Міністерство освіти і науки України
Національний університет водного господарства та природокористування
Навчально-науковий інститут автоматики, кібернетики та обчислювальної техніки
Кафедра комп'ютерних наук

04-05-23

МЕТОДИЧНІ ВКАЗІВКИ

до лабораторних занять та самостійного вивчення навчальної дисципліни

"Моделювання екологічних, економічних та соціальних процесів"

для здобувачів вищої освіти першого (бакалаврського) рівня за спеціальністю 126 «Інформаційні системи та технології» денної та заочної форм навчання

Рекомендовано
науково-методичною комісією
зі спеціальності
126 «Інформаційні системи та технології»
Протокол № 5 від 18.01.2019 р.

Рівне - 2019

Методичні вказівки до лабораторних занять та самостійного вивчення навчальної дисципліни ”Моделювання екологічних, економічних та соціальних процесів” для здобувачів вищої освіти першого (бакалаврського) рівня за спеціальністю 126 «Інформаційні системи та технології» денної та заочної форм навчання / Карпович І. М., Гладка О. М. – Рівне: НУВГП, 2019. – 39 с.

Укладачі: Карпович І. М., кандидат фізико-математичних наук, доцент; Гладка О. М., кандидат технічних наук, доцент.

Відповідальний за випуск – Тулашвілі Ю. Й., доктор педагогічних наук, професор, завідувач кафедри комп’ютерних наук.

ЗМІСТ

Вступ

Математичне моделювання дає змогу розкривати закономірності функціонування технічних, біологічних, соціально-економічних систем та управління ними, з'ясовувати логіку їх внутрішнього розвитку. Вміщені в методичних вказівках лабораторні роботи сприяють оволодінню студентами прийомами розробки і аналізу моделей процесів, які існують в природі, економіці і соціальній сфері, використовуючи основні положення теорії.

Вивчення дисципліни сприятиме розвитку у майбутніх спеціалістів у галузі інформаційних систем і технологій системного мислення, усвідомлення необхідності застосовувати кібернетичні принципи до задач управління і прийняття рішень та розв'язання інших проблем практики.

Підготовка до лабораторної роботи та її захисту передбачає опрацювання відповідного теоретичного матеріалу, який міститься в курсі лекцій, переліку рекомендованої літератури, включаючи електронні джерела інформації та актуальні періодичні видання з інформаційних технологій.

Виконання лабораторних робіт студентами передбачає використання сучасних комп'ютерних технологій. Вибір відповідного програмного засобу здійснюється студентом на основі власного досвіду. При цьому для дослідження моделі розробляється розрахунковий алгоритм і відповідна комп'ютерна програма або використовується наявне прикладне програмне забезпечення (спеціалізовані пакети, математичні пакети загального призначення, електронні таблиці тощо).

Під час співбесіди на захисті лабораторної роботи студент повинен продемонструвати знання мети роботи, теоретичного матеріалу, методів реалізації кожного етапу роботи, змісту основних розділів звіту з демонстрацією результатів на конкретних прикладах, практичних прийомів використання теоретичного матеріалу у виконаних розрахунках. Крім того, для успішного захисту роботи студент повинен відповісти на контрольні запитання, вміщені наприкінці опису відповідної роботи.

Лабораторна робота №1

Тема: Динамічна модель ринку із прогнозованими цінами

Основні відомості

Для моделювання динамічних процесів в економіці застосовують диференціальні рівняння другого порядку. Розглянемо модель ринку із прогнозованими цінами. У простих випадках моделювання ринку зазвичай вважають, що попит і пропозиція залежать тільки від поточної вартості товару. Однак, в реальній ситуації ці фактори залежать ще і від тенденції ціноутворення та від темпів зміни ціни.

В моделях із неперервними і диференційованими за часом функціями ці характеристики подаються відповідно першою та другою похідною функції ціни *P(t)*. Розглянемо це на прикладі.

Нехай функції попиту *D(t)* і пропозиції *S(t)* мають вигляд:

$$D(t) = 3P'' - P' - 2P + 18$$

$$S(t) = 4P'' + P' - 3P + 3$$

Вибрані залежності мають наступний практичний зміст:

1) попит стимулюється темпом зміни ціни: якщо темп зростання збільшується $(P'' > 0)$, ринок збільшує зацікавленість в товарі, і навпаки – швидкий ріст цін відлякує покупців. У зв'язку з цим доданок із P' міститься у функції попиту із знаком «-».

2) пропозиція зростає значною мірою із зростанням темпу зміни ціни. Тому доданок із P'' у функції *S(t)* має більший коефіцієнт, ніж у функції *D(t).* Ріст цін збільшує пропозицію, тому доданок, що містить P' у функції *S(t),* має знак «+».

Як відомо, рівноважний стан ринку характеризується рівністю, де попит дорівнює пропозиції, тобто:

$$S(t) = D(t)$$

$$3P'' - P' - 2P + 18 = 4P'' + P' - 3P + 3$$

$$P'' + 2P' + 5P = 15 \qquad (*)$$

Отримане рівняння – модель ринку із прогнозованими цінами. Проведемо дослідження і аналіз цієї моделі. З математичної точки зору модель ринку із прогнозованими цінами являє собою звичайне лінійне неоднорідне диференціальне рівняння другого порядку. Як відомо з курсу вищої математики, розв'язок такого рівняння – це сума загального розв'язку однорідного та частинного розв'язку неоднорідного рівняння. Однорідне рівняння має вигляд:

$$P'' + 2P' + 5P = 0 \qquad (**)$$

Відповідне характеристичне рівняння:

$$\lambda^2 + 2\lambda + 5 = 0$$

$$D = 4 - 4 \cdot 5 = -16$$

Маємо комплексно-спряжені характеристичні корені:

$$\lambda_{1,2} = \frac{-2 \pm 4i}{2} = -1 \pm 2i$$

$$\mathrm{Re}(\lambda) = \alpha = -1; \mathrm{Im}(\lambda) = \beta = 2$$

Зауваження. При розв'язуванні характеристичного рівняння можливі такі розв'язки:

1. Характеристичні числа λ_1, λ_2 є дійсними числами та не дорівнюють одне одному. Тоді загальний розв'язок однорідного диференціального рівняння має вигляд:

$$P(t) = C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t}$$

2. Характеристичні числа λ_1, λ_2 є дійсними числами та дорівнюють одне одному $\lambda_1 = \lambda. = \lambda$. Тоді загальний розв'язок однорідного диференціального рівняння має вигляд:

$$P(t) = C_1 e^{\lambda t} + C_2 t e^{\lambda t}$$

3. Характеристичні числа є комплексно-спряженими числами $\lambda_{1,2} = \alpha \pm i\beta$. Тоді загальний розв'язок однорідного диференціального рівняння має вигляд:

$$P(t) = e^{\alpha t}(C_1 \cdot \cos\beta t + C_2 \cdot \sin\beta t); C_1, C_2 = const.$$

Загальний розв'язок рівняння $(**)$ має вигляд:

$$P(t) = e^{-t}(C_1 \cdot \cos 2t + C_2 \cdot \sin 2t); C_1, C_2 = const.$$

Частинний розв'язок рівняння $(*)$ знайдемо враховуючи, що права частина є сталим числом. Отже, частинний розв'язок шукаємо у формі *P=A,* де *A - const*.

