 4{,}5x_4 \to \max(\min)$

при обмеженнях:

$2x_1 + x_2 + x_4 = 6,\ \ x_2 + 3x_2 = 4,\ \ x_1 - x_2 - x_3 + x_4 \le -1,\ \ x_i \ge 0,\ i = \overline{1,4}.$

43. При яких значеннях параметра $\beta$ задача лінійного програмування:

$L(x_1, x_2) = x_1 + \beta x_2 \to \max(\min)$

при обмеженнях:

$x_1 - x_2 \le 1,\ \ x_1 + 2x_2 \ge 4,\ \ x_1 \ge 0,\ \ x_2 \ge 0:$

а) має єдиний розв'язок;

б) має безліч розв'язків;

в) не має розв'язків.

44. Знайдіть всі кутові точки та їх базиси для наступних областей:

$X=\left\{x\in R^4 \middle| \ x_1-2x_2-x_3=0,\ -x_1+3x_2+x_4=1,\ x\ge 0\right\}$;

$X=\left\{x\in R^5 \middle| \ x_1+x_2+x_3+x_4=1,\ -x_1+2x_2+x_3+x_5=1,\ x\ge 0\right\}$;

$X=\left\{x\in R^5 \middle| \ 2x_1+3x_3+x_5=3,\ x_1+x_2+2x_3=2,\ x_1+x_3+x_4=1,\ x\ge 0\right\}$.

45. За допомогою симплекс-методу і його варіантів розв'яжіть наступні задачі лінійного програмування:

а) $L(x)=x_1+2x_2+x_4+x_6\to\max(\min)$

при обмеженнях:

$x_1+2x_2-x_3+2x_4-x_5+2x_6=0,$

$x_1-4x_2+2x_3+2x_4-4x_6=1,$

$2x_1-2x_2+x_3+4x_4+x_5-2x_6=3,\ \ x_i\ge 0,\ i=\overline{1,6};$

б) $L(x)=2x_1+2x_2+x_3+2x_4+3x_5+2x_6\to\max(\min)$

при обмеженнях:

$x_1+x_3+x_4-2x_5+2x_6=-1,$

$x_1+2x_2+4x_4+2x_6=5,$

$2x_2+x_3+x_4-x_5+2x_6=0,\ \ x_i\ge 0,\ i=\overline{1,6};$

в) $L(x)=x_1+x_2+x_4+x_4\to\min(\max)$

при обмеженнях:

$x_1-x_2\ge 0,$

$x_1+x_2-x_3+x_4-x_5\ge 1,\ \ x_i\ge 0,\ i=\overline{1,5};$

г) $L(x)=3x_1+2x_2+x_3+x_6\to\min(\max)$

при обмеженнях:

$x_2-x_5-x_6=-1,$

$2x_1+2x_2+x_3+2x_5+3x_7=0,$

$x_1+x_2+x_4+x_5+2x_7=0,\ \ x_i\ge 0,\ i=\overline{1,7};$

д) $L(x)=3x_1+10x_2+8x_3-6x_4\to\min(\max)$

при обмеженнях:

$x_1+3x_2+3x_3-2x_4=-1,$

$3x_1+2x_2+x_3-x_4\ge -1,$

$x_1\le 4,\ \ x_i\ge 0,\ \ i=\overline{1,4}.$

46. Для сформульованих вище задач 45 а) - г) запишіть двоїсті задачі та розв'яжіть їх симплекс-методами.

**Р о з д і л 6**

# МЕТОДИ РОЗВ'ЯЗУВАННЯ ЗАДАЧ НЕЛІНІЙНОГО І СТОХАСТИЧНОГО ПРОГРАМУВАННЯ

## 6.1. Методи проекції градієнта

### 1. Загальний метод

З а д а ч а 1. Знайти $\arg\min\limits_{x\in X} f_0\left(x\right)$ для заданої функції $f_0 : R^n \to R^1$ і заданої множини $X \subset R^n$.

*Припущення 1.* $(\boldsymbol{i})$ – функція $f_0$ опукла донизу і неперервно диференційована на множині $X$;

$(\boldsymbol{ii})$ – $X$ – опукла замкнута множина;

$(\boldsymbol{iii})$ – $\left\|\nabla f_0\left(x\right)-\nabla f_0\left(y\right)\right\| \le \gamma \left\|x-y\right\|,\ 0<\gamma<\infty,\ \ \forall x,\ y\in X.$

***Алгоритм 1***

П о ч а т о к . I. Вибрати початкове наближення $x^0 \in X$.

II. Покласти $k=0$.

О с н о в н и й ц и к л . III. Обчислити $\nabla f_0\left(x^k\right)$.

IV. Вибрати множник $\beta_k$, який задовольняє одну із завчасно вибраних умов

$$0<\beta' \le \beta_k \le \beta'' < +\infty; \tag{6.1}$$

$$0<\beta' \le \beta_k \le \beta'' < 2/\gamma. \tag{6.2}$$

V. Обчислити точку $y^k \in R^n$ за формулою

$$y^k = x^k - \beta_k \nabla f_0\left(x^k\right).$$

VI. Обчислити точку $y_X^k$ – проекцію точки $y^k$ на опуклу замкнену множину $X$.

VII. Якщо $\beta_k$ вибирається у відповідності із нерівністю (6.2), то вибрати множник $\rho_k$, який задовольняє нерівності

$$0<\hat{\rho}<\rho_k \le 1 \tag{6.3}$$

і перейти на крок VIII; якщо $\beta_k$ вибирається згідно (6.1), то обчислити множник $\rho_k$ за однією із формул:

$$f_0\left(x^k-\rho_k\left(x^k-y_X^k\right)\right)=\min_{\rho\in[0;1]} f_0\left(x^k-\rho\left(x^k-y_X^k\right)\right), \tag{6.4}$$

або

$$\rho_k = \min\left\{1,\ \nu_k \frac{\left(\nabla f_0\left(x^k\right),\ x^k - y_X^k\right)}{\left\|x^k - y_X^k\right\|^2}\right\}, \tag{6.5}$$

де $0 < \varepsilon_1 < \nu_k \le \dfrac{2-\varepsilon_2}{\gamma}$; $\varepsilon_2 > 0$ і перейти на крок VIII.

VIII. Покласти $x^{k+1} = x^k - \rho_k\left(x^k - y_X^k\right)$, $k = k+1$ і перейти на крок III.

*Зауваження 1.* Вважається, що $\rho_k$ на кожній ітерації обчислюється за однією і тією ж формулою.

На кроці VI алгоритму 1 на кожній ітерацій необхідно здійснювати операцію проектування точки $y^k$ на випуклу замкнуту множину $X$. В загальному випадку ця операція є складною задачею. Наведемо частинні випадки множини $X$, коли явно виписується вигляд проекції:

1) нехай $X = \left\{x \in R^n \middle| \alpha_j \le x_j \le \beta_j,\ j = \overline{1,n}\right\}$ – гіперпаралелепіпед, тоді

$$\left(\pi_X(y)\right)_j = \begin{cases} \alpha_j, & \text{якщо } y_j < \alpha_j, \\ y_j, & \text{якщо } \alpha_j \le y_j \le \beta_j, \\ \beta_j, & \text{якщо } y_j > \beta_j, \quad j = \overline{1,n}; \end{cases}$$



2) нехай $X = \left\{x \in R^n \middle| x_j \ge 0,\ j = \overline{1,n}\right\}$ – невід’ємний октант, тоді

$$\pi_X(y) = \left(\max(0; y_1), \ldots, \max(0; y_n)\right)^T;$$

3) нехай $X = \left\{x \in R^n \middle| \| x - a \| \le r\right\}$ – куля, тоді

$$\pi_X(y) = a + \frac{y-a}{\| y-a \|} \cdot r;$$

4) нехай $X = \left\{x \in R^n \middle| (c,x) \ge \alpha;\ c \in R^n,\ c \ne 0\right\}$ – півпростір, тоді

$$\pi_X(y) = y + \max\left(0; \alpha - (c,y)\right)\frac{c}{\| c \|^2};$$

5) нехай $X = \left\{x \in R^n \middle| Ax = b\right\}$ – афінна множина (рядки матриці $A$ лінійно незалежні), тоді

$$\pi_X(y) = y - A^T(A \cdot A^T)^{-1}(Ay - b).$$

***Теорема 1.*** *Нехай виконуються припущення 1 і* $(\mathbf{iv})$ – *множина*

$$X_0 \triangleq \left\{f_0(x) \le f_0\left(x^0\right),\ x \in X\right\}$$

*обмежена;*

*(v) –* $\left\|\nabla f_0(x)\right\| \le \tau < +\infty$ *для всіх* $x \in X_0$.

*Тоді алгоритм 1 породжує таку послідовність* $\left\{x^k\right\}_{k=0}^{\infty}$*, що*

$$f_0\left(x^k\right) - \min_{x\in X} f_0(x) \le \alpha / k,\ \ k = 1, 2, \ldots;\ \ \alpha > 0.$$

*При цьому: а) якщо в алгоритмі 1* $\beta_k$ *задовольняє нерівностям (6.1), а* $\rho_k$ *– рівності (6.4), то справедлива оцінка*

$$f_0\left(x^k\right) - f_0\left(x^{k+1}\right) \ge \left(\left(1/\beta''\right)\bar{\rho} - \left(\gamma/2\right)\bar{\rho}^2\right)\left\|x^k - y_X^k\right\|^2,\ \ k = 0, 1, \ldots,$$

*де* $0 < \bar{\rho} < \min\left\{1; 2/\gamma\beta''\right\}$*;*

*б) якщо в алгоритмі 1* $\beta_k$ *задовольняє нерівностям (6.1), а* $\rho_k$ *обчислюється за (6.5), то справедливі оцінки:*

*– при* $\rho_k = 1$*:*

$$f_0\left(x^k\right) - f_0\left(x^{k+1}\right) \ge \left(\varepsilon_2/2\beta''\right)\left\|x^k - y_X^k\right\|^2,\ \ k = 0, 1, \ldots;$$

*– при* $\rho_k = v_k\left(\nabla f_0\left(x^k\right), x^k - y_X^k\right) \Big/ \left\|x^k - y_X^k\right\|^2$ *:*

$$f_0\left(x^k\right) - f_0\left(x^{k+1}\right) \ge \left(\varepsilon_1\varepsilon_2 \Big/ 2\left(\beta''\right)^2\right)\left\|x^k - y_X^k\right\|^2,\ \ k = 0, 1, \ldots;$$

*в) якщо в алгоритмі 1* $\beta_k$ *задовольняє нерівностям (6.2), а* $\rho_k$ *– нерівності (6.3), то справедлива оцінка*

$$f_0\left(x^k\right) - f_0\left(x^{k+1}\right) \ge \hat{\rho}\left(1/\beta'' - \gamma/2\right)\left\|x^k - y_X^k\right\|^2.$$



### 2. Метод проекції градієнта для мінімізації функції при лінійних обмеженнях

Задача 2. Знайти $\arg\min\limits_{x\in X} f_0(x)$ для заданої функції $f_0 : R^n \to R^1$ і заданої множини

$$X = \left\{x \middle| \left(a^j, x\right) - b_j \le 0,\ \ j = \overline{1, m},\ \ x \in R^n\right\},$$

де $a^j \in R^n$; $b_j \in R^1$; $j = \overline{1, m}$.

*Припущення 2.* Функція $f_0$ – опукла донизу і неперервно диференційована.

Позначення і визначення. 1. Довільну підмножину множини $\{1, \ldots, m\}$, яка містить $m'$ елементів, позначимо через $\mathfrak{I}$.

2. Матрицю розмірності $n \times m'$, стовпцями якої є вектори $a^j$, $j \in \mathfrak{I}$, розміщені в порядку зростання $j$, позначимо через $A_{\mathfrak{I}}$.

3. Матрицю проектування $R^n$ на підпростір $L_{\Im}$, породжений векторами $a^j$, $j \in \Im$, позначимо через $P_{\Im}$.

4. Якщо вектори $a^j$, $j \in \Im$, які визначають простір $L_{\Im}$, лінійно незалежні, то матриця, яка проектує $R^n$ на підпростір $L_{\Im}$, обчислюється за формулою

$$P_{\Im} = A_{\Im}\left(A_{\Im}^T A_{\Im}\right)^{-1} A_{\Im}^T . \tag{6.6}$$

Якщо вектори $a^j$, $j \in \Im$ – лінійно залежні, то необхідно знайти підмножину $\Im'$ таку, що вектори $a^j$, $j \in \Im'$ будуть лінійно незалежні і породжують $L_{\Im}$, а проектуюча матриця $P_{\Im} = P_{\Im'}$ буде визначатися згідно (6.6) при $\Im = \Im'$.

5. Матрицю проектування $R^n$ на ортогональне доповнення до $L_{\Im}$ позначимо через

$$P_{\Im}^{\perp} = I - P_{\Im},$$

де $I$ – одинична $n \times n$-матриця.

6. Позначимо через $\Im_{\varepsilon}\left(x\right)$ $\varepsilon$-активну множину індексів, яка визначається за формулою:

$$\Im_{\varepsilon}\left(x\right) = \left\{ j \middle| \left(a^j, x\right) - b^j + \varepsilon \ge 0,\ j = \overline{1,m},\ x \in X \right\},$$

де $\varepsilon \ge 0$.

*Твердження 2.* Точка $\hat{x} \in X$ оптимальна в задачі 2 тоді і тільки тоді, коли

$$\nabla f_0\left(\hat{x}\right) = A_{\Im_0(\hat{x})} y^0\left(\hat{x}\right), \text{ тобто } P_{\Im_0(\hat{x})}^{\perp} \nabla f_0\left(\hat{x}\right) = 0;\ \ y^0\left(\hat{x}\right) \le 0,$$

де $y^0\left(\hat{x}\right)$ обчислюється за наступною формулою (при $\varepsilon = 0$ і $x = \hat{x}$):

$$y^{\varepsilon}\left(x\right) = \left(A_{\Im_{\varepsilon}(x)}^T A_{\Im_{\varepsilon}(x)}\right)^{-1} A_{\Im_{\varepsilon}(x)}^T \nabla f_0\left(x\right). \tag{6.7}$$

Відмітимо, що матриця $P_{\Im_{\varepsilon}(x)}$, яка проектує $R^n$ на підпростір $L_{\Im_{\varepsilon}(x)}$, натягнутий на вектори $a^j$, $j \in \Im_{\varepsilon}\left(x\right)$, обчислюється за формулою

$$P_{\Im_{\varepsilon}(x)} = A_{\Im_{\varepsilon}(x)}\left(A_{\Im_{\varepsilon}(x)}^T A_{\Im_{\varepsilon}(x)}\right)^{-1} A_{\Im_{\varepsilon}(x)}^T,$$

якщо виконується припущення $2'$.

*Припущення* $2'$. Існує таке число $\varepsilon' > 0$, що для будь-якої точки $x \in X$ і будь-якого числа $\varepsilon \in [0;\varepsilon']$ вектори $a^j$, $j \in \Im_{\varepsilon}\left(x\right)$ лінійно незалежні.

***Алгоритм 2***

П о ч а т о к. І. Обчислити $x^0 \in X$; вибрати $\beta \in \left(0;1\right)$ (рекомендується вибирати $\beta = 1/2$), $\bar{\varepsilon} \in \left(0;\varepsilon'\right)$ і $\varepsilon'' \in \left(0;\bar{\varepsilon}\right)$; покласти $k = 0$.

О с н о в н и й  ц и к л.  II. Покласти $x = x^k$.

III. Покласти $\varepsilon_0 = \bar{\varepsilon}$, $j = 0$.

IV. Обчислити множину індексів $\Im_{\varepsilon_j}(x)$ і покласти $\Im = \Im_{\varepsilon_j}(x)$.

V. Обчислити матрицю $P_{\Im}^{\perp}$, яка проектує $R^n$ на ортогональне доповнення до $L_{\Im}$,

$$P_{\Im}^{\perp} = I - A_{\Im}\left(A_{\Im}^T A_{\Im}\right)^{-1} A_{\Im}^T. \quad (6.8)$$

VI. Обчислити градієнт $\nabla f_0(x)$ і вектор

$$h^{\varepsilon_j}(x) = P_{\Im}^{\perp}\nabla f_0(x). \quad (6.9)$$

VII. Якщо $\left\|h^{\varepsilon_j}(x)\right\|^2 > \varepsilon_j$, то покласти $h(x) = -h^{\varepsilon_j}(x)$ і перейти на крок XVII; інакше перейти на крок VIII.

VIII. Якщо $\varepsilon_j \le \varepsilon''$, то обчислити множину індексів $\Im_0(x)$ і перейти на крок IX; інакше перейти на крок XII.

IX. Обчислити вектор:

$$h^0(x) = P_{\Im_0}^{\perp}\nabla f_0(x), \quad (6.10)$$

де $P_{\Im_0}^{\perp}$ – матриця, яка визначається за (6.8) при $\Im = \Im_0$.

X. Обчислити вектор $y^0(x)$ за (6.7) при $\varepsilon = 0$.

XI. Якщо $\left\|h^0(x)\right\|^2 = 0$ і $y^0(x) \le 0$, то покласти $x^* = x$ і припинити обчислення; інакше перейти на крок XII.

XII. Обчислити вектор $y^{\varepsilon_j}(x)$ за формулою (6.7) при $\varepsilon = \varepsilon_j$ і покласти $y = y^{\varepsilon_j}(x)$.

XIII. Якщо $y \le 0$, то покласти $\varepsilon_{j+1} = \beta\varepsilon_j$, $j = j+1$ і перейти на крок IV; інакше перейти на крок XIV.

XIV. Припускаючи, що $\Im = \{i_1, i_2, ..., i_{m'}\}$ та $i_1 < i_2 < ... < i_{m'}$, покласти $\mu_{i_\alpha}(x) = y_\alpha$, $\alpha = \overline{1, m'}$.

XV. Знайти найменше з чисел $i \in \Im$, для яких вектор

$$\bar{h}^{\varepsilon_j}(x) \triangleq P_{\Im - i}^{\perp}\nabla f_0(x) \quad (6.11)$$

задовольняє умову

$$\left\|\bar{h}^{\varepsilon_j}(x)\right\| = \max\left\{\| P_{\Im - l}^{\perp}\nabla f_0(x) \| \;\middle|\; l \in \Im,\ \mu_l(x) > 0\right\} \quad (6.12)$$

(тут і далі $\Im - i \triangleq \left\{ j \middle| j \in \Im,\ j \ne i \right\}$) і покласти $h(x) = -\bar{h}^{\varepsilon_j}(x)$.

XVI. Якщо $\left\|h(x)\right\|^2 \le \varepsilon_j$, то покласти $\varepsilon_{j+1} = \beta\varepsilon_j$, $j = j+1$ і перейти на крок IV; інакше перейти на крок XVII.

XVII. Обчислити найменше з чисел $\rho(x) > 0$, яке задовольняє умову

$$f_0\left(x+\rho(x)h(x)\right)= \\ =\min\left\{ f_0\left(x+\rho h(x)\right) \middle| \rho \ge 0,\ x+\rho h(x)\in X \right\}. \tag{6.13}$$

XVIII. Обчислити наступне наближення

$$x^{k+1} = x + \rho(x)h(x).$$

XIX. Покласти $k = k+1$ і перейти на крок II.

***Теорема 2.*** *Якщо виконані припущення 2, 2′ і множина*

$$X'\left(x^0\right) \triangleq \left\{ x \middle| f_0(x) \le f_0\left(x^0\right),\ \left(a^j, x\right) - b_j \le 0,\ j=\overline{1,m} \right\}$$

*компактна, то послідовність* $x^k$, $k = 0,1,...,$ *яка побудована алгоритмом 2, або скінченна та її останній елемент оптимальний для задачі 2, або нескінченна і кожна її гранична точка оптимальна для задачі 2.*

*Зауваження 2.* Якщо крок XIX алгоритму 2 замінити кроком XIX′ або XIX″, які описані нижче, то висновки теореми 2 залишаються в силі при умові, що величина $\rho(x)$ на кожній ітерації однозначно визначена. (Така заміна є раціональною з обчислювальної точки зору).

XIX′. Покласти $\bar{\varepsilon} = \varepsilon_j$, $k = k+1$ і перейти на крок II.

XIX″. Якщо $k / i'$ – ціле число (тут ціле число $i' \ge 1$ введено для управління швидкістю зменшення величини $\bar{\varepsilon}$), то покласти $\bar{\varepsilon} = \varepsilon_j$, $k = k+1$ і перейти на крок II; інакше покласти $k = k+1$ і перейти на крок II.

*Зауваження 2′.* Щоб отримати версію алгоритму 2, яку можна реалізувати, крок XVII алгоритму замінити на XVII′. Обчислити найменше з цілих чисел $i \ge 0$, для яких

$$f_0\left(x+\bar{\beta}^i \rho h(x)\right) - f_0(x) - \bar{\beta}^i \rho\alpha\left(\nabla f_0(x),\, h(x)\right) \le 0;$$

$$f_l\left(x+\bar{\beta}^i \rho h(x)\right) \le 0,\ \ l=\overline{1,m};\quad f_l(z) \equiv \left(a^l, z\right) - b_l,$$

і покласти $\rho(x) = \bar{\beta}^i \rho$, де $\rho > 0$; $\bar{\beta} \in (0;1)$; $\alpha \in (0;1)$.

***Алгоритм 2′*** **(*пришвидшена версія алгоритму 2*)**

Початок. I. Обчислити $x^0 \in X$; вибрати $\beta \in (0;1)$, $\bar{\varepsilon} \in (0;\varepsilon')$ і $\varepsilon'' \in (0;\bar{\varepsilon})$; покласти $k = 0$.

Основний цикл. II. Покласти $x = x^k$.

III. Покласти $\varepsilon_0 = \bar{\varepsilon}$, $j = 0$.

IV. Обчислити множину індексів $\Im_{\varepsilon_j}(x)$ і покласти $\Im = \Im_{\varepsilon_j}(x)$.

V. Обчислити матрицю проектування $P_{\Im}^{\perp}$ за (6.8).

VI. Обчислити вектор $h^{\varepsilon_j}(x)$ за (6.9).

VII. Якщо $\left\|h^{\varepsilon_j}(x)\right\|^2 > \varepsilon_j$, то перейти на крок VIII, інакше перейти на крок XII.

VIII. Обчислити вектор $y^{\varepsilon_j}(x)$ за формулою (6.7) і покласти $y = y^{\varepsilon_j}(x)$.

IX. Якщо $y \le 0$, то покласти $h(x) = -h^{\varepsilon_j}(x)$ і перейти на крок XIX; інакше перейти на крок X.

X. Припускаючи, що $\Im = \{i_1, i_2, ..., i_{m'}\}$ та $i_1 < i_2 < ... < i_{m'}$, покласти $\mu_{i_\alpha}(x) = y_\alpha$, $\alpha = \overline{1, m'}$.

XI. Обчислити $\bar{h}^{\varepsilon_j}(x)$ згідно (6.11) і (6.12), покласти $h(x) = -\bar{h}^{\varepsilon_j}(x)$ і перейти на крок XIX.

XII. Якщо $\varepsilon_j < \varepsilon''$, то обчислити вектори $h^0(x)$, $y^0(x)$ згідно (6.10) і (6.7) і перейти на крок XIII; інакше перейти на крок XIV.

XIII. Якщо $\left\|h^0(x)\right\|^2 = 0$ і $y^0(x) \le 0$, то покласти $x^* = x$ і припинити обчислення; інакше перейти на крок XIV.

XIV. Обчислити вектор $y^{\varepsilon_j}(x)$ згідно (6.7) і покласти $y = y^{\varepsilon_j}(x)$.

XV. Якщо $y \le 0$, то покласти $\varepsilon_{j+1} = \beta\varepsilon_j$, $j = j+1$ і перейти на крок IV; інакше перейти на крок XVI.

XVI. Припускаючи, що $\Im = \{i_1, i_2, ..., i_{m'}\}$ та $i_1 < i_2 < ... < i_{m'}$, покласти $\mu_{i_\alpha}(x) = y_\alpha$, $\alpha = \overline{1, m'}$.

XVII. Обчислити $\bar{h}^{\varepsilon_j}(x)$ згідно (6.11) і (6.12), покласти $h(x) = -\bar{h}^{\varepsilon_j}(x)$.

XVIII. Якщо $\left\|h(x)\right\|^2 \le \varepsilon_j$, то покласти $\varepsilon_{j+1} = \beta\varepsilon_j$, $j = j+1$ і перейти на крок IV; інакше перейти на крок XIX.

XIX. Обчислити найменше із чисел $\rho(x) \in (0; \infty)$, які задовольняють умову (6.13).

XX. Обчислити наступне наближення

$$x^{k+1} = x + \rho(x) h(x).$$

XXI. Покласти $k = k+1$ і перейти на крок II.

Для послідовності $x^k$, $k = 0, 1, ...,$ яка породжена алгоритмом $2'$, справедлива теорема 2.

**3. Гібридний метод проекції градієнта для мінімізації функції при нелінійних обмеженнях**

З а д а ч а  3. Знайти $\arg\min_{x \in X} f_0(x)$ для заданої функції $f_0 : R^n \to R^1$ і заданої множини

$$X = \left\{ x \middle| f_j(x) \le 0,\ j = \overline{1,m},\ x \in R^n \right\}.$$

*Припущення 3.* $(\boldsymbol{i})$ – функції $f_j$, $j = \overline{0,m}$, неперервно диференційовані та опуклі донизу; $(\boldsymbol{ii})$ – існує таке число $\varepsilon' > 0$, що для будь-якої точки $x \in X$ і будь-якого числа $\varepsilon \in [0;\varepsilon']$ вектори $\nabla f_j(x)$, $j \in \Im_\varepsilon(x)$, лінійно незалежні.

Гібридний метод проекції градієнта об'єднує ідеї методу проекції градієнта (алгоритм 2) і метод можливих напрямків (алгоритм 1 п. 6.6).

***Алгоритм 3***

П о ч а т о к. I. Обчислити $x^0 \in X$; вибрати $\beta \in (0;1)$ (рекомендується вибирати $\beta = 1/2$), $\bar{\varepsilon} \in (0;\varepsilon')$ і $\varepsilon'' \in (0;\bar{\varepsilon})$; покласти $k = 0$.

О с н о в н и й  ц и к л. II. Покласти $x = x^k$.

III. Покласти $\varepsilon_0 = \bar{\varepsilon}$, $j = 0$.

IV. Обчислити $\varepsilon_j$-активну множину індексів $\Im_{\varepsilon_j}(x)$ за формулою

$$\Im_{\varepsilon_j}(x) = \left\{ i \middle| f_i(x) + \varepsilon_j \ge 0,\ i = \overline{1,m},\ x \in X \right\} \tag{6.14}$$

і покласти $\Im = \Im_{\varepsilon_j}(x)$.

V. Знайти матрицю $A_\Im$, стовпцями якої є вектори $\nabla f_j(x)$, $j \in \Im$, лінійно впорядковані по $j$.

VI. Обчислити матрицю проектування

$$P_\Im^\perp = I - A_\Im \left( A_\Im^T A_\Im \right)^{-1} A_\Im^T.$$

VII. Обчислити градієнт $\nabla f_0(x)$ і вектор

$$h^{\varepsilon_j}(x) = P_\Im^\perp \nabla f_0(x).$$

VIII. Якщо $\left\| h^{\varepsilon_j}(x) \right\|^2 > \varepsilon_j$, то покласти $h(x) = -h^{\varepsilon_j}(x)$ і перейти на крок XX, інакше перейти на крок IX.

IX. Якщо $\varepsilon_j \le \varepsilon''$, то обчислити множину індексів $\Im_0(x)$ за (6.14) (при $\varepsilon_j = 0$) і перейти на крок X; інакше перейти на крок XV.

X. Знайти матрицю $A_{\Im_0}$ (як на кроці V при $\Im = \Im_0$).

XI. Обчислити матрицю проектування

$$P_{\Im_0}^{\perp} = I - A_{\Im_0}\left(A_{\Im_0}^T A_{\Im_0}\right)^{-1} A_{\Im_0}^T .$$

XII. Обчислити вектор

$$h^0(x) = P_{\Im_0}^{\perp} \nabla f_0(x).$$

XIII. Обчислити вектор

$$y^0(x) = \left(A_{\Im_0}^T A_{\Im_0}\right)^{-1} A_{\Im_0}^T \nabla f_0(x).$$

XIV. Якщо $\left\|h^0(x)\right\|^2 = 0$ і $y^0(x) \le 0$, то покласти $x^* = x$ і припинити обчислення; інакше перейти на крок XV.

XV. Обчислити вектор $y^{\varepsilon_j}(x)$ за формулою (6.7) при $\varepsilon = \varepsilon_j$ і покласти $y = y^{\varepsilon_j}(x)$.

XVI. Якщо $y \le 0$, то покласти $\varepsilon_{j+1} = \beta\varepsilon_j$, $j = j+1$ і перейти на крок IV; інакше перейти на крок XVII.

XVII. Припускаючи, що $\Im = \{i_1, i_2, ..., i_{m'}\}$ та $i_1 < i_2 < ... < i_{m'}$, покласти $\mu_{i_\alpha}(x) = y_\alpha$, $\alpha = \overline{1,m'}$.

XVIII. Знайти найменше з чисел $i \in \Im$, для яких вектор

$$\bar{h}^{\varepsilon_j}(x) \triangleq P_{\Im-i}^{\perp} \nabla f_0(x)$$

задовольняє умову

$$\left\|\bar{h}^{\varepsilon_j}(x)\right\| = \max\left\{\left\|P_{\Im-l}^{\perp} \nabla f_0(x)\right\| \, l \in \Im,\ \mu_l(x) > 0\right\},$$

і покласти $h(x) = -\bar{h}^{\varepsilon_j}(x)$.

XIX. Якщо $\left\|h(x)\right\|^2 < \varepsilon_j$, то покласти $\varepsilon_{j+1} = \beta\varepsilon_j$, $j = j+1$ і перейти на крок IV; інакше перейти на крок XX.

*Зауваження 3.* В загальному випадку промінь $\{x' = x + \rho h(x) |\ \rho \ge 0\}$ перетинає множину $X$ лише в точці $x$, тобто, вектор $h(x)$ не визначає можливий напрям. Нижче будується вектор $q(x)$, який визначає можливий напрям.

XX. Якщо $h(x) = -h^{\varepsilon_j}(x)$, то покласти $K = \Im$ і перейти на крок XXI; інакше покласти $K = \Im - i$ і перейти на крок XXI.

XXI. Обчислити вектор:

$$q(x) = \lambda(x) h(x) + A_K \left(A_K^T A_K\right)^{-1} g,$$

де $g = -\varepsilon_j (1,1,...,1)^T \in R^n$ і $\lambda(x) \ge 1$ – найменше з чисел на промені $[1;\infty)$, для яких

$$\left(\nabla f_l(x),\ q(x)\right) \le -\varepsilon_j,$$

де $l = 0$ при $K = \Im$ і $l = i$ при $K = \Im - i$.

XXII. Обчислити найменше з чисел $\rho(x) > 0$, які задовольняють умові

$$f_0\left(x + \rho(x)q(x)\right) = \min\left\{f_0\left(x + \rho q(x)\right) \middle| \rho \geq 0,\ x + \rho q(x) \in X\right\}.$$

XXIII. Обчислити наступне наближення

$$x^{k+1} = x^k + \rho(x)q(x).$$

XIX. Покласти $k = k+1$ і перейти на крок II.

***Теорема 3.*** *Нехай виконуються припущення 3 і нехай задача 3 така, що алгоритм 3 коректний для всіх*

$$x \in \left\{x \middle| f_0(x) \leq f_0\left(x^0\right),\ x \in X\right\}.$$

*Тоді, послідовність $x^k$, $k = 0,1,...$, яка породжена алгоритмом 3, або скінченна і її останній елемент оптимальний для задачі 3, або нескінченна і кожна її гранична точка оптимальна для задачі 3.*

*Зауваження 3′.* Алгоритм 3 можна звести до форми, яка реалізується, тим же способом, що і алгоритм 2 (див. зауваження 2′). Можна також отримати пришвидшену версію алгоритму 3, використовуючи для цього ідею побудови алгоритму 2′.



**4. Метод проекції градієнта за наявності збурень**

З а д а ч а 4. Знайти $\arg\min_{x \in X} f_0(x)$ для заданої функції $f_0 : R^n \to R^1$ і заданої множини $X$.

*Припущення 4.* $(\boldsymbol{i})-$ $X$ – замкнута і опукла множина; $(\boldsymbol{ii})-$ $f_0$ – диференційована на $X$ функція.

В алгоритмі 4 припускається, що є можливість обчислювати на кожній ітерації наближене значення $h^k$ градієнта функції $f_0$ в точці $x^k$.

***Алгоритм 4***

П о ч а т о к . І. Обчислити довільне початкове наближення $x^0 \in X$.

ІІ. Покласти $k = 0$.

О с н о в н и й ц и к л . ІІІ. Обчислити напрям руху $h^k$ (який наближено співпадає з градієнтом функції $f_0$ в точці $x^k$).

IV. Знайти кроковий множник $\rho_k$.

V. Обчислити наступне наближення:

$$x^{k+1} = \pi_X\left(x^k - \rho_k h^k\right),$$

де $\pi_X$ – оператор проектування на множину $X$.

VI. Покласти $k = k+1$ і перейти на крок ІІІ.

***Теорема 4.*** *Нехай виконуються припущення 4 і* $(\boldsymbol{iii})-$ *градієнт функції* $f_0$ *на множині*

$$X_0 = \left\{ x \middle| f_0(x) \le f_0(x^0),\ x \in X \right\}$$

*задовольняє умову Ліпшиця з константою* $\gamma$ *(тут* $x^0$ *– початкове наближення в алгоритмі 4);* $(\boldsymbol{iv})-$ *вектор* $h^k$*, який обчислюється на кроці III алгоритму 4, задовольняє умову*

$$\left\| h^k - \nabla f_0(x) \right\| \le \beta \left\| x^{k+1} - x^k \right\|,\ k = 0,1,\dots .$$

*Тоді, для послідовності* $x^k$, $k = 0,1,\dots$, *породженої алгоритмом 4, справедливі твердження:*

1) *якщо* $0 \le \rho_k \le 2/(\gamma + 2\beta)$, $k = 0,1,\dots$ *то* $\left\{ f_0(x^k) \right\}_{k=0}^{\infty}$ *монотонно спадає;*

2) *якщо* $X_0 -$ *обмежена,* $f_0$ *опукла функція і*

$$0 < \varepsilon_1 \le \rho_k \le 2/(\gamma + 2\beta) - \varepsilon_2,\ k = 0,1,\dots,\ (\varepsilon_2 > 0),$$

$$\text{то}\ \lim_{k\to\infty} f_0(x^k) = f_0^* \triangleq \inf_{x\in X} f_0(x),\quad f_0(x^k) - f_0^* = O\left(\frac{1}{k}\right);$$

3) *якщо функція* $f_0$ *– двічі диференційована і*

$$\gamma_1 \|y\|^2 \le \left( \nabla_{xx}^2 f_0(x)\, y, y \right) \le \gamma \|y\|^2,\ \gamma_1 > 0;$$

$$\left\| \nabla_{xx}^2 f_0(x) - \nabla_{xx}^2 f_0(y) \right\| \le \gamma_2 \|x - y\|;$$

$$\delta = \max\left\{ |1 - \rho\gamma_1|, |1 - \rho\gamma| \right\},\ q = \frac{\delta + \beta\rho}{1 - \beta\rho} < 1,$$

*то при* $\rho_k = \rho$, $k = 0,1,\dots$, *послідовність* $\left\{ x^k \right\}_{k=0}^{\infty}$ *збігається до точки мінімуму* $x^*$, *причому* $\left\| x^k - x^* \right\| = O(q^k)$.

**Приклад 1.** Розв’язати методом проекції градієнта задачу:

$$f_0(x_1, x_2) = -x_1^4 - x_2 \to \min$$

при обмеженнях $0 \le x_1 \le 2$ , $0 \le x_2 \le 1$ .

*Розв’язування.* Спочатку знаходимо градієнт функції $f_0$: $\nabla f_0(x) = \begin{pmatrix} -4x_1^3 \\ -1 \end{pmatrix}$.

**Алгоритм 4**

I. Вибираємо початкове наближення $x^0 = (1;1)^T$ .

II. Покладемо $k = 0$ .

*1-а ітерація:*

III. Обчислюємо градієнт функції $f_0$ в початковій точці $x^0$ :

$$\Delta f_0(x^0) = \begin{pmatrix} -4 \cdot 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -4 \\ -1 \end{pmatrix}.$$

IV. Вибираємо кроковий множник $\rho_0 = 0,2$.

V. Обчислюємо наступне наближення

$$x^1 = \pi_X(x^0 - \rho_0 \cdot \nabla f_0(x^0)) = \pi_X\left(\begin{pmatrix}1\\1\end{pmatrix} - 0,2\begin{pmatrix}-4\\-1\end{pmatrix}\right) = \pi_X\begin{pmatrix}1,8\\1,2\end{pmatrix} = \begin{pmatrix}1,8\\1\end{pmatrix}.$$

VI. Покладемо $k = 0+1 = 1$ і переходимо на крок III.

*2-а ітерація*:

III. Обчислюємо градієнт функції $f_0$ в точці $x^1$:

$$\nabla f_0(x^1) = \begin{pmatrix}-4\cdot 1,8^3\\-1\end{pmatrix} = \begin{pmatrix}-23,328\\-1\end{pmatrix}.$$

IV. Вибираємо кроковий множник $\rho_0 = 0,1$.

V. Обчислюємо наступне наближення

$$x^2 = \pi_X(x^1 - \rho_1 \cdot \nabla f_0(x^1)) = \pi_X\left(\begin{pmatrix}1,8\\1\end{pmatrix} - 0,1\begin{pmatrix}-23,238\\-1\end{pmatrix}\right) = \pi_X\begin{pmatrix}4,1328\\1,1\end{pmatrix} = \begin{pmatrix}2\\1\end{pmatrix}.$$

Зупиняємо обчислення.

За дві ітерації обчислили розв'язок задачі $x^* = (2;1)^T$.



## 6.2. Загальний метод штрафних функцій

Задача 1. Знайти $\arg\max_{x\in X} f_0(x)$ для заданої функції $f_0 : R^n \to R^1$ і заданої множини

$$X = \{x \mid f_i(x) \ge 0, \quad i = \overline{1,m}; \quad x \in Y \subset R^n\}.$$

*Припущення 1.* (***i***)– $Y$ – компактна множина в $R^n$; (***ii***)– функції $f_j$, $j = \overline{0,m}$ – неперервні на множині *Y*.

*Означення 1.* Сімейство функцій $\psi(\cdot,\alpha)$, визначених на множині Y, залежних від параметра $\alpha$, і які мають наступні властивості: (***i***)– $\psi(x,\alpha) \ge 0$ при всіх $x \in Y$; (***ii***)– $\psi(x,\alpha) \to 0$ (монотонно спадаючи) при $\alpha \to \infty$ на множині $X^0$, всюди щільній в $X$; (***iii***)– $\lim_{\alpha\to\infty} \psi(x,\alpha) = \infty$ рівномірно по $x \in (Y \setminus V_\delta(X))$ для будь-якого $\delta > 0$ (де $V_\delta(X)$ означає $\delta$-окіл множини X) називають ***штрафними функціями*** або ***функціями штрафу***, а параметр $\alpha$ – ***штрафним коефіцієнтом*** або ***коефіцієнтом штрафу***.

Якщо $\psi(\cdot,\alpha)$ – штрафні функції, то виконується

$$\max_{x\in X} f_0(x) = \lim_{\alpha\to\infty} \sup_{x\in Y} \left( f_0(x) - \psi(x,\alpha) \right)$$

і, таким чином, розв’язування задачі 1 зводиться до допоміжної задачі максимізації по $x \in Y$ функції $f_0(x) - \psi(x,\alpha)$ при достатньо великих $\alpha$. Розв’язування останньої в силу “простої” структури множини Y є простіше від початкової задачі. В загальному випадку розв’язок задачі 1 знаходиться як граничний розв’язок допоміжної задачі при $\alpha \to \infty$. В деяких випадках розв’язок початкової задачі можна отримати при скінченному значенні $\alpha$.

Нижче наведені приклади штрафних функцій:

$$\psi(x,\alpha) = \alpha \sum_{i=1}^{m} \left|\min\{0, f_i(x)\}\right|^{\tau}, \quad \tau > 0; \tag{6.14}$$

$$\psi(x,\alpha) = \alpha \cdot \left|\min\{0, \min_{i\in[1:m]} f_i(x)\}\right|^{\tau}, \quad \tau > 0; \tag{6.15}$$

$$\psi(x,\alpha) = \begin{cases} 0, & \text{якщо} \quad \min\limits_{i\in[1:m]} f_i(x) \geq 0, \\ \alpha \cdot e^{1/\min\limits_i f_i(x)} & \text{в іншому випадку;} \end{cases} \tag{6.16}$$

$$\psi(x,\alpha) = \left[1 + \sum_{i=1}^{m} \left|\min\{0, f_i(x)\}\right|^{\tau}\right]^{\alpha} - 1, \quad \tau > 0; \tag{6.17}$$

$$\psi(x,\alpha) = \begin{cases} \dfrac{1}{\alpha} \sum\limits_{i=1}^{m} f_i^{-1}(x) & \text{при} \quad \min\limits_{i\in[1:m]} f_i(x) > 0, \\ +\infty, & \text{в іншому випадку;} \end{cases} \tag{6.18}$$

$$\psi(x,\alpha) = \sum_{i=1}^{m} \exp(-\alpha \cdot f_i(x)). \tag{6.19}$$

Штрафні функції (6.14)–(6.17) називаються ***зовнішніми***, а (6.18)–(6.19) – ***внутрішніми штрафними функціями***.

Для функції вигляду

$$\psi(x,\alpha) = \begin{cases} -\dfrac{1}{\alpha} \sum\limits_{i=1}^{m} \ln f_i(x), & x \in X^0; \\ +\infty, & x \notin X^0 \end{cases}$$

умова (***i***) у визначенні 1 штрафних функцій у загальному випадку не виконується, але теорема 1 залишається справедливою.

***Алгоритм 1***

I. Вибрати сімейство штрафних функцій $\psi(\cdot,\alpha)$.

II. Задати таку послідовність $\{\alpha_k\}_{k=0}^{\infty}$ штрафних коефіцієнтів, що $\lim\limits_{k\to\infty} 1/\alpha_k = 0$.

III. Задати таку послідовність $\{\varepsilon_k\}_{k=0}^{\infty}$, що

$$\lim_{k\to\infty} \varepsilon_k = 0, \quad \varepsilon_k > 0 \quad , \quad k = 0,1,\ldots.$$

IV. Обчислити при кожному $k$ $(k = 0,1,\ldots)$ точку $x^k$, що задовольняє умову

$$f_0(x^k) - \psi(x^k, \alpha_k) \ge \sup_{x \in Y}(f_0(x) - \psi(x, \alpha_k)) - \varepsilon_k .$$

Тут також вимагається розв'язувати задачу умовної максимізації. Однак на практиці під множиною $Y$ загалом мається на увазі деяка «проста» множина (наприклад, паралелепіпед в просторі $R^n$) і розв'язування цієї задачі, з точки зору наявності обмежень, не викликає особливих складнощів.

***Теорема 1.*** *Якщо виконується припущення 1, то будь-яка гранична точка послідовності* $\{x^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 1, належить допустимій області X і реалізує* $\max_{x \in X} f_0(x)$, *тобто є розв'язком задачі 1.*

В наступній теоремі наводяться оцінки швидкості збіжності алгоритму 1, на основі яких і вибираються штрафні коефіцієнти, що забезпечують задану точність розв'язку по функціоналу.

***Теорема 1'.*** *Нехай виконуються умови припущення 1 та* $(\boldsymbol{iii})-$ *функція* $f_0(x)$ *задовольняє умову Ліпшиця на компакті Y з константою* $\gamma$; $(\boldsymbol{iv})-$ *існують константи* $\delta, \beta, \sigma > 0$ *такі, що при всіх* $x \in (V_\delta(X) \setminus X) \bigcap Y$ *виконується нерівність:*

$$\min_{i \in [1:m]} f_i(x) \le -\beta[\rho_Y(x, X)]^\sigma ,$$

*де* $\rho_Y -$ *метрика в Y;* $V_\delta(X) -$ $\delta$ *- окіл множини X;* $(\boldsymbol{v})-$*штрафна функція* $\psi(x,\alpha)$ *має вигляд*

$$\psi_\tau(x,\alpha) = \alpha \sum_{i=1}^{m} \left|\min\{0, f_i(x)\}\right|^\tau , \quad \tau > 0 .$$

*Тоді, для достатньо великих* $\alpha$ *при* $\tau\sigma > 1$ *справедливі нерівності:*

$$0 \le \Delta(\alpha,\tau) \le \mu(\gamma^{\tau\sigma} / \alpha)^{1/(\tau\sigma - 1)};$$

$$\rho_Y(x^*(\alpha,\tau), X) \le \mu(\gamma / \alpha)^{1/(\tau\sigma-1)},$$

*де*

$$\Delta(\alpha,\tau) = \max_{x \in Y}(f_0(x) - \psi_\tau(x,\alpha)) - \max_{x \in X} f_0(x);$$

$\mu -$ *величина не залежна від* $\alpha, \gamma$;

$$x^*(\alpha,\tau) = \arg\max_{x \in Y}(f_0(x) - \psi_\tau(x,\alpha)) .$$

*Якщо* $\tau\sigma \le 1$, *то існує таке значення штрафного коефіцієнта*

$$\overline{\alpha} = \overline{\alpha}(\gamma,\beta,\tau) = O(\gamma / \beta^\tau),$$

*що* $x^*(\alpha,\tau) \in X$ *і* $\Delta(\alpha,\tau) = 0$ *при всіх* $\alpha \ge \overline{\alpha}(\gamma,\beta,\tau)$.

*Зауваження 1.* Умова $(\boldsymbol{iv})-$теореми $1'$ з довільним $\delta > 0$, $\sigma = 1$ та деяким $\beta > 0$ виконується в кожному із двох наступних випадків:

1) функції $f_i(x)$, $i=\overline{1,m}$ – опуклі доверху на компакті $Y$ і виконується умова Слейтера, тобто, існує така точка $\bar{x}\in Y$, що $f_i(\bar{x})>0$, $i=\overline{1,m}$;

2) $Y$ – опукла множина в $R^n$, а $f_i(x)$, $i=\overline{1,m}$ – лінійні функції.

*Зауваження 1′*. Умова $(\boldsymbol{iv})$ теореми $1'$ з довільним $\delta>0$, $\sigma>0$ і деяким досить малим $\beta$ виконується, якщо $Y$ – скінченна множина.

*Зауваження 1″*. Якщо система нерівностей

$$f_i(x)\geq 0,\ \ i=\overline{1,m},$$

має єдиний розв'язок $x^*\in R^n$, причому

$$f_1(x^*)=\ldots=f_{n+1}(x^*)=0;$$

$$f_{n+2}(x^*)>0,\ldots,\ f_m(x^*)>0,$$

і будь-які $n$ векторів системи $\{\nabla f_1(x^*),\ldots,\nabla f_{n+1}(x^*)\}$ лінійно незалежні, то

$$\min_{i\in[1;m]} f_i(x)\leq -\beta\left\|x-x^*\right\|.$$

*Зауваження 1‴*. З теореми $1'$ випливає, що при $\tau\sigma\leq 1$ можливо точно розв'язати задачу математичного програмування методом штрафних функцій при скінченному значені коефіцієнта штрафу $\alpha$. Забезпечити виконання умови $\tau\sigma\leq 1$ можна вибором параметра $\tau$. Однак слід враховувати, що при $\tau\leq 1$ функція $\psi_\tau(x,\alpha)$ взагалі не є диференційованою по $x$, навіть якщо функції $f_i(x)$, $i=\overline{1,m}$ диференційовані по $x$, що може ускладнити розв'язування задачі $\max\limits_{x\in Y}(f_0(x)-\psi_\tau(x,\alpha))$.

## 6.3. Методи зовнішніх штрафних функцій

### 1. Задачі з обмеженнями у вигляді нерівностей

Задача 1. Знайти $\arg\min\limits_{x\in X} f_0(x)$ для заданої функції $f_0:R^n\to R^1$ і заданої множини

$$X=\left\{x \mid f_j(x)\leq 0;\ \ j=1,2,\ldots,m,\ \ x\in R^n\right\}. \tag{6.20}$$

*Припущення 1*. $(\boldsymbol{i})$ – функція $f_0$ неперервно диференційована; $(\boldsymbol{ii})$ – функції $f_j$, $j=\overline{1,m}$ неперервні; $(\boldsymbol{iii})$ – існує така точка $x'\in X$, що множина

$$X'=\left\{x \mid f_0(x)\leq f_0(x')\right\}$$

компактна.

Ідея методу зовнішніх штрафних функцій базується на зведенні задачі з обмеженнями до розв'язування послідовності задач на безумовний мінімум. При цьому необхідно будувати послідовність зовнішніх штрафних функцій для множини $X$ , яка визначається наступним чином.

*Означення 1.* Послідовність неперервних функцій $p_k\left(x\right)$, $k=0,1,2,...$, називається ***послідовністю зовнішніх штрафних функцій*** для множини $X$ , якщо:

$$\begin{aligned} &p_k\left(x\right)=0 && \text{для всіх } x\in X, \quad k=0,1,2,...; \\ &p_k\left(x\right)>0 && \text{для всіх } x\notin X, \quad k=0,1,2,...; \\ &p_k\left(x\right)\to\infty && \text{для всіх } x\notin X, \quad k\to\infty; \\ &p_{k+1}\left(x\right)>p_k\left(x\right) && \text{для всіх } x\notin X, \quad k=0,1,2,.... \end{aligned}$$

Далі, використовуючи алгоритми розділів 3 і 4, при кожному $k=0,1,2,...$ знаходимо $\tilde{x}^k$ – розв'язок задачі на безумовний мінімум

$$f_0\left(x\right)+p_k\left(x\right)\to\min_{x\in R^n}. \tag{6.21}$$

Граничні точки послідовності $\left\{\tilde{x}^k\right\}_{k=0}^{\infty}$ є оптимальним розв'язком задачі 1. Методи зовнішніх штрафних функцій забезпечують наближений розв'язок задачі 1 поза допустимою областю.

Наведемо приклад послідовності зовнішніх штрафних функцій для множини $X$ , означеної згідно (6.20):

$$p_k(x)=\alpha_k\sum_{j=1}^{m}\left(\max\left\{0,f_j(x)\right\}\right)^{\beta},\quad k=0,1,2,..., \tag{6.22}$$

де $\beta\geq 1$, $\left\{\alpha_k\right\}_{k=0}^{\infty}$ – строго зростаюча послідовність додатних чисел, що прямує до нескінченності при $k\to\infty$ .

***Алгоритм 1***

I. Побудувати $p_k\left(x\right)$, $k=0,1,2,...$ – послідовність зовнішніх штрафних функцій для множини $X$ , означеної згідно (6.20); (для цього можна використати (6.22)).

II. При кожному $k=0,1,2,...$ знайти $\tilde{x}^k$ – розв'язок задачі на безумовний мінімум

$$f_0\left(x\right)+p_k\left(x\right)\to\min_{x\in R^n} \tag{6.23}$$

(для цього використовувати алгоритми, що наведені в розділах 3 і 4).

III. Знайти граничні точки послідовності $\left\{\tilde{x}^k\right\}_{k=0}^{\infty}$.

***Теорема 1.*** *Нехай виконується припущення 1. Тоді довільна гранична точка послідовності* $\left\{\tilde{x}^k\right\}_{k=0}^{\infty}$, *яка породжена алгоритмом 1, є оптимальним розв'язком задачі 1.*

***Теорема 1'.*** *Нехай виконується припущення 1 і, крім того:*

**а**) *функції* $f_j$, $j=\overline{0,m}$ *– неперервно диференційовані;*

**б**) *множина*

$$X_\lambda = \left\{ x \mid f_0(x) + p_k(x) \le \lambda \right\},\quad \lambda < \infty,\quad k = 0,1,... \tag{6.24}$$

*компактна;*

**в**) *задача 1 має єдиний розв'язок* $x^*$;

**г**) *мінімум в (6.23) при великих* $k$ *досягається в єдиній точці* $\tilde{x}^k$;

**д**) *в розв'язку* $x^*$ *задачі 1 градієнти* $\nabla f_i(x^*)$, $i \in \Im(x^*)$, *лінійно незалежні, де множина* $\Im(x^*)$ *визначена за правилом*

$$\Im(x^*) = \left\{ j \mid f_j(x^*) = 0,\ \ j = \overline{1,m} \right\}.$$

*Тоді для точок* $\left\{\tilde{x}^k\right\}_{k=0}^{\infty}$, *які породжені алгоритмом 1, (в якому* $p_k(x)$, $k = 0,1,2,...$, *визначається згідно (6.22) при* $\beta = 2$*), мають місце граничні рівності:*

$$\lim_{k\to\infty} \tilde{x}^k = x^*;\ \lim_{k\to\infty} \alpha_k \left( \max\left\{ 0, f_j\left(\tilde{x}^k\right) \right\} \right) = \frac{1}{2} u_j,\ \ j = \overline{1,m},$$

*де* $u_j$, $j = \overline{1,m}$, *– множники Лагранжа задачі 1.*

У випадку опуклого програмування оцінки наближення $\tilde{x}^k$ до шуканого $x^*$ можуть бути зроблені більш точними.

***Теорема 1''.*** *Нехай функції* $f_j$, $j = \overline{0,m}$, *опуклі донизу і множина* $X_\lambda$, *яка визначається згідно (6.24), компактна. Нехай, крім того, в точці* $x^*$, *що є розв'язком задачі 1, виконуються необхідні умови в формі теореми Куна–Таккера, тобто існують такі числа* $u_j \ge 0$, $j = \overline{1,m}$, *що*

$$f_0(x^*) \le \sum_{j=1}^{m} u_j f_j(x) + f_0(x),\ \ \forall x;$$

$$u_j f_j(x^*) = 0,\quad j = \overline{1,m}.$$

*Тоді*

$$f_j\left(\tilde{x}^k\right) \le \tilde{u}/\alpha_k\text{, якщо } f_j\left(\tilde{x}^k\right) \ge 0;$$

$$f_0\left(\tilde{x}^k\right) \ge f_0(x^*) - (5/4)\left(\overline{u}^2/\alpha_k\right),$$

*де* $\overline{u} = \left( \sum_{j=1}^{m} (u_j)^2 \right)^{1/2}$.

**Приклад 1.** Розв'язати задачу умовної оптимізації методом зовнішніх штрафних функцій:

$$f_0(x) = x^2 - 10x \to \min$$

при обмеженні

$$x \le 1.$$

*Розв'язування.* I. Послідовність зовнішніх штрафних функцій задамо наступним чином

$$p_k(x) = k\left(\max\left\{0, x-1\right\}\right)^2,\ \ k = 1,2,...$$

II. При кожному $k$ знаходимо розв'язок $\tilde{x}^k$ задачі на безумовний мінімум:

$$F\left(x,k\right) = x^2 - 10x + k\left(\max\left\{0, x-1\right\}\right)^2 \to \min.$$

Розв'яжемо цю задачу аналітично.

1) Всередині допустимої області (тобто при $x<1$) маємо

$$F\left(x,k\right) = x^2 - 10x.$$

Скористаємось необхідною умовою екстремуму:

$$\begin{cases} \dfrac{\partial F\left(x,k\right)}{\partial x} = 2x - 10 = 0 \\ x < 1 \end{cases} \Leftrightarrow \begin{cases} x = 5 \\ x < 1 \end{cases} \Leftrightarrow x \in \varnothing.$$

Отже, мінімуму всередині допустимої області не існує.

2) На межі та зовні допустимої області (тобто при $x \ge 1$) маємо

$$F(x,k) = x^2 - 10x + k(x-1)^2.$$

Скористаємось необхідною умовою екстремуму:

$$\begin{cases} \dfrac{\partial F\left(x,k\right)}{\partial x} = 2x - 10 + 2k(x-1) = 0 \\ x \ge 1 \end{cases} \Leftrightarrow \begin{cases} \tilde{x}^k = \dfrac{k+5}{k+1} > 1 \\ x \ge 1 \end{cases}.$$

Отже, на межі та зовні допустимої області існує розв'язок задачі безумовної оптимізації

$$\tilde{x}^k = \frac{k+5}{k+1} > 1.$$

III. Знаходимо границю $\lim\limits_{k\to\infty} \dfrac{k+5}{k+1} = 1 = x^*$.

Таким чином, знайшли (аналітично) розв'язок початкової задачі $x^* = 1$.

**2. Задачі з обмеженнями у вигляді рівностей**

Задача 2. Знайти $\arg\min\limits_{x\in X} f_0\left(x\right)$ для заданої функції $f_0 : R^n \to R^1$ і множини

$$X = \left\{x \mid g_j\left(x\right) = 0;\ \ j = \overline{1,m}\right\}. \tag{6.25}$$

*Припущення 2.* ($\boldsymbol{i}$)– функція $f_0$ неперервно диференційована; ($\boldsymbol{ii}$)– функції $g_j$, $j = \overline{1,m}$ , неперервні; ($\boldsymbol{iii}$)– існує така точка $x'$, що множина

$$X' = \left\{x \mid f_0\left(x\right) \le f_0\left(x'\right)\right\}$$

компактна.

Нижче наводиться приклад послідовності зовнішніх штрафних функцій для множини $X$ , що задана згідно (6.25):

$$p_k(x) = \alpha_k \left\| g(x) \right\|^\beta = \alpha_k \left( \sum_{j=1}^{m} \left( g_j(x) \right)^2 \right)^{\beta/2}, \; k = 0,1,2,\ldots \; , \quad (6.26)$$

де $\beta \geq 1$ і $\alpha_k$, $k = 0,1,2,\ldots$ – строго зростаюча послідовність додатних чисел, що прямує до нескінченності при $k \to \infty$ .

***Алгоритм 2***

I. Побудувати $\left\{ p_k(x) \right\}_{k=0}^{\infty}$ – послідовність зовнішніх штрафних функцій для множини $X$ , визначеної згідно (6.25); (для цього можна використати (6.26)).

II. При кожному $k = 0,1,2,\ldots$ знайти $\tilde{x}^k$ – розв’язок задачі на безумовний мінімум

$$f_0(x) + p_k(x) \to \min_{x \in R^n}$$

(для цього використовувати алгоритми, що наведені в розділах 3 і 4).

III. Знайти граничні точки послідовності $\left\{ \tilde{x}^k \right\}_{k=0}^{\infty}$.

***Теорема 2.*** *Нехай виконуються припущення 2. Тоді довільна гранична точка послідовності* $\left\{ \tilde{x}^k \right\}_{k=0}^{\infty}$, *яка породжена алгоритмом 2, є оптимальним розв’язком задачі 2.*



### 3. Модифікований метод з процедурою переривання

Безпосереднє використання описаних вище методів зовнішніх штрафних функцій не завжди практично здійснене, оскільки не завжди можливо за скінчений проміжок часу знайти глобальний мінімум функцій $f_0 + p_k$ на всьому просторі. Тому опишемо модифіковані методи зовнішніх штрафних функцій, які використовують процедуру переривання при наближенні до мінімуму або стаціонарних точок функцій $f_0 + p_k$ . Послідовність точок $\tilde{x}^k$, $k = 0,1,2,\ldots$, яка породжена цими алгоритмами, має властивість

$$\nabla f_0\left(\tilde{x}^k\right) + \nabla p_k\left(\tilde{x}^k\right) \to 0$$

при $k \to \infty$. При деяких умовах граничні точки цієї послідовності задовольняють необхідним умовам оптимальності. Алгоритм 3 застосовується для розв’язування задачі 2, а алгоритм $3''$ – для розв’язування задачі 1.

***Алгоритм 3***

Початок. I. Вибрати числа $\varepsilon>0$, $\alpha\in(0;1/2)$, $\beta>1$, $\rho>0$ і точку $x^0\in R^n$.

II. Покласти $\varepsilon_0=\varepsilon$, $j=0$, $k=0$.

Основний цикл. III. Обчислити вектор:

$$h\left(x^j,\varepsilon_k\right)=-\left[\nabla f_0\left(x^j\right)+\left(1/\varepsilon_k\right)\nabla p\left(x^j\right)\right],$$

де функція $p$ обчислюється за правилом

$$p(x)=\left\|g(x)\right\|^2/2=\frac{1}{2}\sum_{l=1}^{m}\left(g_l(x)\right)^2.$$

IV. Якщо $\left\|h\left(x^j,\varepsilon_k\right)\right\|>\varepsilon_k$, то перейти на крок V; інакше покласти $\varepsilon_{k+1}=\varepsilon_k/\beta$, $\tilde{x}^k=x^j$, $k=k+1$ і перейти на крок III.

V. Використовуючи алгоритм 3′, обчислити такий кроковий множник $\lambda_j$, що

$$-\lambda_j(1-\alpha)\left\|h\left(x^j,\varepsilon_k\right)\right\|^2\le f_0\left(x^j+\lambda_j h\left(x^j,\varepsilon_k\right)\right)+$$
$$+\frac{1}{\varepsilon_k}p\left(x^j+\lambda_j h\left(x^j,\varepsilon_k\right)\right)-f_0\left(x^j\right)-\frac{1}{\varepsilon_k}p\left(x^j\right)\le-\lambda_j\alpha\left\|h\left(x^j,\varepsilon_k\right)\right\|^2.$$

VI. Покласти $x^{j+1}=x^j+\lambda_j h\left(x^j,\varepsilon_k\right)$, $j=j+1$ і перейти на крок III.

***Алгоритм 3′* (*алгоритм обчислення множника* $\lambda_j$ *для алгоритму 3*)**

I. Покласти $\mu=\rho$.

II. Обчислити

$$\underline{\theta}\left(\mu,x^j\right)=f_0\left(x^j+\mu h\left(x^j,\varepsilon_k\right)\right)+\frac{1}{\varepsilon_k}p\left(x^j+\mu h\left(x^j,\varepsilon_k\right)\right)-$$
$$-f_0\left(x^j\right)-\frac{1}{\varepsilon_k}p\left(x^j\right)+\mu(1-\alpha)\left\|h\left(x^j,\varepsilon_k\right)\right\|^2. \tag{6.27}$$

III. Якщо $\underline{\theta}\left(\mu,x^j\right)=0$, то покласти $\lambda_j=\mu$ і закінчити обчислення; якщо $\underline{\theta}\left(\mu,x^j\right)<0$, то покласти $\lambda_j=\mu+\rho$ і перейти на крок II; якщо $\underline{\theta}\left(\mu,x^j\right)>0$, то перейти на крок IV.

IV. Обчислити

$$\overline{\theta}\left(\mu,x^j\right)=f_0\left(x^j+\mu h\left(x^j,\varepsilon_k\right)\right)+\frac{1}{\varepsilon_k}p\left(x^j+\mu h\left(x^j,\varepsilon_k\right)\right)-$$
$$-f_0\left(x^j\right)-\frac{1}{\varepsilon_k}p\left(x^j\right)+\mu\alpha\left\|h\left(x^j,\varepsilon_k\right)\right\|^2. \tag{6.28}$$

V. Якщо $\overline{\theta}\left(\mu, x^j\right) \le 0$, то покласти $\lambda_j = \mu$ і закінчити обчислення; інакше покласти $a_0 = \mu - \rho$, $b_0 = \mu$ та перейти на крок VI.

VI. Покласти $i = 0$.

VII. Покласти $\upsilon_i = \left(a_i + b_i\right)/2$.

VIII. Обчислити $\underline{\theta}\left(\upsilon_i, x^j\right)$ і $\overline{\theta}\left(\upsilon_i, x^j\right)$ згідно (6.27) і (6.28), відповідно.

IX. Якщо $\underline{\theta}\left(\upsilon_i, x^j\right) \ge 0$ і $\overline{\theta}\left(\upsilon_i, x^j\right) \le 0$, то покласти $\lambda_j = \upsilon_i$ і закінчити обчислення; інакше перейти на крок X.

X. Якщо $\underline{\theta}\left(\upsilon_i, x^j\right) > 0$, то покласти $a_{i+1} = a_i$, $b_{i+1} = \upsilon_i$, $i = i + 1$ і перейти на крок VII; інакше покласти $a_{i+1} = \upsilon_i$, $b_{i+1} = b_i$, $i = i + 1$ та перейти на крок VII.

***Теорема 3.*** *Нехай виконані припущення:* $(\boldsymbol{i})-$ *функції* $f_0$ *і* $g_j$, $j = \overline{1,m}$, *неперервно диференційовані;* $(\boldsymbol{ii})-$ *матриця Якобі* $\partial g\left(x\right)/\partial x$ *має максимальний ранг для всіх* $x$, *що належать досить великій відкритій множині в* $R^n$; $(\boldsymbol{iii})-$ *множина* $\left\{x \mid f_0\left(x\right) \le d_0\right\}$ *компактна для всіх* $d_0 \in R^1$.

*Тоді, якщо послідовність* $\left\{\tilde{x}^k\right\}_{k=0}^{\infty}$, *яка породжена алгоритмом 3, залишається в компактній множині, то* $h\left(\tilde{x}^k, \varepsilon_k\right) \to 0$ *при* $k \to \infty$, *послідовність* $\left\{\tilde{x}^k\right\}_{k=0}^{\infty}$ *буде мати граничні точки* $x^*$, *що лежать в* $X$ *і задовольняють необхідним умовам оптимальності для задачі 2.*

*Зауваження 3.* Із теореми 3 випливає, що для ефективного використання алгоритму 3 (який можна реалізувати), необхідно виконання ряду припущень, і на відміну від чисто принципіальної схеми (алгоритм 2) тут немає простих, доволі загальних умов, які гарантують, що послідовність $\left\{\tilde{x}^k\right\}_{k=0}^{\infty}$ буде залишатися в компактній множині.

***Алгоритм 3′′***

Кроки I, II, IV, V, VI такі, як в алгоритмі 3. Крок III алгоритму 3 замінити на III′.

III′. Обчислити вектор:

$$h\left(x^j, \varepsilon_k\right) = -\left[\nabla f_0\left(x^j\right) + \left(1/\varepsilon_k\right)\nabla p\left(x^j\right)\right],$$

де функція

$$p\left(x\right) = \frac{1}{2}\sum_{l=1}^{m}\left(\max\left\{0, f_l\left(x\right)\right\}\right)^2, \text{ або } p\left(x\right) = \frac{1}{2}\sum_{l \in L(x)}\left(f_l\left(x\right)\right)^2,$$

а множина $L\left(x\right)$ визначається за правилом

$$L(x)=\left\{l \mid f_l(x)\ge 0,\ l\in\left\{\overline{1,m}\right\}\right\}.$$

***Теорема 3′.*** *Нехай виконані припущення:* $(\boldsymbol{i})-$ *функції* $f_0$ *і* $f_l$, $l=\overline{1,m}$ *неперервно диференційовані;* $(\boldsymbol{ii})-$ *множина* $\left\{x \mid f_0(x)\le d_0\right\}$ *компактна для всіх* $d_0\in R^1$; $(\boldsymbol{iii})-$ *для всіх* $x\in R^n$ *вектори* $\nabla f_l(x)$, $l\in L(x)$ *лінійно незалежні.*

*Тоді для послідовності* $\left\{\tilde{x}^k\right\}_{k=0}^{\infty}$, *яка породжена алгоритмом 3″ має місце граничне співвідношення* $h\left(\tilde{x}^k,\varepsilon_k\right)\to 0$ *при* $k\to\infty$.

*Якщо, крім того, послідовність* $\left\{\tilde{x}^k\right\}_{k=0}^{\infty}$ *належить компакту, то її граничні точки* $x^*$ *лежать в* $X$ *і задовольняють необхідним умовам оптимальності для задачі 1.*

*Зауваження 3′.* Із теореми 3′ випливає, що для ефективного використання алгоритму 3″ (який можна реалізувати), немає простих доволі загальних умов, які гарантують, що послідовність $\left\{\tilde{x}^k\right\}_{k=0}^{\infty}$ буде збігатись. Однак на практиці виявляється, що алгоритми типу 3″ добре працюють, оскільки зазвичай послідовність $\left\{\tilde{x}^k\right\}_{k=0}^{\infty}$ збігається.



### 4. Метод зовнішніх штрафних функцій для мінімізації недиференційованих функцій

Задача 4. Знайти $\arg\min\limits_{x\in X_1} f_0(x)$, де

$$X_1=X\bigcap\left\{x \mid f_j(x)\le 0;\ j=\overline{1,m}\right\}. \quad (6.29)$$

Для заданої множини $X\subset R^n$ і заданих функцій $f_j: R^n\to R^1$, $j=\overline{0,m}$.

*Припущення 4.* $(\boldsymbol{i})-$ функції $f_j$, $j=\overline{0,m}$ опуклі вниз і неперервні при $x\in X$; $(\boldsymbol{ii})-$ обмеження (6.29) задовольняє умову Слейтера; $(\boldsymbol{iii})-X$ – опукла і замкнута множина.

Наближений розв’язок даної задачі будемо шукати як розв’язок додаткової (більш «простої») задачі:

знайти $\arg\min\limits_{x\in X} p(x,\alpha)$ для

$$p(x,\alpha)=f_0(x)+\sum_{j=1}^{m}\beta\left(f_j(x)\right)f_j(x), \quad (6.30)$$

де $\alpha$ – деяке додатне число; $\beta\left(f_j(x)\right)=\alpha$, якщо $f_j(x)>0$ і $\beta\left(f_j(x)\right)=0$, якщо $f_j(x)\le 0$.

Відомо, що при $\alpha$, яке перевищує абсолютні значення множників Лагранжа, розв'язок задачі (6.30) співпадає з розв'язком задачі 4.

***Алгоритм 4***

Початок. I. Вибрати: довільне початкове наближення $x^0 \in R^n$; кроковий множник $\rho_0$ і нормуючий множник $\gamma_0$, що задовольняють умовам теореми 4; додатне число $\alpha$.

II. Покласти $k = 0$.

Основний цикл. III. Обчислити числа $\beta_j = \beta(f_j(x^k)),\ \ j = \overline{1,m}$.

IV. Обчислити наступне наближення:

$$x^{k+1} = \pi_X\left(x^k - \rho_k\gamma_k\left(\hat{\nabla}f_0\left(x^k\right) + \sum_{j=1}^{m}\beta_j\hat{\nabla}f_j(x^k)\right)\right),$$

де $\hat{\nabla}f_j\left(x^k\right)$ $\left(j = \overline{0,m}\right)$ – узагальнений градієнт функції $f_j$ в точці $x^k$.

V. Обчислити кроковий множник $\rho_{k+1}$ і нормуючий множник $\gamma_{k+1}$, що задовольняють умовам теореми 4.

VI. Покласти $k = k+1$ і перейти на крок III.

***Теорема 4.*** *Нехай виконані припущення 4, а також:*

$(\boldsymbol{i})$ – *для довільного числа $\delta < \infty$ знайдеться таке скінченне число $\chi = \chi(\delta)$, що*

$$\left\|\hat{\nabla}_x p(x,\alpha)\right\| \le \chi \ \textit{ при } \|x\| \le \delta;$$

$(\boldsymbol{ii})$ –

$$\gamma_k \ge 0,\ \gamma_k\left\|\hat{\nabla}_x p\left(x^k,\alpha\right)\right\| \le \mathrm{const};$$

$(\boldsymbol{iii})$ –

$$\rho_k \ge 0,\ \sum_{k=0}^{\infty}\rho_k = \infty,\ \sum_{k=0}^{\infty}\rho_k^2 < \infty.$$

*Тоді послідовність $\left\{x^k\right\}_{k=0}^{\infty}$, яка породжена алгоритмом 4, збігається до розв'язку задачі (6.30), а при досить великому $\alpha$ – до розв'язку задачі 4.*

*Зауваження 4.* Для пришвидшення збіжності функцію $p\left(x,\alpha\right)$ доцільно мінімізувати з допомогою методів узагальненого градієнтного спуску з розтягненням простору. Щоб отримати гладку функцію цілі $p\left(x,\alpha\right)$ при гладких $f_j$, $j = \overline{0,m}$, замість (6.30) можна взяти функцію

$$p\left(x,\alpha\right) = f_0\left(x\right) + \sum_{j=1}^{m}\beta\left(f_j\left(x\right)\right)\left[f_j\left(x\right)\right]^2. \qquad (6.31)$$

Однак мінімізація функції (6.31) дає лише наближений розв'язок задачі 4, який прямує до точного при $\alpha \to \infty$.

*Зауваження 4′.* Методи зовнішніх штрафних функцій мають наступні переваги: а) в якості початкового наближення можна вибрати довільну точку $x^0 \in R^n$; б) їх можна застосувати до задач з обмеженнями як типу рівностей, так і нерівностей; в) якщо штрафна функція має вигляд (6.22) і в процесі розв'язування задач мінімізації без обмежень знаходимо точку $\tilde{x}^k$, для якої $f_j\left(\tilde{x}^k\right) \le 0$, то з суми (в означені $p_k(x)$) викидається відповідний член з функцією $f_j$ і це зменшує об'єм обчислень.

Недоліки методів зовнішніх штрафних функцій наступні:

а) генеруються недопустимі точки, тобто такі, що не завжди належать множині $X$; б) якщо функція $f_0\left(x\right)$ або деяка з $f_j$, $j=\overline{1,m}$ – неопуклі, то при поганому початковому наближенні може бути обчислений локальний, а не необхідний глобальний мінімум функції $f_0+p_k$ (в таких випадках необхідні додаткові ресурси для пошуку глобального мінімуму); в) штрафні функції вигляду (6.22) диференційовані невелике число разів, що може погано відображатися на швидкості збіжності; г) якщо обчислювальний процес починається з внутрішньої точки $x^0$ множини $X$, то штрафна функція і всі її похідні залишаються рівними нулю доти, доки точка $x^k$ не перетне границю множини $X$ (це робить обчислювальний процес некерованим всередині множини $X$).



## 6.4. Методи внутрішніх штрафних функцій

### 1. Загальна схема

З а д а ч а 1. Знайти $\arg\min\limits_{x\in X} f_0(x)$ для заданої функції $f_0:\ R^n \to R^1$ і для заданої множини

$$X=\{x\,|\,f_j(x)\le 0,\ \ j=\overline{1,m},\ \ x\in R^n\}.$$

*Припущення 1.* (***i***) – функція $f_0$ – неперервно диференційована; (***ii***) – функції $f_j$, $j=\overline{1,m}$, – неперервні; (***iii***) – існує така точка $x'\in X$, що множина

$$X'=\{x\,|\,f_0(x)\le f_0(x')\}$$

компактна; (***iv***) – множина $X$ має внутрішні точки і замикання її внутрішності $X^0$ співпадає з $X$.

Ідея методу полягає у зведенні розв'язування задачі з обмеженнями до розв'язування послідовності задач на безумовний мінімум. При цьому необхідно будувати послідовність внутрішніх штрафних функцій для множини $X$, яка визначається наступним чином.

*Означення 1.* Послідовність неперервних функцій $p'_k: \ X^0 \to R^1$, $k = 0,1,2...$, називається ***послідовністю внутрішніх штрафних функцій*** для множини $X^0$, якщо:

$$0 < p'_{k+1}(x) < p'_k(x) \text{ для всіх } x \in X^0,\ k = 0,1,2...;$$

$$p'_k(x) \to 0 \text{ для всіх } x \in X^0,\ k \to \infty;$$

$p'_k(x^j) \to \infty$ при $j \to \infty$ для будь-якої послідовності $\{x^j\}_{j=0}^{\infty}$, що належить $X^0$, для якої $x^j \to x \in \partial X$ при $j \to \infty$, $k = 0,1,2...$ (тут $\partial X$ означає границю множини $X$ ).

Далі при кожному $k = 0,1,2...$, використовуючи алгоритми розд.2 та 3, знаходять $\tilde{x}^k$ розв'язок задачі

$$f_0(x) + p'_k(x) \to \min_{x \in X^0}. \tag{6.32}$$

Граничні точки послідовності $\{\tilde{x}^k\}_{k=0}^{\infty}$ є оптимальним розв'язком задачі 1.

Нижче наведені приклади послідовностей внутрішніх штрафних функцій для множини $X$ :

(*обернена штрафна функція*)

$$p'_k(x) = -\alpha_k \sum_{j=1}^{m} \frac{1}{f_j(x)},\ k = 0,1,2,...,\ x \in X^0, \tag{6.33}$$

де функції $f_j$, $j = \overline{1,m}$ задовольняють припущенню 1 і $\{\alpha_k\}_{k=0}^{\infty}$ – послідовність додатних чисел, що строго спадає, наближаючись до нуля при $k \to \infty$;

(*логарифмічна штрафна функція*)

$$\begin{aligned} p'_k(x) &= -\alpha_k \sum_{j=1}^{m} \lg(-f_j(x)) = \\ &= -\alpha_k \lg\left((-1)^m f_1(x) f_2(x) ... f_m(x)\right),\quad x \in X^0,\ k = 0,1..., \end{aligned} \tag{6.34}$$

де функції $f_j$, $j = \overline{1,m}$ і послідовність $\{\alpha_k\}_{k=0}^{\infty}$ такі, як в (6.33) і , крім того, виконується

$$\mu = \max_j(-\min_{x \in X} f_j(x)) < \infty.$$

Відмітимо, що у формулі (6.34) замість десяткового логарифма $\lg$ можна використовувати натуральний логарифм $\ln$.

Методи внутрішніх штрафних функцій забезпечують наближений розв'язок всередині допустимої області, який в границі наближається до розв'язку задачі 1.

***Алгоритм 1***

І. Побудувати $p'_k(x)$, $k = 0,1,2,...$ – послідовність внутрішніх штрафних функцій до множини $X$ (можна використовувати (6.33) і (6.34)).

II. При кожному $k = 0,1,2,...,$ знайти $\tilde{x}^k$ – розв'язок задачі мінімізації функції $f_0(x) + p'_k(x)$ на множині $X^0$ (можна використовувати алгоритми, що описані в розділах 3 і 4).

III. Знайти граничні точки послідовності $\{\tilde{x}^k\}_{k=0}^{\infty}$.

***Теорема 1.*** *Нехай виконуються припущення 1. Тоді будь-яка гранична точка послідовності* $\{\tilde{x}^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 1, є оптимальним розв'язком задачі 1.*

**2. Реалізовна схема з процедурою переривання**

Буквальне використання алгоритму 1 практично не здійснене, оскільки неможливо за скінченний час домогтися мінімізації функції $f_0(x) + p'_k(x)$ за змінною $x \in X^0$. Нижче буде наведений модифікований метод внутрішніх штрафних функцій, що використовує процедури переривання при розв'язуванні задачі (6.32).

***Алгоритм 2***

Початок. I. Вибрати числа $\varepsilon > 0$, $\alpha \in (0;1/2)$, $\beta > 1$, $\rho > 0$ і точку $x^0 \in X^0$.

II. Покласти $\varepsilon_0 = \varepsilon$, $j = 0$, $k = 0$.

Основний цикл. III. Обчислити вектор

$$h(x^j, \varepsilon_k) = -(\nabla f_0(x^j) + \varepsilon_k \nabla p'(x^j)),$$

де функція $p'(x)$ така, що $\varepsilon_k p'(x)$, $k = 0,1,2,...$ – послідовність внутрішніх штрафних функцій для множини $X$ ($p'(x)$ можна визначити, наприклад, за правилами:

$$p'(x) = -\sum_{l=1}^{m} \frac{1}{f_l(x)},\ x \in X^0, \text{ або } p'(x) = -\sum_{l=1}^{m} \lg(-f_l(x)),\ x \in X^0).$$

IV. Якщо $\|h(x^j, \varepsilon_k)\| > \varepsilon_k$ та $\left(h(x^j, \varepsilon_k), -\nabla f_0(x^j)\right) > 0,$ то перейти на крок V; інакше покласти $\varepsilon_{k+1} = \varepsilon_k / \beta$, $\tilde{x}^k = x^j$, $k = k+1$ і перейти на крок III.

V. Використовуючи алгоритм 3′ (п. 6.3), обчислити таке число $\lambda_j$, що

$$x^j + \lambda_j h(x^j, \varepsilon_k) \in X$$

і

$$-\lambda_j (1-\alpha) \left\|h(x^j, \varepsilon_k)\right\|^2 \le f_0(x^j + \lambda_j h(x^j, \varepsilon_k)) +$$
$$+\varepsilon_k p'(x^j + \lambda_j h(x^j, \varepsilon_k)) - f_0(x^j) - \varepsilon_k p'(x^j) \le -\lambda_j \alpha \left\|h(x^j, \varepsilon_k)\right\|^2.$$

VI. Покласти $x^{j+1} = x^j + \lambda_j h(x^j, \varepsilon_k)$, покласти $j = j+1$ і перейти на крок III.

***Теорема 2.*** *Нехай виконані припущення 1 і, крім того:* $(\boldsymbol{i})-$ *функції* $f_l$, $l=\overline{1,m}-$ *неперервно диференційовані;* $(\boldsymbol{ii})-$ *для всіх* $x\in R^n$ *вектори* $\nabla f_l(x)$, $l\in L(x)-$ *лінійно незалежні, де* $L(x)=\{l \mid f_l(x)\ge 0,\ l\in\overline{1,m}\}$; $(\boldsymbol{iii})-$ *точка* $x^0\in X^0$ *і число* $\varepsilon>0$ *вибрані так, що множина* $\{x \mid f_0(x)\le f_0(x^0)+\varepsilon p'(x^0)\}$ *– компактна. Тоді граничні точки послідовності* $\{\tilde{x}^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 2, допустимі і задовольняють необхідним умовам оптимальності задачі 1.*

*Зауваження 2.* Перевагою методів внутрішніх штрафних функцій (у порівнянні з методами зовнішніх штрафних функцій) є те, що вони породжують послідовності тільки допустимих точок, і, отже, можуть бути використані в реальному масштабі часу, в основному мають похідні достатньо високого порядку і допускають управління обчислювальним процесом «всередині» множини $X$ .

Перерахуємо тепер їх деякі недоліки. Методи внутрішніх штрафних функцій можна використовувати лише для обмеженого класу множин $X$ (для множин $X$ , що мають внутрішні точки). Процес повинен починатися тільки з внутрішньої точки множини $X$ . Якщо множина $X$ має вигляд

$$X=\{x \mid f_j(x)\le 0,\ \ j=\overline{1,m}\},$$

то всі члени штрафних функцій типу (6.33) і (6.34) завжди дають ненульовий вклад в величину градієнта штрафної функції. Це значно збільшує об'єм обчислень при пошуку напрямку спуску (градієнта внутрішньої штрафної функції). В цьому розумінні методи зовнішніх штрафних функцій більш привабливі.

### 3. Методи внутрішньої точки із застосуванням $Q$ -функцій

З а д а ч а 3. Знайти $\arg\min\limits_{x\in X} f_0(x)$ для заданої функції $f_0$: $R^n\to R^1$ і заданої множини

$$X=\{x \mid f_i(x)\ge 0,\ i=\overline{1,m},\ x\in R^n\}.$$

*Припущення 3.* Функції $f_j$, $j=0,1,...,m$, неперервні, а множина

$$X^0\triangleq\left\{x \mid f_i(x)>0,\ i=\overline{1,m}\right\}$$

не порожня.

Головну роль в даному алгоритмі відіграє характеристична властивість функції, що називається $Q$ -функцією і визначається наступним чином.

*Означення 2.* Нехай $z=(z_0,z_1,...z_m)$ є $m+1$-вимірним додатнім вектором. Функція $Q(z)$ називається ***Q-функцією*** якщо:

а) $Q(z)$ – неперервна по $z > 0$; б) $\varliminf_{i\to\infty} Q(z^i) = \alpha > Q(\bar{z})$, $\bar{z} > 0$ (допускається $\alpha = \infty$), причому нижня межа береться для такої нескінченної послідовності векторів $\{z^i\}$, що $z^i > 0$ для всіх $i$ та $\lim_{i\to\infty} z^i = \bar{z}$, причому $\bar{z}_j = 0$ для деякого *j*.

В якості прикладу $Q$-функції можна навести функцію

$$Q(f_0(x^k) - f_0(x), f_1(x), ..., f_m(x)) = \frac{1}{f_0(x^k) - f_0(x)} + \sum_{i=1}^{m} \frac{1}{f_i(x)}.$$

Алгоритми внутрішньої точки з застосуванням $Q$-функцій складають особливий клас алгоритмів внутрішніх штрафних функцій і мають наступні властивості:

1) не потрібно вибирати строго спадаючу послідовність $\{\alpha_i\}_{i=0}^{\infty}$ невід'ємних параметрів, що використовуються в якості вагових коефіцієнтів у випадку звичайної внутрішньої функції штрафу;

2) процес мінімізації на $k$-й ітерації залежить тільки від значення цільової функції в початковій точці $x^{k-1}$, тобто в точці безумовного мінімуму на ($k-1$)-й ітерації;

3) членам послідовності внутрішніх допустимих точок відповідають значення цільової функції $f_0(x^0)$, $f_0(x^1), ...,$ $f_0(x^k), ...,$ що утворюють, починаючи з початкової точки, строго спадаючу послідовність. При звичайних припущеннях ця послідовність локально сходиться.

***Алгоритм 3***

П о ч а т о к. I. Знайти $x^0 \in X^0$.

II. Покласти $k = 1$.

О с н о в н и й ц и к л. III. Визначити множину

$$X_k^0 = \{x \,|\, f_0(x) \le f_0(x^{k-1}),\ x \in X^0\}.$$

IV. Покласти

$$Q^k(x) = Q(-f_0(x) + f_0(x^{k-1}),\ f_1(x), ..., f_m(x)),$$

де $Q(\cdot)$ – $Q$-функція.

V. Знайти $x^k$ – який-небудь локальний мінімум функції $Q^k(x)$ на множині $X_k^0$.

VI. Покласти $k = k + 1$ і перейти на крок III.

***Теорема 3.*** *Нехай виконані припущення 3 та:* $(\boldsymbol{i})$ – *множина* $X^*$ *локальних мінімумів задачі 3 непорожня, компактна, ізольована множина;* $(\boldsymbol{ii})$ – $X^* \cap X^0 \ne \varnothing$.

*Тоді* $(\boldsymbol{i}')$ – *при* $X_k^0 \ne \varnothing$ *існує точка* $x^k$*, що є безумовним локальним мінімумом функції* $Q^k$ *на* $X_k^0$*, і будь-яка гранична точка рівномірно*

*обмеженої послідовності* $\{x^k\}_{k=0}^{\infty}$ *є локальним мінімумом функції* $f_0$*;*
$(\boldsymbol{i'i'})$*– при* $X_k^0 = \varnothing$ *для деякого скінченного значення* $k$ *точка* $x^{k-1}$*– локальний мінімум функції* $f_0$.

*Зауваження 3.* Алгоритми з $Q$-функцією зберігають для задач опуклого програмування всі переваги алгоритмів внутрішніх штрафних функцій, наприклад такі, як глобальна збіжність і можливість отримання двоїстих допустимих точок, що дають в границі при $k \to \infty$ розв'язок двоїстої опуклої задачі.

**Приклад 1.** Розв'язати методом внутрішніх штрафних функцій з процедурою переривання (використовуючи обернену штрафну функцію) задачу умовної оптимізації:

$$f_0(x) = (x_1 - 4)^2 + (x_2 - 4)^2 \to \min$$

при обмеженні

$$x_1 + x_2 \le 5 .$$

Виконаємо чотири ітерації алгоритму 2.

Вибираємо функцію $p'(x)$ у вигляді:

$$p'(x) = -1/(x_1 + x_2 - 5), \quad x_1 + x_2 < 5 .$$

Знаходимо градієнти:

$$\nabla f_o(x) = \begin{pmatrix} 2x_1 - 8 \\ 2x_2 - 8 \end{pmatrix}, \ \nabla p'(x) = \begin{pmatrix} 1/(x_1 + x_2 - 5)^2 \\ 1/(x_1 + x_2 - 5)^2 \end{pmatrix}.$$

**Алгоритм 2**

I. Вибираємо $\varepsilon = 0{,}1$, $\alpha = 0{,}1$, $\beta = 2$, $\rho = 1$, $x^0 = (2;2)^T$.

II. Покладемо $\varepsilon_0 = 0{,}1$, $j = 0$, $k = 0$.

*1-а ітерація:*

III. Обчислюємо $h(x^0, \varepsilon_0) = -\left( \begin{pmatrix} -4 \\ -4 \end{pmatrix} + 0{,}1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right) = \begin{pmatrix} 3{,}9 \\ 3{,}9 \end{pmatrix}$.