Для знаходження *A* підставимо у рівняння $(*)$ *5A=15, A=3.*

Остаточно додаємо знайдені загальний і частинний розв'язки:

$$P(t) = 3 + e^{-t}(C_1 \cdot \cos 2t + C_2 \cdot \sin 2t); C_1, C_2 = const.$$

Із отриманих результатів видно, що для моделі існує врівноважений стаціонарний стан, до якого прямує система з часом:

$$P_{cm} = \lim_{t \to \infty} P(t) = 3.$$

Розв'язок містить тригонометричні функції, тому графік гармонійний. В загальному випадку динаміка ринку із прогнозованими цінами має коливний характер і з часом наближається до стійкого стану. Сталі C_1, C_2 можна визначити із початкових умов, якщо відома ціна і тенденції її зміни.

Розглянемо задачі.

Задача 1. Нехай у початковий момент часу відома ціна і тенденції зміни ціни. Маємо наступну задачу Коші:

$$P|_{t=0} = 4;$$

$$P'|_{t=0} = 1.$$

На основі початкових даних знайдемо сталі C_1, C_2

$$4 = 3 + C_1;$$

$$C_1 = 1;$$

$$P'(t) = -e^{-t}(\cos 2t + C_2 \cdot sin2t) + e^{-t}(-2sin2t + 2C_2 \cdot \cos 2t);$$

$$1 = -1 + 2C_2;$$

$$C_2 = 1.$$

Аналітичний розв'язок задачі, коли відома ціна і тенденції її зміни, запишеться у вигляді:

$$P(t) = 3 + e^{-t} \cdot \left(\cos 2t + \sin 2t\right) = 3 + e^{-t} \cdot \sqrt{2} \cdot \left(\frac{\sqrt{2}}{2}\cos 2t + \frac{\sqrt{2}}{2}\sin 2t\right) =$$

$$= 3 + \sqrt{2} \cdot e^{-t} \cdot \left(\sin\frac{\pi}{4}\cos 2t + \cos\frac{\pi}{4}\sin 2t\right) = 3 + \sqrt{2} \cdot e^{-t} \cdot \cos(2t - \frac{\pi}{4}).$$

Розв'язування задачі засобами MathCAD може мати вигляд:

Given

P''(t) + 2 · P'(t) + 5 · P(t) = 15

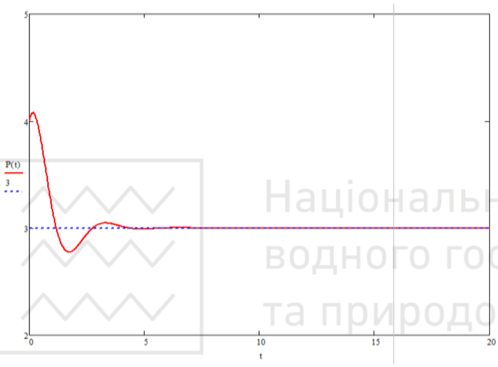
P(0) = 4

P'(0) = 1

P := Odesolve(t, 50)

Врівноважений стаціонарний стан:

P(50) = 3



На основі отриманого розв'язку можна зробити висновок про те, що існує врівноважений стаціонарний стан до якого з часом наближається функція ціни.

$$P_{cm} = \lim_{t\to\infty}\left(3+\sqrt{2}\cdot e^{-t}\cdot\cos(2t-\frac{\pi}{4})\right)=3$$

Задача 2. Дослідити динаміку зміни ціни в умовах із прогнозованими цінами. Встановити існування рівноважного стану, якщо функція попиту і пропозиції мають вигляд:

$$D(t)=5P''-P'-5P+19$$

$$S(t)=15P''+P'+10P+2$$

$$p(0)=3$$

$$p'(0)=1$$

Розв'язання.

$$5P''-P'-5P+19=15P''+P'+10P+2$$
$$-10P''-2P'-15P+17=0$$
$$10P''+2P'+15P=17$$

Розв’язок задачі поданий у вигляді графіка нижче.

Given

$10\cdot P''(t)+2\cdot P'(t)+15\cdot P(t)=17$

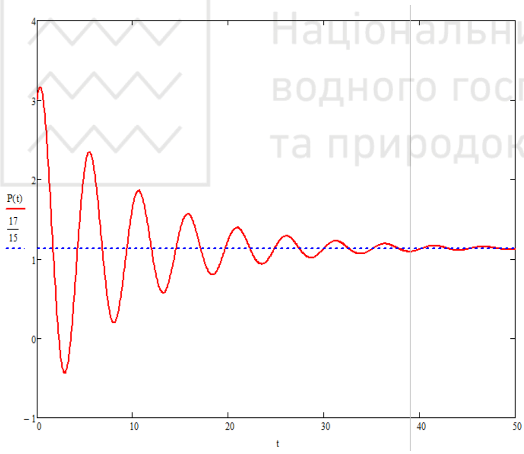
$P(0)=3$

$P'(0)=1$

$P:=\mathrm{Odesolve}(t,70)$

Врівноважений стаціонарний стан:

$P(50)=1.124$



Висновок: В умовах ринку із прогнозованими цінами після затухаючого коливного процесу функція ціни наближається до врівноваженого стаціонарного стану.

Завдання. Дослідити динаміку зміни ціни в умовах із прогнозованими цінами. Встановити існування рівноважного стану, якщо задані функції попиту і пропонування. В початковий момент часу відомі: 1) початкова ціна та тенденція її зміни; 2) початкова ціна та початковий попит. Зробити висновки. Дані для виконання роботи наведені в таблиці 1.

Таблиця 1

№	Функція попиту	Функція пропозиції	Початкові умови
1	$D(t)=3\cdot P''(t)-P'(t)-5\cdot P(t)+25$	$S(t)=8\cdot P''(t)+P'(t)+5\cdot P(t)+5$	$P(0)=5$ $P'(0)=3$
2	$D(t)=5\cdot P''(t)-P'(t)-2\cdot P(t)+15$	$S(t)=15\cdot P''(t)+P'(t)+3\cdot P(t)+5$	$P(0)=7$ $P'(0)=5$
3	$D(t)=7\cdot P''(t)-2\cdot P'(t)-3\cdot P(t)+17$	$S(t)=14\cdot P''(t)+P'(t)+4\cdot P(t)+7$	$P(0)=5$ $P'(0)=4$
4	$D(t)=P''(t)-3\cdot P'(t)-3\cdot P(t)+12$	$S(t)=10\cdot P''(t)+P'(t)+7\cdot P(t)+2$	$P(0)=5$ $P'(0)=2$
5	$D(t)=5\cdot P''(t)-P'(t)-8\cdot P(t)+33$	$S(t)=8\cdot P''(t)+P'(t)+12\cdot P(t)+18$	$P(0)=5$ $P'(0)=5$
6	$D(t)=5\cdot P''(t)-2\cdot P'(t)-3\cdot P(t)+22$	$S(t)=35\cdot P''(t)+3\cdot P'(t)+17P(t)+2$	$P(0)=7$ $P'(0)=5$
7	$D(t)=6\cdot P''(t)-3\cdot P'(t)-12\cdot P(t)+21$	$S(t)=31\cdot P''(t)+4\cdot P'(t)+6\cdot P(t)+2$	$P(0)=9$ $P'(0)=7$
8	$D(t)=5\cdot P''(t)-P'(t)-3\cdot P(t)+20$	$S(t)=23\cdot P''(t)+4\cdot P'(t)+5\cdot P(t)+5$	$P(0)=6$ $P'(0)=2$
9	$D(t)=6\cdot P''(t)-P'(t)-5\cdot P(t)+17$	$S(t)=21\cdot P''(t)+2\cdot P'(t)+5\cdot P(t)+10$	$P(0)=7$ $P'(0)=1$
10	$D(t)=7\cdot P''(t)-P'(t)-3\cdot P(t)+12$	$S(t)=18\cdot P''(t)+P'(t)+7\cdot P(t)+7$	$P(0)=6$ $P'(0)=2$