IV. Оскільки $\left\| h(x^0, \varepsilon_0) \right\| = 3{,}9\sqrt{2} > 0{,}1$ та $\left( h(x^0, \varepsilon_0), -\nabla f_0(x) \right) > 0$, то переходимо на крок V.

V. Вибираємо $\lambda_0 = 0{,}1$ і перевіряємо подвійну нерівність:

1) $x^0 + \lambda_0 h(x^0, \varepsilon_0) = \begin{pmatrix} 2 \\ 2 \end{pmatrix} + 0{,}1 \begin{pmatrix} 3{,}9 \\ 3{,}9 \end{pmatrix} = \begin{pmatrix} 2{,}39 \\ 2{,}39 \end{pmatrix} \in X^0$;

2) $-\lambda_0 0{,}9 \left\| h(x^0, \varepsilon_0) \right\|^2 = -0{,}1 \cdot 0{,}9 \cdot 30{,}42 = -2{,}7378$;

3) $-\lambda_0 0{,}1 \left\| h(x^0, \varepsilon_0) \right\|^2 = -0{,}1 \cdot 0{,}1 \cdot 30{,}42 = -0{,}3042$;

4) $f_0(x^0) = 8$;

$\varepsilon_0 p'(x^0) = 0,1\left(-\frac{1}{2+2-5}\right) = 0,1\cdot 1 = 0,1$;

$f_0(x^0 + \lambda_0 h(x^0, \varepsilon_0\ )) = (2,39-4)^2 + (2,39-4)^2 = 2\cdot 1,61^2 = 5,1842$;

$\varepsilon_0 p'(x^0 + \lambda_0 h(x^0, \varepsilon_0\ )) = 0,1\cdot\left(-\frac{1}{2,39+2,39-5}\right) = 0,1\frac{1}{0,22} = 0,4545$;

$-2,7378 \le 5,1842 + 0,4545 - 8 - 0,1 = -2,4614 \le -0,3042.$

Подвійна нерівність виконується, отже, $\lambda_0 = 0,1$ підібрано правильно.

VI. Покладемо

$$x^1 = x^0 + \lambda_0 h(x^0, \varepsilon_0\ ) = \begin{pmatrix} 2,39 \\ 2,39 \end{pmatrix} \in X^0, \quad j = 0+1 = 1,$$

і переходимо на крок III.

*2-а ітерація:*

III. Обчислюємо

$$h(x^1, \varepsilon_0) = -\left(\begin{pmatrix} 2\cdot 2,39-8 \\ 2\cdot 2,39-8 \end{pmatrix} + 0,1\begin{pmatrix} 1/(2,39+2,39-5)^2 \\ 1/(2,39+2,39-5)^2 \end{pmatrix}\right) = \begin{pmatrix} 1,154 \\ 1,154 \end{pmatrix}.$$

IV. Обчислюємо $\left\| h(x^1, \varepsilon_0\ ) \right\| = 1,154\sqrt{2} \approx 1,632$. Оскільки

$$\left\| h(x^1, \varepsilon_0\ ) \right\| = 1,632 > \varepsilon_0\ = 0,1 \text{ та } \left( h(x^1, \varepsilon_0), -\nabla f_0(x^1) \right) > 0,$$

то переходимо на крок V.

V. Вибираємо $\lambda_1 = 0,01$ і перевіряємо подвійну нерівність:

1) $x^1 + \lambda_1 h(x^1, \varepsilon_0\ ) = \begin{pmatrix} 2,39 \\ 2,39 \end{pmatrix} + 0,01\begin{pmatrix} 1,154 \\ 1,154 \end{pmatrix} = \begin{pmatrix} 2,40154 \\ 2,40154 \end{pmatrix} \in X^0$;

2) $-\lambda_1(1-\alpha)\left\| h(x^1, \varepsilon_0) \right\|^2 = -0,01\cdot 0,9\cdot 2\cdot 1,154^2 = -0,02397$;

3) $-\lambda_1 \alpha \left\| h(x^1, \varepsilon_0) \right\|^2 = -0,01\cdot 0,1\cdot 2\cdot 1,154^2 = -0,00266$;

4) $f_0(x^1) = (2,39-4)^2 + (2,39-4)^2 = 5,1842$;

$\varepsilon_0 p'(x^1) = 0,1\left(-\frac{1}{2,39+2,39-5}\right) = 0,454545$;

$f_0(x^1 + \lambda_1 h(x^1, \varepsilon_0)) = 2\cdot(2,40154-4)^2 = 5,110149$;

$\varepsilon_0 p'(x^1 + \lambda_1 h(x^1, \varepsilon_0)) = 0,1\cdot\left(-\frac{1}{2,40154+2,40154-5}\right) = 0,507820$;

$-0,02397 \le 5,110149 + 0,507820 - 5,1842 - 0,454545 = -0,020776 \le -0,00266.$

Подвійна нерівність виконується, отже, $\lambda_1 = 0,01$ підібрано правильно.

VI. Покладемо $x^2 = x^1 + \lambda_1 h\left(x^1, \varepsilon_0\right) = \begin{pmatrix} 2,40154 \\ 2,40154 \end{pmatrix}$, $j = 1+1 = 2$ і переходимо на крок III.

*3-я ітерація:*

III. Обчислюємо

$$h(x^2, \varepsilon_0) = -\left(\begin{pmatrix} 2\cdot 2,40154 - 8 \\ 2\cdot 2,40154 - 8 \end{pmatrix} + 0,1\begin{pmatrix} 1/(2\cdot 2,40154-5)^2 \\ 1/(2\cdot 2,40154-5)^2 \end{pmatrix}\right) = \begin{pmatrix} 0,6181 \\ 0,6181 \end{pmatrix}.$$

IV. Обчислюємо $\left\| h(x^2, \varepsilon_0) \right\| = 0,6181\sqrt{2} \approx 0,874$.

Оскільки

$$\left\| h(x^2, \varepsilon_0) \right\| = 0,874 > \varepsilon_0 = 0,1 \text{ та } \left( h(x^2, \varepsilon_0), -\nabla f_0(x^2) \right) > 0,$$

то переходимо на крок V.

V. Вибираємо $\lambda_1 = 0,02$ і перевіряємо подвійну нерівність:

1) $x^2 + \lambda_2 h(x^2, \varepsilon_0) = \begin{pmatrix} 2,40154 \\ 2,40154 \end{pmatrix} + 0,02\begin{pmatrix} 0,6181 \\ 0,6181 \end{pmatrix} = \begin{pmatrix} 2,413902 \\ 2,413902 \end{pmatrix} \in X^0$;

2) $-\lambda_2(1-\lambda)\left\| h(x^2, \varepsilon_0) \right\|^2 = -0,02\cdot 0,9\cdot 2\cdot 0,6181^2 = -0,013754$;

3) $-\lambda_2 \alpha \left\| h(x^2, \varepsilon_0) \right\|^2 = -0,02\cdot 0,1\cdot 2\cdot 0,6181^2 = -0,001528$;

4) $f_0(x^2) = (2,40154-4)^2 + (2,40154-4)^2 = 5,110149$;

$$\varepsilon_0 p'(x^2) = 0,1\left(-\frac{1}{2,40154 + 2,40154 - 5}\right) = 0,507820;$$

$$f_0(x^2 + \lambda_2 h(x^2, \varepsilon_0)) = 2\cdot(2,413902-4)^2 = 5,031414;$$

$$\varepsilon_0 p'(x^2 + \lambda_2 h(x^2, \varepsilon_0)) = 0,1\cdot\left(-\frac{1}{2*2,413902-5}\right) = 0,580736;$$

$$-0,013754 \le 5,031414 + 0,580736 - 5,110149 - 0,507820 =$$
$$= -0,005819 \le -0,001528.$$

Подвійна нерівність виконується, отже, $\lambda_2 = 0,02$ підібрано правильно.

VI. $x^3 = x^2 + \lambda_2 h(x^2, \varepsilon_0) = \begin{pmatrix} 2,413902 \\ 2,413902 \end{pmatrix}$, $j = 2+1 = 3$ і переходимо на крок III.

*4-а ітерація:*

III. Обчислюємо

$$h(x^3,\varepsilon_0\ )=-\left(\begin{pmatrix}2\cdot 2,413902-8\\ 2\cdot 2,413902-8\end{pmatrix}+0,1\begin{pmatrix}1/(2\cdot 2,413902-5)^2\\ 1/(2\cdot 2,413902-5)^2\end{pmatrix}\right)=-\begin{pmatrix}0,2003\\ 0,2003\end{pmatrix}.$$

IV. Обчислюємо

$$\left\|h(x^3,\varepsilon_0\ )\right\|=0,2003\sqrt{2}\approx 0,2833>\varepsilon_0\ =0,1\,.$$

Оскільки $\left(h(x^j,\varepsilon_0),-\nabla f_0(x^j)\right)=-0,2003\cdot 3,172136-0,2003\cdot 3,172136<0,$ то покладемо $\varepsilon_1=\varepsilon_0/2=0,1/2=0,05,$ $k=0+1=1$ і переходимо на крок III.

III. Обчислюємо

$$h(x^3,\varepsilon_1)=-\left(\begin{pmatrix}-3,172196\\ -3,172196\end{pmatrix}+0,05\begin{pmatrix}1/(2\cdot 2,413902-5)^2\\ 1/(2\cdot 2,413902-5)^2\end{pmatrix}\right)=\begin{pmatrix}1,485938\\ 1,485938\end{pmatrix}.$$

IV. Обчислюємо $\left\|h(x^3,\varepsilon_1)\right\|=1,485938\sqrt{2}\approx 2,1014\,.$

Оскільки $\left\|h(x^3,\varepsilon_1)\right\|>\varepsilon_1=0,05$ та $\left(h(x^3,\varepsilon_1),-\nabla f_0(x^3)\right)>0$, то переходимо на крок V.

V. Вибираємо $\lambda_3=0,025$ і перевіряємо подвійну нерівність:

1) $x^3+\lambda_3 h(x^3,\varepsilon_1)=\begin{pmatrix}2,413902\\ 2,413902\end{pmatrix}+0,025\begin{pmatrix}1,485938\\ 1,485938\end{pmatrix}=\begin{pmatrix}2,45105\\ 2,45105\end{pmatrix}\in X^0$;

2) $-\lambda_3(1-\lambda)\left\|h(x^3,\varepsilon_1)\right\|^2=-0,025\cdot 0,9\cdot 2\cdot 1,485938^2=-0,099361;$

3) $-\lambda_3\alpha\left\|h(x^3,\varepsilon_1)\right\|^2=-0,01104;$

4) $f_0(x^3)=(2,413902-4)^2+(2,413902-4)^2=5,031413\,;$

$$\varepsilon_1 p'(x^3)=0,05\cdot\left(-\frac{1}{2,413902+2,413902-5}\right)=0,290367;$$

$$f_0(x^3+\lambda_3 h(x^3,\varepsilon_1))=2\cdot(2,45105-4)^2=4,798492;$$

$$\varepsilon_1 p'(x^3+\lambda_3 h(x^3,\varepsilon_1))=0,05\cdot\left(-\frac{1}{2\cdot 2,45105-5}\right)=0,510725;$$

$$-0,099361\le 4,798492+0,510725-5,031413-0,290367=$$
$$=-0,01256\le -0,01104.$$

Подвійна нерівність виконується, отже, $\lambda_3$ вибрано правильно.

VI. Покладемо $x^4=x^3+\lambda_3 h(x^3,\varepsilon_1)=\begin{pmatrix}2,45105\\ 2,45105\end{pmatrix}$ і зупиняємо обчислення.

Таким чином, за чотири ітерації отримали наближений розв'язок $x^4 = (2,45105; 2,45105)^T$ і значення цільової функції $f_0(x^4) = 2\cdot(2,45105-4)^2 = 4,79$.

## 6.5. Комбіновані методи штрафних функцій

З а д а ч а 1. Знайти $\arg\min\limits_{x\in X} f_0(x)$ для заданої функції $f_0 : R^n \to R^1$ і заданої множини $X$.

*Припущення 1.* $(\boldsymbol{i})-$ функція $f_0$ неперервно диференційована; $(\boldsymbol{ii})-$ множина $X$ має вигляд $X = X' \bigcap X''$ де $X'$ – замкнена, а $X''$ – замкнена і замикання її внутрішності $X''^{,0}$ співпадає з $X''$; $(\boldsymbol{iii})-$ існує точка $\tilde{x} \in X$ така, що множина $\tilde{X} = \{x \mid f_0(x) \le f_0(\tilde{x})\}-$ компактна; $(\boldsymbol{iv})-$ щонайменше для однієї точки $\hat{x} \in X$, що є оптимальним розв'язком задачі 1, знайдеться така відкрита куля $V$, що $\hat{x} \in V$ і перетин внутрішності множини $X''$ з множиною $V' = V \bigcap X'$ непорожній.

Комбіновані методи штрафних функцій представляють собою просту комбінацію методів зовнішніх і внутрішніх штрафних функцій. На кожній ітерації необхідно розв'язувати задачу мінімізації функції без обмежень (мінімізується функція, що є сумою цільової функції $f_0$, внутрішньої штрафної функції $p'_k$ і зовнішньої штрафної функції $p_k$). Граничні точки послідовності, що утворена розв'язками цих задач, при певних умовах будуть розв'язками задачі 1.

В загальному випадку комбіновані методи дають кращий результат обчислення умовного мінімуму функції порівняно з кожним методом окремо.

***Алгоритм 1***

I. Покласти $k = 0$.

II. Побудувати функцію $p_k(x)$, що належить послідовності зовнішніх штрафних функцій для множини $X'$ (визначення послідовності зовнішніх штрафних функцій та способи їх побудови наведені в підрозділі 6.3).

III. Побудувати функцію $p'_k(x)$, що належить послідовності внутрішніх штрафних функцій для множини $X''$ (визначення послідовності внутрішніх штрафних функцій і способи їх побудови наведені в підрозділі 6.4).

IV. Знайти $x^k$ – деякий локальний мінімум задачі оптимізації без обмежень:

$$\text{знайти } \arg\min_{x\in X''^{,0}} \left( f_0(x) + p_k(x) + p'_k(x) \right).$$

V. Покласти $k = k+1$ і перейти на крок II.

***Теорема 1.*** *Нехай виконані припущення 1 та існує така точка* $x'' \in X''^{,0}$*, що множина*

$$\{x \mid f_0(x) \le f_0(x'') + p_0(x'') + p_0'(x'')\}$$

*компактна. Тоді кожна гранична точка послідовності* $\{x^k\}_{k=0}^{\infty}$*, яка породжена алгоритмом 1, є оптимальним розв'язком задачі 1.*

## 6.6. Методи можливих напрямків

*Означення 1.* Вектор $h^k$ називають ***можливим напрямком*** у точці $x^k$ (яка є, наприклад, $k$-м наближенням до розв'язку задачі відшукання $\arg\min_{x\in X} f_0(x)$), якщо існує таке число $\rho > 0$, що для всіх $\lambda \in (0;\rho)$ точка $x^k + \lambda h^k$ є допустимою, тобто $x^k + \lambda h^k \in X$ і, крім того, задовольняє нерівності $f_0\left(x^k + \lambda h^k\right) < f_0\left(x^k\right)$.

У методах можливих напрямків покращене $(k+1)$-ше наближення $x^{k+1}$ до $\arg\min_{x\in X} f_0(x)$ обчислюють за формулою

$$x^{k+1} = x^k + \rho_k h^k.$$

Різні варіанти методів можливих напрямків відрізняються способами вибору крокових множників $\rho_k$ і можливих напрямків $h^k$.



#### 1. Методи можливих напрямків розв'язування задач мінімізації з обмеженнями типу нерівностей

З а д а ч а 1. Знайти $\arg\min_{x\in X} f_0(x)$ для заданої функції $f_0 : R^n \to R^1$ і множини

$$X \triangleq \left\{x \middle| f_j(x) \le 0,\quad j \in [1:m],\ x \in R^n\right\},$$

визначеної за допомогою заданих функцій $f_j : R^n \to R^1$, $j \in [1:m]$.

*Припущення 1.* $(\boldsymbol{i})-$ функції $f_j$, $j = \overline{0,m}$ – неперервно диференційовані; $(\boldsymbol{ii})-$ множина $X$ має внутрішність або відносну внутрішність (тобто множина $X$ може знаходитися в лінійному многовиді та мати внутрішність стосовно цього многовиду).

***Алгоритм 1***

П о ч а т о к. I. Обчислити початкове наближення $x^0 \in X$, яке задовольняє умовам теореми 1.

II. Покласти $k = 0$.

О с н о в н и й ц и к л. III. Покласти $x = x^k$.

IV. Знайти розв'язок $\alpha = \alpha_k(x)$, $h = h^k(x)$ задачі лінійного програмування в $(n+1)$-му просторі векторів $(\alpha, h)$:

$$\text{знайти } \min_{(\alpha,h)} \alpha$$

при обмеженнях:

$$\left(\nabla f_0(x), h\right) \le \alpha \ ;$$

$$f_j(x) + \left(\nabla f_j(x), h\right) \le \alpha\,,\ \ j = \overline{1,m};$$

$$\left|h_i\right| \le 1,\ \ i = \overline{1,n}.$$

V. Якщо $\alpha_k(x) = 0$, то покласти $x^* = x^k$ і припинити обчислення; інакше перейти на крок VI.

VI. Обчислити найменше додатне число $\rho_k$, яке задовольняє умову

$$f_0\left(x + \rho_k h^k(x)\right) = \min_{\substack{\rho \ge 0 \\ x + \rho h^k(x) \in X}} f_0\left(x + \rho h^k(x)\right).$$

VII. Обчислити наступне наближення

$$x^{k+1} = x + \rho_k h^k(x).$$

VIII. Покласти $k = k+1$ і перейти на крок III.

***Теорема 1.*** *Якщо виконані припущення 1 і початкове наближення таке, що множина*

$$X^0 \triangleq \left\{ x \middle| f_0\left(x\right) - f_0\left(x^0\right) \le 0,\ \ f_j\left(x\right) \le 0,\ \ j = \overline{1,m},\ \ x \in R^n \right\}$$

*компактна і має внутрішність, то послідовність* $x^0, x^1, \ldots,$ *яка породжена алгоритмом 1, або скінченна та її останній елемент* $x^k$ *задовольняє умову* $\alpha_k(x^k) = 0$, *або нескінченна і кожна її гранична точка* $\overline{x}$ *задовольняє умову* $\alpha_k(\overline{x}) = 0.$

Відмітимо, що якщо $x^*-$ оптимальний розв'язок задачі 1, то $\alpha(x^*) = 0.$

Основний недолік алгоритму 1 полягає в тому, що на кожній ітерації потрібно розв'язувати задачу лінійного програмування, яка може мати велику розмірність (може включати всі функції обмеження $f_j$, $j = \overline{1,m}$).

В наведеному нижче алгоритмі ті з обмежень задачі лінійного програмування

$$f_j\left(x\right) + \left(\nabla f_j\left(x\right), h\right) \le \alpha\,,\ \ j = \overline{1,m}\,,$$

для яких значення $f_j\left(x\right)$ від'ємні й мають достатньо велику абсолютну величину, можуть виявитися несуттєвими і їх виключать із задачі лінійного програмування на $k$-й ітерації.

***Алгоритм 1′***

Початок. I. Обчислити початкове наближення $x^0 \in X$, що задовольняє умовам теореми 1.

II. Вибрати константи $\varepsilon' > 0$, $\varepsilon'' \in (0;\varepsilon')$ і $\beta \in (0;1)$ (рекомендується $\beta = 1/2$).

III. Покласти $\varepsilon = \varepsilon'$, $k = 0$.

Основний цикл. IV. Покласти $x = x^k$.

V. Знайти $\varepsilon$-активну множину індексів

$$\Im_\varepsilon(x) \triangleq \{0\} \cup \left\{ j \middle| f_j(x) \ge -\varepsilon,\ j \in [1:m] \right\}.$$

VI. Якщо $\Im_\varepsilon(x) = \{0\}$, то покласти $h^k = -\nabla f_0(x)$ і перейти на крок XII; інакше знайти розв'язок $\alpha = \alpha_\varepsilon(x)$, $h = h_\varepsilon^k$ задачі лінійного програмування в $(n+1)$-вимірному просторі векторів $(\alpha, h)$: $\min\limits_{(\alpha,h)} \alpha$ при обмеженнях:

$$\left(\nabla f_j(x), h\right) \le \alpha,\quad j \in \Im_\varepsilon(x);$$

$$|h_i| \le 1,\quad i = [1:n].$$

VII. Якщо $\alpha_\varepsilon(x) \le -\varepsilon$, то покласти $h^k = h_\varepsilon^k$ і перейти на крок XII; інакше перейти на крок VIII.

VIII. Якщо $\varepsilon \le \varepsilon''$, то перейти на крок IX; інакше покласти $\varepsilon = \beta\varepsilon$ і перейти на крок V.

IX. Знайти множину індексів

$$\Im_0(x) \triangleq \{0\} \cup \left\{ j \middle| f_j(x) \ge 0,\ j \in [1:m] \right\}.$$

X. Знайти розв'язок $\alpha = \alpha_0(x)$, $h = h_0^k$ задачі лінійного програмування в $(n+1)$-вимірному просторі векторів $(\alpha, h)$:

$$\text{знайти } \min_{(\alpha,h)} \alpha$$

при обмеженнях:

$$\left(\nabla f_j(x), h\right) \le \alpha,\quad j \in \Im_0(x);$$

$$|h_j| \le 1,\quad j = [1:n].$$

XI. Якщо $\alpha_0(x) = 0$, то покласти $x^* = x$ та припинити обчислення; інакше покласти $\varepsilon = \beta\varepsilon$ і перейти на крок V.

XII. Обчислити число

$$\bar\rho(x) = \max\left\{ \rho \middle| f_j\left(x + \lambda h^k\right) \le 0,\ \lambda \in [0;\rho],\ j \in [1:m] \right\}.$$

XIII. Обчислити найменше значення $\rho_k \in [0;\bar\rho(x)]$, що задовольняє умову

$$f_0\left(x + \rho_k h^k\right) = \min_{\rho \in [0;\bar\rho(x)]} f_0\left(x + \rho h^k\right).$$

XIV. Обчислити наступне наближення

$$x^{k+1} = x + \rho_k h^k .$$

XV. Покласти $k = k+1$ і перейти на крок IV.

Для алгоритму $1'$ має місце теорема, аналогічна теоремі 1.

Рекомендується замінити кроки XII і XIII алгоритму $1'$ практично реалізовними кроками XII$'$ і XIII$'$ для обчислення крокового множника $\rho_k$ .

XII$'$. Обчислити найменше із цілих чисел $s \ge 0$, для яких виконуються нерівності:

$$f_0\left(x+\beta^s \delta h^k\right) - f_0(x) - \beta^s \delta\mu\left(\nabla f_0(x), h^k\right) \le 0 ;$$

$$f_j\left(x+\beta^s \delta h^k\right) \le 0 ,\ \ j \in [1:m]$$

(тут $\delta > 0$ і $\mu \in (0;1)$ – наперед задані константи).

XIII$'$. Покласти $\rho_k = \beta^s \delta.$

**Приклад 1.** Розв'язати задачу нелінійного програмування методом можливих напрямків:

$$f_0(x) = x_1^2 + 4x_2^2 - 6x_1 - 32x_2 + 69 \to \min$$

при обмеженнях:

$$f_1(x) = x_1 + x_2 - 4 \le 0, \quad f_2(x) = 4x_1 + x_2 - 12 \le 0,$$
$$f_3(x) = -x_1 \le 0, \qquad f_4(x) = -x_2 \le 0.$$

*Розв'язування.* Знайдемо градієнти функцій $f_j,\ j=\overline{0,4}$:

$$\nabla f_0(x) = \begin{pmatrix} 2x_1 - 6 \\ 8x_2 - 32 \end{pmatrix},\ \nabla f_1(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix},\ \nabla f_2(x) = \begin{pmatrix} 4 \\ 1 \end{pmatrix},$$

$$\nabla f_3(x) = \begin{pmatrix} -1 \\ 0 \end{pmatrix},\ \nabla f_4(x) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

Виконаємо *три* ітерації алгоритму $1'$.

I. Знаходимо початкове наближення $x^0 = (0,5;3)^T$.

II. Задаємо константи $\varepsilon' = 0,1;\ \varepsilon'' = 0,02;\ \beta = 0,5.$

III. Покладемо $\varepsilon = \varepsilon' = 0,1;\ k = 0.$

*1-а ітерація:*

IV. Покладемо $x = x^0 = (0,5;3)^T$.

V. Знайдемо $\varepsilon$-активну множину індексів:

$j = 1: f_1(x) = 0,5 + 3 - 4 = -0,5 < -\varepsilon = -0,1$ (не активний),

$j = 2: f_2(x) = 4 \cdot 0,5 + 3 - 12 = -7 < -\varepsilon = -0,1$ (не активний),

$j = 3: f_3(x) = -0,5 < -\varepsilon = -0,1$ (не активний),

$j = 4: f_4(x) = -3 < -\varepsilon = -0,1$ (не активний).

Отже, $I_\varepsilon(x) = \{0\}$.

VI. Оскільки $I_\varepsilon(x)=\{0\}$, то обчислюємо градієнт

$$\nabla f_0(x)=\begin{pmatrix}2\cdot 0,5-6\\ 8\cdot 3-32\end{pmatrix}=\begin{pmatrix}-5\\ -8\end{pmatrix},$$

покладемо $h^0=(5;8)^T$ і переходимо на крок XII.

XII. Обчислюємо число $\bar\rho(x)$:

$$x+\rho h^0=\begin{pmatrix}0,5\\ 3\end{pmatrix}+\rho\begin{pmatrix}5\\ 8\end{pmatrix}=\begin{pmatrix}0,5+5\rho\\ 3+8\rho\end{pmatrix},$$

$$\begin{cases}f_1(x+\rho h^0)\le 0\\ f_2(x+\rho h^0)\le 0\\ f_3(x+\rho h^0)\le 0\\ f_4(x+\rho h^0)\le 0\end{cases}\Leftrightarrow\begin{cases}0,5+5\rho+3+8\rho-4\le 0\\ 4(0,5+5\rho)+3+8\rho-12\le 0\\ -0,5-5\rho\le 0\\ -3-8\rho\le 0\end{cases}\Leftrightarrow$$

$$\Leftrightarrow\begin{cases}13\rho\le 0,5\\ 2+20\rho+8\rho-9\le 0\\ 5\rho\ge -0,5\\ 8\rho\ge -3\end{cases}\Leftrightarrow\begin{cases}\rho\le 0,0385\\ \rho\le 0,25\\ \rho\ge -0,1\\ \rho\ge -\dfrac{3}{8}\end{cases}\Leftrightarrow \rho\le 0,0385,$$

отже, $\bar\rho(x)=0,0385$.

XIII. Обчислюємо кроковий множник $\rho_0$ із умови

$$f_0(x+\rho_0 h^0)=\min_{\rho\in[0;0,0385]} f_0(x+\rho h^0).$$

Для цього мінімізуємо функцію $f_0(x+\rho h^0)$ по $\rho$:

$$f_0(x+\rho h^0)=f_0\begin{pmatrix}0,5+5\rho\\ 3+8\rho\end{pmatrix}=\left(0,5+5\rho\right)^2+4(3+8\rho)^2-$$

$$-6(0,5+5\rho)-32(3+8\rho)+69=0,25+5\rho+25\rho^2+36+192\rho+$$

$$+256\rho^2-3-30\rho-96-256\rho+69=281\rho^2-89\rho+6,25.$$

Ця функція на відрізку [0;0,0385] приймає найменше значення в точці $\rho_0=0,0385$.

XIV. Обчислюємо наступне наближення

$$x^1=x+\rho_0 h^0\begin{pmatrix}0,5+5\cdot 0,0385\\ 3+8\cdot 0,0385\end{pmatrix}=\begin{pmatrix}0,692\\ 3,308\end{pmatrix}.$$

XV. Покладемо $k=0+1=1$ і переходимо на крок IV.

*2-а ітерація:*

IV. Покладемо $x=x^1=(0,692;3,308)^T$.

V. Знайдемо $\varepsilon$-активну множину індексів:

$j=1: f_1(x)=0,692+3,308-4=0>-\varepsilon=-0,1$ (активний),
$j=2: f_2(x)=4\cdot 0,692+3,308-12=-5,92<-\varepsilon=-0,1$ (не активний),
$j=3: f_3(x)=-0,692<-\varepsilon=-0,1$ (не активний),
$j=4: f_4(x)=-3,308<-\varepsilon=-0,1$ (не активний),
отже, $I_\varepsilon(x)=\{0;1\}$.

VI. Оскільки $I_\varepsilon(x)\neq\{0\}$, то обчислюємо градієнт

$$\nabla f_0(x)=\begin{pmatrix}2\cdot 0,692-6\\ 8\cdot 3,308-32\end{pmatrix}=\begin{pmatrix}-4,616\\ -5,536\end{pmatrix}$$

і розв'язуємо задачу лінійного програмування в просторі $R^3$:

$$\alpha\to\min$$

при обмеженнях:

$$\begin{cases}(\nabla f_0(x),h)\le 2\\ \nabla f_1(x),h)\le 2\\ -1\le h_1\le 1,-1\le h_2\le 1\end{cases}\Leftrightarrow\begin{cases}-4,616h_1-5,536h_2-\alpha\le 0\\ h_1+h_2-\alpha\le 0\\ -1\le h_1\le 1,-1\le h_2\le 1.\end{cases}$$

Графічним способом знаходимо розв'язок $h^1=(-1;0,859)^T$, $\alpha_0(x)=-0,14$.

VII. Оскільки $\alpha_0(x)=-0,14<-\varepsilon=-0,1$ то покладемо $h^1=(-1;0,859)^T$ і переходимо на крок XII.

XII. Обчислюємо число $\overline{\rho}(x)$:

$$x+\rho h^1=\begin{pmatrix}0,692\\ 3,308\end{pmatrix}+\rho\begin{pmatrix}-1\\ 0,859\end{pmatrix}=\begin{pmatrix}0,692-\rho\\ 3,308+0,859\rho\end{pmatrix},$$

$$\begin{cases}f_1(x+\rho h^1)\le 0\\ f_2(x+\rho h^1)\le 0\\ f_3(x+\rho h^1)\le 0\\ f_4(x+\rho h^1)\le 0\end{cases}\Leftrightarrow\begin{cases}0,692-\rho+3,308+0,859\rho-4\le 0\\ 4(0,692-\rho)+3,308+0,859\rho-12\le 0\\ -0,992+\rho\le 0\\ -3,308-0,859\rho\le 0\end{cases}\Leftrightarrow$$

$$\Leftrightarrow\begin{cases}\rho\ge 0\\ \rho\ge -1,9\\ \rho\le 0,992\\ \rho\ge -3,8\end{cases}\Leftrightarrow \rho\le 0,992,$$

отже, $\overline{\rho}(x)=0,992.$

XIII. Обчислюємо кроковий множник $\rho_1$ із умови

$$f_0(x+\rho_1 h^1)=\min_{\rho\in[0;0,992]} f_0(x+\rho h^1).$$

Для цього мінімізуємо функцію $f_0(x+\rho h^1)$ по $\rho$:

$$f_0(x+\rho h^1) = \left(0,692-\rho\right)^2 + 4(3,308+0859\rho)^2 -$$
$$-6(0,692-\rho) - 32(3,308+0,859\rho) + 69 =$$
$$= 0,4789 - 1,3849\rho + \rho^2 + 43,7715 + 22,7326\rho + 2,9515\rho^2 -$$
$$-4,152 + 6\rho - 105,856 - 27,488\rho + 69 = 3,9515\rho^2 - 0,1394\rho + 3,2424.$$

Ця функція на відрізку [0;0,992] приймає найменше значення в точці $\rho_1 = 0,0176$.

XIV. Обчислюємо наступне наближення

$$x^2 = x + \rho_1 h^1 \begin{pmatrix} 0,692 - 0,0176 \\ 3,308 + 0,859 \cdot 0,0176 \end{pmatrix} = \begin{pmatrix} 0,6744 \\ 3,3231 \end{pmatrix}.$$

XV. Покладемо $k = 1+1 = 2$ і переходимо на крок IV.

*3-я ітерація:*

IV. Покладемо $x = x^2 = (0,6744; 3,3231)^T$.

V. Знайдемо $\varepsilon$-активну множину індексів:

$j=1: f_1(x) = 0,6744 + 3,3231 - 4 = -0,0025 > -\varepsilon = -0,1$ (активний),

$j=2: f_2(x) = 4 \cdot 0,6744 + 3,3231 - 12 = -5,97 < -\varepsilon = -0,1$ (не активний),

$j=3: f_3(x) = -0,6744 < -\varepsilon = -0,1$ (не активний),

$j=4: f_4(x) = -3,3231 < -\varepsilon = -0,1$ (не активний),

отже, $I_\varepsilon(x) = \{0;1\}$.

VI. Оскільки $I_\varepsilon(x) \neq \{0\}$, то обчислюємо градієнт

$$\nabla f_0(x) = \begin{pmatrix} 2 \cdot 0,6744 - 6 \\ 8 \cdot 3,3231 - 32 \end{pmatrix} = \begin{pmatrix} -4,6512 \\ -5,4152 \end{pmatrix}$$

і розв'язуємо задачу лінійного програмування в просторі $R^3$:

$$\alpha \to \min$$

при обмеженнях:

$$\begin{cases} (\nabla f_0(x), h) \le \alpha \\ \nabla f_1(x), h) \le \alpha \\ -1 \le h_1 \le 1, -1 \le h_2 \le 1 \end{cases} \Leftrightarrow \begin{cases} -4,6512 h_1 - 5,4152 h_2 - \alpha \le 0 \\ h_1 + h_2 - \alpha \le 0 \\ -1 \le h_1 \le 1, -1 \le h_2 \le 1. \end{cases}$$

Графічним способом знаходимо розв'язок

$$h^2 = (-1; 0,8762)^T, \ \ \alpha_0(x) = -0,1238.$$

VII. Оскільки $\alpha_0(x) = -0,1238 < -\varepsilon = -0,1$, то покладемо $h^2 = (-1; 0,8762)^T$ і переходимо на крок XII.

XII. Обчислюємо число $\overline{\rho}(x)$:

$$x + \rho h^2 = \begin{pmatrix} 0,6744 \\ 3,3231 \end{pmatrix} + \rho \begin{pmatrix} -1 \\ 0,8762 \end{pmatrix} = \begin{pmatrix} 0,6744 - \rho \\ 3,3231 + 0,8762\rho \end{pmatrix},$$

$$\begin{cases} f_1(x+\rho h^2) \le 0 \\ f_2(x+\rho h^2) \le 0 \\ f_3(x+\rho h^2) \le 0 \\ f_4(x+\rho h^2) \le 0 \end{cases} \Leftrightarrow \begin{cases} 0,6744-\rho+3,3231+0,8762\rho-4 \le 0 \\ 4(0,6744-\rho)+3,3231+0,8762\rho-12 \le 0 \\ -0,6744+\rho \le 0 \\ -3,3231-0,8762\rho \le 0 \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} 0,1238\rho \ge 0,0025 \\ 3,1238\rho \ge -5,9793 \\ \rho \le 0,6744 \\ \rho \ge -3,7926 \end{cases} \Leftrightarrow \rho \le 0,6744,$$

отже, $\bar{\rho}(x) = 0,6744$.

XIII. Обчислюємо кроковий множник $\rho_2$ із умови

$$f_0(x+\rho_2 h^2) = \min_{\rho\in[0;0,6744]} f_0(x+\rho h^2).$$

Для цього мінімізуємо функцію $f_0(x+\rho h^2)$ по $\rho$:

$$f_0(x+\rho h^2) = \left(0,6744-\rho\right)^2 + 4(3,3231+0,8762\rho)^2 -$$
$$-6(0,6744-\rho)-32(3,3231+0,8762\rho)+69 =$$
$$= 0,4548-1,3488\rho+\rho^2+44,1720+23,2936\rho+$$
$$+3,0709\rho^2-4,0464+6\rho-106,3392-28,0384\rho+69 =$$
$$= 4,0709\rho^2-0,0936\rho+3,2484.$$

Ця функція на відрізку [0;0,6744] приймає найменше значення в точці

$$\rho_2 = 0,0115.$$

XIV. Обчислюємо наступне наближення

$$x^3 = x+\rho_2 h^2 = \begin{pmatrix} 0,6744-0,0115 \\ 3,3231+0,8762\cdot 0,0115 \end{pmatrix} = \begin{pmatrix} 0,6629 \\ 3,3332 \end{pmatrix}.$$

XV. Покладемо $k = 2+1 = 3$ і зупиняємо обчислення.

Таким чином, за три ітерації алгоритму $1'$ можливих напрямків знайдено наближений розв'язок $x^3 = (0,6629;3,3332)^T$ і значення цільової функції $f_0(x^3) = 3,2405$ (зазначимо, що в оптимальній точці: $x^* = (0,6;3,4)^T$, значення цільової функції $f_0(x^*) = 3,2$).

Нижче наводиться модифікація алгоритму $1'$, у якій на кожній ітерації необхідно розв'язувати більш просту задачу лінійного програмування.

***Алгоритм 1''***

П о ч а т о к. І. Обчислити початкове наближення $x^0 \in X$, що задовольняє умовам теореми $1'$.

ІІ. Вибрати константи $\varepsilon' > 0$, $\varepsilon'' \in \left(0;\varepsilon'\right)$ і $\beta \in (0;1)$ (рекомендується $\beta = 1/2$); покласти $k = 0$.

О с н о в н и й  ц и к л . III. Покласти $x = x^k$, $\varepsilon_0 = \varepsilon'$, $s = 0$.

IV. Обчислити $\varepsilon_s$-активну множину індексів

$$\Im_{\varepsilon_s}(x) \triangleq \left\{ j \middle| f_j(x) \geq -\varepsilon_s ,\ j \in [1:m] \right\}.$$

V. Якщо $\Im_{\varepsilon_s}(x) = \{0\}$, то покласти $h^k = \nabla f_0(x)$ і перейти на крок XIII; інакше знайти розв'язок $h = h^k_{\varepsilon_s}$ задачі лінійного програмування в $n$-вимірному просторі векторів $h$:

$$\text{знайти } \arg\min_h \left( \nabla f_0(x), h \right)$$

при обмеженнях:

$$\left( -\nabla f_j(x), h \right) \geq \varepsilon_s,\ \ j \in \Im_{\varepsilon_s}(x);$$
$$\left| h_i \right| \leq 1,\ \ i \in [1:n].$$

VI. Обчислити значення $\alpha_{\varepsilon_s}(x) = \left( \nabla f_0(x), h^k_{\varepsilon_s} \right)$.

VII. Якщо $\alpha_{\varepsilon_s}(x) \leq -\varepsilon_s$, то покласти $h^k = h^k_{\varepsilon_s}$ і перейти на крок XIII; інакше перейти на крок VIII.

VIII. Якщо $\varepsilon_s \leq \varepsilon''$, то перейти на крок IX; інакше покласти $\varepsilon_{s+1} = \beta \varepsilon_s$, $s = s+1$ і перейти на крок IV.

IX. Обчислити множину індексів

$$\Im_0(x) \triangleq \left\{ j \middle| f_j(x) \geq 0,\ \ j \in [1:m] \right\}.$$



X. Знайти розв'язок $h = h^k_0$ задачі лінійного програмування в $n$-вимірному просторі векторів $h$:

$$\text{знайти } \arg\min_h \left( \nabla f_0(x), h \right)$$

при обмеженнях:

$$\left( -\nabla f_j(x), h \right) \geq 0,\ \ j \in \Im_0(x);$$
$$\left| h_i \right| \leq 1,\ \ i \in [1:n].$$

XI. Обчислити значення $\alpha_0(x)$, $\alpha_0(x) = \left( \nabla f_0(x), h^k_0 \right)$.

XII. Якщо $\alpha_0(x) = 0$, то покласти $x^* = x^k$ і припинити обчислення; інакше покласти $\varepsilon_{s+1} = \beta \varepsilon_s$, $s = s+1$, і перейти на крок IV.

XIII. Обчислити число $\bar{\rho}(x) \geq 0$:

$$\bar{\rho}(x) = \max \left\{ \rho \middle| f_j\left( x + \lambda h^k \right) \leq 0,\ \ \lambda \in [0; \rho],\ \ j \in [1:m] \right\}.$$

XIV. Обчислити найменше значення $\rho_k \in [0; \bar{\rho}(x)]$, що задовольняє умову

$$f_0\left( x + \rho_k h^k \right) = \min_{\rho \in [0; \bar{\rho}(x)]} f_0\left( x + \rho h^k \right).$$

XV. Обчислити наступне наближення

$$x^{k+1} = x + \rho_k h^k.$$

XVI. Покласти $k = k+1$ і перейти на крок III.

***Теорема 1'.*** *Якщо виконані всі припущення теореми 1 і, крім того, для будь-якої точки* $x \in X$ *існує такий вектор* $h \in R^n$, *що*

$$\left(\nabla f_j(x), h\right) < 0, \;\; j \in \Im_0(x),$$

*де* $\Im_0(x)$ *– множина, яка визначається співвідношенням*

$$\Im_0(x) \triangleq \left\{ j \middle| f_j(x) \ge 0, \;\; j \in [1:m] \right\},$$

*то послідовність* $x^0, x^1, \ldots$, *яка породжена алгоритмом 1'', або скінченна та її останній елемент* $x^k$ *задовольняє умову* $\alpha_0(x^k) = 0$, *або нескінченна й кожна її гранична точка* $\bar{x}$ *задовольняє умову* $\alpha_0(\bar{x}) = 0$.

(Якщо $x^*$ – оптимальний розв'язок задачі 1, то $\alpha_0(x^*) = 0$ ).

**Приклад 2.** Знайти наближений розв'язок задачі нелінійного програмування:

$$f_0(x) = 2x_1^2 + 2x_2^2 - 2x_1x_2 - 4x_1 - 6x_2 + 6 \to \min$$

при обмеженнях:

$$f_1(x) = x_1 + 5x_2 - 5 \le 0, \quad f_2(x) = 2x_1^2 - x_2 \le 0,$$
$$f_3(x) = -x_1 \le 0, \quad f_4(x) = -x_2 \le 0.$$

*Розв'язування.* Знайдемо градієнти функцій $f_j$, $j = \overline{0,4}$:

$$\nabla f_0(x) = \begin{pmatrix} 4x_1 - 2x_2 - 4 \\ 4x_2 - 2x_1 - 6 \end{pmatrix}, \;\; \nabla f_1(x) = \begin{pmatrix} 1 \\ 5 \end{pmatrix}, \;\; \nabla f_2(x) = \begin{pmatrix} 4x_1 \\ -1 \end{pmatrix},$$
$$\nabla f_3(x) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \;\; \nabla f_4(x) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

Виконаємо <u>*дві*</u> ітерації алгоритму 1''.

I. Знаходимо початкове наближення $x^0 = (0; 0,8)^T$.

II. Задаємо константи $\varepsilon' = 0,1;$ $\varepsilon'' = 0,02;$ $\beta = 0,5;$ покладемо $k = 0$.

*1-а ітерація:*

III. Покладемо $x = x^0;$ $\varepsilon_0 = \varepsilon' = 0,1;$ $s = 0.$

IV. Обчислюємо $\varepsilon$-активну множину індексів:

$j = 1: f_1(x) = 0 + 5 \cdot 0,8 - 5 = -1 < -\varepsilon = -0,1$ (не активний),

$j = 2: f_2(x) = 2 \cdot 0 - 0,8 = -0,8 < -\varepsilon = -0,1$ (не активний),

$j = 3: f_3(x) = 0 > -\varepsilon = -0,1$ (активний),

$j = 4: f_4(x) = -0,8 < -\varepsilon = -0,1$ (не активний), отже, $\Im_{\varepsilon_0}(x)=\{3\}$.

V. Оскільки $\Im_{\varepsilon_0}(x)\neq\varnothing$, то обчислюємо градієнт

$$\nabla f_0(x)=\begin{pmatrix}4\cdot 0-2\cdot 0,8-4\\ 4\cdot 0,8-2\cdot 0-6\end{pmatrix}=\begin{pmatrix}-5,6\\ -2,8\end{pmatrix}$$

і розв'язуємо задачу лінійного програмування в просторі $R^2$:

$$\begin{cases}\left(\begin{pmatrix}-5,6\\ -2,8\end{pmatrix},\begin{pmatrix}h_1\\ h_2\end{pmatrix}\right)\to\min\\ \left(-\begin{pmatrix}-1\\ 0\end{pmatrix},\begin{pmatrix}h_1\\ h_2\end{pmatrix}\right)\geq 0,1\\ -1\leq h_1\leq 1,-1\leq h_2\leq 1\end{cases}\Leftrightarrow\begin{cases}-5,6h_1-2,8h_2\to\min\\ h_1\geq 0,1\\ -1\leq h_1\leq 1,-1\leq h_2\leq 1.\end{cases}$$

Графічно знаходимо розв'язок $h_{\varepsilon_0}^0=(1;1)^T$.

VI. Обчислюємо значення $\alpha_{\varepsilon_0}=-5,6-2,8=-8,4.$

VII. Оскільки значення $\alpha_{\varepsilon_0}=-8,4<-\varepsilon_0=-0,1$ то покладемо $h^0=h_{\varepsilon_0}^0=(1;1)^T$ і переходимо на крок XIII.

XIII. Обчислюємо число $\bar{\rho}(x)$:

$$x+\rho h^0=\begin{pmatrix}0\\ 0,8\end{pmatrix}+\rho\begin{pmatrix}1\\ 1\end{pmatrix}=\begin{pmatrix}\rho\\ 0,8+\rho\end{pmatrix},$$

$$\begin{cases}f_1(x+\rho h^0)\leq 0\\ f_2(x+\rho h^0)\leq 0\\ f_3(x+\rho h^0)\leq 0\\ f_4(x+\rho h^0)\leq 0\end{cases}\Leftrightarrow\begin{cases}\rho+5(0,8+\rho)-5\leq 0\\ 2\rho^2-0,8-\rho\leq 0\\ -\rho\leq 0\\ -0,8-\rho\leq 0\end{cases}\Leftrightarrow$$

$$\Leftrightarrow\begin{cases}6\rho\leq 1\\ 2\rho^2-\rho-0,8\leq 0\\ \rho\geq 0\\ \rho\geq -0,8\end{cases}\Leftrightarrow\begin{cases}\rho\leq\dfrac{1}{6}\\ \rho\in[-0,41;0,91]\\ \rho\geq 0\end{cases}\Leftrightarrow\rho\leq\frac{1}{6}.$$

Отже, $\bar{\rho}(x)=1/6.$

XIV. Обчислюємо кроковий множник $\rho_0$ із умови

$$f_0(x+\rho_0 h^0)=\min_{\rho\in[0;1/6]}f_0(x+\rho h^0),$$

$$f_0(x+\rho h^0)=2\rho^2+2(0,8+\rho)^2-2\rho(0,8+\rho)-4\rho-$$
$$-6(0,8+\rho)+6=2\rho^2+1,28+3,2\rho+2\rho^2-1,6\rho-2\rho^2-$$
$$-4\rho-4,8-6\rho+6=2\rho^2-8,4\rho+2,48.$$

Ця функція досягає мінімуму на відрізку $[0;1/6]$ в точці $\rho_0=1/6$.

XV. Обчислюємо наступне наближення

$$x^1=x+\rho_0 h^0=\left(1/6;29/30\right)^T.$$

XVI. Покладемо $k=0+1=1$ і переходимо на крок III.

*2-а ітерація:*

III. Покладемо $x=x^1=\left(1/6;29/30\right)^T$; $\varepsilon_0=\varepsilon'=0,1$; $s=0$.

IV. Обчислюємо $\varepsilon$-активну множину індексів:

$j=1$: $f_1(x)=1/6+5\cdot\left(29/30\right)-5=0>-\varepsilon_0=-0,1$ (активний),

$j=2$: $f_2(x)=2\cdot\left(1/36\right)-29/30=-52/90<-\varepsilon_0=-0,1$ (не активний),

$j=3$: $f_3(x)=-1/6<-\varepsilon_0=-0,1$ (не активний),

$j=4$: $f_4(x)=-29/30<-\varepsilon_0=-0,1$ (не активний).

Отже, $\Im_{\varepsilon_0}(x)=\{1\}$.

V. Оскільки $\Im_{\varepsilon_0}(x)\neq\varnothing$, то обчислюємо градієнт

$$\nabla f_0(x)=\begin{pmatrix}4\cdot\left(1/6\right)-2\cdot\left(29/30\right)-4\\4\cdot\left(29/30\right)-2\cdot\left(1/6\right)-6\end{pmatrix}=\begin{pmatrix}-79/15\\-37/15\end{pmatrix}$$

і розв'язуємо задачу лінійного програмування в просторі $R^2$

$$\begin{cases}\left(\begin{pmatrix}-79/15\\-37/15\end{pmatrix},\begin{pmatrix}h_1\\h_2\end{pmatrix}\right)\to\min\\\left(-\begin{pmatrix}1\\5\end{pmatrix},\begin{pmatrix}h_1\\h_2\end{pmatrix}\right)\ge 0,1\\-1\le h_1\le 1,-1\le h_2\le 1\end{cases}\Leftrightarrow\begin{cases}-\dfrac{79}{15}h_1-\dfrac{37}{15}h_2\to\min\\-h_1-5h_2\ge 0,1\\-1\le h_1\le 1,-1\le h_2\le 1.\end{cases}$$

Графічно знаходимо розв'язок $h^1_{\varepsilon_0}=(1;-0,22)^T$.

VI. Обчислюємо $\alpha_{\varepsilon_0}=-79/15\cdot 1-37/15\cdot(-0,22)=-4,72.$

VII. Оскільки $\alpha_{\varepsilon_0}=-4,72<-\varepsilon_0=-0,1$, то покладемо $h^1=h^1_{\varepsilon_0}=(1;-0,22)^T$ і переходимо на крок XIII.

XIII. Обчислюємо число $\bar{\rho}(x)$:

$$x+\rho h^1=\begin{pmatrix}0,1667\\0,9667\end{pmatrix}+\rho\begin{pmatrix}1\\-0,22\end{pmatrix}=\begin{pmatrix}0,1667+\rho\\0,9667-0,22\rho\end{pmatrix},$$

$$\begin{cases}f_1(x+\rho h^1)\le 0\\f_2(x+\rho h^1)\le 0\\f_3(x+\rho h^1)\le 0\\f_4(x+\rho h^1)\le 0\end{cases}\Leftrightarrow\begin{cases}0,1667+\rho+5(0,9667-0,22\rho)-5\le 0\\2(0,1667+\rho)^2-0,9667+0,22\rho\le 0\\-0,1667-\rho\le 0\\-0,9667+0,22\rho\le 0\end{cases}\Leftrightarrow$$

$$\Leftrightarrow\begin{cases}0,1\rho\ge 0,0002\\2\rho^2+0,8868\rho-0,9111\le 0\\\rho\ge -0,1667\\\rho\le 4,4\end{cases}\Leftrightarrow\begin{cases}\rho\ge 0,002\\\rho\in\left[-0,932;0,4887\right]\\\rho\le 4,4\end{cases}\Leftrightarrow \rho\le 0,4887.$$

XIV. Обчислюємо кроковий множник $\rho_1$ із умови

$$f_0(x+\rho_1 h^1)=\min_{\rho\in[0;0,4887]} f_0(x+\rho h^1),$$

$$f_0(x+\rho h^1)=2(0,1667+\rho)^2+2(0,9667-0,22\rho)^2-$$
$$-2(0,1667+\rho)(0,9667-9,22\rho)-4(0,1667+\rho)-6(0,9667-0,22\rho)+$$
$$+6=0,05558+0,3334\rho+2\rho^2+1,86902-0,85070\rho+0,0968\rho^2-0,32230+$$
$$+0,07335\rho-1,9334\rho+0,44\rho^2-0,6668-4\rho-5,8002+1,32\rho+6=$$
$$=2,5368\rho^2-5,05735\rho+1,1353.$$

Ця функція досягає мінімуму на відрізку [0;0,4887] в точці $\rho_1=0,4887$.

XV. Обчислюємо наступне наближення

$$x^2=x+\rho_1 h^1=\begin{pmatrix}0,1667+0,4887\\0,9667-0,22\cdot 0,4887\end{pmatrix}=\begin{pmatrix}0,6554\\0,8592\end{pmatrix}.$$

XVI. Покладемо $k=1+1=2$ і зупиняємо обчислення.

Таким чином, за дві ітерації алгоритму $1''$ можливих напрямків знайдено наближений розв'язок $x^2=(0,6554;0,8592)^T$ і значення цільової функції $f_0(x^2)=-0,5675$ (зазначимо, що в оптимальній точці: $x^*=(0,6589;0,8682)^T$, значення цільової функції $f_0(x^*)=-0,6131$).

**2. Методи можливих напрямків для розв'язування задач мінімізації з обмеженнями змішаного типу**

Задача 2. Знайти $\arg\min\limits_{x\in X} f_0(x)$ для заданої функції $f_0:R^n\to R^1$ і множини $X$, заданої функціями $f_j:R^n\to R^1$, $j=\overline{1,m}$, матрицею $A$ розмірності $l\times n$ та $l$-вимірним вектором $b$,

$$X \triangleq \left\{ x \middle| f_j(x) \le 0,\ \ j \in [1:m],\ \ Ax - b = 0,\ \ x \in R^n \right\}.$$

*Припущення 2.* $(\boldsymbol{i})-$ функції $f_j,\ j = \overline{1,m}$ – опуклі й неперервно диференційовані; $(\boldsymbol{ii})-$ множина $X$ – компактна; $(\boldsymbol{iii})-$ існує така точка $\overline{x} \in R^n$, що

$$A\overline{x} - b = 0,\ \ f_j(\overline{x}) < 0,\ \ j \in [1:m].$$

***Алгоритм 2***

Початок. I. Обчислити початкове наближення $x^0 \in X$.

II. Вибрати константи $\varepsilon_0 > 0$ і $\delta_j > 0,\ \ j = \overline{0,m}.$

III. Покласти $k = 0.$

Основний цикл. IV. Обчислити множину індексів

$$\Im_{\varepsilon_k}(x^k) \triangleq \{0\} \bigcup \left\{ j \middle| f_j(x^k) \ge -\varepsilon_k,\ \ j \in [1:m] \right\}.$$

V. Знайти розв’язок $\alpha = \alpha_k,\ h = h^k$ задачі лінійного програмування в $(n+1)$-вимірному просторі векторів $(\alpha, h)$:

$$\text{знайти } \min_{(\alpha,h)} \alpha$$

при обмеженнях:

$$\left(\nabla f_j(x^k), h\right) \le \delta_j \alpha,\ \ j \in \Im_{\varepsilon_k}(x^k);$$

$$Ah = 0;\ \ |h_i| \le 1,\ \ i = \overline{1,n}.$$

VI. Якщо $\alpha_k = 0,$ то перейти на крок VII; інакше перейти на крок X.

VII. Обчислити множину індексів

$$\Im_0(x^k) \triangleq \{0\} \bigcup \left\{ j \middle| f_j(x^k) \ge 0,\ \ j \in [1:m] \right\}.$$

VIII. Обчислити число

$$\gamma_k = -\max_{j \notin \Im_0(x^k)} f_j(x^k).$$

IX. Якщо $\varepsilon_k < \gamma_k$, то покласти $x^* = x^k$ та припинити обчислення; інакше перейти на крок X.

X. Якщо $\alpha_k < -\varepsilon_k$, то перейти на крок XI; інакше перейти на крок XVII.

XI. Покласти $s = 0.$

XII. Покласти $\rho_k = (1/2)^s$.

XIII. Якщо виконуються нерівності:

$$f_0\left(x^k + \rho_k h^k\right) \le f_0(x^k) + \frac{1}{2}\delta_0 \rho_k \alpha_k;\ \ f_j\left(x^k + \rho_k h^k\right) \le 0,\ \ j = \overline{1,m},$$

то перейти на крок XV; інакше перейти на крок XIV.

XIV. Покласти $s = s + 1$ і перейти на крок XII.

XV. Обчислити наступне наближення

$$x^{k+1} = x^k + \rho_k h^k .$$

XVI. Покласти $\varepsilon_{k+1} = \varepsilon_k$, $k = k+1$ і перейти на крок IV.

XVII. Покласти $x^{k+1} = x^k$, $\varepsilon_{k+1} = \varepsilon_k / 2$, $k = k+1$ і перейти на крок IV.

***Теорема 2.*** *Якщо виконані припущення 2 і, крім того, градієнти функцій $f_j$, $j = \overline{1,m}$, задовольняють умову Ліпшиця*

$$\left\| \nabla f_j(x) - \nabla f_j(y) \right\| \le \gamma \left\| x - y \right\|, \; \gamma < \infty, \quad x, \; y \in R^n,$$

*то нескінченна послідовність $\left\{ x^k \right\}_{k=0}^{\infty}$, яка породжена алгоритмом 2, має таку властивість, що $\left\{ f_0(x^k) \right\}_{k=0}^{\infty}$ монотонно не зростаючи, прямує до значення $f_0(x^*)$, де $x^*$ – точка мінімуму функції $f_0(x)$ в області $X$. Якщо $x^*$ – єдина точка мінімуму, то $x^k \to x^*$ при $k \to \infty$.*

Нижче наводяться два методи можливих напрямків, зручні для реалізації їх на ЕОМ.

***Алгоритм 2′***

Початок. I. Вибрати константи $\varepsilon' > 0$, $\varepsilon'' \in (0; \varepsilon')$, $\lambda > 0$, $\beta' \in (0;1)$, $\beta'' \in (0,5; 0,8)$ і ціле число $\tau$, яке задовольняє умову $5 \le \tau \le 10$.

II. Якщо існує вектор $x'$, який задовольняє умовам:

$$f_j(x') \le 0, \quad j \in [1:m]; \; Ax' = b,$$

то покласти $x^0 = x'$ і перейти на крок VII; інакше перейти на крок III.

III. Обчислити вектор $\bar{x}^0$, який задовольняє умову $A\bar{x}^0 = b$.

IV. Обчислити

$$\alpha_0 = \max \left\{ f_j(\bar{x}^0), \; j \in [1:m] \right\}.$$

V. Використати кроки VII–XXI алгоритму 2′ для розв'язування задачі мінімізації в $(n+1)$-вимірному просторі векторів $(\alpha, \bar{x})$:

$$\text{знайти} \min_{(\alpha, \bar{x})} \alpha$$

при обмеженнях:

$$f_j(\bar{x}) \le \alpha, \; j \in [1:m]; \; A\bar{x} = b$$

із початковим наближенням $(\alpha_0, \bar{x}^0)$ доти, поки при деякому $k = k_0$ не виявиться, що $f_j(\bar{x}^{k_0}) \le 0$, $j \in [1:m]$; $A\bar{x}^{k_0} = 0$.

VI. Покласти $x^0 = \bar{x}^{k_0}$ та перейти на крок VII.

Основний цикл. VII. Покласти $k = 0$.

VIII. Покласти $\varepsilon = \varepsilon'$.

IX. Покласти $x = x^k$.

X. Знайти множини індексів $\Im_\varepsilon^1$, $\Im_\varepsilon^2$, які визначаються співвідношеннями:

$$\Im_\varepsilon^1 \bigcup \Im_\varepsilon^2 = \{0\} \bigcup \left\{ j \middle| f_j(x) + \varepsilon \geq 0,\ \ j \in [1:m] \right\};$$

$$\Im_\varepsilon^1 \bigcap \Im_\varepsilon^2 = \varnothing;\ \ 0 \in \Im_\varepsilon^2;$$

$j \in \Im_\varepsilon^1$, тільки якщо функція $f_j$ – афінна.

XI. Розв'язати задачу лінійного програмування в $(n+1)$-вимірному просторі векторів $(\alpha, h)$:

$$\text{знайти } \min_{(\alpha,h)} \alpha$$

при обмеженнях:

$$\left(\nabla f_j(x), h\right) \leq \alpha,\ \ j \in \Im_\varepsilon^2;$$

$$\left(\nabla f_j(x), h\right) \leq 0,\ \ j \in \Im_\varepsilon^1;$$

$$\left|h_i\right| \leq 1,\ \ i = \overline{1,n};\ Ah = 0$$

і позначити її розв'язок через $\left(\alpha_\varepsilon^0, h_\varepsilon^k\right)$.

XII. Якщо $\alpha_\varepsilon^0 \leq -\lambda\varepsilon$, то покласти $h^k = h_\varepsilon^k$ та перейти на крок XVI; інакше перейти на крок XIII.

XIII. Якщо $\varepsilon \leq \varepsilon''$, то перейти на крок XIV; інакше покласти $\varepsilon = \beta'\varepsilon$ і перейти на крок X.

XIV. Покласти $\overline{\varepsilon} = \varepsilon$.

XV. Покласти $\varepsilon = 0$ та використати кроки X, XI алгоритму $2'$ для обчислення вектора $\left(\alpha_0^0, h_0^k\right)$. Якщо $\alpha_0^0 = 0$, то покласти $x^* = x$ і припинити обчислення; інакше покласти $\varepsilon = \beta'\overline{\varepsilon}$ та перейти на крок X.

XVI. Покласти $s = 0$.

XVII. Покласти $\rho = \left(\beta''\right)^s$.

XVIII. Якщо виконуються нерівності:

$$f_0\left(x + \rho_k h^k\right) - f_0(x) - \frac{1}{2}\rho_k \left(\nabla f_0(x), h^k\right) \leq 0;$$

$$f_j\left(x + \rho_k h^k\right) \leq 0,\ \ j \in [1:m],$$

то перейти на крок XIX; інакше покласти $s = s+1$ і перейти на крок XVII.

XIX. Обчислити наступне наближення

$$x^{k+1} = x + \rho_k h^k.$$

XX. Покласти $k = k+1$.

XXI. Якщо, число $k$ кратне $\tau$ то перейти на крок VIII; інакше перейти на крок IX.

***Алгоритм 2″*** (алгоритм 2″ можна застосовувати для розв'язування задачі 2 лише при виконанні умов регулярності Куна–Таккера)

Початок. I. Вибрати константи $\varepsilon' > 0$, $\varepsilon'' \in (0;\varepsilon')$, $\lambda' > 0$, $\lambda'' > 0$, $\beta' \in (0;1)$, $\beta'' \in (0{,}5;0{,}8)$ і ціле число $\tau$, яке задовольняє умову $5 \le \tau \le 10$.

II. Обчислити початкове наближення $x^0$, яке задовольняє умовам (кроки II–VI алгоритму 2′):

$$f_j(x^0) \le 0,\ \ j \in [1:m];\ \ Ax^0 = b.$$

III. Покласти $k = 0$.

Основний цикл. IV. Покласти $\varepsilon = \varepsilon'$.

V. Покласти $x = x^k$.

VI. Знайти множини індексів $\Im_\varepsilon^1$ і $\Im_\varepsilon^2$ (як на кроці X алгоритму 2′).

VII. Розв'язати задачу лінійного програмування в $(n+1)$-вимірному просторі векторів $h$: мінімізувати $\left(\nabla f_j(x), h\right)$ при обмеженнях:

$$\left(\nabla f_j(x), h\right) + \lambda'\varepsilon \le 0,\ \ j \ne 0,\ j \in \Im_\varepsilon^2;$$
$$\left(\nabla f_j(x), h\right) \le 0\,,\ \ j \in \Im_\varepsilon^1;$$
$$|h_i| \le 1,\ \ i = \overline{1,n};\ Ah = 0$$

і позначити її розв'язок через $h_\varepsilon^k$.

VIII. Обчислити

$$\alpha_\varepsilon = \left(\nabla f_0(x), h_\varepsilon^k\right).$$

IX. Якщо $\alpha_\varepsilon \le -\lambda''\varepsilon$, то покласти $h^k = h_\varepsilon^k$ і перейти на крок XV; інакше перейти на крок X.

X. Якщо $\varepsilon \le \varepsilon''$, перейти на крок XI; інакше покласти $\varepsilon = \beta'\varepsilon$ та перейти на крок VI.

XI. Покласти $\overline{\varepsilon} = \varepsilon$.

XII. Покласти $\varepsilon = 0$ та використати кроки VI, VII алгоритму 2″ для обчислення вектора $h_0^k$.

XIII. Обчислити

$$\alpha_0 = \left(\nabla f_0(x), h_0^k\right).$$

XIV. Якщо $\alpha_0 = 0$, то покласти $x^* = x$ і припинити обчислення; інакше покласти $\varepsilon = \beta'\overline{\varepsilon}$ та перейти на крок VI.

XV. Покласти $s = 0$.

XVI. Покласти $\rho_k = \left(\beta''\right)^s$.

XVII. Якщо виконуються нерівності:

$$f_0\left(x+\rho_k h^k\right)-f_0(x)-\frac{1}{2}\rho_k\left(\nabla f_0(x),h^k\right)\le 0;$$

$$f_j\left(x+\rho_k h^k\right)\le 0,\ j\in[1:m],$$

то перейти на крок XVIII; інакше покласти $s=s+1$ і перейти на крок XVI.

XVIII. Обчислити наступне наближення

$$x^{k+1}=x+\rho_k h^k.$$

XIX. Покласти $k=k+1$.

XX. Якщо, число $k$ кратне $\tau$, то перейти на крок IV; інакше перейти на крок V.

### 3. Метод можливих напрямків із квадратичним пошуком

У методі можливих напрямків із квадратичним пошуком для обчислення вектора руху $h^k$ до покращеного $(k+1)$-го наближення до шуканого розв'язку використовуються члени, які містять другі похідні цільової функції. Очевидно, ці алгоритми мають більшу швидкість збіжності, ніж алгоритми з лінійним пошуком, однак на кожній ітерації доводиться розв'язувати задачу квадратичного програмування. Алгоритм 3 можна застосовувати для розв'язування задачі 2 у тому випадку, коли власні значення матриці Гессе $H(x)\triangleq\partial^2 f_0(x)/\partial x^2$ далеко рознесені, коли виконані умови Куна–Таккера та матриця других похідних $H(x)$ додатньо напіввизначена на множині:

$$X\triangleq\left\{x\middle|f_0(x)\le f_0(x^0),\ f_j(x)\le 0,\ j\in[1:m],\ Ax=b\right\},$$

де $x^0$ – початкове наближення.

***Алгоритм 3***

Початок. I. Вибрати константи $\varepsilon'>0,\ \varepsilon''\in(0;\varepsilon')$, $\lambda'>0,\ \lambda''>0,\ \beta'\in(0;1),\ \beta''\in(0{,}5;0{,}8)$ і ціле число $\tau$, яке задовольняє умову $5\le\tau\le 10$.

II. Обчислити початкове наближення $x^0$, яке задовольняє умовам:

$$f_j(x^0)\le 0,\ j\in[1:m],\ Ax^0=b$$

(про метод обчислення $x^0$ див. кроки II–IV алгоритму $2'$).

III. Покласти $k=0$.

Основний цикл. IV. Покласти $\varepsilon=\varepsilon'$.

V. Покласти $x=x^k$.

VI. Знайти множини індексів $\mathfrak{I}_\varepsilon^1,\ \mathfrak{I}_\varepsilon^2$ (як на кроці X алгоритму $2'$).

VII. Знайти розв'язок задачі квадратичного програмування в $n$-вимірному просторі:

$$\text{знайти } \arg\min_h \left( \frac{1}{2}\left( \frac{\partial^2 f_0(x)}{\partial x^2} h, h \right) + \left(\nabla f_0(x), h\right) \right)$$

при обмеженнях:

$$\left(\nabla f_j(x), h\right) + \lambda'\varepsilon \le 0, \;\; j \ne 0, \;\; j \in \Im_\varepsilon^2;$$

$$\left(\nabla f_j(x), h\right) \le 0, \;\; j \in \Im_\varepsilon^1;$$

$$\left|h_i\right| \le 1, \;\; i = \overline{1,n}; \; Ah = b.$$

VIII. Обчислити

$$\alpha_\varepsilon = \left( \frac{1}{2}\left( \frac{\partial^2 f_0(x)}{\partial x^2} h_\varepsilon^k, h_\varepsilon^k \right) + \left(\nabla f_0(x), h_\varepsilon^k\right) \right).$$

IX. Якщо $\alpha_\varepsilon \le -\lambda''\varepsilon$, то покласти $h^k = h_\varepsilon^k$ і перейти на крок XIV; інакше перейти на крок X.

X. Якщо $\varepsilon \le \varepsilon''$, то перейти на крок XI; інакше покласти $\varepsilon = \beta'\varepsilon$ і перейти на крок VI.

XI. Покласти $\overline{\varepsilon} = \varepsilon$.

XII. Покласти $\varepsilon = 0$ та використати кроки VI–VIII алгоритму 3 для обчислення значення

$$\alpha_0 = \left( \frac{1}{2}\left( \frac{\partial^2 f_0(x)}{\partial x^2} h_0^k, h_0^k \right) + \left(\nabla f_0(x), h_0^k\right) \right).$$

XIII. Якщо $\alpha_0 = 0$, то покласти $x^* = x$ і припинити обчислення; інакше покласти $\varepsilon = \beta'\overline{\varepsilon}$ та перейти на крок VI.

XIV. Покласти $s = 0$.

XV. Покласти $\rho_k = \left(\beta''\right)^s$.

XVI. Якщо виконуються нерівності:

$$f_0\left(x + \rho_k h^k\right) - f_0(x) - \frac{1}{2}\rho_k\left(\nabla f_0(x), h^k\right) \le 0;$$

$$f_j\left(x + \rho_k h^k\right) \le 0, \;\; j \in [1:m],$$

то перейти на крок XVII; інакше покласти $s = s + 1$ і перейти на крок XV.

XVII. Обчислити наступне наближення

$$x^{k+1} = x + \rho_k h^k.$$

XVIII. Покласти $k = k + 1$.

XIX. Якщо число $k$ кратне $\tau$, то перейти на крок IV; інакше перейти на крок V.

#### 4. Модифікований метод можливих напрямків Зойтендейка

Задача 4. Знайти $\arg\min_{x\in X} f_0(x)$ для заданої неперервної функції $f_0 : R^n \to R^1$ та множини:

$$X \triangleq \Big\{ x \Big| f_j(x) = 0,\ \ j = \overline{1,l};\ \ f_j(x) \ge 0,$$
$$j = \overline{l+1,m};\ \ x \in R^n \Big\},$$

де $f_j : R^n \to R^1$, $j = \overline{1,m}$ – задані неперервні функції.

*Припущення 4.* $(\boldsymbol{i})$ – функції $f_j(x)$, $j = \overline{1,m}$ – неперервно диференційовані в $R^n$; $(\boldsymbol{ii})$ – $l < n$.

Розв'язування задачі 4 зводиться до розв'язування наступної допоміжної задачі:

$$\text{знайти } \arg\min_x \left( f_0(x) + \alpha \sum_{j=1}^{l} f_j(x) \right)$$

при обмеженнях:

$$-f_j(x) \le 0,\ \ j = \overline{1,m},$$

де $\alpha > 0$ – деякий параметр.

Будь-який мінімум допоміжної задачі, локальний або глобальний, який досягається в точці $\hat{x}$ з $f_j(\hat{x}) = 0$, $j = \overline{1,l}$, є також локальним або глобальним мінімумом задачі 4. Для розв'язання допоміжної задачі використовується метод можливих напрямків Зойтендейка, доповнений процедурою обчислення необхідного значення параметра $\alpha$.

***Алгоритм 4***

Початок. I. Знайти початкове наближення $x^0 \in X^+$, де

$$X^+ = \Big\{ x \Big| f_j(x) \ge 0,\ \ j = \overline{1,m},\ \ x \in R^n \Big\}. \tag{6.35}$$

II. Вибрати константи $\beta > 0$, $\varepsilon_0 > 0$, $\varepsilon' > 0$, $(\varepsilon' << \varepsilon_0)$, $\eta \in (0;1)$, $\theta \in (0;1)$, $\alpha_{-1} > 0$, $\delta > 0$.

III. Для заданих чисел $\alpha > 0$, $\varepsilon > 0$ визначити функції $\psi_\alpha : R^n \to R^1$ та $\Im_\varepsilon : R^n \to \{1,\dots,m\}$ за правилами:

$$\psi_\alpha \triangleq f_0(x) + \alpha \sum_{j=1}^{l} f_j(x);$$

$$\Im_\varepsilon(x) \triangleq \Big\{ j \Big| -f_j(x) \ge -\varepsilon,\ \ j \in [1:m] \Big\}.$$

IV. Визначити функцію $g_{\alpha\varepsilon} : R^n \to R^1$ за правилом:

$$g_{\alpha\varepsilon}(x) = \min_{h\in S} \max \Big\{ \big(\nabla \psi_\alpha(x), h\big), -\big(\nabla f_j(x), h\big),\ \ j \in \Im_\varepsilon(x) \Big\}, \tag{6.36}$$

де $S = \left\{ h \middle| \, | h_i | \le 1, \; i = \overline{1,n}, \; h \in R^n \right\}$.

V. Покласти $k = 0, \; \varepsilon = \varepsilon_0$.

Основний цикл. VI. Обчислити множники Лагранжа $\lambda(x^k) = \left( \lambda_1(x^k), ..., \lambda_m(x^k) \right)$ за правилом:

$$\lambda(x^k) = [\nabla f(x^k)(\nabla f(x^k))^T]^{-1}[\nabla f(x^k)]\nabla f_0(x^k),$$

де $f(x^k) = \left( f_1(x^k), ..., f_m(x^k) \right)^T$; $\nabla f(x^k)$ – відповідна $m \times n$-матриця.

VII. Обчислити

$$\lambda_0 = \min \left\{ \lambda_j(x^k), \; j \in [1:l] \right\}.$$

VIII. Якщо $\alpha_{k-1} + \lambda_0 \ge \beta$, то покласти $\alpha_k = \alpha_{k-1}$ і перейти на крок IX; якщо $\alpha_{k-1} + \lambda_0 < \beta$, то покласти $\alpha_k = \max\left\{ \beta - \lambda_0, \, \alpha_{k-1} + \delta \right\}$ і перейти на крок IX.

IX. Обчислити множину індексів $\Im_\varepsilon(x^k)$ і вектори:

$$\nabla \psi_{\alpha_k}(x^k), \; \nabla f_j(x^k), \; j \in \Im_\varepsilon(x^k).$$

X. Обчислити значення $g_{\alpha_k \varepsilon}(x^k)$ і знайти можливий напрямок пошуку, розв'язуючи наступну задачу:

$$\text{знайти } \arg\min_{h \in S} \max\left\{ (\nabla \psi_{\alpha_k}(x^k), h); -(\nabla f_j(x^k), h), \; j \in \Im_\varepsilon(x^k) \right\},$$

де $S = \left\{ h \middle| \, | h_i | \le 1, \; i = \overline{1,n}, \; h \in R^n \right\}$.

XI. Якщо $g_{\alpha_k \varepsilon}(x^k) > -\varepsilon$, то покласти $\varepsilon = \varepsilon/2$ і перейти на крок XII; інакше перейти на крок XIII.

XII. Якщо $\varepsilon < \varepsilon'$, то обчислити значення $g_{\alpha_k 0}(x^k)$ і припинити обчислення якщо $g_{\alpha_k 0}(x^k) = 0$; інакше перейти на крок IX.

XIII. Обчислити найменше ціле $i_k \ge 0$ таке, що

$$f_j\left( x^k + \eta^{i_k} h(x^k) \right) \ge 0, \; j = \overline{1,m};$$

$$\psi_{\alpha_k}\left( x^k + \eta^{i_k} h(x^k) \right) - \psi_{\alpha_k}(x^k) \le -\theta \eta^{i_k} \varepsilon.$$

XIV. Покласти $x^{k+1} = x^k + \eta^{i_k} h\left( x^k \right), \; k = k+1$ та перейти на крок VI.

***Теорема 4.*** *Нехай виконуються припущення 4 і* (***iii***)*– для будь-якого* $x \in R^n$ *число лінійно незалежних векторів* $\nabla f_j(x), \; j = [1:m]$ *таких, що* $f_j(x) = 0$ *не перевищує* $n$*;* (***iv***)*– множина*

$$X^+ \triangleq \left\{ x \middle| f_j(x) \ge 0, \; j = 1, ..., m; \; x \in R^n \right\}$$

*має непусту внутрішність, замикання якої співпадає з* $X^+$.

*Тоді:*

1) *якщо алгоритм 4 зупиняється в точці* $x^{\nu}$, *то*

$$g_{\alpha_\nu 0}(x^{\nu})=0 \ \ i \ \ \sum_{j=1}^{l} f_j(x^{\nu})=0;$$

2) *якщо алгоритм 4 генерує нескінченну послідовність* $\left\{x^k\right\}_{k=0}^{\infty}$ *зі зростаючими значеннями* $\alpha_k$, $k=k_1,k_2,...,k_t$ *так, що після* $k_t$ *значення* $\alpha_k$ *зберігаються сталими і рівними* $\bar{\alpha}$, *то будь-яка точка накопичення* $x^*$ *послідовності* $\left\{x^k\right\}_{k=0}^{\infty}$ *із* $\Im_\varepsilon(x^*)\equiv\Im_0(x^*)$ *задовольняє умовам:*

$$g_{\bar{\alpha}0}(x^*)=0\,,\ \sum_{j=1}^{l} f_j(x^*)=0;$$

3) *якщо алгоритм 4 генерує нескінченну послідовність* $\left\{x^k\right\}_{k=0}^{\infty}$, *для якої значення* $\alpha_k$ *зростають нескінченно часто при* $k=k_1,k_2,...,$ *то послідовність* $\left\{x^{k_i}\right\}_{i=1}^{\infty}$ *не має точки накопичення.*

***Теорема 4'.*** *Нехай виконуються всі умови теореми 4, і нехай* $\tilde{X}$ – *компактна підмножина в* $X^+$. *Тоді існує таке число* $\tilde{\alpha}\in(0;\infty)$, *що якщо* $\hat{x}\in\tilde{X}$ *точка локального (глобального) мінімуму наступної задачі:*

$$\text{знайти } \arg\min_{x\in X^+}\left(f_0(x)+\alpha\sum_{j=1}^{l} f_j(x)\right),\ \alpha>\tilde{\alpha}, \tag{6.37}$$

*то* $\hat{x}$ *також точка локального (глобального) мінімуму задачі 4.*

***Теорема 4''.*** *Нехай виконуються всі умови теореми 4. Тоді якщо* $\hat{x}$ *розв'язок задачі (6.37), то* $g_{\alpha 0}(\hat{x})=0$, *де* $g_{\alpha 0}(\hat{x})$ *визначається згідно (6.36) при* $\varepsilon=0$.

#### 5. Аналог методу можливих напрямків у задачах мінімізації майже диференційованих функцій

Задача 5. Знайти $\arg\min_{x\in X} f_0(x)$ для заданої функції $f_0: R^n\to R^1$ і заданої множини обмежень

$$X=\left\{x \,\middle|\, f_j\left(x\right)\le 0,\ \ j=\overline{1,m},\ \ x\in R^n\right\}. \tag{6.38}$$

*Припущення 5.* $(\boldsymbol{i})-$ $f_0$ майже диференційована функція; $(\boldsymbol{ii})-$ функції $f_j$, $j=\overline{1,m}$ – неперервно диференційовані.

***Алгоритм 5***

Початок. I. Вибрати: початкове наближення $x^0\in X$, довільний вектор $z^0\in R^n$; покласти $k=0$.

Основний цикл. II. Знайти числа $\rho_k''$, $\gamma_k$, $\alpha_k$, $\varepsilon_k$, які задовольняють умовам теореми 5.

III. Знайти множину індексів $\Im(x^k, \varepsilon_k)$, яка визначається співвідношенням

$$\Im(x^k, \varepsilon_k) = \left\{ j \middle| -\varepsilon_k \le f_j(x^k) \le 0,\ j = \overline{1,m} \right\}.$$

IV. Обчислити $\tilde{x}_i^k$, $i = \overline{1,n}$ – незалежні реалізації випадкових величин, рівномірно розподілених на відрізках:

$$\left[ x_i^k - \alpha_k/2,\ x_i^k + \alpha_k/2 \right],\ i = \overline{1,n}.$$

V. Обчислити вектор:

$$\theta(x^k, k) = \frac{1}{\alpha_k} \sum_{i=1}^{n} \Big[ f_0(\tilde{x}_1^k, ..., x_i^k + \alpha_k/2, ..., \tilde{x}_n^k) - \\ - f_0(\tilde{x}_1^k, ..., x_i^k - \alpha_k/2, ..., \tilde{x}_n^k) \Big] e^i, \tag{6.39}$$

де $e^i$, $i = \overline{1,n}$ – $i$-й орт.

VI. Знайти вектор $h = h^k$ – розв'язок наступної задачі лінійного програмування:

$$\text{знайти} \max_{(h,\beta)} \beta$$

при обмеженнях:

$$\left( z^k, h \right) + \beta \le 0;$$

$$\left( \nabla f_j(x^k), h \right) + \beta \le 0,\ j \in \Im\left( x^k, \varepsilon_k \right);\ -1 \le h_j \le 1,\ j = \overline{1,n}.$$

VII. Знайти число $\rho_k'$ з умови

$$\rho_k' = \max\left\{ \rho \middle| x^k + \rho h^k \in X \right\}.$$

VIII. Обчислити кроковий множник

$$\rho_k = \min\left\{ \rho_k', \rho_k'' \right\}.$$

IX. Обчислити наступне наближення

$$x^{k+1} = x^k + \rho_k h^k.$$

X. Обчислити вектор

$$z^{k+1} = z^k + \gamma_k \left( \theta(x^k, k) - z^k \right).$$

XI. Покласти $k = k + 1$ і перейти на крок II.

***Теорема 5.*** *Нехай виконані припущення 5 та* (***iii***)– *область* $X$, *що задається обмеженнями (6.38), обмежена;* (***iv***)– *для задачі 5 виконуються умови Слейтера;* (***v***)– *числа* $\rho_k''$, $\gamma_k$, $\alpha_k$, $\varepsilon_k$, $k = 0,1,...$, *задовольняють наступним умовам:*

$$\varepsilon_k > 0,\ \alpha_k > 0,\ \rho_k'' > 0,\ \gamma_k > 0,\ k = 0,1,...;$$

$$\sum_{k=0}^{\infty}\rho_k''=\infty,\ \sum_{k=0}^{\infty}\left(\rho_k''/\alpha_k\right)^2<\infty;\ \sum_{k=0}^{\infty}\gamma_k^2<\infty;$$

$$\rho_k''/\alpha_k\gamma_k\to 0\ \ при\ \ k\to\infty;\ \varepsilon_k\to 0\ \ при\ \ k\to\infty;$$

$$\frac{|\alpha_k-\alpha_{k+1}|}{\rho_k''}\to 0\ при\ k\to\infty;\ \alpha_k\to 0\ \ при\ \ k\to\infty.$$

*Тоді з ймовірністю 1 граничні точки послідовності* $\left\{x^k\right\}_{k=0}^{\infty}$, *яка породжена алгоритмом 5, належать множині* $X^*$ *розв'язків задачі 5 і послідовність* $\left\{f_0\left(x^k\right)\right\}_{k=0}^{\infty}$ *майже напевно збігається.*

*Зауваження 5.* Якщо існує спосіб обчислення майже градієнта $g^0(x)$ функції $f_0$ в точці $x$, то вектор $\theta(x^k,k)$ в (6.39) можна визначати за формулою:

$$\theta(x^k,k)=g^0(\tilde{x}^k)\,,$$

де $\tilde{x}^k$ – реалізація випадкової точки, рівномірно розподіленої в $n$-вимірному кубі зі стороною $\alpha_k$ і центром у точці $x^k$.



**6. Стохастичний аналог методу можливих напрямків**

З а д а ч а 6. Знайти $\arg\min\limits_{x\in X} f_0(x)$ для множини, яка визначається нерівностями:

$$f_i(x)\le b_i,\ i\in\Im_1;\qquad \left(a^i,x\right)\le b_i,\ i\in\Im_2;$$

$$0\le x_i\le c_i,\ i=\overline{1,n},$$

де

$$\Im_1\subset\Im\triangleq\{1,...,m\},\ \Im_2=\Im-\Im_1.$$

*Припущення 6.* $(\boldsymbol{i})$– функції $f_0$ і $f_i$, $i\in\Im_1$ – опуклі й неперервно диференційовані; $(\boldsymbol{ii})$– множина $X$ – обмежена.

Стохастичний аналог методу можливих напрямків можна застосувати в тих випадках, коли або аналітичний вигляд цільової функції не відомий, або значення даних функцій обчислюються з похибками. Для одержання напрямку руху $h^k$ на $k$-й ітерації розв'язують допоміжну задачу лінійного програмування.

***Алгоритм 6***

П о ч а т о к . I. Вибрати: довільні початкові наближення $x^0\in X$, $z^0\in R^n$; початкові значення крокових множників $\rho_0$, $\gamma_0$, які задовольняють умовам теореми 6; додатну константу $\alpha$; $\theta_i$, $i\in\Im_1$ – деякі додатні числа.

II. Покласти $k=0$.

Основний цикл. III. Обчислити $\varepsilon_k = \alpha\rho_k$.

IV. Обчислити наступні множини індексів:

$$\Im_1\left(x^k,\varepsilon_k\right)=\left\{i\in\Im_1 \,\middle|\, b_i-\varepsilon_k\le f_i(x^k)\le b_i\right\};$$

$$\Im_2\left(x^k,\varepsilon_k\right)=\left\{i\in\Im_2 \,\middle|\, b_i-\varepsilon_k\le (a^i,x^k)\le b_i\right\};$$

$$\Im_3\left(x^k,\varepsilon_k\right)=\left\{j\in\{1,...,n\} \,\middle|\, 0\le x_j^k\le\varepsilon_k\right\};$$

$$\Im_4\left(x^k,\varepsilon_k\right)=\left\{j\in\{1,...,n\} \,\middle|\, c_i-\varepsilon_k\le x_j^k\le c_i\right\}.$$

V. Обчислити вектор $h=h^k$ – розв’язок допоміжної задачі лінійного програмування:

$$\text{знайти } \max_{(h,\delta)}\delta$$

при обмеженнях:

$$\left(\nabla f_i(x^k),h\right)+\theta_i\delta\le 0,\ i\in\Im_1(x^k,\varepsilon_k);\quad \left(a^i,h\right)\le 0,\ i\in\Im_2\left(x^k,\varepsilon_k\right);$$

$$h_j\ge 0,\ j\in\Im_3(x^k,\varepsilon_k);\quad h_j\le 0,\ j\in\Im_4\left(x^k,\varepsilon_k\right);$$

$$-1\le h_j\le 1,\ j\in\{1,...,n\};\ \left(z^k,h\right)+\delta\le 0.$$

VI. Обчислити наступне наближення

$$x^{k+1}=x^k+\rho_k h^k.$$

VII. Обчислити вектор $\xi^k$ – реалізацію випадкового вектора $\xi$, умовне математичне сподівання якого дорівнює

$$E\left(\xi/(x^0,z^0),...,(x^k,z^k)\right)=\beta\nabla f_0(x^k)+r^k,$$

*де* $\beta$ – *деяке додатне число;* $r^k$ – *обмежений випадковий вектор похибки, вимірний відносно* $\sigma$*-підалгебри, індукованої величинами* $(x^0,z^0),...,(x^k,z^k)$.