Контрольні запитання

1. Від чого залежить попит і пропозиція на ринку?
2. Умова рівноважного стану ринку.
3. Приклади моделі ринку із прогнозованими цінами.
4. Пошук рівноважного стану ринку засобами MathCAD.

Лабораторна робота № 2

Тема: Аналіз моделей розвитку популяцій

Основні відомості

Приклад 1. Припустимо, що популяція алігаторів нараховує 100 особин, показник смертності популяції δ=0. Нехай коефіцієнт народжуваності β=0,0005*P*. Таким чином, популяція зростає. Що станеться з популяцією в майбутньому?

Розв'язання. Враховуючи загальне рівняння чисельності популяції

$$\frac{dP}{dt} = (\beta - \delta)P,$$

отримуємо таке рівняння для чисельності крокодилів:

$$\frac{dP}{dt} = 0{,}0005P^2,\ P(0) = 100,$$

(t вимірюється в роках). Тоді після розділення змінних отримуємо

$$\int \frac{1}{P^2}\,dP = \int 0{,}0005\,dt;$$

$$-\frac{1}{P} = 0{,}0005t + C.$$

Підстановка $t = 0$, $P = 100$ дає $C = -1/100$. Отже,

$$P(t) = \frac{2000}{20 - t}.$$

Нехай, наприклад, $P(10) = 200$, так що через 10 років чисельність крокодилів подвоюється. При $t \to 20$ значення $P(t) \to +\infty$. Це означає, що через 20 років маємо реальний демографічний вибух.

Приклад 2. Зміна кількості бактерій у популяції описується диференціальним рівнянням

$$\frac{dB}{dt} = aB - cB^2$$

де B - число бактерій, a і c - коефіцієнти. У початковий момент часу кількість бактерій $B(0) = 20$. Побудувати графік залежності кількості бактерій від часу t в діапазоні від 0 до 20 при a =1,2 і c = 0,0002. Розглянемо розв'язування задачі засобами MathCAD.

1) Задаємо значення для коефіцієнтів, інтервалу зміни незалежної змінної і початкової умови. Значення коефіцієнтів:

a:= 1.2 c:= 0.0002

Початкове і кінцеве значення інтервалу зміни параметра *t*:

t_{begin}:= 0 t_{end}: = 20

Початкове значення *B*: B_0:= 20

2) Формуємо блок Given, в якому записуємо диференціальне рівняння і початкову умову. Пам'ятаємо, що B – це функція; в блоці Given слід використовувати знак логічної рівності.

Given

$$\frac{d}{dt}B(t) = a \cdot B(t) - c \cdot B(t)^2$$

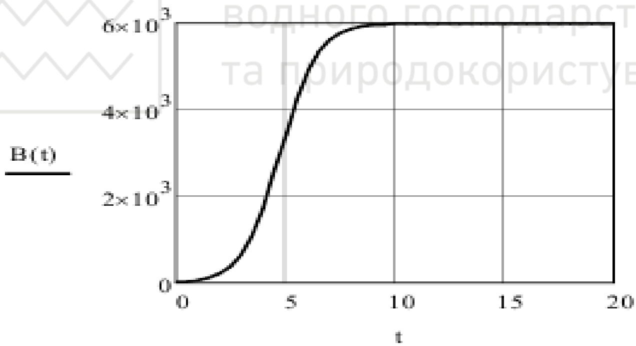
$$B(t_{begin}) = B_0$$

3) Отримаємо розв'язок з допомогою функції Odesolve:

B:= Odesolve(t , t_{end})

4) Будуємо графік розв'язку. Нагадаємо, що для виведення результатів потрібно задати діапазон зміни аргументу.

t: = 0, 0.5 .. 20



Це ж диференціальне рівняння можна розв'язати методом Рунге-Кутта з фіксованим кроком, який у MathCAD реалізується функцією

rkfixed (y, x1, x2, npoints, D)

де y – початкове значення, x1 – ліва межа діапазону зміни незалежної змінної, x2 – права межа діапазону зміни незалежної змінної, npoints – кількість розрахункових точок, D (x, Y) – функція, що визначає праву частину диференціального рівняння, x – ім'я незалежної змінної, Y – ім'я шуканої функції.

Результат обчислень повертається у вигляді матриці, перший стовпець якої містить значення незалежної змінної в розрахункових точках, а другий – значення функції в тих же точках.

1) Для задання початкового значення і діапазону зміни незалежної змінної скористаємося даними із завдання 1.

t_{begin}:=0 t_{end}:=20 B_0:=20

Кількість розрахункових точок рекомендується здавати рівною 1000, тобто

m:=1000

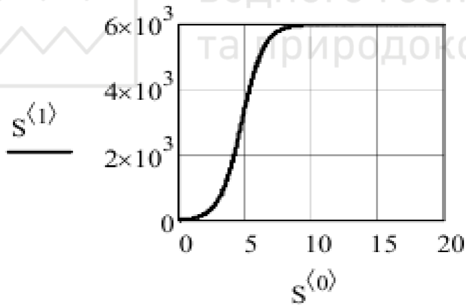
2) Задаємо функцію D, яка в нашому випадку матиме вигляд

$D(t,B):=a \cdot B - c \cdot B^2$

3) Знаходимо розв'язок:

S: = rkfixed (B0, t_{begin}, t_{end}, m, D)

Будуємо графік, враховуючи, що перший стовпець матриці S містить значення незалежної змінної t, а другий - функції B. (Нагадаємо, що за замовчуванням нумерація стовпців у матриці починається з нуля).