VIII. Обчислити вектор $z^{k+1}=z^k+\gamma_k(\xi^k-z^k)$.

IX. Обчислити крокові множники $\rho_{k+1}$ та $\gamma_{k+1}$, які задовольняють умовам теореми 6.

X. Покласти $k=k+1$ і перейти на крок III.

***Теорема 6.*** *Нехай виконані припущення 6 і, крім того, мають місце умови:* $(\boldsymbol{i})$ – *градієнт функції* $f_0$ *задовольняє умову Ліпшиця*

$$\left\|\nabla f_0(x)-\nabla f_0(y)\right\|\le\mu\left\|x-y\right\|,\ \mu<\infty;$$

$(\boldsymbol{ii})$ – *для задачі 6 виконуються умови Слейтера;* $(\boldsymbol{iii})$ – *константа* $\alpha$ *така, що* $x^k$, $k=0,1,...,$ *належать допустимій області* $X$; $(\boldsymbol{iv})$ – *крокові множники* $\rho_k$ *і* $\gamma_k$, *вимірні відносно* $\sigma$*-підалгебри* $\mathcal{B}_k$, *індукованої сімейством випадкових величин* $(x^0,z^0,...,x^k,z^k)$ *такі, що для деяких чисел* $\lambda_k$ *:*

$$\gamma_k \left\| r^k \right\| + 2\rho_k - \lambda_k \gamma_k \le 0 \text{ *майже напевно;*}$$

$$\sum_{k=0}^{\infty} E\rho_k^2 < \infty, \ \sum_{k=0}^{\infty} E\gamma_k^2 < \infty; \ \ \sum_{k=0}^{\infty} E\left(\gamma_k \left\| r^k \right\| + \rho_k\right)\lambda_k < \infty;$$

$$\sum_{k=0}^{\infty} \rho_k = \infty, \ E\left(\left\|\xi^k\right\|^2 \Big/ \mathcal{B}_k\right) < \infty, \ E\left\|z^0\right\|^2 < \infty.$$

*Тоді послідовність* $\left\{x^k\right\}_{k=0}^{\infty}$, *яка породжена алгоритмом 6, є такою, що* $\left\{f_0(x^k)\right\}_{k=0}^{\infty}$ *збігається майже напевно до* $f_0^* \triangleq \min_{x \in X} f_0(x).$

**7. Методи можливих напрямків для відшукання точок локальних мінімумів неопуклої функції на неопуклій множині**

З а д а ч а 7. Знайти $\arg\min_{x \in X} f_0(x)$ для заданої функції $f_0 : R^n \to R^1$ і заданої множини

$$X = \left\{ x \middle| f_j(x) \ge 0, \ j = \overline{1,m}, \ x \in R^n \right\}.$$

*Припущення* 7. $(\boldsymbol{i})-$ функції $f_i(x), \ i = \overline{0,m}-$ неперервно диференційовані; $(\boldsymbol{ii})- \inf_{x \in X} f_0(x) > -\infty.$

На кожній ітерації наведеного тут алгоритму розв'язується (наближено) деяка допоміжна задача опуклого програмування (яка може бути замінена задачею лінійного програмування). Нескінченна послідовність наближень $\left\{x^k\right\}_{k=0}^{\infty}$ збігається до множини стаціонарних точок задачі 7, що містить точки, у яких виконуються необхідні умови локального мінімуму.

***Алгоритм*** *7*

П о ч а т о к . I. Вибрати довільне початкове наближення $x^0 \in X$.

II. Вибрати величину $\varepsilon_0$ з півінтервалу $\left(0;1\right]$.

III. Покласти $k = 0$.

О с н о в н и й ц и к л . IV. Обчислити множину індексів

$$\Im\left(x^k, \varepsilon_k\right) = \left\{ j \middle| 0 \le f_j(x^k) \le \varepsilon_k, \ j \in \left[1 : m\right] \right\},$$

і покласти $\Im = \Im\left(x^k, \varepsilon_k\right)$.

V. Розв'язуючи допоміжну задачу:

$$\text{знайти } \max_{(h,\sigma)} \sigma \tag{6.40}$$

при обмеженнях:

$$\left(\nabla f_j(x^k),h\right)+\sigma\le 0,\;\; j\in\Im;$$
$$-\left(\nabla f_0(x^k),h\right)+\sigma\le 0;\;\left(h,h\right)\le 1,\qquad(6.41)$$

обчислити допустимі значення $\sigma_k$ та $h^k$, $\left\|h^k\right\|=1$ такі, що

$$\sigma_k\ge\xi_k\tilde{\sigma}\left(x^k,\varepsilon_k\right),$$

де $0<\xi\le\xi_k\le 1$; $\tilde{\sigma}\left(x^k,\varepsilon_k\right)$, $\tilde{h}\left(x^k,\varepsilon_k\right)-$ точні розв'язки допоміжної задачі (6.40) – (6.41).

VI. Якщо $\sigma_k\ge\varepsilon_k$, то перейти на крок VII; якщо $\sigma_k=0$, то перейти на крок X; інакше (якщо $0<\sigma_k<\varepsilon_k$) покласти:

$$x^{k+1}=x^k;\;\varepsilon_{k+1}=\gamma_k\varepsilon_k;\;k=k+1,$$

де $0<\gamma_k\le\gamma<1$, і перейти на крок IV.

VII. Розв'язуючи задачу одновимірної мінімізації функції $f_0\left(x^k-\rho h^k\right)$ на відрізку $\left[0;\zeta_k\right]$, обчислити величину $\rho_k$, яка задовольняє умовам:

$$f_0\left(x^k-\rho_k h^k\right)\le\left(1-\lambda_k\right)f_0(x^k)+\lambda_k w_k,\;\;0<\lambda\le\lambda_k\le 1;$$
$$\omega_k=\inf_{\rho\in[0;\zeta_k]}f_0\left(x^k-\rho h^k\right),$$

де $\zeta_k-$ довжина найбільшого відрізка $\left[x^k;x^k-\zeta_k h^k\right]$, який належить множині $X$.

VIII. Обчислити наступне наближення

$$x^{k+1}=x^k-\rho_k h^k.$$

IX. Покласти $\varepsilon_{k+1}=\varepsilon_k$, $k=k+1$ і перейти на крок IV.

X. Обчислити множину індексів

$$\Im\left(x^k,0\right)=\left\{j\,\middle|\,f_j(x^k)=0,\;\;j\in\left[1:m\right]\right\}$$

і покласти $\Im=\Im\left(x^k,0\right)$.

XI. Знайти величину $\tilde{\sigma}\left(x^k,0\right)-$ розв'язок допоміжної задачі (6.40) – (6.41).

XII. Якщо $\tilde{\sigma}\left(x^k,0\right)=0$, то покласти $x^*=x^k$ і припинити обчислення; інакше перейти на крок XIII.

XIII. Покласти $x^{k+1}=x^k$, $\varepsilon_{k+1}=\gamma_k\varepsilon_k$, де $0<\gamma_k\le\gamma<1$; $k=k+1$ і перейти на крок IV.

***Теорема 7.*** *Якщо виконуються припущення 7 та (**iii**)– функції $f_i(x)$, $i=\overline{0,m}$ належать класу $C^{1,1}(X)$ (якщо $\psi\in C^{1,1}(X)$, то існує таке число $\alpha>0$, що для будь-якого відрізка $\left[x;y\right]$, який цілком належить* *множині* $X$, *буде виконуватися* $\|\nabla\psi(x)-\nabla\psi(y)\|\le\alpha\|x-y\|$); $(\boldsymbol{iv})-$ *існує таке число* $\delta>0$, *що* $\|\nabla f_j(x)\|\le\delta$ *для всіх* $x\in X$, $j=\overline{1,m}$; $(\boldsymbol{v})-$ *множина*

$$X^*=\left\{x^*\middle|\tilde{\sigma}\left(x^*,0\right)=0,\ x^*\in X\right\}$$

*непорожня (тут* $\tilde{\sigma}\left(x^*,0\right)-$ *розв'язок допоміжної задачі (6.40) – (6.41) при* $x^k=x^*$, $\Im=\Im\left(x^*,0\right)$, *де*

$$\Im\left(x^*,0\right)=\left\{j\middle|f_j\left(x^*\right)=0,\ j\in[1:m]\right\});$$

$(\boldsymbol{vi})-$ *послідовність* $x^k$, $k=0,1,...,$ *яка породжена алгоритмом 7, компактна, то у випадку нескінченної послідовності* $x^k$, $k=0,1,...,$ *справедливе граничне співвідношення*

$$\lim_{k\to\infty}\inf_{x^*\in X^*}\left\|x^k-x^*\right\|=0.$$

*Зауваження 7.* Якщо в допоміжній задачі (6.40) – (6.41) замість обмеження $(h,h)\le1$ поставити обмеження

$$\max_{j\in[1:n]}\left|h_j\right|\le1 \tag{6.42}$$

(в такому випадку допоміжна задача стає задачею лінійного програмування), тоді для такого алгоритму теорема 7 залишається в силі.

## 6.7. Методи центрів

Задача 0. Знайти $\arg\min\limits_{x\in R^n}f_0(x)$ для заданої функції $f_0:R^n\to R^1$ і заданої множини

$$X\triangleq\left\{x\middle|f_j(x)\le0,\ j=\overline{1,m},\ x\in R^n\right\}.$$

*Припущення 0.* $(\boldsymbol{i})-$ функції $f_j$, $j=\overline{0,m}$ – неперервно диференційовані; $(\boldsymbol{ii})-$ існує така точка $x^0\in X$, що множина

$$X'(x^0)=\left\{x\middle|f_0(x)-f_0(x^0)\le0,\ f_j(x)\le0,\ j=\overline{1,m}\right\},$$

компактна і має внутрішність.

Методи центрів за своєю ідеєю тісно пов'язані як з методами штрафних функцій, так і з методами можливих напрямків – їх можна трактувати як бар'єрні методи, що не містять параметрів і засновані на внутрішніх штрафних функціях, або як методи можливих напрямків, що не містять параметрів.

У методах центрів за заданою точкою $x^k$ будується наступна точка $x^{k+1}$ у «центрі» множини $X'(x^k)$, тобто у деякій множині точок, найбільш «віддалених» від границі множини $X'(x^k)$. У якості функції відстані між точками $x$ та $x'$ можна взяти функцію

$$d(x',x)=\max\left\{f_0(x')-f_0(x);\ f_j(x'),\ \ j=\overline{1,m}\right\}. \qquad (6.43)$$

Наступна точка $x^{k+1}$ обчислюється з умови

$$d(x^{k+1},x^k)=\min_x\{d(x,x^k)\mid x\in R^n\}. \qquad (6.44)$$

Таким чином, як і методи штрафних функцій, методи центрів розбивають процедуру розв'язування задачі 0 на послідовність задач мінімізації без обмежень.

**1. Основний варіант**

***Алгоритм 1***

Початок. I. Знайти точку $x^0\in X$, яка задовольняє умову (***ii***) припущення 0.

II. Покласти $k=0$.

Основний цикл. III. Знайти таку точку $x^1\in R^n$, що

$$d(x^1,x^k)=\min_x\left\{d(x,x^k)\right\}. \qquad (6.45)$$

IV. Якщо $d(x^1,x^k)=0$, то покласти $x^*=x^k$ і припинити обчислення; інакше перейти на крок V.

V. Покласти $x^{k+1}=x^1$, $k=k+1$ і перейти на крок III.

***Теорема 1.*** *Нехай виконані припущення 0, і нехай, крім того,* (***iii***)*– множина X має внутрішність та її замикання співпадає з X, тобто* $\overline{\operatorname{int}X}=X$*;* (***iv***)*–* $f_j(x)<0$ *для всіх* $x\in\operatorname{int}X$, $j=\overline{1,m}$*. Якщо послідовність* $\left\{x^k\right\}_{k=0}^{\infty}$*, яка породжена алгоритмом 1 – нескінченна, то вона має граничні точки і всі вони оптимальні для задачі 0, а якщо послідовність* $\left\{x^k\right\}$ *– скінченна, то її останній елемент також є оптимальним для задачі 0.*

На практиці було з'ясовано, що алгоритм 1 – неефективний здебільшого через складнощі обчислення точки $x^1$, яка визначається відповідно до (6.45), що вимагає дуже великих витрат часу. Це призвело до необхідності розробки методів, які переривали б пошук точки $x^1$, коли вже знайдена точка $x''$, що є „$\varepsilon$-центром”, де $\varepsilon$ обирається по можливості малим. Опишемо модифікований метод, що перериває пошук центру $x^1$ множини $X'(x^k)$ після єдиної ітерації процедури пошуку.

**2. Модифікований метод центрів**

***Алгоритм 2***

Початок. I. Знайти точку $x^0\in X$, яка задовольняє умову (***ii***) припущення 0.

II. Покласти $k = 0$.

Основний цикл. III. Покласти $x = x^k$.

IV. Обчислити розв'язок $h_0 = h_0(x)$, $h = h(x)$ задачі лінійного програмування:

$$\text{знайти } \min_{(h_0, h)} h_0 \tag{6.46}$$

при обмеженнях:

$$-h_0 + (\nabla f_0(x), h) \le 0; \tag{6.47}$$

$$-h_0 + f_j(x) + (\nabla f_j(x), h) \le 0, \quad j = \overline{1,m}; \tag{6.48}$$

$$|h_i| \le 1, \; i = \overline{1,n}, \;\; (h = (h_1, ..., h_n)^T). \tag{6.49}$$

V. Якщо $h_0(x) = 0$, то покласти $x^* = x^k$ і припинити обчислення (якщо $x^*$ – оптимальний розв'язок задачі 0, то:

$$h_0(x^*) \triangleq \min_{h \in S} \left( \max \left\{ (f_0(x^*), h); \; f_j(x^*) + (\nabla f_j(x^*), h), \; j = \overline{1,m} \right\} \right) = 0,$$

де $S \triangleq \left\{ h \in R^n \middle| \, | h_i | \le 1, \; i = \overline{1,n} \right\}$); інакше перейти на крок VI.

VI. Обчислити найменше додатне значення $\mu(x)$, яке задовольняє умову:

$$d\left(x + \mu(x)h(x), \, x\right) = \min_{\mu \ge 0} d\left(x + \mu h(x), \, x\right), \tag{6.50}$$

де $d(\cdot,\cdot)$ визначається за формулою (6.43).

VII. Покласти $x^k = x + \mu(x)h(x)$, $k = k + 1$ і перейти на крок III.

*Зауваження 2*. Вектор $h(x)$, який обчислюється на кроці IV алгоритму 2, неоднозначно визначається з розв'язку (6.46) – (6.49), оскільки розв'язок задачі лінійного програмування може не бути єдиним. Вважається, що $h(x)$ – це будь-який вектор, який є оптимальним розв'язком задачі; його можна отримати методами, описаними у розділі 5.

*Зауваження 2'*. Точку $x^0 \in X$, тобто початкове наближення на кроці I алгоритму 2 можна обчислювати, застосовуючи алгоритм 2 до наступної задачі (в просторі векторів $(x_0, x)$):

$$\min_{(x_0, x)} \{ x_0 \,|\, f_j(x) - x_0 \le 0, \quad j = \overline{1,m} \},$$

причому замість допустимого початкового розв'язку для цієї задачі можна взяти $\overline{x}^0 \triangleq (x_0^0, x^0) \in R^{n+1}$, де точка $x^0 \in R^n$ – довільна і $x_0^0 = \max\{ f_j(x^0), \; j \in \{1, \ldots, m\} \}$. Знайдеться таке скінченне ціле число $i$, що $\overline{x}^i = (x_0^i, x^i)$ буде задовольняти умовам $x_0^i < 0$, $x^i \in X$. Тоді покласти $x^0 = x^i$ і перейти до розв'язку задачі 0 за допомогою алгоритму 2.

***Теорема 2.*** *Нехай виконані припущення 0. Тоді послідовність $x^0, x^1, ...$, яка побудована алгоритмом 2, або скінченна та її останній елемент*

*$x^* = x^k$ задовольняє умову $h_0(x^*) = 0$, або нескінченна і кожна її гранична точка $x^*$ задовольняє умові $h_0(x^*) = 0$.*

**3. Реалізація модифікованого методу центрів з використанням методу золотого перерізу**

Розглянемо, як замінити принципову операцію (6.50) в алгоритмі 2 іншою, яку можна було б реалізувати на ЕОМ. Найпростіший – це випадок, коли всі функції $f_j(\cdot)$, $j = \overline{0,m}$ – опуклі. Легко бачити, що в цьому випадку функція $d(\cdot, x)$, яка визначається відповідно (6.43) при $x \in X$, також опукла.

***Алгоритм 3***

Початок. I. Вибрати числа $\bar{\varepsilon}_0 > 0$, $\rho \ge 1$, $\eta > 1$ (рекомендується $\eta = 2$).

II. Обчислити початкове наближення $x^0 \in X$.

III. Покласти $k = 0$.

Основний цикл. IV. Покласти $x = x^k$.

V. Розв'язати задачу лінійного програмування (6.46–6.49) і отримати $(n+1)$-вимірний вектор $(h_0(x), h(x))$.

VI. Якщо $h_0(x) = 0$, то покласти $x^{k+1} = x^k$ і припинити обчислення; інакше перейти на крок VII.

(Наступні кроки VII–X замінюють крок VI алгоритму 2 для обчислення величини $\mu(x)$).

VII. Для $\lambda \ge 0$ покласти

$$\theta(\lambda) = d(x + \lambda h(x), x). \tag{6.51}$$

VIII. Покласти $\varepsilon_i = \bar{\varepsilon}_0$, $i = 0$.

IX. Використати для обчислення $\bar{\mu}$ алгоритм пошуку мінімуму одновимірної функції за методом золотого перерізу (алгоритм 2Б, п. 3.4), де $\theta(\cdot)$ визначається рівністю (6.51), $\rho$ визначається на кроці I і $\varepsilon = \varepsilon_i$.

X. Якщо $d(x + \bar{\mu} h(x), x) \le -\varepsilon_i$, то покласти $\mu(x) = \bar{\mu}$ і перейти на крок XI; інакше покласти $\varepsilon_{i+1} = \varepsilon_i / \eta$, $i = i + 1$ і перейти на крок IX.

XI. Покласти $x^{k+1} = x + \mu(x) h(x)$, $k = k + 1$ і перейти на крок IV.

***Теорема 3***. *Нехай виконані припущення 0, і нехай функції $f_j(\cdot)$, $j = \overline{0,m}$ – опуклі донизу. Тоді послідовність $x^0, x^1, \ldots$, яка побудована алгоритмом 3, або скінченна і її останній елемент $x^{k+1} = x^k$ задовольняє умову $h_0(x^k) = 0$, або нескінченна і кожна її гранична точка $x^*$ задовольняє умову $h_0(x^*) = 0$.*

*Зауваження 3.* Можна побудувати інші реалізації алгоритму 2, змінюючи операцію (6.50) на кроці VI довільним алгоритмом пошуку мінімуму (або стаціонарної точки) дійсної функції однієї змінної.

**4. Реалізація модифікованого методу центрів з використанням функцій переривання**

На практиці алгоритми наведені нижче ефективні в ситуації, коли вони по суті вирішують багато разів одну і ту ж задачу. У цьому випадку корисно завчасно витратити певний час на ретельний підбор «функції переривання», яка буде використовуватись, і зробити спробу отримати алгоритм, який працює швидше ніж алгоритм 3.

*Означення 1.* Нехай $N = \{0,1,2,...\}$. Будемо називати функцію $t : N \to N$ ***функцією переривання***, якщо для кожного $i \in N$ знайдеться таке число $i' \in N$, що $t(k) > i$ при всіх $k > i'$.

***Алгоритм 4***

Початок. I. Обчислити початкове наближення $x^0 \in X$ та покласти $k = 0$.

II. Вибрати функції переривання $t_1(\cdot),\ t_2(\cdot)$.

Основний цикл. III. Покласти $x = x^k,\ i = t_1(k)$.

IV. Розв'язати задачу лінійного програмування (6.46)–(6.49) і отримати $(n+1)$ - вимірний вектор $(h_0(x),\ h(x))$.

V. Якщо $h_0(x) = 0$, то покласти $x^{k+1} = x^k$ і припинити обчислення; інакше перейти на крок VI.

VI. Покласти $\theta(\mu) = d(x + \mu h(x), x)$ та використати для обчислення $\bar{\mu}$ алгоритм $4'$ – модифікацію алгоритму пошуку мінімуму одновимірної функції за методом золотого перерізу.

VII. Якщо $f_0\left(x + \bar{\mu}h(x)\right) < f_0(x)$, то покласти $x^{k+1} = x + \bar{\mu}h(x),\ \ k = k+1$ і перейти на крок III; інакше покласти $x^{k+1} = x^k,\ i = t_2(k),\ k = k+1$ і перейти на крок VI.

***Алгоритм 4′*** **(*модифікація алгоритму пошуку за методом золотого перерізу*)**

I. Вибрати $\rho > 0$, обчислити дроби Фібоначі

$$F_1 = (3-\sqrt{5})/2,\ \ F_2 = (\sqrt{5}-1)/2.$$

II. Обчислити $\theta(0),\ \theta(\rho)$.

III. Якщо $\theta(\rho) \ge \theta(0)$, то покласти $a_0 = 0,\ b_0 = \rho$ і перейти на крок VIII; інакше перейти на крок IV.

IV. Покласти $l = 0,\ \ \mu_0 = 0.$

V. Покласти $\mu_{l+1} = \mu_l + \rho$.

VI. Обчислити $\theta(\mu_{l+1})$.

VII. Якщо $\theta(\mu_{l+1}) \geq \theta(\mu_l)$, то покласти $a_0 = \mu_{l-1}$, $b_0 = \mu_{l+1}$ та перейти на крок VIII; інакше покласти $l = l+1$ і перейти на крок V.

*Примітка.* Тепер точка мінімуму належить $[a_0;b_0]$.

VIII. Покласти $j = 0$.

IX. Покласти $v_j = a_j + F_1(b_j - a_j)$, $w_j = a_j + F_2(b_j - a_j)$.

X. Якщо $\theta(v_j) < \theta(w_j)$, то покласти $a_{j+1} = a_j$, $b_{j+1} = w_j$ та перейти на крок XI; інакше покласти $a_{j+1} = v_j$, $b_{j+1} = b_j$ та перейти на крок XI.

XI. Якщо $j < i$, то покласти $j = j+1$ та перейти на крок IX; інакше покласти $\overline{\mu} = (a_j + b_j)/2$ і припинити обчислення.

***Теорема 4.*** *Нехай виконані припущення 0 і функції $f_j(\cdot)$, $j = \overline{0,m}$ – опуклі донизу. Тоді послідовність $x^0, x^1, \ldots$, яка побудована алгоритмом 4, або скінченна і її останній елемент $x^{k+1} = x^k$ задовольняє умову $h_0(x^k) = 0$, або нескінченна і тоді кожна її гранична точка $x^*$ задовольняє умову $h_0(x^*) = 0$.*



## 6.8. Методи типу Ньютона

### 1. Методи типу Ньютона з регулюванням кроку

Задача 1. Знайти $\arg\min_{x \in X} f_0(x)$ для заданої функції $f_0 : R^n \to R^1$ і заданої опуклої компактної множини $X \subset R^n$.

*Припущення 1.* Функція $f_0$ – двічі неперервно диференційована на опуклій компактній множині $X$.

***Алгоритм 1***

Початок. I. Вибрати початкове наближення $x^0 \in X$.

II. Вибрати константи $\varepsilon \in (0;1)$, $\beta \in (0;1)$ (рекомендується $\varepsilon = 1/2$, $\beta = 0{,}8$).

III. Покласти $k = 0$.

Основний цикл. IV. Покласти $x = x^k$.

V. Обчислити градієнт $\nabla f_0(x)$ і матрицю других похідних $\nabla^2_{xx} f_0(x)$.

VI. Обчислити точку $y^k$, яка забезпечує мінімум квадратичної функції $\psi_k(y)$ на множині $X$, де

$$\psi_k(y) \triangleq \left(\nabla f_0(x), y - x\right) + \frac{1}{2}\left(\nabla^2_{xx} f_0(x)(y-x), y - x\right). \quad (6.52)$$

Якщо $\psi_k(y^k)=0$, то завершити обчислення (в цьому випадку $y^k$ – розв'язок задачі 1); інакше перейти на крок VII.

V. Покласти $h^k = y^k - x$.

VI. Покласти $\rho^k = 1$.

VII. Якщо виконується нерівність

$$f_0(x+\rho_k h^k) - f_0(x) \le \varepsilon \rho_k \psi_k(y^k), \tag{6.53}$$

то перейти на крок XI; інакше перейти на крок X.

VIII. Покласти $\rho_k = \beta\rho_k$ і перейти на крок IX.

IX. Покласти $x^{k+1} = x + \rho_k h^k$, $k = k+1$ і перейти на крок IV.

***Теорема 1.*** *Якщо функція $f_0$ – опукла донизу і двічі неперервно диференційована на опуклій компактній множині $X$, то для нескінченної послідовності $\{x^k\}_{k=0}^{\infty}$, яка породжується алгоритмом 1, виконуються умови:*

а) $\{f_0(x^k)\}_{k=0}^{\infty}$ *монотонно спадає;*

б) $\lim\limits_{k\to\infty} f_0(x^k) = \min\limits_{x\in X} f_0(x)$.

*Якщо, крім того,*

$$a_1\|y\|^2 \le \left(\nabla_{xx}^2 f_0(x)y, y\right) \le a_2\|y\|^2,\ a_1>0,\ x\in X,\ y\in R^n, \tag{6.54}$$

*то послідовність $\{x^k\}_{k=0}^{\infty}$ збігається до розв'язку задачі 1 із надлінійною швидкістю, тобто:*

$$\left\|x^{l+i} - x^*\right\| \le \nu_1\lambda_l \ldots \lambda_{l+i}, \tag{6.55}$$

*де $l, \nu_1 < \infty$, $\lambda_{l+i} < 1$ при довільному $i \ge 0$, $\lambda_j \to 0$ при $j \to \infty$.*

***Теорема 1'.*** *Якщо виконується умова (6.54) і матриця $\nabla_{xx}^2 f_0(x)$ на опуклій компактній множині X задовольняє умову Ліпшиця з константою $\gamma$, тобто*

$$\left\|\nabla_{xx}^2 f_0(x) - \nabla_{xx}^2 f_0(y)\right\| \le \gamma\|x-y\|,\ \forall x,y\in X,\ \gamma<\infty,$$

*то нескінченна послідовність $\{x^k\}_{k=0}^{\infty}$, яка породжується алгоритмом 1, збігається до розв'язку $x^*$ задачі 1 з квадратичною швидкістю, тобто*

$$\left\|x^{l+i} - x^*\right\| \le \nu_2\beta_l^{2^i},\ \nu_2<\infty,\ l<\infty,\ \beta_l<1.$$

Наведений нижче алгоритм є принциповою схемою для алгоритму 1.

***Алгоритм 1'***

I–VII. Кроки I–VII такі, як і в алгоритмі 1.

VIII. Обчислити число $\rho_k \in [0;1]$, яке задовольняє умову

$$f_0(x+\rho_k h^k) = \min_{\rho\in[0;1]} f_0(x+\rho h^k).$$

IX. Покласти $x^{k+1} = x + \rho_k h^k$, $k = k+1$ і перейти на крок IV.

***Теорема 1''.*** *Якщо виконується умова (6.54) і множина $X$ – опуклий компакт, то нескінченна послідовність $\{x^k\}_{k=0}^{\infty}$, яка породжується алгоритмом 1', збігається до розв'язку задачі 1 із надлінійною швидкістю, тобто має місце оцінка (6.55), причому $\rho_k \to 1$ при $k \to \infty$.*

**2. Метод типу Ньютона при наявності збурень**

Задача 2. Знайти $\arg\min_{x\in X} f_0(x)$ для заданої функції $f_0 : R^n \to R^1$ і заданої множини $X$.

*Припущення 2.* $(\boldsymbol{i})-$ множина $X$ – замкнута і опукла; $(\boldsymbol{ii})-$ функція $f_0(x)$ – двічі неперервно диференційована на множині $X$.

На $k$-й ітерації алгоритму 2 функція $f_0(x) - f_0(x^k)$ в точці $x^k$ апроксимується квадратичним функціоналом:

$$\overline{f}_k(x) = \frac{1}{2}\left(A_k\left(x - x^k\right), x - x^k\right) + \left(h^k, x - x^k\right), \tag{6.56}$$

де $A_k$ – наближення для $\nabla^2_{xx} f_0(x^k)$; $h^k$ – наближення для $\nabla f_0(x^k)$; в якості $x^{k+1}$ вибирається точка, для якої наближено виконується умова мінімуму $\overline{f}_k(x)$ на множині $X$:

$$\left(\nabla \overline{f}_k(x^{k+1}), \frac{x^{k+1} - x}{\left\|x^{k+1} - x\right\|}\right) \triangleq \left(A_k(x^{k+1} - x^k) + h^k, \frac{x^{k+1} - x}{\left\|x^{k+1} - x\right\|}\right) \le \sigma_k \text{ для } \forall x \in X.$$

***Алгоритм 2***

Початок. I. Вибрати початкове наближення $x^0 \in X$.

II. Покласти $k = 0$.

Основний цикл. III. Обчислити матрицю $A_k -$ наближення для $\nabla^2_{xx} f_0(x^k)$.

IV. Обчислити вектор $h^k -$ наближення для $\nabla f_0(x^k)$.

V. Обчислити точку $x^{k+1} \in X$, для якої виконується нерівності:

$$\left(A_k\left(x^{k+1} - x^k\right) + h^k, \frac{x^{k+1} - x}{\left\|x^{k+1} - x\right\|}\right) \le \sigma_k, \quad \sigma_k \ge 0 \quad \forall x \in X.$$

Обчислення точки $x^{k+1}$ зводиться до знаходження наближеного мінімуму квадратичного функціоналу (6.56) на множині $X$.

VI. Покласти $k = k+1$ і перейти на крок III.

***Теорема 2.*** *Нехай виконуються припущення 2 та $(\boldsymbol{iii})-$ матриця других похідних $\nabla^2_{xx} f_0(x)$ задовольняє умову Ліпшиця з константою $\gamma$ в* *області*

$$\bar{X}_0 = \left\{x \middle| \| x - x^0 \| \le \rho,\ x \in X\right\};$$

$(\boldsymbol{iv})$ –

$$\left(\nabla^2_{xx} f_0(x) y, y\right) \ge m_1 \|y\|^2, m_1 > 0.$$

*Тоді:* 1) *якщо виконуються умови:*

$(\boldsymbol{v})$ – $\left\|A_k - \nabla^2_{xx} f_0\left(x^k\right)\right\| \le c_1$;

$(\boldsymbol{vi})$ – $\left\|h^k - \nabla f_0\left(x^k\right)\right\| \le c_2 \left\|x^k - x^{k+1}\right\|$;

$(\boldsymbol{vii})$ – $\sigma_k \le c_3 \left\|x^k - x^{k+1}\right\|,\ m_1 > c_4,\ c_4 = c_1 + c_2 + c_3$;

$(\boldsymbol{viii})$ – $q = \beta + \gamma c_4^2 \left\|x^1 - x^0\right\| / 2\left(m_1 - c_4\right) < 1,\ \beta = c_4 / \left(m_1 - c_4\right)$;

$(\boldsymbol{ix})$ – $\rho \ge r_1 \triangleq \left\|x^1 - x^0\right\| / \left(1 - q\right)$,

*то існує така точка мінімуму* $x^* \in \bar{X}_0$, *що для послідовності* $\left\{x^k\right\}_{k=0}^{\infty}$, *яка породжується алгоритмом 2, виконуються співвідношення:*

$$\left\|x^k - x^*\right\| \le r_1 q^k,\ \left\|x^k - x^*\right\| = O(\beta^k);$$

2) *якщо виконуються умови:*

$(\boldsymbol{x})$ – $\left\|A_k - \nabla^2_{xx} f_0\left(x^k\right)\right\| \le c_1 \left\|x^k - x^{k+1}\right\|$;

$(\boldsymbol{xi})$ – $\left\|h^k - \nabla f_0\left(x^k\right)\right\| \le c_2 \left\|x^k - x^{k+1}\right\|^2$;

$(\boldsymbol{xii})$ – $\sigma_k \le c_3 \left\|x^k - x^{k+1}\right\|^2,\ m_1 > c_4 \left\|x^1 - x^0\right\|,\ c_4 = c_1 + c_2 + c_3$;

$(\boldsymbol{xiii})$ – $q \triangleq \left(c_4 + \gamma / 2\right) \left\|x^1 - x^0\right\| / \left(m_1 - c_4 \left\|x^1 - x^0\right\|\right) < 1$;

$(\boldsymbol{xiv})$ – $\rho \ge \left(\left\|x^1 - x^0\right\| / q\right) \psi_0(q)$,

*то існує така точка мінімуму* $x^* \in \bar{X}_0$, *що для послідовності* $\left\{x^k\right\}_{k=0}^{\infty}$, *яка породжується алгоритмом 2, виконується:*

$$\left\|x^k - x^*\right\| \le \left(\left\|x^1 - x^0\right\| \big/ q\right) \psi_k(q),$$

*де* $\psi_k(q) = \sum_{i=k}^{\infty} q^{2^i}$, $k = 0, 1, \ldots$.

### 3. Квазіньютоновські методи

З а д а ч а 3. Знайти $\arg\min_{x \in X} f_0(x)$ для заданої функції $f_0 : R^n \to R^1$ і множини:

$$X \triangleq \left\{x \mid f_i(x) \le 0,\ i = \overline{1,m},\ x \in R^n\right\},$$

де $f_i : R^n \to R^1$, $i = \overline{1,m}$ – задані функції.

*Припущення 3.* Функції $f_j(x)$, $j = \overline{0,m}$ – двічі диференційовані в $R^n$.

На $k$-й ітерації алгоритму необхідно обчислювати матрицю, близьку до матриці других похідних по *x*, функції Лагранжа задачі 3 (також можна використовувати скінченно-різницеву апроксимацію гесіана по *x* функції Лагранжа). Наступне $(k+1)$-ше наближення $\left(x^{k+1}, u^{k+1}\right)$ по основним і двоїстим змінним задачі 3 знаходиться як розв'язок деякої допоміжної задачі квадратичного програмування.

***Алгоритм 3***

Початок. I. Вибрати будь-яке початкове наближення $(x^0, u^0) \in R^n \times R^m$.

II. Покласти $k = 0$.

Основний цикл. III. Обчислити матрицю $G(x^k, u^k)$, яка задовольняє умовам теореми 3.

IV. Знайти точку Куна-Таккера $(x^{k+1}, u^{k+1})$, розв'язуючи наступну задачу квадратичного програмування:

$$\text{знайти } \arg\min_x \left[ \left( \nabla f_0(x^k), x - x^k \right) + \frac{1}{2} \left( x - x^k, G(x^k, u^k)(x - x^k) \right) \right] \quad (6.57)$$

при обмеженнях: $f_i(x^k) + (\nabla f_i(x^k), x - x^k) \le 0, \quad i = \overline{1,m}.$

Якщо розв'язок задачі (6.57) не єдиний, то в якості $(x^{k+1}, u^{k+1})$ обирають точку, найближчу до $(x^k, u^k)$.

V. Покласти $k = k + 1$ і перейти на крок III.

***Теорема 3.*** *Нехай* $(\boldsymbol{i})$ – $\left(\bar{x}, \bar{u}\right)$ *– точка Куна–Таккера, яка задовольняє достатнім умовам другого порядку для задачі 3 (точка* $\left(\bar{x}, \bar{u}\right)$ *із* $R^n \times R^m$ *задовольняє достатнім умовам другого порядку, якщо вона задовольняє умовам Куна–Таккера першого порядку і якщо* $\nabla^2_{xx} \varphi(\bar{x}, \bar{u}) y > 0$ *для кожного ненульового вектора* $y \in R^n$, *який задовольняє умовам:*

$$\nabla f_j(\bar{x}) y = 0, \quad j \in \left\{ j \mid \bar{u}_j > 0, \quad j = \overline{1,m} \right\};$$

$$\nabla f_j(\bar{x}) y \le 0, \quad j \in \left\{ j \mid \bar{u}_j = 0, \ f_j(\bar{x}) = 0, \ j \in [1:m] \right\},$$

*де* $\nabla^2_{xx} \varphi\left(\bar{x}, \bar{u}\right)$ – *матриця других похідних по x функції Лагранжа*

$$\varphi(x, u) = f_0(x) + \sum_{i=1}^{m} u_i f_i(x), \quad u = (u_1, \ldots, u_m),$$

*обчислена в точці* $\left(\bar{x}, \bar{u}\right)$), *а також умову строгої доповнюючої нежорсткості та умову лінійної незалежності градієнтів активних обмежень (точка* $\left(\bar{x}, \bar{u}\right)$ *задовольняє умову строгої доповнюючої* *нежорсткості, якщо* $\bar{u}_j > 0$ *або* $f_j(\bar{x}) < 0$ *для* $j = \overline{1,m}$); (***ii***)$-$ $f_j(x)$, $j = \overline{0,m}$ *мають другі похідні, неперервні за Ліпшицем у відкритому околі точки* $\bar{x}$; (***iii***)$-$ *матриці* $G(x^k, u^k)$ *на кроці III алгоритму 3 задовольняють одній з умов:*

а)

$$\left\| G(x^k, u^k) - \nabla^2_{xx}\phi(x^k, u^k) \right\| \le 1/10\beta,$$

*де*

$$\beta = \frac{3}{2}\left\| \left( \nabla_z h\left(z^k\right) \right)^{-1} \right\|;$$

$$h(z) = \begin{pmatrix} \nabla f_0(x) + \nabla f(x)u \\ u_1 f_1(x) \\ \dots \\ u_m f_m(x) \end{pmatrix}, \qquad \begin{array}{c} f(x) = \left(f_1(x), \dots, f_m(x)\right), \\ z = (x, u), \\ \\ z^k = (x^k, u^k); \end{array}$$

б)

$$\left\| G(x^k, u^k) - \nabla^2_{xx}\phi(x^k, u^k) \right\| \le 1/10\beta, \tag{6.58}$$

$$\left\| \frac{(G(x^k, u^k) - \nabla^2_{xx}\varphi(x^k, u^k))(x^{k+1} - x^k)}{\left\| z^{k+1} - z^k \right\|} \right\| \to 0 \;\; \textit{при} \;\; k \to \infty;$$

в)

$$G(x^k, u^k) = \nabla^2_{xx}\varphi(x^k, u^k). \tag{6.59}$$

*Тоді існують такі додатні числа* $\delta_1$ *і* $\delta_2$, *що якщо початкове наближення* $(x^0, u^0)$ *вибрати з умов:*

$$\left\| (x^0, u^0) - (\bar{x}, \bar{u}) \right\| < \delta_1;$$

$$\left\| G(x^0, u^0) - \nabla^2_{xx}\phi(x^0, u^0) \right\| < \delta_2,$$

*то послідовність* $\left\{(x^k, u^k)\right\}_{k=0}^{\infty}$ *існує і збігається до точки* $\left(\bar{x}, \bar{u}\right)$:

1) *з лінійною швидкістю збіжності, тобто існують числа* $q \in (0;1)$, $a > 0$, $\bar{k} \ge 0$ *такі, що*

$$\left\| (x^k, u^k) - (\bar{x}, \bar{u}) \right\| \le \alpha q^k \;\; \textit{для усіх} \;\; k \ge \bar{k};$$

2) *із надлінійною швидкістю збіжності, тобто для кожного досить малого* $q \in (0;1)$ *існують* $a > 0$, $\bar{k} \ge 0$ *такі, що*

$$\left\| (x^k, u^k) - (\bar{x}, \bar{u}) \right\| \le \alpha q^k \;\; \textit{для усіх} \;\; k \ge \bar{k},$$

*при умові, що матриці* $G(x^k, u^k)$, $k = 0, 1, \dots,$ *задовольняють* (6.58);

3) *з квадратичною швидкістю збіжності, тобто існують числа* $q \in (0;1)$, $a > 0$, $\bar{k} \ge 0$ *такі, що*

$$\left\|(x^k, u^k) - (\bar{x}, \bar{u})\right\| \le \alpha q^{2^k} \ \textit{для усіх} \ k \ge \bar{k},$$

*при умові, що матриці* $G(x^k, u^k)$, $k = 0,1,...,$ *задовольняють* (6.59).

## 6.9. Методи лінеаризації

Методи лінеаризації застосовуються для відшукання розв'язку $x^*$ задачі мінімізації нелінійної функції $f_0(x)$ при досить загальних обмеженнях у вигляді рівностей $f_j(x) = 0$, $j = \overline{1,m}$ і нерівностей $f_j(x) \le 0$, $j = \overline{m+1, m+l}$. Розв'язок $x^*$ визначається як границя побудованої послідовності $x^0, x^1, \ldots, x^k, x^{k+1}, \ldots$. Покращене $(k+1)$-ше наближення $x^{k+1}$ до розв'язку $x^*$ визначається за допомогою розв'язування значно простішої допоміжної задачі мінімізації функції

$$\hat{f}_0(x) \triangleq (\nabla f_0(x^k), x) + \left(1/2\right) \| x - x^k \|^2 \tag{6.60}$$

при лінійних обмеженнях:

$$\hat{f}_j(x) \triangleq f_j(x^k) + \left(\nabla f_j(x^k), x - x^k\right) \le \varepsilon_k, \quad j \in \Im_{\delta_k}. \tag{6.61}$$

Допоміжною множиною $\Im_{\delta_k}$ виділяють тільки ті з початкових обмежень, які бажано задовольнити на $(k+1)$-му наближенні $x^{k+1}$ (наприклад, ті з обмежень, які «найбільш порушуються» на $k$-му наближенні $x^k$). Квадратична добавка $\left(1/2\right)\| x - x^k \|^2$ гарантує існування розв'язку $z(\varepsilon, \delta_k)$ допоміжної задачі на непорожній допустимій множині. Покращене $(k+1)$-ше наближення $x^{k+1}$ обчислюють за формулою

$$x^{k+1} = x^k + \rho_k (z(\varepsilon, \delta_k) - x^k).$$

Різні способи визначення крокових множників $\rho_k$ і констант $\varepsilon, \delta_k$ визначають різні варіанти методів лінеаризації.

#### 1. Обмеження типу нерівностей

Задача 1. Знайти $\arg\min_{x \in X} f_0(x)$ для заданої функції $f_0 : R^n \to R^1$ і множини

$$X = \{x \,|\, f_j(x) \le 0, \quad j \in \Im, \quad x \in R^n\},$$

яка визначена заданими функціями $f_j : R^n \to R^1, \quad j \in \Im$.

*Припущення 1.* $(\boldsymbol{i})$– функції $f_j,\ j \in \{0\} \cup \Im$ – неперервно диференційовані; $(\boldsymbol{ii})$– градієнти функцій $f_j,\ j \in \{0\} \cup \Im$ задовольняють умову Ліпшиця

$$\| \nabla f_j(x) - \nabla f_j(y) \| \le \gamma \| x - y \|, \quad j \in \{0\} \cup \Im, \quad \gamma < \infty .$$

***Алгоритм 1***

Початок. I. Вибрати довільне наближення $x^0 \in R^n$.

II. Вибрати достатньо велику константу $\alpha > 0$ і константу $\delta > 0$, які задовольняють умовам теореми 1.

III. Вибрати константу $0 < \varepsilon < 1$.

IV. Покласти $k = 0$.

Основний цикл. V. Покласти $x = x^k$.

VI. Знайти множину індексів

$$\Im_\delta(x) = \{ j \mid f_j(x) \ge \max_{j \in \Im} f_j(x) - \delta, \quad j \in \Im \} .$$

VII. Знайти розв'язок $h = h(x)$ задачі квадратичного програмування:

$$\text{знайти } \arg\min_h \left( \left( \nabla f_0(x), h \right) + \left( 1/2 \right) \| h \|^2 \right)$$

при обмеженнях: $\left( \nabla f_j(x), h \right) + f_j(x) \le 0, \quad j \in \Im_\delta(x)$.

VIII. Якщо $h(x) = 0$, то покласти $x^* = x$ і припинити обчислення; інакше перейти на крок IX.

IX. Покласти $i = 0$.

X. Покласти $\rho_k = \left( 1/2 \right)^i$.

XI. Якщо виконується нерівність

$$f_0(x + \rho_k h(x)) + \alpha \max_{j \in \Im} \{0, f_j(x + \rho_k h(x))\} \le$$

$$\le f_0(x) + \alpha \max_{j \in \Im} \{0, f_j(x)\} - \rho_k \varepsilon \| h(x) \|^2,$$

тоді перейти на крок XII; інакше покласти $i = i + 1$ і перейти на крок X.

XII. Обчислити наступне наближення

$$x^{k+1} = x + \rho_k h(x) .$$

XIII. Покласти $k = k + 1$ і перейти на крок V.

***Теорема 1.*** *Якщо виконані припущення 1 та існують константи $\alpha > 0$ і $\delta > 0$ такі, що ($\boldsymbol{iv}$) – множина*

$$X_\alpha = \{ x \mid f_0(x) + \alpha \max_{j \in \Im} f_j(x) \le f_0(x^0) + \alpha \max_{j \in \Im} f_j(x^0), \; x \in R^n \}$$

*обмежена*;

($\boldsymbol{v}$) *– задача квадратичного програмування:*

$$\textit{знайти } \arg\min_h \left( (\nabla f_0(x), h) + \frac{1}{2}(h, h) \right)$$

*при обмеженнях: $(\nabla f_j(x), h) + f_j(x) \le 0, \quad j \in \Im_\delta(x)$,*

*розв'язується відносно $h \in R^n$ при будь-якому $x \in X_\alpha$ та існують такі* *множники Лагранжа* $\lambda_j(x),\ j \in \Im_\delta(x)$, *що* $\sum_{j\in\Im_\delta(x)} \lambda_j(x) \le \alpha$; $(\boldsymbol{vi})$ *– існує такий індекс* $j \in \Im$, *що функція* $f_j(x) \equiv 0$, *тоді нескінченна послідовність* $\{x^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 1, має наступні властивості:*

$$\max_{j\in\Im} f_j(x^k) \to 0 \ \textit{при}\ k \to \infty;$$

*будь-яка гранична точка* $x^*$ *послідовності* $\{x^k\}_{k=0}^{\infty}$ *належить множині* $X$ *і в цій точці виконуються необхідні умови мінімуму функції* $f_0$ *на множині* $X$.

***Алгоритм 1'***

Початок. I. Вибрати початкове наближення $x^0 \in R^n$.

II. Вибрати константи $\delta_0 > 0$ та $\rho > 0$.

III. Покласти $k = 0$.

Основний цикл. IV. Покласти $x = x^k$.

V. Покласти $\delta = \delta_0$.

VI. Знайти множину індексів $\Im_\delta(x)$ (крок VI алгоритму 1).

VII. Знайти вектор $h(x)$ (крок VII алгоритму 1).

VIII. Якщо $h(x) = 0$, то покласти $x^* = x$ і припинити обчислення; інакше перейти на крок IX.

IX. Обчислити наступне наближення

$$x^{k+1} = x + \rho h(x).$$

X. Покласти $k = k + 1$ і перейти на крок IV.

***Теорема 1'.*** *Нехай* $x^*$ *– розв'язок задачі 1 і виконуються наступні умови:* $(\boldsymbol{i})$ *– для довільного достатньо малого* $\delta > 0$ *допоміжна задача квадратичного програмування:*

$$\textit{знайти}\ \arg\min_h \left( \left(\nabla f_0(x), h\right) + \left(1/2\right) \| h \|^2 \right)$$

*при обмеженнях:*

$$(\nabla f_j(x), h) + f_j(x) \le 0, \quad j \in \Im_\delta(x),$$

*де* $\Im_\delta(x) = \{ j \,|\, f_j(x) \ge \max_{j\in\Im} f_j(x) - \delta, \quad j \in \Im \}$ *є розв'язною;*

$(\boldsymbol{ii})$ *– функції* $f_j$ *– двічі неперервно диференційовані та градієнти* $\nabla f_j(x^*),\ j \in \Im_0(x^*)$, *де*

$$\Im_0(x^*) = \{ j \,|\, f_j(x^*) = 0, \quad j \in \Im \},$$

*лінійно незалежні;* $(\boldsymbol{iii})$ *– в точці* $x^*$ *виконуються необхідні умови мінімуму у формі:*

$$\nabla f_0(x^*) + \sum_{j\in\Im_0(x^*)} \lambda_j^0 \nabla f_j(x^*) = 0,$$

*де* $\lambda_j^0 > 0,\ j \in \Im_0(x^*)$; $(\boldsymbol{iv})$ *– виконується достатня умова локального* *мінімуму, тобто*

$$\left(h, \frac{\partial^2 \varphi(x^*, \lambda^0)}{\partial x^2} h\right) > 0$$

*для всіх* $h \neq 0$*, і котрі задовольняють умову:*

$$\left(h, \nabla f_j(x^*)\right) = 0, \quad j \in \Im_0(x^*),$$

*де*

$$\varphi(x, \lambda) = f_0(x) + \sum_{j \in \Im_0(x^*)} \lambda_j f_j(x);$$

*вектор* $\lambda^0$ *складається із компонент* $\lambda_j^0$, $j \in \Im_0(x^*)$.

*Тоді існує такий окіл* $V$ *точки* $x^*$ *і константи* $\delta_0 > 0, \alpha > 0$*, що нескінченна послідовність* $\{x^k\}_{k=0}^{\infty}$*, яка породжена алгоритмом 1', збігається до точки* $x^*$ *з довільного початкового наближення* $x^0 \in V$ *з геометричною швидкістю, тобто існує таке число* $0 \le q_1 \le 1$*, що*

$$\| x^k - x^* \| \le \beta_1 (q_1)^k, \quad \beta_1 < \infty$$

*для всіх достатньо великих значень* $k$.

***Теорема 1''.*** *Якщо виконані всі припущення теореми 1' і, крім того, кількість індексів у множині* $\Im_0(x^*)$ *дорівнює n (розмірності простору), то нескінченна послідовність* $\{x^k\}_{k=0}^{\infty}$*, яка породжена алгоритмом 1' при* $\rho = 1$*, збігається в деякому околі точки* $x^*$ *із квадратичною швидкістю, тобто*

$$\| x^{k+1} - x^* \| \le \beta_2 \| x^k - x^* \|^2, \quad \beta_2 < \infty.$$

**2. Обмеження типу рівностей**

З а д а ч а 2. Знайти $\arg\min_{x \in X} f_0(x)$ для даної функції $f_0 : R^n \to R^1$ і множини *X*, яка задається співвідношенням

$$X = \{x \mid f_j(x) = 0, \quad j = \overline{1,m}, \quad x \in R^n\}.$$

*Припущення 2.* Функції $f_j$, $j = \overline{0,m}$ – двічі неперервно диференційовані.

***Алгоритм 2***

П о ч а т о к. І. Вибрати початкове наближення $x^0$ та числа $\lambda_i^0$, $i = \overline{1,m}$.

II. Покласти $k = 0$.

О с н о в н и й ц и к л. III. Якщо виконуються рівності:

$$\nabla f_0(x^k) + \sum_{i=1}^{m} \lambda_i^k \nabla f_i(x^k) = 0;$$

$$f_i(x^k) = 0, \qquad i = \overline{1, m},$$

то покласти $x^* = x^k$ та припинити обчислення; інакше перейти на крок IV.

IV. Обчислити вектор похідних $\partial\varphi(x^k, \lambda^k)/\partial x$ функції Лагранжа за формулою

$$\partial\varphi(x^k, \lambda^k)/\partial x = \partial f_0(x^k)/\partial x + \sum_{i=1}^{m} \lambda_i^k \partial f_i(x^k)/\partial x.$$

V. Обчислити матрицю других похідних $\partial^2\varphi(x^k, \lambda^k)/\partial x^2$ функції Лагранжа за формулою

$$\partial^2\varphi(x^k, \lambda^k)/\partial x^2 = \partial^2 f_0(x^k)/\partial x^2 + \sum_{i=1}^{m} \lambda_i^k \partial^2 f_i(x^k)/\partial x^2 .$$

VI. Знайти вектор $h^k$ та числа $\delta_i^k$, $i = \overline{1, m}$, які задовольняють рівностям:

$$\frac{\partial^2\phi(x^k, \lambda^k)}{\partial x^2} h^k + \sum_{i=1}^{m} \delta_i^k \nabla f_i(x^k) + \frac{\partial\phi(x^k, \lambda^k)}{\partial x} = 0;$$

$$(\nabla f_i(x^k), h^k) + f_i(x^k) = 0, \qquad i = \overline{1, m}.$$



VII. Обчислити наступне наближення

$$x^{k+1} = x^k + h^k.$$

VIII. Обчислити

$$\lambda_i^{k+1} = \lambda_i^k + \delta_i^k, \quad i = \overline{1, m}.$$

IX. Покласти $k = k + 1$ та перейти на крок III.

***Теорема 2.*** *Нехай виконується припущення 2, а також наступні додаткові припущення:* $(\boldsymbol{i})$ *– другі похідні функцій* $f_j$, $j = \overline{0, m}$, *задовольняють умову Ліпшиця;* $(\boldsymbol{ii})$ *– в точці* $x^*$*, яка є розв'язком задачі 2, градієнти* $\nabla f_j(x^*)$, $j = \overline{1, m}$ *– лінійно незалежні, так що існують множники Лагранжа* $\lambda_i^*$, $i = \overline{1, m}$, *які задовольняють системі:*

$$\nabla f_0(x^*) + \sum_{j=1}^{m} \lambda_j^* \nabla f_j(x^*) = 0; \qquad f_j(x^*) = 0, \quad j = \overline{1, m};$$

$(\boldsymbol{iii})$ *– виконуються умови локального мінімуму, тобто*

$$\left( y, \frac{\partial^2\varphi(x^*, \lambda^*)}{\partial x^2} y \right) > 0, \text{ якщо } y \neq 0; \quad (\nabla f_j(x^*), y) = 0, \quad j = \overline{1, m},$$

(*тут* $\varphi(x, \lambda) = f_0(x) + \sum_{j=1}^{m} \lambda_j f_j(x)$).

*Тоді, якщо початкове наближення* $x^0$ *та числа* $\lambda_i^0$, $i = \overline{1, m}$, *достатньо* *близькі до* $x^*$ *та* $\lambda_i^*,\ i=\overline{1,m},$ *відповідно, то послідовність* $\{x^k;\ \lambda_i^k\ \ i=\overline{1,m}\}_{k=0}^{\infty},$ *яка породжена алгоритмом 2, збігається до* $(x^*;\ \lambda_i^*,\ i=\overline{1,m})$ *з квадратичною швидкістю.*

**3. Обмеження змішаного типу**

З а д а ч а 3. Знайти $\arg\min_{x\in X} f_0(x)$ для заданої функції $f_0 : R^n \to R^1$ і множини

$$X = \{x | f_j(x) \le 0, \quad j \in \Im, \quad x \in Y\},$$

де $Y$ – задана опукла замкнута множина;
$f_j : R^n \to R^1$ $(j \in \Im)$ – задані неперервні функції.

*Припущення 3.* $(\boldsymbol{i})$ – функції $f_j,\ j \in \Im \cup \{0\}$ – неперервно диференційовані.

***Алгоритм 3***

Всі кроки, за винятком I та VII, такі, як в алгоритмі 1. Кроки I та VII в алгоритмі 1 відповідно змінюються на I′ та VII′.

I′. Вибрати початкове наближення $x^0 \in Y$.

VII′. Знайти розв’язок $h = h(x)$ задачі квадратичного програмування:

$$\text{знайти } \arg\min_h \left( (\nabla f_0(x),\ h) + (1/2)\|h\|^2 \right)$$

при обмеженнях: $(\nabla f_j(x),\ h) + f_j(x) \le 0,\ \ j \in \Im_\delta(x),\ \ x, h \in Y.$

***Теорема 3.*** *Якщо виконані припущення 3 та існують такі константи* $\delta > 0$ *та* $\alpha > 0$, *що*: $(\boldsymbol{iii})$ – *множина*

$$X_\alpha = \{x \mid f_0(x) + \alpha \max_{j\in\Im} f_j(x) \le f_0(x^0) + \alpha \max_{j\in\Im} f_j(x^0),\ \ x \in Y\},$$

*обмежена;* $(\boldsymbol{iv})$ – *градієнти функцій* $f_j,\ j \in \Im \cup \{0\}$ *в* $X_\alpha$ *задовольняють умову Ліпшиця, тобто*

$$\left\|\nabla f_j(x) - \nabla f_j(\bar{x})\right\| \le \gamma \|x - \bar{x}\|,\ \ \gamma < \infty,\ \ \forall\, x,\ \bar{x} \in X_\alpha;$$

$(\boldsymbol{v})$ – *задача квадратичного програмування:*

$$\textit{знайти } \arg\min_h \left( (\nabla f_0(x),\ h) + (1/2)\|h\|^2 \right) \qquad (5.62)$$

*при обмеженнях:*

$$\left(\nabla f_j(x),\ h\right) + f_j(x) \le 0,\ \ j \in \Im_\delta(x),\ \ x, h \in Y, \qquad (5.63)$$

*розв’язується відносно* $h$ *при довільних* $x \in X_\alpha$ *та існують такі множники Лагранжа* $\lambda_j(x),\ \ j \in \Im_\delta(x)$, *що*

$$\sum_{j\in\Im_\delta(x)} \lambda_j(x) \le \alpha;$$

$(\boldsymbol{vi})$ – *серед функцій* $f_j,\ j\in\Im$, *є функція* $f_j$, *тотожно рівна нулю* $(f_j(x)\equiv 0)$, *тоді довільна гранична точка нескінченної послідовності* $\{x^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 3, належить множині* $X$ *і в цій точці виконується необхідна умова мінімуму функції* $f_0$ *на множині* $X$ *та, крім того, виконується*

$$\max_{j\in\Im} f_j(x^k)\to 0,\ \text{при}\ k\to\infty;\quad x^k\in Y,\quad k=0,1,\dots .$$

(Множники Лагранжа для задачі (6.62)–(6.63) – це такі невід'ємні числа, які задовольняють нерівності

$$\big(\nabla f_0(x),\, h(x)\big)+\|h(x)\|^2+\sum_{j\in\Im_\delta(x)}\lambda_j(x)\Big[\big(\nabla f_j(x),\, h(x)\big)+f_j(x)\Big]\le$$

$$\le\big(\nabla f_0(x),\, h\big)+\big(h(x),\, h\big)+\sum_{j\in\Im_\delta(x)}\lambda_j(x)\Big[\big(\nabla f_j(x),\, h\big)+f_j(x)\Big]$$

для всіх $h$, які задовольняють умову $x+h\in Y$ та рівностям:

$$\lambda_j(x)\Big[\big(\nabla f_j(x),\, h(x)\big)+f_j(x)\Big]=0,\quad j\in\Im_\delta(x)).$$

#### 4. Конструктивний метод лінеаризації

Задача 4. Знайти $\arg\min\limits_{x\in X} f_0(x)$ для заданої функції $f_0:R^n\to R^1$ і множини

$$X=\Big\{x\Big|\, f_j(x)\le 0,\quad j=\overline{1,m},\quad x\in R^n\Big\}\bigcap\Big\{x\Big|\, f_j(x)=0,\ \ j=\overline{m+1,m+l},\quad x\in R^n\Big\}$$

визначеної за допомогою заданих неперервно диференційованих функцій $f_j:R^n\to R^1$.

Наведений нижче алгоритм не потребує апріорного знання констант $\alpha$ і $\delta$, які забезпечують збіжність алгоритму. Ці константи обчислюються в процесі обчислень. Для реалізації даного алгоритму необхідна ефективна стандартна підпрограма розв'язування задачі квадратичного програмування.

***Алгоритм 4***

Початок. I. Вибрати довільне початкове наближення $x^0\in R^n$.

II. Вибрати константи $\delta_0>0,\ 0<\varepsilon<1,\ \alpha_0$ ($\alpha_0$ – досить велике число).

III. Покласти $k=0$.

Основний цикл. IV. Покласти $\delta=\delta_k,\ x=x^k,\ \alpha=\alpha_k$.

V. Обчислити

$$\psi(x)=\max\{0,\ f_1(x),\ \dots,\ f_m(x), |\, f_{m+1}(x)\,|,\ \dots, |\, f_{m+l}(x)\,|\}.$$

VI. Знайти множини індексів:

$$\Im_\delta^-(x)=\{\,j\mid f_j(x)\ge\psi(x)-\delta,\quad j=\overline{1,m}\};$$

$$\Im_\delta^0(x) = \{ j \mid f_j(x) \geq \psi(x) - \delta, \quad j = \overline{m+1, m+l} \}.$$

VII. Якщо задача квадратичного програмування:

$$\text{знайти } \arg\min_h \Big( \big(\nabla f_0(x),\, h\big) + (1/2)\|h\|^2 \Big)$$

при обмеженнях:

$$\big(\nabla f_j(x),\, h\big) + f_j(x) \leq 0, \quad j \in \Im_\delta^-(x),$$

$$\big(\nabla f_j(x),\, h\big) + f_j(x) = 0, \quad j \in \Im_\delta^0(x)$$

сумісна, то знайти її розв'язок $h = h^k$, множники Лагранжа $\lambda_j$, $j \in \Im_\delta^-(x) \bigcup \Im_\delta^0(x)$ та перейти на крок IX; інакше перейти на крок VIII.

VIII. Покласти $x^{k+1} = x$, $\alpha_{k+1} = \alpha$, $\delta_{k+1} = \delta / 2$ та перейти на крок XVI.

IX. Покласти $s = 0$.

X. Обчислити $\rho_k = (½)^s$.

XI. Якщо виконується нерівність

$$f_0(x + \rho_k h^k) + \alpha\psi(x + \rho_k h^k) \leq f_0(x) + \alpha\psi(x) - \varepsilon\rho_k \left\|h^k\right\|^2,$$

то перейти на крок XII; інакше покласти $s = s + 1$ та перейти на крок X.

XII. Обчислити наступне наближення

$$x^{k+1} = x + \rho_k h^k.$$

XIII. Покласти $\delta_{k+1} = \delta$.

XIV. Якщо виконується нерівність

$$\alpha > \sum_{j \in \Im_\delta^-(x)} \lambda_j + \sum_{j \in \Im_\delta^0(x)} \left|\lambda_j\right|,$$

то покласти $\alpha_{k+1} = \alpha$ та перейти на крок XVI; інакше перейти на крок XV.

XV. Обчислити

$$\alpha_{k+1} = 2\left( \sum_{j \in \Im_\delta^-(x)} \lambda_j + \sum_{j \in \Im_\delta^0(x)} \left|\lambda_j\right| \right).$$

XVI. Покласти $k = k + 1$ та перейти на крок IV.

Для алгоритму 4 справедливі теореми збіжності, аналогічні теоремам пункту 1. Відмітимо, що починаючи з деякого значення $k$, величини $\alpha_k$ і $\delta_k$ не будуть змінюватися.

### 5. Аналог методу лінеаризації в детермінованих задачах мінімізації майже диференційованих функцій

Задача 5. Знайти $\arg\min_{x \in X} f_0(x)$ для заданої функції $f_0 : R^n \to R^1$ і заданої множини $X \subset R^n$.

*Припущення 5.* $(\boldsymbol{i})$ – функція $f_0$ – майже диференційована; $(\boldsymbol{ii})$ – $X$ –

опукла, обмежена і замкнута множина, яка утворена нерівностями:

$$f_j(x) \le 0, \quad j = \overline{1,m},$$

де $f_j : R^n \to R^1$ – опуклі донизу функції.

Наведені нижче методи використовують загальну ідею методу лінеаризації та основані на використанні майже градієнта цільової функції.

***Алгоритм 5***

Початок. I. Вибрати: початкове наближення $x^0 \in X$, початкове значення $z^0 \in R^n$, початкове значення крокових множників $\rho_0, \alpha_0$ і величину зміщення $\delta_0$, які задовольняють умовам теореми 5.

II. Покласти $k = 0$.

Основний цикл. III. Обчислити $\overline{x}_i^k, i = \overline{1,n}$ – незалежні випадкові величини, рівномірно розподілені на відрізках:

$$\left[x_i^k - \delta_k/2,\, x_i^k + \delta_k/2\right], \qquad i = \overline{1,n}.$$

IV. Обчислити вектор:

$$\theta(x^k, k) = \frac{1}{\delta_k}\sum_{i=1}^{n}\Big[f_0\left(\overline{x}_1^k, \ldots, x_i^k + \delta_k/2, \ldots, \overline{x}_n^k\right) - \\ - f_0\left(\overline{x}_1^k, \ldots, x_i^k - \delta_k/2, \ldots, \overline{x}_n^k\right)\Big]e^i,$$

де $e^i$, $i = 1, \ldots, n$ $-i$-й орт.

V. Обчислити вектор $y^k \in X$, який задовольняє умову

$$(z^k, y^k) = \min_{x \in X}(z^k, x).$$

VI. Обчислити вектор

$$x^{k+1} = x^k - \rho_k(x^k - y^k).$$

VII. Обчислити вектор

$$z^{k+1} = z^k + \alpha_k\left(\theta(x^k, k) - z^k\right).$$

VIII. Знайти значення крокових множників $\rho_{k+1}, \alpha_{k+1}$ та значення зміщення $\delta_{k+1}$, які задовольняють умовам теореми 5.

IX. Покласти $k = k + 1$ та перейти на крок III.

***Теорема 5.*** *Нехай виконуються припущення 5 і, крім того, мають місце умови:*

$$0 \le \rho_k \le 1 \text{ } ma \text{ } \delta_k > 0 \text{ } npu \text{ } k = 0, 1, \ldots;$$

$$\sum_{k=0}^{\infty}\rho_k = \infty, \ \sum_{k=0}^{\infty}(\rho_k/\delta_k)^2 < \infty, \ \sum_{k=0}^{\infty}\alpha_k^2 < \infty;$$

$$\rho_k/\alpha_k\delta_k \to 0, \ \left|\delta_k - \delta_{k+1}\right|/\rho_k \to 0, \ \delta_k \to 0 \text{ } npu \text{ } k \to \infty.$$

*Тоді з ймовірністю 1 граничні точки послідовності* $\{x^k\}_{k=0}^{\infty}$, *яка утворена алгоритмом 5, належать множині* $X^*$ *розв'язків задачі 5 і* *послідовність* $\{f_0(x^k)\}_{k=0}^{\infty}$ *майже напевно збігається.*

*Зауваження 5.* Якщо існує конструктивний спосіб обчислення майже градієнта функції $f_0$ в кожній точці $x$, то на кроці IV алгоритму 5 потрібно обчислити $g(\bar{x}^k)-$ майже градієнт функції $f_0$ в точці $\bar{x}^k$, і на кроці VII вектор $z^{k+1}$ обчислювати за формулою

$$z^{k+1} = z^k + \alpha_k (g(\bar{x}^k) - z^k).$$

### 6. Аналог методу лінеаризації в стохастичних задачах мінімізації майже диференційованих функцій

Задача 6. Знайти $\arg\min\limits_{x\in X} E\,F_0(x,\omega)$ для заданої функції $F_0 : X\times\Omega \to R^1$ і заданої множини $X \subset R^n$ $(EF_0(x,\omega) \triangleq \int\limits_{\Omega} F_0(x,\omega)\mu(d\omega);$ міра $\mu$ може бути невідомою).

*Припущення 6.* $(\boldsymbol{i})$ – функція $F_0(x,\omega)$ в деякій області $X' \supset X$ задовольняє по $x$ умову Ліпшиця з константою $\gamma(\omega)$; $(\boldsymbol{ii}) - X -$ опукла, обмежена та замкнута множина, яка утворена нерівностями:

$$f_j(x) \le 0, \quad j = \overline{1,m},$$

де $f_j : R^n \to R^1$, $j = \overline{1,m}$ – опуклі донизу функції.

***Алгоритм 6***

Початок. I. Задати: початкові наближення $x^0 \in X$, $z^0 \in R^n$, початкові значення крокових множників $\rho_1(0), \rho_2(0)$ і величину зміщення $\delta_0$, які задовольняють умовам теореми 6.

II. Покласти $k = 0$.

Основний цикл. III. Обчислити $\bar{x}_i^k$, $i = \overline{1,n}-$ незалежні випадкові величини, рівномірно розподілені на відрізках:

$$\left[x_i^k - \delta_k/2,\ x_i^k + \delta_k/2\right], \qquad i = \overline{1,n}.$$

IV. Обчислити вектор:

$$\xi\left(x^k, k\right) = \frac{1}{\delta_k}\sum_{i=1}^{n}\Big[F_0\left(\bar{x}_1^k, \ldots, x_i^k + \delta_k/2, \ldots, \bar{x}_n^k, \omega^k\right) -$$
$$-F_0\left(\bar{x}_1^k, \ldots, x_i^k - \delta_k/2, \ldots, \bar{x}_n^k, \omega^k\right)\Big]e^i,$$

де $\omega^k -$ незалежні по $k$ спостереження випадкового параметра $\omega$, $e^i - i$-й орт.

V. Обчислити вектор $y^k \in X$, який задовольняє умову

$$(z^k, y^k) = \min_{x\in X}(z^k, x).$$

VI. Обчислити вектор

$$x^{k+1} = x^k - \rho_1(k)(x^k - y^k).$$

VII. Обчислити вектор

$$z^{k+1} = z^k + \rho_2(k)(\xi(x^k, k) - z^k).$$

VIII. Обчислити значення крокових множників $\rho_1(k+1), \rho_2(k+1)$ і зміщення $\delta_{k+1}$, які задовольняють умовам теореми 6.

IX. Покласти $k = k+1$ і перейти на крок III.

***Теорема 6.*** *Нехай виконані припущення 6 і, крім того, мають місце умови:*

$$E\gamma^2(\omega) < \infty;$$

$$0 \le \rho_1(k) \le 1,\ \delta_k > 0 \ \text{при}\ k = 0, 1, \ldots;$$

$$\sum_{k=0}^{\infty} \rho_1(k) = \infty,\ \sum_{k=0}^{\infty} \left(\rho_1(k)/\delta_k\right)^2 < \infty,\ \sum_{k=0}^{\infty} \rho_2^2(k) < \infty;$$

$$\rho_1(k)/\rho_2(k)\delta_k \to 0,\ \left|\delta_k - \delta_{k+1}\right|/\rho_1(k) \to 0 \ \text{та}\ \delta_k \to 0 \ \text{при}\ k \to \infty.$$

*Тоді з ймовірністю 1 граничні точки послідовності $\{x^k\}_{k=0}^{\infty}$, яка утворена алгоритмом 6, належать множині $X^*$ розв'язків задачі 6 і послідовність $\left\{f_0(x^k)\right\}_{k=0}^{\infty}$ майже напевно збігається.*



**7. Стохастичний метод лінеаризації**

Задача 7. Знайти $\arg\min_{x \in X} f_0(x),\ f_0 : R^n \to R^1,\ X \subset R^n$.

*Припущення 7.* (***i***) – функція $f_0$ – неперервно диференційована; (***ii***) – множина $X \subset R^n$ – обмежена і замкнута.

Якщо множина $X$ має не простий вигляд (наприклад, $X$ не є $n$-вимірним паралелепіпедом), то розв'язування задачі 7 з використанням операції проектування на область $X$ потребує ефективних способів розв'язування задачі мінімізації суми квадратів при наявності обмежень, що не завжди можливо.

В стохастичному методі лінеаризації звичайна операція проектування на множину $X$ замінюється операцією (іноді простішою, ніж проектування) мінімізації на $X$ лінійної функції $(z^k, x)$, яка визначена деяким допоміжним вектором $z^k$.

Позначимо через $Z$ опуклу, обмежену та замкнуту множину, для якої

$$Z \supseteq \left\{\nabla f_0(x) \mid x \in X^*\right\},$$

де $X^*$ – множина розв'язків задачі 7.

***Алгоритм 7***

Початок. I. Вибрати довільні початкові точки $x^0 \in X$ та $z^0 \in Z$.

II. Задати правила побудови послідовностей крокових множників

$$\{\rho_k\}_{k=0}^{\infty} \ \text{і}\ \{\delta_k\}_{k=0}^{\infty}.$$

III. Покласти $k = 0$.

О с н о в н и й   ц и к л.  IV. Обчислити вектор $\bar{x}^k$, який задовольняє умову

$$(z^k,\, \bar{x}^k) = \min_{x\in X}(z^k,\, x).$$

V. Обчислити вектор

$$x^{k+1} = x^k + \rho_k(\bar{x}^k - x^k).$$

VI. Знайти випадковий вектор $\xi^k$, умовне математичне сподівання якого

$$E(\xi^k /(x^0,\, z^0),\, \ldots,\, (x^k,\, z^k)) = \nabla f_0(x^k) + b^k,$$

де $b^k$ – вектор, вимірний відносно $\sigma$-алгебри, індукованої величинами $(x^0,\, z^0),\, \ldots,\, (x^k,\, z^k)$.

Способи побудови вектора $\xi^k$ наведені у прикладах підрозділу 6.19.

VII. Обчислити вектор

$$z^{k+1} = \pi_Z(z^k + \delta_k(\xi^k - z^k)).$$

VIII. Покласти $k = k+1$ та перейти на крок IV.

***Теорема 7.*** *Нехай мають місце припущення 7 і, крім того, $\nabla f_0(x)$ задовольняє локальну умову Ліпшиця. Тоді, якщо виконані умови:*

$$\|\xi^k\| + \|\nabla f_0(x^k)\| + \|b^k\| \le c_1,\; c_1 < \infty;$$

$$\rho_k \ge 0,\;\; \delta_k \ge 0,\;\; \rho_k / \delta_k \to 0 \;\; \textit{при} \;\; k \to \infty;$$

$$\sum_{k=0}^{\infty}\rho_k = \infty,\; \sum_{k=0}^{\infty}\delta_k = \infty,\; \sum_{k=0}^{\infty}\rho_k\|b^k\| < \infty \;\; \textit{майже напевно};$$

$$\sum_{k=0}^{\infty}E(\rho_k^2 + \delta_k^2) < \infty,$$

*то послідовність $\{x^k\}_{k=0}^{\infty}$, яка утворена алгоритмом 7, є такою, що $\left\{f_0(x^k)\right\}_{k=0}^{\infty}$ збігається майже напевно і кожна гранична точка $\{x^k\}_{k=0}^{\infty}$ належить множині*

$$X^* = \left\{x^* \mid \min_{x\in X}(\nabla f_0(x^*),\; x - x^*) = 0\right\}.$$

## 6.10. Методи відсікання

### 1. Лінійний випадок

З а д а ч а 1 . Знайти $\arg\min_{x\in X}(c, x)$ для заданого вектора $c \in R^n$ і множини

$$X \triangleq \left\{x \mid f_1(x) \le 0, \quad x \in R^n\right\},$$

визначеної за допомогою заданої опуклої функції $f_1 : R^n \to R^1$.

Зауважимо, що більш загальна задача опуклого програмування в просторі $R^n$: знайти $\arg\min_{x\in X} f_0(x)$, де

$$X \triangleq \left\{x \mid \hat{f}_j(x) \le 0, \quad j=\overline{1,m}\right\},$$

$\hat{f}_j$, $j=\overline{0,m}$ – опуклі функції, легко приводиться до вигляду задачі 1 за допомогою додаткової координати $x_{n+1}$ і нерівності

$$\overline{f}_{m+1}(x, x_{n+1}) \triangleq f_0(x) - x_{n+1} \le 0:$$

знайти $\arg\min_{(x,x_{n+1})\in\overline{X}} x_{n+1}$,

де

$$\overline{X} = \left\{(x, x_{n+1}) \mid \overline{f}(x, x_{n+1}) \le 0, \quad (x, x_{n+1}) \in R^{n+1}\right\};$$

$$\overline{f}(x, x_{n+1}) \triangleq \max_{i\in[1;m+1]} \overline{f}_i(x, x_{n+1}); \quad \overline{f}_i(x, x_{n+1}) \triangleq \hat{f}_i(x) + 0\cdot x_{n+1}, \quad i=\overline{1,m}.$$

*Припущення 1.* Множина $X$ – непорожня компактна множина. Сутність методу відсікаючих гіперплощини полягає в розв'язуванні задачі 1 шляхом розв'язування послідовності спеціальних задач лінійного програмування, отриманих апроксимацією множини обмежень $X$ послідовністю багатогранників $X_k$, $k=0,1,\ldots$, кожна з яких містить $X$.

***Алгоритм 1***

Початок. I. Вибрати ціле число $l>0$, $n$-вимірні вектори $a^i$, $i=-l,-(l-1),\ldots,-1,0$ і числа $\beta_i$, $i=\overline{-l,0}$, такі, що множина

$$X_0 \triangleq \left\{x \mid (a^i, x) - \beta_i \le 0, \quad i=\overline{-l,0}, \quad x\in R^n\right\}$$

компактна і містить $X$.

II. Покласти $k=0$.

Основний цикл. III. Знайти розв'язок $x=x^k$ задачі лінійного програмування:

$$\text{знайти } \arg\min_x (c, x)$$

при обмеженнях:

$$(a^i, x) - \beta_i \le 0, \quad i=\overline{-l,k}. \tag{6.64}$$

IV. Якщо $f_1(x^k) \le 0$, то покласти $x^* = x^k$ і зупинити обчислення; інакше перейти на крок V.

V. Обчислити опорний вектор $a^{k+1}$ до функції $f_1(x)$ у точці $x^k$.

VI. Обчислити

$$\beta_{k+1} = (a^{k+1}, x^k) - f_1(x^k).$$

VII. Побудувати множину

$$X_{k+1} \triangleq \left\{x \mid (a^{k+1}, x) - \beta_{k+1} \le 0, \quad x\in R^n\right\} \cap X_k.$$

VIII. Покласти $k=k+1$ і перейти на крок III.

*Зауваження 1.* На кроці III алгоритму 1 потрібно розв'язувати задачу лінійного програмування, число обмежень якої росте зі збільшенням $k$. Це вимагає великої оперативної пам'яті ЕОМ для збереження векторів $a^i$, $i=\overline{-l,k}$, і чисел $\beta_i$, $i=\overline{-l,k}$. Для спрощення алгоритму на кроці III замість задачі лінійного програмування (6.64) можна розв'язувати двоїсту до неї задачу:

$$\text{знайти } \arg\max_u \left(-\sum_{i=-l}^{k} u_i\beta_i\right)$$

при обмеженнях:

$$\sum_{i=-l}^{k} u_i\alpha^i + c = 0,\ \ u_i \geq 0,\quad i=\overline{-l,k}.$$

При розв'язуванні цих двоїстих задач симплекс-методом дуже важливо, що розв'язок попередньої задачі служить добрим допустимим розв'язком для наступної, так що їхній розв'язок отримується досить швидко.

***Теорема 1.*** *Якщо:* $(\boldsymbol{i})-$ $f_1(x)$ *– опукла неперервна функція;* $(\boldsymbol{ii})-$ *множина X – непорожня і компактна;* $(\boldsymbol{iii})-$ *існує таке число* $\delta<\infty$, *що при кожному* $\bar{x}\in X_0$ *вектор* $a$, *опорний до* $f_1(x)$ *в точці* $\bar{x}$, *задовольняє нерівності* $\|a\|\leq\delta$, *то будь-яка гранична точка* $x^*$ *нескінченної послідовності* $\left\{x^k\right\}_{k=0}^{\infty}$, *яка породжена алгоритмом 1, є розв'язком задачі 1 і* $f_1(x^k)\to 0$ *при* $k\to\infty$.

**2. Загальний випадок**

З а д а ч а  2. Знайти $\arg\min\limits_{x\in X} f_0(x)$ для заданої опуклої донизу функції $f_0: R^n \to R^1$ і множини $X$, що задається співвідношенням

$$X=\left\{x \mid f_1(x)\leq 0,\ x\in R^n\right\},$$

де $f_1: R^n \to R^1$ – неперервна опукла функція.

***Алгоритм 2***

П о ч а т о к . I. Вибрати опуклу замкнену множину $X_0$, яка містить $X$.

II. Покласти $k=0$.

О с н о в н и й  ц и к л . III. Розв'язати задачу опуклого програмування

$$x^k = \arg\min_{x\in X_k} f_0(x).$$

IV. Якщо $f_1(x^k)\leq 0$, то покласти $x^*=x^k$ і зупинити обчислення; інакше перейти на крок V.

V. Обчислити опорний вектор $\hat{\nabla} f_1(x^k)$ функції $f_1$ у точці $x^k$.

VI. Побудувати множину

$$X_{k+1} \triangleq X_k \cap \left\{ x \mid f_1(x^k) + (\hat{\nabla} f_1(x^k), x - x^k) \le 0, \quad x \in R^n \right\}.$$

VII. Покласти $k = k + 1$ і перейти на крок III.

Множини $X_k$ будують таким чином, щоб розв'язування цієї задачі було простішим, ніж розв'язування початкової задачі 2.

***Теорема 2.*** *Якщо:* $(\boldsymbol{i})-$ $f_0(x)$ *– рівномірно опукла донизу функція на множині* $X_0$; $(\boldsymbol{ii})-$ $X_0-$ *опукла і замкнена множина;* $(\boldsymbol{iii})-$ $f_1(x)-$ *неперервна опукла функція, що задовольняє умову Ліпшиця на* $X_0$, *то для нескінченної послідовності* $\left\{x^k\right\}_{k=0}^{\infty}$, *яка породжена алгоритмом 2, справедливо:*

1) $\lim\limits_{k\to\infty} f_0(x^k) \le f_0^* \triangleq \min\limits_{x\in X} f_0(x), \quad \overline{\lim\limits_{k\to\infty}} f_1(x^k) \le 0;$

2) *якщо, крім того,* $f_1(x)$ *– коректне обмеження і* $f_0(x)$ *задовольняє умову Ліпшиця на множині* $X_0$, *то*

$$\lim_{k\to\infty} x^k = x^*, \quad x^* \in X\ ; \quad f_0(x^*) = f_0^*; \quad \lim_{k\to\infty} f_0(x^k) = f_0^*;$$

3) *якщо, крім того,* $f_0(x)$ *– сильно опукла донизу функція й існує внутрішня точка* $\bar{x}$ *множини* $X_0$ *така, що* $f_1(\bar{x}) < 0$, *то мають місце наступні оцінки збіжності:*

$$f_0(x^k) - f_0^* \le \alpha_1 / k, \quad \alpha_1 < \infty;$$

$$x^k - x^* \le \alpha_2 \big/ \sqrt{k}, \quad \alpha_2 < \infty.$$

### 3. Метод відсікання з розтягненням простору для розв'язування задач опуклого програмування

З а д а ч а 3. Знайти $\arg\min\limits_{x\in X} f_0(x)$ для заданої функції $f_0 : R^n \to R^1$ і заданої множини

$$X \triangleq \left\{ x \mid f_i(x) \le 0, \ i = \overline{1,m}, \quad x \in R^n \right\}.$$

*Припущення 3.* $(\boldsymbol{i})-$ функції $f_i$, $i = \overline{0,m}$ – неперервні і опуклі донизу.

Наведений нижче метод відсікання з розтягненням простору гарантує зменшення об'єму області, у якій локалізується оптимум, зі швидкістю геометричної прогресії, причому знаменник прогресії залежить лише від розмірності задачі. Цей алгоритм відноситься до класу алгоритмів узагальненого градієнтного спуску з розтягненням простору в напрямку градієнта. Крім того, його можна розглядати як метод відсікання, у якому операція розтягнення простору використовується для симетризації області, у якій локалізований оптимум.

***Алгоритм 3***

Початок. I. Вибрати початкове наближення $x^0 \in R^n$ і число $\gamma > 0$, які задовольняють умовам теореми 3; покласти $B_0 = I$ ($I-$ одинична матриця), а також

$$\rho_0 = \gamma / (n+1), \; \beta = \sqrt{(n-1)/(n+1)}, \; k = 0.$$

Основний цикл. II. Обчислити вектор:

$$g(x^k) = \begin{cases} g_0(x^k), & \text{якщо} \quad \max\limits_{i \in [1;m]} f_i(x^k) \le 0; \\ g_{i^*}(x^k), & \text{якщо} \quad \max\limits_{i \in [1;m]} f_i(x^k) \triangleq f_{i^*}(x^k) > 0, \end{cases}$$

де $g_i(x^k)$, $i = \overline{0,m}$ – субградієнти (узагальнені градієнти) функцій $f_i$, $i = \overline{0,m}$, відповідно.

III. Якщо $g(x^k) = 0$, то покласти $x^* = x^k$ і зупинити обчислення; інакше перейти на крок IV.

IV. Обчислити новий напрямок $\xi^k$ для розтягнення простору:

$$\xi^k = B_k^T g(x^k) \Big/ \left\| B_k^T g(x^k) \right\|,$$

де $B_k^T -$ матриця, транспонована до матриці $B_k$.

V. Обчислити наступне наближення

$$x^{k+1} = x^k - \rho_k B_k \xi^k.$$

VI. Обчислити матрицю $B_{k+1}$, яка по суті є оберненою до результуючої матриці $A_{k+1}$ розтягнення простору після $(k+1)$-го кроку з коефіцієнтом розтягнення $\alpha = 1/\beta$,

$$B_{k+1} = B_k G_\beta(\xi^k),$$

де $G_\beta(\xi^k) -$ оператор розтягнення простору в напрямку $\xi^k$ із коефіцієнтом $\beta$ (визначення оператора розтягання простору наведене в п. 4.4).

VII. Обчислити значення крокового множника

$$\rho_{k+1} = \rho_k \frac{n}{\sqrt{n^2 - 1}}.$$

VIII. Покласти $k = k+1$ і перейти на крок II.

***Теорема 3.*** *Якщо виконане припущення 3 й існує оптимальна точка $x^*$ (не обов'язково єдина), що знаходиться в кулі радіусом $\gamma$ з центром у точці $x^0$, то послідовність $\left\{ x^k \right\}_{k=0}^{\infty}$, яка породжена алгоритмом 3, задовольняє нерівностям:*

$$\left\| A_k (x^k - x^*) \right\| \le \rho_k (n+1), \tag{6.65}$$

*де* $A_k \triangleq B_k^{-1}$, $k = 0, 1, \ldots$

*Зауваження 3.* Нерівність (6.65) дає корисну геометричну інтерпретацію методу, тому що множина точок $x$, які задовольняють нерівності

$$\left\| A_k (x^k - x) \right\| \le (n+1)\rho_k = \gamma \left( n \big/ \sqrt{n^2 - 1} \right)^k ,$$

являє собою еліпсоїд $\Phi_k$, об'єм якого

$$V(\Phi_k) = V_0 \gamma^k \left( n \big/ \sqrt{n^2 - 1} \right)^{nk} \Big/ \det A_k ,$$

де $V_0$ – об'єм одиничної $n$-вимірної кулі.

У зв'язку з цим об'єм еліпсоїда, у якому локалізується оптимальна точка $x^*$, зменшується зі швидкістю геометричної прогресії зі знаменником

$$q_n = \sqrt{\frac{n-1}{n+1}} \left( \frac{n}{\sqrt{n^2 - 1}} \right)^n < 1.$$

*Зауваження 3′.* До алгоритму 3 призводить також схема послідовних відсікань. Дійсно, якщо розглянути кулю $S_0 \triangleq \left\{ x \mid \left( \| x - x^0 \| \right) \le \gamma \right\}$ і провести гіперплощину $\left( g(x^0), x - x^0 \right) = 0$, то область локалізації $x^*$, очевидно, звужується до половини кулі $\overline{S} = S_0 \cap \Pi_0$, де $\Pi_0 \triangleq \left\{ x \middle| \left( g(x^0), x - x^0 \right) \le 0 \right\}$.

Якщо тепер описати навколо кулі $\overline{S}$ еліпсоїд мінімального об'єму, то його центр буде

$$x^0 - \frac{\gamma}{n+1} \frac{g(x^0)}{\left\| g(x^0) \right\|} = x^0 - \frac{\gamma}{n+1} \xi^0,$$

тобто буде співпадати з $x^1$.

Якщо тепер зробити розтягнення простору в напрямку $\xi^0 = g(x^0) \Big/ \left\| g(x^0) \right\|$, то в розтягнутому просторі образом зазначеного еліпсоїда буде куля радіусом $\gamma\, n \big/ \sqrt{n^2 - 1}$ з центром у точці $G_\alpha(\xi^0) x^1$ і т.д.

Нижче наводиться узагальнення алгоритму 3.

*Припущення 3′.* Нехай $g(x)$ – таке векторне поле, визначене на $R^n$, що

$$\left( g(x), x - x^* \right) \ge 0, \qquad \forall x \in R^n,$$

де $x^*$ – шукана оптимальна точка функції $f_0$ (причому $g(x) \ne 0$ при $x \ne x^*$).

Наведемо приклад визначення векторного поля $g(z)$. Нехай задана опукло-вгнута функція $f(x, y)$ двох векторних змінних $x \in R^n$, $y \in R^m$. Позначимо $z = (x, y) \in R^n \times R^m \triangleq R^{n+m}$. Сформуємо векторне поле $g(z)$

наступним чином:

$$g(z)=\left\{g_f^x(z),-g_f^y(z)\right\},\quad g_f^x(z)\in G_f^x(z),\ g_f^y(z)\in G_f^y(z),$$

де $G_f^x(x,y)$ – множина частинних субградієнтів (узагальнених градієнтів) функції $f(x,y)$, яка розглядається як функція від $x$ при фіксованому $y$; $G_f^x(x,y)$– множина частинних субградієнтів від функції $f(x,y)$ по $y$ при фіксованому $x$.

***Алгоритм 3′***

Початок. I. Вибрати початкове наближення $x^0\in R^n$, що лежить на відстані, яка не перевищує $\gamma$, від шуканої оптимальної точки $x^*$; покласти $B_0=I$; $\rho=\gamma/(n+1)$; $\beta=\sqrt{(n-1)/(n+1)}$; $k=0$.

Основний цикл. II. Обчислити вектор $g(x^k)$ – значення векторного поля, визначеного відповідно до припущення 3′, у точці $x^k$.

III. Якщо $g(x^k)=0$, то покласти $x^*=x^k$ і зупинити обчислення; інакше перейти на крок IV.

IV. Обчислити вектор $\xi^k$ напрямку розтягнення простору

$$\xi^k=B_k^T g(x^k)\Big/\left\|B_k^T g(x^k)\right\|.$$

V. Обчислити наступне наближення

$$x^{k+1}=x^k-\rho_k B_k\xi^k.$$

VI. Обчислити $n\times n$-матрицю $B_{k+1}$ (обернену до результуючої матриці розтягнення простору після $(k+1)$-ї ітерації з коефіцієнтом розтягнення $\alpha=1/\beta$):

$$B_{k+1}=B_k G_\beta(\xi^k),$$

де $G_\beta(\xi^k)$– оператор розтягнення простору в напрямку $\xi^k$ з коефіцієнтом $\beta$.

VII. Обчислити значення крокового множника

$$\rho_{k+1}=\rho_k\, n\Big/\sqrt{n^2-1}.$$

VIII. Покласти $k=k+1$ і перейти на крок II.

***Теорема 3′.*** *Нехай виконане припущення 3′ і точка $x^0$ лежить на відстані, що не перевищує $\gamma$, від оптимальної точки $x^*$.*

*Тоді при $n>1$ послідовність $\left\{x^k\right\}_{k=0}^{\infty}$, яка породжена алгоритмом 3′, задовольняє нерівностям:*

$$\left\|A_k(x^k-x^*)\right\|\le\rho_k(n+1),\quad A_k=B_k^{-1},\quad k=0,1,...$$

Теорема 3′ дає оцінки збіжності алгоритму 3′, які виражені в термінах зменшення об’єму області локалізації мінімуму.

Наведена нижче теорема 3″ характеризує швидкість збіжності алгоритму 3′ по функціоналу.

***Теорема 3″.*** *Якщо функція $f_0$ – опукла донизу в $R^n$ і в процесі застосування алгоритму 3′ виконуються наступні умови:*

$(\boldsymbol{i})-\ \left\|x^0-x^*\right\|\le\gamma;$

$(\boldsymbol{ii})-\ \left\|g(x^k)\right\|\le\sigma_k,\quad k=0,1,...,$

*тут $g(x^k)-$ узагальнений градієнт функції $f_0$ у точці $x^k$, то для послідовності $\left\{x^k\right\}_{k=0}^{\infty}$, яка породжена алгоритмом 3′, справедливі нерівності:*

$$\min_{r\in[0;k]}(f_0(x^r)-f_0(x^*))\le\gamma\sigma_k\sqrt{k(\alpha^2-1)}q_n^{k/n}\sqrt{1-\beta^{2k/n}},\quad k=1,2,...,$$

*де*

$$\sigma_k=\max_{r\in[0;k]}\left\|g(x^r)\right\|,\quad \alpha=1/\beta;\quad q_n=\sqrt{\frac{n-1}{n+1}}\left(\frac{n}{\sqrt{n^2-1}}\right)^n.$$

*Зауваження 3″.* Оскільки, для досить великих $n-\ q_n\approx1-1/(2n)$, то теорема 3″ гарантує швидкість збіжності «рекордів відхилення» функціоналів від оптимального значення, що відповідає швидкості збіжності геометричної прогресії зі знаменником $q_n\approx1-1/(2n^2)$,

## 6.11. Методи, які використовують функцію Лагранжа

### 1. Аналітичний метод

Спочатку розглянемо аналітичний метод розв’язування класичної задачі на умовний екстремум, який базується на необхідних і достатніх умовах локального екстремуму. Класична задача записується так:

$$f_0(x)\to\min$$

при обмеженнях:

$$x\in X\triangleq\left\{x\in R^n\,\middle|\,f_i(x)=0,\ i=\overline{1,m},\ m<n\right\}.$$

Основна ідея методу полягає в переході від задачі на умовний екстремум початкової функції $f_0(x)$ до задачі на безумовний екстремум спеціально побудованої функції Лагранжа:

$$\varphi(x,y)=f_0(x)+\sum_{i=1}^{m}y_if_i(x),\qquad(6.66)$$

де $y=(y_1,y_2,...,y_m)\in R^m-$ вектор множників Лагранжа.

***Необхідна умова локальної оптимальності.***

Нехай точка $x^*\in R^n$ є точкою локального екстремуму і функції

$f_j(x),\ j=\overline{0,m}-$ неперервно диференційовані в точці $x^*$. Тоді існує такий вектор $y^*=(y^*_1, y^*_2,..., y^*_m)$, компоненти якого не дорівнюють нулю одночасно, що

$$\nabla_x\varphi(x^*, y^*)=0 \text{ і } \nabla_y\varphi(x^*, y^*)=0 \tag{6.67}$$

(при цьому повинна виконуватися умова регулярності: градієнти $\nabla f_i(x^*),\ i=\overline{1,m}$ повинні бути лінійно незалежні). Будь-яка точка $x^*$, яка задовольняє умови (6.67) при деякому ненульовому $y^*$, називається ***стаціонарною точкою*** задачі (6.66).

Для вияснення характеру стаціонарних точок використовується достатня умова оптимальності, яка використовує матрицю других похідних $\nabla^2_{xx}\varphi(x, y)$ функції Лагранжа.

***Достатня умова локальної оптимальності.***

Нехай функції $f_i(x),\ i=\overline{0,m}$ – двічі неперервно диференційовані в точці $x^*$, причому при деякому $y^*\neq 0$ виконуються умови (6.67). Тоді, якщо

$$(\nabla^2_{xx}\varphi(x^*, y^*)u,u)>0 \tag{6.68}$$

при всіх $u\in R^n$ таких, що $\left(\nabla f_i(x),u\right)=0,\ i=\overline{1,m}$, то $x^*-$ точка локального мінімуму функції $f_0(x)$ на множині $X$.

(Відмітимо, що коли замість (6.68) виконується умова $(\nabla^2_{xx}\varphi(x^*, y^*)u,u)<0$, то $x^*-$ точка локального максимуму).

Сформулюємо аналітичний метод відшукання точок умовних локальних екстремумів класичної задачі .

***Алгоритм 1***

Початок. І. Скласти функцію Лагранжа $\varphi(x, y)$ за формулою (6.66).

ІІ. Знайти $\nabla_x\varphi(x, y)$.

ІІІ. Розв'язати систему алгебраїчних рівнянь з $n+m$ змінними $x_j,\ j=\overline{1,n}$; $y_i,\ i=\overline{1,m}$:

$$\begin{cases}\dfrac{\partial\varphi(x, y)}{\partial x_j}=0, \quad j=\overline{1,n};\\[2mm] \dfrac{\partial\varphi(x, y)}{\partial y_i}=f_i(x)=0,\ i=\overline{1,m}.\end{cases}$$

В результаті цього знаходяться стаціонарні точки $x^{(k)},\ k=\overline{1,M}$ і відповідні їм значення $y^{(k)},\ k=\overline{1,M}$.

IV. Покласти $k = 1$.

Основний цикл. V. Знайти $\nabla^2_{xx}\varphi(x^{(k)}, y^{(k)})$ і скласти квадратичну форму $G_k(u)$:

$$G_k(u) = (\nabla^2_{xx}\varphi(x^{(k)}, y^k)u, u).$$

VI. Обчислити вектори $u^{(r_k)}$, які задовольняють системі рівнянь

$$\left(\nabla f_i(x^{(k)}), u\right) = 0,\ \ i = \overline{1, m}.$$

VII. Обчислити значення квадратичних форм $G_k(u^{(r_k)})$ при всіх $u^{(r_k)}$ і проаналізувати їх знаки:

якщо $G_k(u^{(r_k)}) > 0$ для всіх ненульових $u^{(r_k)}$, то $x^{(k)}$ – точка умовного локального мінімуму;

якщо $G_k(u^{(r_k)}) < 0$ для всіх ненульових $u^{(r_k)}$, то $x^{(k)}$ – точка умовного локального максимуму.

VIII. Якщо $k < M$, покласти $k = k + 1$ і перейти на крок V; інакше зупинитись.



**Приклад 1.** Визначити точки локальних екстремумів функції

$$f_0(x) = 0{,}5\alpha x_1^2 + 0{,}5\beta x_2^2,\ \ \alpha > 0,\ \ \beta > 0$$

при обмеженнях

$$x_1^3 + x_2^3 = 1.$$

*Розв'язування.*

I. Складемо функцію Лагранжа

$$\varphi(x, y) = 0{,}5\alpha x_1^2 + 0{,}5\beta x_2^2 + y(x_1^3 + x_2^3 - 1).$$

II. Знаходимо $\nabla_x\varphi(x, y)$:

$$\frac{\partial\varphi(x, y)}{\partial x_1} = \alpha x_1 + 3yx_1^2,\ \ \frac{\partial\varphi(x, y)}{\partial x_2} = \beta x_2 + 3yx_2^2.$$

III. Розв'язуємо систему алгебраїчних рівнянь з трьома змінними $x_1,\ x_2,\ y$:

$$\begin{cases} \alpha x_1 + 3yx_1^2 = 0, \\ \beta x_2 + 3yx_2^2 = 0, \\ x_1^3 + x_2^3 - 1 = 0, \end{cases} \Leftrightarrow \begin{cases} x_1(\alpha + 3yx_1) = 0 \\ x_2(\beta + 3yx_2) = 0 \\ x_1^3 + x_2^3 - 1 = 0. \end{cases}$$

Можливі <u>*три*</u> варіанти :

а) $x_1 = 0, x_2 = 1,\ \beta + 3yx_2 = 0 \Leftrightarrow y = -\beta / 3.$

Знайшли першу стаціонарну точку $x^{(1)} = (0;1)$ і відповідний множник

Лагранжа $y^{(1)} = -\beta / 3$;

б) $x_2 = 0, x_1 = 1, \alpha + 3yx_1 = 0 \Leftrightarrow y = -\alpha / 3.$

Знайшли другу стаціонарну точку $x^{(2)} = (1;0)$ і відповідний множник Лагранжа $y^{(2)} = -\alpha / 3;$

в) при $x_1 \neq 0, x_2 \neq 0$ повинно виконуватися

$$\begin{cases} \alpha + 3yx_1 = 0, \\ \beta + 3yx_2 = 0, \end{cases} \Leftrightarrow \begin{cases} y = -\alpha / 3x_1 \\ y = -\beta / 3x_2, \end{cases} \Leftrightarrow x_1 = \alpha x_2 / \beta.$$

Підставимо це значення $x_1$ в третє рівняння системи

$$\frac{\alpha^3 x_2^3}{\beta^3} + x_2^3 - 1 = 0 \Leftrightarrow x_2^3 (\frac{\alpha^3 + \beta^3}{\beta^3}) = 1 \Leftrightarrow x_2 = \frac{\beta}{\sqrt[3]{\alpha^3 + \beta^3}};$$

тепер можемо знайти $x_1$ і $y$:

$$x_1 = \frac{\alpha}{\beta} \frac{\beta}{\sqrt[3]{\alpha^3 + \beta^3}} = \frac{\alpha}{\sqrt[3]{\alpha^3 + \beta^3}}, y = -\frac{\alpha \sqrt[3]{\alpha^3 + \beta^3}}{3\alpha} = -\frac{\sqrt[3]{\alpha^3 + \beta^3}}{3}.$$

Отже, отримали третю стаціонарну точку

$x^{(3)} = (\frac{\alpha}{\sqrt[3]{\alpha^3 + \beta^3}}, \frac{\beta}{\sqrt[3]{\alpha^3 + \beta^3}})$ і відповідний множник Лагранжа

$y^{(3)} = -\sqrt[3]{\alpha^3 + \beta^3} / 3$. При цьому $M$ =3.

IV. Покладемо $k = 1$.

*1-а ітерація:*

V. Знаходимо $\nabla_{xx}^2 \varphi(x^{(1)}, y^{(1)})$:

$$\nabla_{xx}^2 \varphi(x^{(1)}, y^{(1)}) = \begin{pmatrix} \alpha + 6y^{(1)} x_1^{(1)} & 0 \\ 0 & \beta + 6y^{(1)} x_2^{(1)} \end{pmatrix} =$$

$$= \begin{pmatrix} \alpha + 6(-\beta / 3) \cdot 0 & 0 \\ 0 & \beta + 6(-\beta / 3) \cdot 1 \end{pmatrix} = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}.$$

Складемо квадратичну форму

$$G_1(u) = \left( \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \right) = \left( \begin{pmatrix} \alpha u_1 \\ -\beta u_2 \end{pmatrix}, \begin{matrix} u_1 \\ u_2 \end{matrix} \right) = \alpha u_1^{\ 2} - \beta u_2^{\ 2}.$$

VI. Розв'язуємо систему рівнянь:

$$\left( \begin{pmatrix} 3x_1^2 \\ 3x_2^2 \end{pmatrix}, \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \right) = 0 \quad \text{при} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1^{\ (1)} \\ x_2^{\ (1)} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Отримуємо $3x_1^2u_1 + 3x_2^2u_2 = 0 \Leftrightarrow 3u_2^{(1)} = 0 \Leftrightarrow u_2^{(1)} = 0$,

$u_1^{(1)} = \delta \in R^1$, тобто $u^{(r_1)} = \begin{pmatrix} \delta \\ 0 \end{pmatrix}, \delta \in R^1$.

VII. Обчислюємо значення квадратичної форми $G_1(u^{(r_1)}) = \alpha \cdot \delta^2 - \beta \cdot 0 = \alpha \cdot \delta^2$. Оскільки $G_1(u^{(r_1)}) > 0$ при $\delta \neq 0$, то стаціонарна точка $x^{(1)}$ є точкою умовного локального мінімуму.

VIII . Оскільки $k < 3$, то покладемо $k = 1 + 1 = 2$ і переходимо на крок V.

*2-а ітерація:*

V. Знаходимо матрицю

$$\nabla_{xx}^2 \varphi(x^{(2)}, y^{(2)}) = \begin{pmatrix} \alpha + 6(-\alpha/3)\cdot 1 & 0 \\ 0 & \beta + 6(-\alpha/3)\cdot 0 \end{pmatrix} = \begin{pmatrix} -\alpha & 0 \\ 0 & \beta \end{pmatrix}.$$

Складаємо квадратичну форму

$$G_2(u) = \left( \begin{pmatrix} -\alpha & 0 \\ 0 & \beta \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \right) = \left( \begin{pmatrix} -\alpha u_1 \\ \beta u_2 \end{pmatrix}, \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \right) = -\alpha u_1^2 + \beta u_2^2.$$

VI. Розв'язуємо систему рівнянь

$$\left( \begin{pmatrix} 3x_1^2 \\ 3x_2^2 \end{pmatrix}, \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \right) = 0 \text{ при } \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1^{(2)} \\ x_2^{(2)} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Отримуємо $3x_1^2u_1 + 3x_2^2u_2 = 0 \Leftrightarrow 3u_1^{(2)} = 0 \Leftrightarrow u_1^{(2)} = 0$, $u_2^{(2)} = \delta \in R^1$, тобто $u^{(r_2)} = \begin{pmatrix} 0 \\ \delta \end{pmatrix}$, $\delta \in R^1$.

VII. Обчислюємо значення квадратичної форми

$$G_2(u^{(r_2)}) = -\alpha \cdot 0 + \beta \cdot \delta^2 = \beta \cdot \delta^2.$$

Оскільки $G_2(u^{(r_2)}) > 0$ при $\delta \neq 0$, то стаціонарна точка $x^{(2)}$ теж є точкою умовного локального мінімуму.

VIII . Оскільки $k < 3$, то покладемо $k = 2 + 1 = 3$ і переходимо на крок V.

*3-я ітерація*:

V. Знаходимо матрицю

$$\nabla_{xx}^2 \varphi(x^{(3)}, y^{(3)}) =$$

$$= \begin{pmatrix} \alpha + 6\left(-\frac{\sqrt[3]{\alpha^3 + \beta^3}}{3}\right)\frac{\alpha}{\sqrt[3]{\alpha^3 + \beta^3}} & 0 \\ 0 & \beta + 6\left(-\frac{\sqrt[3]{\alpha^3 + \beta^3}}{3}\right)\frac{\beta}{\sqrt[3]{\alpha^3 + \beta^3}} \end{pmatrix} = \begin{pmatrix} -\alpha & 0 \\ 0 & -\beta \end{pmatrix}.$$

Складаємо квадратичну форму

$$G_3(u) = \left(\begin{pmatrix} -\alpha & 0 \\ 0 & -\beta \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}\right) = \left(\begin{pmatrix} -\alpha u_1 \\ -\beta u_2 \end{pmatrix}, \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}\right) = -\alpha {u_1}^2 - \beta {u_2}^2.$$

VI. Розв’язуємо систему рівнянь

$$\left(\begin{pmatrix} 3x_1^2 \\ 3x_2^2 \end{pmatrix}, \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}\right) = 0 \text{ при } \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1^{(3)} \\ x_2^{(3)} \end{pmatrix} = \begin{pmatrix} \alpha \big/ \sqrt[3]{\alpha^3 + \beta^3} \\ \beta \big/ \sqrt[3]{\alpha^3 + \beta^3} \end{pmatrix}.$$

Отримуємо $\quad 3x_1^2 u_1 + 3x_2^2 u_2 = 0 \Leftrightarrow 3\dfrac{\alpha^2}{(\alpha^3+\beta^3)^{2/3}} {u_1}^{(3)} + 3\dfrac{\beta^2}{(\alpha^3+\beta^3)^{3/2}} {u_2}^{(3)} =$

$= \alpha^2 {u_1}^{(3)} + \beta^2 {u_2}^{(3)} = 0 \Leftrightarrow {u_1}^{(3)} = -\beta^2 {u_2}^{(3)} / \alpha^2$, тобто $u^{(r_3)} = \begin{pmatrix} -\beta^2 \delta / \alpha^2 \\ \delta \end{pmatrix}$, $\delta \in R^1$.

VII. Оскільки значення квадратичної форми $G_3(u^{(r_3)}) < 0$ при $\delta \neq 0$, то $x^{(3)}$ є точкою умовного локального максимуму.

### 2. Градієнтний метод для задач із обмеженнями типу нерівностей

З а д а ч а 2. Знайти $\arg\max\limits_{x \in X} f_0(x)$ для заданої функції $f_0 : R^n \to R^1$ та множини

$$X \triangleq \left\{ x \,\middle|\, f_i(x) \geq 0,\ \ i = \overline{1,m},\ \ x \in R^n \right\}.$$

*Припущення 2.* $(\boldsymbol{i})$ – функції $f_j(x)$, $j = \overline{0,m}$, – опуклі доверху та неперервно диференційовані в $R^n$ ($f_0(x)$ до того ж строго опукла); $(\boldsymbol{ii})$ – виконується умова Слейтера, тобто існує така точка $\bar{x}$, що $f_j(\bar{x}) > 0,\ j = \overline{1,m}$, $(\boldsymbol{iii})$ – існує така стаціонарна точка $x^*$ задачі 1, що

$$f_1(x^*) = 0,\ \ldots,\ f_p(x^*) = 0,\ \ f_{p+1}(x^*) > 0,\ \ldots,\ f_m(x^*) > 0;$$

$(\boldsymbol{iv})$ – функції $f_j(x)$, $j = \overline{0,m}$ мають в точці $x^*$ другу похідну.

Введемо функцію

$$\tilde{\varphi}(x, \tilde{y}) \triangleq f_0(x) + \sum_{i=1}^{p} y_i f_i(x), \quad \tilde{y} = (y_1, \ldots, y_p) \in R_+^p,$$

яка є обмеженням функції Лагранжа

$$\varphi(x, y) \triangleq f_0(x) + \sum_{i=1}^{m} y_i f_i(x), \ \ y = (y_1, \ldots, y_m) \in R_+^m, \tag{6.69}$$

на множині $R^n \times R_+^p$.

Оскільки $x^*$ – стаціонарна точка задачі 2, то існує такий вектор $\tilde{y}^* = (y_1^*, \ldots, y_p^*) \in R_+^p$, що $\nabla_x \tilde{\varphi}(x^*, \tilde{y}^*) = 0.$

Нижче приводиться метод множників Лагранжа з постійним кроковим множником, який при певних умовах збігається глобально до сідлової точки функції Лагранжа зі швидкістю геометричної прогресії.

***Алгоритм 2***

Початок. I. Вибрати будь-яке початкове наближення $(x^0, y^0) \in R^n \times R^m_+$.

II. Вибрати постійний кроковий множник $\rho > 0$.

III. Покласти $k = 0$.

IV. Визначити функцію Лагранжа $\varphi(x, y)$ по (6.69).

Основний цикл. V. Обчислити вектор $\nabla_x \varphi(x^k, y^k)$.

VI. Покласти $h^k = (f_1(x^k), ..., f_m(x^k))$.

VII. Обчистити наступні наближення $x^{k+1}$ та $y^{k+1}$ за формулами:

$$x^{k+1} = x^k + \rho \nabla_x \varphi(x^k, y^k);$$

$$y^{k+1} = \max\left\{0,\ y^k - \rho h^k\right\}.$$

VIII. Покласти $k = k + 1$ та перейти на крок V.

***Теорема 2.*** *Нехай виконуються припущення 2 і нехай:* $(\boldsymbol{v})-$ *задача 2 строго регулярна, тобто*

1) *не існує такого ненульового вектора* $u \in R^n$*, що*

$$-\nabla^2_{xx}\tilde{\varphi}(x^*, \tilde{y}^*)u = 0,\ \ \nabla^2_{\tilde{y}x}\tilde{\varphi}(x^*, \tilde{y}^*)u = 0;$$

2) *не існує такого ненульового вектора* $\upsilon \in R^p$*, що*

$$\nabla^2_{\tilde{y}x}\tilde{\varphi}(x^*, \tilde{y}^*)^T \upsilon = 0$$

*(тобто вектори* $\nabla f_1(x^*), ..., \nabla f_p(x^*)$ *– лінійно незалежні);*

3) *виконується умова строгої доповнюючої нежорсткості*

$$\tilde{y}^* \in \operatorname{Int} R^p_+,\ \text{тобто}\ \ y_i^* > 0,\ \ i = \overline{1, p};$$

$(\boldsymbol{vi})-$ *не існує числа* $\lambda \neq 0$ *та ненульового вектора* $u \in R^n$ *таких, що:*

$$-\nabla^2_{xx}\tilde{\varphi}(x^*, \tilde{y}^*)u = 0,$$

$$\left(\nabla^2_{\tilde{y}x}\tilde{\varphi}(x^*, \tilde{y}^*)\right)^T \nabla^2_{\tilde{y}x}\tilde{\varphi}(x^*, \tilde{y}^*)u = \lambda^2 u.$$

*Тоді для будь-якого* $\delta > 0$ *знайдеться число* $\rho(\delta) > 0$ *таке, що алгоритм 2 при будь-якому початковому наближенні* $(x^0, y^0)$*, що задовольняє умову* $\left\|(x^0, y^0) - (x^*, y^*)\right\| \leq \delta$ *та при будь-якому постійному кроковому множнику* $\rho \leq \rho(\delta)$*, буде породжувати послідовність* $\left\{(x^k, y^k)\right\}_{k=0}^{\infty}$*, яка лінійно (тобто зі швидкістю геометричної прогресії) буде збігатися до сідлової точки* $(x^*, y^*)$ *задачі 2.*

*Зауваження 2.* Умова $(\boldsymbol{vi})$ теореми 2 є не лише достатньою, але й, в деякому сенсі, необхідною умовою для збіжності алгоритму 2 в строго регулярній задачі.

**3. Градієнтний метод для задач з обмеженнями типу рівностей**

Задача 3. Знайти $\arg\max\limits_{x\in X} f_0(x)$ для заданої функції $f_0 : R^n \to R^1$ та множини

$$X \triangleq \left\{ x \middle| f_i(x) = 0,\ i = \overline{1,m},\ x \in R^n \right\}.$$

*Припущення 3.* $(\boldsymbol{i})-$ задача 3 має розв'язок $x^*$; $(\boldsymbol{ii})-$ функції $f_j(x)$, $j = \overline{0,m}$ – неперервно диференційовані в околі точки $x^*$; $(\boldsymbol{iii})-$ функції $f_j(x)$, $j = \overline{0,m}$ – двічі диференційовані в околі точки $x^*$; $(\boldsymbol{iv})-$ вектори $\nabla f_i(x^*)$, $i = \overline{0,m}$ – лінійно незалежні.

Припущення 3 гарантують існування та єдиність множників Лагранжа $y^* = (y_1^*, ..., y_m^*)$ задачі 3, тобто:

$$\nabla_x \varphi(x^*, y^*) = 0,\ \nabla_y \varphi(x^*, y^*) = 0,$$

де $\varphi(x, y) \triangleq f_0(x) + \sum\limits_{i=1}^{m} y_i f_i(x)$, $y = (y_1, ..., y_m) \in R^m$ – функція Лагранжа задачі 3.

Введемо позначення $A \triangleq \nabla_{xx}^2 \varphi(x^*, y^*)$, $C \triangleq \nabla_{yx}^2 \varphi(x^*, y^*)$, $C^T$ – матриця транспонована до $C$.

Наведений нижче алгоритм породжує послідовність $\left\{(x^k, y^k)\right\}_{k=0}^{\infty}$, яка локально (тобто при початковому наближенні $(x^0, y^0)$ з околу точки $(x^*, y^*)$) збігається зі швидкістю геометричної прогресії до точки $(x^*, y^*)$. На $k$-й ітерації алгоритму рух до наступного наближення $x^{k+1}$ здійснюється в напрямку антиградієнту по $x$ функції Лагранжа, а рух до $y^{k+1}$ – в напрямку градієнта по $y$ цієї ж функції.

***Алгоритм 3***

Початок. I. Вибрати будь-яке початкове наближення $(x^0, y^0)$ з околу точки $(x^*, y^*)$.

II. Вибрати постійний кроковий множник $\rho > 0$.

III. Визначити функцію Лагранжа $\varphi(x, y)$ для задачі 3, тобто покласти

$$\varphi(x, y) \triangleq f_0(x) + \sum_{i=1}^{m} y_i f_i(x),\ \ y = (y_1, ..., y_m) \in R^m. \tag{6.70}$$

IV. Покласти $k = 0$.

Основний цикл. V. Обчислити вектори:

$$\nabla_x \varphi(x^k, y^k) = \nabla f_0(x^k) + \sum_{i=1}^{m} y_i^k \nabla f_i(x^k),$$

$$\nabla_y \varphi(x^k, y^k) = \left(f_1(x^k), \ldots, f_m(x^k)\right).$$

VI. Обчислити наступні наближення $x^{k+1}$ та $y^{k+1}$ за формулами:

$$x^{k+1} = x^k - \rho \nabla_x \varphi(x^k, y^k);$$

$$y^{k+1} = y^k + \rho \nabla_y \varphi(x^k, y^k).$$

VII. Покласти $k = k+1$ і перейти на крок V.

***Теорема 3.*** *Нехай виконуються припущення 3 та нехай:* $(\boldsymbol{v})-$ *матриця* $AC^T$ *має ранг m, тобто із* $AC^T u = 0$ *випливає* $u = 0$; $(\boldsymbol{vi})-$ *матриця A – невід'ємно визначена, причому* $Ax \neq 0$ *при* $Cx = 0,\ x \neq 0.$

*Тоді існує таке число* $\bar{\rho} > 0$, *що при* $\rho < \bar{\rho}$ *алгоритм 3 локально збігається до* $(x^*, y^*)$ *зі швидкістю геометричної прогресії (тобто існують такі числа* $\varepsilon > 0,\ 0 \le q_1 < 1,$ *що для будь-яких* $\left\|x^0 - x^*\right\| \le \varepsilon$, $\left\|y^0 - y^*\right\| \le \varepsilon$ *будуть виконуватись наступні нерівності*

$$\left\|x^k - x^*\right\| \le \beta_1(\varepsilon)(q_1)^k; \quad \left\|y^k - y^*\right\| \le \beta_1(\varepsilon)(q_1)^k).$$

### 4. Метод квадратичної апроксимації для задач з обмеженнями типу рівностей



В методі квадратичної апроксимації на $k$-й ітерації функція Лагранжа $\varphi(x, y)$ квадратично апроксимується по $x$, в околі точки $x^k$ і наступне наближення $x^{k+1}$ знаходять з умов мінімуму цієї квадратичної апроксимації. Напрямок спуску по змінній $y$ здійснюється аналогічно як і в градієнтному методі.

***Алгоритм 4***

Початок. I. Вибрати будь-яке початкове наближення $(x^0, y^0)$ з околу точки $(x^*, y^*)$.

II. Вибрати постійний кроковий множник $\rho > 0$.

III. Визначити функцію Лагранжа $\varphi(x, y)$ для задачі 3 за формулою (6.70)

IV. Покласти $k = 0$.

Основний цикл. V. Обчислити вектори:

$$\nabla_y \varphi(x^k, y^k) = \left(f_1(x^k), \ldots, f_m(x^k)\right);$$

$$\nabla_x \varphi(x^k, y^k) = \nabla f_0(x^k) + \sum_{i=1}^{m} y_i^k \nabla f_i(x^k).$$

VI. Обчислити матрицю

$$\nabla_{xx}^2 \varphi(x^k, y^k) = \nabla_{xx}^2 f_0(x^k) + \sum_{i=1}^{m} y_i^k \nabla_{xx}^2 f_i(x^k).$$

VII. Обчислити наступні наближення $x^{k+1}$ та $y^{k+1}$ із системи:

$$\nabla^2_{xx}\varphi(x^k, y^k)(x^{k+1} - x^k) = -\rho\nabla_x\varphi(x^k, y^k);\ y^{k+1} = y^k + \rho\nabla_y\varphi(x^k, y^k).$$

VII. Покласти $k = k + 1$ та перейти на крок V.

***Теорема 4.*** *Нехай виконуються припущення 3. Нехай:* $(\boldsymbol{vii})-$ *матриця* $A \triangleq \nabla^2_{xx}\varphi(x^*, y^*)$ *– додатньо визначена, тобто* $(Ax, x) > 0$ *при* $x \neq 0$.

*Тоді існує таке число* $\bar{\rho} > 0$, *що при* $\rho < \bar{\rho}$ *алгоритм 4 локально збігається до* $(x^*, y^*)$ *зі швидкістю геометричної прогресії (тобто існують числа* $\varepsilon > 0,\ 0 \le q_2 < 1$ *такі, що для будь-яких* $\left\|x^0 - x^*\right\| \le \varepsilon$, $\left\|y^0 - y^*\right\| \le \varepsilon$ *будуть виконуватись наступні нерівності:*

$$\left\|x^k - x^*\right\| \le \beta_2(\varepsilon)(q_2)^k;\quad \left\|y^k - y^*\right\| \le \beta_2(\varepsilon)(q_2)^k).$$

#### 5. Двоїстий метод для задач з обмеженнями типу рівностей

В двоїстому методі на $k$-й ітерації обчислюють $(k+1)$-ше наближення $x^{k+1}$ шляхом розв'язування задачі безумовної оптимізації по $x$ функції Лагранжа $\varphi(x, y^k)$ (припускається, що задача безумовної оптимізації розв'язується).

Напрям спуску по змінній $y$ здійснюється аналогічно як і в градієнтному методі.

***Алгоритм 5***

Початок. I. Вибрати будь-яке початкове наближення $(x^0, y^0)$ з околу точки $(x^*, y^*)$.

II. Вибрати постійний кроковий множник $\rho > 0$.

III. Визначити функцію Лагранжа $\varphi(x, y)$ за формулою (6.70).

IV. Покласти $k = 0$.

Основний цикл. V. Обчислити вектор

$$\nabla_y\varphi(x^k, y^k) = \left(f_1(x^k), \ldots, f_m(x^k)\right).$$

VI. Обчислити наступні наближення $x^{k+1}$ та $y^{k+1}$ за формулами:

$$\varphi(x^{k+1}, y^k) = \min_{x \in R^n}\varphi(x, y^k);$$

$$y^{k+1} = y^k + \rho\nabla_y\varphi(x^k, y^k).$$

VII. Покласти $k = k + 1$ та перейти на крок V.

Для алгоритму 5 має місце теорема, аналогічна теоремі 4.

#### 6. Метод Ньютона для задач з обмеженнями типу рівностей

*Припущення 6.* Виконуються всі умови припущення 3, і до того ж $(\boldsymbol{ix})-$ функції $f_j(x)$, $j = \overline{0, m}$ – двічі диференційовані в околі точки $x^*$, причому другі похідні задовольняють умову Ліпшиця.

***Алгоритм 6***

П о ч а т о к. I. Вибрати будь-яке початкове наближення $(x^0, y^0)$ з околу точки $(x^*, y^*)$.

II. Визначити функцію Лагранжа $\varphi(x, y)$ за формулою (6.70).

III. Покласти $k = 0$.

О с н о в н и й  ц и к л. IV. Обчислити вектори:

$$\nabla_x \varphi(x^k, y^k) = \nabla f_0(x^k) + \sum_{i=1}^{m} y_i^k \nabla f_i(x^k);$$

$$\nabla_y \varphi(x^k, y^k) = \left(f_1(x^k), \ldots, f_m(x^k)\right).$$

V. Обчислити $n \times n$-матрицю $\nabla_{xx}^2 \varphi(x^k, y^k)$ та $n \times m$-матрицю $\nabla_{xy}^2 \varphi(x^k, y^k)$ за формулами:

$$\nabla_{xx}^2 \varphi(x^k, y^k) = \nabla_{xx}^2 f_0(x^k) + \sum_{i=1}^{m} y_i^k \nabla_{xx}^2 f_i(x^k);$$

$$\nabla_{xy}^2 \varphi(x^k, y^k) = \left(\nabla f_1(x^k), \ldots, \nabla f_m(x^k)\right).$$

VI. Обчислити розв'язок $(x^{k+1}, y^{k+1})$ із системи лінійних рівнянь:

$$\nabla_{xx}^2 \varphi(x^k, y^k)(x^{k+1} - x^k) + \nabla_{xy}^2 \varphi(x^k, y^k)(y^{k+1} - y^k) = -\nabla_x \varphi(x^k, y^k);$$

$$\nabla_{xy}^2 \varphi(x^k, y^k)^T (x^{k+1} - x^k) = -\nabla_y \varphi(x^k, y^k).$$

VII. Покласти $k = k + 1$ та перейти на крок IV.

***Теорема 6.*** *Нехай виконуються припущення 6 та нехай:* $(\boldsymbol{x})$ – *матриця* $A \triangleq \nabla_{xx}^2 \varphi(x^*, y^*)$ – *задовольняє умову* $(Ax, x) \neq 0$ *при* $\nabla_{yx}^2 \varphi(x^*, y^*) x = 0$, $x \neq 0$. *Тоді алгоритм 6 локально збігається до* $(x^*, y^*)$ *з квадратичною швидкістю, тобто*

$$\left\|x^k - x^*\right\| \leq \beta_3(\varepsilon)(q_3)^{2^k}, \ \left\|y^k - y^*\right\| \leq \beta_3(\varepsilon)(q_3)^{2^k}, \ q_3 < 1,$$

*при*

$$\left\|x^0 - x^*\right\| \leq \varepsilon, \ \left\|y^0 - y^*\right\| \leq \varepsilon.$$

*Зауваження 6.* Найбільш простим для реалізації є градієнтний метод, але збіжність в алгоритмі 3 не є монотонною, тому визначити на практиці, чи має місце збіжність при зафіксованому $\rho$, зробити важко. Як правило, алгоритми квадратичної апроксимації та двоїстий збігаються швидше ніж градієнтний метод, але вони більш трудомісткі з обчислювальної точки зору. Метод Ньютона за складністю обчислень можна порівняти з методами пункту 4 та 5, але він збігається з квадратичною швидкістю та не потребує знаходити кроковий множник $\rho$, який забезпечує збіжність.

*Зауваження 6′.* Для початку роботи алгоритмів, які наведені в пунктах 3–6, необхідно мати початкове наближення в околі точки $(x^*, y^*)$. Для отримання вдалого початкового наближення пропонується використовувати метод штрафних функцій. Спочатку необхідно знайти добре наближення $x^k$ для $x^*$, мінімізуючи функцію:

$$f_0(x)+\sum_{i=1}^{m}\frac{1}{2}\alpha_k(f_i(x))^2$$

по $x$ в просторі $R^n$ при достатньо великих значеннях $\alpha_k$. Тоді вектор $\left(\alpha_k f_1(x^k),...,\alpha_k f_m(x^k)\right)$ буде вдалим наближенням для $y^*$.

## 6.12. Методи, які використовують модифіковані функції Лагранжа

**1. Градієнтний метод**

З а д а ч а 1. Знайти $\arg\max\limits_{x\in X} f_0(x)$ для заданої функції $f_0:R^n\to R^1$ та множини

$$X\triangleq\left\{x\middle|f_j(x)\geq 0,\ \ j=\overline{1,m}\right\}.$$

*Припущення 1.* $(\boldsymbol{i})-$ функції $f_j(x),\ \ j=\overline{0,m}$ – диференційовані, опуклі догори та їх похідні задовольняють умову Ліпшиця на будь-якому компакті; $(\boldsymbol{ii})-$ множина сідлових точок $Z^*=\left\{(x^*,u^*)\right\}$ функції Лагранжа

$$\varphi(x,u)\triangleq f_0(x)+\sum_{j=1}^{m}f_j(x)u_j \tag{6.72}$$



не порожня.

Наведений метод базується на побудові модифікованого лагранжіана $\psi(x,u)$ задачі 1, тобто функції:

$$\psi(x,u)\triangleq f_0(x)-\frac{1}{2\gamma}\sum_{j=1}^{m}\left(\left([u_j-\gamma f_j(x)]_+\right)^2-u_j^2\right),\ \ \gamma>0, \tag{6.73}$$

де $[\alpha]_+=(\alpha+|\alpha|)/2$, множина сідлових точок якого співпадає з множиною сідлових точок класичного лагранжіана $\varphi(x,u)$.

***Алгоритм 1***

П о ч а т о к . І. Вибрати початкове наближення $(x^0,u^0)\in R^n\times R^m$.

ІІ. Вибрати константу $\gamma>0$.

ІІІ. Покласти $k=0$.

IV. Знайти додатне число $\rho_0$, яке задовольняє нерівність:

$$\rho_0\leq\min\left\{\gamma,2^{-1}\left\|\nabla\psi(x^0,u^0)\right\|^{-2}\right\},$$

де $\nabla\psi(x^0,u^0)=\left(\nabla_x\psi(x^0,u^0)\nabla_u\psi(x^0,u^0)\right)$, а функція $\psi: R^n\times R^m\to R^1$ визначається за формулою (6.73).

Основний цикл. V. Знайти вектор

$$x^{k+1}=x^k+\rho_k\nabla_x\psi(x^k,u^k).$$

VI. Обчислити вектор

$$u^{k+1}=u^k-\rho_k\nabla_u\psi(x^k,u^k).$$

V. Обчислити градієнт $\nabla\psi(x^{k+1},u^{k+1})$ функції $\psi$ в точці $\left(x^{k+1},u^{k+1}\right)$.

VI. Обчислити значення крокового множника

$$\rho_{k+1}=\min\{\rho_k(1-\rho_k\,\|\nabla\psi(x^k,u^k)\|^2),\ 2^{-1}\|\nabla\psi(x^{k+1},u^{k+1})\|^{-2}\}$$

(випадок збіжності алгоритму за скінчене число ітерацій, коли $\|\nabla\psi(x^{k+1},u^{k+1})\|=0$ тут не розглядається).

IX. Покласти $k=k+1$ та перейти на крок V.

***Теорема 1.*** *Якщо виконані припущення 1, то для алгоритму 1 справедливі наступні твердження:* 1) – *існує такий номер* $k_0$, *що*

$$\rho_{k+1}=\rho_k(1-\rho_k\left\|\nabla\psi(x^k,u^k)\right\|^2),\ k>k_0;$$

2) – $\lim\limits_{k\to\infty}\rho_k>0;$ 3) – *послідовність* $\left\{x^k,u^k\right\}_{k=0}^{\infty}$ *збігається до деякої точки* $(x^*,u^*)$ *з множини сідлових точок функції Лагранжа (6.72).*

***Теорема 1'.*** *Якщо виконані припущення 1 та:* (***iii***) – $z^*=(x^*,u^*)$ – *деяка сідлова точка функції Лагранжа;* (***iv***) – *функції* $f_j(x),\ j=\overline{0,m}$ *в точці* $x^*$ *тричі диференційовані;* (***v***) – *в точці* $(x^*,u^*)$ *виконуються достатні умови максимуму другого порядку та умови строгої регулярності, то послідовність* $\left\{\left(x^k,u^k\right)\right\}_{k=0}^{\infty}$, *яка породжена алгоритмом 1, збігається до точки* $(x^*,u^*)$ *зі швидкістю геометричної прогресії.*

Умови (***v***) означають наступне:

1) матриця $\nabla^2_{ux}\varphi^0(x^*,u^*)$ має ранг *l*, де $\varphi^0(x,u)\triangleq f_0(x)+\sum\limits_{j=1}^{l}f_j(x)u_j$, а число *l* таке, що обмеження $f_j(x)\ge 0,\ \ j=\overline{1,l}$ – активні в точці $x^*$, а при $j=\overline{l+1,m}$ – пасивні;

2) $u_j^*>0$ при $j=\overline{1,l}$;

3) не існує ненульового вектора *h*, який би задовольняв умовам:

$\nabla^2_{xx}\varphi^0(x^*,u^*)h=0,\ \nabla^2_{ux}\varphi^0(x^*,u^*)h=0$ (відмітимо, що при таких умовах $(x^*,u^*)$ – єдина сідлова точка функції $\varphi$ на $R^n\times R^m$).

**2. Метод, який використовує штрафні функції експоненціального типу**

З а д а ч а 2. Знайти $\arg\min\limits_{x\in X} f_0(x)$, де

$$X\triangleq\left\{x\,\middle|\,f_i(x)\le 0,\ i=\overline{1,m},\ x\in R^n\right\};$$

$f_i:R^n\to R^1,\ i=\overline{0,m}$ – неперервні функції.

*Припущення 2.* $(\boldsymbol{i})$ – функції $f_i(x),\ i=\overline{0,m}$ – неперервно диференційовані в $R^n$.

В методі штрафних функцій експоненціального типу будується послідовність $\left\{(x^k,y^k)\right\}$, яка збігається при відповідних припущеннях лінійно до точки $(x^*,y^*)$, де $x^*$ – розв'язок задачі 2, а $u^*=((y_1^*)^2,...,(y_m^*)^2)$ – множники Лагранжа задачі 2.

***Алгоритм 2***

П о ч а т о к. I. Вибрати початкове наближення $x^0\in R^n$, числа $y_j^0>0,\ \ j=\overline{1,m}$ та коефіцієнт штрафу $\alpha>0$ такі, щоб:

$$\nabla_x\psi(x^0,y^0,\alpha)=0\qquad\left(y^0=(y_1^0,...,y_m^0)\right),$$

де експоненціальна штрафна функція $\psi:R^n\times R_+^m\times R_+^1\to R^1$ визначається за формулою

$$\psi(x,y,\alpha)\triangleq f_0(x)+(1/\alpha)\sum_{j=1}^m(y_j)^2\left[\exp\left(\alpha f_j(x)\right)-1\right];\ \left(y=(y_1,...,y_m)\right).$$

II. Покласти $k=0$.

О с н о в н и й ц и к л. III. Обчислити вектор $x^{k+1}$ так, щоб

$$\nabla_x\psi(x^{k+1},y^k,\alpha)=0\,. \qquad (6.74)$$

Якщо точка $x^{k+1}$, яка задовольняє (6.74), не єдина, то вибрати з них точку $x^{k+1}$, яка є найближчою до точки $x^k$ за деякою нормою в $R^n$.

IV. Обчислити вектор $y^{k+1}=(y_1^{k+1},...,y_m^{k+1})$ за формулою:

$$y_j^{k+1}=y_j^k\exp\left(\alpha f_j(x^{k+1})/2\right),\quad j=\overline{1,m}\,.$$

V. Покласти $k=k+1$ та перейти на крок III.

***Теорема 2.*** *Нехай виконується припущення 2 та:* $(\boldsymbol{ii})$ – *функції* $f_i(x),\ i=\overline{0,m}$ – *опуклі донизу;* $(\boldsymbol{iii})$ – *для будь-яких* $\Im_1,\ \Im_2\subset\{1,...,m\},\ \Im_1\ne\Im_2$, *підзадачі:*

*знайти:* $\arg\min_{x\in X_1} f_0(x)$, $X_1 \triangleq \left\{x \,\middle|\, f_j(x)\le 0,\ j\in\Im_1,\ x\in R^n\right\}$;

*і знайти:* $\arg\min_{x\in X_2} f_0(x)$, $X_2 \triangleq \left\{x \,\middle|\, f_j(x)\le 0,\ j\in\Im_2,\ x\in R^n\right\}$,

*мають, відповідно, оптимальні розв'язки* $x^{1,*}$ *та* $x^{2,*}$, *причому* $f_0(x^{1,*}) \ne f_0(x^{2,*})$.

*Тоді будь-яка гранична точка* $(x^*, y^*)$ *послідовності* $\left\{(x^k, y^k)\right\}_{k=0}^{\infty}$, *яка породжена алгоритмом 2, є парою, що складається з оптимального розв'язку* $x^*$ *задачі 2 та множників Куна-Таккера* $(y^*)^2$ *для задачі 2.*

*Зауваження 2.* Теорема 2 залишається справедливою і в тому випадку, якщо умову $(\boldsymbol{iii})$ замінити однією з наступних умов: $(\boldsymbol{iii'})$ – послідовність $\left\{(x^k, y^k)\right\}_{k=0}^{\infty}$ збігається до точки $(x^*, y^*)$; $(\boldsymbol{iii''})$ – гранична точка $x^*$ послідовності $\left\{x^k\right\}_{k=0}^{\infty}$ є допустимою точкою для задачі 2, тобто:

$$f_j(x^*)\le 0,\quad j=\overline{1,m}.$$



***Теорема 2'.*** *Якщо припущення 2 виконується, то для послідовності* $\left\{(x^k, y^k)\right\}_{k=0}^{\infty}$, *породженої алгоритмом 2, справедливі наступні твердження:*

*1) для всіх* $i\in K$ *виконується нерівність:*

$$\liminf_{k\to\infty} f_i(x^k)\le 0,$$

*де* $K \triangleq \left\{ i \,\middle|\, \liminf_{k\to\infty} y_i^k = 0,\ i=\overline{1,m}\right\}$;

2) *якщо послідовність* $\left\{y^k\right\}_{k=0}^{\infty}$ *збігається до точки* $\overline{y}$, *то будь-яка гранична точка* $(\overline{x}, \overline{y})$ *послідовності* $\left\{(x^k, y^k)\right\}_{k=0}^{\infty}$ *така, що*

$$\nabla_x \varphi(\overline{x},\overline{u})=0,\quad \overline{y}_j f_j(\overline{x})=0,\quad j=\overline{1,m},$$

*де*

$$\overline{u}=\left((\overline{y}_1)^2,\ldots,(\overline{y}_m)^2\right),\quad \varphi(x,u)\triangleq f_0(x)+\sum_{j=1}^{m}u_j f_j(x),\quad u\in R_+^m;$$

3) *якщо послідовність* $\left\{x^k\right\}_{k=0}^{\infty}$ *збігається до точки* $\overline{x}$, *то будь-яка гранична точка* $(\overline{x}, \overline{y})$ *послідовності* $\left\{(x^k, y^k)\right\}_{k=0}^{\infty}$ *така, що*

$$\nabla_x \varphi(\overline{x},\overline{u})=0\quad \left(\overline{u}=\left((\overline{y}_1)^2,\ldots,(\overline{y}_m)^2\right)\right);$$

$$\overline{y}_j f_j(\overline{x}) = 0, \quad f_j(\overline{x}) \le 0, \quad j = \overline{1,m}.$$

*Зауваження 2′.* Якщо припущення 2 виконується, і якщо послідовність $\left\{(x^k, y^k)\right\}_{k=0}^{\infty}$, яка породжена алгоритмом 2, збігається до точки $(\overline{x}, \overline{y})$, то точка $(\overline{x}, \overline{u})$, де $\overline{u} = \left((\overline{y}_1)^2, \ldots, (\overline{y}_m)^2\right)$ задовольняє умови Куна–Таккера для задачі 2.

***Теорема 2″.*** *Нехай виконуються умови:* $(\boldsymbol{iv})$ *– існує оптимальний локальний розв'язок* $x^*$ *початкової задачі 2;* $(\boldsymbol{v})$ *–*

$$\{1,\ldots,l\} = \left\{ j \middle| f_j(x^*) = 0, \ j = \overline{1,m} \right\}, \ \ l \le m;$$

$(\boldsymbol{vi})$ *– функції* $f_i$, $i = \overline{0,m}$ *– двічі неперервно диференційовані у відкритому околі точки* $x^*$; $(\boldsymbol{vii})$ *– існує лише єдиний вектор множників Лагранжа* $u^* \in R_+^m$ *такий, що пара* $(x^*, u^*)$ *– є точкою Куна-Таккера, яка задовольняє достатні умови оптимальності другого порядку та* $x^*$ *– регулярна точка для* $f_j$, $j \in \{1,\ldots,l\}$; $(\boldsymbol{viii})$ *– виконуються умови строгої доповнюючої нежорсткості в точці* $(x^*, u^*)$, *тобто* $u_j^* > 0$, *якщо* $f_j(x^*) = 0$.

*Тоді для достатньо великого скінченного числа* $\alpha$ *існують околи* $V(x^*)$ *та* $W(y^*)$ *точок* $x^*, y^*$ *такі, що для будь-якої точки* $y^0 \in W(y^*)$ *існує єдина точка* $x^0 \in \overline{V}(x^*)$ *(де* $\overline{V}(x^*)$ *– замикання множини* $V(x^*)$*) така, що* $\nabla_x \psi(x^0, y^0, \alpha) = 0$ *і послідовність* $\left\{(x^k, y^k)\right\}_{k=0}^{\infty}$ *збігається до* $(x^*, y^*)$ *лінійно. Більш того, якщо* $\alpha$ *прямує до нескінченності, то послідовність* $\left\{y^k\right\}_{k=0}^{\infty}$ *збігається до точки* $y^*$ *надлінійно.*



## 6.13. Методи навантаженого функціоналу

В методах, наведених в цьому підрозділі на $k$-й ітерації потрібно шукати абсолютний мінімум по $x$ навантаженого функціоналу:

$$\psi(x, \alpha_k) \triangleq \left[ f_0(x) - \alpha_k \right]^2 + \sum_{i=1}^{m} \left( f_i(x) \right)^2,$$

де числова послідовність $\left\{\alpha_k\right\}_{k=0}^{\infty}$ будується в алгоритмах таким чином, щоб виконувалась така рівність:

$$\lim_{k \to \infty} \alpha_k = \min_{x \in X} f_0(x).$$

**1. Загальний випадок**

Задача 1. Знайти $\arg\min\limits_{x\in X} f_0(x)$ для заданої функції $f_0 : R^n \to R^1$ та заданої множини:

$$X \triangleq \left\{ x \middle| f_i(x)=0,\ i=\overline{1,m},\ x\in R^n \right\},$$

де $f_i : R^n \to R^1$, $m<n$.

В алгоритмі 1 на $k$-й ітерації обчислюють точку:

$$x^k = \arg\min_{x\in R^n} \psi(x,\alpha_k),$$

де $\psi(x,\alpha_k) \triangleq \left[ f_0(x)-\alpha_k \right]^2 + \sum\limits_{i=1}^{m} \left( f_i(x) \right)^2$, $\alpha_k \to \min\limits_{x\in X} f_0(x)$ при $k\to\infty$.

Граничні точки послідовності $\left\{x^k\right\}_{k=0}^{\infty}$ є розв'язком задачі 1. До того ж, на кожній ітерації обчислюється відрізок $[\alpha^-;\alpha^+]$, який містить значення $\min\limits_{x\in X} f_0(x)$, та довжина якого при $k\to\infty$ прямує до нуля.

***Алгоритм 1***

Початок. I. Вибрати будь-яке початкове наближення $x^0 \in X$.

II. Покласти $\alpha_0 = f_0(x^0)$.

III. Покласти $k=1$.

IV. Визначити функцію $\psi(x,\alpha): R^n \times R^1 \to R^1$ за таким правилом

$$\psi(x,\alpha) = \left( f_0(x)-\alpha \right)^2 + \sum_{i=1}^{m} \left( f_i(x) \right)^2 .$$

V. Вибрати константу $\delta_0 > 0$ ($\delta_0$ – достатньо мала величина) та покласти $\delta = \delta_0$.

Основний цикл. VI. Обчислити

$$\alpha_k = \alpha_{k-1} - \delta .$$

VII. Обчислити вектор $x^k$, який задовольняє умову

$$\psi(x^k,\alpha_k) = \min_{x\in R^n} \psi(x,\alpha_k).$$

VIII. Якщо виконується нерівність

$$\left[ \psi(x^k,\alpha_k) \right]^{1/2} \le \text{eps}\ \ (\text{eps} – «машинний нуль»),$$

то покласти $\delta = 2\delta$, $k=k+1$ та перейти на крок VI; інакше покласти $\alpha^+ = \alpha_{k-1}$ та перейти на крок IX.

IX. Покласти $\alpha^- = \alpha_k$ та перейти на крок X.

X. Обчислити значення $\sigma_{k+1}$ та $\tau_{k+1}$, за формулами:

$$\sigma_{k+1} = \alpha_k + \left( \psi(x^k,\alpha_k) \right)^{1/2},$$

$$\tau_{k+1} = \alpha_k + \psi(x^k, \alpha_k) \Bigg/ \left[ \psi(x^k, \alpha_k) - \sum_{i=1}^{m} \left( f_i(x^k) \right)^2 \right]^{1/2}.$$

XI. Покласти $k = k + 1$.

XII. Покласти $a^- = \sigma_k$.

XIII. Якщо $\tau_k < \alpha^+$, тоді покласти $\alpha_k = \tau_k$ та перейти на крок XIV; інакше покласти $\alpha_k = \sigma_k$ та перейти на крок XIV.

XIV. Обчислити вектор $x^k$, який задовольняє умову

$$\psi(x^k, \alpha_k) = \min_{x \in R^n} \psi(x, \alpha_k).$$

XV. Якщо $\left[ \psi(x^k, \alpha_k) \right]^{1/2} \geq \text{eps}$, то перейти на крок X; інакше покласти $\alpha^+ = \alpha_k$ та перейти на крок XVI.

XVI. Якщо $\alpha^+ - \alpha^- < \text{eps}$, тоді припинити обчислення (в цьому випадку знаходять розв'язок $x^k$ задачі 1); інакше покласти $\alpha_k = \alpha^-$ та перейти на крок XIV.

***Теорема 1.*** *Нехай* $(\boldsymbol{i})-$ *функції* $f_j(x)$, $j = \overline{0, m}-$ *неперервні в будь-якій замкненій обмеженій області Y простору* $R^n$; $(\boldsymbol{ii})-$ *задача 1 має розв'язок* $x^*$; $\text{eps} = 0$; $(\boldsymbol{iii})-$ $\psi(x^k, \alpha_k) = \min_{x \in R^n} \psi(x, \alpha_k)$ *може бути знайдене алгоритмами безумовної оптимізації;* $(\boldsymbol{iv})-$ $\psi(x^k, \alpha_k) > 0$ *еквівалентно* $f_0(x^k) < f_0(x^*)$. *Тоді алгоритм 1 збігається та граничні точки послідовності* $\{x^k\}_{k=0}^{\infty}$ *є розв'язками задачі 1.*

**2. Опуклий випадок**

Задача 2. Знайти $\arg\max_{x \in X} f_0(x)$ для заданої функції $f_0 : R^n \to R^1$ та заданої множини

$$X \triangleq \left\{ x \middle| f_i(x) \leq 0,\ i = \overline{1, m},\ x \in R^n \right\}.$$

*Припущення 2.* $(\boldsymbol{i})-$ функція $f_0(x)$ – опукла доверху, функції $f_j(x)$, $j = \overline{1, m}$ – опуклі донизу в $R^n$; $(\boldsymbol{ii})-$ функції $f_j(x)$, $j = \overline{0, m}$ – двічі неперервно диференційовані в $R^n$; $(\boldsymbol{iii})-$ множина $X$ – обмежена та непорожня.

Метод навантаженого функціоналу в опуклому випадку при певних умовах збігається до розв'язку задачі 2 із надлінійною швидкістю. На кожній ітерації алгоритму потрібно точно розв'язати задачу безумовної мінімізації навантаженого функціоналу. Для початку роботи алгоритму необхідно знати верхню оцінку величини $\max\limits_{x\in X} f_0(x)$.

***Алгоритм 2***

Початок. I. Знайти число $\alpha_1$, яке задовольняє умову:

$$\alpha_1 \ge f_0(x^*),$$

де $x^*-$ розв'язок задачі 2.

II. Визначити функцію $\psi(x,\alpha)$ (навантажений функціонал) за правилом

$$\psi(x,\alpha) \triangleq \left(\max\left\{0, \alpha - f_0(x)\right\}\right)^2 + \sum_{i=1}^{m}\left(\max\left\{0, f_i(x)\right\}\right)^2 .$$

III. Покласти $k=1$.

Основний цикл. IV. Обчислити вектор $x^k$, який задовольняє умову

$$\psi(x^k,\alpha_k) = \min_{x\in R^n}\psi(x,\alpha_k). \qquad (6.75)$$

V. Якщо $\psi(x^k,\alpha_k)=0$, тоді обчислення припиняються; інакше перейти на крок VI.

VI. Обчислити

$$\alpha_{k+1} = \alpha_k - \psi(x^k,\alpha_k)\left(\alpha_k - f_0(x_k)\right)^{-1}.$$

VII. Покласти $k=k+1$ та перейти на крок IV.

***Теорема 2.*** *Нехай виконуються припущення 2 та існує розв'язок $x^*$ задачі 2 такий, що має розв'язок наступна задача:*

$$\textit{знайти } \arg\max_{x\in Y}(\nabla f_0(x^*), x - x^*), \qquad (6.76)$$

*де* $Y = \left\{x \middle| (\nabla f_i(x^*), x-x^*) \le 0, \quad i \in \left\{j \middle| f_j(x^*)=0\right\},\ x\in R^n\right\}.$

*Тоді для нескінченної послідовності $\left\{\alpha_k\right\}_{k=1}^{\infty}$, яка породжена алгоритмом 2, існує константа $\beta$ така, що $\lim\limits_{k\to\infty}\alpha_k = f_0(x^*)$ та при достатньо великому $k$ виконуються наступні нерівності:*

$$0 < \alpha_{k+1} - f_0(x^*) \le \beta\left(\alpha_k - f_0(x^*)\right)^{3/2}.$$

***Теорема 2′.*** *Нехай виконуються припущення 2 та:* $(\boldsymbol{iv})-$ *існує розв'язок $x^*$ задачі 2 такий, що вектори $\nabla f_i(x^*),\ i\in\left\{j \middle| f_j(x^*)=0,\ j=\overline{1,m}\right\}$ – лінійно незалежні;* $(\boldsymbol{v})-$ *існує число $\delta>0$ для якого*

$$\left( y^T \frac{\partial^2 \varphi(p^*, x^*)}{\partial x^2}, y \right) \le -\delta \|y\|^2, \ \forall y \in R^n,$$

*де*

$$\varphi(p, x) = f_0(x) - \sum_{i=1}^{m} p_i f_i(x), \quad p = (p_1, ..., p_m);$$

$$(\nabla f(x^*))^T p^* = \nabla f_0(x^*), \quad f(x) = (f_1(x), .., f_m(x));$$

$$(p^*, f(x^*)) = 0, \quad p^* \ge 0.$$

*Тоді для нескінченної послідовності* $\{x^k\}_{k=1}^{\infty}$, *яка породжена алгоритмом 2, існує константа* $\eta \ge 0$ *така, що при достатньо великих* $k$ *виконується нерівність:*

$$\|x^k - x^*\| \le \eta \left(\alpha_k - f_0(x^*)\right)^{1/4} \|x^{k-1} - x^*\|,$$

*тобто послідовність* $\{x^k\}_{k=1}^{\infty}$ *збігається до* $x^*$ *надлінійно.*

**3. Наближена схема**



В наведеному нижче алгоритмі при відомій верхній оцінці значення $\max\limits_{x \in X} f_0(x)$ за скінченне число ітерацій обчислюється наближений розв’язок $\bar{x}$ задачі 2, такий, що:

$$\left(f_0(\bar{x}) - f_0(x^*)\right)^2 - \sum_{i=1}^{m} \left(\max\{0, f_i(\bar{x})\}\right)^2 \le \varepsilon_2,$$

де $x^*$– розв’язок задачі 2; $\varepsilon_2$– наперед задана константа. При цьому на кожній ітерації алгоритму задачу безумовної мінімізації навантаженого функціоналу достатньо розв’язувати також наближено.

***Алгоритм 3***

Початок. I. Вибрати будь-яке початкове наближення $x^0 \in R^n$.

II. Вибрати будь-яке число $\alpha_1$ (як правило, $\alpha_1 \ge f_0(x^*)$, де $x^*$– розв’язок задачі 2).

III. Задати будь-які константи $\varepsilon_1 > 0$ та $\varepsilon_2 > 0$, які визначають точність розв’язку допоміжної задачі та задачі 2 (рекомендується $\varepsilon_1 \in [10^{-5}; 10^{-1}]$, $\varepsilon_2 \in [10^{-5}; 10^{-1}]$).

IV. Визначити навантажений функціонал

$$\psi(x, \alpha) \triangleq \left(\max\{0, \alpha - f_0(x)\}\right)^2 + \sum_{i=1}^{m} \left(\max\{0, f_i(x)\}\right)^2.$$

V. Покласти $k = 1$.

Основний цикл. VI. Обчислити вектор $x^k$, для якого виконується нерівність

$$\psi(x^k,\alpha_k)\le\psi(x^{k-1},\alpha_k), \qquad (6.77)$$

а також або нерівності:

$$\left\|\nabla_x\psi(x^k,\alpha_k)\right\|<\varepsilon_1(\alpha_k-f_0(x^k)),\ \ \psi(x^k,\alpha_k)>\varepsilon_2, \qquad (6.78)$$

або

$$\psi(x^k,\alpha_k)\le\varepsilon_2\,. \qquad (6.79)$$

VII. Якщо виконуються нерівності (6.77) та (6.79), то обчислення припиняються; якщо (6.77) та (6.78), то перейти на крок VIII.

VIII. Обчислити

$$\alpha_{k+1}=\alpha_k-\psi(x^k,\alpha_k)\left(\alpha_k-f_0(x^k)\right)^{-1}. \qquad (6.80)$$

IX. Покласти $k=k+1$ та перейти на крок VI.

***Теорема 3.*** *Нехай виконуються умови* (***i***), (***iii***) *припущення 2 та функції* $f_j(x)$, $j=\overline{0,m}$ *– неперервно диференційовані. Тоді для будь-якого початкового наближення* $x^0\in R^n$ *та будь-яких значень* $\alpha_1$, $\varepsilon_1>0$, $\varepsilon_2>0$ *алгоритм 3 завершиться за скінченне число кроків, причому точки* $x^k$ *його зупинки, отримані при фіксованих* $x^0$, $\alpha_1\ge f_0(x^*)$ *та різних* $\varepsilon_1$, $\varepsilon_2$ *будуть, при* $\varepsilon_1\to0$, $\varepsilon_2\to0$, *збігатися до множини розв'язків задачі 2.*



***Теорема 3'.*** *Нехай виконуються припущення теореми 3 та крім цього:* (***i***)*–* *функція* $f_0(x)$ *– двічі неперервно диференційована;* (***ii***)*–* *існує константа* $\mu$ *така, що для будь-яких* $x$, $y\in R^n$ *виконується нерівність*

$$\left(y^T\frac{\partial^2 f_0(x)}{\partial x^2},y\right)\le-\mu\|y\|^2;$$

(***iii***)*–* *похідна* $\partial f_0(x^*)/\partial x$ *не дорівнює нулю;* (***iv***)*–* *в алгоритмі 3 нерівності (6.78) замінені на наступні:*

$$\alpha_k>f_0(x^k);$$

$$\left\|\nabla_x\psi(x^k,\alpha_k)\right\|\le\varepsilon_1\min\left\{\sum_{i=1}^m\left(\max\left\{0,f_i(x^k)\right\}\right)^2;\right.$$

$$\left.\left[\sum_{i=1}^m\left(\max\left\{0,f_i(x^k)\right\}\right)^2\left(\alpha_k-f_0(x^k)\right)\right]^{1/2}\right\}; \qquad (6.81)$$

$$\psi(x^k,\alpha_k)>\varepsilon_2;$$

$(\boldsymbol{v})$ – *функція* $f_0(x)$ – *строго опукла догори.*

*Тоді при достатньо малих* $\varepsilon_1 > 0$ *та будь-яких* $x^k, \alpha_k, \varepsilon_2 > 0$, *які задовольняють умовам* (6.81), *значення* $\alpha_{k+1}$, *обчислене на кроці* VIII *алгоритму 3 по* (6.80), *буде не менше, ніж* $f_0(x^*)$.

## 6.14. Методи штрафних оцінок

Задача 0. Знайти $\arg\min_{x \in X} f_0(x)$ для заданої функції $f_0 : R^n \to R^1$ і заданої множини

$$X \triangleq \{x \mid f_j(x) = 0, \quad j = \overline{1,m}, \quad x \in R^n\}.$$

Методи штрафних оцінок є комбінацією методів штрафних функцій і множників Лагранжа і оснований на побудові модифікованої функції Лагранжа $\psi : R^n \times R^m \times R^1_+ \to R^1$ :

$$\psi(x, y, \alpha) \triangleq f_0(x) + \sum_{j=1}^{m} y_j f_j(x) + \frac{\alpha}{2} \sum_{j=1}^{m} f_j^2(x), \tag{6.82}$$

де $\alpha \geq 0$, $y = (y_1, \ldots, y_m) \in R^m$.

В методі проводяться ітерації як по початковим змінним $x$, так і по двоїстим змінним $y$. В детермінованому випадку на кожній ітерації розв’язується допоміжна задача на безумовний мінімум. Метод штрафних оцінок в детермінованому випадку локально збігається зі швидкістю геометричної прогресії, знаменник якої тим менший, чим більший коефіцієнт штрафу $\alpha$. При наявності випадкових збурень методи штрафних оцінок збігаються до локального розв’язку з ймовірністю $1-\delta$ (оскільки без припущення опуклості функцій $f_j$ збіжності з ймовірністю 1 не існує).

#### 1. Детермінований випадок

### *Алгоритм 1*

Початок. І. Вибрати початкове наближення $y^0 \in R^m$ двоїстої змінної.

ІІ. Покласти $k = 0$.

Основний цикл. ІІІ. Вибрати коефіцієнт штрафу $\alpha_k$.

IV. Обчислити наступне наближення $x^{k+1}$ як точку безумовного мінімуму функції $\psi(x, y^k, \alpha_k)$ (де $\psi(x, y, \alpha)$ визначається за формулою (6.82)), тобто

$$\psi(x^{k+1}, y^k, \alpha_k) = \min_x \psi(x, y^k, \alpha_k).$$

V. Обчислити наступне наближення $y^{k+1}$ двоїстої змінної за формулою:

$$y^{k+1} = y^k + \alpha_k f(x^{k+1}),$$

де

$$f(x^{k+1}) = (f_1(x^{k+1}),..., f_m(x^{k+1})).$$

VI. Покласти $k = k+1$ і перейти на крок III.

***Теорема 1.*** *Нехай виконуються умови:* $(\boldsymbol{i})-$ *існує локальна точка мінімуму* $x^*$, *тобто:*

$$f_j(x^*) = 0,\ \ j = \overline{1,m},\ \text{і}\ f_0(x^*) \le f_0(x)$$

при

$$f_j(x) = 0,\ \ j = \overline{1,m},\ \| x - x^* \| \le \varepsilon,\ \ \varepsilon > 0;$$

$(\boldsymbol{ii})-$ в $\varepsilon$*-околі точки* $x^*$ *існують перша і друга похідні функцій* $f_j(x),\ j = \overline{0,m},$ *причому* $\nabla^2_{xx} f_j(x),\ j = \overline{0,m},$ *задовольняють умовам Ліпшиця:*

$$\| \nabla^2_{xx} f_j(x) - \nabla^2_{xx} f_j(\bar{x}) \| \le \gamma_1 \| x - \bar{x} \|,\quad \gamma_1 < \infty;$$

$(\boldsymbol{iii})-$ *градієнти* $\nabla f_j(x^*),\ \ j = \overline{1,m}$ *– лінійно незалежні*;

$(\boldsymbol{iv})-$ *виконується достатня умова мінімуму:*

$$(\nabla^2_{xx}\varphi(x^*, y^*)x, x) \ge \delta_1 \| x \|^2,\ \delta_1 > 0\ \ \textit{при}\ \ (\nabla^2_{yx}\varphi(x^*, y^*))x = 0.$$

*Тоді для будь-якого* $\rho > 0$ *знайдеться число* $\beta = \beta(\rho)$ *таке, що при* $\alpha_k \ge \beta,\ k = 0,1,\ldots,$ *алгоритм 1 (в якому під* $x^{k+1}$ *розуміється точка локального мінімуму функції* $\psi(x, y^k, \alpha_k)$ *із околу x*) збігається при будь-якому початковому наближенні* $y^0 \in S_\rho$, *де*

$$S_\rho \triangleq \{ y \,|\, \| y - y^* \| \le \rho,\ \ y \in R^m \},$$

*причому для всіх* $k \ge 1$ *справедливі оцінки:*

$$\| x^k - x^* \| \le (\theta_0)^k \| y^0 - y^* \| / (\alpha_0 \alpha_1 \ldots \alpha_{k-1});$$

$$\| y^k - y^* \| \le (\theta_0)^k \| y^0 - y^* \| / (\alpha_0 \alpha_1 \ldots \alpha_{k-1}),$$

*де* $\theta_0 > 0$ *– деяка константа.*

При умовах (***i***) – (***iii***) можна застосувати правило множників Лагранжа, тобто знайдеться $y^* \in R^m$ такий, що $\nabla_x \varphi(x^*, y^*) = 0$, де $\nabla_x \varphi(x^*, y^*)$ – похідна по $x$ функції Лагранжа

$$\varphi(x, y) \triangleq f_0(x) + \sum_{j=1}^{m} f_j(x) y_j.$$

Умови теореми (умови мінімуму) виражають умови невиродженості локального мінімуму і являються мінімально обмежуючими.

*Зауваження 1.* Із теореми 1 випливає, що при $\alpha_k = \bar{\alpha}$ для достатньо великих $\bar{\alpha}$ алгоритм 1 збігається не повільніше геометричної прогресії зі знаменником $q = \theta_0 / \bar{\alpha}$. Причому за рахунок збільшення $\bar{\alpha}$ знаменник може бути зроблений як завгодно малим.

*Зауваження 1′.* Алгоритм 1 можна модифікувати таким чином, щоб на кроці IV лише приблизно мінімізувати функцію $\psi(x, y^k, \alpha_k)$. А саме, в якості $x^{k+1}$ вибирати будь-яку точку з околу $x^*$, в якій:

$$\left\|\nabla_x \psi(x^{k+1}, y^k, \alpha_k)\right\| \le \mu \left\|f(x^{k+1})\right\|,$$

де $\mu \ge 0$ – наперед задана константа.

В умовах теореми 1 для модифікованого алгоритму справедливі оцінки:

$$\left\|x^k - x^*\right\| \le (\theta_0(1+\mu))^k \left\|y^0 - y^*\right\| / (\alpha_0 \alpha_1 \ldots \alpha_{k-1});$$

$$\left\|y^k - y^*\right\| \le (\theta_0(1+\mu))^k \left\|y^0 - y^*\right\| / (\alpha_0 \alpha_1 \ldots \alpha_{k-1})$$

(константа $\theta_0$ та ж сама, що і в теоремі 1).

*Зауваження 1″.* Якщо відоме наближення $x^0$ для $x^*$, то початкове наближення $y^0$ в алгоритмі 1 можна знайти з умови мінімуму квадратичного функціоналу

$$q_0(y) \triangleq \left\|\nabla f_0(x^0) + \left(\nabla^2_{yx} \varphi(x^0, 0)\right)^T y\right\|^2.$$

Чим ближче $x^0$ до $x^*$, тим ближче буде знайдене $y^0$ до $y^*$.

*Зауваження 1‴.* Найбільш істотною трудністю в застосуванні алгоритму 1 є проблема локальних мінімумів функції $\psi(x, y^k, \alpha_k)$: незрозуміло, як вибрати з них той, який знаходиться в околі $x^*$.

**2. Стохастичний випадок**

***Алгоритм 2***

Початок. I. Вибрати довільне початкове наближення $x^0 \in R^n$ і $y^0 \in R^m$ відповідно основної і двоїстої змінних.

II. Вибрати коефіцієнт штрафу $\alpha_0 > 0$.

III. Покласти $k = 0$.

Основний цикл. IV. Знайти кроковий множник $\rho_k$, який задовольняє умовам теореми 2.

V. Обчислити реалізацію $\xi^k$ випадкового вектора $\bar{\xi}^k$, який задовольняє умову

$$E(\bar{\xi}^k / x^0, y^0, \ldots, x^k, y^k) = \nabla f_0(x^k).$$

VI. Обчислити реалізацію $\zeta^k$ випадкового вектора $\bar{\zeta}^k$, який задовольняє умову:

$$E(\bar{\zeta}^k / x^0, y^0, \ldots, x^k, y^k) = f(x^k),$$

де $f(x^k)$ – вектор-стовпець з компонентами $f_1(x^k),\ f_2(x^k), \ldots, f_m(x^k)$.

VII. Обчислити матрицю $A_k$ розмірності $m \times n$, рядками якої є реалізації $(v_i^k)^T$ випадкових векторів $(\bar{v}_i^k)^T$, $i = \overline{1,m}$, які задовольняють умовам:

$$E(\bar{v}_i^k / x^0, y^0, \ldots, x^k, y^k) = \nabla f_i(x^k),\ \ i = \overline{1,m}.$$

VIII. Обчислити наступне наближення $x^{k+1}$ для основної змінної

$$x^{k+1} = x^k - \rho_k \left( \xi^k + (A_k)^T y^k + \alpha_0 (A_k)^T \zeta^k \right).$$

IX. Обчислити наступне наближення $y^{k+1}$ для двоїстої змінної за формулою

$$y^{k+1} = y^k + \rho_k \zeta^k.$$

X. Покласти $k = k+1$ і перейти на крок IV.

***Теорема 2.*** *Нехай виконуються умови* $(\boldsymbol{i})$ – $(\boldsymbol{iv})$ *теореми 1 та:* $(\boldsymbol{v})$ – *випадкові вектори* $\bar{\xi}^k$, $\bar{\zeta}^k$, $\bar{v}_i^k$, $i = \overline{1,m}$ – *взаємно незалежні при різних* $k$; $(\boldsymbol{vi})$ – *випадкові вектори* $\bar{\zeta}^k$ *і* $\bar{v}_i^k$, $i = \overline{1,m}$, – *взаємно незалежні,* $(\boldsymbol{vii})$ – *сума дисперсій компонент кожного з векторів* $\bar{\xi}^k$, $\bar{\zeta}^k$, $\bar{v}_i^k$, $i = \overline{1,m}$ – *обмежена числом* $\sigma^2$; $(\boldsymbol{viii})$ – *крокові множники* $\rho_k$ *такі, що:*

$$\rho_k \ge 0, \quad k = 0,1,\ldots, \quad \sum_{k=0}^{\infty} \rho_k = \infty, \quad \sum_{k=0}^{\infty} \rho_k^2 < \infty;$$

$(\boldsymbol{ix})$ – $E(\|x^0\|^2 + \|y^0\|^2) < \infty$.

*Тоді знайдеться таке число* $\bar{\alpha} > 0$, *що при* $\alpha_0 \ge \bar{\alpha}$ *послідовність* $\left\{(x^k, y^k)\right\}_{k=0}^{\infty}$, *яка породжена алгоритмом 2, задовольняє умову:*

$$P\left\{x^k \to x^*, y^k \to y^*\right\} \ge 1 - \delta,$$

*де*

$$\delta = \gamma_2 E\left(\|x^0 - x^*\|^2 + \|y^0 - y^*\|^2\right) + \gamma_3 \sigma^2 \sum_{k=0}^{\infty} \rho_k^2,$$

$\gamma_2, \gamma_3$ *– константи, що залежить від детермінованої частини задачі 1.*

*При цьому, якщо* $k\rho_k \to \rho$, $k\rho_k$ *– монотонно зростає, то для всіх* $k$ *при* $\gamma_4 > 0$ *виконується*

$$P\left\{\left\|x^k - x^*\right\|^2 + \left\|y^k - y^*\right\|^2 \le \gamma_4 \rho_k\right\} \ge 1 - \delta - 1/\gamma_4 .$$

*Зокрема, для будь-якого* $\delta_0 > 0$ *можна вибрати* $\overline{\beta}, \rho, \overline{\mu}, \gamma_4$ *так, що при*

$$\rho_k = \rho/(k+\overline{\mu}),\quad E\left(\left\|x^0 - x^*\right\|^2 + \left\|y^0 - y^*\right\|^2\right) \le \overline{\beta}$$

*для алгоритму 2 справедлива нерівність*

$$P\left\{\left\|x^k - x^*\right\|^2 + \left\|y^k - y^*\right\|^2 \le \gamma_4 / (k+\overline{\mu})\right\} \ge 1 - \delta_0 .$$

### 3. Метод штрафних оцінок для задач опуклого програмування

Задача 3. Знайти $\arg\min\limits_{x\in X} f_0(x)$ для заданої функції $f_0 : R^n \to R^1$ і множини

$$X \triangleq \{x \mid f_j(x) \le 0,\ \ j = \overline{1,m},\ \ x \in Q\},$$

де $Q$ – задана множина в $R^n$; $f_j : R^n \to R^1$, $j = \overline{1,m}$ – задані функції.

*Припущення 3.* $(\boldsymbol{i})$ – множина $Q$ – опукла і замкнута, $(\boldsymbol{ii})$ – функції $f_j(x)$, $j = \overline{0,m}$ – опуклі донизу і неперервні.

В опуклому випадку при звичних припущеннях метод штрафних оцінок збігається до множини розв'язків прямої і двоїстої задач. При виконанні умов регулярності розв'язків задачі 3 метод локально збігається зі швидкістю геометричної прогресії, причому за рахунок збільшення штрафного коефіцієнта знаменник геометричної прогресії може бути зроблений як завгодно малим.

***Алгоритм 3***

Початок. I. Вибрати довільне початкове наближення двоїстої змінної $y^0 \in R_+^m$.

II. Визначити модифіковану функцію Лагранжа $\psi(x, y, \alpha) : R^n \times R_+^m \times R_+^1 \to R^1$ ,

$$\psi(x, y, \alpha) \triangleq f_0(x) + \frac{1}{2\alpha}\sum_{i=1}^{m}\left(\left[y_i + \alpha f_i(x)\right]_+\right)^2 - \frac{1}{2\alpha}\sum_{i=1}^{m}(y_i)^2 .$$

(Тут і далі $[t]_+ \triangleq \max\{0, t\}$).

III. Покласти $k = 0$.

Основний цикл. IV. Знайти штрафний коефіцієнт $\alpha_k > 0$, який задовольняє умові $\alpha_k \ge \alpha > 0$, де $\alpha > 0$ – довільна константа.

V. Обчислити точку $x^k \in Q$, яка задовольняє умову

$$\psi(x^k, y^k, \alpha_k) = \min_{x\in Q} \psi(x, y^k, \alpha_k) .$$

VI. Покласти $h^k = \left( f_1(x^k), \dots f_m(x^k) \right)$.

VII. Обчислити точку

$$y^{k+1} = [y^k + \alpha_k h^k]_+ .$$

VIII. Покласти $k = k+1$ і перейти на крок IV.

***Теорема 3.*** *Нехай виконуються припущення 3 та:* $(\boldsymbol{iii})$ *– множина розв'язків* $X^*$ *задачі 3 непорожня і обмежена;* $(\boldsymbol{iv})$ *– множина розв'язків* $Y^*$ *двоїстої до задачі 3 непорожня.*

*Тоді алгоритм 3 при будь-якому* $y^0 \ge 0$ *(і будь-якому* $\alpha > 0$*) породжує обмежені послідовності* $\{x^k\}_{k=0}^{\infty}$, $\{y^k\}_{k=0}^{\infty}$ *такі, що:*

$$\min_{x^* \in X^*} \left\| x^k - x^* \right\| \to 0 \quad \text{при} \quad k \to \infty, \quad \min_{y^* \in Y^*} \left\| y^k - y^* \right\| \to 0 \quad \text{при} \quad k \to \infty,$$

$$\psi(x^k, y^k, \alpha_k) \le f_0(x^*), \quad x^* \in X^*.$$

Двоїстою до задачі 3 є задача:

$$\text{знайти} \quad \arg\max_{y \ge 0} g_0(y),$$

де $g_0(y) \triangleq \inf_{x \in Q} \phi(x, y)$, а $\varphi(x, y)$ – функція Лагранжа задачі 3, тобто

$$\varphi(x, y) \triangleq f_0(x) + \sum_{i=1}^{m} y_i f_i(x), \quad y = (y_1, \dots, y_m) \in R_+^m, \quad x \in Q .$$

Множина $X^* \times Y^*$ є множиною сідлових точок функції $\varphi(x, y)$.

***Теорема 3'.*** *Нехай виконуються припущення 3 і нехай* $(\boldsymbol{v})$ *– в задачі 3 множина* $Q = R^n$; $(\boldsymbol{vi})$ *– задача 3 має розв'язок* $x^*$; $(\boldsymbol{vii})$ *– в деякому околі* $x^*$ *функції* $f_j(x)$, $j = \overline{0, m}$ *– двічі неперервно диференційовані,* $(\boldsymbol{viii})$ *– вектори* $\nabla f_i(x^*)$, *при* $i \in \Im^* \triangleq \left\{ i \mid f_i(x^*) = 0, \ i \in [1:m] \right\}$ *– лінійно незалежні;* $(\boldsymbol{ix})$ *– якщо* $\left( \nabla f_i(x^*), x \right) = 0$ *для всіх* $i \in \Im^*$, *то:*

$$(\nabla_{xx}^2 \varphi(x^*, y^*) x, x) \ge \beta_1 \left\| x \right\|^2, \quad \beta_1 > 0,$$

*де* $\nabla_{xx}^2 \varphi(x^*, y^*)$ *– матриця других частинних похідних функції* $\varphi(x, y^*)$ *в точці* $x^*$, $\varphi(x, y)$ *– функція Лагранжа задачі 3;* $y^*$ *– множники Лагранжа задачі 3;* $(\boldsymbol{x})$ *– виконані умови строгої доповнюючої нежорсткості, тобто* $y_i^* > 0$ *при* $i \in \Im^*$.

*Тоді для будь-якого* $\alpha > 0$ *знайдеться* $\delta > 0$ *таке, що при* $\left\| y^0 - y^* \right\| < \delta$ *послідовності* $\{x^k\}_{k=0}^{\infty}$, $\{y^k\}_{k=0}^{\infty}$, *які породжені алгоритмом 3, задовольняють при всіх* $k \ge 0$ *нерівності:*

$$\begin{aligned} \left\| x^k - x^* \right\| &\le \gamma_1 \left\| y^k - y^* \right\|, \quad \gamma_1 > 0, \\ \left\| y^{k+1} - y^* \right\| &\le q_1 \left\| y^k - y^* \right\|, \quad 0 < q_1 < 1, \end{aligned} \tag{6.83}$$

*тобто послідовності* $\{x^k\}_{k=0}^{\infty}$, $\{y^k\}_{k=0}^{\infty}$ *збігаються зі швидкістю геометричної прогресії до* $x^*$ *і* $y^*$, *відповідно.*

Наслідок. При умовах теореми $3'$ для будь-яких $\alpha > 0$, $y^0 \geq 0$ знайдеться таке $k_0$, що при $k \geq k_0$ послідовності $\{x^k\}_{k=0}^{\infty}$, $\{y^k\}_{k=0}^{\infty}$, які породжені алгоритмом 3, задовольняють нерівностям (6.83).

***Теорема 3''.*** *Нехай виконуються всі припущення теореми 3' і, крім цього,* $(\boldsymbol{xi})-$ *матриці* $\nabla_{xx}^2 f_0(x)$ *і* $\nabla_{xx}^2 f_i(x)$, $i \in \Im^*$, *задовольняють умову Ліпшиця в деякому околі точки* $x^*$.

*Тоді для будь-якого* $y^0 \geq 0$ *знайдеться достатньо велике число* $\alpha = \alpha\left(\left\|y^0 - y^*\right\|\right) > 0$, *що при* $\alpha_k \geq \alpha$ *послідовності* $\{x^k\}_{k=0}^{\infty}$, $\{y^k\}_{k=0}^{\infty}$, *які породжені алгоритмом 3, задовольняють при всіх* $k \geq 0$ *нерівностям:*

$$\left\|x^k - x^*\right\| \leq \frac{\gamma_2}{\alpha_k}\left\|y^k - y^*\right\|, \quad \left\|y^{k+1} - y^*\right\| \leq \frac{\gamma_2}{\alpha_k}\left\|y^k - y^*\right\|,$$

*де постійна* $\gamma_2 > 0$ *не залежить від* $\alpha_k$.

## 6.15. Методи проекції узагальнених градієнтів



### 1. Основний метод

Задача 1. Знайти $\arg\min_{x \in X} f_0(x)$ для заданої функції $f_0 : R^n \to R^1$ і заданої множини $X \subset R^n$.

*Припущення 1.* (***i***) – функція $f_0$ – опукла донизу; (***ii***) – множина $X$ опукла і замкнута.

Цей метод схожий на метод проекції градієнтів, якщо замість градієнта використовувати узагальнений градієнт $\hat{\nabla} f_0$ функції $f_0$.

***Алгоритм 1***

Початок. І. Вибрати довільне початкове наближення $x^0 \in R^n$ і покласти $k = 0$.

Основний цикл. II. Обчислити узагальнений градієнт $\hat{\nabla} f_0(x^k)$ функції $f_0$ в точці $x^k$.

III. Обчислити значення крокового множника $\rho_k$ і нормуючого множника $\gamma_k$, які задовольняють умовам теореми 1.