Приклад 3. Розглянемо розв'язок системи диференціальних рівнянь Вольтера-Лотки, яка описує динаміку кількості хижаків та жертв у замкнутому ареалі і є одною з базових моделей в екології

$$\frac{dN_1}{dt} = N_1(\varepsilon_1 - \gamma_2 N_2)$$

$$\frac{dN_2}{dt} = N_2(\varepsilon_2 - \gamma_1 N_1)$$

У прикладі значення origin = 1, тобто нумерація елементів масиву починається від 1, а не з 0, як це прийнято за замовчуванням. У початковий момент число хижаків N_1 = 7 і жертв – N_2 = 9. Фрагмент виконання завдання.

$$\text{origin} := 1$$

$$N := \begin{pmatrix} 7 \\ 9 \end{pmatrix} \qquad \varepsilon := \begin{pmatrix} 0.11 \\ 0.28 \end{pmatrix} \qquad \gamma := \begin{pmatrix} 0.03 \\ 0.05 \end{pmatrix} \qquad t_{max} := 200$$

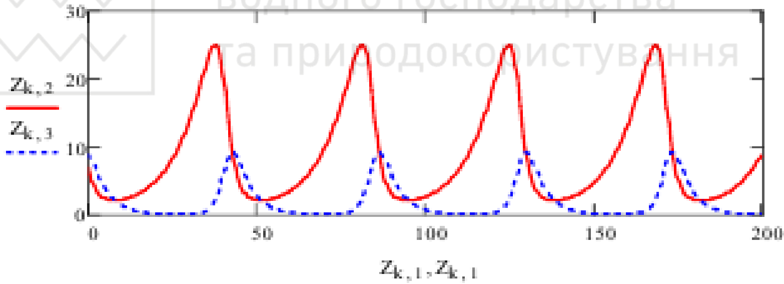
$$\text{npoints} := 400$$

$$D(t, N) := \begin{bmatrix} N_1 \cdot (\varepsilon_1 - \gamma_2 \cdot N_2) \\ -N_2 \cdot (\varepsilon_2 - \gamma_1 \cdot N_1) \end{bmatrix}$$

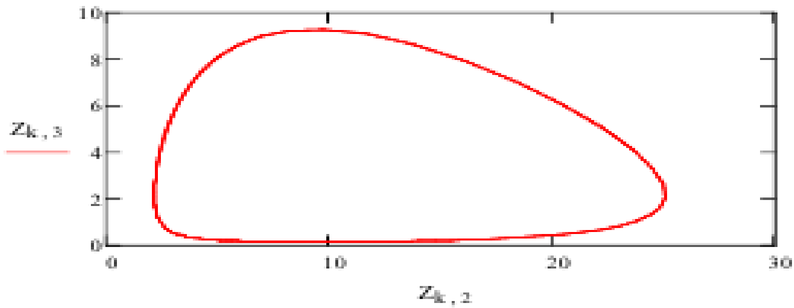
$$Z := \text{rkfixed}(N, 0, t_{max}, \text{npoints}, D)$$

Результати розрахунків подамо графічно у вигляді залежності чисельності популяцій хижаків та жертв від часу а) та залежності числа жертв від числа хижаків б).

а) k := 1.. npoints



б) k := 1.. npoints



Завдання.

1. Швидкість зміни чисельності популяції кролів *N(t)* пропорційна квадратному кореню з *N*. В момент часу $t = 0$ (місяці) популяція кролів має 100 особин і збільшується із швидкістю 20 кролів на місяць. Скласти математичну модель розвитку популяції і визначити чисельність кролів через рік.

2. Два зоопарки вирішили вирощувати хом'яків на продаж. Кількість хом'яків *P(t)* (*t* виражається в місяцях) задовольняє диференціальне рівняння

$$\frac{dP}{dt} = 0{,}001P(t)[125 - P(t)]$$

Перший зоопарк закупив 175 хом'яків, другий – 100. Розв'яжіть диференціальне рівняння та визначте, що станеться з популяцією хом'яків через 20 місяців. Який зоопарк вчинив раціональніше з точки зору витрат на придбання першої партії хом'яків? Побудуйте графіки зміни чисельності популяцій для двох зоопарків.

3. Маємо такі системи диференціальних рівнянь, що описують взаємодію двох популяцій типу «хижак–жертва»:

$$\begin{cases} \dfrac{dx}{dt} = (0{,}2)x - (0{,}005)xy, \\ \dfrac{dy}{dt} = -(0{,}5)y + (0.01)xy. \end{cases}$$

$$\begin{cases} \dfrac{dx}{dt} = (0{,}2)x - (0{,}005)xy - (0{,}0035)x^2, \\ \dfrac{dy}{dt} = -(0{,}5)y + (0.01)xy. \end{cases}$$

$$\begin{cases} \dfrac{dx}{dt} = (0{,}2)x - (0{,}005)xy - (0{,}0035)x^2, \\ \dfrac{dy}{dt} = -(0{,}5)y + (0.01)xy - (0{,}002)y^2. \end{cases}$$

Побудувати фазовий портрет в додатних осях *(x, y)* та зобразити періодичні коливання чисельності популяції хижака та жертви в осях *(t,x)* та *(t,y)* на одному графіку. Рекомендовані параметри*:* $x(0)=70$; $y(0)=40$; $t=0..100$.

4. Знайти розв'язок системи диференціальних рівнянь при *t,* що змінюється в діапазоні від 1 до $4N/(N+1)$, де N – номер варіанту і побудувати графіки функцій.

$$\frac{dx}{dt} = x - y + xy,$$

$$\frac{dy}{dt} = x + y - xy,$$

$$x(1) = 1, \qquad y(1) = 20$$

5. Розв'язати систему диференціальних рівнянь і побудувати графіки функцій $y_1(x)$ і $y_2(x)$ на відрізку [a, b]:

$$\begin{cases} y_1' = f_1(x, y_1, y_2), \\ y_2' = f_2(x, y_1, y_2), \quad y_1(a) = A_1, \; y_2(a) = A_2. \end{cases}$$

$$f_1(x, y_1, y_2) = \sin y_1 y_2,$$
$$f_2(x, y_1, y_2) = \cos(x y_1 y_2),$$
$$A_1 = -1, \; A_2 = 2, \; a = 0, \; b = 2.$$

Контрольні запитання

1. Яка різниця між задачею Коші та крайовою задачею?
2. Назвіть основні функції для розв'язання звичайних диференціальних рівнянь.
3. Призначення та особливість використання блоку розв'язку given…odesolve.
4. Опишіть структуру та аргументи блоку розв'язку given…odesolve.
5. Аргументи та використання функції *rkfixed(…).*

Лабораторна робота № 3

Тема: Динамічна модель Леонтьєва з дискретним часом

Основні відомості

Макроекономіка функціонування багатогалузевого господарства вимагає балансу між окремими галузями. Кожна галузь з одного боку є споживачем продукції іншої галузі, а з іншого - виробником продукції. Виникає задача розрахунку зв'язку між галузями через випуск та споживання продукції різного виду.

Розглядається економіка, що має в своєму складі *n* галузей, які виробляють і споживають *n* типів продуктів. Кожна галузь займається виробництвом тільки одного виду продукту і відповідно вважається, що різні галузі є виробниками різних продуктів.