IV. Обчислити наступне наближення

$$x^{k+1} = \pi_X(x^k - \rho_k \gamma_k \hat{\nabla} f_0(x^k)).$$

V. Покласти $k = k + 1$ і перейти на крок II.

***Теорема 1.*** *Нехай виконуються припущення 1 і умови:* $(\boldsymbol{i})-$ *існує* *точка* $x^*$, *яка належить множині мінімумів* $X^*$ *функції* $f_0$ *в області* $X$, *для якої* $\|x^*\| \le \text{const}$; $(\boldsymbol{ii})-$ *для будь-якого числа* $\alpha < \infty$ *знайдеться таке число* $\beta(\alpha) < \infty$, *що* $\left\|\hat{\nabla} f_0(x)\right\| \le \beta(\alpha)$ *при* $\|x\| \le \alpha$; $(\boldsymbol{iii})-$ *нормуючі множники* $\gamma_k$, $k = 0,1,\ldots$ *такі, що* $\gamma_k > 0$ *і* $\gamma_k \left\|\hat{\nabla} f_0(x^k)\right\| \le \text{const}$, $k = 0,1,\ldots$, *де* $\{x^k\}_{k=0}^{\infty}$ – *послідовність, яка породжена алгоритмом 1;* $(\boldsymbol{iv})$ – *крокові множники* $\rho_k$, $k = 0,1,\ldots$ *задовольняють умовам:*

$$\rho_k \ge 0,\ \rho_k \to 0,\ \text{при } k \to \infty;\ \sum_{k=0}^{\infty} \rho_k = \infty .$$

*Тоді знайдеться така підпослідовність* $\left\{f_0(x^{k_s})\right\}_{s=0}^{\infty}$ *послідовності* $\left\{f_0(x^k)\right\}_{k=0}^{\infty}$, *що*

$$\lim_{s\to\infty} f_0(x^{k_s}) = f_0(x^*),$$

*тобто*

$$\lim_{k\to\infty} \min_{s \le k} f_0(x^s) = f_0(x^*).$$

***Теорема 1'.*** *Якщо умови теореми 1 доповнити умовою, що множина мінімумів* $X^*$ *функції* $f_0$ *в області* $X$ *обмежена, то послідовність* $\left\{f_0(x^k)\right\}_{k=0}^{\infty}$, *яка породжена алгоритмом 1, збігається до* $f_0(x^*)$ *і*

$$\inf_{x^* \in X^*} \left\|x^* - x^k\right\| \to 0 \text{ при } k \to \infty .$$

*Зауваження 1.* Якщо в теоремі 1 вимагати, щоб замість $(\boldsymbol{iv})$ виконувалися умови:

$$\rho_k \ge 0, \quad \sum_{k=0}^{\infty} \rho_k = \infty, \quad \sum_{k=0}^{\infty} \rho_k^2 < \infty ,$$

то послідовність $\{x^k\}_{k=0}^{\infty}$, яка породжена алгоритмом 1, буде збігатися до розв’язку $x^* \in X^*$ без припущення про обмеженість множини $X^*$.

**2. Багатокроковий метод узагальненого градієнтного спуску**

Задача 2. Знайти $\arg\min_{x\in X} f_0(x)$ для заданої функції $f_0 : R^n \to R^1$ і заданої множини

$$X \triangleq \{x \mid f_1(x) \le 0,\ x \in R^n\}.$$

*Припущення 2.* $(\boldsymbol{i})-$ функції $f_0$ і $f_1$ – опуклі донизу на $R^n$; $(\boldsymbol{ii})-$ існує точка $\bar{x} \in R^n$ така, що $f_1(\bar{x}) < 0$; $(\boldsymbol{iii})-$ множина $X$ – обмежена.

В наведеному нижче алгоритмі напрямок спуску вибирається з використанням узагальнених градієнтів і значень функцій $f_0$ і $f_1$ на попередніх ітераціях. Крокові множники $\rho_k$ задовольняють класичним умовам.

***Алгоритм 2***

Початок. I. Вибрати довільне натуральне число $m \geq 1$.

II. Вибрати довільний набір точок $\{x^{-m+1}, \ldots, x^0\}$.

III. Вибрати константу $\alpha$, яка задовольняє умову $\alpha > f_0(x^*)$, де $x^*$ – розв'язок задачі 2.

IV. Вибрати константи $\delta_1 > 0$, $\delta_2 > 0$, причому $\delta_2$ вибрати з умови $\delta_2 < -f_1(\bar{x})$ (точка $\bar{x}$ визначається в умові (***ii***) в припущенні 2).

V. Покласти $k = 0$.

Основний цикл. VI. Обчислити кроковий множник $\rho_k$, який задовольняє умовам теореми 2.

VII. Якщо $f_0(x^k) > \alpha$, то покласти $\Im_k = \{k\}$ і перейти на крок VIII; якщо $f_0(x^k) \leq \alpha$, то покласти $\Im_k = \{k - m + 1, \ldots, k\}$ і перейти на крок VIII.

VIII. Обчислити $\hat{\nabla} f_0(x^k)$ і $\hat{\nabla} f_1(x^k)$ – узагальнені градієнти функцій $f_0$ і $f_1$ в точці $x^k$.

IX. Обчислити вектор $h^k$ з умови:

$$\min_{\|h\| \leq \rho_k} \varphi_k(h) = \varphi_k(h^k),$$

де $\varphi_k$ визначається за правилом:

$$\varphi_k(h) \triangleq \max_{j \in \Im_k} \psi_{kj}(h),$$

$$\psi_{kj}(h) = \begin{cases} f_0(x^j) - f_0(x^k) + \left(x^k - x^j, \hat{\nabla} f_0(x^j)\right) + \left(h, \hat{\nabla} f_0(x^j)\right), & f_1(x^j) < -\delta_2, \\ \max\left\{ f_0(x^j) - f_0(x^k) + \left(x^k - x^j, \hat{\nabla} f_0(x^j)\right) + \left(h, \hat{\nabla} f_0(x^j)\right), \right. & \\ \left. f_1(x^j) + \left(x^k - x^j, \hat{\nabla} f_1(x^j)\right) + \left(h, \hat{\nabla} f_1(x^j)\right) \right\}, & \delta_1 \geq f_1(x^j) \geq -\delta_2, \\ f_1(x^j) + \left(x^k - x^j, \hat{\nabla} f_1(x^j)\right) + \left(h, \hat{\nabla} f_1(x^j)\right), & f_1(x^j) > \delta_1. \end{cases}$$

X. Обчислити наступне наближення

$$x^{k+1} = x^k + h^k.$$

XI. Покласти $k = k + 1$ і перейти на крок VI.

***Теорема 2.*** *Нехай виконуються припущення 2 і крокові множники* $\rho_k$, $k = 0, 1, \ldots$ *задовольняють умовам:*

$$\rho_k \to +0 \;\; при \;\; k \to \infty, \quad \sum_{k=0}^{\infty} \rho_k = \infty.$$

*Тоді нескінченна послідовність* $\{x^k\}_{k=0}^{\infty}$*, яка породжена алгоритмом 2, така, що:*

$$\min_{x\in X^*}\left\|x^k - x\right\| \to 0 \ \text{при} \ k\to\infty;$$

$$f_0(x^k) \to \min_{x\in X} f_0(x) \ \text{при} \ k\to\infty,$$

*де* $X^* = \left\{x \middle| f_0(x) = \min_{y\in X} f_0(y), \ x\in X\right\}$.

**3. Методи проекції узагальненого градієнта для розв'язування граничних екстремальних задач**

Задача 3. Знайти $\arg\min_{x\in X} f_0(x)$, де $f_0(x) \triangleq \lim_{k\to\infty}\varphi_k(x)$, $\varphi_k : R^n \to R^1, \quad k = 0,1,\ldots$ – задані функції, $X \subset R^n$.

*Припущення 3.* $(\boldsymbol{i})$ – функції $\varphi_k(x),\ k = 0,1,\ldots$ – опуклі донизу; $(\boldsymbol{ii})$ – $X$ – опуклий компакт.

***Алгоритм 3***

Початок. I. Вибрати довільне початкове наближення $x^0 \in X$ .

II. Покласти $k = 0$.

Основний цикл. III. Обчислити вектор $\hat{\nabla}\varphi_k(x^k)$ – узагальнений градієнт функції $\varphi_k(x)$ в точці $x^k$.

IV. Знайти кроковий множник $\rho_k$, який задовольняє умовам теореми 3.

V. Обчислити наступне наближення

$$x^{k+1} = \pi_X(x^k - \rho_k\hat{\nabla}\varphi_k(x^k)).$$

VI. Покласти $k = k+1$ і перейти на крок III.

***Теорема 3.*** *Нехай виконуються припущення 3 та:* $(\boldsymbol{iii})$ – *послідовність* $\{\varphi_k(x)\}_{k=0}^{\infty}$ *збігається на X рівномірно до функції* $f_0(x)$;

$(\boldsymbol{iv})$ –

$$\rho_k > 0, \quad k = 0,1,\ldots; \ \ \rho_k \to 0 \ \text{при} \ k\to\infty; \ \ \sum_{k=0}^{\infty}\rho_k = \infty.$$

*Тоді для будь-якої збіжної підпослідовності* $\{x^{k_i}\}_{i=0}^{\infty}$ *послідовності* $\{x^k\}_{k=0}^{\infty}$*, яка породжена алгоритмом 3, справедливе граничне співвідношення:*

$$\lim_{i\to\infty} x^{k_i} = x^* \in X^*,$$

*де* $X^*$ – *множина розв'язків задачі 3.*

Нижче наводиться стохастичний аналог алгоритму 3.

*Алгоритм 3'*

Початок. I. Вибрати довільне початкове наближення $x^0 \in X$ .

II. Покласти $k = 0$.

Основний цикл. III. Обчислити реалізацію $\xi^k$ випадкового вектора $\overline{\xi}^k$, умовне математичне сподівання якого задовольняє співвідношенню

$$E(\overline{\xi}^k / x^0, x^1, \ldots x^k) = \hat{\nabla}\varphi_k(x^k) .$$

IV. Знайти кроковий множник $\rho_k$ , який задовольняє умовам теореми 3′.

V. Обчислити наступне наближення

$$x^{k+1} = \pi_X(x^k - \rho_k \xi^k) .$$

VI. Покласти $k = k + 1$ і перейти на крок III.

***Теорема 3'.*** *Нехай виконуються всі припущення теореми 3 і умова*

$$\sum_{k=0}^{\infty}(\rho_k)^2 < \infty .$$

*Тоді з ймовірністю 1 граничні точки послідовності* $\{x^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 3', належать множині розв'язків* $X^*$ *задачі 3.*

## 6.16. Методи умовного градієнта

### 1. Реалізовний метод умовного градієнта

Задача 1. Знайти $\arg\min_{x \in X} f_0(x)$ для заданої функції $f_0 : R^n \to R^1$ і заданої опуклої компактної множини $X$ .

*Припущення 1.* Функція $f_0$ – неперервно диференційована.

Методи умовного градієнта можуть застосовуватись для розв'язування задачі мінімізації нелінійної функції в області, в якій задача мінімізації лінійної функції розв'язується без великих обчислювальних затрат. В даному методі на $k$-й ітерації за напрямок руху до наступного наближення $x^{k+1}$ вибирається вектор $h^k = z^k - x^k$, де $z^k \in X$ задовольняє умову

$$\left(\nabla f_0(x^k), z^k\right) = \min_{z \in X}\left(\nabla f_0(x^k), z\right).$$

***Алгоритм 1***

Початок. I. Вибрати довільне початкове наближення $x^0 \in X$ .

II. Вибрати константу $\beta \in (0;1)$ (рекомендується $\beta = 1/2$ ).

III. Покласти $k = 0$.

Основний цикл. IV. Обчислити $\nabla f_0(x^k)$ .

V. Знайти точку $z^k \in X$ , яка задовольняє умову

$$\left(\nabla f_0(x^k), z^k\right) = \min_{z\in X}\left(\nabla f_0(x^k), z\right).$$

VI. Обчислити вектор $h^k = z^k - x^k$.

VII. Обчислити $\delta_k = (\nabla f_0(x^k), h^k)$.

VIII. Якщо $\delta_k = 0$, то покласти $x^* = x^k$ і зупинити обчислення; інакше перейти на крок IX.

IX. Покласти $j = 0$.

X. Обчислити $\rho_k = \beta^j$.

XI. Якщо виконується нерівність

$$f_0(x^k + \rho_k h^k) \le f_0(x^k) + \rho_k(1-\beta)\delta_k,$$

то перейти на крок XII; інакше покласти $j = j+1$ і перейти на крок X.

XII. Обчислити наступне наближення

$$x^{k+1} = x^k + \rho_k h^k.$$

XIII. Покласти $k = k+1$ і перейти на крок IV.

***Теорема 1.*** *Якщо виконується припущення 1 та:* $(\boldsymbol{i})-$ *множина* $X$ *опукла і компактна;* $(\boldsymbol{ii})-$ *градієнт функції* $f_0$ *в області* $X$ *задовольняє умову Ліпшиця*

$$\|\nabla f_0(x) - \nabla f_0(y)\| \le \gamma\|x-y\|, \quad \gamma < \infty, \quad \forall x, y \in X,$$

*то будь-яка гранична точка* $x^*$ *нескінченної послідовності* $\{x^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 1, задовольняє необхідну умову мінімуму функції* $f_0$ *на множині* $X$

$$\left(\nabla f_0(x^*), x^*\right) \le \left(\nabla f_0(x^*), x\right), \quad \forall x \in X.$$

Умова $\delta_k = 0$ також є необхідною умовою мінімуму функції $f_0$ на множині $X$.

***Теорема 1'.*** *Якщо виконуються умови теореми 1 і функція* $f_0$ *– опукла донизу, то нескінченна послідовність* $\{x^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 1 така, що*

$$\lim_{k\to\infty} f_0(x^k) = \min_{x\in X} f_0(x),$$

*причому справедлива наступна оцінка швидкості збіжності:*

$$f_0(x^k) - \min_{x\in X} f_0(x) \le \nu_1 / k,$$

*де* $\nu_1 > 0$ *– деяка константа.*

***Теорема 1''.*** *Якщо виконуються умови теореми 1' та:* $(\boldsymbol{i})-$ *область* $X$ *сильно опукла, тобто існує число* $\beta_0 > 0$ *таке, що для всіх* $x, y \in X$ *точки* $z = (x+y)/2 + \omega$ *належать* $X$ *для всіх* $\omega$, *які задовольняють умову* $\|\omega\| \le \beta_0\|x-y\|^2$; $(\boldsymbol{ii})-$ *існує константа* $\varepsilon_0 > 0$ *така, що* $\|\nabla f_0(x)\| \ge \varepsilon_0$ *для* *всіх* $x \in X$. *Тоді нескінченна послідовність* $\{x^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 1, збігається до точки мінімуму* $x^*$ *функції* $f_0$ *на множині* $X$ *зі швидкістю геометричної прогресії:*

$$\left\|x^k - x^*\right\| \le \nu_2 q_0^k, \quad q_0 < 1,$$

*де*

$$\nu_2 = \left(\left(f_0(x^0) - f_0(x^*)\right) \big/ 2\beta_0\varepsilon_0\right)^{1/2}, \quad q_0 = \left(1 - \beta_0\varepsilon_0 / 4\gamma\right)^{1/2}.$$

**2. Алгоритм Франка–Вульфа**

*Припущення 2.* $(\boldsymbol{i})-$ функція $f_0$ – опукла донизу і неперевно диференційована; $(\boldsymbol{ii})-$ $X-$ опукла множина.

***Алгоритм 2***

Початок. I. Вибрати довільне початкове наближення $x^0 \in X$ .

II. Покласти $k = 0$.

Основний цикл. III. Знайти точку $y^k \in X$ таку, що для всіх $x \in X$ виконується нерівність

$$\left(\nabla f_0(x^k), y^k\right) \le \left(\nabla f_0(x^k), x\right).$$

IV. Знайти число $\rho_k \in [0;1]$ таке, що

$$f_0\left((1-\rho_k)x^k + \rho_k y^k\right) \le f_0\left((1-\rho)x^k + \rho y^k\right) \text{ для всіх } \rho \in [0;1].$$

V. Покласти $x^{k+1} = (1-\rho_k)x^k + \rho_k y^k$ .

VI. Якщо $f_0(x^{k+1}) < f_0(x^k)$, то покласти $k = k+1$ і перейти на крок III; якщо $f_0(x^{k+1}) = f_0(x^k)$, то покласти $x^* = x^{k+1}$ і зупинити обчислення (в цьому випадку знаходять розв'язок задачі 1).

***Теорема 2.*** *Нехай виконуються припущення 2. Тоді граничні точки нескінченної послідовності* $\{x^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 2, є розв'язками задачі 1.*

**3. Прискорений алгоритм Франка–Вульфа**

***Алгоритм 3***

Початок. I. Вибрати довільне початкове наближення $x^1 \in X$ і ціле число $\tau \ge 0$.

II. Покласти $y^0 = x^1$, $\mu_1^0 = 1$, $k = 1$.

Основний цикл. III. Знайти точку $y^k \in X$ таку, що

$$\left(\nabla f_0(x^k), y^k\right) \le \left(\nabla f_0(x^k), x\right) \quad \forall\, x \in X.$$

IV. Знайти число $\rho_k \in [0;1]$ таке, що

$$f_0\left((1-\rho_k)x^k+\rho_k y^k\right)\le f_0\left((1-\rho)x^k+\rho y^k\right)\quad \forall\rho\in[0;1].$$

V. Покласти $j=0$.

VI. Обчислити числа:

$$\lambda_{k,j}^i=(1-\rho_k)\mu_k^i \text{ для } i=\overline{0,k-1};\quad \lambda_{k,j}^k=\rho_k.$$

VII. Покласти

$$z^{k,j}=(1-\rho_k)x^k+\rho_k y^k.$$

VIII. Якщо $j\ge\tau$, то перейти на крок XIV; якщо $j<\tau$, то перейти на крок IX.

IX. Обчислити множину індексів

$$\Im_{k,j}\triangleq\left\{i\middle|\left(\nabla f_0(z^{k,j}),\,y^i-z^{k,j}\right)\le 0,\quad i\in\{0,1,\ldots,k\}\right\}.$$

X. Обчислити число $\psi_{k,j}$ і вектор $\nu^{k,j}$ за формулами:

$$\psi_{k,j}=1\Big/\sum_{i\in\Im_{k,j}}\lambda_{k,j}^i;\quad \nu^{k,j}=\psi_{k,j}\sum_{i\in\Im_{k,j}}\lambda_{k,j}^i y^i.$$

XI. Обчислити число $\chi_{k,j}\in[0;1]$ таке, що

$$f_0\left((1-\chi_{k,j})z^{k,j}+\chi_{k,j}\nu^{k,j}\right)\le f_0\left((1-\chi)z^{k,j}+\chi\nu^{k,j}\right)\ \ \forall\chi\in[0;1].$$

XII. Обчислити числа $\lambda_{k,j+1}^i$, $i=\overline{0,k}$ за формулами:

$$\lambda_{k,j+1}^i=(1-\chi_{k,j}+\psi_{k,j}\chi_{k,j})\lambda_{k,j}^i \text{ для } i\in\Im_{k,j};$$

$$\lambda_{k,j+1}^i=(1-\chi_{k,j})\lambda_{k,j}^i \text{ для } i\notin\Im_{k,j}.$$

XIII. Покласти $z^{k,j+1}=(1-\chi_{k,j})z^{k,j}+\chi_{k,j}\nu^{k,j}$, $j=j+1$ і перейти на крок VIII.

XIV. Покласти $x^{k+1}=z^{k,\tau}$ і $\mu_{k+1}^i=\lambda_{k,\tau}^i$, $i=\overline{0,k}$.

XV. Якщо $f_0(x^{k+1})<f_0(x^k)$, то покласти $k=k+1$ і перейти на крок III, інакше покласти $x^*=x^{k+1}$ і зупинити обчислення (тобто в цьому випадку знаходять розв'язок $x^*$ задачі 1).

Для алгоритму 3 має місце теорема аналогічна теоремі 2.

## 6.17. Метод спряжених градієнтів

### 1. Метод спряжених градієнтів

Задача 1. Знайти $\arg\min\limits_{x\in X} f_0(x)$ для заданої функції $f_0:R^n\to R^1$ і заданої множини

$$X\triangleq\{x\mid f_j(x)=0,\ \ j=\overline{1,m},\ \ x\in R^n\}.$$

*Припущення 1.* $(\boldsymbol{i})-$ функції $f_i(x)$, $i=\overline{0,m}-$ двічі неперервно диференційовані в околі деякої точки $x^*$, в якій виконуються достатні умови локального мінімуму:

$$f_j(x^*) = 0,\ \ j = \overline{1,m};\ \ \nabla f_0(x^*) + \sum_{j=1}^{m} \lambda_j^* \nabla f_j(x^*) = 0;$$

$$\left(\left(\nabla_{xx}^2 f_0(x^*) + \sum_{j=1}^{m} \lambda_j^* \nabla_{xx}^2 f_j(x^*)\right)h,\ h\right) \ge \alpha \| h \|^2,\ \ \alpha > 0;\ \left(\nabla f_j(x^*), h\right) = 0.$$

Для застосування методу спряжених градієнтів до задач умовної мінімізації необхідно мати добрі початкові наближення точки локального мінімуму $x^*$. По суті, наведений тут метод являється методом спряжених градієнтів, який застосований для безумовної мінімізації деякої допоміжної функції. На кожній ітерації алгоритму потрібно лише обчислювати градієнти функцій $f_j(x),\ j = \overline{0,m}$ та розв'язувати задачу мінімізації одновимірної функції. Алгоритм 1 будує послідовність $\{x^k\}_{k=0}^{\infty}$, яка локально збігається до $x^*$ із квадратичною швидкістю.

***Алгоритм 1***

Початок. I. Вибрати довільне початкове наближення $x^0$ із околу точки локального мінімуму $x^*$.

II. Покласти $k = 0$.

Основний цикл. III. Обчислити вектори $\nabla f_0(x^k)$, $\nabla f_1(x^k),...,\nabla f_m(x^k)$.

IV. Покласти $i = 1$.

V. Покласти $e_{(j)}^0 = \nabla f_j(x^k),\ \ j = \overline{1,m}.$

VI. Обчислити вектори $e_{(j)}^i,\ \ j = \overline{1,m},$ за правилами:

$$e_{(i)}^i = (e_{(i)}^{i-1}, e_{(i)}^0)^{-1} e_{(i)}^{i-1};$$

$$e_{(j)}^i = e_{(j)}^{i-1} - (e_{(j)}^{i-1}, e_{(i)}^0) e_{(i)}^i,\ \ j = \overline{1,m};\ \ j \ne i.$$

VII. Якщо $i = m$, то покласти $e^j = e_{(j)}^m,\ \ j = \overline{1,m},$ і перейти на крок VIII; інакше (якщо $i < m$) покласти $i = i + 1$ і перейти на крок VI.

VIII. Обчислити числа

$$\bar{\lambda}_j = -\left(\nabla f_0(x^k), e^j\right),\ \ \ j = \overline{1,m}.$$

IX. Визначити функцію

$$\varphi(x) = f_0(x) + \sum_{j=1}^{m} \bar{\lambda}_j f_j(x).$$

X. Покласти $s = 0$.

XI. Обчислити $n$-вимірні вектори:

$$u^0 = x^k - \sum_{j=1}^{m} f_j(x^k) e^j;$$

$$g^0 = \nabla\varphi(u^0) - \sum_{j=1}^{m}\left(\nabla\varphi(u^0),\, e^j\right)\nabla f_j(x^k); \quad h^0 = -g^0.$$

XII. Якщо $s = n - m$, то перейти на крок XVI; інакше (якщо $s < n - m$) перейти на крок XIII.

XIII. Знайти значення $\rho_s \geq 0$ з умови:

$$\varphi(u^s + \rho_s h^s) = \min_{\rho \geq 0} \varphi(u^s + \rho h^s).$$

XIV. Обчислити вектори $u^{s+1}$, $g^{s+1}$, $h^{s+1}$:

$$u^{s+1} = u^s + \rho_s u^s; \quad g^{s+1} = \nabla\varphi(u^{s+1}) - \sum_{j=1}^{m}\left(\nabla\varphi(u^{s+1}), e^j\right)\nabla f_j(x^k);$$

$$h^{s+1} = -g^{s+1} + \left(\| g^{s+1} \| / \| g^s \|\right)^2 h^s.$$

XV. Покласти $s = s + 1$ і перейти на крок XII.

XVI. Покласти $x^{k+1} = u^{n-m}$, $k = k + 1$ і перейти на крок III.

***Теорема 1.*** *Нехай виконуються припущення 1 та: (**ii**) – другі похідні функцій $f_0(x), f_1(x), ..., f_m(x)$ задовольняють в околі точки $x^*$ умову Ліпшиця; (**iii**) – градієнти $\nabla f_1(x^*), ..., \nabla f_m(x^*)$ – лінійно незалежні.*

*Тоді послідовність $\{x^k\}_{k=0}^{\infty}$, яка породжена алгоритмом 1, локально збігається до точки $x^*$ із квадратичною швидкістю, тобто:*

$$\| x^k - x^* \| \leq \beta^{-1}(\beta \| x^0 - x^* \|)^{2^k}, \quad k = 0, 1, \dots,$$

де $\beta > 0$ – деяка константа.

### 2. Стохастичний аналог методу спряжених градієнтів

З а д а ч а 2. Знайти $\arg\min_{x \in X} f_0(x)$ для заданої функції $f_0: R^n \to R^1$ і заданої множини $X$.

*Припущення 2.* (***i***) – функція $f_0$ – неперервно диференційована та опукла донизу; (***ii***) – $X$ – опукла, обмежена і замкнута множина.

Стохастичний аналог методу спряжених градієнтів може бути застосований і в тому випадку, коли функція (або її градієнт) точно не обчислюються. На кожній ітерації будується вектор стохастичного квазіградієнта $\xi^k$ і напрямок $z^{k+1}$ руху до наступного наближення $x^{k+1}$ вибирається за допомогою визначеного усереднення векторів $\xi^k$, $\xi^{k-1}$,

$\xi^{k-2},\ldots$ Усереднення реалізується таким чином, щоб забезпечити (з ростом $k$) наближення значення $z^k$ до градієнта цільової функції в точці $x^k$.

***Алгоритм 2***

Початок. I. Вибрати: довільне початкове наближення $x^0 \in R^n$; крокові множники $\rho_1(0)$ і $\rho_2(0)$, які задовольняють умовам теореми 2; довільний вектор $z^0 \in R^n$ і довільне число $\delta \geq \| z^0 \|$.

II. Покласти $k=0$.

Основний цикл. III. Обчислити реалізацію $\xi^k$ випадкового вектора $\bar{\xi}^k$, умовне математичне сподівання якого:

$$E\left(\bar{\xi}^k /(x^0,\, z^0),\, \ldots,\, (x^k,\, z^k)\right) = \alpha \nabla f_0(x^k) + b^k, \tag{6.84}$$

де $\alpha -$ деяке невід’ємне число; $b^k -$ вектор «похибки», вимірний відносно $\sigma$-підалгебри $\mathcal{B}_k$, індукованої величинами $(x^0,\, z^0),\, \ldots,\, (x^k,\, z^k)$.

IV. Обчислити вектор

$$\bar{z}^{k+1} = z^k + \rho_2(k)(\xi^k - z^k).$$

V. Обчислити вектор

$$z^{k+1} = \begin{cases} \bar{z}^{k+1}, & \text{якщо } \| \bar{z}^{k+1} \| \leq \delta; \\ \delta \bar{z}^{k+1} / \| \bar{z}^{k+1} \|, & \text{якщо } \| \bar{z}^{k+1} \| > \delta. \end{cases}$$

VI. Обчислити наступне наближення

$$x^{k+1} = \pi_X \left( x^k - \rho_1(k) z^{k+1} \right).$$

VII. Обчислити значення крокових множників $\rho_1(k+1)$ і $\rho_2(k+1)$, які задовольняють умовам теореми 2.

VII. Покласти $k = k+1$ і перейти на крок III.

***Теорема 2.*** *Нехай виконуються припущення 2 і умови:* $(\boldsymbol{i}) -$ *градієнт функції* $f_0$ *задовольняє умову Ліпшиця*

$$\| \nabla f_0(x) - \nabla f_0(y) \| \leq \bar{\beta}_1 \| x - y \|, \quad \bar{\beta}_1 < \infty;$$

$(\boldsymbol{ii}) -$ $E\left(\| \bar{\xi}^k \|^2 /(x^0,\, z^0),\, \ldots,\, (x^k,\, z^k)\right) \leq \beta_2 < \infty, \quad k = 0,1,\ldots;$

$(\boldsymbol{iii}) - \rho_1(k),\ \rho_2(k),\ b^k -$ *вимірні відносно* $\sigma$*-підалгебри* $\mathcal{B}_k$ *і такі, що:*

$$\sum_{k=0}^{\infty} E\rho_1^2(k) < \infty, \quad \sum_{k=0}^{\infty} E\rho_1(k) \| b^k \| < \infty;$$

$$\sum_{k=0}^{\infty} E\rho_2^2(k) < \infty, \quad \sum_{k=0}^{\infty} \rho_1(k) = \infty \text{ майже напевно;}$$

$(\boldsymbol{iv}) -$ *для деяких чисел* $\lambda_k$ *виконуються умови:*

$$\rho_1(k) + \rho_2(k) \| b^k \| - \lambda_k \rho_2(k) \leq 0 \text{ майже напевно;}$$

$$\sum_{k=0}^{\infty} E\lambda_k \left(\rho_1(k) + \rho_2(k) \| b^k \|\right) < \infty.$$

*Тоді послідовність* $\{x^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 2, майже напевно збігається до розв'язку задачі 2 (а якщо розв'язок не єдиний, то* $\{x^k\}_{k=0}^{\infty}$ *збігається до одного з розв'язків).*

*Зауваження 2.* У випадку $\| b^k \| \equiv 0$ умови теореми 2 (***iii***) і (***iv***) виконуються, якщо тільки вибрати:

$$\rho_1(k) = k^{\tau+\varepsilon},\ \rho_2(k) = k^{\varepsilon};\ \lambda_k = k^{\tau},$$

де числа $\tau$ і $\varepsilon$ визначаються нерівностями

$$-1 < \varepsilon < -1/2; \quad -(1+\varepsilon) \le \tau < -(1+\varepsilon)/2.$$

## 6.18. Методи покоординатного спуску

З а д а ч а 0. Знайти $\arg\min_{x \in X} f_0(x)$ для заданої функції $f_0 : R^n \to R^1$ і паралелепіпеда:

$$X = \{x \mid \alpha_i \le x_i \le \beta_i,\ \ i = \overline{1,n},\ \ x \in R^n\},$$

де $\alpha_i$, $\beta_i$, $i = \overline{1,n}$ – задані дійсні числа.

*Припущення 0.* (***i***)– функція $f_0(x)$ – неперервно диференційована і для будь-яких $x, y \in X$ виконується

$$\|\nabla f_0(x) - \nabla f_0(y)\| \le \gamma \|x - y\|, \quad 0 < \gamma < \infty.$$



### 1. Детермінований покоординатний спуск

В методі детермінованого покоординатного спуску за напрямок руху послідовно вибираються орти. Параметри методу $\rho$ і $\delta$ обчислюються в процесі роботи алгоритму. Алгоритм включає «великі» ітерації по індексу $k$ і «малі» – по індексу $j$, причому при кожному $k$ «малі» ітерації закінчуються при скінченному значенні індексу $j$.

***Алгоритм 1***

П о ч а т о к. I. Вибрати довільне початкове наближення $x^0 \in X$ і обчислити $f_0(x^0)$.

II. Вибрати довільні константи $\rho_0$, $\delta_0$, $\beta$, які задовольняють нерівності:

$$\rho_0 > 0;\ \delta_0 > 0;\ \beta > 1.$$

III. Задати вектори $e^1, ..., e^n$ (тут $e^i$, $i = \overline{1,n}$ – вектор, $i$-та координата якого дорівнює одиниці, а інші дорівнюють нулю).

IV. Покласти $k = 0$.

Основний цикл. V. Покласти $j=0$.

VI. Обчислити індекс $i_k$ $(i_k \in [1:n])$ за правилом:

$$i_k = k - n\,\mathrm{Ent}(k/n) + 1,$$

де $\mathrm{Ent}(t)$ – ціла частина числа $t$.

VII. Покласти $h^k = e^{i_k}$.

VIII. Обчислити вектори:

$$x^{k(+)} = x^k + (\rho_k / 2^j)h^k; \quad x^{k(-)} = x^k - (\rho_k / 2^j)h^k.$$

IX. Якщо $x^{k(+)} \in X$ і $x^{k(-)} \in X$, то перейти на крок X; якщо $x^{k(+)} \in X$ і $x^{k(-)} \notin X$, то обчислити $f_0(x^{k(+)})$ і перейти на крок XV; якщо $x^{k(+)} \notin X$ і $x^{k(-)} \in X$, то обчислити $f_0(x^{k(-)})$ і перейти на крок XVI; якщо $x^{k(+)} \notin X$ і $x^{k(-)} \notin X$, то покласти $j = j+1$ і перейти на крок VIII.

X. Обчислити $f_0(x^{k(+)})$.

XI. Якщо виконується нерівність

$$f_0(x^{k(+)}) - f_0(x^k) \le -\rho_k \delta_k / 2^j, \tag{6.85}$$

то покласти:

$$x^{k+1} = x^{k(+)}; \quad \rho_{k+1} = \rho_k; \quad \delta_{k+1} = \delta_k; \quad f_0(x^{k+1}) = f_0(x^{k(+)})$$

і перейти на крок XVIII; інакше перейти на крок XII.

XII. Обчислити $f_0(x^{k(-)})$.

XIII. Якщо виконується нерівність

$$f_0(x^{k(-)}) - f_0(x^k) \le -\rho_k \delta_k / 2^j, \tag{6.86}$$

то покласти:

$$x^{k+1} = x^{k(-)}; \quad \rho_{k+1} = \rho_k; \quad \delta_{k+1} = \delta_k; \quad f_0(x^{k+1}) = f_0(x^{k(-)})$$

і перейти на крок XVIII; інакше перейти на крок XIV.

XIV. Якщо виконується нерівність

$$| f_0(x^{k(+)}) - f_0(x^{k(-)}) | \le 2\rho_k \beta \delta_k / 2^j,$$

то перейти на крок XVII; інакше покласти $j = j+1$ і перейти на крок VIII.

XV. Якщо виконується нерівність (6.85), то покласти:

$$x^{k+1} = x^{k(+)}; \quad f_0(x^{k+1}) = f_0(x^{k(+)}); \quad \rho_{k+1} = \rho_k; \quad \delta_{k+1} \le \delta_k$$

і перейти на крок XVIII; інакше перейти на крок XVII.

XVI. Якщо виконується нерівність (6.86), то покласти:

$$x^{k+1} = x^{k(-)}; \quad f_0(x^{k+1}) = f_0(x^{k(-)}); \quad \rho_{k+1} = \rho_k; \quad \delta_{k+1} \le \delta_k$$

і перейти на крок XVIII; інакше перейти на крок XVII.

XVII. Якщо виконуються рівності:

$$i_k = n; \quad x^k = x^{k-n+1},$$

то покласти:

$$x^{k+1} = x^k; \quad f_0(x^{k+1}) = f_0(x^k); \quad \rho_{k+1} = \rho_k / 2; \quad \delta_{k+1} = \delta_k / 2$$

і перейти на крок XVIII; якщо $i_k \neq n$ або $x^k \neq x^{k-n+1}$, то покласти:

$$x^{k+1} = x^k;\ f_0(x^{k+1}) = f_0(x^k);\ \ \rho_{k+1} = \rho_k;\ \ \delta_{k+1} = \delta_k$$

і перейти на крок XVIII.

XVIII. Покласти $k = k+1$ і перейти на крок V.

***Теорема 1.*** *Якщо виконуються припущення 0, функція* $f_0(x)$ *– опукла донизу на X і множина*

$$X_0 = \{x \mid f_0(x) \le f_0(x^0),\ \ x \in X\}$$

*обмежена (тут* $x^0$ *– початкове наближення в алгоритмі 1), то алгоритм 1 породжує послідовність* $\{x^k\}_{k=0}^{\infty}$, *для якої:*

$$\lim_{k\to\infty} \inf_{x^*\in X^*} \left\| x^k - x^* \right\| = 0,$$

*де*

$$X^* = \{x^* \mid f_0(x^*) = \min_{x\in X} f_0(x),\ \ \ x^* \in X\}.$$

## 2. Випадковий покоординатний спуск



В методі випадкового покоординатного спуску за напрямок руху на $k$-й ітерації випадково обирається $i_k$-й орт. Для обчислення крокових множників $\rho_k$ наводяться три різних правила, які легко реалізуються на практиці.

***Алгоритм 2***

Початок. I. Вибрати довільне початкове наближення $x^0 \in X$.

II. Задати вектори $e^1,...,e^n$ (тут $e^i$, $i = \overline{1,n}$ – $i$-й орт).

III. Покласти $k = 0$.

Основний цикл. IV. Знайти незалежну реалізацію $i_k$ випадкової величини $\tilde{i}$, яка приймає значення з множини $\{1,...,n\}$ з ймовірностями $p_1 = 1/n,..., p_n = 1/n$, відповідно.

V. Обчислити частинну похідну $\dfrac{\partial}{\partial x_{i_k}} f_0(x^k)$.

VI. Обчислити кроковий множник

$$\rho_k = \begin{cases} -\min\left\{ x_{i_k}^k - \alpha_{i_k}, \dfrac{\partial}{\partial x_{i_k}} f_0(x^k) \Big/ \gamma \right\}, & \text{якщо } \dfrac{\partial}{\partial x_{i_k}} f_0(x^k) \ge 0; \\ -\min\left\{ \beta_{i_k} - x_{i_k}^k, \left| \dfrac{\partial}{\partial x_{i_k}} f_0(x^k) \right| \Big/ \gamma \right\}, & \text{якщо } \dfrac{\partial}{\partial x_{i_k}} f_0(x^k) \le 0. \end{cases}$$

Тут припускається, що відома константа Ліпшиця $\gamma$ в умові (***i***) припущення 0. Якщо константа $\gamma$ невідома, то $\rho_k$ визначається за формулами, наведеними в зауваженні 2.

VII. Обчислити наступне наближення:

$$x^{k+1} = x^k - \rho_k e^{i_k}.$$

VIII. Покласти $k = k+1$ і перейти на крок IV.

***Теорема 2.*** *Якщо виконуються припущення 0 і множина*

$$X_0 = \{x \mid f_0(x) \le f_0(x^0),\ x \in X\}$$

*обмежена (тут $x^0$ – початкове наближення в алгоритмі 2), то алгоритм 2 породжує послідовність $\{x^k\}_{k=0}^{\infty}$, для якої з ймовірністю 1:*

$$\lim_{k\to\infty} \inf_{x^*\in X^*} \left\| x^k - x^* \right\| = 0,$$

*де*

$$X^* \triangleq \left\{ x^* \mid \left(\nabla f_0(x^*),\ x - x^*\right) \ge 0 \ \textit{для всіх}\ x \in X \right\}$$

*множина стаціонарних точок функції $f_0(x)$. Якщо, окрім цього, функція $f_0(x)$ опукла донизу на X, то з ймовірністю 1*

$$\lim_{k\to\infty} f_0(x^k) = \min_{x\in X} f_0(x).$$

*Зауваження 2.* Кроковий множник $\rho_k$ на кроці VI алгоритму 2 можна вибирати з умови:

$$f_0(x^k - \rho_k e^{i_k}) \le \lambda_k f_0(x^k) + (1-\lambda_k)\omega_k,$$

де $0 < \lambda \le \lambda_k \le 1$; $\beta_k = \inf_{\rho} f_0(x^k - \rho e^{i_k})$; $x^k - \rho e^{i_k} \in X$.

В якості $\rho_k$ також може бути вибране будь-яке число:

$$\rho \ge \delta \mu_k,$$

де $\delta \in (0;1]$; $\mu_k$ – найбільше з чисел, які задовольняють нерівність

$$f_0(x^k - \mu_k e^{i_k}) - f_0(x^k) \le -\frac{1}{2}\mu_k \left| \frac{\partial}{\partial x_{i_k}} f_0(x^k) \right|.$$

Для обчислення $\mu_k$ можуть бути використані наступні нерівності:

$$\mu_k \ge \min\left\{ x_{i_k}^k - \alpha_{i_k}, \frac{\partial}{\partial x_{i_k}} f_0(x^k) \Big/ \gamma \right\}, \text{ якщо } \frac{\partial}{\partial x_{i_k}} f_0(x^k) \ge 0;$$

$$\mu_k \ge -\min\left\{ \beta_{i_k} - x_{i_k}^k, \left| \frac{\partial}{\partial x_{i_k}} f_0(x^k) \right| \Big/ \gamma \right\}, \text{ якщо } \frac{\partial}{\partial x_{i_k}} f_0(x^k) \le 0.$$

## 6.19. Стохастичні квазіградієнтні методи

З а д а ч а  0. Знайти $\arg\min_{x\in X} f_0$ для заданої функції $f_0 : R^n \to R^1$ й заданої множини $X \subset R^n$.

*Припущення 0.* $(\boldsymbol{i})$– функція $f_0$ опукла донизу; $(\boldsymbol{ii})$– множина $X$ опукла й замкнута.

Стохастичні квазіградієнтні методи застосовуються для розв'язування задач оптимізації, у яких немає достатньо повної інформації про цільову функцію, функції обмежень та їх похідні. Ці методи основані на понятті стохастичного квазіградієнта – випадкового вектора $\xi^k$, який задовольняє рівність:

$$E(\xi^k / x^0,\ldots,x^k) = \alpha_k \hat{\nabla} f_0(x^k) + b^k,$$

де $\alpha_k$ – невід'ємна випадкова величина, $b^k$ – випадковий вектор, вимірні відносно $\sigma$-підалгебри, індукованої сімейством випадкових величин $(x^0,\ldots,x^k)$; $\hat{\nabla} f_0(x^k)$ – узагальнений градієнт функції $f_0$ в точці $x^k$.

Наведемо приклади обчислення стохастичних квазіградієнтів для функції $f_0$.

### 1. Приклади обчислення стохастичних квазіградієнтів

**Приклад 1.** Нехай функція $f_0(x)$ має обмежені другі похідні. Розглянемо вектор $\beta = (\beta_1,\ldots,\beta_n)^T$ з незалежними й рівномірно розподіленими на $[-1;1]$ компонентами. Тоді вектор стохастичного квазіградієнта на $k$-й ітерації можна обчислювати за формулою:

$$\xi^k = \sum_{s=1}^{s_k} \frac{f_0\left(x^k + \Delta_k \beta^{k,s}\right) - f_0(x^k)}{\Delta_k} \beta^{k,s},$$

де $\beta^{k,s}$, $s = \overline{1, s_k}$ – серія незалежних спостережень $\beta$ на $k$-й ітерації, причому $s_k \geq 1$; $\Delta_k > 0$. Якщо випадкові величини $s_k$, $\Delta_k$ вимірні відносно $\sigma$-підалгебри $\mathcal{B}_k$, індукованої величинами $x^0,\ldots,x^k$, то умовне математичне сподівання такого вектора дорівнює:

$$E(\xi^k / x^k) = \frac{s_k}{3} \nabla f_0(x^k) + v^k \Delta_k,$$

де $v^k$ – деякий випадковий вектор, вимірний відносно $\mathcal{B}_k$, $\left\|v^k\right\| \leq \text{const}$.

**Приклад 2.** Нехай функція $f_0(x)$ має вигляд:

$$f_0(x) = \sum_{i=1}^{r} p_i f_i(x),$$

де $p_i \geq 0$, $\sum_{i=1}^{r} p_i = 1$, і нехай функції $f_i(x)$, $i = \overline{1, r}$ – двічі неперервно диференційовані. Введемо випадкову величину $v$, яка приймає значення

$1, 2, \ldots, r$ з ймовірностями $p_1, p_2, \ldots, p_r$. За вектор стохастичного квазіградієнта на $k$-й ітерації можна взяти:

$$\xi^k = \sum_{s=1}^{s_k} \frac{f_{\nu_k}(x^k + \Delta_k \beta^{k,s}) - f_{\nu_k}(x^k)}{\Delta_k} \beta^{k,s},$$

де $\nu_k$ – реалізація випадкової величини $\nu$ в $k$-й ітерації, вектор $\beta^{k,s}$ і числа $s_k$, $\Delta_k$ такі, як і у прикладі 1. Умовне математичне сподівання такого вектора:

$$E(\xi^k / x^k) = \frac{s_k}{3} \nabla f_0(x^k) + v^k \Delta_k,$$

де $v^k$ – деякий випадковий вектор, вимірний відносно $\mathcal{B}_k$, $\|v^k\| \le \text{const}$.

**Приклад 3.** Нехай функція $f_0(x)$ має вигляд:

$$f_0(x) = \max_{y \in Y} f(x, y),$$

де функція $f(x, y)$ – опукла донизу по $x$, двічі неперервно диференційована по $x$ при кожному $y$; $Y$ – опукла компактна множина.

Узагальненим градієнтом функції $f_0$ в точці $x$ є вектор:

$$\hat{\nabla} f_0(x) = \left( \frac{\partial f(x, y)}{\partial x_1}, \ldots, \frac{\partial f(x, y)}{\partial x_n} \right)^T \Bigg|_{y = y(x)},$$

де вектор $y(x)$ такий, що $f(x, y(x)) = \max_{y \in Y} f(x, y)$.

За вектор стохастичного квазіградієнта на $k$-й ітерації можна взяти:

$$\xi^k = \sum_{s=1}^{s_k} \frac{f(x^k + \Delta_k \beta^{k,s}, y(x^k)) - f(x^k, y(x^k))}{\Delta_k} \beta^{k,s},$$

де вектор $\beta^{k,s}$ і числа $s_k, \Delta_k$ визначаються як і в прикладі 1. Умовне математичне сподівання такого вектора дорівнює:

$$E(\xi^k / x^k) = \frac{s_k}{3} \hat{\nabla} f_0(x^k) + v^k \Delta_k,$$

де $v^k$ – деякий випадковий вектор, вимірний відносно $\mathcal{B}_k$, $\|v^k\| \le \text{const}$.

**Приклад 4.** Якщо $f_0(x) = EF(x, \omega)$ і виконуються умови, достатні для диференціювання під знаком математичного сподівання, то в якості $\xi^k$ можна брати $\nabla_x F(x^k, \omega)$, тому що $E(\xi^k / x^k) = \nabla f_0(x^k)$, або

$$\xi^k = \sum_{s=1}^{s_k} \frac{F(x^k + \Delta_k \beta^{k,s}, \omega) - F(x^k, \omega)}{\Delta_k} \beta^{k,s},$$

де $\beta^{k,s}$ визначені в прикладі 1.

***Алгоритм 2***

П о ч а т о к . I. Вибрати довільну початкову точку $x^0 \in R^n$, для якої $E\left\|x^0\right\|^2 < \infty$; задати послідовність крокових множників $\{\rho_k\}_{k=0}^{\infty}$ і послідовність нормуючих множників $\{\gamma_k\}_{k=0}^{\infty}$ .

II. Покласти $k = 0$.

О с н о в н и й ц и к л . III. Обчислити реалізацію випадкового вектора $\xi^k(\omega)$, умовне математичне сподівання якого:

$$E(\xi^k / x^0, \ldots, x^k) = \alpha_k \hat{\nabla} f_0(x^k) + b^k,$$

де $\alpha_k$ – невід'ємна випадкова величина й $b^k = (b_1^k, \ldots, b_n^k)^T$ – випадковий вектор, що залежать від послідовності $x^0, \ldots, x^k$, або, більш точно, вимірні відносно $\sigma$-підалгебри $\mathcal{B}_k$, індукованої сімейством випадкових величин $(x^0, \ldots, x^k)$; $\hat{\nabla} f_0(x^k)$ – узагальнений градієнт функції $f_0(\cdot)$ в точці $x^k$.

IV. Обчислити вектор:

$$x^{k+1} = \pi_X(x^k - \rho_k \gamma_k \xi^k(\omega)), \tag{6.87}$$

де $\pi_X$ – оператор проектування на множину *X*.

V. Покласти $k = k + 1$ і перейти на крок III.

***Теорема 2.*** *Нехай мають місце припущення 0 і нехай, крім цього, для будь-якого числа* $w < \infty$ *знайдеться таке число* $c_w < \infty$, *що:*

$$E\left(\| \xi^k \|^2 / x^0, \ldots, x^k\right) \le \eta_k^2 \le c_w \text{ при } \left\|x^i\right\| \le w,\ i = \overline{0, k};$$

*величини* $\rho_k$, $\gamma_k$ *вимірні відносно* $\mathcal{B}_k$ *і такі, що:*

$$\gamma_k > 0,\ \gamma_k\left(\eta_k + \tau_k\left\|x^k\right\|\right) \le \text{const},$$

*де*

$$\tau_k = 1 \text{ при } \left\|b^k\right\| \ne 0,\ \tau_k = 0 \text{ при } \left\|b^k\right\| = 0;$$

$$\rho_k \ge 0,\ \sum_{k=0}^{\infty} \rho_k \alpha_k = \infty,\ \sum_{k=0}^{\infty} E\rho_k \left\|b^k\right\| < \infty.$$

*Тоді знайдеться підпослідовність* $x^{k_i}$ *послідовності* $\{x^k\}_{k=0}^{\infty}$, *яка породжена алгоритмом 2, для якої:*

$$\lim_{i \to \infty} f_0(x^{k_i}) = f_0(x^*) \text{ майже напевно,}$$

*або*

$$\lim_{k \to \infty} \min_{i \le k} f_0(x^i) = f_0(x^*) \text{ майже напевно.}$$

***Теорема 2'.*** *Нехай на додаток до умов теореми 2 та множина X**

*обмежена. Тоді*

$$\inf_{x^*\in X^*} E\left\|x^*-x^k\right\|^2 \to 0 \;\; при \;\; k\to\infty .$$

Для послідовності точок $x^0, x^1, \ldots,$ які породжені алгоритмом 2, у якому на кроці IV формула (6.87) замінена формулою

$$x^{k+1} = \pi_X\left(x^k - \rho_k \xi^k(\omega)\right),$$

має місце наступна теорема 2″, яка дає оцінку швидкості збіжності.

***Теорема 2''.*** *Нехай виконуються припущення 0 і нехай існують такі постійні* $\overline{\beta}$ *і* $\lambda$*, що:*

$$f_0(x) \ge f_0(x^*) + \beta\left\|x^*-x\right\|^2$$

*при* $x\in X$*, і з ймовірністю 1*

$$E(\|\xi^k\|^2 / x^0, \ldots, x^k) \le \lambda .$$

*Крім цього,* $\rho_k$ *детерміновані та:*

$$\alpha_k \ge \alpha > 0, \quad \left\|b^k\right\| \le r_k, \quad r_k \le \rho_k, \quad r_k < r < 2\overline{\beta}\alpha .$$

*Тоді існує таке число* $c_k \le \mathrm{const}$*, що при* $\rho_k = c_k / k$ *послідовність* $\left\{x^k\right\}_{k=0}^{\infty}$ *із ймовірністю 1 збігається до єдиного розв'язку* $x^*$*, причому*

$$E\left\|x^*-x^k\right\|^2 = O(1/k) .$$

*Означення 1.* Нехай $G$ – деяка підмножина простору $R^n$. Послідовність $z^s(\omega)$, $s=0,1,\ldots,$ випадкових векторів називається ***випадковою квазіфейєровською послідовністю*** щодо множини $G$, якщо для довільної точки $y\in G$:

$$E\left(\left\|y-z^{s+1}\right\|^2 / z^0, \ldots, z^s\right) \le \left\|y-z^s\right\|^2 + W_s, \quad s=0,1,\ldots,$$

де випадкові величини $W_s(\omega)\ge 0$ є вимірними відносно $\sigma$-підалгебри $\mathcal{B}_k$, індукованої сімейством $(z^0, \ldots, z^s)$ і такими, що $\sum_{s=0}^{\infty} EW_s < \infty$.

***Теорема 2'''.*** *Нехай мають місце припущення 0 і нехай* $\eta_k$ *– випадкова величина, вимірна відносно* $\sigma$*-підалгебри* $\mathcal{B}_k$*, індукованої величинами* $(x^0,\ldots,x^k)$*, така, що для будь-якого* $w$ *і деякого числа* $c_w$*:*

$$E\left(\|\xi^k\|^2 / x^0, \ldots, x^k\right) \le \eta_k^2 \le c_w \tag{6.88}$$

*при* $\left\|x^s\right\| \le w,\; s=\overline{0,k}$*; нормуючий множник* $\gamma_k$ *для деяких чисел* $\overline{\gamma}, \underline{\gamma}$ *задовольняє нерівностям:*

$$0 < \underline{\gamma} \le \gamma_k(\eta_k + \tau_k \|x^k\|) \le \overline{\gamma} < \infty,$$

*де* $\tau_k = 1$, *якщо* $\|b^k\| > 0$ *і* $\tau_k = 0$, *якщо* $\|b^k\| = 0$; *величини* $\rho_k$, $\alpha_k$, $b^k$ *такі, що:*

$$\rho_k \ge 0, \ \ \alpha_k \ge 0, \ \ \sum_{k=0}^{\infty} E(\rho_k \|b^k\| + \rho_k^2) < \infty; \tag{6.89}$$

*початкова точка* $x^0$ *така, що* $E\left(\| x^0 \|\right)^2 < \infty$. *Тоді послідовність точок* $x^k$, $k = 0, 1, \ldots,$ *яка породжена алгоритмом 2, є випадковою квазіфейєровською щодо множини* $X^*$. *Якщо ж, крім цього, з ймовірністю 1*

$$\sum_{k=0}^{\infty} \rho_k \alpha_k = \infty,$$

*то майже для кожного* $\omega$ *послідовність* $\left\{x^k(\omega)\right\}_{k=0}^{\infty}$ *збігається до деякого елемента* $x^*(\omega) \in X^*$.

*Зауваження 2.* Умови теореми $2'''$ легко перевіряються при розв'язанні конкретних задач. Відмітимо, що

$$E\left(\|\xi^k\|^2 / \mathcal{B}_k\right) = \sum_{j=1}^{n} D\left(\xi_j^k \big/ \mathcal{B}_k\right) + \alpha_k^2 \left\|\hat{\nabla} f_0(x^k)\right\|^2 + 2\alpha_k \left(\hat{\nabla} f_0(x^k), b^k\right) + \|b^k\|^2,$$

звідси, наприклад, випливає, що якщо сума дисперсій $\sum_{j=1}^{n} D(\xi_j^k \big/ \mathcal{B}_k)$ компонент вектора $\xi^k \triangleq (\xi_1^k, \ldots, \xi_n^k)^T$ обмежена в області $X$, а також обмежені $\alpha_k$, $\left\|\hat{\nabla} f_0(x^k)\right\|$, $\|b^k\|$, то $\eta_k = \text{const}$, тобто (6.88) виконується. Справедливість цієї умови в реальних задачах є наслідком обмеженості області *X*.

*Зауваження 2'.* Якщо область *X* – обмежена, то твердження теореми $2'''$ залишається справедливим при $\gamma_k = \text{const}$.

*Зауваження 2''.* Твердження теореми $2'''$ залишається справедливим також і при наступних (іноді корисних) змінах умов: $\|b^k\|$ задовольняє умовам, аналогічним (6.88) тобто $\|b^k\| \le c_w$ при $\|x^s\| \le w$, $s = \overline{0, k}$; нормуючий множник $\gamma_k$ задовольняє умовам:

$$0 < \underline{\gamma} \le \gamma_k \eta_k \le r_k; \quad \underline{\gamma} \le \gamma_k (1 + \|x^k\|) \|b^k\| \le \delta_k,$$

де $\underline{\gamma}$ – деяке число; $\delta_k$, $r_k$ – величини, вимірні відносно $\mathcal{B}_k$; замість (6.89) справедливі умови:

$$\rho_k \geq 0,\ \ \alpha_k \geq 0;\ \ \sum_{k=0}^{\infty} E(\rho_k \delta_k + \rho_k^2 r_k^2) < \infty .$$

**3. Стохастичний метод зменшення нев'язок у детермінованих задачах**

Задача 3. Знайти $\arg\min_{x \in X'} f_0(x)$, де $X' = X \bigcap \left\{ x \middle| f_j(x) \leq 0, \quad j = \overline{1,m} \right\}$, $X \subset R^n$ – задана множина, а, $f_i : R^n \to R^1$, $i = \overline{0,m}$ – задані функції.

*Припущення 3.* $(\boldsymbol{i})$– функції $f_j(\cdot)$, $j = \overline{0,m}$ – неперервні й опуклі донизу в області $X$; $(\boldsymbol{ii})$– множина $X$ опукла і замкнута; $(\boldsymbol{iii})$– обмеження задачі 3 задовольняють умові Слейтера, тому функція Лагранжа

$$\varphi(x,u) = f_0(x) + \sum_{j=1}^{m} u_j f_j(x) \tag{6.90}$$

має сідлову точку $(x^*, u^*)$ в області $x \in X$, $u \geq 0$ $\left(u = (u_1, u_2, \ldots, u_m)\right)$, причому множина $\{u^*\}$ компонент сідлових точок $W = \left\{(x^*, u^*)\right\}$ обмежена. Стохастичний метод зменшення нев'язок є своєрідним стохастичним варіантом градієнтного методу Ерроу–Гурвиця. Однак, тут не робиться припущення про точне обчислення функцій $f_j(x)$, $j = \overline{0,m}$. Стохастичний метод зменшення нев'язок оснований на застосуванні функції Лагранжа (6.90), сідлові точки якої обчислюються за допомогою методів стохастичних квазіградієнтів.

***Алгоритм 3***

Початок. I. Вибрати початкову точку $(x^0, u^0) \in R^n \times R_+^m$, для якої $E\left(\left\|x^0\right\|^2 + \left\|u^0\right\|^2\right) < \infty$.

II. Вибрати послідовність крокових множників $\left\{\rho_k\right\}_{k=0}^{\infty}$ та нормуючих множників $\left\{\gamma_k\right\}_{k=0}^{\infty}$ (вимірних відносно $\mathcal{B}_k$).

III. Покласти $k = 0$.

Основний цикл. IV. Обчислити реалізацію випадкового вектора $\xi^k$, який задовольняє рівність:

$$E(\xi^k / (x^0, u^0),\ \ldots, (x^k, u^k)) = \alpha_k \hat{\nabla}_x \varphi(x^k, u^k) + b^k ,$$

де випадкова величина $\alpha_k > 0$ і випадковий вектор $b^k$ вимірні відносно $\sigma$-

підалгебри $\mathcal{B}_k$, індукованої сімейством величин $\left((x^0, u^0), \ldots, (x^k, u^k)\right)$; $\hat{\nabla}_x \varphi(x^k, u^k)$ – вектор узагальненого градієнта функції Лагранжа (6.90) по змінній $x$ при фіксованому $u$ в точці $(x^k, u^k)$.

V. Обчислити реалізацію випадкового вектора $\zeta^k$, умовне математичне сподівання якого:

$$E\left(\zeta^k / (x^0, u^0),\ \ldots, (x^k, u^k)\right) = \alpha_k \nabla_u \varphi(x^k, u^k) + d^k,$$

де $d^k$ – випадковий вектор, вимірний відносно $\sigma$-підалгебри $\mathcal{B}_k$; $\nabla_u \varphi(x^k, u^k)$ – градієнт функції Лагранжа по змінній $u$ при фіксованому $x$ в точці $(x^k, u^k)$, який дорівнює вектору $\left(f_1(x^k), f_2(x^k), \ldots, f_m(x^k)\right)$.

VI. Знайти точку $(x^{k+1}, u^{k+1})$ за формулами:

$$x^{k+1} = \pi_X (x^k - \rho_k \gamma_k \xi^k),\ \ u^{k+1} = \pi_U (u^k + \rho_k \gamma_k \zeta^k),$$

де $\pi_U$ – оператор проектування на опуклу і замкнуту множину $U$, яка містить компоненти $u^*$ сідлових точок $(x^*, u^*)$ функції Лагранжа $\varphi(x, y)$ в області $x \in X$, $u \geq 0$.

VII. Покласти $k = k + 1$ і перейти на крок IV.

***Теорема 3.*** *Нехай мають місце припущення 3 і, крім того,* $f_0(x)$ – *строго опукла донизу функція. Нехай* $\eta_k$ – *випадкова величина, вимірна відносно* $\sigma$*-підалгебри* $\mathcal{B}_k$, *індукованої величинами* $(x^0, u^0), \ldots, (x^k, u^k)$ *така, що:*

$$E\left(\left\|\xi^k\right\|^2 + \left\|\zeta^k\right\|^2 / (x^0, u^0), \ldots, (x^k, u^k)\right) \leq \eta_k^2 \leq c_w < \infty$$

*при* $\left\|x^s\right\| + \left\|u^s\right\| \leq w < \infty$, $s = \overline{0, k}$; *нормуючий множник* $\gamma_k$ *для деяких чисел* $\overline{\gamma}, \underline{\gamma}$ *задовольняє нерівностям:*

$$0 < \underline{\gamma} \leq \gamma_k \left(\eta_k + \tau_k \left\|x^k\right\| + \theta_k \left\|u^k\right\|\right) \leq \overline{\gamma} < \infty,$$

*де* $\tau_k = 1$ *при* $\left\|b^k\right\| > 0$ *і* $\tau_k = 0$ *при* $\left\|b^k\right\| = 0$; $\theta_k = 1$ *при* $\left\|d^k\right\| > 0$ *і* $\theta_k = 0$ *при* $\left\|d^k\right\| = 0$; *величини* $\rho_k, \alpha_k$ *і вектори* $b^k$, $d^k$ *такі, що:*

$$\rho_k \geq 0,\ \alpha_k \geq 0,\ \sum_{k=0}^{\infty} E\left(\rho_k \left\|b^k\right\| + \rho_k \left\|d^k\right\| + \rho_k^2\right) < \infty. \qquad (6.91)$$

*Тоді послідовність точок* $\left\{(x^k, u^k)\right\}_{k=0}^{\infty}$ *яка породжена алгоритмом 2, є випадковою квазіфейєровською послідовністю щодо множини сідлових точок W. Якщо при цьому з ймовірністю 1*

$$\sum_{k=0}^{\infty} \rho_k \alpha_k = \infty ,$$

*то з ймовірністю 1 одна із граничних точок послідовності* $\left\{x^k\right\}_{k=0}^{\infty}$ *належить* $X^*$, *тобто майже напевно*

$$\lim_{s\to\infty}\left(\min_{k\in[0;s]} f_0(x^k)\right) = f_0(x^*), \quad x^* \in X^* .$$

*Зауваження 3.* Твердження теореми 2 залишається справедливим при умовах:

$$\rho_k \ge 0, \quad \alpha_k \ge 0, \quad \sum_{k=0}^{\infty} E(\rho_k \delta_k + \rho_k^2 r_k^2) < \infty ;$$

$$\left\|b^k\right\| + \left\|d^k\right\| \le c_w \text{ при } \left\|x^s\right\| + \left\|u^s\right\| \le w, \quad s = \overline{0,k};$$

нормуючий множник $\gamma_k$ задовольняє умовам:

$$0 < \underline{\gamma} \le \gamma_k \eta_k \le r_k < \infty;$$

$$\underline{\gamma} \le (1 + \left\|x^k\right\|)\left\|b^k\right\| + (1 + \left\|u^k\right\|)\left\|d^k\right\| \le \delta_k ,$$

де $\underline{\gamma}$ – деяке число; величини $r_k, \delta_k$ вимірні відносно $\mathcal{B}_k$ .

*Зауваження 3'.* Якщо припустити, що функції $f_j(\cdot)$, $j = \overline{0,m}$ обчислюються точно і функція Лагранжа має по $x$ другі похідні, обмежені в області $x \in X$, $u \in U$ , то на кроці IV алгоритму 2 можна взяти вектор:

$$\xi^k = \sum_{s=1}^{s_k} \frac{\varphi(x^k + \Delta_k \beta^{k,s}, u^k) - \varphi(x^k, u^k)}{\Delta_k} \beta^{k,s} ,$$

де $\beta^{k,s}$, $s = \overline{1, s_k}$ – серія незалежних спостережень вектора $\beta = (\beta_1, \ldots, \beta_n)^T$ з незалежними й рівномірно розподіленими на $[-1;1]$ компонентами на $k$ -й ітерації, причому $s_k \ge 1$, $\Delta_k > 0$ ;

на кроці V алгоритму 3 у якості $\zeta^k$ можна взяти вектор

$$\zeta^k = s_k / 3\left(f_1(x^k), \ldots, f_m(x^k)\right).$$

Тоді

$$E\left(\xi^k / \mathcal{B}_k\right) = \left(s_k / 3\right) \nabla_x \varphi(x^k, u^k) + v^k \Delta_k ,$$

де $\left\|v^k\right\| \le \text{const}$ ;

$$E\left(\zeta^k / \mathcal{B}_k\right) = \left(s_k / 3\right) \nabla_u \varphi(x^k, u^k) .$$

## 4. Метод зменшення нев'язок у задачах стохастичного програмування

З а д а ч а 4. Знайти $\arg\min\limits_{x \in X'} f_0(x)$, де:

$$X' \triangleq X \bigcap \left\{ x \middle| f_j(x) \le 0,\ j = \overline{1,m} \right\};$$

$$f_j(x) \triangleq EF_j(x,\omega),\ \ j = \overline{0,m};$$

$X \subset R^n$ – задана множина; $F_j : R^n \times \Omega \to R^1$ – задані функції.

Для функцій $f_j$, $j = \overline{0,m}$ і множини $X$ виконуються припущення 3.

***Алгоритм 4***

П о ч а т о к. I. Вибрати довільну початкову точку $(x^0, u^0) \in R^n \times R_+^m$.

II. Задати кроковий множник $\rho_0$, нормуючий множник $\gamma_0$ і величину зміщення $\Delta_0$, які задовольняють умовам теореми 4.

III. Покласти $k = 0$.

О с н о в н и й ц и к л. IV. Обчислити вектор:

$$\xi^k = \sum_{i=1}^{n} \frac{\Phi(x^k + \Delta_k e^i, u^k, \omega^{k,i}) - \Phi(x^k, u^k, \omega^{k,0})}{\Delta_k} e^i,$$

де $\omega^{k,\nu}$, $\nu = \overline{0,n}$ – серія незалежних по $k$ спостережень «стану природи $\omega$» (зокрема, можна взяти $\omega^{k,0} = \omega^{k,1} = \ldots = \omega^{k,n} = \omega^k$).

Тут функція $\Phi(x,u,\omega)$ визначається за правилом:

$$\Phi(x,u,\omega) = F_0(x,\omega) + \sum_{j=1}^{m} u_j F_j(x,\omega),$$

де $u = (u_1, \ldots, u_m)$; $e^i$, $i = \overline{1,m}$ – *i-й* орт.

Умовне математичне сподівання вектора $\xi^k$ дорівнює:

$$E(\xi^k / x^k, u^k) = \nabla_x \varphi(x^k, u^k) + v^k \Delta_k,$$

де функція $\varphi(x,u) \triangleq E\Phi(x,u,\omega)$;

$v^k$ – деякий вектор, для якого $\left\| v^k \right\| \le \text{const}$, тобто вектор $\xi^k$ є стохастичним квазіградієнтом функції $\varphi(x,u)$ в точці $(x^k, u^k)$.

V. Обчислити вектор

$$x^{k+1} = \pi_X(x^k - \rho_k \gamma_k \xi^k).$$

VI. Обчислити вектор:

$$u^{k+1} = \pi_U\left(u^k + \rho_k \gamma_k F(x^k, \omega^{k,0})\right),$$

де $F = (F_1, \ldots, F_m)$; множина $U$ визначена в пункті 3.

VII. Обчислити кроковий множник $\rho_{k+1}$, нормуючий множник $\gamma_{k+1}$ та величину зміщення по координатних осях $\Delta_{k+1}$, які задовольняють умовам теореми 4.

VIII. Покласти $k = k+1$ і перейти на крок IV.

***Теорема 4.*** *Нехай виконуються припущення 3 та функції* $f_j$, $j=\overline{0,m}$, *мають обмежені в області X другі похідні і* $f_0$ *– строго опукла донизу функція. Нехай, крім цього:*

$(\boldsymbol{i}) - \eta_k -$ *випадкова величина, вимірна відносно* $\sigma$*-підалгебри індукованої величинами* $(x^0,u^0),\dots,(x^k,u^k)$ *така, що:*

$$E\left(\left\|\xi^k\right\|^2+\left\|F(x^k,\omega^{k,0})\right\|^2 \Big/ (x^0,u^0),\dots,(x^k,u^k)\right) \le \eta_k^2 \le \beta(\chi) < \infty$$

*при* $\left\|x^s\right\|+\left\|u^s\right\| \le \chi < \infty$, $s=\overline{0,k}$ (тут $\beta(\chi)$ та $\chi$ *– обмежені константи*);

$(\boldsymbol{ii}) -$ *нормуючий множник* $\gamma_k$ *для деяких чисел* $\gamma', \gamma''$ *задовольняє нерівностям:*

$$0<\gamma' \le \gamma_k\left(\eta_k+\tau_k\left\|x^k\right\|\right) \le \gamma'' < \infty,$$

*де* $\tau_k = 1$ *при* $\left\|v^k\right\| > 0$, $\tau_k = 0$ *при* $\left\|v^k\right\| = 0$;

$(\boldsymbol{iii}) -$ *кроковий множник* $\rho_k$ *і величина зміщення по координатних осях* $\Delta_k$ *детерміновані і такі, що:*

$$\rho_k \ge 0; \quad \sum_{k=0}^{\infty}\rho_k = \infty; \quad \sum_{k=0}^{\infty}(\rho_k\left|\Delta_k\right| + \rho_k^2) < \infty .$$

*Тоді з ймовірністю 1 одна із граничних точок послідовності* $\left\{x^k\right\}_{k=0}^{\infty}$, *яка породжена алгоритмом 3, належить* $X^*$, *тобто майже напевно*

$$\lim_{s\to\infty}\left(\min_{k\in[0;s]} f_0(x^k)\right) = f_0(x^*), \quad x^* \in X^* .$$

#### 5. Гібридний стохастичний метод

З а д а ч а 5. Знайти $\arg\min\limits_{x\in X\bigcap A} f_0(x)$ для заданих множин $X \subset R^n$ та $A=\left\{x \middle| f(x) \le 0\right\}$.

*Припущення 5.* $(\boldsymbol{i}) -$ функція $f_0(\cdot)$ – неперервна та опукла донизу; $(\boldsymbol{ii}) -$ множини *X* і *A* – опуклі й замкнуті; $(\boldsymbol{iii}) -$ $A \cap X \ne \varnothing$.

Наведений нижче гібридний стохастичний метод пошуку екстремуму функції $f_0(\cdot)$ поєднує в собі ідеї стохастичного релаксаційного методу для розв’язування систем нерівностей і стохастичного квазіградієнтного методу для розв'язування задач нелінійного програмування.

***Алгоритм 5***

Початок. I. Вибрати початкову точку $x^0 \in R^n$, для якої $E\left\|x^0\right\|^2 < \infty$.

II. Задати послідовності крокових множників $\{\delta_k\}_{k=0}^{\infty}$; $\{\rho_k\}_{k=0}^{\infty}$.

III. Задати послідовності нормуючих множників $\{\beta_k\}_{k=0}^{\infty}$; $\{\gamma_k\}_{k=0}^{\infty}$.

IV. Покласти $k = 0$.

Основний цикл. V. Обчислити реалізацію випадкового вектора $\xi^k$, умовне математичне сподівання якого:

$$E(\xi^k / x^0, \ldots, x^k) = \alpha_k \hat{\nabla} f_0(x^k) + b^k, \tag{6.92}$$

де $\alpha_k$ – невід'ємна випадкова величина і $b^k$ – випадковий вектор, вимірні відносно $\sigma$-підалгебри $\mathcal{B}_k$, індукованої сімейством випадкових величин $(x^0, \ldots, x^k)$; $\hat{\nabla} f_0(x^k)$ – вектор узагальненого градієнта функції $f_0$ в точці $x^k$.

VI. Обчислити випадковий вектор $\zeta^k$, умовне математичне сподівання якого:

$$E(\zeta^k / x^0, \ldots, x^k) = \lambda_k g(x^k) + d^k, \tag{6.93}$$

де $\lambda_k$ – невід'ємна випадкова величина і $d^k$ – випадковий вектор, вимірні відносно $\sigma$-підалгебри $\mathcal{B}_k$; $g(x^k)$ – вектор, для якого півпростір, що відповідає нерівності

$$\left(g(x), z - x\right) + f(x) \le 0, \tag{6.94}$$

при $x = x^k$ містить множину *A*, якщо $x^k \in A$.

VII. Знайти вектор

$$x^{k+1} = \begin{cases} \pi_X(x^k - \delta_k \beta_k f(x^k) \zeta^k - \rho_k \gamma_k \xi^k), & f(x^k) > 0; \\ \pi_X(x^k - \rho_k \gamma_k \xi^k), & f(x^k) \le 0. \end{cases}$$

VIII. Покласти $k = k + 1$ та перейти на крок V.

***Теорема 5.*** *Нехай мають місце припущення 5 і нехай $\eta_k$ випадкова величина, вимірна відносно $\sigma$-підалгебри $\mathcal{B}_k$, така, що для будь-якого числа $w < \infty$ знайдеться число $c_w$, для якого:*

$$E\left(\left\|\xi^k\right\|^2 + \left\|\zeta^k\right\|^2 / x^0, \ldots, x^k\right) \le \eta_k^2 \le c_w$$

*як тільки $\left\|x^s\right\| \le w,\ s = \overline{0,k}$; для деяких чисел $\underline{\gamma}, \overline{\gamma}$ нормуючі множники $\beta_k$ і $\gamma_k$ задовольняють умовам:*

$$\beta_k \left( f(x^k) \left( \tau_1(k) \left\|x^k\right\| + 1 \right) + \eta_k \right) = 1;$$

$$\underline{\gamma} \le \gamma_k \left( \tau_2(k) \left\|x^k\right\| + \eta_k^2 \right) \le \overline{\gamma},$$

*де* $\tau_1(k)=1$, *якщо* $\left\|d^k\right\|>0$, $\tau_1(k)=0$, *якщо* $\left\|d^k\right\|=0$; $\tau_2(k)=1$, *якщо* $\left\|b^k\right\|>0$; $\tau_2(k)=0$, *якщо* $\left\|b^k\right\|=0$; *величини* $\rho_k$, $\delta_k$, $\alpha_k$, $\lambda_k$ *і вектори* $b^k, d^k$ *такі, що:*

$$\rho_k \ge 0, \quad 0 \le \delta_k < 2\lambda_k - \varepsilon_k, \quad \varepsilon_k \ge 0, \quad \alpha_k \ge 0, \quad \lambda_k \ge 0;$$

$$\sum_{k=0}^{\infty} E(\rho_k \left\|b^k\right\| + \rho_k^2 + \rho_k\delta_k + \delta_k \left\|d^k\right\|) < \infty .$$

*Тоді випадкова послідовність* $\left\{x^k\right\}_{k=0}^{\infty}$, *яка породжена алгоритмом 5, є випадковою квазіфейєровською щодо множини* $X^*$. *Якщо ж, крім цього, з ймовірністю 1:*

$$\sum_{k=0}^{\infty} \delta_k\varepsilon_k = \infty; \quad \sum_{k=0}^{\infty} \rho_k\alpha_k = \infty ,$$

*то вона збігається до деякого елемента* $x^* \in X^*$ *майже напевно.*

*Зауваження 5.* Якщо, наприклад,

$$f(x) = \max_i f_i(x) \triangleq f_{i(x)}(x),$$

де функції $f_i(x)$ – опуклі донизу і неперервно диференційовані; $i(x)$ – індекс, на якому досягається $\max\limits_i f_i(x)$ при заданому $x$, то вектор $g(x) = \nabla f_i(x)\big|_{i=i(x)}$ задовольняє нерівності (6.94).

*Зауваження 5′.* Відзначимо, що на кроці VII алгоритму 5 використовується стохастичний релаксаційний метод

$$x^{k+1} = \pi_X\left(x^k - \delta_k\beta_k f(x^k)\zeta^k\right)$$

для розв’язування системи нерівностей:

$$f_i(x) \le 0, \quad i = \overline{1,m}; \quad x \in X ,$$

де $f(x) = \max\limits_{i\in[0;m]} f_i(x)$.

## 6.20. Комбінований метод стохастичних градієнтів і штрафних функцій

Задача 1. Знайти $\arg\max\limits_{x\in X} f_0(x)$ для заданої функції $f_0 : R^n \to R^1$ і заданої множини

$$X = \left\{x \middle| f_i(x) \ge 0, \ i = \overline{1,m}, \quad x \in Y\right\}.$$

*Припущення 1.* $(\boldsymbol{i})$ – функції $f_j(x)$, $j = \overline{0,m}$ – неперервні разом зі своїми похідними; $(\boldsymbol{ii})$ – функції $f_j(x)$, $j = \overline{0,m}$ – опуклі доверху;

$(\boldsymbol{iii}) - Y -$ опуклий компакт із $R^n$.

У наведеному нижче методі визначаються функції:

$$g_\tau(x,\alpha) = f_0(x) - \alpha \sum_{i=1}^{m} p_i \left| \min\{0, f_i(x)\} \right|^\tau$$

(тут $\tau > 1$ – константа; $\alpha$ – коефіцієнти штрафу, які прямують до нескінченності; $p_i$, $i = \overline{1,m}$ – вибрані ймовірності) і для відшукання їхнього максимуму застосовується метод стохастичних градієнтів.

***Алгоритм 1***

П о ч а т о к . I. Вибрати довільне початкове наближення $x^1 \in Y$.

II. Задати числа $p_i$, $i = \overline{1,m}$ $(p_i > 0,\ i = \overline{1,m};\ \sum_{i=1}^{m} p_i = 1)$, які, відповідно, характеризують необхідну відносну точність виконання обмежень-нерівностей у задачі 1 (рекомендується $p_i = 1/m$, $i = \overline{1,m}$).

III. Вибрати параметр $\tau > 1$.

IV. Покласти $k = 1$.

О с н о в н и й  ц и к л . V. Знайти незалежну реалізацію $i_k$ випадкової величини $\tilde{i}$, що приймає значення із множини $\Im = \{1, \ldots, m\}$, відповідно, з ймовірностями $p_1, \ldots, p_m$.

VI. Знайти значення крокового множника $\rho_k$ й коефіцієнта штрафу $\alpha_k$, що задовольняють умовам теореми 1.

VII. Обчислити стохастичний градієнт $\xi^k$ функції $g_\tau(x, \alpha_k)$ в точці $x^k$

$$\xi^k = \nabla f_0(x^k) + \alpha_k \tau \left| \min\{0, f_{i_k}(x^k)\} \right|^{\tau-1} \nabla f_{i_k}(x^k).$$

VIII. Обчислити наступне наближення

$$x^{k+1} = \pi_Y(x^k + \rho_k \xi^k).$$

IX. Покласти $k = k+1$ і перейти на крок V.

***Теорема 1.*** *Якщо виконуються припущення 1, множина X – непорожня і числові послідовності* $\{\rho_k\}_{k=1}^{\infty}$, $\{\alpha_k\}_{k=1}^{\infty}$ *задовольняють умовам:*

$$\rho_k > 0,\ \lim_{k\to\infty} \rho_k = 0,\ \rho_{k+1} \le \rho_k,\ \sum_{k=1}^{\infty} \rho_k = \infty;$$

$$\alpha_k > 0,\ \lim_{k\to\infty} \alpha_k = \infty,\ \alpha_{k+1} \ge \alpha_k,\ \sum_{k=1}^{\infty} (\rho_k \alpha_k)^2 < \infty,$$

*то для будь-якого початкового наближення* $x^1$ *послідовність* $\{x^k\}_{k=1}^{\infty}$, *яка породжена алгоритмом 1, з ймовірністю 1 збігається до* $\operatorname{Arg}\max_{x\in X} f_0(x)$ – *множини розв'язків задачі 1.*

*Зауваження 1.* Алгоритм 1 рекомендується застосовувати для розв’язування задач математичного програмування з великим числом обмежень (особливо для тих задач, обмеження яких можуть формуватися в процесі розв’язування задач на ЕОМ) і для розв’язування задач із блочною структурою.

## 6.21. Методи усереднення напрямків спуску

### 1. Детермінована задача

З а д а ч а  1. Знайти $\arg\min\limits_{x\in X} f_0\left(x\right)$ для заданої функції $f_0 : R^n \to R^1$ і множини $X$ заданої співвідношенням:

$$X = X_1 \bigcap X_2 ,$$

де $X_1 = \left\{ x \mid f_j\left(x\right) \le 0,\ j = \overline{1,m},\ x \in R^n \right\}$, $X_2$ – опукла, замкнена, обмежена множина.

*Припущення 1.* $(\boldsymbol{i})-$ функції $f_j$, $j = \overline{0,m}$ – опуклі донизу; $(\boldsymbol{ii})-$ множина $X_2$ має непорожній перетин з множиною внутрішніх точок множини $X_1$.



Нижче наводяться два алгоритми усереднення напрямків спуску, перший з яких потребує знаходження субградієнтів функцій $f_j$, $j = \overline{0,m}$, а другий – лише знаходження значень функцій $f_j$, $j = \overline{0,m}$.

***Алгоритм 1***

П о ч а т о к . І. Вибрати довільне початкове наближення $x^0 \in R^n$.

ІІ. Вибрати довільне натуральне число $l$.

ІІІ. Покласти $k = 0$.

О с н о в н и й   ц и к л . IV. Знайти кроковий множник $\rho_k$, який задовольняє умови теореми 1.

V. Якщо виконується нерівність

$$\max_{j\in[1;m]} f_j\left(x^k\right) \le 0 ,$$

то покласти $v_k = 0$ та перейти на крок VII; інакше перейти на крок VI.

VI. Визначити індекс $i \in \left[1:m\right]$, який задовольняє умову

$$f_i\left(x^k\right) = \max_{j\in[1;m]} f_j\left(x^k\right),$$

покласти $v_k = i$ та перейти на крок VII.

VII. Знайти субградієнт $\hat{\nabla} f_{v_k}\left(x^k\right)$ функції $f_{v_k}$ в точці $x^k$.

VIII. Якщо $k \ge l$, то обчислити вектор

$$h^k = \sum_{i=k-l}^{k} \hat{\nabla} f_{v_i}\left(x^i\right)$$

і перейти на крок IX; інакше обчислити вектор $h^k$ за формулою

$$h^k = \sum_{i=0}^{k} \hat{\nabla} f_{v_i}\left(x^i\right)$$

і перейти на крок IX.

IX. Знайти наступне наближення:

$$x^{k+1} = \pi_{X_2}\left(x^k - \rho_k h^k\right),$$

де $\pi_{X_2}\left(x\right)$ – оператор проектування на множину $X_2$ точки $x \in R^n$.

X. Покласти $k = k+1$ і перейти на крок IV.

***Теорема 1.*** *Якщо виконуються припущення 1 і умови:*

$$\rho_k \geq 0,\ k = 0,1,...;\ \sum_{k=0}^{\infty} \rho_k = \infty;\ \rho_k \to 0 \text{ при } k \to \infty,$$

*то при довільному початковому наближені $x^0 \in R^n$ знайдеться підпослідовність $\left\{x^{k_i}\right\}_{i=0}^{\infty}$ послідовності $\left\{x^k\right\}_{k=0}^{\infty}$, яка породжена алгоритмом 1, така, що*

$$\lim_{i\to\infty} f_0(x^{k_i}) = \min_{x\in X} f_0\left(x\right).$$

*Якщо, крім цього, функція $f_0$ – строго опукла донизу, то:*

$$\lim_{k\to\infty} x^k = x^*,$$

*де* $f_0\left(x^*\right) = \min_{x\in X} f_0\left(x\right),\ \ x^* \in X$.

***Алгоритм 1′***

Початок. I. Вибрати довільне початкове наближення $x^0 \in R^n$.

II. Вибрати довільне натуральне число $l$.

III. Покласти $k = 0$.

Основний цикл. IV. Знайти кроковий множник $\rho_k$ і параметри $\delta_k$, $\alpha_k$, які задовольняють умови теореми 1′.

V. Якщо виконується нерівність

$$\max_{j\in[1;m]} f_j\left(x^k\right) \leq 0,$$

то покласти $v_k = 0$ і перейти на крок VII; інакше перейти на крок VI.

VI. Визначити індекс $i \in \left[1:m\right]$, який задовольняє умову

$$f_i\left(x^k\right) = \max_{j\in[1;m]} f_j\left(x^k\right),$$

покласти $v_k = i$ та перейти на крок VII.

VII. Знайти реалізацію $\tilde{x}^k$ випадкової точки, рівномірно розподіленої в

$n$ -вимірному кубі з центром в точці $x^k$ і стороною $\alpha_k$.

VIII. Обчислити вектор:

$$\theta^k = \sum_{i=1}^{n} \frac{f_{v_k}(\tilde{x}^k + \delta_k e^i) - f_{v_k}(\tilde{x}^k)}{\delta_k} e^i,$$

де $e^i$, $i = \overline{1,n}$ – $i$ -й орт.

IX. Якщо $k \ge l$, то обчислити вектор

$$h^k = \sum_{i=k-l}^{k} \theta^i$$

і перейти на крок X; інакше обчислити вектор $h^k$ за формулою

$$h^k = \sum_{i=0}^{k} \theta^i$$

і перейти на крок X.

X. Знайти наступне наближення

$$x^{k+1} = \pi_{X_2}\left(x^k - \rho_k h^k\right).$$

XI. Покласти $k = k+1$ і перейти на крок IV.

***Теорема 1'.*** *Якщо виконуються припущення 1 і умови:*

$$\rho_k \ge 0,\ k = 0,1,...;\ \sum_{k=0}^{\infty} \rho_k = \infty,\ \sum_{k=0}^{\infty} \rho_k^2 < \infty;$$

$$\alpha_k \ge 0,\ k = 0,1,...;\ \delta_k / \alpha_k \to 0 \ \textit{при}\ k \to \infty,\ \alpha_k \to 0\ \textit{при}\ k \to \infty,$$

*то з ймовірністю 1 існує підпослідовність* $\left\{x^{k_i}\right\}_{i=0}^{\infty}$ *послідовності* $\left\{x^k\right\}_{k=0}^{\infty}$, *яка породжена алгоритмом 1, така, що*

$$\lim_{i\to\infty} f_0(x^{k_i}) = \min_{x\in X} f_0(x).$$

*Якщо, крім цього, функція* $f_0$ *– строго опукла донизу, то з ймовірністю 1*

$$\lim_{k\to\infty} x^k = x^*,$$

*де* $f_0\left(x^*\right) = \min_{x\in X} f_0\left(x\right),\ x^* \in X$.

**2. Стохастична задача**

З а д а ч а 2. Знайти

$$\arg\min_{x\in X} EF_0\left(x,\omega\right) \quad \left( f_0(x) = EF_0\left(x,\omega\right) \triangleq \int_{\Omega} F_0\left(x,\omega\right)\mu\left(d\omega\right) \right)$$

для заданої функції $F_0 : X \times \Omega \to R^1$ і заданої множини $X \subset R^n$.

*Припущення 2.* $(\boldsymbol{i})$ – функція $f_0$ – неперервно диференційована і опукла донизу в області $X$; $(\boldsymbol{ii})$ – $X$ – опукла і замкнута множина.

У наведеному нижче непошуковому алгоритмі адаптивного керування на $k$-й ітерації $\left(k \geq p_0\right)$ робиться тільки одне вимірювання (обчислення) функції $F_0$. Для знаходження напрямку руху $\xi^k$ до наступного наближення $x^{k+1}$ використовується відома інформація із $p_k$ $\left(p_k \geq 1\right)$ попередніх ітерацій:

$$\xi^k = \sum_{s=k-p_k+1}^{k} \frac{1}{\Delta_s}\Big(F_0(x^s + \Delta_s\theta^s, \omega^s) - F_0(x^{s-1} + \Delta_{s-1}\theta^{s-1}, \omega^{s-1})\Big)\theta^s \, ,$$

де $\Delta_s > 0$; $\left\{\theta^s\right\}_{s=k-p_k}^{k}$ – серія незалежних спостережень випадкового вектора $\theta$ з незалежними і рівномірно розподіленими на $\left[-1;1\right]$ компонентами.

***Алгоритм 2***

Початок. I. Вибрати початкове значення параметра $p_0$, яке задовольняє умовам теореми 2.

II. Знайти: початкові точки $x^k$, $k = \overline{0, p_0}$, які належать множині $X$; вектори $\theta^k$, $k = \overline{0, p_0}$ – серію незалежних спостережень випадкового вектора $\theta$ з незалежними і рівномірно розподіленими на відрізку $\left[-1;1\right]$ компонентами; реалізації $\omega^k$, $k = \overline{0, p_0}$ випадкової величини $\omega$.

III. Задати зміщення $\Delta_k$, $k = \overline{0, p_0}$, які задовольняють умовам теореми 2.

IV. Покласти $k = p_0$.

Основний цикл. V. Обчислити параметр $p_k$ і вектор

$$\xi^k = \sum_{i=k-p_k+1}^{k} \frac{1}{\Delta_i}\Big(F_0(x^i + \Delta_i\theta^i, \omega^i) - F_0(x^{i-1} + \Delta_{i-1}\theta^{i-1}, \omega^{i-1})\Big)\theta^i.$$

VI. Обчислити кроковий множник $\rho_k$ і нормуючий множник $\gamma_k$, які задовольняють умовам теореми 2.

VII. Обчислити наступне наближення

$$x^{k+1} = \pi_X(x^k - \rho_k\gamma_k\xi^k).$$

VIII. Знайти незалежне спостереження $\theta^{k+1}$ випадкового вектора $\theta$ з незалежними і рівномірно розподіленими на відрізку $\left[-1;1\right]$ компонентами.

IX. Знайти реалізацію $\omega^{k+1}$ випадкової величини $\omega$.

X. Знайти зміщення $\Delta_{k+1}$, яке задовольняє умовам теореми 2.

XI. Покласти $k = k+1$ і перейти на крок V.

***Теорема 2.*** *Нехай виконуються припущення 2 і мають місце умови:* $(\boldsymbol{i})$ *– градієнт функції* $f_0$ *задовольняє глобальну умову Ліпшиця, тобто при* $x$, $y \in R^n$

$$\left\|\nabla f_0\left(x\right)-\nabla f_0\left(y\right)\right\| \le \alpha_1 \left\|x-y\right\|, \ \alpha_1 < \infty;$$

$(\boldsymbol{ii})-$ *відома величина* $\beta_k$ *така, що:*

$$E\left(\sum_{i=k-p_k+1}^{k}\left\|\left(F_0(x^i+\Delta_i\theta^i,\omega^i)-\right.\right.\right.$$

$$\left.\left.\left.-F_0(x^{i-1}+\Delta_{i-1}\theta^{i-1},\omega^{i-1})\right)\theta^i\right\|^2 / x^{k-p_k+1},\dots,x^k\right) \le \beta_k^2 \le \delta(\tau) < \infty$$

*при* $\| x^s \| \le \tau < \infty$, $s=\overline{k-p_k,k}$;

$(\boldsymbol{iii})-$ *нормуючий множник* $\gamma_k$ *задовольняє нерівностям:*

$$0<\gamma' \le \gamma_k\left(\| x^k \| + \beta_k\right) \le \gamma'' < \infty;$$

$(\boldsymbol{iv})-$ *величини* $\Delta_k$, $p_k$ *і* $\rho_k$ *такі, що:*

$$1 \le p_k \le \alpha_2 < \infty, \ k-p_k \ge 0, \ \rho_k \ge 0, \ \Delta_k > 0;$$

$$\rho_k \to 0, \ \Delta_k \to 0 \ \ при \ \ k \to \infty;$$

$$\sum_{k=p_0}^{\infty}\rho_k\Delta_{k-p_k+1} < \infty, \ \sum_{k=0}^{\infty}\left(\rho_k/\Delta_k\right)^2 < \infty.$$

*Тоді послідовність випадкових точок* $\left\{x^k\left(\omega\right)\right\}_{k=0}^{\infty}$, *яка породжена алгоритмом 2, є випадковою квазіфейеровською відносно множини розв'язків* $X^*$ *задачі 2.*

*Якщо, крім цього, виконується умова* $(\boldsymbol{v})-\ \sum_{k=0}^{\infty}\rho_k=\infty$, *то майже для кожного* $\omega$ *послідовність* $\left\{x^k(\omega)\right\}_{k=0}^{\infty}$ *збігається до розв'язку задачі 2.*

## 6.22. Прямий метод розв'язування задач стохастичного програмування

Задача 1. Знайти $\arg\min\limits_{x\in X} E_\omega f_0\left(x,\omega\right)$ для заданої функції $f_0: X\times\Omega\to R^1$ і заданої множини

$$X=\left\{x \mid E_\omega f_i\left(x,\omega\right)\le 0, \ i=\overline{1,m}, \ \ x\in X^1\right\},$$

де $X^1$ – деяка множина простору $R^n$.

*Припущення 1.* $(\boldsymbol{i})-X^1-$ опукла, замкнута і обмежена множина; $(\boldsymbol{ii})-$ $E_\omega f_i\left(x,\omega\right)$, $i=\overline{0,m}$ – неперервні, опуклі донизу функції.

***Алгоритм 1***

Початок. I. Вибрати довільне початкове наближення $x^0\in R^n$;

довільні числа $z_i^0$, $i=\overline{1,m}$; достатню велику константу $\delta>0$.

II. Покласти $k=0$.

Основний цикл. III. Знайти кроковий множник $\rho_k$ і множник $\alpha_k$, які задовольняють умови теореми 1.

IV. Якщо виконується нерівність

$$\max_{i\in[1;m]} z_i^k \le 0,$$

то покласти $j=0$ і перейти на крок VI; інакше перейти на крок V.

V. Знайти індекс $j\in[1:m]$, який задовольняє умову

$$z_j^k = \max_{i\in[1;m]} z_i^k.$$

VI. Обчислити реалізацію $\xi^k$ випадкового вектора $\overline{\xi}^k$, для якого:

$$E_\omega\left(\overline{\xi}^k / x^0, z^0, ..., x^k, z^k\right) = \hat{\nabla}\varphi_j\left(x^k\right),$$

де $\hat{\nabla}\varphi_j\left(x^k\right)$ – узагальнений градієнт функції $\varphi_j(x) \triangleq E_\omega f_j(x,\omega)$ в точці $x=x^k$.

VII. Обчислити наступне наближення

$$x^{k+1} = \pi_{X^1}\left(x^k - \rho_k \xi^k\right).$$

VIII. Знайти випадкові величини $\theta_i^k$, $i=\overline{1,m}$, для яких:

$$E_\omega\left(\theta_i^k / x^0, z^0, ..., x^k, z^k\right) = E_\omega f_i\left(x^k, \omega\right),\ \ i=\overline{1,m}.$$

IX. Покласти $i=1$.

X. Якщо виконується нерівність

$$\left|z_i^k + \alpha_k\left(\theta_i^k - z_i^k\right)\right| \le \delta,$$

то покласти

$$z_i^{k+1} = z_i^k + \alpha_k\left(\theta_i^k - z_i^k\right)$$

і перейти на крок XI; інакше покласти

$$z_i^{k+1} = \left(\delta z_i^k + \delta\alpha_k\left(\theta_i^k - z_i^k\right)\right) \Big/ \left|z_i^k + \alpha_k\left(\theta_i^k - z_i^k\right)\right|$$

і перейти на крок XI.

XI. Якщо $i<m$, то покласти $i=i+1$ і перейти на крок X; інакше перейти на крок XII.

XII. Покласти $k=k+1$ і перейти на крок III.

***Теорема 1.*** *Якщо виконуються припущення 1 і крім цього:* (***iii***)*– область $Y$, яка визначена обмеженнями:*

$$E_\omega f_i(x,\omega) \le 0,\ \ i=\overline{1,m},$$

*задовольняє умову Слейтера;*

$(\boldsymbol{iv})-\ \sum_{k=0}^{\infty}\rho_k=\infty,\ \sum_{k=0}^{\infty}{\rho_k}^2<\infty\ \ \rho_k/\alpha_k\to 0\ \ при\ \ k\to\infty;$

$(\boldsymbol{v})-\ \ E\left|\theta_i^k\right|\le\delta<\infty,\ \ i=\overline{1,m}\ \ при\ \ k=0,1,...;$

$(\boldsymbol{vi})-\ \ E\,\|\,\overline{\xi}^k\,\|^2\le\delta<\infty,\ при\ \ k=0,1,...;$

*то з ймовірністю 1, одна з граничних точок послідовності* $\left\{x^k\right\}_{k=0}^{\infty}$, *яка породжена алгоритмом 1, належить множині розв'язків задачі 1. Якщо* $Ef_0\left(x,\omega\right)$ *– строго опукла донизу функція, то з ймовірністю 1*

$$\lim_{k\to\infty}x^k=x^*,$$

*де* $x^*$ *– розв'язок задачі 1.*

*Зауваження 1.* Якщо функції $f_i\left(x,\omega\right),\ i=\overline{0,m},$ при кожному $\omega$ опуклі донизу, то на кроці VI алгоритму 1 можна взяти

$$\xi^k=\hat{\nabla}f_j\left(x^k,\omega^k\right),$$

а на кроці VIII:

$$\theta_i^k=f_i\left(x^k,\omega^k\right),\ \ i=\overline{1,m},$$

де $\hat{\nabla}f_j\left(x^k,\omega^k\right)$ – узагальнений градієнт функції $f_j\left(x,\omega\right)$ по $x$; $\omega^k$ – незалежні спостереження випадкової величини $\omega$.



## 6.23. Метод випадкового пошуку в опуклих задачах мінімізації

З а д а ч а 1. Знайти $\arg\min_{x\in X}f_0\left(x\right)$ для заданої функції $f_0:R^n\to R^1$ і множини $X\subset R^n$.

*Припущення 1.* $(\boldsymbol{i})-$ функція $f_0\left(x\right)$ – опукла донизу і неперервно диференційована на множині $X$; $(\boldsymbol{ii})-$ градієнт функції $f_0\left(x\right)$ задовольняє на $X$ умову Ліпшиця із константою $0<\gamma<\infty$, тобто

$$\left\|\nabla f_0\left(x\right)-\nabla f_0\left(y\right)\right\|\le\gamma\left\|x-y\right\|,\ \ \forall x,y\in X;$$

$(\boldsymbol{iii})-$ множина $X$ – опукла і замкнена.

В методі випадкового пошуку на $k$-й ітерації по відомому наближенню $x^k\in X$ обчислюється наступне наближення $x^{k+1}$, як точка в деякому сенсі близька до точки мінімуму функції $f_0\left(\cdot\right)$ на відрізку прямої $x^k-\rho h^k$ $\left(\rho\in\left(-\infty;\infty\right)\right)$, який належить множині $X$, де $h^k$ – незалежна реалізація одиничного випадкового вектора, рівномірно розподіленого на одиничній сфері з центром в початку координат. Для знаходження крокового множника $\rho_k$ необхідно застосувати скінченну кількість ітерацій деякого алгоритму одновимірної оптимізації.

***Алгоритм 1***

Початок. I. Вибрати довільне початкове наближення $x^0 \in X$.

II. Покласти $k = 0$.

Основний цикл. III. Якщо $\nabla f_0\left(x^k\right) = 0$, то покласти $x^* = x^k$ і завершити обчислення; інакше перейти на крок IV.

IV. Знайти незалежну реалізацію $h^k$ одиничного випадкового вектора $\xi$ рівномірно розподіленого на одиничній $n$-вимірній сфері з центром в початку координат.

V. Якщо множина

$$L_k = \left\{\rho \middle| x^k - \rho h^k \in X, \quad \rho \in (-\infty; \infty)\right\}$$

містить хоча б одну точку $\rho \neq 0$, то перейти на крок VI; інакше покласти $x^{k+1} = x^k$ і перейти на крок VIII.

VI. Знайти кроковий множник $\rho_k \in L_k$, що задовольняє умову:

$$f_0\left(x^k - \rho_k h^k\right) \leq \left(1 - \lambda_k\right) f_0\left(x^k\right) + \lambda_k w_k,$$

де $\lambda_k$ – довільна точка фіксованого відрізку $[\lambda; 1]$; $\lambda$ – довільна константа з півінтервалу $(0; 1]$; $w_k = \min\limits_{\rho \in L_k} f_0\left(x^k - \rho h^k\right)$.

VII. Знайти наступне наближення

$$x^{k+1} = x^k - \rho_k h^k.$$

VIII. Покласти $k = k + 1$ і перейти на крок III.

***Теорема 1.*** *Нехай виконуються припущення 1 і* (***iv***)*− множина* $X$ *регулярна, тобто в* $R^n$ *існує сфера радіусу* $\chi$*, яка повністю належить множині* $X$; (***v***)*− для довільного елемента* $x^0 \in X$ *множина*

$$X_0 = \left\{x \middle| f_0(x) \leq f_0\left(x^0\right), \ x \in X\right\}$$

*обмежена. Тоді нескінченна послідовність* $\left\{x^k\right\}_{k=0}^{\infty}$*, яка породжена алгоритмом 1, така, що*

$$P\left\{f_0\left(x^k\right) \to f_0(x^*), \ k \to \infty\right\} = 1.$$

## 6.24. Методи розв'язування задач оптимізації з нескінченим числом обмежень

Задача 0. Знайти $\arg\min\limits_{x \in X} f_0(x)$ для заданої функції $f_0 : R^n \to R^1$ і заданої компактної множини $X \subset R^n$ при наявності додаткових обмежень

$$f_1(x, y) \leq 0 \text{ для всіх } y \in Y,$$

де $f_1 : R^n \times R^m \to R^1$, $Y \subset R^m$.

Багато задач оптимізації може бути представлено у формі задачі 0. Наприклад, якщо в мінімаксній задачі вигляду: знайти $v^*$,

$$v^* = \min_{x \in X} \max_{y \in Y} \varphi(x, y)$$

покласти $v_1 < v^*$, $v_2 > v^*$ і $X_0 = [v_1; v_2]$, то її можна представити у формі задачі 0 наступним чином: знайти $\min\limits_{(x_0, x) \in X_0 \times X} x_0$ при обмежені $\varphi(x, y) - x_0 \le 0$ для всіх $y \in Y$.

*Припущення 0.* $(\boldsymbol{i})-$ функції $f_0$, $f_1$ є неперервними, відповідно, на множинах $X$ і $X \times Y$; $(\boldsymbol{ii})-$ $X$ і $Y$ – непорожні компактні множини.

На $k$-й ітерації знаходять наближення $x^k$ як розв'язок апроксимуючої задачі:

$$\text{знайти } \arg\min_{x \in G_k} f_0(x),$$

де

$$G_k = \{x | f_1(x, y) \le 0 \text{ для всіх } y \in Y_k, \ x \in X\}.$$

Скінчені множини $Y_k$, $k = 0,1,\ldots$ апроксимують множину $Y$ та їх знаходять за допомогою розв'язування задачі максимізації по $y \in Y$ функції $f_1(x^k, y)$.



### 1. Загальний метод

***Алгоритм 1***

П о ч а т о к . І. Вибрати довільну скінченну підмножину $Y_0 \subset Y$.

ІІ. Покласти $k = 0$.

О с н о в н и й ц и к л . ІІІ. Знайти допустиму множину $G_k$ $k$-ї апроксимуючої задачі за правилом

$$G_k = \{x | f_1(x, y) \le 0 \text{ для всіх } y \in Y_k, \ x \in X\}.$$

IV. Знайти розв'язок $x^k \in G_k$ наступної $k$-ї апроксимуючої задачі:

$$\text{знайти } \arg\min_{x \in G_k} f_0(x).$$

(Відмітимо, що виконуються включення:

$$G \subset \ldots \subset G_{k+1} \subset G_k \subset \ldots \subset G_0, \tag{6.95}$$

де $G \triangleq \left\{x \middle| f_1(x, y) \le 0 \text{ для всіх } y \in Y, \ x \in X\right\}$, і справедливі нерівності:

$$f_0(x^0) \le f_0(x^1) \le \ldots \le f_0(x^k) \le \ldots \le f_0(x^*), \tag{6.96}$$

де $x^*-$ розв'язок задачі 0).

V. Знайти розв'язок $y^{k+1} \in Y$ наступної $k$ -ї допоміжної задачі:

$$\text{знайти } \arg\max_{y\in Y} f_1(x^k, y).$$

VI. Якщо виконується нерівність

$$f_1(x^k, y^{k+1}) \le 0,$$

то покласти $x^* = x^k$ і зупинити обчислення (розв'язок задачі 0 знаходиться в цьому випадку за скінченну кількість ітерацій); інакше покласти $Y_{k+1} = Y_k \bigcup \{y^{k+1}\}$ і перейти на крок VII.

VII. Покласти $k = k+1$ і перейти на крок III.

***Теорема 1.*** *Нехай виконуються припущення 0. Якщо $x^*$ є граничною точкою нескінченної послідовності $\{x^k\}_{k=0}^{\infty}$, яка породжена алгоритмом 1, то $x^*$ є розв'язком задачі 0.*

***Теорема 1'.*** *Нехай виконуються припущення 0 та: (**iii**) – функція $f_1(x, y)$ – опукла догори по $y$ при всіх $x \in X$; (**iv**) – множина $Y$ є багатокутником. Тоді алгоритм 1 збігається до розв'язку $x^*$ задачі 0 за скінченну кількість ітерацій.*

### 2. Послаблений метод



У послабленому методі на кожній ітерації відкидаються всі незв'язні обмеження. Збіжність послабленого алгоритму має місце при деяких додаткових припущеннях.

***Алгоритм 2***

I–V (кроки I–V такі ж самі, як в алгоритмі 1).

Якщо виконуються тільки припущення 0, то співвідношення (6.95) і (6.96) для алгоритму 2 загалом не мають місця.

VI. Якщо виконується нерівність

$$f_1(x^k, y^{k+1}) \le 0,$$

то покласти $x^* = x^k$ і зупинити обчислення (в цьому випадку розв'язок задачі 0 знаходять за скінченну кількість ітерацій); інакше перейти на крок VII.

VII. Покласти

$$Z_k = \{y \mid f_1(x^k, y) < 0,\ \ y \in Y_k\}.$$

VIII. Покласти

$$Y_{k+1} = Y_k \bigcup \{y^{k+1}\} \setminus Z_k.$$

IX. Покласти $k = k+1$ і перейти на крок III.

***Теорема 2.*** *Нехай виконуються припущення 0. Тоді:*

1) *якщо нескінченна послідовність $\left\{x^k\right\}_{k=0}^{\infty}$, яка породжена алгоритмом 2, збігається до деякої точки $x^*$, то $x^*$ є розв'язком задачі 0;*

2) *якщо функції $f_0(x)$ і $f_1(x, y)$ опуклі донизу по $x$ (при кожному*

*$y \in Y$), а $X$ – опукла множина, то послідовність $\left\{ f_0(x^k) \right\}_{k=0}^{\infty}$ – монотонно неспадна;*

3) *якщо функція $f_0(x)$ – строго опукла донизу, кожна з функцій $f_1(x, y)$ – опукла донизу по $x$ для $y \in Y$, $X$ – опукла множина, то послідовність $\left\{ x^k \right\}_{k=0}^{\infty}$ збігається до розв'язку $x^*$ задачі 0.*

## 6.25. Методи розв'язування задач квадратичного програмування

В цьому підрозділі наводяться деякі методи розв'язування задачі квадратичного програмування, яка є допоміжною задачею для багатьох алгоритмів. ***Задачею квадратичного програмування*** називають задачу мінімізації квадратичної функції при лінійних обмеженнях.

### 1. Метод спряжених градієнтів для мінімізації квадратичної функції на підпросторі



З а д а ч а 1. Знайти $\arg\min\limits_{x \in X} \left[ \dfrac{1}{2}(Cx, x) + (d, x) \right]$,

де $X = \left\{ x \,|\, (a^i, x) - b_i = 0, \ \ i \in \Im \right\}$;

$C -$ симетрична додатньо визначена матриця розмірності $n \times n$; $d \in R^n$, $a^i \in R^n$, $i \in \Im$, $\Im -$ скінченна множина індексів; $b_i \in R^1$, $i \in \Im$.

*Припущення 1.* $(\boldsymbol{i}) -$ $C -$ симетрична додатньо визначена матриця розмірності $n \times n$, тобто $(Cx, x) > 0$ для всіх ненульових $x \in R^n$; $(\boldsymbol{ii}) -$ вектори $a^i$, $i \in \Im -$ лінійно незалежні.

***Алгоритм 1***

П о ч а т о к . I. Знайти початкове наближення $x^0 \in X$.

II. Обчислити $n \times n$-матрицю $H$ за формулою:

$$H = A^T (AA^T)^{-1} A,$$

де $A$ – матриця, рядками якої є $n$-вимірні вектори-рядки $(a^i)^T$, $i \in \Im$.

III. Обчислити градієнт цільової функції в точці $x = x^0$ за формулою

$$\nabla f_0(x^0) = Cx^0 + d.$$

IV. Обчислити $n$-вимірний вектор:

$$h^1 = -(I - H) \nabla f_0(x^0),$$

де $I -$ одинична $n \times n$-матриця.

Якщо $h^1 = 0$, то зупинити обчислення; інакше перейти на крок V.

V. Обчислити кроковий множник

$$\rho_1 = -\left(\nabla f_0(x^0), h^1\right) \Big/ \left(h^1, Ch^1\right).$$

VI. Покласти $x^1 = x^0 + \rho_1 h^1$.

VII. Покласти $k = 1$, $g^0 = h^1$.

Основний цикл. VIII. Обчислити $n$-вимірний вектор:

$$g^k = (I - H)\nabla f_0(x^k),$$

де $\nabla f_0(x^k) = Cx^k + d$.

Якщо $g^k = 0$, то зупинити обчислення; інакше перейти на крок IX.

IX. Обчислити $n$-вимірний вектор

$$h^{k+1} = -g^k + \left(\left\|g^k\right\|^2 \Big/ \left\|g^{k-1}\right\|^2\right) h^k.$$

X. Обчислити кроковий множник

$$\rho_{k+1} = -\left(\nabla f_0(x^k), h^{k+1}\right) \Big/ \left(h^{k+1}, Ch^{k+1}\right).$$

XI. Обчислити $x^{k+1} = x^k + \rho_{k+1} h^{k+1}$.

XII. Покласти $k = k + 1$ і перейти на крок VIII.

***Теорема 1.*** *Якщо виконуються припущення 1, то алгоритм 1 розв'язує задачу 1 за скінченну кількість ітерацій, що не перевищує n. При цьому, якщо цільова функція* $f_0(x) \triangleq (1/2)(Cx, x) + (d, x)$ *має на множині* $X$ *скінченний мінімум, то послідовність* $x^0, x^1, \ldots,$ *яка породжена алгоритмом 1, збігається до точки мінімуму* $x^*$ *задачі 1 за скінченну кількість ітерацій* $k \le n$. *Якщо функція* $f_0(x)$ *– необмежена (знизу) на множині* $X$, *то при деякому* $k \le n$ *буде виконуватися рівність*

$$(h^{k+1}, Ch^{k+1}) = 0,$$

*з чого випливає* $\rho_{k+1} = -\infty$.

**Приклад 1.** Використовуючи алгоритм 1, розв'язати задачу квадратичного програмування:

$$f_0(x) = 2x_1^2 + 2x_1 x_2 + 2x_2^2 - 4x_1 - 6x_2 \to \min$$

при обмеженні $x_1 + 2x_2 = 2$.

*Розв'язування.*

Спочатку знайдемо матрицю $C$ і вектор $d$.

$$f_0(x) = \frac{1}{2}(Cx, x) + (d, x) =$$

$$= \frac{1}{2}\left(\begin{pmatrix} c_{11} & c_{12} \\ c_{12} & c_{22} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) + \left(\begin{pmatrix} d_1 \\ d_2 \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) =$$

$$= \frac{1}{2}\left(\begin{pmatrix} c_{11}x_1 + c_{12}x_2 \\ c_{12}x_1 + c_{22}x_2 \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) + d_1 x_1 + d_2 x_2 = 2x_1^2 + 2x_1 x_2 + 2x_2^2 - 4x_1 - 6x_2.$$

Отримуємо систему, розв'язавши яку знайдемо $C$ і $d$:

$$\begin{cases} (1/2)c_{11} = 2 \\ c_{12} = 2 \\ (1/2)c_{22} = 2 \\ d_1 = -4 \\ d_2 = -6 \end{cases} \Leftrightarrow \begin{cases} c_{11} = 4 \\ c_{12} = 2 \\ c_{22} = 4 \\ d_1 = -4 \\ d_2 = -6, \end{cases} \qquad C = \begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix},\ d = \begin{pmatrix} -4 \\ -6 \end{pmatrix}.$$

Тепер з обмеження знаходимо вектор $a$ і матрицю $A$: $a = \begin{pmatrix} 1 \\ 2 \end{pmatrix},\ A = a^T = (1; 2).$

**Алгоритм 1**

I. Задаємо початковий допустимий розв'язок $x^0 = \begin{pmatrix} 1 \\ 0,5 \end{pmatrix} \in X.$

II. Обчислюємо матрицю $H$,

$$H = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \left( (1 \quad 2) \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right)^{-1} (1 \quad 2) = \begin{pmatrix} 1 \\ 2 \end{pmatrix} (5)^{-1} (1 \quad 2) = \frac{1}{5} \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} 1/5 & 2/5 \\ 2/5 & 4/5 \end{pmatrix}.$$

III. Обчислюємо градієнт квадратичної функції в точці $x^0$

$$\nabla f_0(x^0) = \begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 0,5 \end{pmatrix} + \begin{pmatrix} -4 \\ -6 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}.$$

IV. Обчислюємо вектор

$$h^1 = (H - I)\nabla f_0(x^0) = \begin{pmatrix} 1/5 - 1 & 2/5 \\ 2/5 & 4/5 - 1 \end{pmatrix} \begin{pmatrix} 1 \\ -2 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} -8 \\ 4 \end{pmatrix}.$$

V. Послідовно обчислюємо кроковий множник $\rho_1$:

$$\left( \nabla f_0(x^0), h^1 \right) = \left( \begin{pmatrix} 1 \\ -2 \end{pmatrix}, \begin{pmatrix} -8 \\ 4 \end{pmatrix} \right) \frac{1}{5} = \frac{1}{5}(-8 \quad -8) = -\frac{16}{5};$$

$$Ch^1 = \begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} -8 \\ 4 \end{pmatrix} \frac{1}{5} = \frac{1}{5} \begin{pmatrix} -32 + 8 \\ -16 + 16 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} -24 \\ 0 \end{pmatrix};$$

$$(h^1, Ch^1) = \left( \begin{pmatrix} -8 \\ 4 \end{pmatrix}, \begin{pmatrix} -24 \\ 0 \end{pmatrix} \right) \frac{1}{25} = \frac{1}{25} \cdot 8 \cdot 24 = \frac{192}{25}; \qquad \rho_1 = -\frac{-16 \cdot 25}{5 \cdot 192} = \frac{5}{12}.$$

VI. Обчислюємо наступне наближення

$$x^1 = \begin{pmatrix} 1 \\ 0,5 \end{pmatrix} + \frac{5}{12} \cdot \frac{1}{5} \begin{pmatrix} -8 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 \\ 0,5 \end{pmatrix} + \begin{pmatrix} -2/3 \\ 1/3 \end{pmatrix} = \begin{pmatrix} 1/3 \\ 5/6 \end{pmatrix} = \frac{1}{6} \begin{pmatrix} 2 \\ 5 \end{pmatrix}.$$

VII. Покладемо $k = 1,\ g^0 = h^1 = \frac{1}{5} \cdot \begin{pmatrix} -8 \\ 4 \end{pmatrix}.$

*1-а ітерація:*

VIII. Обчислюємо вектор $g^1$ за формулою

$$g^1 = (I - H)\nabla f_0(x^1).$$

Спочатку обчислимо градієнт цільової функції:

$$\nabla f_0(x^1) = Cx^1 + d = \begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix}\begin{pmatrix} 2 \\ 5 \end{pmatrix}\frac{1}{6} + \begin{pmatrix} -4 \\ -6 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \end{pmatrix} + \begin{pmatrix} -4 \\ -6 \end{pmatrix} = \begin{pmatrix} -1 \\ -2 \end{pmatrix}.$$

$$g^1 = \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 1/5 & 2/5 \\ 2/5 & 4/5 \end{pmatrix}\right)\begin{pmatrix} -1 \\ -2 \end{pmatrix} =$$

$$= \begin{pmatrix} 4/5 & -2/5 \\ -2/5 & 1/5 \end{pmatrix}\begin{pmatrix} -1 \\ -2 \end{pmatrix} = \begin{pmatrix} -4/5 + 4/5 \\ 2/5 - 2/5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Оскільки $g^1 = 0$, то покладемо $x^* = x^1 = \begin{pmatrix} 1/3 \\ 5/6 \end{pmatrix}$ і зупиняємо обчислення.

Таким чином, знайшли оптимальний розв'язок задачі квадратичного програмування $x_1^* = 1/3$ і $x_2^* = 5/6$ і оптимальне значення цільової функції

$$f_0(x^*) = 2\cdot\frac{1}{9} + 2\cdot\frac{5}{18} + 2\cdot\frac{25}{36} - 4\cdot\frac{1}{3} - 6\cdot\frac{5}{6} = \frac{2}{9} + \frac{5}{9} + \frac{25}{18} - \frac{4}{3} - 5 = -4\frac{1}{6}.$$

**2. Метод спряжених градієнтів розв'язування для загальної задачі квадратичного програмування**

Задача 2. Знайти $\arg\min\limits_{x\in X}\left[\frac{1}{2}(Cx, x) + (d, x)\right]$,

де $X = \{x \mid (a^i, x) - b_i \le 0,\ \ i \in \Im^-;\ \ (a^i, x) - b_i = 0,\ \ i \in \Im^0\}$;

$C-$ симетрична додатньо визначена матриця $n\times n$; $d \in R^n$, $\Im^-$ і $\Im^0$ – скінченні множини індексів; $a^i \in R^n$, $i \in \Im^0 \bigcup \Im^-$; $b_i \in R^1$, $i \in \Im^0 \bigcup \Im^-$.

*Припущення 2.* $(\boldsymbol{i})-$ $C-$ симетрична додатньо визначена матриця; $(\boldsymbol{ii})-$ виконується наступна умова невиродженості: при будь-якому $x \in R^n$ вектори $a^i$, $i \in \Im(x)-$ лінійно незалежні, де

$$\Im(x) \triangleq \left\{i \mid (a^i, x) - b_i = 0,\ \ i \in \Im^- \bigcup \Im^0\right\}. \qquad (6.97)$$

***Алгоритм 2***

Початок. І. Знайти початкове наближення $x^0 \in X$.

ІІ. Знайти множину індексів $\Im(x^0)$ і покласти $\Im_0 = \Im(x^0)$ (тут множина $\Im(x)$ визначається згідно (6.97)).

ІІІ. Обчислити $n\times n$-матрицю:

$$H_{\Im_0} = A_{\Im_0}^T (A_{\Im_0} A_{\Im_0}^T)^{-1} A_{\Im_0},$$

де $A_{\Im_0}$ – матриця, рядками якої виступають $n$-вимірні вектор-рядки $(a^i)^T$, $i \in \Im_0$.

IV. Обчислити вектор:

$$u^0 = -(A_{\Im_0} A_{\Im_0}^T)^{-1} A_{\Im_0} \nabla f_0(x^0),$$

де $\nabla f_0(x^0) = Cx^0 + d$.

V. Обчислити $n$-вимірний вектор

$$g^0 = (I - H_{\Im_0}) \nabla f_0(x^0) = \nabla f_0(x^0) + A_{\Im_0}^T u^0. \qquad (6.98)$$

Якщо $g^0 = 0$, то перейти на крок VI; інакше перейти на крок XXII.

(У випадку, коли $g^0 = 0$, точка $x^0$ є точкою мінімуму функції $f_0(x) \triangleq (1/2)(Cx, x) + (d, x)$ на грані, яка визначається системою рівнянь $(a^i, x) - b_i = 0$, $i \in \Im_0$).

VI. Визначити компоненти вектора $u^0$, які в рівностях (6.98) множаться на вектори $a^i$, $i \in \Im^- \bigcap \Im_0$. Нехай вони будуть $u_{j_1}^0, u_{j_2}^0, \ldots, u_{j_{m_0}}^0$

VII. Якщо $u_{j_1}^0 \geq 0, \ldots, u_{j_{m_0}}^0 \geq 0$, то покласти $x^* = x^0$ і зупинити обчислення (в цьому випадку $x^*$ – розв'язок задачі 2); інакше перейти на крок VIII.

VIII. Знайти індекс $j_{m'} \in \{j_1, \ldots, j_{m_0}\}$, для якого виконується нерівність $u_{j_{m'}}^0 < 0$.

IX. Знайти індекс $i' \in \Im_0 \bigcap \Im^-$ такий, що вектор $a^{i'}$ в (6.98) множиться на компоненту $u_{j_{m'}}^0$.

X. Утворити нову індексну множину $\Im_0{}'$, яка утворюється з множини $\Im_0$ відкиданням індексу $i'$.

XI. Утворити матрицю $A_{\Im_0{}'}$, рядками якої виступають $n$-вимірні вектор-рядки $(a^i)^T$, $i \in \Im_0{}'$.

XII. Обчислити $n \times n$-матрицю

$$H_{\Im_0{}'} = A_{\Im_0{}'}^T (A_{\Im_0{}'} A_{\Im_0{}'}^T)^{-1} A_{\Im_0{}'}.$$

XIII. Обчислити вектор

$$h^1 = -(I - H_{\Im_0{}'}) \nabla f_0(x^0).$$

XIV. Покласти $k = 0$ і перейти на крок XVII.

Основний цикл. XV. Обчислити вектор $(I - H_{\Im_0{}'}) \nabla f_0(x^k)$.

Якщо $(I - H_{\Im_0{}'}) \nabla f_0(x^k) = 0$, то покласти $x^0 = x^k$ і перейти на крок II; інакше перейти на крок XVI.

XVI. Обчислити вектор

$$h^{k+1} = -(I - H_{\Im_0'})\nabla f_0(x^k) + \frac{\left\|(I - H_{\Im_0'})\nabla f_0(x^k)\right\|^2}{\left\|(I - H_{\Im_0'})\nabla f_0(x^{k-1})\right\|^2} h^k.$$

XVII. Обчислити

$$\rho_{k+1} = -\left(\nabla f_0(x^k), h^{k+1}\right) \Big/ \left(h^{k+1}, Ch^{k+1}\right).$$

XVIII. Знайти підмножину $\Im_0'' \subseteq \Im_0'$ таку, що для всіх $i \in \Im_0''$ виконується нерівність $(a^i, h^{k+1}) > 0.$

XIX. Обчислити

$$\bar{\rho}_{k+1} = \min_{i \in \Im_0''} \frac{b_i - (a^i, x^k)}{(a^i, h^{k+1})}.$$

XX. Якщо $\rho_{k+1} < \bar{\rho}_{k+1}$, то перейти на крок XXI; інакше покласти $x^{k+1} = x^k + \bar{\rho}_{k+1} h^{k+1}$, $x^0 = x^{k+1}$ і перейти на крок II.

XXI. Покласти $x^{k+1} = x^k + \rho_{k+1} h^{k+1}$, $k = k+1$ і перейти на крок XV.

XXII. Покласти $\Im_0' = \Im_0$ і перейти на крок XI.



***Теорема 2.*** *Якщо виконуються припущення 2, то алгоритм 2 розв'язує задачу 2 за скінченну кількість ітерацій, при цьому або отримується точка* $x^*$, *яка мінімізує квадратичну функцію*

$$f_0(x) \triangleq \frac{1}{2}(Cx, x) + (d, x)$$

*на множині X, або встановлюється той факт, що* $f_0(x) -$ *необмежена знизу на множині X.*

*Зауваження 2.* У випадку виродженості матриці *C* може виникнути ситуація, коли на кроці XVII алгоритму 2 буде виконуватися $\left(\nabla f_0(x^k), h^{k+1}\right) \neq 0$, але $(h^{k+1}, Ch^{k+1}) = 0$ з чого випливає $\rho_{k+1} = \infty$. Тоді, якщо $\bar{\rho}_{k+1} < \infty$, то здійснюється перехід на крок XX. Якщо ж $\bar{\rho}_{k+1}$ – необмежене, тобто на кроці XVIII буде виконуватись

$$(a^i, h^{k+1}) \leq 0 \text{ для всіх } i \in \Im_0',$$

то це означає, що задача 2 не має розв'язку, оскільки нижня границя функції $f_0(x)$ на множині *X* дорівнює $-\infty$.

*Зауваження 2'.* Наведені в пунктах 1 і 2 алгоритми містять, по суті, одну складну обчислювальну операцію: проектування градієнта на підпростір, тобто обчислення вектора $(I - H_{\Im(x)})\nabla f_0(x)$, що вимагає знаходження оберненої матриці $(A_{\Im(x)} A_{\Im(x)}^T)^{-1}$.

Щоб уникнути обчислення оберненої матриці, можна спочатку знайти розв'язок $\bar{u}^0$ задачі безумовної мінімізації:

$$\text{знайти } \arg\min_u \frac{1}{2}\left\|\nabla f_0(x) + A^T_{\Im(x)} u\right\|^2 . \qquad (6.99)$$

(Задачу (6.99) можна розв'язати, наприклад, за допомогою методу спряжених напрямків). Крім того, розв'язком задачі (6.99) є вектор

$$\bar{u}^0 = -(A_{\Im(x)} A^T_{\Im(x)})^{-1} A_{\Im(x)} \nabla f_0(x).$$

Знаючи $\bar{u}^0$, можна знайти вектор $(I - H_{\Im(x)})\nabla f_0(x)$ за формулою

$$(I - H_{\Im(x)})\nabla f_0(x) = \nabla f_0(x) + A^T_{\Im(x)} \bar{u}^0.$$

### 3. Метод спряжених градієнтів для розв'язування задачі квадратичного програмування з простими обмеженнями

Задача 3. Знайти $\arg\min\limits_{x\in X}\left[\frac{1}{2}(Cx, x) + (d, x)\right]$,

де $X \triangleq \left\{x \mid x_i \geq 0,\ i \in \Im^+\right\}$; $\Im^+$ – деяка підмножина множини $\{1, 2, \ldots, n\}$; $d \in R^n$.

*Припущення 3.* Матриця $C$ – додатньо визначена.

*Означення 3.* Визначимо множину $\Im(x)$ і функцію $f_0(x)$:

$$\Im(x) \triangleq \{i \mid x_i = 0,\ i \in \Im^+\};\quad f_0(x) \triangleq \left(1/2\right)(Cx, x) + (d, x).$$

***Алгоритм 3***

Початок. I. Вибрати початкове наближення $x^0 \in X$.

II. Знайти множину $\Im(x^0)$.

III. Обчислити градієнт

$$\nabla f_0(x^0) \triangleq \left(\frac{\partial f_0(x^0)}{\partial x_1}, \ldots, \frac{\partial f_0(x^0)}{\partial x_n}\right)^T .$$

IV. Якщо $\dfrac{\partial f_0(x^0)}{\partial x_i} = 0,\ \forall i \notin \Im(x^0)$,

то перейти на крок V; інакше перейти на крок XVI.

V. Якщо $\dfrac{\partial f_0(x^0)}{\partial x_i} \geq 0,\quad \forall i \in \Im(x^0)$,

то покласти $x^* = x^0$ і зупинити обчислення (в цьому випадку знаходять оптимальний розв'язок $x^*$ задачі 3); інакше покласти

$$\Im'(x^0) = \left\{ i \left| \frac{\partial f_0(x^0)}{\partial x_i} \geq 0,\quad i \in \Im(x^0) \right. \right\}$$

і перейти на крок VI.

*Примітка.* На наступних кроках застосовується метод спряжених градієнтів для мінімізації функції $f_0(x)$, де в якості змінних беруть тільки $x_i$, $i \notin \Im'(x^0)$, а всі $x_i$, $i \in \Im'(x^0)$ прирівнюються до нуля.

VI. Покласти $h^1 = -\nabla f_0(x^0)$.

VII. Покласти $k = 0$ і перейти на крок X.

Основний цикл. VIII. Обчислити

$$\nabla f_0(x^k) \triangleq \left( \frac{\partial f_0(x^k)}{\partial x_1}, ..., \frac{\partial f_0(x^k)}{\partial x_n} \right)^T.$$

Якщо $\dfrac{\partial f_0(x^k)}{\partial x_i} = 0$, $\forall i \notin \Im'(x^0)$, то покласти $x^0 = x^k$ і перейти на крок II; інакше перейти на крок IX.

IX. Обчислити вектор

$$h^{k+1} = -\nabla f_0(x^k) + \frac{\left\| \nabla f_0(x^k) \right\|^2}{\left\| \nabla f_0(x^{k-1}) \right\|^2} h^k.$$

X. Обчислити

$$\rho_{k+1} = -(\nabla f_0(x^k), h^{k+1}) / (h^{k+1}, Ch^{k+1}).$$

XI. Визначити $\bar{\Im}(x^0) -$ множину всіх $i \notin \Im'(x^0)$, для яких $h_i^{k+1} < 0$.

XII. Обчислити

$$\bar{\rho}_{k+1} = \min_{i \in \bar{\Im}(x^0)} (-x_i^k / h_i^{k+1}).$$

XIII. Якщо $\rho_{k+1} < \bar{\rho}_{k+1}$, то покласти:

$$x_i^{k+1} = x_i^k + \rho_{k+1} h_i^{k+1}, \; i \notin \Im'(x^0); \quad x_i^{k+1} = x_i^k = 0, \; i \in \Im'(x^0),$$

і перейти на крок XIV; інакше покласти:

$$x_i^{k+1} = x_i^k + \bar{\rho}_{k+1} h_i^{k+1}, \; i \notin \Im'(x^0); \quad x_i^{k+1} = x_i^k = 0, \; i \in \Im'(x^0)$$

і перейти на крок XV.

XIV. Покласти $k = k + 1$ і перейти на крок VIII.

XV. Покласти $x^0 = x^{k+1}$ і перейти на крок II.

XVI. Покласти $\Im'(x^0) = \Im(x^0)$ і перейти на крок VI.

Для алгоритму 3 мають місце теорема і зауваження, аналогічні теоремі 2 і зауваженню 2.

### 4. Модифікація методу спряжених напрямків для розв’язування задач квадратичного програмування великої розмірності

Задача 4. Знайти $\arg\min_x f_0(x)$, де $f_0(x) = \left( (1/2)(x, Cx) + (d, x) \right)$, при обмеженнях:

$$(a^j, x) = b_j, \;\; j = \overline{1, m}; \tag{6.100}$$

$$\upsilon_i \le x_i \le w_i, \quad i = \overline{1,n}, \qquad (6.101)$$

де $C - n \times n$-матриця $(C \ge 0)$; $d - n$-вимірний вектор; $a^j$, $j = \overline{1,n} - n$-вимірний рядок, тобто $a^j = (a_1^j, \ldots, a_n^j)$; $b_j$, $j = \overline{1,m}$; $\upsilon_i$, $w_i$, $i = \overline{1,n} -$ дійсні числа.

Наведена в цьому пункті модифікація методу спряжених напрямків пристосована для задач програмування з великим числом змінних $x_i$, $i = \overline{1,n}$, і відносно малим числом $m$ обмежень вигляду (6.100). При відсутності обчислювальної похибки наведений алгоритм визначає ті ж самі точки, що і метод спряжених напрямків.

Відмінністю наведеної модифікації від методу спряжених напрямків є те, що вона вимагає меншої кількості обчислень, меншого об’єму машинної пам’яті і дозволяє уникати накопичення похибок машинних обчислень.

***Алгоритм 4***

Початок. I. Задати початкове наближення $x^0$, що задовольняє обмеженням-рівностям (6.100) і наступним нерівностям:

$$\upsilon_i < x_i^0 < w_i, \quad i = \overline{1,n}.$$

II. Позначити через $A$ $m \times n$-матрицю, $j$-м рядком якої є $a^j$, $j = \overline{1,m}$.

III. Покласти $k = 0$.

Основний цикл. IV. Обчислити множини індексів $\Im^+$, $\Im^-$ за формулами:

$$\Im^+ = \{ i \mid x_i^k \ge \omega_i, \ i = \overline{1,n} \}; \quad \Im^- = \{ i \mid x_i^k \le \upsilon_i, \ i = \overline{1,n} \}$$

і покласти $\Im = \Im^+ \bigcup \Im^-$.

V. Обчислити $\nabla f_0(x^k) -$ градієнт цільової функції в точці $x = x^k$ за формулою

$$\nabla f_0(x^k) = Cx^k + d.$$

VI. Для довільної матриці $B$ з $m$ рядками визначити матрицю $\tilde{B}$, яка утворюється, якщо в $B$ замінити всі рядки з номерами $j \in \Im^+ \bigcup \Im^-$ нульовими рядками, і визначити матрицю $B_+$ (або $B_-$), рядками якої є рядки матриці $B$ з номерами $j \in \Im^+$ (відповідно, $j \in \Im^-$).

VII. Для довільного $n$-вимірного вектор-стовпця $g$ визначити оператори:

$$Q_{\Im} g = \bar{g} - \tilde{A}^T (\tilde{A}\tilde{A}^T)^{-1} \tilde{A}\tilde{g}; \qquad (6.102)$$

$$G_{\Im} g = \begin{pmatrix} G_{\Im}^+ g \\ G_{\Im}^- g \end{pmatrix}, \qquad (6.103)$$

де

$$G_{\Im}^{+} g = (A_{+})^{T} (\tilde{A}\tilde{A}^{T})^{-1} \tilde{A}\tilde{g} - g_{+};$$

$$G_{\Im}^{-} g = g_{-} - (A_{-})^{T} (\tilde{A}\tilde{A}^{T})^{-1} \tilde{A}\tilde{g}$$

(тут $\tilde{A}^{T}$ – матриця, транспонована до $\tilde{A}$).

VIII. Обчислити вектори $Q_{\Im}\nabla f_0(x^k)$, $G_{\Im}\nabla f_0(x^k)$ і покласти $q = G_{\Im}\nabla f_0(x^k)$.

IX. Якщо $Q_{\Im}\nabla f_0(x^k) \neq 0$, то перейти на крок XII; інакше перейти на крок X.

X. Якщо при всіх $i = \overline{1,n}$ виконується $q_i \geq 0$, то зупинити обчислення (в цьому випадку $x^k$ є розв'язком задачі 4); інакше перейти на крок XI.

XI. Побудувати нові множини $\Im^{+}$ і $\Im^{-}$ шляхом видалення з них одного з номерів $i$ таких, що $q_i < 0$ і перейти на крок XII.

Відмітимо, що при зміні множин $\Im^{+}$ і $\Im^{-}$ змінюються також оператори $Q_{\Im}$, $G_{\Im}$.

На кроках XII–XVII здійснюється мінімізація квадратичної функції $f_0(x)$ на підпросторі, що задається рівняннями:

$$x_i^k = w_i,\ i \in \Im^{+};\ x_i^k = \upsilon_i,\ i \in \Im^{-};\quad (a^j, x) = b_j,\ j = \overline{1,m},$$

причому мінімум $f_0(x)$ на вказаному підпросторі знаходять не більш, ніж за $n - m - X$ «малих» ітерацій, що визначається кроками XII-XVII, де $X$ – число елементів у множині $\Im^{+} \bigcup \Im^{-}$.

XII. Покласти $s = 0$.

XIII. Покласти $y^0 = x^k$, $h^0 = 0$.

XIV. Обчислити вектор

$$h^{s+1} = -Q_{\Im}\nabla f_0(y^s) + \left( \left\| Q_{\Im}\nabla f_0(y^s) \right\|^2 \Big/ \left\| Q_{\Im}\nabla f_0(y^{s-1}) \right\|^2 \right) h^s.$$

XV. Обчислити

$$\rho_{s+1} = -\left( \nabla f_0(y^s), h^{s+1} \right) \Big/ \left( h^{s+1}, Ch^{s+1} \right).$$

Якщо $\rho_{s+1} = 0$, то покласти $x^{k+1} = y^s$ і перейти на крок XVIII; інакше перейти на крок XVI.

XVI. Обчислити величину $\bar{\rho}_{s+1}$ за правилом

$$\bar{\rho}_{s+1} = \min \left\{ \min_{h_i^{s+1} > 0} \frac{w_i - y_i^k}{h_i^{s+1}}, \min_{h_i^{s+1} > 0} \frac{y_i^k - \upsilon_i}{h_i^{s+1}} \right\}.$$

XVII. Якщо $\rho_{s+1} < \bar{\rho}_{s+1}$, то покласти:

$$y^{s+1} = y^s + \rho_{s+1} h^{s+1},\ s = s + 1$$

і перейти на крок XIV; якщо $\rho_{s+1} \geq \bar{\rho}_{s+1}$, то покласти

$$x^{k+1} = y^s + \bar{\rho}_{s+1} h^{s+1}$$

і перейти на крок XVIII.

XVIII. Покласти $k = k+1$ і перейти на крок IV.

***Теорема 4.*** *Нехай ітеративний процес, що задається алгоритмом 4, реалізований на ЕОМ і нехай $y^s$ – точка, побудована на $s$-й ітерації. Тоді виконуються наступні нерівності:*

$$y_i^s - w_i \leq c_1 u; \quad \upsilon_i - y_i^s \leq c_1 u;$$

$$\left\| Ay^s - b \right\| \leq c_2 u (1 + \lambda^s) \left\| A \right\| + \left\| Ay^0 - b \right\|$$

(*тут* $c_1$, $c_2$ – *константи;* $u = 2^{1-\tau}$ – *одинична похибка округлення* ($\tau$ – *кількість значущих розрядів мантиси*); $b = (b_1, \ldots, b_m)$; $\lambda$ – *константа, яка задовольняє нерівностям:*

$$\left(1/4\right)(\operatorname{cond}\tilde{C} + (\operatorname{cond}\tilde{C})^{-1} - 2) \leq \lambda \leq \left(1/2\right)(\operatorname{cond}\tilde{C} + (\operatorname{cond}\tilde{C})^{-1} - 2),$$

*де* $\tilde{C}$ – *матриця звуження функції* $f_0(x)$ *на гіперплощину* (6.100); $\operatorname{cond}\tilde{C} \triangleq \left\| \tilde{C} \right\| \cdot \left\| \tilde{C}^{-1} \right\|$ – *число обумовленості матриці* $C$).

*Зауваження 4.* Оскільки похибка виконання рівності (6.100) може зростати доволі швидко, то рекомендується застосовувати на деяких ітераціях по $s$ наступну процедуру корегування обчислювальної похибки:

$$y^s_{\text{кор}} = \theta Q y^s + (1-\theta) y^0, \tag{6.104}$$

де $Q$ – оператор $Q_{\mathfrak{I}}$, що визначається за (6.102) при $\mathfrak{I}^+ = \mathfrak{I}^- = \varnothing$; $\theta$ – максимальне число з відрізку $[0;1]$ таке, що права частина в (6.104) задовольняє обмеженням задачі.

### 5. Стійкий метод для розв'язування задач квадратичного програмування

З а д а ч а  5. Знайти $\arg\min\limits_x \frac{1}{2} \left\| Gx - c \right\|^2$ при обмеженнях:

$$Ax = b, \ \ x \geq 0,$$

де $G$ – $l \times n$-матриця, яка складається із стовпців $g^j$, $j = \overline{1,n}$; $c$ – $l$-вектор; $A$ – $m \times n$-матриця стовпців $a^j$, $j = \overline{1,n}$; $b$ – $m$-вектор.

В наведеному нижче алгоритмі на кожній ітерації вимагається розв'язувати системи лінійних рівнянь. Алгоритм за скінченну кількість ітерацій приводить до розв'язку задачі 5 або встановлює відсутність допустимих розв'язків задачі 5 і, що важливо, може бути застосований і в (майже) виродженому випадку задачі 5.

***Алгоритм 5***

П о ч а т о к . I. Знайти множину індексів $\Im \subseteq \{1,2,\ldots,n\}$ таку, що початкова матриця $\begin{pmatrix} B_{\Im} \\ Q_{\Im} \end{pmatrix}$ (тут і надалі $B_{\Im}$ – підматриця матриці *A*, яка складається із стовпців $a^j$, $j \in \Im$ матриці *A*; $Q_{\Im}$ – підматриця матриці *G*, яка складається із стовпців $g^j$, $j \in \Im$ матриці *G*) має повний стовпцевий ранг і система лінійних рівнянь:

$$B_{\Im} y = b; \quad B_{\Im}^T u + Q_{\Im}^T Q_{\Im} y = Q_{\Im}^T c \tag{6.105}$$

має розв'язок $(y,u)$, для якого $y > 0$.

О с н о в н и й  ц и к л . II. Якщо $y > 0$, то перейти на крок III; інакше перейти на крок VII.

III. Якщо для всіх $j \in \{1,2,\ldots,n\} \setminus \Im$ виконується нерівність

$$(a^j, u) + (g^j, Q_{\Im} y - c) \geq 0,$$

то обчислити оптимальний розв'язок $x^* = (x_1^*, \ldots, x_n^*)^T$ задачі 5:

$$x_j^* = \begin{cases} y_j, & \text{якщо} \quad j \in \Im, \\ 0, & \text{якщо} \quad j \notin \Im, \end{cases} \quad j = \overline{1,n}$$

і зупинити обчислення; інакше покласти $\bar{\Im} = \Im$ і перейти на крок IV.

IV. Знайти індекс $k \in \{1,2,\ldots,n\} \setminus \bar{\Im}$, для якого виконується нерівність

$$(a^k, u) + (g^k, Q_{\Im} y - c) < 0.$$

V. Покласти $\Im = \bar{\Im} \bigcup \{k\}$.

VI. Знайти вектор $\bar{y}$, координати якого $\bar{y}_j$, $j \in \Im$, обчислюються за правилом

$$\bar{y}_i = y_i \text{ при всіх } i \in \bar{\Im}; \quad \bar{y}_k = 0$$

і перейти на крок XI.

VII. Обчислити максимальне значення $\bar{\lambda} \in [0;1]$, при якому

$$\bar{\lambda} y + (1 - \bar{\lambda}) \bar{y} \geq 0,$$

і покласти

$$z_j = \bar{\lambda} y_j + (1 - \bar{\lambda}) \bar{y}_j, \quad j \in \Im.$$

VIII. Покласти $\bar{\Im} = \Im$.

IX. Знайти підмножину $\tilde{\Im}$, яка складається з тих індексів $k \in \bar{\Im}$, для яких $z_k = 0$ і покласти $\Im = \bar{\Im} \setminus \tilde{\Im}$.

X. Обчислити вектор $\bar{y}$, координати якого знаходяться за правилом

$$\bar{y}_j = z_j, \quad j \in \Im,$$

і перейти на крок XI.

XI. У відповідності із множиною $\Im$ скласти підматрицю $B_\Im$ матриці $A$ і підматрицю $Q_\Im$ матриці $G$.

XII. Знайти розв’язок $(y, u)$ системи лінійних рівнянь:

$$B_\Im y = b; \qquad B_\Im u + Q_\Im^T Q_\Im y = Q_\Im^T c \tag{6.106}$$

і перейти на крок II.

***Теорема 5.*** *Нехай існує множина індексів* $\Im \subseteq \{1, 2, \ldots, n\}$ *така, що матриця* $\begin{pmatrix} B_\Im \\ Q_\Im \end{pmatrix}$ *має повний стовпчиковий ранг і система (6.105) має розв’язок* $(y, u)$, *для якого* $y > 0$, *і нехай кожна матриця* $B_\Im$, *одержана в результаті застосування алгоритму, має повний рядковий ранг. Тоді за скінченну кількість ітерацій алгоритму 5 обчислюється оптимальний розв’язок* $x^*$ *задачі 5.*

*Зауваження 5.* Для одержання початкової матриці $\begin{pmatrix} B_\Im \\ Q_\Im \end{pmatrix}$, визначеної на кроці I алгоритму 5, необхідно спочатку отримати розв’язок $(\tilde{t}^*, \tilde{x}^*)$ наступної задачі квадратичного програмування:

$$\text{знайти } \arg\min_{(\tilde{t}, \tilde{x})} \left\| (-b \,\vdots\, A) \begin{pmatrix} \tilde{t} \\ \tilde{x} \end{pmatrix} \right\|^2 \tag{6.107}$$

при обмеженнях

$$\tilde{t} = 1, \ (\tilde{t}, \tilde{x}) \geq 0. \tag{6.108}$$

Якщо виконується нерівність

$$\left\| (-b \,\vdots\, A) \begin{pmatrix} \tilde{t}^* \\ \tilde{x}^* \end{pmatrix} \right\|^2 > 0,$$

то задача 5 не має допустимого розв’язку;
якщо

$$\left\| (-b \,\vdots\, A) \begin{pmatrix} \tilde{t}^* \\ \tilde{x}^* \end{pmatrix} \right\|^2 = 0,$$

то $\tilde{x}^*$ задовольняє умовам $\tilde{x}^* \geq 0$ і $A\tilde{x}^* = b$.

Якщо $\tilde{x}^* \neq 0$, то множина стовпців $a^j$, для яких $\tilde{x}_j^* > 0$, лінійно незалежна і разом із відповідними стовпцями $g^j$ утворює початкову матрицю $\begin{pmatrix} B_{\Im} \\ Q_{\Im} \end{pmatrix}$ алгоритму 5.

Відмітимо, що для розв'язування задачі (6.107) – (6.108) можна використовувати алгоритм 5 із початковою матрицею $\begin{pmatrix} B_{\Im} \\ Q_{\Im} \end{pmatrix}$, яка дорівнює $(1;-b)^T$.

*Зауваження 5′.* Для розв'язування систем лінійних рівнянь (6.105) і (6.106) пропонується наступний спосіб.

Нехай $\begin{pmatrix} B_{\Im} \\ Q_{\Im} \end{pmatrix}$ має розмірність $(m+l)\times r$. Тоді $y$ і $u$ визначаються за допомогою розв'язування трикутних лінійних систем

$$S(b-u)=h;\ \ Vy=d$$

для $(b-u)$ і $y$, де $d=W_Q^T c+Zh;$ $h=S^{-T}b-Z^T W_Q^T c;$ $W_B^T=ZS;$ $Z-$ $r\times m$-матриця з ортогональними стовпцями і $S-$ верхня трикутна невироджена $m\times m$-матриця, $\begin{pmatrix} B_{\Im} \\ Q_{\Im} \end{pmatrix}=WV$, $W-(m+l)\times r$-матриця з ортогональними стовпцями, $V-$ верхня трикутна невироджена $r\times r$-матриця, $W=\begin{pmatrix} W_B \\ W_Q \end{pmatrix}$, $W_B-$ має $m$ рядків, $W_Q-$ має $l$ рядків.

**Приклад 2.** Розв'язати аналітичним способом задачу квадратичного програмування:

$$L(x)=5(x_1-3)^2+15(x_2-3{,}5)^2 \to \min$$

при обмеженнях:

$$3x_2+2x_1 \le 12,$$
$$2x_2-3x_1 \le 6,$$
$$x_1 \ge 0, x_2 \ge 0.$$

*Розв'язування.* Спочатку зобразимо допустиму область і лінії рівня цільової функції (еліпси).

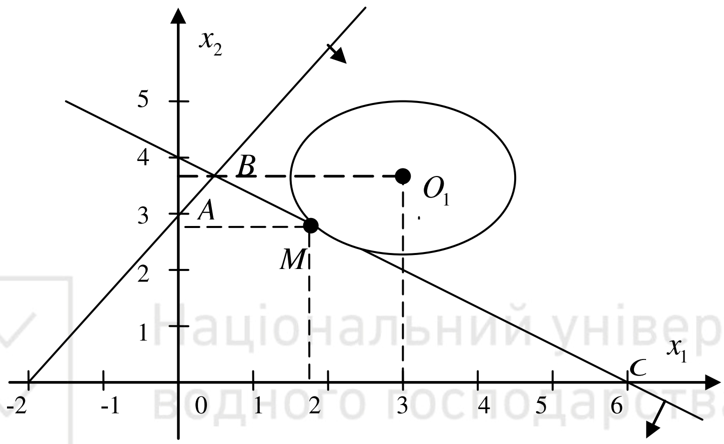
Допустима область задачі є чотирикутником $ABCO$.

Без врахування обмежень цільова функція $L(x)$ досягає мінімуму в точці (3;3,5) причому $L((3;3{,}5))=0$. Найменше значення цільова функція в чотирикутнику $ABCO$ приймає в точці дотику еліпса і прямої $BC$, отже координати точки $M(x_1^*, x_2^*)$ задовольняють рівнянню прямої $BC$

$$3x_2^* + 2x_1^* = 12$$

і рівнянню еліпса з невідомим значенням $L(x^*)$. Друге рівняння отримаємо з таких міркувань. Бачимо, що кутовий коефіцієнт дотичної до еліпса в точці $M$ дорівнює кутовому коефіцієнту прямої $BC$, тобто $(-2/3)$. Рівняння еліпса, що проходить через точку $M$ наступне:

$$\overline{L}(x_1, x_2) = 5(x_1 - 3)^2 + 15(x_2 - 3{,}5)^2 - L(x^*) = 0.$$




Вважаємо $x_2$ неявною функцією від $x_1$, тоді маємо:

$$\overline{L'}_{x_1} + \overline{L'}_{x_2} \cdot \frac{dx_2}{dx_1} = 0 \Leftrightarrow \frac{dx_2}{dx_1} = -\frac{\overline{L'}_{x_1}}{\overline{L'}_{x2}} = -\frac{10(x_1 - 3)}{30(x_2 - 3{,}5)} = -\frac{x_1 - 3}{3(x_2 - 3{,}5)}.$$

Звідси отримаємо друге рівняння для $(x_1^*, x_2^*)$:

$$-\frac{x_1^* - 3}{3(x_2^* - 3{,}5)} = -\frac{2}{3}.$$

Розв’язуємо систему:

$$\begin{cases} 3x_2^* + 2x_1^* = 12 \\ \dfrac{x_1^* - 3}{3(x_2^* - 3{,}5)} = -\dfrac{2}{3} \end{cases} \Leftrightarrow \begin{cases} 3x_2^* + 2x_1^* = 12 \\ -2x_2^* + x_1^* = -4 \,|\cdot(-2) \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} 7x_2^* = 20 \\ x_1^* = -4 + 2x_2^* \end{cases} \Leftrightarrow \begin{cases} x_2^* = 20/7 \\ x_1^* = -4 + \dfrac{40}{7} \end{cases} \Leftrightarrow \begin{cases} x_2^* = 20/7 \\ x_1^* = 12/7 \end{cases}.$$

Отже, знайшли оптимальний розв'язок $x^* = (12/7; 20/7)^T$ і оптимальне значення цільової функції $L(x^*) = 405/28 \approx 14{,}464$.

## Запитання та вправи для самостійної роботи до розділу 6.

1. Сформулюйте необхідну і достатню умову локальної оптимальності.
2. Запишіть формули обчислення крокового множника $\rho_k$ в методі проекції градієнта.
3. В чому полягає сутність гібридного методу проекції градієнта?
4. Сформулюйте означення штрафних функцій та штрафних коефіцієнтів (коефіцієнтів штрафу).
5. В чому полягає сутність штрафних методів?
6. Наведіть приклади зовнішніх і внутрішніх штрафних функцій.
7. Сформулюйте означення послідовності зовнішніх штрафних функцій для множини $X$.
8. Методи зовнішніх штрафних функцій забезпечують наближений розв’язок задачі оптимізації зовні чи всередині допустимої області?
9. Назвіть переваги і недоліки методів зовнішніх штрафних функцій.
10. Сформулюйте означення послідовності внутрішніх штрафних функцій для множини $X^0$.
11. Назвіть переваги і недоліки методів внутрішніх функцій.
12. В чому полягає сутність і перевага комбінованих методів штрафних функцій?
13. Сформулюйте означення можливого напрямку в точці $x^k$.
14. Чим відрізняються різні варіанти методів можливих напрямків?
15. Яку допоміжну задачу потрібно розв’язувати на кожній ітерації в методі можливих напрямків з квадратичним пошуком?
16. В яких випадках рекомендується використовувати стохастичний аналог методу можливих напрямків?
17. В чому полягає сутність методу центрів?
18. Як можна трактувати метод центрів?
19. В чому полягає сутність модифікованого методу центрів?
20. Яка швидкість збіжності методів типу Ньютона?
21. Яка допоміжна оптимізаційна задача розв’язується на кожній ітерації в методі лінеаризації?
22. В чому полягає сутність методів відсікання?
23. Запишіть функції Лагранжа для оптимізаційних задач з обмеженнями типу рівностей.
24. Запишіть модифіковану функцію Лагранжа для оптимізаційної задачі з обмеженнями типу нерівностей.
25. Запишіть «навантажений» функціонал для оптимізаційних задач з обмеженнями типу рівностей і з обмеженнями типу нерівностей.

26. Яка швидкість збіжності методу «навантаженого» функціоналу в опуклому випадку?
27. В чому полягає сутність методів штрафних оцінок?
28. Яка швидкість збіжності методів штрафних оцінок в детермінованому випадку?
29. В чому схожість і відмінність методів проекції градієнта і узагальненого градієнта?
30. Яким умовам задовольняє кроковий множник $\rho_k$ в багатокроковому методі узагальненого градієнтного спуску?
31. У яких випадках рекомендується використовувати метод умовного градієнта?
32. В якому напрямку здійснюється рух до наступного наближення $x^{k+1}$ в методі умовного градієнта?
33. Яким способом шукають напрямок руху до наступного наближення $x^{k+1}$ в стохастичному аналізі методу спряжених градієнтів?
34. В яких напрямках здійснюється рух до наступного наближення $x^{k+1}$ в методах покоординатного спуску?
35. Дайте означення стохастичного квазіградієнта.
36. Запишіть приклади формул для обчислення стохастичних квазіградієнтів.
37. В чому полягає сутність стохастичного методу скорочення нев'язок?
38. Для розв'язування яких задач рекомендується використовувати комбінований метод стохастичних градієнтів і штрафних функцій?
39. З яких умов визначається кроковий множник $\rho_k$ в методах усереднення напрямків спуску?
40. Як знаходиться напрямок руху в методах випадкового пошуку?
41. Як обчислюють кроковий множник $\rho_k$ в методах випадкового пошуку?
42. Сформулюйте задачу квадратичного програмування.
43. За яку кількість ітерацій гарантує розв'язок задачі квадратичного програмування метод спряжених градієнтів?
44. Реалізуйте декілька ітерацій методу умовного градієнту для функції $f(x) = x_1{}^2 + x_1 x_2 + x_2{}^2$ при $X = \left\{ x \in R^2 : 0 \le x_1 \le 1, -1 \le x_2 \le 0 \right\}$, вибираючи за початкове наближення $x^{(0)} = (1;-1)^T$, $(-1;0)^T$, $(1;0)^T$, $(0;0)^T$.

45. Знайдіть умовний антиградієнт $h$ функції $f(x) = -\frac{1}{8}x_1^2 - \frac{1}{3}x_2^2 + 2$ у точці $(2,3)^T$ на множині $X = \{x \in R^2 : x_1 + x_2 \le 6, x_1 - x_2 \le 1,$ $2x_1 + x_2 \ge 6,\ \ (1/2)x_1 - x_2 \ge -4,\ \ x_1 \ge 0,\ x_2 \ge 0\}$. Знайти довжину кроку $\alpha$ в напрямку $h$ з умови одновимірної мінімізації та по правилу дроблення кроку.

46. Розв'яжіть методом проекції градієнту та методом можливих напрямків таку задачу:

$$f_0(x) = \exp(x_1) + x_1^2 - x_1x_2 - 3x_2^2 + 4x_1 - 6x_2 \to \min,$$
$$2x_1 + x_2 \le 8,$$
$$-x_1 + x_2 \le 2,$$
$$x_1 \ge 1, x_2 \ge 3.$$

47. Розв'яжіть методом можливих напрямків таку задачу:

$$f_0(x) = x_1^3 - 2x_2^3 + x_1 - 2x_2 - x_1^2 \to \min,$$
$$x_1 + 2x_2 \le 6,$$
$$-x_1 + 2x_2 \le 3,$$
$$x_1 \ge 0, x_2 \ge 0.$$



48. Розв'яжіть модифікованим методом можливих напрямків таку задачу:

$$f_0(x) = \exp(-x_1 - x_2) + (1 - x_1)^2 - 10(x_1 - x_2)^2 + x_1^2 - 2x_1x_2 \to \min,$$
$$x_1^2 + x_2^2 \le 16,$$
$$(x_2 - x_1)^2 + x_1 \le 6,$$
$$x_1 + x_2 \ge 2.$$

49. Розв'яжіть таку задачу методом проекції градієнту:

$$f_0(x) = \exp(-x_1 - x_2) + (1 - x_1)^2 - 10(x_1 - x_2)^2 + x_1^2 - 2x_1x_2 \to \min,$$
$$2x_1 + 5x_2 \le 25,$$
$$-x_1 + 2x_2 + x_3 \le 8,$$
$$x_1 \ge 0, x_2 \ge 0, x_3 \ge 0.$$

50. Розв'яжіть методом штрафних функцій такі задачі:

- $f_0(x) = x_1^2 + x_2^2 \to \min,\ \ x_1^2 - x_2 \le 0, -x_1 \le 0,$
- (з використанням логарифмічної функції штрафу)
  $f_0(x) = x^2 - 10x \to \min,\ \ x - 1 \le 0,$
- (з використанням квадратичної функції штрафу)
  $f_0(x) = (x_1 - x_2)^2 + (x_2 - x_3)^2 \to \min,\ \ x_1^2 + x_1x_2^2 + x_3^4 = 3.$

51. Розв'яжіть задачу методом послідовної безумовної оптимізації:

$$f_0(x) = \exp(-x_1 - x_2) - x_1{}^2 + 2x_2{}^2 + 2x_1x_2 \to \min,$$

$$x_1 + x_2 - 4 = 0, \quad x_1 + x_2 \le 1.$$

52. Розв'яжіть задачу методом умовного градієнту:

$$f_0(x) = 2x_1{}^2 + x_2{}^2 + 2x_3{}^2 - x_1x_3 - 6 \to \max,$$

$$x_1 - 3x_2 + 2x_3 \le 1,$$

$$2x_1 + x_2 \le 10,$$

$$x_1 \ge 0, x_2 \ge 0, x_3 \ge 0$$

$$x^{(0)} = (1;1;1)^T.$$

53. Розв'яжіть наступні задачі квадратичного програмування:

а) $f_0(x) = -x_1 - 2x_2 + 2x_2^2 \to \min$

при обмеженнях:

$3x_1 + 2x_2 \le 6,\ x_1 + 2x_2 \le 4,\ x_1 \ge 0,\ x_2 \ge 0;$

б) $f_0(x) = x_1^2 + x_2^2 + x_3^2 - 2x_1 + x_2 \to \min$

при обмеженнях:

$x_1 + 2x_2 + 3x_3 \le 12,\ 2x_1 + x_2 + x_3 \le 6;$

в) $f_0(x) = x_1^2 + 2x_2^2 - 2x_1x_2 + 5x_1 - 6x_2 \to \min$

при обмеженнях:

$2x_1 + x_2 \le 8,\ x_1 + 3x_2 \le 15,\ x_1 \ge 0,\ x_2 \ge 0;$

г) $f_0(x) = -x_1^2 - x_2^2 + 3x_1 + 4x_2 \to \max$

при обмеженнях:

$x_1 + 2x_2 \le 20,\ x_1 + x_2 \ge 8,\ x_1 \ge 0,\ x_2 \ge 0;$

д) $f_0(x) = -x_1^2 - x_2^2 + x_1x_2 + 5x_1 + 2x_2 \to \max$

при обмеженнях:

$2x_1 + 3x_2 \le 15,\ x_1 + 2x_2 = 8,\ x_1 \ge 0,\ x_2 \ge 0;$

е) $f_0(x) = -x_1^2 - 3x_2^2 + x_1x_2 + 15x_1 + 8x_2 \to \max$

при обмеженнях:

$3x_1 + x_2 \le 15,\ 0,5x_1 + x_2 \le 5,\ x_1 \ge 0,\ x_2 \ge 0;$

є) $f_0(x) = x_1^2 + x_2^2 + x_3^2 - x_1x_2 - 5x_1 + 6x_2 \to \min$

при обмеженнях:

$x_1 + x_2 + x_3 \le 20,\ x_1 + x_2 \le 15,\ x_2 + x_3 \le 10,\ x_1 \ge 0,\ x_2 \ge 0,\ x_3 \ge 0.$

**Р о з д і л  7**

# СПЕЦІАЛЬНІ МЕТОДИ РОЗВ'ЯЗУВАННЯ МІНІМАКСНИХ ЗАДАЧ І ВІДШУКАННЯ СІДЛОВИХ ТОЧОК

## 7.1. Методи послідовних наближень розв'язування дискретних мінімаксних задач

### 1. Мінімаксна задача з обмеженнями простої структури

З а д а ч а  1. Знайти $\arg\min\limits_{x\in X}\max\limits_{i\in\Im}\varphi_i\left(x\right)$ для заданих функцій $\varphi_i : R^n \to R^1,\ i\in\Im$, заданої множини індексів $\Im$ і заданої множини обмежень $X\subset R^n$.

*Припущення 1.* $(\boldsymbol{i})-$ функції $\varphi_i,\ i\in\Im$ – неперервно диференційовані; $(\boldsymbol{ii})-$ множина $X-$ опукла і замкнута.

***Алгоритм 1***

П о ч а т о к . I. Вибрати початкове наближення $x^0\in X$.

II. Вибрати константи $\varepsilon_0>0$ і $\alpha_0>0$.

III. Покласти $k=0$.

О с н о в н и й  ц и к л . IV. Знайти множину точок:

$$X^k=\left\{x\middle|\left\|x-x^k\right\|_1\le 1,\ x\in X\right\},$$

де $\|y\|_1=\max\limits_{i\in[1:n]}|y_i|$.

V. Обчислити множину індексів

$$K_0(x^k)=\left\{i\mid\varphi_i(x^k)=\max_{i\in\Im}\varphi_i(x^k),\ i\in\Im\right\}.$$

VI. Обчислити

$$\psi(x^k)=\min_{x\in X^k}\max_{i\in K_0(x^k)}\left(\nabla\varphi_i(x^k),\ x-x^k\right).$$

VII. Якщо $\psi(x^k)=0$, то покласти $x^*=x^k$ і припинити обчислення; якщо $\psi(x^k)<0$, то перейти на крок VIII.

VIII. Покласти $j=0$.

IX. Покласти $\varepsilon=\varepsilon_j$.

X. Обчислити множину індексів

$$K_\varepsilon(x^k)=\left\{i\mid\max_{i\in\Im}\varphi_i(x^k)-\varphi_i(x^k)\le\varepsilon,\ \ i\in\Im\right\}.$$

XI. Обчислити

$$\psi_\varepsilon(x^k)=\min_{x\in X^k}\max_{i\in K_\varepsilon(x^k)}\left(\nabla\varphi_i(x^k),\ x-x^k\right).$$

XII. Якщо виконується нерівність $\psi_{\varepsilon}\left(x^{k}\right) \leq -\alpha_{0}\varepsilon / \varepsilon_{0}$, то перейти на крок XIII; інакше покласти $\varepsilon_{j+1} = \varepsilon_{j} / 2, \;\; j = j+1$ і перейти на крок IX.
(Вихід із циклу, який визначається кроками IX – XII, буде здійснюватися за скінченне число ітерацій при кожному $k = 0,1,...$).

XIII. Обчислити вектор $y^{k}$, який задовольняє умову

$$\psi_{\varepsilon}\left(x^{k}\right) = \max_{i \in K_{\varepsilon}\left(x^{k}\right)} \left(\nabla \varphi_{i}\left(x^{k}\right), y^{k} - x^{k}\right).$$

XIV. Обчислити кроковий множник $\rho_{k}$ з умови

$$\max_{i \in \Im} \varphi_{i}\left(x^{k} + \rho_{k}\left(y^{k} - x^{k}\right)\right) = \min_{\rho \in [0;1]} \max_{i \in \Im} \varphi_{i}\left(x^{k} + \rho\left(y^{k} - x^{k}\right)\right).$$

XV. Обчислити наступне наближення

$$x^{k+1} = x^{k} + \rho_{k}\left(y^{k} - x^{k}\right).$$

XVI. Покласти $k = k+1$ і перейти на крок IV.

***Теорема 1.*** *Якщо виконані припущення 1 і початкове наближення* $x^{0}$ *в алгоритмі 1 таке, що множина*

$$X\left(x^{0}\right) = \left\{ x \,\middle|\, \max_{i \in \Im} \varphi_{i}\left(x\right) \leq \max_{i \in \Im} \varphi_{i}\left(x^{0}\right), \;\; x \in X \right\}$$

*обмежена, то будь-яка гранична точка* $x^{*}$ *нескінченної послідовності* $\left\{x^{k}\right\}_{k=0}^{\infty}$, *яка породжена алгоритмом 1, є стаціонарною точкою функції* $\max_{i \in \Im} \varphi_{i}\left(x\right)$ *на множині* $X$.

*Зауваження 1.* Основна складність в застосуванні алгоритму 1 пов'язана з обчисленням значення $\psi_{\varepsilon}\left(x^{k}\right)$ на кроці XI. Якщо $X^{k}$ – строго опукла множина і початок координат не належить опуклій оболонці $L_{\varepsilon}\left(x^{k}\right)$, яка натягнута на вектори $\nabla \varphi_{i}\left(x^{k}\right), \; i \in K_{\varepsilon}\left(x^{k}\right)$, то задача обчислення $\psi_{\varepsilon}\left(x^{k}\right)$ зводиться до простішої задачі максимізації на $L_{\varepsilon}\left(x^{k}\right)$ неперервно диференційованої функції

$$\theta\left(z\right) \triangleq \min_{v \in X^{k}} \left(z, v - x^{k}\right),$$

градієнт якої дорівнює:

$$\nabla \theta\left(z\right) = v\left(z\right) - x^{k},$$

де $v\left(z\right) \in X^{k}$ задовольняє умову $\theta\left(z\right) = \left(z, v\left(z\right) - x^{k}\right)$.

**2. Перший метод послідовних наближень розв'язування мінімаксної задачі з обмеженнями типу нерівностей**

З а д а ч а 2. Знайти $\arg\min_{x \in X} \max_{i \in \Im} \varphi_{i}\left(x\right)$ для заданих функцій $\varphi_{i} : R^{n} \rightarrow R^{1}, \; i \in \Im$, заданої множини індексів $\Im$ і множини

$$X \triangleq \left\{ x \middle| f_j(x) \le 0, \quad j \in \Im_1, \quad x \in R^n \right\}.$$

*Припущення 2.* ($\boldsymbol{i}$)– функції $\varphi_i$, $i \in \Im$– неперервно диференційовані; ($\boldsymbol{ii}$)– функції $f_j$, $j \in \Im_1$– опуклі і неперервно диференційовані; ($\boldsymbol{iii}$)– виконується умова Слейтера

$$\inf_{x \in R^n} \max_{i \in \Im_1} f_i(x) < 0.$$

В наступних пунктах 2 – 4 наводяться методи послідовних наближень для знаходження стаціонарних точок $x^*$ функції $\max\limits_{i \in \Im} \varphi_i(x)$ на множині $X$.

*Означення 1.* Точка $x^* \in X$ називається ***стаціонарною*** точкою функції $\max\limits_{i \in \Im} \varphi_i(x)$ на множині $X$, якщо многогранник $\tilde{L}(x^*)$ містить початок координат (тут

$$\tilde{L}(x^*) = \mathrm{co}\, \hat{H}(x^*), \ \hat{H}(x^*) = H(x^*) \cup H_1(x^*);$$

$$H(x^*) = \left\{ \nabla \varphi_i(x^*), \ i \in K_0(x^*) \right\}; \ H_1(x^*) = \left\{ \nabla f_j(x^*), \ j \in Q_0(x^*) \right\};$$

$$K_0(x^*) = \left\{ i \middle| \max_{i \in \Im} \varphi_i(x^*) = \varphi_i(x^*), \ i \in \Im \right\}; \ Q_0(x^*) = \left\{ j \mid f_j(x^*) = 0, \ j \in \Im_1 \right\}).$$

Якщо функція $\max\limits_{i \in \Im} \varphi_i(x)$ – опукла донизу на $X$, то стаціонарна точка $x^*$ є точкою мінімуму.

***Алгоритм 2***

Початок. I. Вибрати початкове наближення $x^0 \in X$.

II. Вибрати константи $\varepsilon_0 > 0$, $\mu_0 > 0$, $\alpha_0 > 0$.

III. Покласти $k = 0$.

Основний цикл. IV. Знайти множини індексів:

$$K_0(x^k) = \left\{ i \mid \varphi_i(x^k) = \max_{i \in \Im} \varphi_i(x^k), \ i \in \Im \right\}; \ Q_0(x^k) = \left\{ j \mid f_j(x^k) = 0, \ j \in \Im_1 \right\}.$$

V. Обчислити множину векторів:

$$\tilde{H}_{00}(x^k) = H_0(x^k) \cup H_0'(x^k),$$

де $H_0(x^k) = \left\{ \nabla \varphi_i(x^k), \ i \in K_0(x^k) \right\}$; $H_0'(x^k) = \left\{ \nabla f_j(x^k), \ j \in Q_0(x^k) \right\}$.

VI. Використовуючи алгоритм $1'$ (п. 4.10), визначити, чи належить початок координат многограннику $\tilde{L}_{00}(x^k)$, який є опуклою оболонкою, натягнутою на множину векторів $\tilde{H}_{00}(x^k)$. Якщо початок координат належить многограннику $\tilde{L}_{00}(x^k)$, то покласти $x^* = x^k$ і припинити обчислення; інакше перейти на крок VII.

VII. Покласти $s = 0$.

VIII. Покласти $\varepsilon = \varepsilon_s$, $\mu = \mu_s$, $\alpha = \alpha_s$.

IX. Знайти множини індексів:

$$K_\varepsilon\left(x^k\right)=\left\{i \middle| \max_{i\in\Im}\varphi_i\left(x^k\right)-\varphi_i\left(x^k\right)\le\varepsilon,\ i\in\Im\right\};$$

$$Q_\varepsilon\left(x^k\right)=\left\{j\,|-\mu\le f_j\left(x^k\right)\le 0,\ j\in\Im_1\right\}.$$

X. Обчислити множину векторів:

$$\tilde{H}_{\varepsilon\mu}=H_\varepsilon\left(x^k\right)\cup H'_\mu\left(x^k\right),$$

де

$$H_\varepsilon\left(x^k\right)=\left\{\nabla\varphi_i\left(x^k\right),\ \ i\in K_\varepsilon\left(x^k\right)\right\};\ H'_\mu\left(x^k\right)=\left\{\nabla f_j\left(x^k\right),\ \ j\in Q_\varepsilon\left(x^k\right)\right\}.$$

XI. Використовуючи алгоритм $1''$ (п. 4.10), обчислити точку $z_{\varepsilon\mu}$ – найближчу до початку координат точку многогранника $\tilde{L}_{\varepsilon\mu}\left(x^k\right)$, який є опуклою оболонкою, натягнутою на множину векторів $\tilde{H}_{\varepsilon\mu}\left(x^k\right)$.

XII. Якщо виконується нерівність $\left\|z_{\varepsilon\mu}\right\|\ge\alpha$, то перейти на крок XIII; інакше покласти $\varepsilon_{s+1}=\varepsilon_s/2$, $\mu_{s+1}=\mu_s/2$, $\alpha_{s+1}=\alpha_s/2$, $s=s+1$ і перейти на крок VIII.

XIII. Обчислити вектор $h_{\varepsilon\mu}\left(x^k\right)$ – напрям $(\varepsilon,\mu)$-квазінайшвидшого спуску функції $\max\limits_{i\in\Im}\varphi_i\left(x\right)$ в точці $x^k$ за формулою

$$h_{(\varepsilon,\mu)}\left(x^k\right)=-\left(1/\left\|z_{\varepsilon\mu}\right\|\right)z_{\varepsilon\mu}.$$

XIV. Обчислити кроковий множник $\rho_k$ з умов:

$$\max_{i\in\Im}\varphi_i\left(x^k+\rho_k h_{\varepsilon\mu}\left(x^k\right)\right)=\min_{\rho\ge0}\max_{i\in\Im}\varphi_i\left(x^k+\rho h_{\varepsilon\mu}\left(x^k\right)\right);$$

$$x^k+\rho_k h_{\varepsilon\mu}\left(x^k\right)\in X\ .$$

XV. Обчислити наступне наближення

$$x^{k+1}=x^k+\rho_k h_{\varepsilon\mu}\left(x^k\right).$$

XVI. Покласти $k=k+1$ і перейти на крок IV.

***Теорема 2.*** *Якщо виконані припущення 2 і початкове наближення $x^0\in X$ таке, що множина*

$$X^0\triangleq\left\{x\middle|\max_{i\in\Im}\varphi_i\left(x\right)\le\max_{i\in\Im}\varphi_i\left(x^0\right),\ \ x\in X\right\}$$

*обмежена, то будь-яка гранична точка нескінченної послідовності $\left\{x^k\right\}_{k=0}^{\infty}$, яка породжена алгоритмом 2, є стаціонарною точкою функції $\max\limits_{i\in\Im}\varphi_i\left(x\right)$ на множині $X$.*

### 3. Другий метод послідовних наближень розв'язування мінімаксної задачі з обмеженнями типу нерівностей

***Алгоритм 3***

П о ч а т о к . I. Вибрати початкове наближення $x^0 \in X$ .

II. Вибрати константи $\varepsilon > 0, \quad \mu > 0$ .

III. Покласти $k = 0$.

О с н о в н и й ц и к л . IV. Знайти множини індексів:

$$K_\varepsilon(x^k) = \{ i \mid \max_{i \in \Im} \varphi_i(x^k) - \varphi_i(x^k) \le \varepsilon, \quad i \in \Im \};$$

$$Q_\varepsilon(x^k) = \{ j \mid -\mu \le f_j(x^k) \le 0, \quad j \in \Im_1 \}.$$

V. Обчислити множину векторів:

$$\widetilde{H}_{\varepsilon\mu}(x^k) = H_\varepsilon(x^k) \bigcup H'_\mu(x^k),$$

де $H_\varepsilon(x^k) = \{ \nabla\varphi_i(x^k), \quad i \in K_\varepsilon(x^k) \}$; $H'_\mu(x^k) = \{ \nabla f_j(x^k), \quad j \in Q_\varepsilon(x^k) \}$.

VI. Використовуючи алгоритм $1'$ (п. 4.10), визначити, чи належить початок координат многограннику $\widetilde{L}_{\varepsilon\mu}(x^k)$, який є опуклою оболонкою, натягнутою на множину векторів $\widetilde{H}_{\varepsilon\mu}(x^k)$.

Якщо початок координат належить многограннику $\widetilde{L}_{\varepsilon\mu}(x^k)$, то покласти $x^*_{\varepsilon\mu} = x^k$ і зупинити обчислення; інакше перейти на крок VII.

VII. Використовуючи алгоритм $1''$ (п. 4.10), обчислити точку $z^k_{\varepsilon\mu}$ – найближчу до початку координат точку многогранника $\widetilde{L}_{\varepsilon\mu}(x^k)$. Покласти $\delta^k_{\varepsilon\mu} = \left\| z^k_{\varepsilon\mu} \right\|$.

VIII. Обчислити вектор $h_{\varepsilon\mu}(x^k)$ – напрямок $(\varepsilon, \mu)$-найшвидшого спуску функції $\max_{i \in \Im} \varphi_i(x)$ в точці $x^k$ за формулою

$$h_{\varepsilon\mu}(x^k) = -(1/\delta^k_{\varepsilon\mu}) z^k_{\varepsilon\mu}.$$

IX. Обчислити кроковий множник $\rho_k$, який задовольняє умовам:

$$\max_{i \in \Im} \varphi_i(x^k + \rho_k h_{\varepsilon\mu}(x^k)) = \min_{\rho \ge 0} \max_{i \in \Im} \varphi_i(x^k + \rho h_{\varepsilon\mu}(x^k));$$

$$x^k + \rho_k h_{\varepsilon\mu}(x^k) \in X.$$

X. Обчислити наступне наближення

$$x^{k+1} = x^k + \rho_k h_{\varepsilon\mu}(x^k).$$

XI. Покласти $k = k + 1$ і перейти на крок IV.

***Теорема 3.*** *Якщо виконуються всі умови теореми 2, то будь-яка гранична точка нескінченної послідовності* $\{x^k\}_{k=0}^\infty$, *яка породжена* *алгоритмом 3, є* $(\varepsilon,\mu)$*-квазістаціонарною точкою функції* $\max\limits_{i\in\Im}\varphi_i(x)$ *на множині X.*

(Точка $x^*_{\varepsilon\mu}$ називається $(\varepsilon,\mu)$-***квазістаціонарною точкою*** функції $\max\limits_{i\in\Im}\varphi_i(x)$ на множині $X$, якщо $0\in\tilde{L}_{\varepsilon\mu}(x^*_{\varepsilon\mu})$).

### 4. Модифікація другого методу послідовних наближень

Алгоритм 3 для знаходження $(\varepsilon,\mu)$-квазістаціонарних точок використовується у наведеному нижче алгоритмі для пошуку стаціонарних точок функції $\max\limits_{i\in\Im}\varphi_i(x)$ на множині X.

***Алгоритм 4***

Початок. I. Вибрати початкове наближення $\bar{x}^0\in X$.