В дискретній моделі час змінюється із періодом 1 рік: $t = 1,2,3,\ldots T$, $\Delta t=1$ *рік*. Модель зображена у матричній формі. Нижня індексація визначає номер року. В моделі, що вивчається, використовуються наступні змінні, які характеризують стан економіки в динаміці (всього $3n+1$ змінна).

$$X_t = \begin{pmatrix} x_t^1 \\ x_t^2 \\ \ldots \\ x_t^n \end{pmatrix}$$ - вектор-стовпчик валових випусків галузей,

$$\overline{X}_t = \begin{pmatrix} \bar{x}_t^1 \\ \bar{x}_t^2 \\ \ldots \\ \bar{x}_t^n \end{pmatrix}$$ - вектор-стовпчик галузевих потужностей,

$$V_t = \begin{pmatrix} v_t^1 \\ v_t^2 \\ \ldots \\ v_t^n \end{pmatrix}$$ - вектор-стовпчик введення потужностей.

$L_t = (L_t^1, L_t^2, \ldots, L_t^n)$ - вектор-рядок трудових ресурсів.

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$ - матриця прямих витрат.

Коефіцієнти a_{ij} не залежать від часу і масивів виробництва ($i=\overline{1,n}$, $j=\overline{1,n}$)

Крім того, задані такі матриці з постійними коефіцієнтами:

$$B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \dots & \dots & \dots & \dots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{pmatrix}$$ - матриця, що характеризує фондоємність.

$$C = \begin{pmatrix} c^1 \\ c^2 \\ \dots \\ c^n \end{pmatrix}$$ - вектор-стовпчик споживання в розрахунку на одного працюючого.

$l = (l^1, l^2, ..., l^n)$ - вектор-рядок трудоємності.

Враховуючи вказані позначення, економічна модель міжгалузевого балансу (МГБ) з дискретним часом приймає наступний вигляд:

$$\begin{cases} X_t \geq A \cdot X_t + B \cdot V_t + L_t \cdot C & (4) \\ X_t \leq \overline{X}_{t-1} & (5) \\ \overline{X_t} \leq \overline{X}_{t-1} + V_t & (6) \\ l \cdot X_t \leq L_t & (7) \\ X_t \geq 0, \overline{X}_t \geq 0, V_t \geq 0, L_t \geq 0 \\ t = 0, 1, 2, \dots, T \end{cases}$$

Цю систему співвідношень називають моделлю Леонтьєва в матричній формі.

У запропонованій моделі (4) - (7) лаг капіталовкладень складає 1 рік, що означає: інвестиції здійснені в рік t починають працювати в рік $t+1$ (сукупність нерівностей (5) - (6)).

Послідовність векторів $X_t, \overline{X}_t, t = 1, 2, ..., T$ називають допустимою траєкторією, якщо в кожен рік t виконуються всі умови моделі, при умові, що в базовому році ($t=0$) потужності є заданими і дорівнюють $\overline{X_0}$.

***Приклад* 1.** На основі динамічної моделі Леонтьєва з дискретним часом дослідити зміни вектора валового випуску впродовж двох років, якщо матриця прямих витрат A, матриця скінченного споживання $y(t) = f_0$, вектор валового випуску в початковий момент часу x_0 мають наступні значення:

$$A = \begin{pmatrix} 0{,}05 & 0{,}35 & 0{,}4 \\ 0{,}1 & 0{,}1 & 0{,}4 \\ 0{,}2 & 0{,}1 & 0{,}2 \end{pmatrix} \qquad f_0 = \begin{pmatrix} 40 \\ 60 \\ 10 \end{pmatrix} \qquad x_0 = \begin{pmatrix} 100 \\ 100 \\ 50 \end{pmatrix}$$

Крім того, відомо, що всі компоненти вектора скінченного споживання зростають на 30% кожного року.

Розв'язання. В розрахунках будемо користуватись наступною базовою моделлю: $x_{t+1} = Ax_t + f_t$

Визначимо на основі моделі Леонтьєва всі формули, що дозволяють обчислити значення вектора у будь-який момент часу на основі тільки початкових даних і значень матриці прямих витрат A.

$$t = 0$$

$$x_1 = Ax_0 + f_0$$

$$t = 1$$

$$x_2 = Ax_1 + f_1 = A(Ax_0 + f_0) + f_1 = A^2 x_0 + Af_0 + f_1$$

$$t = 2$$

$$x_3 = Ax_2 + f_2 = A(A^2 x_0 + Af_0 + f_1) + f_2 = A^3 x_0 + A^2 f_0 + Af_1 + f_2$$

$$t = n$$

$$x_n = A^n x_0 + \sum_{k=0}^{n-1} A^{n-k-1} \cdot f_k$$

$$f_1 = 1{,}3 f_0$$

$$f_2 = 1{,}3 f_1 = 1{,}3 \cdot 1{,}3 f_0 = (1{,}3)^2 f_0$$

$$f_k = (1{,}3)^k f_0$$

$$x_n = A^n x_0 + \sum_{k=0}^{n-1} A^{n-k-1} \cdot (1{,}3)^k f_0$$

Маємо універсальну формулу для обчислення валового випуску на основі початкових даних.

В нашому випадку обчислення проведемо безпосередньо за моделлю Леонтьєва.

$$x_1 = Ax_0 + f_0$$

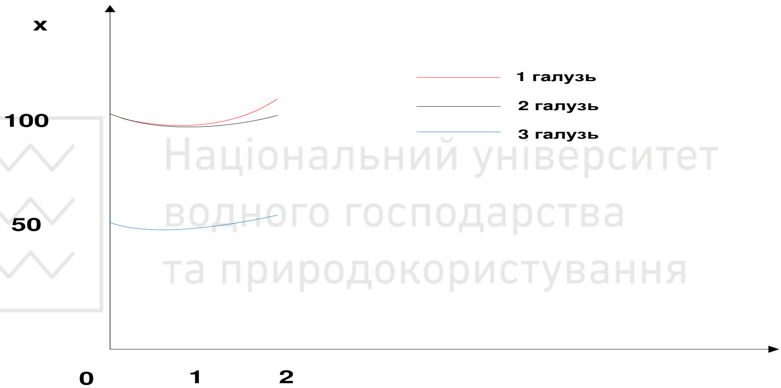
$$Ax_0 = \begin{pmatrix} 0.05 & 0.35 & 0.4 \\ 0.1 & 0.1 & 0.4 \\ 0.2 & 0.1 & 0.2 \end{pmatrix} \times \begin{pmatrix} 100 \\ 100 \\ 50 \end{pmatrix} = \begin{pmatrix} 5+35+20 \\ 10+10+20 \\ 20+10+10 \end{pmatrix} = \begin{pmatrix} 60 \\ 40 \\ 40 \end{pmatrix}$$

$$x_1 = \begin{pmatrix} 60 \\ 40 \\ 40 \end{pmatrix} + \begin{pmatrix} 40 \\ 60 \\ 10 \end{pmatrix} = \begin{pmatrix} 100 \\ 100 \\ 50 \end{pmatrix}$$

$$x_2 = Ax_1 + f_1$$

$$Ax_1 = \begin{pmatrix} 60 \\ 40 \\ 40 \end{pmatrix} \qquad f_1 = 1.3 f_0 = \begin{pmatrix} 52 \\ 78 \\ 13 \end{pmatrix}$$

$$x_2 = \begin{pmatrix} 60 \\ 40 \\ 40 \end{pmatrix} + \begin{pmatrix} 52 \\ 78 \\ 13 \end{pmatrix} = \begin{pmatrix} 112 \\ 118 \\ 53 \end{pmatrix}$$



Завдання.

1. На основі динамічної моделі Леонтьєва з дискретним часом дослідити зміни вектора валового випуску впродовж трьох років, якщо матриця прямих витрат *A,* матриця скінченного споживання f_t , вектор валового випуску в початковий момент часу x_0 мають наступні значення:

$$A = \begin{pmatrix} 0.05 & 0.1 & 0.85 \\ 0.8 & 0.7 & 0.65 \\ 0.4 & 0.7 & 0.8 \end{pmatrix} \qquad f_0 = \begin{pmatrix} 100 \\ 150 \\ 200 \end{pmatrix} \qquad x_0 = \begin{pmatrix} 500 \\ 700 \\ 800 \end{pmatrix}$$

Крім того, відомо, що всі компоненти вектора скінченного споживання зростають на 25% кожного року, $t=1,2,3$.

2. На основі динамічної моделі міжгалузевого балансу Леонтьева розрахувати вектор валового випуску *x(t)* на момент часу *T=5*, якщо всі компоненти вектора скінченого споживання *y(t)* збільшуються на *(10+N)*% за кожний період. В початковий момент часу *t=0* вектор валового випуску *x(0)*, вектор скінченого споживання *y(0)* та матриця виробничих витрат мають наступний вигляд:

$$\overline{x_0} = \begin{pmatrix} 100+2\cdot N \\ 200-2\cdot N \\ 150+3\cdot N \\ 75+4\cdot N \\ 90+5\cdot N \\ 120-2\cdot N \\ 180+2\cdot N \\ 50+5\cdot N \\ 100+3\cdot N \\ 220+4\cdot N \end{pmatrix} \qquad \overline{y_0} = \begin{pmatrix} 40+N \\ 45+N \\ 50-N \\ 60-N \\ 100-3\cdot N \\ 20+N \\ 30+N \\ 35+N \\ 50-N \\ 70-2\cdot N \end{pmatrix}$$

$$A = \begin{pmatrix} 0.05 & 0.35 & 0.45 & 0.51 & 0.52 & 0.34 & 0.38 & 0.8 & 0.35 & 0.7 \\ 0.1 & 0.9 & 0.1 & 0.85 & 0.35 & 0.2 & 0.35 & 0.2 & 0.24 & 0.4 \\ 0.2 & 0.85 & 0.15 & 0.7 & 0.45 & 0.2 & 0.36 & 0.1 & 0.25 & 0.7 \\ 0.3 & 0.8 & 0.2 & 0.6 & 0.48 & 0.1 & 0.37 & 0.2 & 0.27 & 0.54 \\ 0.4 & 0.75 & 0.25 & 0.5 & 0.87 & 0.1 & 0.38 & 0.3 & 0.28 & 0.37 \\ 0.45 & 0.7 & 0.3 & 0.2 & 0.25 & 0.15 & 0.4 & 0.4 & 0.3 & 0.35 \\ 0.46 & 0.65 & 0.27 & 0.2 & 0.35 & 0.12 & 0.45 & 0.6 & 0.34 & 0.6 \\ 0.47 & 0.6 & 0.31 & 0.1 & 0.38 & 0.3 & 0.47 & 0.7 & 0.38 & 0.58 \\ 0.48 & 0.7 & 0.35 & 0.4 & 0.5 & 0.5 & 0.48 & 0.5 & 0.8 & 0.7 \\ 0.43 & 0.1 & 0.8 & 0.6 & 0.51 & 0.6 & 0.5 & 0.22 & 0.4 & 0.33 \end{pmatrix}$$

Дослідити динаміку зростання компонентів вектора валового випуску протягом всього періоду при заданих темпах зростання продукту скінченного споживання та зробити відповідні висновки (N – номер варіанту).

Контрольні запитання

1. Суть балансового методу дослідження економічних систем.
2. Принципова схема міжгалузевого балансу. Основні розділи схеми і їх економічна суть.
3. Економіко-математична модель міжгалузевого балансу.
4. Зміст коефіцієнтів прямих і повних матеріальних витрат.
5. Економічний сенс коефіцієнтів прямої та повної фондомісткості.
6. Навести схему та послідовність обчислення коефіцієнтів трудомісткості та фондомісткості на підставі економіко-математичної моделі.
7. Основні сфери використання в економіці моделей міжгалузевого балансу. Приклади.

Лабораторна робота № 4

Тема: Побудова однофакторних регресійних моделей

Основні відомості

Приклад 1. Визначити залежність між видобутком вугілля Y (т) на одного робітника і потужністю пласта вугілля X (м). Вхідні дані наведено у таблиці. Знайти рівняння регресії Y на X.

i	1	2	3	4	5	6	7	8	9	10
x_i	8	11	12	9	8	8	9	9	8	12
y_i	5	10	10	7	5	6	6	5	6	8

Розв'язання. Обчислимо необхідні суми:

$$\sum_{i=1}^{10} x_i = 8+11+12+9+8+8+9+9+8+12 = 94;$$

$$\sum_{i=1}^{10} x_i^2 = 8^2+11^2+12^2+9^2+8^2+8^2+9^2+9^2+8^2+12^2 = 908;$$

$$\sum_{i=1}^{10} y_i = 5+10+10+7+5+6+6+5+6+8 = 68;$$

$$\sum_{i=1}^{10} x_i y_i = 8\cdot 5+11\cdot 10+12\cdot 10+9\cdot 7+8\cdot 5+8\cdot 6+9\cdot 6+9\cdot 5+8\cdot 6+12\cdot 8 = 664;$$

Обчислимо вибіркові характеристики та параметри рівняння регресії:

$$\bar{x} = 94/10 = 9{,}4\ м;\ \bar{y} = 68/10 = 6{,}8\ т,$$

$$s_x^2 = 908/10 - 9{,}4^2 = 2{,}44,$$

$$K(X,Y) = 664/10 - 9{,}4 \cdot 6{,}8 = 2{,}48;\ b_1 = 2{,}48/2{,}44 = 1{,}016.$$

Отримаємо рівняння регресії

$\hat{y} - 6{,}8 = 1{,}016(x - 9{,}4)$ або $\hat{y} = 2{,}75 + 1{,}016x$.

З одержаного рівняння випливає, що при збільшенні потужності пласта на 1 м видобуток вугілля на одного робітника збільшується в середньому на 1,016 т.

Приклад 2. Для побудови лінійної моделі створимо масив даних, що містить 11 точок. Значення змінної x сформуємо у вигляді арифметичної прогресії з першим членом –5 та різницею 1. Значення змінної y сформуємо за формулою: $y = x^2 + 1 + \varepsilon$, де ε – рівномірно розподілена випадкова величина, задана на відрізку [–2; 2].

Нижче показано робочі вікна з двома варіантами програми побудови моделі та результат побудови моделі (однаковий для обох випадків).

Варіант 1.