II. Вибрати константи $\varepsilon_0>0,\quad \mu_0>0,\quad \alpha_0>0$.

III. Покласти $s=0$.

Основний цикл. IV. Знайти множини індексів:

$$K_0(\bar{x}^s)=\{i\,|\max_{i\in\Im}\varphi_i(\bar{x}^s)=\varphi_i(\bar{x}^s),\quad i\in\Im\};$$

$$Q_0(\bar{x}^s)=\{j\,|\,f_j(\bar{x}^s)=0,\ \ j\in\Im_1\}.$$

V. Обчислити множину векторів:

$$\tilde{H}_{00}(\bar{x}^s)=H_0(\bar{x}^s)\bigcup H_0'(\bar{x}^s),$$

де $H_0(\bar{x}^s)=\{\nabla\varphi_i(\bar{x}^s),\quad i\in K_0(\bar{x}^s)\}$; $H_0'(\bar{x}^s)=\{\nabla f_j(\bar{x}^s),\quad j\in Q_0(\bar{x}^s)\}$.

VI. Використовуючи алгоритм $1'$ (п. 4.10), визначити, чи належить початок координат многограннику $\tilde{L}_{00}(\bar{x}^s)$, який є опуклою оболонкою, натягнутою на множину векторів $\tilde{H}_{00}(\bar{x}^s)$. Якщо початок координат належить многограннику $\tilde{L}_{00}(\bar{x}^s)$, то покласти $x^*=\bar{x}^s$ і зупинити обчислення; інакше перейти на крок VII.

VII. Використовуючи алгоритм 3, в якому покласти

$$\mu=\mu_s,\quad \varepsilon=\varepsilon_s,\quad x^0=\bar{x}^s,$$

знайти перший індекс $k$, при якому виконується нерівність $\delta^k_{\varepsilon\mu}\le\alpha_s$ і покласти $\bar{x}^{s+1}=x^k$.

VIII. Покласти $\varepsilon_{s+1}=\varepsilon_s/2,\ \mu_{s+1}=\mu_s/2,\ \alpha_{s+1}=\alpha_s/2$.

IX. Покласти $s=s+1$ і перейти на крок IV.

***Теорема 4.*** *Якщо виконуються всі умови теореми 2, то будь-яка гранична точка нескінченної послідовності* $\{\bar{x}^s\}_{s=0}^{\infty}$, *яка породжена алгоритмом 4, є стаціонарною точкою функції* $\max\limits_{i\in\Im}\varphi_i(x)$ *на множині X.*

## 7.2. Узагальнений безпараметричний метод зовнішньої точки розв'язування дискретних мінімаксних задач

Задача 0. Знайти $\arg\min\limits_{x\in X}\max\limits_{i\in[1:m]} f_i(x)$, де

$$X \triangleq \left\{x \mid g_j(x)\le 0,\ j\in\Im,\ h_l(x)=0,\ l\in L,\ x\in R^n\right\};$$

$$g_j : R^n \to R^1,\ j\in\Im,\ h_l : R^n \to R^1,\ l\in L;$$

$\Im$ та $L$ – скінченні множини індексів.

*Припущення 0.* (***i***) – функції $f_i(x)$, $i=\overline{1,m}$, $g_j(x)$, $j\in\Im$, $h_l(x)$, $l\in L$ – неперервно диференційовані в $R^n$; (***ii***) – існує розв'язок $x^*$ мінімаксної задачі 0.

Позначимо

$$f_0(x)=\max_{i\in[1:m]} f_i(x);\ \Im(x)=\left\{j \mid g_j(x)>0,\ j\in\Im\right\}.$$

### 1. Основний метод

На $k$-й ітерації алгоритму обчислюється точка $x^k$ безумовного мінімуму деякої допоміжної функції і оцінка знизу $\alpha_k$ для значення $f_0(x^*)$ такі, що:

$$f_0\left(x^k\right)\to f_0\left(x^*\right) \text{ при } k\to\infty;\quad \alpha_k\to f_0\left(x^*\right) \text{ при } k\to\infty;$$

$$\alpha_{k+1}\ge\alpha_k,\ k=0,1,\ldots;\ f_0\left(x^*\right)\ge\alpha_k,\ k=0,1,\ldots.$$

Для початку роботи алгоритму необхідна нижня оцінка $\alpha_0$ для значення $f_0\left(x^*\right)$.

***Алгоритм 1***

Початок. І. Вибрати константу $\alpha_0\le f_0\left(x^*\right)$ (тобто нижню оцінку $\alpha_0$ оптимального значення $f_0\left(x^*\right)$ задачі 0).

ІІ. Задати довільні ваги $w_j$, $j\in\Im$ і $\upsilon_l$, $l\in L$.

ІІІ. Для фіксованого числа $\alpha$ визначити функцію:

$$\varphi(x,\alpha)=\sum_{i\in K(x)}\left[f_i(x)-\alpha\right]^2+\sum_{i\in\Im(x)} w_i\left(g_i(x)\right)^2+\sum_{l\in L}\upsilon_l\left(h_l(x)\right)^2,$$

де $K(x)=\left\{i \mid f_i(x)>\alpha,\ i=\overline{1,m}\right\}$.

IV. Покласти $k = 0$.

О с н о в н и й  ц и к л .  V. Знайти точку $x^k$, яка задовольняє умову

$$\varphi(x^k, \alpha_k) = \min_{x \in R^n} \varphi(x, \alpha_k),$$

тобто розв'язати задачу безумовної мінімізації функції $\varphi(x, \alpha_k)$, яка є диференційованою по $x$.

VI. Обчислити

$$\alpha_{k+1} = \alpha_k + \left[\varphi(x^k, \alpha_k)/m\right]^{1/2}. \tag{7.1}$$

VII. Покласти $k = k+1$ і перейти на крок V.

***Теорема 1.*** *Нехай виконуються припущення 0 і нехай (**iii**) – $f_0(\bar{x}(u))$ неперервна функція аргументу $u$ в точці $u = 0$, де $\bar{x}(u)$ є розв'язком наступної задачі:*

$$\textit{знайти } \arg\min_{x} \max_{i \in [1:m]} f_i(x)$$

*при обмеженнях:*

$$g_j(x) \le u_j,\ \ j \in \Im;\ \ h_l(x) = u_l,\ \ l \in L,$$

*тут $u_j$, $j \in \Im$, $u_l$, $l \in L$ – компоненти вектора $u$.*

*Тоді алгоритм 1 породжує послідовності $\{x^k\}_{k=0}^{\infty}$, $\{\alpha_k\}_{k=0}^{\infty}$ такі, що:*

$$\alpha_k \to f_0(x^*),\ \ f_0(x^k) \to f_0(x^*),\ \ \varphi(x^k, \alpha_k) \to 0 \text{ при } k \to \infty.$$

**2. Модифікація основного методу**

В удосконаленому варіанті алгоритму 1 на деяких ітераціях обчислення величини $\alpha_{k+1}$ проводиться не за формулою (7.1), а за формулою

$$\alpha'_{k+1} = \alpha'_k + \varphi(x^k, \alpha_k) \Big/ \sum_{i \in K(x^k)} \left(f_i(x^k) - \alpha'_k\right), \tag{7.2}$$

при цьому справедлива нерівність $\alpha'_{k+1} \ge \alpha_{k+1}$, яка забезпечує прискорення збіжності $\{\alpha'_k\}_{k=1}^{\infty}$ до $f_0(x^*)$ (принаймні при $\alpha'_k$ близьких до $f_0(x^*)$).

***Алгоритм 2***

П о ч а т о к .  I–III. Кроки I–III такі ж, як в алгоритмі 1.

IV. Вибрати верхню оцінку $\bar{\alpha}_0$ оптимального значення $f_0(x^*)$ задачі 0.

V. Вибрати константу $\varepsilon_f > 0$, яка характеризує точність обчислення $f_0(x^*)$ і константу $\varepsilon_\varphi$, яка характеризує близькість $\varphi(x, \alpha)$ до нуля.

VI. Покласти $\tau_1 = \alpha_0$.

VII. Покласти $k = 1$.

О с н о в н и й  ц и к л .  VIII. Знайти точку $x^k$, яка задовольняє умову

$$\varphi\left(x^k,\tau_k\right)=\min_{x\in R^n}\varphi\left(x,\tau_k\right).$$

IX. Обчислити значення $t_1$ і $t_2$, відповідно,

$$t_1=\tau_k+\left[\varphi\left(x^k,\tau_k\right)/m\right]^{1/2};\ \ t_2=\tau_k+\varphi\left(x^k,\tau_k\right)\Big/\sum_{i\in K\left(x^k\right)}\left(f_i\left(x^k\right)-\tau_k\right).$$

X. Якщо $t_2\le\bar{\alpha}_0$, то покласти $\tau_{k+1}=t_2$ і перейти на крок XI; інакше покласти $\tau_{k+1}=t_1$ і перейти на крок XI.

XI. Покласти $\tau_0=\tau_{k+1}-\tau_k$; $\alpha_0=t_1$; $s=\varphi\left(x^k,\tau_k\right)$.

XII. Покласти $k=k+1$ і перейти на крок XIII.

XIII. Знайти точку $x^k$, яка задовольняє умову

$$\varphi\left(x^k,\tau_k\right)=\min_{x\in R^n}\varphi\left(x,\tau_k\right).$$

XIV. Якщо $\bar{\alpha}_0-\alpha_0\le\varepsilon_f$ або $\tau_0\le\varepsilon_f$, то зупинити обчислення (у цьому випадку знаходять наближене значення для $f_0\left(x^*\right)$, яке можна обчислити, наприклад, за правилом $f_0\left(x^*\right)\cong\left(\bar{\alpha}_0+\alpha_0\right)/2$); інакше перейти на крок XV.

XV. Якщо $\varphi\left(x^k,\tau_k\right)>0$, то перейти на крок IX; якщо $s<\varepsilon_\varphi$, то зупинити обчислення (у цьому випадку знаходять наближене значення для $f_0(x^*)$); інакше покласти $\bar{\alpha}_0=\tau_k$, $s=0$, $\tau_k=\alpha_0$ і перейти на крок XIII.

*Зауваження 2.* Алгоритм 2 базується на тому, що нерівність $\tau\ge f_0\left(x^*\right)$ справедлива тоді і тільки тоді, коли $\inf\limits_{x\in R^n}\varphi\left(x,\tau\right)=0$.



## 7.3. Сітковий метод послідовних наближень розв’язування неперервних мінімаксних задач

Задача 1. Знайти $\arg\min\limits_{x\in X}\max\limits_{y\in Y}\varphi\left(x,y\right)$ для заданої функції $\varphi:R^n\times R^m\to R^1$, обмеженої і замкненої множини $Y\subset R^m$, опуклої і замкненої множини $X\subset R^n$.

*Припущення 1.* Функція $\varphi\left(x,y\right)$ неперервна разом з $\nabla_x\varphi\left(x,y\right)$ по сукупності змінних в $X'\times Y$, де $X'$ – довільна відкрита множина, що містить $X$.

Нижче описується метод обчислення стаціонарних точок функції $\max\limits_{y\in Y}\varphi\left(x,y\right)$ на множині $X$, тобто точок $x^*$, для яких:

$$\inf_{x\in X}\max_{y\in K(x^*)}\left(\nabla_x\varphi\left(x^*,y\right),x-x^*\right)=0,$$

де $K\left(x\right)=\left\{y\,|\,\varphi\left(x,y\right)=\max\limits_{y\in Y}\varphi\left(x,y\right),\ \ y\in Y\right\}$.

***Алгоритм 1***

П о ч а т о к . І. Вибрати довільне початкове наближення $x^0 \in X$ .

ІІ. Покласти $k = 0$ .

О с н о в н и й  ц и к л . ІІІ. Побудувати сітку

$$Y_{N_k} = \left\{ y^i \mid i \in \left[0 : N_k\right], \quad y^i \in Y \right\},$$

що задовольняє умовам теореми 1.

IV. Визначити функції $\varphi_i : R^n \to R^1$, $i \in \left[0 : N_k\right]$ за правилом

$$\varphi_i\left(x\right) \triangleq \varphi\left(x, y^i\right), \ \ i \in \left[0 : N_k\right].$$

V. Використовуючи алгоритми п. 7.1, знайти хоча б одну стаціонарну точку $\bar{x}^k \in X$ функції $\max\limits_{i \in [0;N_k]} \varphi_i\left(x\right)$ на множині $X$ .

VI. Покласти $k = k + 1$ і перейти на крок ІІІ.

***Теорема 1.*** *Нехай виконується припущення 1 і, крім того:* ($\boldsymbol{i}$)– *множина* $Y$ *обмежена і замкнена;* ($\boldsymbol{ii}$)– *множина* $X$ *опукла і замкнена;* ($\boldsymbol{iii}$)– *послідовність сіток* $\left\{Y_{N_k}\right\}_{k=0}^{\infty}$ *всюди щільна на множині* $Y$, *тобто для довільного* $\varepsilon > 0$ *можна вказати такий індекс* $k_\varepsilon$, *що для всіх* $k > k_\varepsilon$ *відстань між довільною точкою* $y \in Y$ *і найближчою до неї точкою сітки* $Y_{N_k}$ *буде менша* $\varepsilon$ *. Тоді кожна гранична точка послідовності* $\left\{x^k\right\}_{k=0}^{\infty}$, *яка породжена алгоритмом 1, є стаціонарною точкою функції* $\max\limits_{y \in Y} \varphi\left(x, y\right)$ *на множині* $X$ .

## 7.4. Методи стохастичного квазіградієнта в задачі пошуку максиміну

З а д а ч а  0 . Знайти $x^* = \arg\max\limits_{x \in X} \min\limits_{y \in Y} \varphi(x, y)$ для заданої функції $\varphi : R^n \times R^m \to R^1$, компактної множини $Y$ простору $R^m$ та множини

$$X = \left\{ x \middle| f_i(x) \geq 0, \quad i = \overline{1, m}, \quad x \in R^n \right\}. \tag{7.3}$$

*Припущення 0.* ($\boldsymbol{i}$)– функція $\varphi(x, y)$ – обмежена та неперервно диференційована по $x$ на $R^n \times Y$ ; ($\boldsymbol{ii}$)– функції $f_i(x)$, $i = \overline{1, m}$ – неперервно диференційовані та обмежені знизу на $R^n$ .

Задача обчислення вектора $x^*$ та величини максиміну

$$u^* = \max_{x \in X} \min_{y \in Y} \varphi(x, y)$$

зводиться до максимізації по $(x, u) \in X \times U$ функції:

$$g_{\tau_1,\tau_2}(x,u,\alpha) = u - \alpha\int\limits_Y \left|\min\left\{0,\varphi(x,y)-u\right\}\right|^{\tau_1} d\mu(y) -$$

$$-\sum_{i=1}^{m} p_i \left|\min\left\{0, f_i(x)\right\}\right|^{\tau_2},$$

де $\tau_1,\ \tau_2,\ p_i > 0;\ \sum\limits_{i=1}^{m} p_i = 1;\ \alpha -$ коефіцієнт штрафу; $\mu -$ деяка міра на множині $Y$; $U -$ будь-який відрізок, який містить точку $u^*$.

Функцію $g_{\tau_1,\tau_2}(x,u,\alpha)$ можна розглядати як математичне сподівання $E_{y,i} q_{\tau_1,\tau_2}(x,u,\alpha / y,i)$ функції

$$q_{\tau_1,\tau_2}(x,\ u,\ \alpha / y,\ i) = u - \alpha\left|\min\left\{0,\ \varphi(x,y)-u\right\}\right|^{\tau_1} -$$

$$-\alpha\left|\min\left\{0,\ f_i(x)\right\}\right|^{\tau_2}$$

від двох незалежних випадкових величин $y, i,$ причому $y$ розподілена на компакті $Y$ у відповідності до міри $\mu$, а $i$ приймає значення із множини $\Im = \{1,\ldots,m\}$ з ймовірностями $p_1, p_2,..., p_m$.

Для розв'язування задачі максимізації застосовується метод стохастичного градієнту. В алгоритмах із зростанням ітерацій коефіцієнт штрафу $\alpha$ прямує до нескінченності.

**1. Основний метод**

***Алгоритм 1***

П о ч а т о к . І. Вибрати довільне початкове наближення $(x^1, u^1) \in R^n \times R^1$.

II. Задати міру $\mu$ у просторі, що містить $Y$, таку, що будь-який непорожній перетин $Y$ з будь-якою відкритою множиною має додатну міру.

III. Вибрати параметри $\tau_1 \ge 2,\ \tau_2 \ge 2$ та константу $\alpha_0 > 0$ (рекомендується $\tau_1 = 2,\ \tau_2 = 2,\ \alpha_0 \in [1;10^2]$).

IV. Задати числа $p_i,\ i = \overline{1,m}$ $(p_i > 0,\ i = \overline{1,m},\ \sum\limits_{i=1}^{m} p_i = 1)$, які, відповідно, характеризують відносну точність виконання обмежень (7.3) у задачі 0 (рекомендується $p_i = 1/m,\ i = \overline{1,m}$).

V. Покласти $k = 1$.

О с н о в н и й ц и к л . VI. Знайти незалежну реалізацію $y^k$ випадкової величини $\tilde{y}$, що розподілена на компакті $Y$ у відповідності до міри $\mu$.

VII. Знайти незалежну реалізацію $i_k$ випадкової величини $\tilde{i}$, яка приймає значення з множини $\Im = \{1, ..., m\}$ з ймовірностями $p_1, p_2, ..., p_m$, відповідно.

VIII. Обчислити значення крокового множника $\rho_k$, яке задовольняє умову теореми 1.

IX. Обчислити приріст $\Delta_k$ коефіцієнта штрафу, який задовольняє умову теореми 1.

X. Покласти $\alpha_k = \alpha_{k-1} + \Delta_k$.

XI. Обчислити стохастичний градієнт $(\xi^k, \eta^k)$ функції $g_{\tau_1,\tau_2}(x, u, \alpha_k)$ в точці $(x^k, u^k)$:

$$\xi^k = \alpha_k \tau_1 \left| \min\left\{0, \varphi(x^k, y^k) - u^k\right\} \right|^{\tau_1 - 1} \nabla_x \varphi(x^k, y^k) + \\ + \alpha_k \tau_2 \left| \min\left\{0, f_{i_k}(x^k)\right\} \right|^{\tau_2 - 1} \nabla_x f_{i_k}(x^k); \tag{7.4}$$

$$\eta^k = 1 - \alpha_k \tau_1 \left| \min\left\{0, \varphi(x^k, y^k) - u^k\right\} \right|^{\tau_1 - 1}. \tag{7.5}$$

XII. Обчислити наступне наближення $(x^{k+1}, u^{k+1})$ за формулами:

$$x^{k+1} = x^k + \rho_k \xi^k;\ u^{k+1} = u^k + \rho_k \eta^k.$$

XIII. Покласти $k = k + 1$ і перейти на крок VI.

***Теорема 1.*** *Нехай виконуються припущення 0 та:* $(\boldsymbol{iii})$– $\nabla_x \varphi(x, y)$ *обмежений на* $R^n \times Y$ *і задовольняє умову Ліпшиця по* $x$ *на* $R^n$ *рівномірно відносно* $y \in Y$; $(\boldsymbol{iv})$– $\nabla_x f_i(x)$, $i = \overline{1, m}$ *обмежені та задовольняють умову Ліпшиця на* $R^n$; $(\boldsymbol{v})$– *числові послідовності* $\{\rho_k\}_{k=1}^{\infty}$, $\{\Delta_k\}_{k=1}^{\infty}$ *в алгоритмі 1 такі, що:*

$$\rho_k > 0,\ \sum_{k=1}^{\infty} \rho_k = \infty,\ \lim_{k\to\infty} \alpha_k = \infty,\ \lim_{k\to\infty} \Delta_k = 0,\ \Delta_k > 0;$$

$$\lim_{k\to\infty} (\Delta_k / \rho_k) = 0,\ \lim_{k\to\infty} (\rho_k \alpha_k^3) = 0 \quad (\rho_k = k^{-1/2},\ \Delta_k = k^{-1}).$$

1. *Тоді для будь-якого початкового наближення* $(x^1, u^1)$ *послідовність* $\left\{(x^k, u^k)\right\}_{k=1}^{\infty}$, *яка породжена алгоритмом 1, містить підпослідовність таку, що всі її скінченні граничні точки* $(x', u')$ *з ймовірністю 1 задовольняють наступним умовам: існують числа*

$$\beta_s,\ \ s = \overline{1, n+1},\ \ \lambda_i \geq 0,\ \ i = \overline{1, m},$$

*і точки*

$$y_s,\ \ s = \overline{1, n+1},\ \ y_s \in \left\{\arg\min_{y \in Y} \varphi(x', y)\right\},$$

*для яких справедливі співвідношення:*

$$u' = \min_{y \in Y} \varphi(x', y);$$

$$\sum_{s=1}^{n+1} \beta_s + \sum_{i=1}^{m} \lambda_i > 0; \qquad (7.6)$$

$$\sum_{s=1}^{n+1} \beta_s \nabla_x \varphi(x', y_s) + \sum_{i=1}^{m} \lambda_i \nabla_x f_i(x') = 0.$$

2. *Якщо, крім цього,*

$$0 \notin \mathrm{co}\{\nabla_x f_i(x),\ i \in \Im^-(x)\}, \qquad (7.7)$$

*де* $\Im^-(x) = \{i \mid f_i(x) \le 0,\ i \in \Im\}$,

*то* $x' \in X$, $\sum_{s=1}^{n+1} \beta_s > 0$ *і справедливі умови доповнюючої нежорсткості*

$$\lambda_i f_i(x') = 0, \quad i = \overline{1,m},$$

*тобто точка* $x'$ *є стаціонарною в задачі 0.*

*Зауваження 1.* Умова (7.7) виконується для $x \in R^n$, якщо вектори $\nabla_x f_i(x),\ i \in \Im^-(x)$ лінійно незалежні для будь-яких $x$ таких, що $\Im^-(x) \ne \varnothing$. Дана умова виконується також у тому випадку, коли функції $f_i(x)$ опуклі догори та задовольняють умову Слейтера в просторі $R^n$.

**2. Модифікація основного методу**

Для практичного знаходження стаціонарних точок у задачі 0 за допомогою алгоритму 1 необхідно виділяти спеціальну підпослідовність, усі граничні точки якої є стаціонарними. У наведеному нижче алгоритмі необхідна підпослідовність виділяється при реалізації ітераційного процесу.

***Алгоритм 2***

Початок. I–IV. Кроки I–IV такі, як і в алгоритмі 1.

V. Вибрати довільне значення $\varepsilon > 0$.

VI. Вибрати довільний вектор $v^1 \in R^{m+1}$ (як правило, у якості $v^1$ обирають наближене значення градієнту функції $g_{\tau_1,\tau_2}(x,u,\alpha_1)$ по $(x,u)$ у точці $(x^1,u^1)$).

VII. Покласти $k = 1,\ j = 1$.

Основний цикл. VIII. Якщо $\|v^k\| < \varepsilon$, то покласти $k_j = k,\ \varepsilon = \varepsilon / 2,\ j = j + 1$ та перейти на крок IX; якщо $\|v^k\| \ge \varepsilon$, то перейти на крок IX.

IX. Знайти незалежну реалізацію $y^k$ випадкової величини $\tilde{y}$, що розподілена на компакті $Y$ у відповідності до міри $\mu$.

X. Знайти незалежну реалізацію $i_k$ випадкової величини $\tilde{i}$, яка приймає значення з множини $\Im = \{1,...,m\}$ з ймовірностями $p_1, p_2,..., p_m$, відповідно.

XI. Обчислити значення крокових множників $\rho_k$, $\rho'_k$ та значення коефіцієнта штрафу $\alpha_k$, які задовольняють умовам теореми 2.

XII. Обчислити стохастичний градієнт $\zeta^k = (\xi^k, \eta^k)$ функції $g_{\tau_1,\tau_2}(x,u,\alpha_k)$ в точці $(x^k,u^k)$ згідно (7.4), (7.5).

XIII. Обчислити наступні наближення:

$$x^{k+1} = x^k + \rho_k \xi^k;\ u^{k+1} = u^k + \rho_k \eta^k .$$

XIV. Обчислити вектор

$$v^{k+1} = v^k + \rho'_k(\zeta^k - v^k) .$$

XV. Покласти $k = k+1$ та перейти на крок VIII.

***Теорема 2.*** *Нехай виконуються умови припущення 0, умови* (***iii***), (***iv***) *теореми 1 та числові послідовності* $\{\rho_k\}_{k=1}^{\infty}$, $\{\rho'_k\}_{k=1}^{\infty}$, $\{\Delta_k\}_{k=1}^{\infty}$ *такі, що:*

$$\rho_k > 0,\ \lim_{k\to\infty}\rho_k = 0,\ \sum_{k=1}^{\infty}\rho_k = \infty;$$

$$\lim_{k\to\infty}\alpha_k = \infty;\ \ \alpha_{k+1} = \alpha_k + \Delta_k;\ \ \Delta_k > 0,\ \ \lim_{k\to\infty}\Delta_k = 0,\ \ \sum_{k=1}^{\infty}(\Delta_k)^2 < \infty;$$

$$\lim_{k\to\infty}\Delta_k\alpha_k = 0,\ \lim_{k\to\infty}\rho'_k = 0,\ \sum_{k=1}^{\infty}\rho'_k = \infty,\ \lim_{k\to\infty}\Delta_k / \rho_k = 0;\ \lim_{k\to\infty}\rho_k(\alpha_k)^3 = 0,$$

$$\sum_{k=1}^{\infty}(\rho'_k\alpha_k)^2 < \infty;\ \sum_{k=1}^{\infty}(\rho'_k)^2\alpha_k(\sum_{s=1}^{k}\rho'_s\alpha_s) < \infty,\ \ \sum_{k=1}^{\infty}(\rho'_k)^2(\sum_{s=1}^{k}\rho'_s\alpha_s)^2 < \infty;$$

$$\sum_{k=1}^{\infty}(\rho_k^2\alpha_k^4 / \rho'_k) < \infty,\ \rho_k\alpha_k^4 / \rho'_k > \rho_{k+1}\alpha_{k+1}^4 / \rho'_{k+1}.$$

*Тоді для будь-яких векторів* $(x^1,u^1)$, $v^1$:

1) *послідовність* $\left\{(x^k,u^k)\right\}_{k=1}^{\infty}$, *яка породжена алгоритмом 2, містить таку підпослідовність* $\left\{(x^{s_i},u^{s_i})\right\}_{i=1}^{\infty}$, *що майже напевно*

$$\lim_{i\to\infty}\nabla_{(x,u)}g_{\tau_1,\tau_2}(x^{s_i},u^{s_i},\alpha_{s_i}) = 0.$$

*Будь-яка гранична точка* $(x',u')$ *цієї підпослідовності задовольняє умовам (7.6) і є стаціонарною в задачі 0, якщо* $0 \notin \mathrm{co}\left\{\nabla_x f_i(x'),\ i \in \Im^-(x')\right\}$;

2) $\lim_{k\to\infty}\left\|v^k - \nabla_{(x,u)}g_{\tau_1,\tau_2}(x^k,u^k,\alpha_k)\right\| = 0$ *майже напевно;*

3) *гранична точка* $(\overline{x},\overline{u})$ *послідовності* $\left\{(x^{k_j},u^{k_j})\right\}_{j=1}^{\infty}$, *яка породжена алгоритмом 2, є стаціонарною точкою в задачі 0.*

Наведені вище умови будуть виконуватися, якщо, наприклад, покласти:

$$\rho_k = 1/k;\ \rho_k' = k^{-12/13};\ \Delta_k = (k\ln k)^{-1},$$

або

$$\rho_k = k^{-49/50};\ \rho_k' = k^{-21/25};\ \Delta_k = k^{-99/100}.$$

### 3. Опуклий випадок

Задача 3. Знайти $\arg\max_{x\in X}\min_{y\in Y}\varphi(x,y)$ для заданої функції $\varphi: R^n\times R^m \to R^1$, компактної множини $Y\subset R^m$ та множини

$$X \triangleq \left\{x\middle| f_i(x)\ge 0,\ i=\overline{1,m};\ x\in Z\right\},$$

де $Z-$ опуклий компакт в $R^n$; $f_i: R^n\to R^1$, $i=\overline{1,m}$ – опуклі догори функції на $Z$.

***Алгоритм 3***

Початок. I–V. Кроки I–V такі, як і в алгоритмі 1.

Основний цикл. VI. Знайти незалежну реалізацію $y^k$ випадкової величини $\tilde{y}$, що розподілена на компакті $Y$ відповідно до міри $\mu$.

VII. Знайти незалежну реалізацію $i_k$ випадкової величини $\tilde{i}$, яка приймає значення з множини $\Im=\{1,...,m\}$ з ймовірностями $p_1,p_2,...,p_m$, відповідно.

VIII. Знайти значення крокового множника $\rho_k$ та коефіцієнта штрафу $\alpha_k$, які задовольняють умовам теореми 3.

IX. Обчислити стохастичний градієнт $(\xi^k,\eta^k)$ функції $g_{\tau_1,\tau_2}(x,u,\alpha_k)$ в точці $(x^k,u^k)$ за формулами (7.4), (7.5).

X. Обчислити наступне наближення $(x^{k+1},u^{k+1})$:

$$x^{k+1}=\pi_Z(x^k+\rho_k\xi^k);\ u^{k+1}=\pi_U(u^k+\rho_k\eta^k),$$

де $\pi_Q$ – оператор проектування на $Q$; $U-$ достатньо великий відрізок, який містить

$$[\min_{(x,y)\in Z\times Y}\varphi(x,y), \max_{(x,y)\in Z\times Y}\varphi(x,y)].$$

XI. Покласти $k=k+1$ і перейти на крок VI.

***Теорема 3.*** *Нехай виконуються умови:* $(\boldsymbol{i})-$ $Z-$ *опуклий компакт,* $Y-$ *компакт;* $(\boldsymbol{ii})-$ *функції* $\varphi(x,y)$, $\nabla_x\varphi(x,y)$ *неперервні на* $Z\times Y$; $(\boldsymbol{iii})-$ $\varphi(x,y)$ *опукла догори по* $x$ *на* $Z$ *для будь-якого* $y\in Y$; $(\boldsymbol{iv})-$ *функції*

$f_i(x),\ i=\overline{1,m}$, *мають неперервні частинні похідні та опуклі догори на* $Z$; $(\boldsymbol{v})-$ *числові послідовності* $\{\rho_k\}_{k=1}^{\infty}$, $\{\alpha_k\}_{k=1}^{\infty}$ *такі, що:*

$$\rho_k > 0,\ \lim_{k\to\infty}\rho_k = 0,\ \ \rho_{k+1} \le \rho_k,\ \ \sum_{k=1}^{\infty}\rho_k = \infty;$$

$$\alpha_k > 0,\ \lim_{k\to\infty}\alpha_k = \infty,\ \alpha_{k+1} \ge \alpha_k,\ \sum_{k=1}^{\infty}(\rho_k\alpha_k)^2 < \infty$$

*(умови* $(\boldsymbol{v})$ *будуть виконуватися, якщо покласти* $\rho_k = k^{-8/4},\ \alpha_k = k^{1/5},\ k=1,2,...$*).*

*Тоді для будь-якого початкового наближення* $(x^1,u^1)$ *послідовність* $\left\{(x^k,u^k)\right\}_{k=1}^{\infty}$*, яка породжена алгоритмом 3, з ймовірністю 1 збігається до множини* $G^*$ *розв'язків задачі 3:*

$$G^* \triangleq \left\{\arg\max_{x\in X}\min_{y\in Y}\varphi(x,y)\right\}\times\{u^*\}.$$

*Зауваження 3.* Якщо у максимінній задачі 3 $Z\equiv R^n$ і функції $\varphi(x,y),\ \ f_i(x),\ \ i=\overline{1,m}$ задовольняють умовам теореми 1, то на кроці X алгоритму 3 точку $(x^{k+1},u^{k+1})$ можна обчислювати за формулами:

$$x^{k+1} = x^k + \rho_k\xi^k;\ \ u^{k+1} = u^k + \rho_k\eta^k,$$

де $\rho_k$, $\alpha_k$ обчислюють у відповідності до вимог теореми 3.

*Зауваження 3'.* Алгоритми, наведені в даному параграфі, порівняно прості для реалізації, стійкі відносно похибок обчислень; їх ефективність порівняно з іншими методами розв'язування задачі 0 зростає із збільшенням розмірностей змінних $x$ та $y$. Розв'язування тестових задач на ЕОМ свідчить про більш швидку збіжність алгоритмів по змінній $x$, ніж по змінній $u$, тому рекомендується один крок по змінній $x$ чергувати з декількома кроками по змінній $u$.

## 7.5. Квазіградієнтні методи розв'язування неперервних мінімаксних задач стохастичного програмування

З а д а ч а 0. Знайти $\arg\min_{x\in X}\max_{y\in Y}E\varphi(x,y,\omega)$ для заданої функції $\varphi: R^n\times R^m\times\Omega\to R^1$ і заданих множин $X\subset R^n$, $Y\subset R^m$.

*Припущення 0.* $(\boldsymbol{i})-$ функція $f(x,y)\triangleq E\varphi(x,y,\omega)$ і її градієнт по $x$ неперервні по сукупності змінних $x,y$; $(\boldsymbol{ii})-X-$ опукла, замкнута, обмежена множина; $(\boldsymbol{iii})-Y-$ замкнута, обмежена множина; $(\boldsymbol{iv})-$ маємо можливість обчислювати тільки значення функції $\varphi(x, y, \omega)$ і її градієнта $\nabla_x \varphi(x, y, \omega)$ в кожній точці $(x, y, \omega) \in X \times Y \times \Omega$.

Наведені далі методи розв’язування задачі 0 базуються на побудові спеціальних множин $Y_k$ для визначення на $k$-й ітерації напрямку руху $h^k$ до $(k+1)$-го наближення $x^{k+1}$. Вектор $h^k$ визначається як стохастичний квазіградієнт функції $\max\limits_{y \in Y_k} E\varphi(x, y, \omega)$ в точці $x^k$.

### 1. Стохастичний квазіградієнтний метод

***Алгоритм 1***

Початок. I. Вибрати початкове наближення $x^0 \in X$.

II. Вибрати натуральне число $N_s$, покласти $N = N_s$.

III. Вибрати довільні дійсні числа $z_i^0$, $i \in [0:N]$.

IV. Побудувати сітку

$$Y_N = \left\{ y^i \mid y^i \in Y,\ i \in [0:N] \right\}.$$

V. Покласти $k = 0$.

Основний цикл. VI. Знайти індекс $i_k$ з умови

$$z_{i_k}^k = \max_{i \in [0:N]} z_i^k .$$

VII. Обчислити незалежну реалізацію $\omega^k$ випадкового параметра $\omega$.

VIII. Обчислити $\nabla_x \varphi(x^k(\omega), y^{i_k}, \omega^k)$.

IX. Обчислити крокові множники $\rho_k$ і $\sigma_k$, які задовольняють умовам теореми 1.

X. Обчислити наступне наближення

$$x^{k+1}(\omega) = \pi_X(x^k(\omega) - \rho_k \nabla_x \varphi(x^k(\omega), y^{i_k}, \omega^k)).$$

XI. Обчислити значення функції $\varphi(x^{k+1}(\omega), y^i, \omega^k)$, $i \in \left[0:N\right]$.

XII. Обчислити величини:

$$z_i^{k+1}(\omega) = z_i^k(\omega) + \sigma_k(\varphi(x^{k+1}(\omega), y^i, \omega^k) - z_i^k(\omega)),\ i \in \left[0:N\right].$$

XIII. Покласти $k = k+1$ і перейти на крок VI.

***Теорема 1.*** *Якщо виконані припущення 0 та* $(\boldsymbol{i})-$ *функція* $E\varphi(x, y, \omega)$ *опукла донизу по* $x$, $x \in X$ *при кожному* $y \in Y$; $(\boldsymbol{ii})-$ *градієнт функції* $f(x, y) \triangleq E\varphi(x, y, \omega)$ *по* $x$ *задовольняє умову Ліпшиця*

$$\left\| \nabla_x f(x, y) - \nabla_x f(\bar{x}, y) \right\| \le \alpha_1 \left\| x - \bar{x} \right\|,\quad y \in Y,\quad x,\ \bar{x} \in X,\ \alpha_1 < \infty;$$

$(\boldsymbol{iii})-$

$$E\left|\varphi\left(x, y, \omega\right)\right|^2 < \infty,\ \ E\left\|\nabla_x \varphi(x, y, \omega)\right\|^2 < \infty,\quad y \in Y,\ x \in X;$$

$(\boldsymbol{iv})-$ *крокові множники* $\rho_k$ *і* $\sigma_k$ *задовольняють умовам:*

$$\sum_{k=0}^{\infty}\sigma_k^2 < \infty,\ \sum_{k=0}^{\infty}\rho_k = \infty,\ \rho_k/\sigma_k \to 0,\ \rho_{k+1}/\rho_k \to 1\ \ при\ k \to \infty,$$

*то майже при кожному $\omega$ граничні точки $\overline{x}^{N_s}$ послідовності $\left\{x^k(\omega)\right\}_{k=0}^{\infty}$, породженої алгоритмом 1, являються точками мінімуму функції $\max\limits_{i\in[0:N]} E\varphi(x, y^i, \omega)$ на множині $X$.*

Для розв’язування початкової задачі 0, необхідно при кожному фіксованому $N_s$ ($N_s \to \infty$ при $s \to \infty$) розв’язувати дискретну мінімаксну задачу за допомогою алгоритму 1, причому послідовність сіток $\left\{Y_{N_s}\right\}_{s=0}^{\infty}$ повинна «мати властивість щільності» на множині $Y$. Кожна гранична точка послідовності $\left\{\overline{x}^{N_s}\right\}_{s=0}^{\infty}$ являється розв’язком задачі 0.

Послідовність сіток $\left\{Y_{N_s}\right\}_{s=0}^{\infty}$ «має властивість щільності» на множині $Y$, якщо для кожного $\varepsilon > 0$ можна вказати такий номер $\overline{N}$, що для всіх $N_s > \overline{N}$ відстань між довільною точкою $\overline{y} \in Y$ і найближчою до $\overline{y}$ точкою сітки $Y_{N_s}$ буде менше $\varepsilon$.



**2. Модифікований стохастичний квазіградієнтний метод**

***Алгоритм 2***

Початок. I. Вибрати початкове наближення $x^0 \in R^n$.

II. Вибрати натуральне число $N_0$ і побудувати сітку

$$Y_{N_0} = \left\{y^i \mid y^i \in Y,\ i \in \left[0 : N_0\right]\right\}.$$

III. Вибрати довільні дійсні числа $z_i^0$, $i \in [0 : N_0]$.

IV. Вибрати початкове значення крокового множника $\rho_0 > 0$ і константи $\alpha > 0$, $\beta > 0$.

V. Покласти $k = 0$.

Основний цикл. VI. Знайти індекс $i_k$ з умови

$$z_{i_k}^k(\omega) = \max_{i\in[0:N_k]} z_i^k(\omega).$$

VII. Знайти незалежну реалізацію $\omega^k$ випадкового параметра $\omega$.

VIII. Обчислити $\nabla_x(\varphi(x^k(\omega), y^{i_k}, \omega^k))$.

IX. Обчислити наступне наближення

$$x^{k+1}(\omega) = \pi_X(x^k(\omega) - \rho_k \nabla_x \varphi(x^k(\omega), y^{i_k}, \omega^k)).$$

X. Обчислити значення $\varphi(x^{k+1}(\omega), y^i, \omega^k)$, $i \in \left[0 : N_k\right]$.

XI. Знайти крокові множники $\rho_{k+1}$ і $\sigma_{k+1}$, які задовольняють умовам теореми 2.

XII. Побудувати сітку

$$Y_{N_{k+1}} = \left\{ y^i \mid y^i \in Y,\ i \in \left[0 : N_{k+1}\right] \right\}$$

так, щоб вона містила усі точки сітки $Y_{N_k}$ із збереженням їх нумерації і щоб виконувалася наступна умова: для довільного $\bar{y} \in Y$ відстань від $\bar{y}$ до найближчої точки сітки $Y_{N_{k+1}}$ не перевищувала значення $\alpha\rho_{k+1}^{1+\beta}$.

XIII. Обчислити величини

$$z_i^{k+1}(\omega) = z_i^k(\omega) + \sigma_{k+1}(\varphi(x^{k+1}(\omega), y^i, \omega^k) - z_i^k(\omega)),\ i \in \left[0 : N_k\right].$$

XIV. Покласти $i = N_k + 1$.

XV. Знайти точку $y^{j_i}$ сітки $Y_{N_k}$, яка належить кулі радіусом $\alpha\rho_{k+1}$ з центром у точці $y^i$ і має найменший індекс $j_i \in [0 : N_k]$.

XVI. Покласти $z_i^{k+1}(\omega) = z_{j_i}^k(\omega)$.

XVII. Якщо $i < N_{k+1}$, то покласти $i = i + 1$ і перейти на крок XV; інакше перейти на крок XVIII.

XVIII. Покласти $k = k + 1$ і перейти на крок VI.

***Теорема 2.*** *Якщо виконані умови теореми 1 та:* $(\boldsymbol{i})-$ *функція* $E\varphi(x, y, \omega)$ *для довільного* $x \in X$ *задовольняє умову Ліпшиця по змінній* $y$

$$\left|E\varphi(x, y, \omega) - E\varphi(x, \bar{y}, \omega)\right| \le \alpha_1 \left\|y - \bar{y}\right\|,\ \alpha_1 < \infty;$$

$(\boldsymbol{ii})-$ *послідовність* $\left\{\sigma_k\right\}_{k=0}^{\infty}$ *додатково задовольняє умову*

$$\lim_{s\to\infty} \sum_{j=s}^{m} \sigma_j = \infty \quad \text{при } m - s \to \infty,$$

(*наприклад,* $\sigma_k = k^{-\nu},\ \nu \in (1/2;1)$ ),

*то майже при кожному* $\omega$ *граничні точки послідовності* $\left\{x^k(\omega)\right\}_{k=0}^{\infty}$, *породженої алгоритмом 2, являються точками мінімуму функції* $\max_{y\in Y} E\varphi(x, y, \omega)$ *на множині* $X$.

### 7.6. Градієнтні методи знаходження сідлових точок

Задача 0. Знайти точку $(x^*,\ y^*) \in X \times Y$, яка задовольняє нерівності:

$$\varphi(x,\ y^*) \le \varphi(x^*,\ y^*) \le \varphi(x^*,\ y),\quad \forall x \in X,\ y \in Y, \tag{7.8}$$

де $\varphi : R^n \times R^m \to R^1$; $X$ та $Y$ – опуклі замкнуті тілесні підмножини евклідових просторів $R^n$ та $R^m$, відповідно.

*Означення 1.* Множина точок $(x^*, y^*)$, які задовольняють (7.8), називають ***множиною сідлових точок*** $X^* \times Y^*$ функції $\varphi(x, y)$ на множині $X \times Y$.

*Припущення 0.* $(\boldsymbol{i})-$ функція $\varphi(x, y)-$ неперервно диференційована на множині $X \times Y$, опукла догори по $x$ і опукла донизу по $y$; $(\boldsymbol{ii})-$ множина сідлових точок $X^* \times Y^*$ функції $\varphi$ непорожня і обмежена; $(\boldsymbol{iii})-$ виконується умова стійкості множини сідлових точок по $x$, тобто для довільного $x \in (X \setminus X^*)$ виконується строга нерівність

$$\varphi(x,\ y^*) < \varphi(x^*,\ y^*).$$

Припущення 0 виконуються для функції Лагранжа задачі опуклого програмування:

$$\text{знайти} \quad \arg\max_{x \in X_1} f_0(x), \tag{7.9}$$

де $X_1 = \{x \mid f_i(x) \geq 0,\ \ i = \overline{1,m},\ \ x \in R^n\}$, із диференційованими (та опуклими догори) функціями $f_j,\ j = \overline{0,m},$ якщо функція $f_0$ строго опукла і виконується умова Слейтера.

Умова $(\boldsymbol{iii})$ відіграє важливу роль, так як при її відсутності градієнтні методи пошуку сідлових точок, в загальному, не збігаються. Прикладом є функція

$$\varphi(x,\ y) = (1-y)x,\ \ X = R^1,\ \ Y = R^1_+,$$

яка являється функцією Лагранжа задачі:

$$\text{знайти} \quad \arg\max_{x \in X_2} x,$$

де $X_2 = \{x \mid -x \geq 0, \quad x \in R^1\}$; $R^m_+ -$ підпростір $m$-вимірних векторів із невід'ємними компонентами.

#### 1. Основний метод

***Алгоритм 1***

П о ч а т о к . I. Вибрати початкове наближення $(x^0,\ y^0) \in\ X \times Y$.

II. Покласти $k = 0$.

О с н о в н и й  ц и к л . III. Вибрати кроковий множник $\rho_k$, який задовольняє умовам теореми 1.

IV. Обчислити вектори $\nabla_x \varphi(x^k,\ y^k)$ і $\nabla_y \varphi(x^k,\ y^k)-$ градієнти, відповідно, по $x$ та $y$ функції $\varphi(x,\ y)$ в точці $(x^k,\ y^k)$.

V. Обчислити наступні наближення:

$$x^{k+1} = \pi_X(x^k + \rho_k \nabla_x \varphi(x^k,\ y^k));$$
$$y^{k+1} = \pi_Y(y^k - \rho_k \nabla_y \varphi(x^k,\ y^k)),$$

де $\pi_Q -$ оператор проектування на множину $Q$.

VI. Покласти $k = k+1$ і перейти на крок III.

***Теорема 1.*** *Нехай виконуються припущення 0. Тоді для довільного додатного числа $\delta$ існує таке додатне число $\rho(\delta)$, що при довільному початковому наближенні $(x^0,\ y^0) \in Z(\delta)$,*

$$Z(\delta) \triangleq \{(x,\ y) \mid \min_{(x^*,\ y^*) \in X^* \times Y^*} \|(x,\ y) - (x^*,\ y^*)\| \le \delta,\ (x,\ y) \in X \times Y\}, \quad (7.10)$$

*і послідовності $\{\rho_k\}_{k=0}^{\infty}$, яка задовольняє умовам:*

$$\rho_k > 0;\ \rho_k \to 0 \ \textit{при}\ k \to \infty;\ \sum_{k=0}^{\infty} \rho_k = \infty;\ \rho_k \le \rho(\delta),$$

*алгоритм 1 породжує послідовність $\{(x^k,\ y^k)\}_{k=0}^{\infty}$, яка збігається до множини сідлових точок $X^* \times Y^*$.*

**Приклад 1.** Знайти сідлову точку $\left(x^*, y^*\right)$ функції $\varphi\left(x, y\right) = -2x^2 + 4x + 4y^2 - y + 3xy - 1$ на множині $X \times Y$, де $X = \left\{x \mid 0 \le x \le 2\right\}$, $Y = \left\{y \mid -2 \le y \le 0\right\}$.

*Розв'язування.* Для знаходження наближеного значення сідлової точки використаємо 5 ітерацій алгоритму 1. Спочатку знайдемо градієнти цільової функції $\varphi\left(x, y\right)$ по змінних $x$ і $y$:

$$\nabla_x \varphi\left(x, y\right) = -4x + 3y + 4;\ \nabla_y \varphi\left(x, y\right) = 8y + 3x - 1.$$

**Алгоритм 1**

I. Вибираємо початкове наближення $\left(x^0; y^0\right)^T = \left(1; -1\right)^T$.

II. Покладемо $k = 0$.

*1-а ітерація:*

III. Задаємо $\rho_0 = 0,5$.

IV. Обчислюємо градієнти:

$$\nabla_x \varphi\left(x^0, y^0\right) = -4 \cdot 1 + 3 \cdot \left(-1\right) + 4 = -3;\ \nabla_y \varphi\left(x^0, y^0\right) = 8 \cdot \left(-1\right) + 3 \cdot 1 - 1 = -6.$$

V. Обчислюємо наступні наближення:

$$x^1 = \pi_X\left(x^0 + \rho_0 \cdot \nabla_x \varphi\left(x^0, y^0\right)\right) = \pi_X\left(1 + 0,5 \cdot \left(-3\right)\right) = \pi_X\left(-0,5\right) = 0;$$

$$y^1 = \pi_Y\left(y^0 - \rho_0 \cdot \nabla_y \varphi\left(x^0, y^0\right)\right) = \pi_Y\left(-1 - 0,5 \cdot \left(-6\right)\right) = \pi_Y\left(2\right) = 0.$$

VI. Покладемо $k = 0 + 1 = 1$ і переходимо на крок III.

*2-а ітерація:*

III. Задаємо $\rho_1 = 0,4$.

IV. Обчислюємо градієнти:

$$\nabla_x \varphi\left(x^1, y^1\right) = -4\cdot 0 + 3\cdot 0 + 4 = 4; \quad \nabla_y \varphi\left(x^1, y^1\right) = 8\cdot 0 + 3\cdot 0 - 1 = -1.$$

V. Обчислюємо наступні наближення:

$$x^2 = \pi_X\left(0 + 0,4\cdot 4\right) = \pi_X\left(1,6\right) = 1,6; \quad y^2 = \pi_Y\left(0 - 0,4\cdot\left(-1\right)\right) = \pi_Y\left(0,4\right) = 0.$$

VI. Покладемо $k = 1 + 1 = 2$ і переходимо на крок III.

*3-я ітерація:*

III. Задаємо $\rho_2 = 0,3$.

IV. Обчислюємо градієнти:

$$\nabla_x \varphi\left(x^2, y^2\right) = -4\cdot 1,6 + 3\cdot 0 + 4 = -2,4;$$

$$\nabla_y \varphi\left(x^2, y^2\right) = 8\cdot 0 + 3\cdot 1,6 - 1 = 3,8.$$

V. Обчислюємо наступні наближення:

$$x^3 = \pi_X\left(1,6 + 0,3\cdot\left(-2,4\right)\right) = \pi_X\left(0,88\right) = 0,88;$$

$$y^3 = \pi_Y\left(0 - 0,3\cdot 3,8\right) = \pi_Y\left(-0,114\right) = -0,114.$$

VI. Покладемо $k = 2 + 1 = 3$ і переходимо на крок III.

*4-а ітерація:*

III. Задаємо $\rho_3 = 0,1$.

IV. Обчислюємо градієнти:

$$\nabla_x \varphi\left(x^3, y^3\right) = -4\cdot 0,88 + 3\cdot\left(-0,114\right) + 4 = 0,138;$$

$$\nabla_y \varphi\left(x^3, y^3\right) = 8\cdot\left(-0,114\right) + 3\cdot 0,88 - 1 = 0,728.$$

V. Обчислюємо наступні наближення:

$$x^4 = \pi_X\left(0,88 + 0,1\cdot 0,138\right) = \pi_X\left(0,894\right) = 0,894;$$

$$y^4 = \pi_Y\left(-0,114 - 0,1\cdot 0,728\right) = \pi_Y\left(-0,187\right) = -0,187.$$

VI. Покладемо $k = 3 + 1 = 4$ і переходимо на крок III.

*5-а ітерація:*

III. Задаємо $\rho_4 = 0,05$.

IV. Обчислюємо градієнти:

$$\nabla_x \varphi\left(x^4, y^4\right) = -4\cdot 0,894 + 3\cdot\left(-0,187\right) + 4 = -0,137;$$

$$\nabla_y \varphi\left(x^4, y^4\right) = 8\cdot\left(-0,187\right) + 3\cdot 0,894 - 1 = 0,186.$$

V. Обчислюємо наступні наближення:

$$x^5 = \pi_X\left(0{,}894 + 0{,}05\cdot\left(-0{,}137\right)\right) = \pi_X\left(0{,}887\right) = 0{,}887;$$

$$y^5 = \pi_Y\left(-0{,}187 - 0{,}05\cdot 0{,}186\right) = \pi_Y\left(-0{,}196\right) = -0{,}196.$$

VI. Покладемо $k = 4+1 = 5$ і зупиняємо обчислення.

Таким чином, за п'ять ітерацій алгоритму 1 знайшли наближене значення сідлової точки $\left(x^5; y^5\right)^T = \left(0{,}887; -0{,}196\right)^T$.

### 2. Градієнтний метод відшукання сідлових точок із постійним кроковим множником

***Алгоритм 2***

Початок. I. Вибрати довільне початкове наближення $(x^0,\ y^0) \in X \times Y$.

II. Вибрати постійний кроковий множник $\rho > 0$.

III. Покласти $k = 0$.

Основний цикл. IV. Обчислити вектори $\nabla_x \varphi(x^k,\ y^k)$ і $\nabla_y \varphi(x^k,\ y^k)$.

V. Обчислити наступні наближення:

$$x^{k+1} = \pi_X(x^k + \rho\nabla_x \varphi(x^k,\ y^k));\quad y^{k+1} = \pi_Y(y^k - \rho\nabla_y \varphi(x^k,\ y^k)),$$

де $\pi_Q -$ оператор проектування на множину $Q$.

VI. Покласти $k = k+1$ і перейти на крок IV.

***Теорема 2.*** *Нехай виконуються припущення 0. Тоді для довільних чисел $\varepsilon$ і $\delta$, які задовольняють нерівність $0 < \varepsilon < \delta$, існують такі додатні числа $\rho(\varepsilon,\ \delta)$ і $k_0(\varepsilon,\ \delta)$, що алгоритм 2 при довільному $(x^0,\ y^0) \in Z(\delta)$ і довільному сталому кроковому множнику $\rho \le \rho(\varepsilon,\ \delta)$ породжує послідовність $\{(x^k,\ y^k)\}_{k=0}^{\infty}$, для якої при всіх $k > k_0(\varepsilon,\ \delta)\rho^{-1}$ буде виконуватись $(x^k,\ y^k) \in Z(\varepsilon)$ (тут і далі множини $Z(\delta)$ і $Z(\varepsilon)$ визначаються згідно (7.10)).*

***Теорема 2'.*** *Нехай виконуються припущення 0 і нехай: $(\boldsymbol{iv})-$ множина сідлових точок $X^* \times Y^*$ функції $\varphi(x,\ y)$ складається з однієї точки $(x^*,\ y^*)$; $(\boldsymbol{v})-$ функція $\varphi(x,\ y)$ має в точці $(x^*,\ y^*)$ другу похідну; $(\boldsymbol{vi})-$ точка $(x^*,\ y^*)$ являється внутрішньою точкою множини $X \times Y$; $(\boldsymbol{vii})-$ не існує числа $\lambda$ і ненульового вектора $(u,\ v) \in R^n \times R^m$ таких, що:*

$$\nabla^2_{xx}\varphi(x^*,\ y^*)u = 0,\quad \nabla^2_{yy}\varphi(x^*,\ y^*)v = 0;$$

$$\nabla^2_{yx}\varphi(x^*,\ y^*)u = \lambda v,\quad \left(\nabla^2_{yx}\varphi(x^*,\ y^*)\right)^T v = \lambda u.$$

*Тоді для довільного числа $\delta > 0$ знайдеться таке число $\rho(\delta) > 0$, що алгоритм 2 при всякому початковому наближенні $(x^0,\ y^0) \in Z(\delta)$ і довільному сталому кроковому множнику $\rho \le \rho(\delta)$ утворює послідовність $\{(x^k,\ y^k)\}_{k=0}^{\infty}$, яка лінійно (зі швидкістю геометричної прогресії) збігається до точки $(x^*,\ y^*)$.*

**3. Узагальнений градієнтний метод**

Задача 3. Знайти точку $(x^*,\ y^*) \in X \times Y$, яка задовольняє умову

$$\varphi(x,\ y^*) \le \varphi(x^*,\ y^*) \le \varphi(x^*,\ y),\quad \forall x \in X,\ y \in Y,$$

де $\varphi(x, y): R^n \times R^m \to R^1$; $X$ та $Y-$ опуклі та замкнуті множини евклідових просторів $R^n$ і $R^m$, відповідно.

*Припущення 3.* $(\boldsymbol{i})-$ функція $\varphi(x,\ y)$ при довільному фіксованому $y \in Y$ опукла догори по $x$ в деякому околі $\tilde{X}$ множини $X$; функція $\varphi(x,\ y)$ при довільному фіксованому $x \in X$ опукла донизу по $y$ в деякому околі $\tilde{Y}$ множини $Y$; $(\boldsymbol{ii})-$ множина сідлових точок $X^* \times Y^*$ функції $\varphi(x,\ y)$ відносно множин $X$ та $Y$ непорожня і обмежена; $(\boldsymbol{iii})-$ виконується умова стійкості множини сідлових точок $X^* \times Y^*$ функції $\varphi(x,\ y)$, тобто для довільних $x^* \in X^*,\ y^* \in Y^*$ виконується:

$$X^* = \{x \mid \varphi(x,\ y^*) = \max_{x' \in X} \varphi(x',\ y^*),\quad x \in X\};$$

$$Y^* = \{y \mid \varphi(x^*,\ y) = \min_{y' \in Y} \varphi(x^*,\ y'),\quad y \in Y\}.$$

У наведеному нижче алгоритмі в якості векторів, які визначають напрямок руху по $x$ та $y$ на $k$-й ітерації, вибирають, відповідно, узагальнені градієнти функції $\varphi(x,\ y)$ по $x$ та $y$, обчислені в точці $(x^k,\ y^k)$. Крокові множники задовольняють класичним умовам.

***Алгоритм 3***

Початок. I. Вибрати довільне початкове наближення $(x^0,\ y^0) \in X \times Y$.

II. Покласти $k = 0$.

Основний цикл. III. Обчислити узагальнені градієнти $\hat{\nabla}_x \varphi(x^k,\ y^k)$ та $\hat{\nabla}_y \varphi(x^k,\ y^k)$ функції $\varphi(x,\ y)$ по $x$ та $y$ в точці $(x^k,\ y^k)$, тобто вектори, які задовольняють умовам:

$$\left(\hat{\nabla}_x \varphi(x^k,\ y^k),\ x' - x^k\right) \geq \varphi(x',\ y^k) - \varphi(x^k,\ y^k), \quad \forall x' \in \tilde{X};$$

$$\left(\hat{\nabla}_y \varphi(x^k,\ y^k),\ y' - y^k\right) \leq \varphi(x^k,\ y') - \varphi(x^k,\ y^k), \quad \forall y' \in \tilde{Y}.$$

IV. Обчислити кроковий множник $\rho_k$.

V. Обчислити наступні наближення $(x^{k+1},\ y^{k+1})$:

$$x^{k+1} = \pi_X(x^k + \rho_k \hat{\nabla}_x \varphi(x^k,\ y^k)); \quad y^{k+1} = \pi_Y(y^k - \rho_k \hat{\nabla}_y \varphi(x^k,\ y^k)).$$

VI. Покласти $k = k+1$ та перейти на крок III.

***Теорема 3.*** *Нехай виконуються припущення 3 і нехай:* $(\boldsymbol{iv})-$ *множини узагальнених градієнтів функції* $\varphi(x,\ y)$ *по* $x$ *та* $y$, *відповідно, рівномірно обмежені, тобто існує стала* $\alpha$ *така, що:*

$$\left\|\hat{\nabla}_x \varphi(x,\ y)\right\| \leq \alpha, \quad \left\|\hat{\nabla}_y \varphi(x,\ y)\right\| \leq \alpha, \ \forall (x,\ y) \in X \times Y;$$

$(\boldsymbol{v})-$ *крокові множники* $\rho_k$ *в алгоритмі 3 задовольняють умовам:*

$$\rho_k > 0, \quad k = 0, 1, \ldots; \quad \lim_{k\to\infty} \sum_{i=0}^{k} \rho_i = \infty.$$

*Тоді послідовність* $\left\{(x^k,\ y^k)\right\}_{k=0}^{\infty}$, *яка утворена алгоритмом 3, така, що*

$$\min_{(x^*,\ y^*)\in X^*\times Y^*} \left\|(x^k,\ y^k) - (x^*,\ y^*)\right\| \to 0 \ \text{ при } \ k \to \infty. \tag{7.11}$$

***Теорема 3′.*** *Нехай виконуються всі припущення теореми 3 за винятком умови* $(\boldsymbol{iv})$. *Тоді існує число* $\rho > 0$ *таке, що при виконанні додаткової умови*

$$\rho_k < \rho, \quad k = 0, 1, \ldots,$$

*послідовність* $\left\{(x^k,\ y^k)\right\}_{k=0}^{\infty}$, *яка породжена алгоритмом 3, задовольняє граничному співвідношенню (7.11).*

## Запитання та вправи для самостійної роботи до розділу 7.

1. Сформулюйте дискретну мінімаксну задачу з обмеженнями простої структури.
2. Запишіть означення стаціонарної точки функції $\max\limits_{i\in I} \varphi_i(x)$ на множині $X = \left\{x \middle| f_j(x) \leq 0,\ j \in J_1,\ x \in R^n\right\}$.
3. Запишіть умови, з яких визначається кроковий множник $\rho_k$ в другому методі послідовних наближень.
4. Сформулюйте неперервну мінімаксну задачу.

5. Дайте означення стаціонарної точки функції $\max\limits_{y \in Y} \varphi(x, y)$ на множині $X$ .
6. Яким умовам повинна задовольняти послідовність сіток $\left\{Y_{N_k}\right\}_{k=0}^{\infty}$ в методі послідовних наближень розв'язування неперервних мінімаксних задач?
7. В чому полягає сутність методів стохастичного квазіградієнта в задачі пошуку максиміну?
8. Сформулюйте неперервну мінімаксну задачу стохастичного програмування.
9. Сформулюйте означення множини сідлових точок $X^* \times Y^*$ функції $\varphi(x, y)$ на множині $X \times Y$ .
10. В яких напрямках здійснюється рух по змінних $x$ і $y$ :

    а) в градієнтних методах відшукання сідлових точок;

    б) в узагальнених градієнтних методах?
11. З яких умов визначаються крокові множники в узагальненому градієнтному методі відшукання сідлових точок?
12. Знайдіть сідлову точку $\left(x^*, y^*\right)$ функції $\varphi\left(x, y\right) = -3x^2 + 3x + 5y^2 - 2y + 1{,}5xy + 10$ на множині $X \times Y$, де $X = \left\{x \mid -1 \le x \le 3\right\}$, $Y = \left\{y \mid -3 \le y \le 5\right\}$.

**Розділ 8**

# МЕТОДИ ОПТИМІЗАЦІЇ В НЕСКІНЧЕННОВИМІРНИХ ПРОСТОРАХ

## 8.1. Диференціали, субдиференціали та їх властивості

Нехай $X$ та $Y$ лінійні нормовані простори над полем дійсних чисел, а $F$ відображення, яке діє із $X$ в $Y$ та визначене на деякій відкритій множині $U$ простору $X$. Позначимо через $L(X,Y)$ множину лінійних обмежених операторів, що переводять простір $X$ в $Y$.

*Означення 1.* Відображення $F$ називається ***диференційованим за Фреше*** в точці $x \in U$, якщо існує оператор $D_x \in L(X,Y)$ такий, що

$$F(x+h) - F(x) = D_x h + \alpha(x,h), \qquad (8.1)$$

для довільного $h$ такого, що $x+h \in U$, причому $\|h\|^{-1} \alpha(x,h) \to 0$ при $\|h\| \to 0$.

Відображення $F$ називається ***диференційованим*** на множині $U$, якщо воно диференційоване в кожній точці $U$.

Вираз $D_x h$ називають ***диференціалом Фреше***, а оператор $D_x$ – ***похідною Фреше***, яку ми будемо позначати через $F'(x)$.

*Зауваження 1.* Якщо $Y = R^1$, то $F'(x) \in X^*$, тобто є лінійним неперервним функціоналом.

*Зауваження 2.* Відмітимо, що рівність (8.1) можна записати також у вигляді

$$\lim_{\|h\|\to 0} \|F(x+h) - F(x) - D_x h\| \cdot \|h\|^{-1} = 0, \qquad (8.2)$$

або на мові нескінченно малих:

для $\forall \varepsilon > 0$ $\exists \delta > 0$, при якому $\forall h$ таких, що $\|h\| < \delta$ виконується нерівність

$$\|F(x+h) - F(x) - F'(x)h\| \le \varepsilon \|h\|. \qquad (8.2')$$

*Означення 2.* Відображення $F(x)$ називається ***строго диференційованим за Фреше*** в точці $x$, якщо $\exists\ F'(x) \in L(X,Y)$ при якому $\forall\, \varepsilon > 0$ $\exists\, \delta > 0$ таке, що $\forall\ x_1, x_2$, які задовольняють нерівностям $\|x_1 - x\| < \delta$, $\|x_2 - x\| < \delta$, виконується нерівність

$$\|F(x_1) - F(x_2) - F'(x)(x_1 - x_2)\| \le \varepsilon \|x_1 - x_2\|.$$

*Означення 3.* ***Диференціалом Гато*** відображення $F(x)$ в точці $x$ називають границю (якщо вона існує для $\forall\ h \in X$)

$$D_\Gamma F(x,h) = \lim_{t\to 0} \frac{F(x+th) - F(x)}{t}. \qquad (8.3)$$

У тому випадку, коли існує оператор $F'_{\Gamma}(x) \in L(X,Y)$ такий, що $D_{\Gamma}F(x,h) = F'_{\Gamma}(x)h,$ вираз $F'_{\Gamma}(x)$ називають ***похідною Гато***.

*Зауваження 3.* Із співвідношення (8.3) можемо одержати наступний вираз для обчислення диференціала Гато

$$D_{\Gamma}F(x,h) = \frac{d}{dt}F(x+th)\bigg|_{t=0}.$$

Можна показати, що із диференційованості за Фреше випливає диференційованість за Гато, але обернене твердження невірне, навіть для випадку $X = R^2, Y = R^1$.

В той же час має місце наступна теорема.

***Теорема 1.*** *Нехай похідна Гато $F'(x)$ існує в деякому околі точки $x$ і є в цьому околі неперервною функцією. Тоді в цій точці існує і похідна Фреше, причому $F'_{\Gamma}(x) = F'(x)$.*

Нехай $X$, $Y$, $Z$ – лінійні нормовані простори, $U$ – окіл точки $x \in X$, $V$ – окіл точки $y \in Y$, причому $y = \phi(x)$, а $F$ – відображення $V$ в $Z$, причому $l(x) = F(\phi(x))$. Має місце наступна теорема.

***Теорема 2.*** *Припустимо, що відображення $F(y)$ диференційоване за Фреше в точці $y$, а відображення $\phi(x)$ диференційоване за Фреше (Гато) в точці $x$. Тоді відображення $l(x)$ диференційоване за Фреше (Гато), причому відповідний диференціал має вигляд*

$$Dl(x,h) = F'(\phi(x))D\phi(x,h). \tag{8.4}$$

*Означення 4.* Нехай відображення $F(x,y)$ визначено в околі $U$ точки $(x,y) \in X \times Y$ і переводить окіл $U$ в окіл $Z$. Якщо відображення $F(x,y)$ при фіксованому $y$ диференційоване в точці $x$ за Фреше (Гато), то його похідна називається ***частковою похідною*** по $x$ відображення $F$ в точці $(x,y)$ і позначається $F_x(x,y)$.

Аналогічно визначається часткова похідна по $y$.

***Теорема 3 (про повний диференціал).*** *Нехай відображення $F(x,y)$ має в кожній точці околу $U$ часткові похідні $F_x(x,y)$, $F_y(x,y)$ за Гато, які є неперервними відображеннями в $U$ (в рівномірній операторній топології). Тоді відображення $F(x,y)$ диференційоване за Фреше (навіть строго диференційоване), причому*

$$F'(x,y)(h_1,h_2) = F_x(x,y)h_1 + F_y(x,y)h_2.$$

Нехай $\Pi$ – відображення $\underbrace{X \times ... \times X}_{n}$ в $Y$.

*Означення 5.* Відображення $\Pi(x_1,...,x_n)$ називається ***полілінійним***, якщо воно лінійне по кожному аргументу.

*Означення 6.* Полілінійне відображення називається ***обмеженим***, якщо існує константа $C > 0$ така, що $\|F(x_1,...x_n)\| \le C\|x_1\|...\|x_n\|$, $\forall x_i$, $i = \overline{1,n}$.

Простір полілінійних обмежених відображень є нормованим, де норма визначається за правилом

$$\|\Pi\| = \sup_{\|x_1\|\le 1,...\|x_n\|\le 1} \|\Pi(x_1,...,x_n)\|.$$

*Означення 7.* Полілінійне відображення називається ***симетричним***, якщо його значення не змінюється при довільній перестановці аргументів.

Нехай $\Pi(x_1,...,x_n)$ – симетричне неперервне полілінійне відображення. Тоді неперервна $n$-лінійна форма $\Pi(x,...,x)$ визначається як $\Pi(x,...,x) = \Pi(x_1,...,x_n)\big|_{x_1=x,...,x_n=x}$.

Введемо далі поняття похідних вищих порядків для відображення $F(x)$, $x \in U$.

Припустимо, що існує функціонал:

$$P_n(x) = \Pi_1(h) + \Pi_2(h,h) + ... + \Pi_n(h,...,h),$$

де $\Pi_s(h,...,h)$, $s = \overline{1,n}$ неперервні $s$-лінійні форми, такий що:

$$F(x+h) - F(x) = \Pi_1(h) + ... + \Pi_n(h,...,h) + \alpha_n(h)\|h\|^2,$$

де $\lim\limits_{\|h\|\to 0} \|\alpha_n(h)\| = 0.$

Вираз $n!\Pi_n(h,...,h)$ називають ***диференціалом n-го порядку*** для $F(x)$ і позначають $D^n F(x,h)$. Відповідне симетричне нелінійне неперервне відображення має вигляд

$$D^n F(x,h_1,...,h_n) = n!\Pi(h_1,...,h_n).$$

Це полілінійне відображення називають ***n-ю похідною*** $F(x)$ в точці $x$ та позначають $F^{(n)}(x)$, тоді має місце співвідношення:

$$F(x+h) - F(x) = F'(x)h + \frac{1}{2}F''(x)(h,h) + ... +$$

$$+\frac{1}{n!}F^{(n)}(x)(h,...,h) + \alpha_n(x,h)\|h\|^n,$$

яке називають ***формулою Тейлора***.

Зауважимо також, що справедлива рівність $D^n F(x,h) = \frac{d^n}{dt^n}F(x+th)\big|_{t=0}$.

Введемо далі інше визначення $n$-го диференціалу. Нехай перший диференціал відображення існує в околі $U$ точки $x$. Диференціал $DF(x,h)$ може бути диференційованим по $x$. Таким чином, ми приходимо до другого послідовного диференціалу

$$D[D(x,h),h_1] = D\left[F'(x)(h),h_1\right] = DF'(x,h_1)(h).$$

Позначимо $DF'(x,h_1) = F''(x)(h_1)$.

Вираз $F''(x)$ називається ***другою послідовною похідною***. Має місце рівність

$$D[DF(x,h),h_1] = F_p''(x)(h_1,h) = \frac{\partial^2}{\partial \tau_0 \partial t} F\left(x + th + \tau h_1\right)\Big|_{\tau_0 = t = 0}.$$

Аналогічно визначається $n$-й послідовний диференціал та $n$-та послідовна похідна. Має місце наступна теорема.

***Теорема 4.*** *Нехай існує $n$-й послідовний диференціал відображення $F(x)$, причому $n$-а послідовна похідна $F_p^{(n)}(x)$ є рівномірно неперервною в деякому околі $U$ точки $x$. Тоді в околі $U$ існує диференціал*

$$D^n F(x,h) = F^{(n)}(x)(h,...,h).$$

Нехай $F(x)$ функціонал, що визначений на лінійному нормованому просторі $X$ та приймає значення із множини $\overline{R} = R^1 \cup \{-\infty;\infty\}$. Введемо множини:

$$\operatorname{dom} F = \{x \in X : F(x) < \infty\},$$

$$\operatorname{epi} F = \{(x,\alpha) \in X \times R :\ \alpha \ge F(x),\ x \in \operatorname{dom} F\}.$$

Множина $\operatorname{dom} F$ називається ***ефективною*** множиною, а $\operatorname{epi} F$ – ***надграфіком*** функціоналу $F(x)$.

*Означення 8.* Функціонал, у якого $\operatorname{dom} F \ne \varnothing$ та скрізь $F(x) > -\infty$, називається ***власним***, а інші функціонали називаються ***невласними***.

Функціонал $F(x)$ називається ***опуклим***, якщо $\operatorname{epi} F$ – опукла множина в $X \times R$.

**Твердження 1.** Для опуклості власного функціоналу необхідно і достатньо, щоб для довільних $x_i \in \operatorname{dom} F$ та для довільних $\alpha_i \ge 0, i = \overline{1,n}, \sum_{i=1}^{n} \alpha_i = 1$ виконувалась нерівність

$$F\left(\sum_{i=1}^{n} \alpha_i x_i\right) \le \sum_{i=1}^{n} \alpha_i F(x_i). \qquad (8.5)$$

*Зауваження 4.* Сума скінченного числа та верхня границя довільної сім'ї опуклих функцій є опуклою функцією.

*Зауваження 5.* Функціонал $F(x)$, що визначений на опуклій множині $U$ банахового простору називається ***опуклим*** на цій множині, якщо для нього виконується нерівність (8.5), та ***сильно опуклим***, якщо нерівність (8.5) виконується лише коли одне із $\alpha_i = 1$.

*Означення 9.* Функціонал $F(x)$ називається ***надсильно опуклим*** на опуклій множині $U$ гільбертового простору $X$, якщо існує постійна величина $\mu > 0$ така, що виконується нерівність

$$F(\alpha x_1 + (1-\alpha)x_2) \le \alpha F(x_1) + (1-\alpha)F(x_2) - \alpha(1-\alpha)\mu\|x_1 - x_2\|^2$$
$$\forall\ x_1, x_2 \in U \text{ та } 0 \le \alpha \le 1.$$

**Твердження 2.** Функціонал $F(x)$ є опуклим тоді і тільки тоді, коли функція $f(t) = F(x+ty)$ опукла по $t \in R^1, \forall x, y \in X$.

***Теорема 5.*** *Нехай $U$ – опукла й відкрита множина із банахового простору $X$ і функціонал $F(x)$ є неперервно диференційованим в $U$. Тоді для опуклості $F(x)$ необхідно і достатньо, щоб виконувалася одна із наступних нерівностей:*

$$F(x_1) \ge F(x_2) + F'(x_2)(x_1 - x_2),$$
$$(F'(x_1) - F'(x_2))(x_1 - x_2) \ge 0.$$

*Якщо ж $F(x)$ двічі неперервно диференційована, то для опуклості $F(x)$ необхідно і достатньо, щоб виконувалась нерівність*

$$F''(x)(h, h) \ge 0, \quad \forall h \in X.$$

***Теорема 6.*** *Нехай $U$ – відкрита й опукла множина гільбертового простору $X$, а функціонал $F(x)$ неперервно диференційований в $U$. Тоді для того, щоб $F(x)$ був надсильно опуклим, необхідно і достатньо, щоб виконувалася одна із умов:*

*1) існує постійна величина $\mu > 0$ така, що*

$$F(x_1) \ge F(x_2) + (F'(x_2), x_1 - x_2) + \mu\|x_1 - x_2\|^2 \ \ \forall x_1, x_2 \in U;$$

*2) існує постійна величина $\mu > 0$ така, що*

$$(F'(x_1) - F'(x_2), x_1 - x_2) \ge \mu\|x_1 - x_2\|^2 \ \ \forall x_1, x_2 \in U.$$

*Якщо ж $F(x)$ двічі неперервно диференційований, то необхідно і достатньо, щоб існувало $\mu > 0$ таке, що*

$$(F''(x)h, h) \ge \mu\|h\|^2 \quad \forall h \in X,\ x \in X.$$

*Означення 10.* Функціонал $F(x)$ називається ***напівнеперервним знизу*** в точці $x_0 \in X$, якщо $\underline{\lim}_{x \to x_0} F(x) \ge F(x_0)$, та ***слабо неперервним знизу***, якщо нерівність виконується тоді, коли $x$ слабо збігається до $x_0$. Функціонал $F(x)$ називається ***просто напівнеперервним (слабо напівнеперервним) знизу***, якщо нерівність виконується для кожної точки простору $X$.

Зауважимо, що якщо функціонал $F(x)$ є слабо неперервним знизу, то він є і напівнеперервним знизу, оскільки із збіжності за нормою простору $X$ випливає і слаба збіжність. Обернене твердження справедливо для опуклих функціоналів.

**Твердження 3.** Якщо опуклий функціонал є напівнеперервним знизу, то він є і слабо напівнеперервним знизу.

Приведемо далі один із критеріїв напівнеперервності знизу опуклих функціоналів.

**Твердження 4.** Диференційований за Гато опуклий функціонал є напівнеперервним знизу.

***Доведення.***

Дійсно, оскільки $F(x)$ – опуклий функціонал, то

$$F((1-\lambda)x+\lambda y)\le(1-\lambda)F(x)+\lambda F(y) \text{ при } 0\le\lambda\le 1,$$

або

$$F(x+\lambda(y-x))\le F(x)+\lambda(F(y)-F(x)),$$

$$\frac{F(x+\lambda(y-x))-F(x)}{\lambda}\le F(y)-F(x).$$

Звідси при $\lambda\to 0$ одержимо

$$F'(x)(y-x)\le F(y)-F(x).$$

Нехай $y\to x.$ Тоді

$$F'(x)(y-x)\to 0 \text{ і } \varliminf_{y\to x}[F(y)-F(x)]\ge 0,$$

що і потрібно було показати.

**Твердження 5.** Для того, щоб власна опукла функція була напівнеперервна знизу, необхідно і достатньо, щоб її ефективна множина була замкнена.

*Означення 11.* Функціонал $o(x)=l(x)-b$ називається ***опорним*** для функціонала $F(x)$, якщо

1) $o(x)\le F(x) \quad \forall x$;

2) $b=\sup\limits_{x}\{l(x)-F(x)\}$.

Має місце наступна теорема.

***Теорема 7 (Мінковського).*** *Власний функціонал опуклий та замкнений, тобто надграфік його є замкненою множиною, тоді і тільки тоді, коли він є верхньою границею всіх своїх опорних функціоналів.*

*Зауваження 6.* Із теореми Мінковського випливає, що напівнеперервний знизу опуклий функціонал є верхньою границею опорних функціоналів.

**Твердження 6.** Якщо функціонал $F(x)$ опуклий та диференційований за Гато в точці $x_0$, то функціонал $o(x)=F'_{\Gamma}(x_0)(x-x_0)+F(x_0)$ є для нього опорним.

*Означення 12.* Функціонал $F^*(p)$ на $X^*$ називається ***спряженим*** до функціоналу $F(x)$ або ***перетворенням Юнга-Фенхеля***, якщо він визначається наступним чином

$$F^*(p)=\sup_{x\in X}\{p(x)-F(x)\},$$

а функціонал

$$F^{**}(x)=\sup_{p\in X^*}\{(p,x)-F^*(p)\}$$

називається ***другим спряженим*** функціоналом.

*Зауваження 7.* Між опорними і спряженими функціоналами є взаємозв'язок, а власне, якщо $F^*(p)$ – спряжений функціонал, то функціонал $o_p(x)=p(x)-F^*(p)$ буде опорним і навпаки, якщо $o_p(x)$ є опорним функціоналом, то $F^*(p)=p(x)-o_p(x)$ буде спряженим функціоналом.

**Твердження 7.** 1) Функціонали $F^*$ та $F^{**}$ є опуклими та напівнеперервними знизу;

2) для довільних $x\in X$, $p\in X^*$ виконується нерівність Юнга

$$F(x)+F^*(p)\geq p(x);$$

3) $F^{**}(x)\leq F(x)\quad \forall x$.

**Твердження 8.** $F^{**}(x)=F(x)$ тоді і тільки тоді, коли функціонал $F(x)$ опуклий і напівнеперервний знизу.

Нехай функціонал $F(x):X\to\overline{R}$.

*Означення 13.* ***Субдиференціалом*** $F$ в точці $x_0$ називається підмножина $X^*$, що складається із таких елементів $x^*\in X^*$, для яких виконується нерівність

$$F(x)-F(x_0)\geq x^*(x-x_0)\quad \forall x\in X.$$

Будемо позначати субдиференціал в точці $x_0$ через $\partial F(x_0)$.

**Твердження 9.** Субдиференціал $\partial F(x_0)$ є опуклою множиною в $X^*$.

***Теорема 8 (Моро-Рокафелара).*** *Нехай $F_i$ опуклі власні функції на $X$. Тоді*

$$\partial\left(\sum_{i=1}^{n}F_i\right)(x)\supseteq\sum_{i=1}^{n}\partial F_i(x).$$

*Якщо ж в деякій точці $\overline{x}$ всі функції, окрім можливо однієї, неперервні, а ця остання скінченна, то в довільній точці маємо рівність*

$$\partial\left(\sum_{i=1}^{n}F_i\right)(x)=\sum_{i=1}^{n}\partial F_i(x).$$

Наслідок 1. Якщо $F(x)$ опукла й неперервна в точці $x_0$ функція, то $\partial F(x_0)\neq\varnothing$.

***Теорема 9 (Дубовицького-Мілютіна).*** *Нехай $F_i(x)$ опуклі функціонали на $X$, які неперервні в точці $x_0$, $F(x)=\max\limits_{1\leq i\leq n}F_i(x)$, $I=\{i_1,...,i_s\}$ – набір індексів такий, що $F_i(x_0)=F(x_0)$, $i\in I$ та $F_i(x_0)<F(x_0)$, $i\notin I$. Тоді*

$$\partial F(x_0)=\left\{x^*: x^*=\sum_{k=1}^{s}\lambda_k x_k^*,\, x_k^*\in\partial F_{i_k}(x_0),\, \lambda_k\geq 0,\, \sum_{k=1}^{s}\lambda_k=1\right\}.$$

Наслідок 2. Нехай функціонал $F_k$ диференційований за Гато. Тоді в теоремі Дубовицького-Мілютіна $x_k^*=F'_{i_k}(x_0)$.

## 8.2. Теореми про існування екстремумів. Необхідні умови екстремумів

*Означення 1.* Множина $U$ із лінійного нормованого простору називається ***компактною***, якщо із довільної послідовності $x_n\in U$ можемо виділити підпослідовність $x_{n_k}$, яка збігається до елемента $x_0\in U$. Множина $U$ називається ***слабо компактною***, якщо із довільної послідовності можемо виділити підпослідовність, яка слабо збігається до деякого елемента $x_0$.

В скінченновимірному просторі будь-яка обмежена замкнена множина є компактною. В нескінченновимірних просторах це твердження, взагалі кажучи, є неправильним.

**Приклад 1**. Нехай $X$ – гільбертів простір, а $U=\{u:\|u\|\leq 1\}$ – одинична куля в $X$, яка є обмеженою та замкненою множиною в $X$ (доведіть це).

Нехай $\{l_k\}$ – ортонормований базис в $X$. Оскільки $(l_k,x)\to 0$ при $k\to\infty\ \ \forall x\in X$, тому що $\sum_{k=1}^{\infty}(l_k,x)^2=\|x\|^2$, то $l_k$ слабо збігається до нуля. Якщо б із послідовності можна було б виділити сильно збіжну підпослідовність, то вона б збігалася сильно до нуля, але в той же час $\lim_{k\to\infty}\|l_k\|=1$, оскільки $\|l_k\|=1$. Тобто, одержали протиріччя.

***Теорема 1****. Нехай $U$ – слабо компактна множина із лінійного нормованого простору $X$. Функціонал $F(x)$ скінченний та слабо напівнеперервний на $U$. Тоді $F^*=\inf_{x\in U}F(x)>-\infty$, множина $U_*=Arg\min F(x)$ – непорожня, слабо компактна і довільна мінімізуюча послідовність слабо збігається до одного із елементів $U_*$.*

*Означення 2.* Множина $U$ із лінійного нормованого простору $X$ називається ***слабо замкненою***, якщо довільна слабо збіжна послідовність із цієї множини збігається до елемента множини $U$.

***Теорема 2****. Всяка обмежена слабо замкнена множина із рефлексивного лінійного нормованого простору $X$ є слабо компактною.*

*Зауваження 1.* Якщо ж простір не рефлексивний, то обмежена слабо замкнена множина не може бути слабо компактною.

***Теорема 3.*** *Всяка обмежена, опукла, замкнена множина $U$ із рефлексивного лінійного нормованого простору є слабо компактною.*

Нехай $G$ – вимірна за Лебегом множина із евклідового простору $R^n$. Тоді банахів простір вимірних вектор-функцій $x(t)=(x_1(t),...,x_m(t))$ із скінченною нормою

$$\|x\|_p=\left\{\int\limits_G\left(\sum_{i=1}^m x_i^2(t)\right)^{\frac{p}{2}}dt\right\}^{\frac{1}{p}}$$

є рефлексивним при $1<p<\infty$, і значить, куля $S_r=\{x:\|x\|_p\le r\},\ r>0$ в такому просторі є слабо компактною (але не компактною) множиною.

***Теорема 4.*** *Нехай $U$ опукла, замкнена й обмежена множина із рефлексивного лінійного нормованого простору $X$, функціонал $F(x)$ опуклий та напівнеперервний знизу. Тоді $F^*=\inf\limits_U F(x)>-\infty,\ \ U_*=Arg\min F(x)$ – непорожня, опукла, замкнена, обмежена множина, а довільна мінімізуюча послідовність збігається до одного із елементів $U_*$.*

***Теорема 5.*** *Нехай $U$ – опукла й замкнена множина із рефлексивного лінійного нормованого простору $X$, функціонал $F(x)$ опуклий, напівнеперервний знизу на $U$ і для деякої фіксованої точки $y\in U$ множина Лебега*

$$L(y)=\{x\in U: F(x)\le F(y)\}$$

*обмежена. Тоді $\inf F(x)>-\infty,\ \ U_*=Arg\min F(x)\ne\varnothing$ – опукла, замкнена, обмежена, а довільна мінімізуюча послідовність слабо збігається до одного із елементів множини $U_*$.*

*Зауваження 2.* Достатньою умовою для обмеженості множини $L(y)$ є умова

$$\lim_{k\to\infty}F(x_k)\to\infty\quad\forall x_k,\ \ \|x_k\|\to\infty.$$

*Зауваження 3.* Якщо в умовах теореми 5 функціонал $F(x)$ сильно опуклий, то $U_*$ складається із однієї точки.

Дійсно, якщо $x_1,x_2\in U_*$, то $\frac{1}{2}x_1+\frac{1}{2}x_2\in U_*$, $F^*=F(\frac{1}{2}x_1+\frac{1}{2}x_2)>$ $>\frac{1}{2}F(x_1)+\frac{1}{2}F(x_2)=F^*$, а це протиріччя.

Якщо функціонал $F(x)$ є надсильно опуклим, то теореми про існування точок мінімуму можна уточнити.

***Теорема 6.*** *Нехай $U$ – опукла й замкнена множина із гільбертового простору $X$, а функціонал $F(x)$ є сильно опуклим та напівнеперервним знизу на $X$. Тоді*

*1) Множина Лебега $L(y)$ є надсильно опуклою, замкненою, обмеженою при всіх $y$;*

*2) $F^* > -\infty$, $U_* \neq \varnothing$, причому $U_*$ складається із однієї точки;*

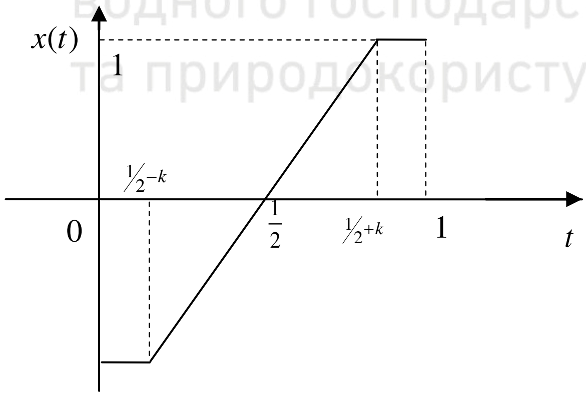
*3) довільна мінімізуюча послідовність $\{x_n\}$ збігається до точки $x_*$ за нормою, причому $\mu\|x_n - x_*\|^2 \leq F(x_n) - F(x_*)$.*

**Приклад 2.** Розглянемо задачу мінімізації функціоналу $F(x) = \int_0^1 sign(\frac{1}{2} - t)x(t)dt$ при:

$$x = x(t) \in U = \left\{x(t) \in C_{[0;1]},\ |x(t)| \leq 1,\ 0 \leq t \leq 1\right\},$$

де $C_{[0;1]}$ – простір неперервних функцій на відрізку $[0;1]$. Очевидно, що $F(x) \geq -1$. Зауважимо, що існує послідовність функцій $x_k(t)$ таких, що $\lim_{k\to\infty} F(x_k) = 1$, тобто $x_k(t)$ – мінімізуюча послідовність. За послідовність $x_k(t)$ візьмемо наступну

$$x_k(t) = \begin{cases} k(t - \frac{1}{2}), & \text{при } |t - \frac{1}{2}| \leq \frac{1}{k}; \\ sign(t - \frac{1}{2}), & \text{при } |t - \frac{1}{2}| > \frac{1}{k}. \end{cases}$$




При $k \to \infty$ послідовність $x_k(t)$ збігається до функції

$$x_0(t) = \begin{cases} 0, & \text{при } t \neq \frac{1}{2}; \\ 1, & \text{при } t = \frac{1}{2}. \end{cases}$$

$$U = \left\{x(t) \middle|\ |x(t)| \leq 1, -1 \leq t \leq 1\right\}.$$

Множина $U$ є опуклою, замкненою; функціонал $F(x)$ є опуклим напівнеперервним знизу, але простір $C_{[-1;1]}$ не є рефлексивним, і тому, як видно із прикладу, при відсутності рефлексивності простору теорема не справедлива.