$x := (-5\ \ -4\ \ -3\ \ -2\ \ -1\ \ 0\ \ 1\ \ 2\ \ 3\ \ 4\ \ 5)^T$

$y := (-9.5\ \ -8.6\ \ -4.6\ \ -3.4\ \ 0.5\ \ 2.8\ \ 1.1\ \ 4.6\ \ 8.8\ \ 7.4\ \ 9.9)^T$

$\mathrm{line}(x, y) = \begin{pmatrix} 0.818 \\ 1.98 \end{pmatrix}$

$f(t) := \mathrm{line}(x, y)_0 + \mathrm{line}(x, y)_1 \cdot x$

Варіант 2.

$x := (-5\ \ -4\ \ -3\ \ -2\ \ -1\ \ 0\ \ 1\ \ 2\ \ 3\ \ 4\ \ 5)^T$

$y := (-9.5\ \ -8.6\ \ -4.6\ \ -3.4\ \ 0.5\ \ 2.8\ \ 1.1\ \ 4.6\ \ 8.8\ \ 7.4\ \ 9.9)^T$

$\mathrm{intercept}(x, y) = 0.818$

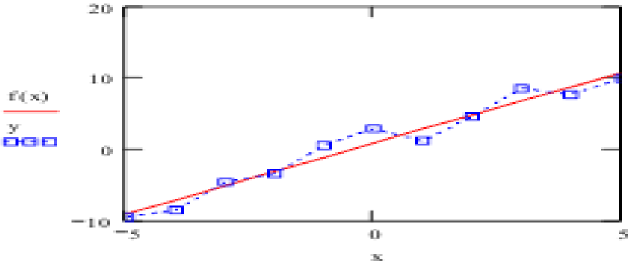
$\mathrm{slope}(x, y) = 1.98$

$g(x) := \mathrm{intercept}(x, y) + \mathrm{slope}(x, y)\,x$

Функція line(x, y) видає вектор коефіцієнтів лінійної моделі

$f(x) = ax+b$ у вигляді вектора $\begin{pmatrix} b \\ a \end{pmatrix}$.

Значенням функції intercept(x, y) є ордината точки перетину моделі з віссю ординат, а значенням функції slope(x, y) – тангенс кута нахилу моделі до осі абсцис.



Приклад 3.

Для побудови поліноміальної моделі згенеруємо вектор x1, який містить 21 елемент, що утворюють арифметичну прогресію з початковим значенням –2,25 і різницею 0,25. Елементи вектора y1 розрахуємо за формулою:

$$y1=-2x1^4+3x1^3+1{,}5x1^2-7x1+2+\varepsilon,$$

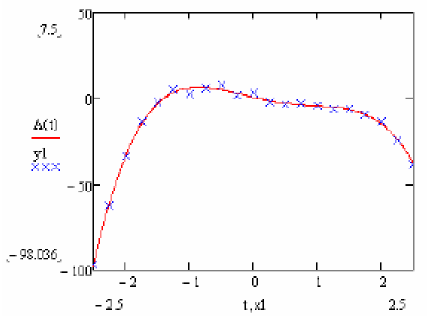
де ε – елементи рівномірної випадкової послідовності, заданої на відізку [–3; 3]. Нижче наведено фрагмент програми, що використовується для побудови **поліноміальної моделі**.

x1:=(-2.5 -2.25 -2 -1.75 -1.5 -1.25 -1 -0.75 -0.5 -0.25 0 0.25 0.5 0.75 1 1.25 1.5 1.75 2 2.25 2.5)^T

y1:=(-96.8 -62.5 -33.8 -13.6 -23 5.1 2.6 5.6 7.5 1.6 2.8 -2.3 -3.7 -3.1 -4.2 -6.3 -6.4 -9.3 -13.7 -14.1 -38.1) ^{T}

k:=4

<u>s:= regress(x1,y1,4)</u>



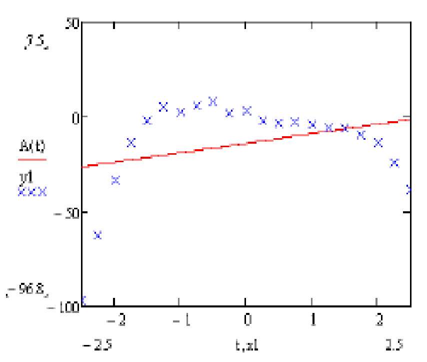
$$s = \begin{pmatrix} 3 \\ 3 \\ 4 \\ 0.644 \\ -8.548 \\ 2.366 \\ 3.294 \\ -2.134 \end{pmatrix}$$

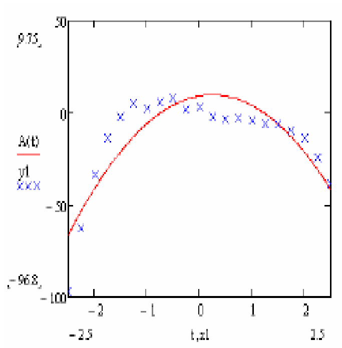
A(t):=interp(s,x1,y1,t)

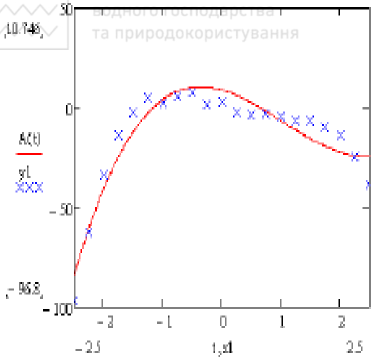
Останні п’ять елементів вектора *s* є оцінками коефіцієнтів *a0, a1, a2, a3, a4* полінома $A(x) = a0 + a1x + a2x^2 + a3x^3 + a4x^4$. На рис. нижче показано графік отриманої моделі у вигляді полінома четвертого степеня, який добре узгоджується із вхідними даними.

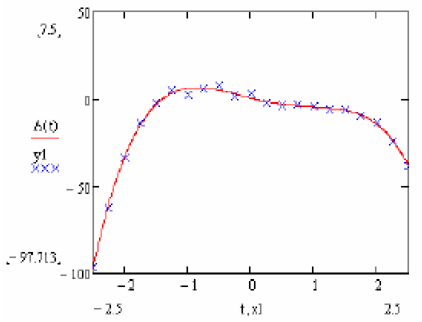
Національний університет водного господарства та природокористування

Завдання 1. Побудувати графіки моделей у вигляді поліномів **першого**, **другого, третього та п’ятого** степенів для тих же вхідних даних.









Порівняння результатів дає підстави зробити висновок, що для наведених даних вже **модель третього порядку** добре відображає основні особливості вхідних даних, а модель п'ятого порядку є надлишковою, оскільки немає жодних переваг перед моделлю четвертого порядку, але є складнішою.