**Приклад 3.** Нехай множина альтернатив $U$ має вигляд $U = \left\{ u(t) \middle| \ u(t) \in L_2(0;1),\ |u(t)| \le 1 \text{ майже скрізь на } (0;1) \right\}$.

Оцінку альтернативи $u(t) \in U$ задано у вигляді вектора $\left(F_1(u),\ F_2(u)\right)$, де $F_1(u) = \int\limits_0^1 x^2(t)dt,\ \ F_2(u) = -\int\limits_0^1 u^2(t)dt,\ \ x(t) = \int\limits_0^t u(s)ds,\ 0 \le t \le 1$.

Для знаходження оптимальної за Парето альтернативи, застосуємо метод адитивної згортки.

Для цього введемо функціонал

$$F(u) = \lambda_1 F_1(u) + \lambda_2 F_2(u),\ \ \lambda_1,\ \lambda_2 > 0,$$

та знайдемо функцію $\hat{u}(t)$ таку, що $\hat{u} \in \arg\min F(u)$.

Не обмежуючи загальності подальших міркувань, покладемо $\lambda_1 = \lambda_2 = 1$. Тоді функціонал $F(u)$ буде мати вигляд

$$F(u) = \int\limits_0^1 \left( x^2(t) - u^2(t) \right) dt\ .$$

Зауважимо, що $F(u) \ge -1$, оскільки

$$F(u) \ge -\int\limits_0^1 u^2(t)dt \ge -\max_{0 \le t \le 1} u^2(t) = -1\ .$$

Множина $U$ є опуклою та слабо компактною (але не компактною множиною). Для слабої компактності опуклої множини достатньо показати її обмеженість та замкненість. Оскільки обмеженість очевидна, покажемо її замкненість.

Нехай послідовність $u_k(t)$ збігається в метриці простору $L_2(0;1)$ до функції $u_0(t)$. Тоді існує підпослідовність $u_{k_n}(t)$, яка збігається до $u_0(t)$ майже скрізь. Оскільки $\left|u_{k_n}(t)\right| \le 1$, то майже скрізь $\left|u_0(t)\right| \le 1$, що й потрібно було довести.

Функціонал $F(u)$ – неперервний в метриці простору $L_2(0;1)$. Дійсно, якщо $u_n(t)$ збігається в метриці $L_2(0;1)$ до $u_0(t)$, тобто

$$\lim_{n \to \infty} \int\limits_0^1 \left( u_n(t) - u_0(t) \right)^2 dt = 0\,,$$

то

$$\lim_{n \to \infty} \int\limits_0^1 u_n^2(t)dt = \int\limits_0^1 u_0^2(t)dt\ .$$

Крім цього, $x_n(t) = \int\limits_0^t u_n(s)ds$ збігається до $x_0(t)$ в метриці $L_2(0;1)$ (довести це), тому

$$\lim_{n\to\infty}\int_0^1 x_n^2(t)dt=\int_0^1 x_0^2(t)dt .$$

З цих рівностей одержимо неперервність функціоналу $F(u)$. Якби множина $U$ була компактна, то існувала б функція $u_0(t)$ така, що $F(u_0)=-1$, так як $\inf_{u\in U} F(u)=-1$. Зауважимо, що оскільки простір $L_2(0;1)-$ гільбертів, а значить рефлексивний, то для існування $u_0(t)$ ($u_0 \in Arg\min F(u)$) достатньо показати слабу напівнеперервність знизу функціоналу $F(u)$.

Але, як ми покажемо нижче, функціонал $F(u)$ не є слабо напівнеперервним знизу (незважаючи на те, що він є диференційованим і неперервним). Дійсно, якщо ми візьмемо послідовність $\sin \pi kt = u_k(t)$ $0\le t\le 1$, то $u_k(t)$ слабо збігається до нуля, оскільки для довільної функції $x(t)$ із $L_2(0;1)$

$$\int_0^1 \sin \pi kt\, x(t)dt = c_k\text{ , } c_k - \text{ коефіцієнти Фур'є функції } x(t),$$

тому $\sum_{k=1}^{\infty} {c_k}^2 < \infty$, тобто $c_k \to 0,\ k\to\infty$.

Обчислимо:

$$F(u_k)=\int_0^1 ({x_k}^2(t)-{u_k}^2(t))dt ,$$

$$x_k(t)=\int_0^t \sin \pi ks\, ds = \frac{1}{\pi k}\left(\cos \pi kt - 1\right),$$

$$\int_0^1 {u_k}^2(t)dt=\int_0^1 \sin^2 \pi ks\, ds=\int_0^1 \frac{1-\cos 2\pi ks}{2}ds=\frac{1}{2}-\int_0^1 \frac{\cos 2\pi ks}{2}ds=\frac{1}{2},$$

$$\int_0^1\left(1-\frac{\cos \pi ks}{\pi k}\right)^2 ds=\int_0^1\left(1-\frac{2\cos \pi ks}{\pi k}+\frac{\cos^2 \pi ks}{(\pi k)^2}\right)^2 ds=\frac{1}{\pi k}+\int_0^1 \frac{\cos^2 \pi ks}{(\pi k)^2}ds \to 0 .$$

Отже,

$$F(u_k)\to -1/2\ ,\ k\to\infty ,$$

але $F(0)=0$. Тому $\lim_{k\to\infty} F(u_k)=-1/2<0$, тобто функціонал $F(u)$ не є слабо напівнеперервним знизу.

Покажемо, що функціонал $F(u)$ не досягає своєї нижньої границі на жодній із функцій множини $U$. Дійсно, якщо $\int_0^1 x^2(t)dt=0$, то $x(t)=0$, а значить $u(t)=0$ і $F(0)=0>-1.$ Якщо ж $\int_0^1 x^2(t)dt \neq 0,$ то $F(u)>-\int_0^1 u^2(t)dt \geq -1,$ що й потрібно було показати.

Цікаво відмітити, що якщо розглянути скінченновимірний аналог цієї задачі, тобто ввести функцію:

$$F(u_1,u_2,...,u_N)=\frac{1}{N}\sum_{k=1}^{N}({x_k}^2-{u_k}^2),$$

де $x_k$ є розв’язком рівняння $x_{k+1}=x_k+u_k$, то на множині $|u_k|\leq 1$ існує точка $x_0=0$, де досягається мінімум.

*Означення 3.* Нехай функціонал $F(x)$ визначається в деякому околі точки $x_0$ лінійного нормованого простору $X$. Точка $x_0$ називається ***точкою мінімуму*** для функціонала $F(x)$, якщо існує окіл точки $x_0$ такий, що $F(x)\geq F(x_0)$ $\forall x$ із цього околу.

**Твердження 1.** Нехай існує диференціал Гато функціоналу $F(x)$ в околі точки $x_0$ та нехай $F(x_0)=\inf F(x)$. Тоді необхідно, щоб виконувалася умова $DF(x_0,h)=0$ $\forall\ h\in X$. Якщо існує похідна Гато, то $F_\Gamma{}'(x_0)=0$. Нехай в точці $x_0$ існує друга похідна від функції $f(t)=F(x_0+th),\ t\in R$. Тоді необхідною умовою мінімуму є:

$$\left.\frac{d^2}{dt^2}F(x_0+th)\right|_{t=0}\geq 0 \quad \forall\, h\in X\ .$$

*Зауваження 4.* Якщо в скінченновимірному просторі існують другі похідні від $F(x)$, то

$$\left.\frac{d^2}{dt^2}F(x_0+th)\right|_{t=0}=(F''(x_0)h,\ h)$$

і наша умова переходить в необхідну умову для невід’ємно означеної матриці $F''(x_0)$. Якщо в скінченновимірному просторі $F'(x_0)=0$, то умова $F''(x_0)>0$ є достатньою для того, щоб $x_0$ була точкою мінімуму.

Наведемо приклад, який показує, що в нескінченновимірному просторі ця умова не є достатньою.

**Приклад 4.** Нехай в сепарабельному гільбертовому просторі $H$ із базисом $e_1,...,e_k,...$ визначається функціонал

$$F(x)=\sum_{k=1}^{\infty}\frac{(x,e_k)^2}{k^3}-\sum_{k=1}^{\infty}(x,e_k)^4\ .$$

Тоді

$$\left.\frac{dF(x+th)}{dt}\right|_{t=0} = 2\sum_{k=1}^{\infty}\frac{(x,e_k)(h,e_k)}{k^3} - 4\sum_{k=1}^{\infty}(x,e_k)^3(h,e_k),$$

при $x=0$:

$$\left.\frac{dF(th)}{dt}\right|_{t=0} = 0, \quad \left.\frac{d^2}{dt^2}F(th)\right|_{t=0} = 2\sum_{k=1}^{\infty}\frac{(h,e_k)^2}{k^3} > 0,$$

$$\left.\frac{d^2}{dt^2}F(x+th)\right|_{t=0} = 2\sum_{k=1}^{\infty}\frac{(h,e_k)^2}{k^3} - 12\sum_{k=1}^{\infty}(x,e_k)^2(h,e_k)^2.$$

Візьмемо точку $\bar{x}$, для якої $(\bar{x},e_j)=0$ при $j\neq n$, $(\bar{x},e_n)=1/n$. Отже, $F(\bar{x})=\frac{1}{n^5}-\frac{1}{n^4}<0=F(0)$ і оскільки $\|\bar{x}\|^2=\sum_{k=1}^{\infty}(x,e_k)^2=\frac{1}{n^2}$, то в будь якому околі точки 0 існує точка $\bar{x}$ така, що $F(\bar{x})<F(0)$, а це значить, що 0 не є точкою мінімуму.

**Твердження 2.** Нехай $F_i(x)$ $i=1,2$ – опуклі функціонали, які визначені на опуклій множині $U$ лінійного нормованого простору $X$, причому для $F_1(x)$ існує диференціал Гато. Позначимо через $U_* = Arg\min_{x\in U} F(x)$, де $F(x)=F_1(x)+F_2(x)$. Тоді наступні дві умови еквівалентні:

1) $x\in U_*$,
2) $DF_1(x,y-x)+F_2(y)-F_2(x)\geq 0 \ \ \forall y\in U$.

***Доведення.***

Нехай $x\in U_*$. Тоді якщо $y$ довільна точка із $U$, то

$$F((1-\lambda)x+\lambda y)=F(x+\lambda(y-x))\geq F(x) \text{ при } 0\leq\lambda\leq 1,$$

або

$$F_1(x)+F_2(x)\leq F_1(x+\lambda(y-x))+F_2((1-\lambda)x+\lambda y)\leq$$
$$\leq F_1(x+\lambda(y-x))+(1-\lambda)F_2(x)+\lambda F_2(y)$$

або

$$\frac{F_1(x+\lambda(y-x))-F_1(x)}{\lambda}+(F_2(y)-F_2(x))\geq 0.$$

При $\lambda\to 0$ отримаємо справедливість нерівності пункту 2).

***Теорема 7* (*про достатні умови екстремуму*)**. *Нехай для функціоналу $F(x)$ існує друга похідна в сенсі Фреше. Тоді якщо*

$$F'(x)=0, \ F''(x)(h,h)\geq\alpha\|h\|^2 \ \ \forall h\in X$$

*для деякого $\alpha>0$, то $x\in Arg\min F(x)$.*

Нехай $F(x)$ – опукла функція, визначена на лінійному нормованому просторі $X$ і для якої існує субдиференціал.

**Твердження 3.** Для того, щоб $x_0\in Arg\min F(x)$ необхідно і достатньо, щоб $0\in\partial F(x_0)$.

***Доведення.***

Згідно означення субдиференціала, якщо $x^* \in \partial F(x_0)$, то $F(x) \ge F(x_0) + x^*(x - x_0) \quad \forall x \in X$. Нехай $0 \in \partial F(x_0)$, тоді $F(x) \ge F(x_0)$, тобто $x_0$ – точка мінімуму. Навпаки, із того, що $x_0$ – точка мінімуму випливає нерівність $F(x) \ge F(x_0) + o(x - x_0)$, а значить $0 \in \partial F(x_0)$.

Нехай $X$ та $Y$ – банахові простори $\Lambda \in L(X,Y)$, $F_i : X \to \overline{R}$, $i = \overline{0,m}$ – опуклі функціонали, $U$ опукла множина із $X$. Розглянемо наступну задачу:

$$F_0(x) \to \inf,\ F_i(x) \le a_i \ \ i = \overline{1,m},$$

$$\Lambda x = b,\ x \in U \text{, де } a_i \in R^1. \tag{8.6}$$

Позначимо через $U_1$ множину $U_1 = \{x \mid \Lambda x = b\}$.

***Теорема 8* (*аналог теореми Куна–Таккера*).** *Нехай функціонали $F_i(x)$, $i = \overline{0,m}$ – неперервні в точці $\hat{x} \in U \cap U_1$, в якій досягається мінімум в задачі (8.6). Тоді існують числа $\lambda_i \ge 0$, $i = \overline{1,m}$ такі, що $\sum_{i=1}^{m} \lambda_i = 1$, $\lambda_i F_i(\hat{x}) = 0$, $i \ge 1$ та точка $\hat{x}_0 \in \partial\delta(U \cap U_1)(\hat{x})$, для яких $0 \in \sum_{i=1}^{m} \lambda_i \partial F_i(\hat{x}) + \hat{x}_0$, де*

$$\delta(U)(x) = \begin{cases} 1, & x \in U, \\ 0, & x \notin U. \end{cases}$$

Н а с л і д о к 1 (***принцип Лагранжа***). Нехай в умовах теореми 8 $U = X$ та образ $X$ при відображені $\Lambda$ замкнений в $Y$. Тоді існують множники Лагранжа $\lambda_i \ge 0$, $i \ge 0$ такі, що:

$$\lambda_i F_i(\hat{x}) = 0,\ i \ge 1,\ \hat{y} \in Y^*,\ \min_x L(x, \hat{y}, \lambda) = L(\hat{x}, \hat{y}, \lambda),$$

де $L(x, y, \lambda) = \lambda_0 F_0(x) + \sum_{i=1}^{m} \lambda_i (F_i(x) - a_i) + y(\Lambda x - b)$.

***Доведення.***

Якщо образ $\Lambda$ є власним підпростором $Y$, то можна покласти $\lambda_i = 0$, $\hat{y} \in (\operatorname{Im}\Lambda)^{\perp}$, де $\operatorname{Im}\Lambda$ – образ $\Lambda$, $(\operatorname{Im}\Lambda)^{\perp}$ – ортогональне доповнення до образу, а значить умови теореми 8 виконуються. Нехай $\operatorname{Im}\Lambda = Y$. Тоді $\partial\delta(U_1)(x) = (Ker\Lambda)^{\perp} = \operatorname{Im}\Lambda^*$, де $Ker\Lambda$ – ядро оператора $\Lambda$, $\Lambda^*$ – спряжений оператор до $\Lambda$. Тоді, згідно теореми 8, $0 \in \partial \sum_{i=1}^{m} \lambda_i \partial F_i(\hat{x}) + \operatorname{Im}\Lambda^* = \partial_x L(\hat{x}, \hat{y}, \lambda)$, а значить, згідно означення субдиференціала $L(x, \hat{y}, \lambda) \ge L(\hat{x}, \hat{y}, \lambda) + o(x - \hat{x}) = L(x, \hat{y}, \lambda) \ \ \forall x$.

## 8.3. Застосування методів оптимізації в нескінченновимірних просторах

### 1. Проблема мінімізації квадратичних функціоналів та варіаційні рівняння

В теорії крайових задач для рівнянь з частинними похідними еліптичного типу важливими є питання про існування та єдиність розв'язків. У випадку розривних правих частин, такі розв'язки не є класичними і для визначення узагальнених розв'язків вводиться спеціальні простори і функціонали на них. Точки мінімуму цих функціоналів, що одержуються як розв'язки варіаційних рівнянь, і називають ***узагальненими розв'язками задач***, що породжують відповідні функціонали.

Приведемо далі відповідні викладки для деяких випадків.

Нехай $H$ – гільбертів простір над полем дійсних чисел, який співпадає із своїм спряженим, $V$ – підпростір простору $H$, який є гільбертовим відносно свого скалярного добутку і який неперервно та щільно вкладений в $H$. На просторі $V$ визначимо функціонал:

$$F(\phi) = a(\phi,\phi) - 2l(\phi),$$

де $a(\phi,\phi)$ – квадратична форма, що відповідає симетричній білінійній формі $a(\phi,\psi)$, $l(\phi)$ – лінійний функціонал.

**Твердження 1.** Припустимо, що форма $a(\phi,\psi)$ – неперервна та існує таке число $\gamma > 0$, що для всіх $\phi$ виконується нерівність $a(\phi,\phi) \geq \gamma \|\phi\|_V^{\ 2}$. Тоді, якщо $l(\phi)$ – неперервний функціонал на $V$, то існує єдиний розв'язок варіаційного рівняння

$$a(\phi,\psi) = l(\psi),\ \forall \psi \in V \text{ і } \phi \in Arg\min F(\phi).$$

***Доведення.***

Очевидно, що функціонал $F(\phi)$ є опуклим та неперервним і $\lim\limits_{\|\phi\|\to\infty} F(\phi) = \infty$. В силу теореми 4 п. 8.2 існує точка мінімуму. Ця точка мінімуму може бути знайдена із умови

$$\left.\frac{d}{dt} F(\phi + t\psi)\right|_{t=0} = 0 \quad \forall \psi \in V.$$

Але $\dfrac{1}{2}\dfrac{d}{dt} F(\phi + t\psi) = a(\phi,\psi) - l(\psi)$. Звідси отримаємо необхідне рівняння.

Припустимо далі, що існують два розв'язки такого рівняння $\phi_1,\ \phi_2$.

Тоді

$$\gamma \|\phi_1 - \phi_2\|_V^{\ 2} \leq a(\phi_1 - \phi_2, \phi_1 - \phi_2) = a(\phi_1, \phi_1) - 2a(\phi_1, \phi_2) + a(\phi_2, \phi_2).$$

Але $a(\phi_1,\phi_1)=l(\phi_1),\ a(\phi_2,\phi_2)=l(\phi_2)$, $a(\phi_1,\phi_2)=l(\phi_2)$ та $a(\phi_1,\phi_2)=a(\phi_2,\phi_1)=l(\phi_1)$.

Звідси одержимо, що

$$\|\phi_1-\phi_2\|_V^2=0,\ \text{тобто}\ \phi_1=\phi_2 .$$

Нехай далі $H=L_2(0,1)$, $V=\overset{0}{W}{}_2^1(0;1)$ – простір функцій Соболева, які визначені на відрізку $[0;1]$ і такі, що $\phi(0)=\phi(1)=0$.

Введемо білінійну форму

$$a(\phi,\psi)=\int_0^1 a(t)\phi'(t)\psi'(t)dt+\int_0^1 c(t)\phi(t)\psi(t)dt .$$

Нехай $a(t),c(t)$ – вимірні, обмежені та невід'ємні функції, причому існує $\gamma>0$ таке, що $a(t)\ge\gamma$ майже скрізь. Тоді для форми $a(\phi,\psi)$ виконуються всі умови твердження 1. Якщо ввести функціонал $l(\phi)=\int_0^1 f(t)\phi(t)dt$, де $f(t)\in L_2(0;1)$, то він буде неперервним і лінійним, а значить буде існувати єдиний розв'язок варіаційного рівняння

$$a(\phi,\psi)=l(\psi)\quad \forall\psi\in\overset{0}{W}{}_2^1(0;1) .$$



Зауважимо, що функція $\phi(t)$ є узагальненим розв'язком крайової задачі:

$$-\frac{d}{dt}a(t)\frac{d}{dt}\phi+c(t)\phi=f(t),$$
$$\phi(0)=\phi(1)=0 .$$

**2. Багатокритеріальні задачі для варіаційних рівнянь**

Нехай $U$ – деяка множина альтернатив гільбертового простору $H_1$. Оцінка кожної альтернативи $u\in U$ задається вектором $(F_1(\phi_1,u),F_2(\phi_2,u),\ldots,F_m(\phi_m,u))$, де $\phi_j$ є розв'язком варіаційного рівняння:

$$a_j(\phi_j,\psi)=(B_ju,\psi)_H+l_j(\psi)\quad\forall\psi\in V,\ \ j=\overline{1,m}; \tag{8.7}$$

$a_j(\phi,\psi)$ – білінійні, неперервні, симетричні форми на $V$, причому існують $\gamma_j>0$ такі, що $a_j(\phi,\phi)\ge\gamma_j^2\|\phi\|_V^2$, $B_j\in L(H_1,H)$, $F_j(\phi_j,u)$ – функціонали на $V\times H_1$.

На множині оцінок введемо відношення Парето. Оцінка $(F_1(\phi_1^{(1)},u_1),\ldots,F_m(\phi_m^{(1)},u_1))$ знаходиться у відношенні Парето до оцінки $(F_1(\phi_1^{(2)},u_2),\ldots,F_m(\phi_m^{(2)},u_2))$, якщо виконуються нерівності:

$$F_i(\phi_i^{(1)},u_1)\le F_i(\phi_i^{(2)},u_2)),\ i=\overline{1,m}$$

та існує $i_0$ таке, що

$$F_{i_0}(\phi_{i_0}^{(1)}, u_1) < F_{i_0}(\phi_{i_0}^{(2)}, u_2)) .$$

Оцінки, для яких не виконується відповідна система нерівностей, називають ***Парето-оптимальними***.

Для знаходження Парето-оптимальних оцінок можна застосовувати різні методи. Одним із таких методів є метод ідеальної точки. В нашому випадку все зводиться до наступної процедури.

Знаходяться величини $F_i^* = \min\limits_{u \in U} F_i(\phi_i, u)$, $i = \overline{1,m}$. Далі обчислюється відстань в деякій метриці простору $R^m$ від точки $F = (F_1(\phi_1, u), \ldots, F_m(\phi_m, u))$ до точки $F^* = (F_1^*(\phi_1^*, u^*), \ldots, F_m^*(\phi_m^*, u^*))$, тобто вводиться функціонал

$$I(u) = \rho(F, F^*) .$$

Оптимальна за Парето альтернатива $u^*$ визначається із умови $u^* \in Arg \min I(u)$.

Приведемо конкретний приклад обчислення $u^*$.

**Приклад 1**. Нехай $F_i(\phi_i, u) = (g_i, \phi_i)_H$, $i = \overline{1,m}$, $U = \{u : \|u\| \le \beta^2\}$, $g_i$ – система лінійно незалежних векторів в просторі $H$.

Обчислимо спочатку $F_i^*$. Оскільки:

$$(g_i, \phi_i) = (p_i, B_i u) + \delta_i = (B^* p_i, u) + \delta_i ,$$

де $\delta_i = (g_i, \bar{\phi}_i)$, а $\bar{\phi}_i$ – розв'язок рівняння $a_i(\bar{\phi}_i, \psi) = l_i(\psi)$, $p_i$ – розв'язок рівняння $a_i(p_i, \psi) = (g_i, \psi)$, то

$$\min_u (g_i, \phi_i) = -\beta^2 \left\| B_i^* p_i \right\| + \delta_i .$$

Розглянемо далі розв'язок задачі $(g_i, \phi_i) = F_i^*$, тобто

$$(B_i^* p_i, u) = -\beta^2 \left\| B_i^* p_i \right\|, \ i = \overline{1,m} .$$

Нехай система векторів $B_i^* p_i$ – лінійно незалежна в просторі $H_1$. Тоді розв'язок, який має найменшу норму, має вигляд:

$$\hat{u} = \sum_{j=1}^m \hat{x}_j B_j^* p_j ,$$

де $\hat{x}_j$ – є розв'язок системи лінійних алгебраїчних рівнянь

$$\sum_{j=1}^m (B_i^* p_i, B_j^* p_j) \hat{x}_j = -\beta^2 \left\| B_i^* p_i \right\| .$$

В силу лінійної незалежності векторів $B_i^* p_i$ матриця системи додатньо визначена, а значить невироджена. Таким чином, існує єдиний її розв'язок – $\hat{x}_j$.

Якщо $(\hat{u}, \hat{u}) \le \beta^2$, то альтернатива $\hat{u}$ є розв'язком, який має ту властивість, що на $\hat{u}$ одночасно досягають мінімуму всі функціонали. В іншому випадку, ми повинні розв'язувати задачу на умовний екстремум, тобто:

знайти мінімум $I(u)$ при умові $(u, u) = \beta^2$.

Задамо метрику

$$\rho(F, F^*) = \sum_{j=1}^{m} \left| F_j - F_j^* \right| = \sum_{j=1}^{m} F_j - \sum_{j=1}^{m} F_j^* ,$$

тобто

$$I(u) = \sum_{j=1}^{m} (B_j^* p_j, u) - \sum_{j=1}^{m} F_j^* = \left( \sum_{j=1}^{m} B_j^* p_j, u \right) - \sum_{j=1}^{m} F_j^* .$$

В силу нерівності Коші-Буняковського

$$\min_u I(u) = \left( \sum_{j=1}^{m} B_j^* p_j, \hat{u} \right) - \sum_{j=1}^{m} F_j^* ,$$

причому

$$\hat{u} = -\beta \sum_{j=1}^{m} B_j^* p_j \left\| \sum_{j=1}^{m} B_j^* p_j \right\|^{-1} .$$

Розглянемо далі інший вигляд метрики:

$$\rho(F, F^*) = \max_{i \in [1;m]} \left| F_i - F_i^* \right| = \max_{i \in [1;m]} (F_i - F_i^*) = \max_{q_i} \sum_{i=1}^{m} q_i (F_i - F_i^*) ,$$

де $q_i$ – невід'ємні числа, що задовольняють умові $\sum_{i=1}^{m} q_i = 1$.

Із теореми про сідлову точку маємо:

$$\min_{u \in U} \rho(F, F^*) = \max_{q_i} \min_{u \in U} \sum_{i=1}^{m} q_i (F_i - F_i^*) =$$

$$= \max_{q_i} \left( \min_{u \in U} \left( \sum_{i=1}^{m} q_i B_i^* p_i, u \right) - \sum_{i=1}^{m} q_i F_i^* \right) = \max_{q_i} \left( -\beta \left\| \sum_{i=1}^{m} q_i B_i^* p_i \right\| - \sum_{i=1}^{m} q_i F_i^* \right),$$

де $\hat{u} = (\beta \sum_{i=1}^{m} q_i B_i^* p_i) \left\| \sum_{i=1}^{m} q_i B_i^* p_i \right\|^{-1}$.

Нехай $(\hat{q}_1, \ldots, \hat{q}_m) \in Arg \min L(q_1, \ldots, q_m)$, де $L(q_1, \ldots, q_m) = \beta \left\| \sum_{i=1}^{m} q_i B_i^* p_i \right\| + \sum_{i=1}^{m} q_i F_i^*$. Тоді оптимальна альтернатива має вигляд

$$\hat{u} = (\beta\sum_{i=1}^{m}\hat{q}_i B_i^* p_i)\left\|\sum_{i=1}^{m}\hat{q}_i B_i^* p_i\right\|^{-1}.$$

Далі розглянемо приклад розв'язування багатокритеріальної задачі методом головного критерію.

Нехай є два критерії, тобто $m=2$ і головним є перший, причому

$$F_i(\phi_i,u)=\left\|\phi_i-\overline{\phi}_i\right\|_H^2+\alpha_i^2\left\|u\right\|^2,\ i=\overline{1,2}.$$

Згідно методу головного критерію ми повинні назначити число $c>0$ і за оптимальну альтернативу взяти вектор

$$\hat{u}\in Arg\min_{u\in U}F_1(\phi_i,u),\ U=\{u \mid F_2(\phi_2,u)\le c\}.$$

Зауважимо, що для того, щоб задача була змістовною необхідно, щоб константа $c$ вибиралася з умови $c\ge\min F_2(\phi_2,u)$.

**Твердження 2.** Множина $U$ є опуклою, обмеженою, непорожньою множиною для чисел $c$, що задовольняють нерівності $c\ge\min F_2(\phi_2,u)$.

***Доведення.***

Оскільки $F_2(\phi_2,u)\ge\alpha_2^2\left\|u\right\|^2$, то множина $U$ міститься у множині $U_1=\{u \mid \alpha_2^2\left\|u\right\|^2\le c\}$, яка очевидно є обмеженою, а значить обмеженою є і множина $U$. Опуклість функціоналу $I_2(u)=F_2(\phi_2,u)$ випливає із опуклості функції $f(t)=I_2(u+tv)$ тому, що $f''(t)\ge\alpha_2^2(v,v)\ge0$. Крім того функціонал $I_2(u)$ є неперервним, а значить множина $U$ є замкненою. Очевидно, що також важливо знайти $\min I_2(u)=I_2(\hat{u})$.

Зауважимо, що функціонал $F_2(\phi_2,u)$ є неперервним, сильно опуклим та таким, що $\lim\limits_{\|u\|\to\infty}F_2(\phi_2,u)=\infty$.

Тоді існує єдина точка $\hat{u}_1\in Arg\min F_2(\phi_2,u)$, яка може бути знайдена із умови

$$\frac{d}{dt}I_2(u+tv)\Big|_{t=0}=0\quad\forall v\in H_1.$$

Оскільки $\dfrac{1}{2}\dfrac{d}{dt}I_2(u+tv)\Big|_{t=0}=(\phi_2-\overline{\phi}_2,\hat{\phi}_2)_H+a_2^2(u,v)$, де $\phi_2$ – розв'язок рівняння

$$\alpha_2(\hat{\phi}_2,\psi)=(\beta_2 v,\psi)_H.$$

Введемо вектор $p$ як розв'язок рівняння

$$a_2(p,\psi)=(\phi_2-\overline{\phi}_2,\psi)_H.$$

Тоді мають місце рівності:

$$(\phi_2 - \overline{\phi}_2, \hat{\phi}_2) = a_2(p, \hat{\phi}_2) = a_2(\hat{\phi}_2, p) = (B_2 v, p)_H = (v, B_2^* p)_{H_1} .$$

Враховуючи ці викладки одержимо, що $\hat{u}_1 = -a_2^{-2} B_2^* p$, де вектор $p$ знаходиться із системи рівнянь:

$$\begin{cases} a_2(\phi_2, \psi_1) = (B\hat{u}, \psi_1)_H + l_2(\psi_1); \\ a_2(p, \psi_2) = (\phi_2 - \overline{\phi}_2, \psi_2)_H , \end{cases} \tag{8.8}$$

для $\forall \psi_1, \psi_2 \in V$.

Таким чином, $\min I_2(u) = I_2(\hat{u}_1)$.

Введемо далі вектори $p_1, p_2 \in V$ як розв'язки системи рівнянь:

$$\begin{cases} a_i(\phi_i, \psi_i) = (B\hat{u}(\hat{\lambda}), \psi_i) + l_i(\psi_i); \\ a_i(p_i, \psi_{1i}) = (\phi_i - \overline{\phi}_i, \psi_{1i}), \forall \psi_i, \psi_{1i} \in V, i = \overline{1, 2}, \end{cases} \tag{8.9}$$

де

$$\hat{u}(\hat{\lambda}) = (-B_1^* p_1 - \hat{\lambda} B_2^* p_2)(\alpha_1^2 + \hat{\lambda}\alpha_2^2)^{-1} , \tag{8.10}$$

а $\hat{\lambda} = 0$, якщо виконується умова $I_2(\hat{u}(0)) \le c$, інакше число $\hat{\lambda}$ знаходимо з умови $I_2(\hat{u}(\hat{\lambda})) = c$.

**Твердження 3**. Нехай $\phi_1$, $\phi_2$, $p_1$, $p_2$ – розв'язки системи рівнянь (8.8), (8.9) і $c \le \min I_2(u)$. Тоді оптимальна альтернатива $u^*$ за методом головного критерію має вигляд $u^* = \hat{u}(\hat{\lambda})$.

***Доведення***.

Оскільки $U$ – обмежена, опукла, замкнена множина, функціонал $I_1(u) = F_1(\phi, u)$ є сильно опуклим, то існує єдина альтернатива $u^*$ така, що $I_2(u^*) = \min\limits_{u \in U} I_1(u)$, тобто існує число $\hat{\lambda} \ge 0$ таке, що функція Лагранжа

$$I_\lambda(u) = I_1(u) + \hat{\lambda} I_2(u)$$

досягає мінімуму на векторі $u^*$, тобто на $u^*$ виконується тотожність

$$\frac{d}{dt} I_\lambda(u^* + tv)\Big|_{t=0} = 0 \quad \forall v \in H_1 .$$

Аналогічно тому, як це було зроблено в твердженні 2, можливо показати, що

$$\frac{d}{dt} I_\lambda(u^* + tv) = (B_1^* p_1 + \hat{\lambda} B_2^* p_2 + (\alpha_1^2 + \hat{\lambda}\alpha_2^2)u^*, v) = 0 .$$

Звідси одержимо $u^* = \hat{u}(\hat{\lambda})$, що й потрібно було показати.

*Зауваження 1*. Із методу доведення твердження 3 випливає, що система рівнянь (8.8), (8.9) має єдиний розв'язок.

## 8.4. Рівновага за Штакельбергом

Нехай існує множина альтернатив $U_1 \times U_2$, де $U_i \subseteq H_i$, $i = \overline{1,2}$, $H_i$ – гільбертові простори, і є оцінки альтернатив вигляду $(F_1(\phi_1, u_1, u_2), F_2(\phi_2, u_1, u_2))$, де $F_i$ – деякі функціонали на добутку просторів $V_i \times H_1 \times H_2$. Гільбертові простори $V_i$, $i = \overline{1,2}$ – неперервно та щільно вкладені в гільбертовий простір $H = H^*$, а вектори $\phi_1$ та $\phi_2$ є розв'язками варіаційних рівнянь:

$$a_i(\phi_i, \psi_i) = (B_{1i}u_1, \psi_i) + (B_{2i}u_2, \psi_i) + l_i(\psi_i) \quad \forall \psi_i \in V_i$$

$$B_{1i} \in L(H_1, H),\ B_{2i} \in L(H_2, H),\ l_i \in V_i^*,$$

білінійні форми $a_i(\phi_i, \psi_i)$ такі ж як і в рівнянні (8.7).

Альтернативи $u_1 \in U_1$ та $u_2 \in U_2$ обираються двома особами незалежно, причому кожна з них прагне мінімізувати свій функціонал $F_i(\phi_i, u_1, u_2)$.

Припустимо, що одна із осіб є лідером і знає вигляд функціоналу $F_2(\phi_2, u_1, u_2)$.

*Означення 1.* Альтернатива $(u_1^*, u_2^*)$ називається ***рівновагою за Штакельбергом***, якщо $u_1^*$ знаходиться із умови

$$u_1^* \in Arg \min_{u_1 \in U_1} \max_{u_2 \in S(u_1)} I_1(u_1, u_2),$$

а вектор $u_2^*$ із умови $u_2^* \in S(u_1^*)$, де

$$S(u_1) = Arg \min_{u_2 \in U_2} I_2(u_1, u_2),\ I_i(u_1, u_2) = F_i(\phi_i, u_1, u_2),\ i = 1,2.$$

Припустимо, що функціонали $I_1$ та $I_2$ мають вигляд:

$$I_1(u_1, u_2) = \left\|\phi_1 - \bar{\phi}_1\right\|_H^2 + \alpha_{11}^2 \left\|u_1\right\|^2 + \alpha_{12}^2 \left\|u_2\right\|^2,$$

$$I_2(u_1, u_2) = \left\|\phi_2 - \bar{\phi}_2\right\|_H^2 + \alpha_{21}^2 \left\|u_1\right\|^2 + \alpha_{22}^2 \left\|u_2\right\|^2,$$

а $U_1 = H_1$, $U_2 = H_2$.

Знайдемо спочатку $S(u_1)$. Враховуючи попередні міркування, легко бачити, що множина $S(u_1)$ буде складатись із одного елемента $\hat{u}_2$, причому $\hat{u}_2 = -\alpha_{22}^{-2} B_{22} p_2$, де вектор $p_2$ знаходиться із розв'язку системи рівнянь:

$$\begin{cases} a_2(\phi_2, \psi_2) = (B_{12}u_1, \psi_2) + (B_{22}\hat{u}_2, \psi_2) + l_2(\psi_2), \\ a_2(p_2, \psi_3) = (\phi_2 - \bar{\phi}_2, \psi_3). \end{cases} \tag{8.11}$$

Зауважимо далі, що вектори $\phi_2$ та $p_2$ можливо подати у вигляді:

$$\phi_2 = \phi_2^{(1)} + \phi_2^{(2)},\ p_2 = p_2^{(1)} + p_2^{(2)},$$

де $\phi_2^{(i)}$, $p_2^{(i)}$, $i=\overline{1,2}$ є розв'язками систем рівнянь, що отримуються при $l_2=0$, $\overline{\phi}_2=0$ та $u_1=0$, відповідно.

Враховуючи неперервну залежність розв'язку системи (8.11) від векторів з правої частини, можна записати, що $p_2^{(1)}=Du_1$, де $D\in L(H_1,V_2)$.

Таким чином, будемо мати рівність

$$I_1(u_1,\hat{u}_2)=\left\|\phi_1-\overline{\phi}_1\right\|_H^2+\alpha_{11}^2\left\|u_1\right\|^2+\alpha_{12}^2\alpha_{22}^{-4}\left\|B_{22}^*Du_1+B_{22}^*p_2^{(2)}\right\|^2.$$

Очевидно, що існує єдиний вектор $\hat{u}_1$ такий, що $\hat{u}_1\in Arg\min I_1(u_1,\hat{u}_2)$ і для якого виконується тотожність

$$\frac{d}{dt}I_2(\hat{u}_1+tv_1,\hat{u}_2)\Big|_{t=0}\equiv 0\quad \forall v_1\in H_1. \tag{8.12}$$

Введемо вектор $p_1$ як розв'язок рівняння

$$a_1(p_1,\psi)=(\phi_1-\overline{\phi}_1,\psi)_H\quad \forall\psi\in V_1.$$

Тоді із умови (8.12) одержимо рівняння для визначення $\hat{u}_1$

$$(\alpha_{11}^2E+\alpha_{12}^2\alpha_{21}^{-4}D^*B_{22}B_{22}^*D)\hat{u}_1=-B_{11}^*p_1-D^*B_{22}B_{22}^*p_2^{(2)}. \tag{8.13}$$

Таким чином, одержали, що $\hat{u}_1$ визначається із рівняння (8.13), де $p_1$ визначається із розв'язку системи рівнянь:

$$\begin{cases}a_1(\phi_1,\psi_1)=(B_{11}\hat{u}_1,\psi_1)-\alpha_{22}^{-2}(B_{22}^*D\hat{u}_1,\psi_1)-\alpha_{22}^{-2}(B_{22}^*p_{22}^{(2)}\hat{u}_1,\psi_1)+l_1(\psi_1),\\ a_1(p_1,\psi_2)=(\phi_1-\overline{\phi}_1,\psi_2)\qquad \forall\psi_1,\psi_2\in V_1.\end{cases}$$

Якщо позначити $\hat{\hat{u}}_2=\hat{u}_2\big|_{u_1=\hat{u}_1}$, то точка рівноваги Штакельберга буде мати вигляд $(\hat{u}_1,\hat{\hat{u}}_2)$.

## 8.5. Наближені методи розв'язування екстремальних задач

Наближені методи апроксимації, що розглядалися раніше, зводились до ітераційних алгоритмів в нескінченновимірних просторах. Реалізація алгоритмів оптимізації функціоналів на ЕОМ вимагає зведення таких задач до проблем оптимізації в скінченновимірних просторах. Розглянемо далі відповідні процедури апроксимації та їх збіжність.

Нехай $X$ – нескінченновимірний простір, а $F(x)$ – функціонал, що визначений на $X$. Позначимо через $X_n$, $n=1,2,...$ послідовність множин із $X$.

**Твердження 1.** Нехай $\forall x\in X$ існує точка $x_n=x_n(x)\in X_n$, така що

1) $\overline{\lim\limits_{n\to\infty}}[F(x_n)-F(x)]\le 0$;

2) існує послідовність $\beta_n$ невід'ємних чисел таких, що $\beta_n\to 0$ при $n\to\infty$, та $F(x_n)-F(x)\le\beta_n$.

Тоді якщо позначимо через $F_n^*$, $F^*$ числа

$$F_n^* = \inf_{x_n \in X_n} F(x_n),\ F^* = \inf_{x \in X} F(x),$$

то виконується нерівність

$$\left|F_n^* - F^*\right| \le \beta_n .$$

***Доведення.*** Нехай $x$ – довільна точка із $X$. Тоді

$$F_n^* \le F(x_n) \le F(x) + \beta_n ,$$

тобто $F_n^* \le F^* + \beta_n$, або $-\beta_n + F_n^* \le F^* + \beta_n$, звідси $\left|F_n^* - F^*\right| \le \beta_n$, що й потрібно було довести.

Наслідок 1. Нехай $F(x)$ – неперервний функціонал, що визначений на множині $U$ нормованого простору $X$. $X_n$ – послідовність множин така, що $\forall x \in U$ знайдеться послідовність векторів $x_n \in U$ така, що $\|x_n - x\| \to 0$ при $n \to \infty$. Тоді $\lim\limits_{n\to\infty} F_n^* = F^*$, де $F_n^* = \inf\limits_{x \in X_n} F_n(x)$, $F^* = \inf\limits_{x \in U} F(x)$.

***Доведення.*** В силу неперервності функціоналу $F(x)$ одержимо

$$\overline{\lim_{n\to\infty}}[F(x_n) - F(x)] = \lim_{n\to\infty}[F(x_n) - F(x)] = 0 .$$

Крім цього, покладемо

$$\beta_n = \left|F(x_n) - F(x)\right|.$$

Тоді $\beta_n \to 0$ при $n \to \infty$ та $F(x_n) - F(x) \le \beta_n$, тобто виконуються умови 1) та 2) твердження 1.

Нехай $X$ – рефлексивний нормований простір, $X_n$ – послідовність множин із $X$, причому із того, що послідовність векторів $x_n \in X$ слабо збігається до вектора $x_0$ випливає, що $x_0$ належить деякій множині $U \subseteq X$.

Припустимо, що $\forall x \in U$ існує послідовність $x_n \in X$ така, що $x_n$ збігається до вектора $x$. Крім того якщо $x_n \in X$, $n = 1, 2, \ldots$, то існує константа $c > 0$ така, що $\|x_n\| \le c$.

**Твердження 2.** Нехай множини $\hat{X}_n = Arg\min\limits_{x \in X_n} F(x)$ та $\hat{U} = Arg\min\limits_{x \in U} F(x)$ – непорожні. Тоді якщо $\hat{x}_n \in \hat{X}_n$, то існує вектор $\hat{x} \in \hat{U}$ такий, що $\hat{x}_n$ слабо збігається до $\hat{x}$.

***Доведення.*** Нехай $\hat{x}_n \in \hat{X}_n$. Виділимо із послідовності $\hat{x}_n$ слабо збіжну підпослідовність $\hat{x}_{n_k}$, що збігається до вектора $x_0$. Тоді в силу слабої напівнеперервності функціоналу $F(x)$ виконується

$$\lim_{k\to\infty} F(\hat{x}_{n_k}) \ge F(x_0).$$

Нехай $\hat{x} \in \hat{U}$. Тоді існує послідовність $x_n \in X_n$ така, що $x_n$ збігається до $\hat{x}$. В силу нерівності $F(\hat{x}_{n_k}) < F(x_n)$, одержимо

$$F(x_0) \le \underline{\lim}_{k\to\infty} F(\hat{x}_{n_k}) \le \lim_{k\to\infty} F(x_n) = F(\hat{x}).$$

Оскільки $F(x_0) \ge F(\hat{x}) = \inf_{x\in U} F(x)$, то $x_0 \in Arg\min_{x\in U} F(x)$, що й потрібно було довести.

**Твердження 3.** Нехай функціонал $F(x)$ визначений на множині $U \subseteq X$ та існує друга похідна $F''(x)$, причому для $x \in U$ виконується нерівність:

$$\left|F''(x)(h)\right| \ge \rho\left(\|h\|\right),$$

де $h \in X$, $\rho(\cdot)$ – деяка невід'ємна функція, що визначена на $R_+$. Тоді якщо $\hat{x} \in \hat{U}$ і є внутрішньою точкою множини $U$, то має місце нерівність

$$\left|F(x) - F(\hat{x})\right| \ge \frac{1}{2}\rho(\|x-\hat{x}\|).$$

***Доведення.*** Розглянемо функцію

$$g(t) = F(\hat{x} + t(x-\hat{x})),\ t \in [0;1].$$

Тоді має місце рівність

$$g(t) = g(0) + g'(0)t + \frac{1}{2}g''(\theta)t^2, \text{ де } \theta \in [0,1].$$

Оскільки $\hat{x} \in Arg\min F(x)$, то $g'(0) = 0$. Крім того, оскільки $g(1) = F(x)$, $g(0) = F(\hat{x})$, то одержимо

$$\left|F(x) - F(\hat{x})\right| = \frac{1}{2}\left|F''(\hat{x} + \theta(x-\hat{x}))(x-\hat{x})\right| \ge \frac{1}{2}\rho(\|x-\hat{x}\|),$$

що й потрібно було довести.

Наслідок 2. Нехай функція $\rho(t)$ – неперервна, а рівняння $\rho(t) = 0$ має єдиний розв'язок $t = 0$. Тоді якщо

$$\lim_{n\to\infty} F(x_n) = F(\hat{x}),$$

то $x_n$ збігається сильно до вектора $\hat{x}$.

Нехай $X$ – рефлексивний нормований простір, $X_n$ – послідовність скінченновимірних підпросторів простору $X$ із скалярним добутком $(x_1, x_2)_n$. Позначимо через $p_n$ лінійні неперервні оператори, що діють із простору $X_n$ в $X$ і припустимо, що існує константа $c > 0$ така, що $\|p_n\| \le c$.

*Означення 1.* Кажуть, що ***послідовність*** $X_n$ ***апроксимує простір*** $X$, якщо послідовність $p_n x_n$ слабо збігається до вектора $x_0$ простору $X$ та, якщо $x \in X$, то існує послідовність $x_n \in X_n$ така, що норма вектора $x_n$ рівномірно обмежена і така, що виконується умова: $p_n x_n$ сильно збігається до вектора $x$.

Нехай тепер $U$ – замкнута й опукла множина простору $X$.

*Означення 2.* Кажуть, що ***послідовність замкнених опуклих множин*** $U_n \subset X_n$ ***апроксимує множину*** $U$, якщо із того, що $p_n x_n$ (де $x_n \in U_n$) слабо збігається в $X$ до вектора $x_0$, то $x_0 \in U$, та для $\forall x \in U$ існує послідовність $x_n \in U_n$ з рівномірно обмеженою нормою така, що $p_n x_n$ сильно збігається до $x$ в просторі $X$.

Задамо далі послідовність функцій $F_n(x_n)$ на скінченновимірних просторах $X_n$.

*Означення 3.* Кажуть, що ***послідовність власних опуклих функцій*** $F_n(x_n)$ ***апроксимує власний опуклий функціонал***, якщо:

1) із того, що $p_n x_n$ слабо збігається в $X$ до вектора $x_0$ випливає, що

$$\underline{\lim}_{n\to\infty} F_n(x_n) \geq F(x);$$

2) із того, що $p_n x_n$ сильно збігається до вектора $x \in X$ випливає, що

$$\lim_{n\to\infty} F_n(x_n) = F(x).$$

Нехай $F(x)$ власний опуклий функціонал, що визначений на замкненій опуклій множині $U \subseteq X$ і такий, що

$$\hat{U} = Arg\min_{x\in U} F(x) \neq 0.$$

Припустимо також, що $F_n(x_n)$ – власні опуклі функціонали на скінченновимірних просторах $X_n, n = 1, 2, ...$ та

$$\hat{U}_n = Arg\min_{x_n \in U_n} F_n(x_n),$$

де $U_n$ – послідовність замкнених опуклих множин із $X_n$.

*Означення 4.* Будемо казати, ***що послідовність множин*** $\hat{U}_n$, $n = 1, 2, ...$ ***із скінченновимірного простору*** $X_n$ ***апроксимує множину*** $\hat{U} \subset X$, якщо:

1) послідовність $U_n$ апроксимує множину $U$;
2) послідовність функцій $F_n(x_n)$ апроксимує функціонал $F(x)$.

*Зауваження 1.* Для того, щоб знайти множину $\hat{U}_n$ потрібно розв'язати задачу на мінімум функції в скінченновимірному просторі. Позначимо через $p_n U_n$ множину векторів із простору $X$, кожен із яких має вигляд $p_n x_n$, $x_n \in U_n$.

*Означення 5.* Будемо казати, що ***множина*** $p_n U_n$ ***слабо (сильно) збігається до множини*** $U$, якщо послідовність $p_n x_n$, де $x_n \in U_n$, слабо (сильно) збігається до деякого елемента $x \in U$.

**Твердження 4.** Нехай послідовність власних опуклих функціоналів $F_n(x_n)$, $x_n \in X_n$, $n = 1, 2, ...$ апроксимує власний опуклий функціонал $F(x)$, а опуклі замкнені рівномірно обмежені множини $U_n$ – апроксимують опуклу замкнену множину $U$. Тоді множина $\hat{U}_n$ слабо апроксимує множину $\hat{U}$.

***Доведення.*** Нехай $\hat{x}_n \in \hat{U}_n$. Оскільки $\hat{U}_n$ є рівномірно обмежена множина, то існує константа $c > 0$ така, що $\|\hat{x}_n\| \le c$. В силу рівномірної обмеженості оператора $p_n$ маємо, що $\|p_n \hat{x}_n\| \le c_1$, $c_1 > 0$. Оскільки простір $X$ – рефлексивний, то можливо виділити слабо збіжну підпослідовність $p_{n_k} \hat{x}_{n_k}$, що збігається до елемента $x_0 \in U$.

Покажемо, що $x_0 \in \hat{U}$. Для довільного вектора $x \in U$ існує послідовність векторів $x_n \in U_n$ таких, що $p_n x_n$ сильно збігається до $x_0$. Звідси одержуємо, що $F_n(x_n)$ збігається до $F(x_0)$.

Крім цього із того, що $p_{n_k} \hat{x}_{n_k}$ слабо збігається до $x_0$ випливає, що

$$\underline{\lim} F_{n_k}(\hat{x}_{n_k}) \ge F(x_0),$$

але

$$F_{n_k}(\hat{x}_{n_k}) \le F_{n_k}(x_{n_k}),$$

тобто

$$F(x_0) \le \underline{\lim} F_{n_k}(\hat{x}_{n_k}) \le \lim_{k \to \infty} F_{n_k}(x_{n_k}) = F(x).$$

Отже, для довільного вектора $x \in U$ виконується нерівність $F(x_0) \le F(x)$, а це значить, що $x_0 \in \hat{U}$.

*Зауваження 2.* Умову рівномірної обмеженості множини $U_n$ можливо замінити наступною умовою: існує монотонно неспадна функція $\rho(t) \ge 0$ $t \in [0; \infty)$ така, що для $\forall n = 1, 2, ...$ виконується умова

$$F_n(x_n) \ge \rho\left(\|x_n\|_n\right).$$

Дійсно, якщо $\hat{x}_n \in \hat{U}_n$, а $x$ – довільний елемент із $U$, то існує послідовність $x_n$ така, що $p_n x_n$ сильно збігається до $x$. З іншої сторони

$$F_n(x_n) \ge F_n(\hat{x}_n) \ge \rho\left(\|\hat{x}_n\|_n\right).$$

Оскільки $F_n(x_n)$ збігається до $F(x)$, то $F_n(x_n)$ – обмежена константою $\beta$, тобто $\rho(\|\hat{x}_n\|_n) \le \beta$, а значить виконується нерівність

$$\|\hat{x}_n\|_n \le \rho^{-1}(\beta).$$

**Приклад 1.** Нехай множина альтернатив співпадає із гільбертовим простором $U$ з елементами $u$. Позначимо через $(F_1(u), F_2(u))$ оцінку альтернативи $u \in U$, де $F_1(u)$, $F_2(u)$ функціонали, що мають вигляд:

$$F_i(u) = \|y_i - C_i\phi_i\|^2 q_{1i}^2 + q_{2i}^2(u,u), \quad i = 1, 2,$$

де $y_i$ – вектори із гільбертових просторів $H_1$, $H_2$, відповідно, $\phi_i$ – вектори із гільбертового простору $H_+$, що неперервно та щільно вкладений в $H_0$, причому $H_0 = H_0'$, та які задовольняють варіаційним рівнянням:

$$a_i(\phi, \phi) = (B_i u, \phi) \;\; \forall \psi \in H_+,$$

де $a_i(\phi, \psi)$ – неперервні на $H_+$ форми такі, що існують числа $\alpha_i^2$, $\alpha_i \ne 0$, що виконуються нерівності

$$a_i(\phi, \psi) \ge \alpha_i^2(\phi, \psi), \quad i = 1, 2,$$

$$B_i \in L(U, H_-), \quad H_- = H_+',$$

$$C_i \in L(H_+, H_i), \quad i = 1, 2,$$

$q_{1i}, q_{2i}, \quad i = 1, 2$ – відмінні від нуля числа.

*Означення 6.* Альтернатива $\hat{u}$ називається ***оптимальною за Слейтером***, якщо не існує альтернативи $v$ такої, що виконуються нерівності:

$$F_1(v) < F_1(\hat{u}),\; F_2(v) < F_2(\hat{u}).$$

Будемо шукати оптимальні за Слейтером альтернативи методом головного критерію. Введемо множину

$$U(\alpha) = \{u \,|\, F_2(u) \le \alpha\}.$$

Очевидно, що для того, щоб множина $U(\alpha)$ була непорожньою, достатньо, щоб число $\alpha$ вибиралося із умови

$$\alpha > \inf_{u \in U} F_2(u) = F_2(\hat{u}).$$

Оскільки $F_2(u)$ – опуклий функціонал, то множина $U(\alpha)$ – опукла. Легко бачити, що вона є також обмеженою та замкненою. Тоді існує вектор $\hat{u}_\alpha$ (він є єдиним), для якого

$$\min_{u \in U(\alpha)} F_1(u) = F_1(\hat{u}_\alpha).$$

Вектор $\hat{u}_\alpha$ називають ***оптимальною альтернативою за методом головного критерію***. Приведемо алгоритм наближеного обчислення вектора

$\hat{u}_\alpha$, що зводиться до задачі мінімізації функції в скінченновимірному просторі. Нехай простір $U$ – сепарабельний, а $\{l_i\}_{i=1}^{\infty}$ – ортонормований базис в $U$. Через $F_n$ позначимо лінійний підпростір простору $U$, що породжений векторами $l_1, ..., l_n$.

Введемо множину

$$U_n(\alpha) = \{u_n \in F_n \big| F_2(u_n) \le \alpha\}.$$

***Наближеним значенням вектора*** $\hat{u}_\alpha$ назвемо вектор $\hat{u}_n \in Arg \min\limits_{u_n \in U_n(\alpha)} F_1(u_n)$.

Можна показати, що вектор $\hat{u}_n$ єдиний і, що $\hat{u}_n$ сильно збігається до $\hat{u}$.

*Зауваження 3.* Вектор $u_n$ можна подати у вигляді

$$u_n = \sum_{k=1}^{n} \alpha_k l_k .$$

Тоді вектори $\phi_i(u_n)$ будуть мати вигляд:

$$\phi_i(u_n) = \sum_{k=1}^{n} \alpha_k \phi_i(l_k),$$

де $\phi_i(l_k)$ – розв'язки рівнянь:

$$a_i(\phi_i(l_k), \psi) = (B_i l_k, \psi) \quad \forall \psi \in H_+, \quad i = 1, 2 .$$

Функціонали $F_i(u_n),\ i = 1, 2$ будуть мати вигляд:

$$F_i(u_n) = \left\| y_i - \sum_{k=1}^{n} C_i \phi_i(l_k) \alpha_k \right\|^2 q_{1i}^2 + q_{2i}^2 \sum_{k=1}^{n} \alpha_k = f_i(\alpha_1, ..., \alpha_n) .$$

Таким чином, $\hat{u}_n = \sum\limits_{k=1}^{n} \hat{\alpha}_k l_k$, де $\hat{\alpha}_1, ..., \hat{\alpha}_n$ знаходяться як розв'язок задачі на мінімум функції $f_1(\alpha_1, ..., \alpha_n)$ при обмеженнях $f_2(\alpha_1, ..., \alpha_n) \le \alpha$.

**Приклад 2.** Задамо функціонали $F_i(u), i = 1, 2$ у вигляді:

$$F_i(u) = \sum_{k=1}^{N_i} \left( y_k - \int_0^1 g_{ki}(x) \phi_i(x) dx \right)^2 q_{1i}^2 + q_{2i}^2 \int_0^1 u^2(x) dx , \quad i = 1, 2 ,$$

де $g_{ki}(x)$ – відомі неперервні функції на відрізку $[0; 1]$,

$\phi_i(x)$ – узагальнені розв'язки рівнянь:

$$-\frac{d}{dx} a_i(x) \frac{d}{dx} \phi_i(x) + b_i(x) \phi_i(x) = c_i(x) u(x) ,$$

$$\phi_i(0) = \phi_i(1) = 0 ,$$

$a_i(x)$, $b_i(x)$, $c_i(x)$ – неперервні функції на відрізку $[0;1]$, причому $a_i(x)$, $b_i(x)$ – невід'ємні та існує $\gamma > 0$ таке, що виконується нерівність $a_i(x) \geq \gamma > 0$. Функція $u(x)$ належить простору $L_2(0;1)$.

Розіб'ємо відрізок [0;1] точками $x_k = k/n,\ k = \overline{0, n-1}$, та позначимо через $F_{in}(v_n)$ функції:

$$F_{in}(v_n) = \sum_{k=1}^{N_i} \left( y_k - n \sum_{s=0}^{n} g_{ki}^{(s)} \cdot \phi_{is} \right)^2 q_{1i}^2 + q_{2i}^2 n \sum_{k=0}^{n} u_k^2 ,$$

де $g_{ki}^{(s)} = g_{ki}\left(s/n\right)$, $v_n = (u_0, ..., u_n)$, $\phi_{is} = \phi_i\left(s/n\right)$ – розв'язки крайової задачі для різницевого рівняння:

$$\nabla^- a_{is} \nabla^+ \phi_{is} + b_{is} \phi_s = c_{is} u_s, \quad s = \overline{1, n-1};\ \phi_i(0) = \phi_i(1) = 0 ,$$

$$a_{is} = a_i\left(s/n\right),\ b_{is} = b_i\left(s/n\right),\ c_{is} = c_i\left(s/n\right),$$

$$\Delta^+ f\left(s/n\right) = \left(f\left((s+1)/n\right) - f\left(s/n\right)\right) n ,$$

$$\Delta^- f\left(s/n\right) = \left(f\left((s-1)/n\right) - f\left(s/n\right)\right) n .$$

Апроксимуємо задачу:

$$\min_{u \in U(\alpha)} F_1(u) = F_1(\hat{u}_\alpha)$$

задачею:

$$\min_{v_n \in U_n(\alpha)} F_{1n}(v_n) = F_{1n}(\hat{v}_n) ,$$

де $u_n(\alpha) = \{ v_n \mid F_{2n}(v_n) \leq \alpha \}$.

Якщо $\hat{v}_n = (\hat{u}_0, ..., \hat{u}_{n-1})$, то наближене значення функції $\hat{u}_\alpha(x)$ визначимо як функцію:

$$\hat{u}_{n\alpha}(x) = \sum_{k=0}^{n-1} \hat{u}_k \theta_k(x) ,$$

де $\theta_k(x) = \begin{cases} 1, & x \in \left[\frac{k}{n}, \frac{k+1}{n}\right], \\ 0, & x \notin \left[\frac{k}{n}, \frac{k+1}{n}\right]. \end{cases}$

Можна показати, що $\hat{u}_{n\alpha}(x)$ збігається в просторі $L_2(0;1)$ до функції $\hat{u}(x)$.

*Зауваження 4.* Позначимо через $P_{ki}(x)$ розв'язки рівнянь:

$$-\frac{d}{dx} a_i(x) \frac{d}{dx} P_{ki}(x) + b_i(x) P_{ki}(x) = g_{ki}(x) ,$$

$$P_{ki}(0) = P_{ki}(1) = 0 .$$

Тоді функціонали $F_i(u),\ i = 1, 2$ запишуться у вигляді:

$$F_i(u)=\sum_{k=1}^{N_i}\left(y_k-\int_0^1 P_{ki}(x)c_i(x)u(x)dx\right)^2 q_{1i}^2+q_{2i}^2\int_0^1 u^2(x)dx\,,\quad i=1,2\,.$$

*Зауваження 5.* Функції $F_{in}(v_n)$, $i=\overline{1,n}$ можна записати у вигляді:

$$F_{in}(v_n)=\sum_{k=1}^{N_i}\left(y_k-n\sum_{s=0}^{n}P_{ki}\left(\frac{s}{n}\right)c_i\left(\frac{s}{n}\right)u_s\right)^2 q_{1i}^2+n\sum_{s=0}^{n}u_s^2\,.$$

Таким чином, при чисельному розв'язуванні екстремальних задач в нескінченновимірних просторах потрібно зробити наступні дії:

або

1) сформулювати необхідні (а якщо можливо і достатні) умови екстремуму у вигляді рівнянь або нерівностей в нескінченновимірних просторах;
2) вибрати наближений алгоритм розв'язування цих рівнянь та нерівностей;
3) апроксимувати ці наближені алгоритми алгоритмами в скінченновимірному просторі;
4) довести збіжність цих алгоритмів;

або

1) апроксимувати функціонали та множини, на яких розглядається екстремальна задача, функціями та множинами в скінченновимірному просторі;
2) вибрати один із алгоритмів наближеного розв'язування апроксимованої задачі;
3) довести збіжність.

## Запитання та вправи для самостійної роботи до розділу 8.

1. Доведіть, що функціонал $F(u)=\int_0^1 u^2(t)dt$ є неперервним в метриці простору неперервних функцій $C[0;1]$. Покажіть, що на множині $U=\{u\,|\,u(0)=0,u(1)=1\}$ в просторі $C[0;1]$ функціонал $F(u)$ не досягає своєї нижньої границі $F_*=0$.
2. Доведіть, що функціонал $F(u)=\|u-\bar{u}\|$ на довільній замкненій опуклій множині рефлексивного простору Банаха досягає своєї нижньої границі.
3. Доведіть, що функціонал $F(u)=\int_0^1 u^4(t)dt$ не є неперервним в метриці простору $L_2(0;1)$. Чи буде він напівнеперервним знизу в цій метриці?

4. Нехай $x(t)$ така функція, що $\dot{x}(t) = u(t)$, де $u(t) \in L_2(0;1)$, $x(0) = 0$. Покажіть, що функціонал $F(u) = x^2(t) + \int_0^1 x^2(t)dt$ є опуклим в просторі $L_2(0;1)$ та диференційованим. Знайдіть похідну Гато цього функціоналу.
5. Нехай $F(u) = \int_0^1 \left(u^2(t) - ax^2(t)\right)dt$, $\dot{x}(t) = u(t)$, $x(0) = 0$, $u(t) \in L_2(0;1)$. При якому значенні параметра $a$ функціонал $F(u)$ буде опуклим?
6. Знайдіть похідну Фреше оператора $F(u)$ в точці $u_0 = \cos x$ в просторі неперервних на $[0;\pi]$ функцій, де $F(u) = \sin u(x)$.
7. Знайдіть похідну Фреше оператора $F(u) = u(x) + \int_0^1 f(x, u(t), t)dt$ в просторі неперервних на $[0;1]$ функцій.
8. Нехай $X$ – простір Гільберта. Покажіть, що функціонал $f(x) = \|x\|$ є опуклим та неперервним в $X$ і для нього в точці $0$ не існує похідної. Знайдіть субдиференціал функціоналу $f(x)$.
9. Нехай $x(t) = (x_1(t), \ldots, x_n(t))$, причому $\dot{x}(t) = u(t)$, $x(0) = 0$, $u(t) \in L_2(0;1)$. Покладемо $F(u) = \|x(T)\|$, де $\|x(T)\| = \left\{\sum_{i=1}^{n} x_i^2(T)\right\}^{\frac{1}{2}}$. Покажіть, що функціонал $F(u)$ – опуклий та недиференційований. Знайдіть його субдиференціал.
10. Нехай оператор $F(x)$ заданий в просторі $C(0;1)$ наступним чином:

$$F(x) = \int_0^1 x(t)x(s)ds + f(t),\ f(t) \in C(0;1).$$

При якій умові на функцію $f(x)$ існує нерухома точка оператора $F(x)$? Знайдіть цю точку.

11. Нехай функція $x(t)$ є розв'язком рівняння $\dot{x}(t) = -x(t)a^2(t)$, $x(0) = x_0$, $x_0 > 0$, $a(t)$ – неперервна на відрізку $[0;T]$ функція. Застосуйте модифікований алгоритм Ньютона. Порівняйте функції, одержані з допомогою ітераційного процесу з точним розв'язком рівняння. ***Вказівка:*** В просторі $C(0;T)$ ввести нелінійний оператор, що діє за правилом $F(x) = (\dot{x}(t) + x^2(t)a^2(t), x(0))$.
12. Нехай функція $x(t)$ є розв'язком рівняння $\dot{x}(t) = u(t)$, $x_0 = 0$, де $u(t)$ – функція із простору $L_4(0;1)$ інтегрованих з четвертою степеню функцій.

Введемо функціонал $F(u)=(x(1)-x_0)^4+\gamma^2\int\limits_0^1 u^4(t)dt$. Покажіть, що в $L_4(0;1)$ існує єдина точка мінімуму цього функціоналу і функціонал $F(u)$ є диференційованим. Застосуйте:

а) модифікований метод Ньютона для розв'язування рівняння $F'(u)=0$;

б) метод градієнтного спуску.

Доведіть збіжність цих методів.

13. Нехай функціонал $F(u)$ заданий у вигляді $F(u)=(x(1)-x_0)^4$, де $x(t)$ є розв'язком рівняння $\dot{x}(t)=u(t)$, $x_0=0$, $u(t)\in L_4(0;1)$. Введіть множину $U=\left\{u\left|\int\limits_0^1 u^4(t)dt\le\gamma^2\right.\right\}$. Доведіть, що на множині $U$ існує точка мінімуму функціоналу $F(u)$. Для наближеного знаходження точки мінімуму застосуйте метод проекції градієнта та метод умовного градієнта. Доведіть збіжність цих методів.

14. В просторі $L_4(0;1)$ виберіть базис $e_1(t),...,e_n(t),...$ Наближене значення функції $u(t)$ знайдіть у вигляді $u_n(t)=\sum\limits_{k=1}^n\alpha_k e_k(t)$. На множині функцій $u_n(t)$ розв'яжіть задачу про мінімум функціоналу $F(u)$. Доведіть збіжність $\hat{u}_n(t)$ до функції $\hat{u}(t)$, де $\hat{u}_n(t)\in Arg\min F(u_n)$, $\hat{u}(t)\in Arg\min F(u)$. Функція $F(u)$ визначена в задачі 12.

# Література

1. *Аввакумов В.Г.* Методы нескалярной оптимизации и их приложения / Аввакумов В.Г. – К.: Выща школа, 1990. – 188 с.
2. *Алексеев В.М.* Сборник задач по оптимизации. Теория. Примеры. Задачи: учебник для студентов вузов / В.М. Алексеев, Э.М. Галеев, В.М. Тихомиров. –[3-е изд., испр.]. – М.: ФИЗМАТЛИТ, 2008. – 255 с. – (Класич. ун-кий учебник / Моск. гос. ун-т им. М.В. Ломоносова).
3. *Аттеков А В.* Введение в методы оптимизации: учеб. пособие / Аттеков А.В., Зарубин В.С., Капатников А.Н. – М.: Финансы и статистика: ИНФРА-М, 2008. – 270с.
4. *Ашманов С.А.* Линейное программирование / Ашманов С.А. – М.: Наука, 1981. – 304 с.
5. *Ашманов С.А.* Теория оптимизации в задачах и упражнениях / С.А. Ашманов, А.В. Тимохов. – М.: Наука, 1991. – 446 с.
6. *Банди Б.* Методы оптимизации. Вводный курс / Банди Б.; пер. с англ. О.В. Шихеевой; под ред. В.А. Волынского. – М.: Радио и связь, 1988. – 128 с.
7. *Бартіш М.Я.* Методи оптимізації. Теорія і алгоритми / Бартіш М.Я.: [навч. посібник]. – Львів: Видавничий центр ЛНУ імені Івана Франка, 2006. – 223 с.
8. *Бейко И.В.* Методы и алгоритмы решения задач оптимизации / И.В. Бейко, Б.Н. Бублик, П.Н. Зинько. – К.: Вища школа, 1983. – 512 с.
9. *Бирюков С. И.* Оптимизиция. Элементы теории. Численные методы: [учеб. пособие] / Бирюков С.И. – М.: МЗ–Пресс, 2003. – 248с.
10. *Благодатских В.И.* Введение в оптимальное управление: Линейная теория: учебник [для студ. вузов] / Благодатских В.И. – М.: –Высшая школа, 2001. – 239 с.
11. *Бобылев Н.А.* Методы нелинейного анализа в задачах негладкой оптимизации / Н.А. Бобылев, В.С. Климов. – М.: Наука, 1992. – 204 с.
12. *Бондаренко М.Ф.* Оптимізаційні задачі в системах прийняття рішень / М.Ф. Бондаренко, А.М. Гвоздинський. – Х.: 1998. – 215 с. – (Ун-т змісту та методів навчання, Харків. держ. техн. ун-т радіоелектроніки).
13. *Васильев Ф.П.* Численные методы решения экстремальных задач: учеб. пособие [для студ. вузов, обуч. по спец. «Прикладная матем.». –2-е изд. переб. и доп.] / Васильев Ф.П. – М.: Наука, 1988. – 552 с.
14. *Васильев Ф.П.* Методы решения экстремальных задач / Васильев Ф.П. – М.: Наука, 1981. – 400 с.
15. *Васильев Ф.П.* Линейное программирование / Ф.П. Васильев, А.Ю. Иваницкий. – М.: Факториал Пресс, 2008. – 328 с.
16. Введение в методы оптимизации и теорию технических систем: Учеб. пособие / А.В. Усов, Г.А. Оборский, Ю.А. Морозов, К.А. Дубров. – О.: Астропринт, 2005. – 493 с.

17. *Верес М.М.* Мінімаксні методи оцінювання в лінійних задачах із параметром / М.М. Верес, О.Г. Наконечний. – К.: ВПЦ «Київський університет», 2007. – 115 с.
18. *Волошин О.Ф.* Моделі та метоли прийняття рішень: навч. посібник [для студ. вищих навч. закл.; –2-ге вид., перероб. та допов.] / О.Ф. Волошин, С.О. Мащенко. – К.: ВПЦ «Київський університет», 2010. – 336 с.
19. *Габасов Р.Ф.* Конструктивные методы оптимизации / Р.Ф. Габасов, Ф.М. Кирилова, А.И. Тятюшкин. – Минск: Из-во «Университетское», 1984. –ч. 1. Линейные задачи, 1984. – 214 с.
20. *Габасов Р.Ф.* Конструктивные методы оптимизации / Р.Ф. Габасов, Ф.М. Кирилова, О.И. Костюкова, В.М. Ракецкий. – Минск: Из-во «Университетское», 1984. –ч. 4. Выпуклые задачи, 1987. – 222 с.
21. *Габасов Р.Ф.* Методы оптимизации: учеб. пособие. – [2-е издание] / Р.Ф. Габасов, Ф.М. Кириллова. – Минск: Изд-во БГУ, 1981. – 350 с. ил.
22. *Гайна Г.А.* Методи оптимізації: алгоритми, приклади, задачі: навч. посібник [для студ. усіх спеціальн. напряму підготовки «Комп'ютерні науки»] / Гайна Г.А. – К.: КНУБА, 2005. – 144 с. – (Київ. націон. ун-т будівн. і архітектури).
23. *Гилл Ф.* Практическая оптимизация / Ф. Гилл, У. Мюррей, М. Райт. – М.: Мир, 1985. – 509 с.
24. *Горчакова І.А.* Математичне програмування: [навч. посібник] / Горчакова І.А. – Донецьк: ІПШІ «Наука і освіта», 2004. – 87с. – (Донецький держ. ін-т штучного інтелекту).
25. *Грешилов А.А.* Метематические методы принятия решений: [учеб. пособие для вузов] / Грешилов А.А. – М.: Из-во МГТУ им. Н.Э. Баумана, 2006. – 584 с.
26. *Губанов В.А.* Введение в системный анализ: [учебное пособие] / В.А. Губанов, В.В. Захаров, А.Н. Коваленко. – Санкт-Пет. ГУ, 1988. – 232 с.
27. *Гупал А.М.* Оптимальные процедуры распознавания / А.М. Гупал, И.В. Сергиенко. – К.: Наук.думка, 2008. – 232 с. – (НАН Украины. Ин-т кибернетики им. В.М. Глушкова).
28. *Давыдов Э.Г.* Исследование операций / Давыдов Э.Г. – М.: Высшая школа, 1990. – 383 с.
29. *Демьянов В.Ф., Васильев Л.В.* Недифференцируемая оптимизация / В.Ф. Демьянов, Л.В. Васильев. – М.: Наука, 1981. – 384 с.
30. *Демьянов В.Ф.* Основы негладкого анализа и квазидифференциальное исчисление / В.Ф. Демьянов, А.М. Губинов. – М.: Наука, 1990. – 431 с.
31. Дослідження операцій і методи оптимізації: навч. посібник [для студ. вищих навч. закладів] / М.Є. Корольов, В.І. Павленко, О.В. Савіна, А.Г. Тимошенко. – К.: Університет «Україна», 2007. – 177 с. – (Відкритий міжнародний ун-т розвитку людини «Україна».).
32. *Дякон В.М.* Математичне програмування: навч. посібник [для студ. вищих навч. закл.] / В.М. Дякон, Л.Є. Ковальов; В.М. Михайленко заг.

ред.; Європейський ун-т. – К.: Вид-во Європейського ун-ту, 2004. – 497 с.

33. *Евтушенко Ю.Г.* Методы решения экстремальных задач и их применение в системах оптимизации / Евтушенко Ю.Г. – М.: Наука, 1982. – 432 с.
34. *Еремин И.И.* Линейная оптимизация и системы линейных неравенств: учеб. пособие [для студ. вузов, обучающихся по спец. «Математика», «Прикладная математика и информатика», «Математические методы в економике»] / Еремин И.И. – М.: Академия, 2007. – 250 с. – (Университетский учебник. Серия «Прикладная математика и информатика»).
35. *Жиглявский А.А.* Методы поиска глобального экстремума / Жиглявский А.А. – М.: Наука, 1991. – 248 с.
36. *Жиляков В.И.* Основы оптимального управления: [учеб. пособие]. ч.1. / В.И. Жиляков, Р.Ю. Ткачев; Донбасский гос. техн. ун-т –Алчевск: ДОНГТУ, 2008. – 112 с.
37. *Жилинскас А.Г.* Глобальная оптимизация. Аксиоматика статических моделей, алгоритмы, применения / Жилинскас А.Г. – Вильнюс: Мокслас, 1986. – 166 с.
38. *Жилинскас А.Г.* Поиск оптимума: компьютер расширяет возможности / Жилинскас А.Г. – М.: Наука, 1989. – 128 с.
39. *Зайченко Ю.П.* Дослідження операцій / Зайченко Ю.П. – Київ: ЗАТ «Віпол», 2000. – 688 с.
40. *Зайченко Ю.П.* Дослідження операцій. Підручник / Зайченко Ю.П. – К.: Видавничий Дім «Слово», 2006. – 816 с.
41. *Зайченко О.Ю.* Дослідження операцій. Збірник задач / О.Ю. Зайченко, Ю.П. Зайченко. – К.: Видавничий Дім «Слово», 2007. – 472 с.
42. *Згуровський М.З.* Системний аналіз: Проблеми, методологія, застосування / М.З. Згуровський, Н.Д. Панкратова. – К.: Наук. думка, 2005. – 743 с.
43. *Згуровський М.З., Панкратова Н.Д.* Основи системного аналізу. Навч. підручник / М.З. Згуровський, Н.Д. Панкратова. – К.: Видавнича група BHV, 2007. – 544 с.
44. *Зеленський К.Х.* Математичне програмування: навч. посібник [для дистанц. навчання] / Зеленський К.Х. – К.: Університет «Україна», 2007. – 241 с. – (Відкритий міжнародний ун-т розвитку людини «Україна». Інститут дистанційного навчання).
45. *Івченко І.Ю.* Математичне програмування: навч. посібник [для студ. вищих навч. закладів] / Івченко І.Ю. – К.: Центр учбової літератури, 2007. – 232 с.
46. *Ириарт-Уррути Ж.Б.* Оптимизация и выпуклый анализ. Сборник задач и упражнений / Ириарт-Уррути Ж.Б. – К.: КИТ, 2004. – 376 с.
47. *Карманов В. Г.* Математическое программирование: [учеб. пособие; –5-е изд. испр.] / Карманов В.Г. – М.: ФИЗМАТЛИТ, 2000, –263 с.
48. *Катренко А.В.* Системний аналіз: [підручник] / Катренко А.В. – Львів:

«Новий світ – 2000», 2009. – 396 с.

49. *Киселева Е.М.* Численные методы оптимизации / Киселева Е.М., Шевелева А.Е.; Дніпропетровський держ. ун-т. – Д.: ДДУ, 1997. – 160 с.
50. *Ковалев В.Г.* Математическое программирование (линейные задачи): [учеб. пособие] / В.Г. Ковалев, А.Р. Наринян, В.А. Поздеев; Европейский ун-т. – К.: Из-во Европейского университета, 2004. – 169 с.
51. *Коваленко И.Н.* Вероятностный расчет и оптимизация / Коваленко И. Н. – К.: Наук. думка, 1989. – 192 с.
52. *Колмогоров А.Н.* Элементы теории функцій и функционального анализа / А.Н. Колмогоров, С.В. Фомин. – М.: Наука, 1989. – 459 с.
53. *Косолап А.И.* Выпуклый анализ и многоэкстремальные задачи: [монография] / Косолап А.И.; Днепропетровский национальный университет. – Д.: Из-во ДНУ, 2007. – 278 с.
54. *Кутковецький В.Я.* Дослідження операцій: [навч. посібник] / Кутковецький В.Я. – К.: Вид-во ТОВ «Видавничий Дім «Професіонал»», 2004. – 350 с.
55. *Ларіонов Ю.І.* Дослідження операцій: навч. посібник [для студ. вищих навч. закл.] / Ю.І. Ларіонов, Л.С. Марченко, М.А. Хашмуридов. – Х.: «ІНЖЕК», 2004. ч.2. – 2005. – 287 с.
56. *Ларичев О.И.* Теория и методы принятия решений / Ларичев О.И. – М.: Логос, 2000. – 296 с.
57. *Лесин В.В.* Основы методов оптимизации: учеб. пособие [для студентов вузов] / В.В. Лесин, Ю.П. Лисовец. – М.: Изд-во МАИ, 1995. – 344 с.
58. *Любчак В.О.* Комп'ютерна реалізація методів оптимізації: навч. посібник [для студ. та асп. фіз.-мат., інж. та екон. спец.] / В.О. Любчак, Л.Г. Острівна; Сумський держ. ун-т. – Суми: Видавн. сумського держ. ун-ту, 2002. – 161 с.
59. Математичне програмування: [навч. посібник] / В.В. Бурий, І.В. Шевченко; Національний авіаційний ун-т. Модуль 1: Лінійне програмування. – К.: НАУ, 2007. – 168 с.
60. Математичне програмування: [навч. посібник] / В.В. Бурий, О.С. Давидов, І.В. Шевченко; Національний авіаційний ун-т. Модуль 2: Спеціальні методи математичного програмування. – К.: НАУ, 2008. – 124 с.
61. Математичні методи дослідження операцій: [навч. посібник] / В.П. Лавренчук та ін. (уклад.). – Чернівці: Рута, 2005. – 351 с.
62. *Макаров И.М.* Теория выбора и принятия решений. Учебное пособие / И.М. Макаров, Т.М. Виноградская. – М.: Наука, 1982. – 328 с.
63. *Марчук Г.И.* Методы вычислительной математики: учеб. пособие [для студ. вузов, обуч.по спец. «Прикладная математика»; –3-е изд., перер. и дополн.] / Марчук Г.И. – М: Наука, 1989. – 608 с.
64. Методи оптимізації: метод. вказ. до виконання практ. робіт для студ. техн. спец. / [уклад. А.П. Яковлева. ч.1. – К.: НТУУ «КПІ», 2007. – 56 с.
65. Методи оптимізації складних систем: навч. посібник [для студ. спец. «комп'ютеризовані системи управління і автоматика»] / І.В. Кузьмін,

М.М. Биков, С.М. Москвіна, А.І. Кузьмін; Вінницький держ. техн. ун-т. – Вінниця: ВДТУ, 2003. – 165 с.

66. *Мику М.* Математическое программирование: теория и алгоритмы / Мику М.; пер. с франц. и предисл. А.И. Штерна. – М.: Наука, 1990. – 486 с.
67. *Михалевич В.М.* Математичне програмування разом з Maple: навч. посібник [для студ. усіх спеціальностей] / Михалевич В.М.; Вінницький націон. техн. ун-т. ч.1. Методи розв'язування задач лінійного програмування. – Вінниця: ВНТУ, 2008. – 158 с.
68. *Михалевич В.С.* Методы невыпуклой оптимизации / В.С. Михалевич, А.М. Гупал, В.И. Норкин. – М.: Наука, 1987. – 279 с.
69. *Моисеев Н.Н.* Математические задачи системного анализа / Моисеев Н.Н. – М.: Наука, 1981. – 488 с.
70. *Моклячук М.П.* Варіаційне числення. Екстремальні задачі. Підручник / Моклячук М.П. – К.: «ТВ і МС», 2004. – 304 с.
71. *Моклячук М.П.* Негладкий аналіз та оптимізація: навч. посібник [для студ. вищих навч. закладів] / Моклячук М.П.; Київський націон. ун-т ім. Тараса Шевченка. – К.: ВПЦ «Київський ун-т», 2008. – 400 с.
72. *Морозов В.В., Сухарев А.Г., Федоров В.В.* Исследование операций в задачах и упражнениях: [учебное пособие] / В.В. Морозов, А.Г. Сухарев, В.В. Федоров. – М.: Высшая школа, 1986. – 287 с.
73. *Муртаф Б.* Современное линейное программирование / Муртаф Б. – М.: Мир, 1984. – 224 с.
74. *Наконечний О.Г.* Оптимальне керування та оцінювання в рівняннях із частинними похідними: [навч. посібник] / Наконечний О.Г. – К.: ВПЦ «Київський університет», 2004. – 103 с.
75. *Наконечний О.Г.* Оцінювання розв'язків алгебраїчно-диференціальних рівнянь в умовах невизначеності / О.Г. Наконечний, С.М. Жук. – Рівне: НУВГП, 2009. – 120 с.
76. *Нестеров Ю.Е.* Эффективные методы в нелинейном программировании / Нестеров Ю.Е. – М.: Радио и связь, 1989. – 301 с.
77. *Нефьодов Ю.М.* Методи оптимізації в прикладах і задачах: навч. посібник [для студ. вищих навч. закл.] / Ю.М. Нефьодов, Т.Ю. Балицька; Східноукраїнський націон. ун-т ім. Володимира Даля. – Луганськ: Ноулідж, 2009. – 324 с.
78. *Нурминский Е.А.* Численные методы выпуклой оптимизации / Нурминский Е.А.; ДВО. Ин-т прикладной математики. – М.: Наука, 1991. – 164 с.
79. *Перегудов Ф.И.* Введение в системный анализ: [учебное пособие] / Ф.И. Перегудов, Ф.П. Тарасенко. – М.: Высшая школа, 1989. – 367 с.
80. *Подиновский В.В.* Парето-оптимальные решения многокритериальных задач / В.В. Подиновский, В.Д. Ногин. – М.: Наука, 1982. – 254 с.
81. *Поляк Б.Т.* Введение в оптимизацию / Поляк Б. Т. – М.: Наука, 1983. – 384 с.
82. *Пономаренко О.І.* Системні методи а економіці, менеджменті та бізнесі:

[навч. посібник] / О.І. Пономаренко, В.О. Пономаренко. – К.: Либідь, 1995. – 240 с.

83. *Попов Ю.Д., Тюптя В.І., Шевченко В.І.* Методи оптимізації: [навч. посібник] / Ю.Д. Попов, В.І. Тюптя, В.І. Шевченко. – К.: Абрис, 1999. – 217 с.

84. *Пшеничный Б.Н.* Методы линеаризации / Пшеничный Б.Н. – М.: Наука, 1983. – 136 с.

85. *Ржевський С.В.* Дослідження операцій: [підручник] / С.В. Ржевський, В.М. Александрова. – К.: Академвидав, 2006. – 558 с.

86. *Ржевский С.В.* Монотонные методы выпуклого программирования / Ржевский С.В. – К.: Наук. думка, 1993. – 356 с.

87. *Сергієнко І.В*. Методи оптимізації та системного аналізу для задач трансобчислювальної складності / І.В. Сергієнко; НАН України, Ін-т кібернетики ім. В.М. Глушкова. – К.: Академперіодика, 2010. – 293 с.

88. *Струченков В.И.* Методы оптимизации: основы теории, задачи, обучающие компьютерные программы; [учеб. пособие для студентов вузов] / Струченков В.И. – М.: Экзамен, 2005. – 124 с.

89. *Сухарев А.Г.* Курс методов оптимизации: [учеб. пособие] / А.Г. Сухарев, А.В. Тимохов, В.В. Федоров. – [2-е изд.]. – М.: ФИЗМАТЛИТ, 2008. – 367 с. –(Серия «Классический университетский учебник») (Москов. гос. ун-т им. М.В. Ломоносова).

90. *Таха Хемди А*. Введение в исследование операций / Таха Хемди А; пер. с англ. и ред. А.А. Минько . –[7-е изд.]. – М.: Изд-во дом «Вильямс», 2005. – 901 с.

91. *Толбатов Ю.А.* Математичне програмування: підруч. [для студ. вищ. навч. закл.] / Ю.А. Толбатов, Є.Ю. Толбатов. – Т.: Підручники і посібники, 2008. – 432 с.

92. *Треногин В.А.* Функциональный анализ / Треногин В.А. – М.: Наука, 1980. – 495 с.

93. *Цегелик Г.Г.* Лінійне програмування / *Цегелик Г.Г.* – Львів: Світ, 1995. – 160 с.

94. *Жалдак М.І.* Основи теорії і методів оптимізації: навч. посібник [для студ. мат. спец. вищих навч. закл.] / М.І. Жалдак, Ю.В. Триус. – Черкаси: Брама–Україна, 2005. – 608 с.

95. *Шикин Е.В.* Исследование оперций: [учебник] / Е.В. Шикин, Г.Е. Шикина; Москов. гос. ун-т им. М.В. Ломоносова. – М.: Проспект, 2008. – 276 с.

96. *Щитов И.Н.* Введение в методы оптимизации: учеб. пособие [для студ. вузов, обуч. по направлению «физ.-мат. образование»] / Щитов И. Н. – М.: Высшая школа, 2008. – 206 с. – (Матем. для высших учебных заведений).

97. *Юдин Д.Б.* Вычислительные методы теории принятия решений / Юдин Д. Б. – М.: ФИЗМАТЛИТ, 1989. – 320 с.