Приклад 3. Альтернативний варіант передбачає побудову моделі у вигляді відрізків поліномів. Фрагмент програми для побудови такої моделі наведено нижче. При цьому використовували ті самі дані, що і для побудови звичайної поліноміальної моделі.

s1:= loess(x1,y1,0.75)

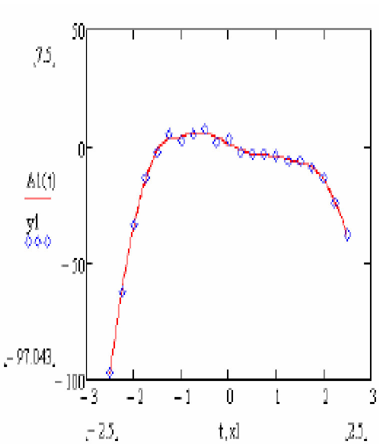
A1(t):=interp(s1,x1,y1,t)

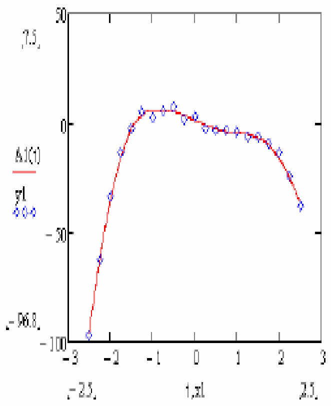
Нижче розміщено результати побудови моделі для різних значень фактора span, що задає довжину відрізків поліномів (у наведеному прикладі span=0.75). Аналіз наведених даних показує, що при малих значеннях фактора span модель краще відображає наявні дані. Але занадто добра відповідність моделі і вхідних даних може бути зайвою, оскільки дані містять певну похибку. Крім того, для малих значень фактора span істотно збільшується обсяг розрахунків і при span $\leq$ 0,28 з'являється повідомлення про нестачу пам'яті для завершення операції. Разом з тим, із збільшенням параметра span понад 0,75..0,8 якість моделі погіршується, а при span = 1 відхилення моделі від даних вже є неприпустимо великими.

Результат побудови графіків регресійної моделі з відрізків поліномів для:

span = 0.3

span = 0.5





Завдання 2. Побудуйте графіки регресійної моделі для span=0.75 і span=1 та оцініть якість моделі для цих параметрів.

Приклад 4. Для побудови спеціальних моделей можна використовувати спеціальні функції пакету MathCad. Для ілюстрації цього візьмемо попередній вектор x1, а вектор z1 сформуємо за формулою:

$$z1 = \frac{5\varepsilon}{1 + 2\exp(-3x1)},$$

де ε – випадкова величина, рівномірно розподілена на відрізку [0,9; 1,1]. Фрагмент програми побудови експоненційної та логістичної регресійної моделі:

$x1 := (-2.5\ -2.25\ -2\ -1.75\ -1.5\ -1.25\ -1\ -0.75\ -0.5\ -0.25\ 0\ 0.25\ 0.5\ 0.75\ 1\ 1.25\ 1.5\ 1.75\ 2\ 2.25\ 2.5)^T$

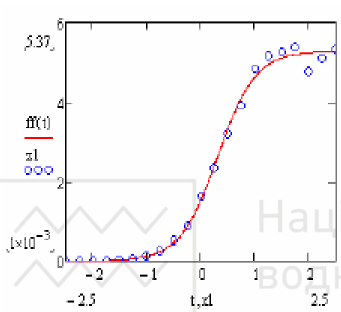
$z1 := (0.001\ 0.003\ 0.006\ 0.014\ 0.026\ 0.061\ 0.12\ 0.27\ 0.53\ 0.88\ 1.64\ 2.34\ 3.19\ 3.91\ 4.83\ 5.16\ 5.25\ 5.37\ 4.76\ 5.09\ 5.33)^T$

$$g := \begin{pmatrix} 5 \\ 2 \\ 2 \end{pmatrix} \qquad V := \mathrm{expfit}(x1, z1, g) \qquad V = \begin{pmatrix} 10.233 \\ 0.135 \\ -8.123 \end{pmatrix}$$

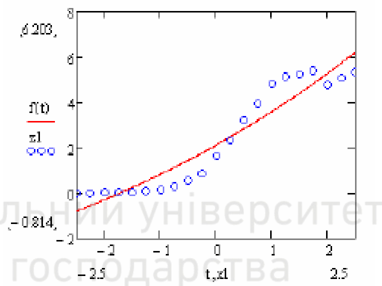
f(t):=V0 exp(V1t)+V2

$$S:=\mathrm{lgsfit}(x1,z1,g) \qquad S=\begin{pmatrix}5.266\\2.461\\2.865\end{pmatrix}$$

$$ff(t):=\frac{S_0}{1+S_1\exp(-S_2 t)}$$



а)



б)

На рисунках наведено результати побудови експоненційної а) та логістичної б) регресійних моделей. Як бачимо, логістична модель краще відображає емпіричні дані, що у даному випадку є цілком природним, оскільки вхідні дані побудовано саме на основі логістичної моделі.

Завдання 3. Розвязати задачі:

Задача 1. В таблиці наведено дані про рівень механізації робіт X(%) та продуктивність праці Y(т/год) для 14 однотипних підприємств.

xi	32	30	36	40	41	47	56	54	60	55	61	67	69	76
yi	20	24	28	30	31	33	34	37	38	40	41	43	45	48

Необхідно:

а) оцінити тісноту і напрям зв'язку між змінними за допомогою коефіцієнта кореляції;

б) знайти рівняння регресії.

Задача 2. За результатом досліджень кореляційної залежності між ціною на нафту X та індексом нафтових компаній Y одержані наступні дані: x =16,2 (грош. од), y = 4000 (умовн. од), s_x^2 =4, s_y^2 =500, Cov(X,Y)=40.

Необхідно:

а) скласти рівняння регресії Y(X);

б) використовуючи рівняння регресії, знайти середнє значення індексу при ціні на нафту 16,5 грош. од.

Задача 3. За даними задачі 1:

а) знайти рівняння регресії Y(X);

б) знайти коефіцієнт детермінації R^2 і пояснити його зміст;

в) перевірити значущість рівняння регресії на 5%-му рівні за F-критерієм;

г) оцінити середню продуктивність праці на підприємствах з рівнем механізації робіт 60% та побудувати для неї 95%-ий довірчий інтервал;

д) аналогічний довірчий інтервал знайти для індивідуальних значень продуктивності праці на тих самих підприємствах.

Задача 4. За даними 30 нафтових компаній одержане наступне рівняння регресії між оцінкою Y (грош. од.) та фактичною вартістю X (грош. од.) цих компаній: $y_x = 0{,}8750x + 295$. Знайти: 95%-ні довірчі інтервали для середнього та індивідуального значень оцінки підприємств, фактична вартість яких склала 1300 грош. од., якщо коефіцієнт кореляції між змінними дорівнює 0,76, а середнє квадратичне відхилення змінної X дорівнює 270 грош. од.

Контрольні запитання

1. Основне завдання регресійного аналізу.
2. Основні припущення класичного регресійного аналізу.
3. Як формулюється задача побудови регресійної моделі?
4. Сформулюйте суть методу найменших квадратів.
5. Що являє собою коефіцієнт регресії, коефіцієнт кореляції, коефіцієнт детермінації?

6. Які функціонали використовують для визначення параметрів регресійних моделей? У чому полягають переваги й недоліки різних типів таких функціоналів?

7. Якими є основні типи функцій, що використовуються для побудови однофакторних регресійних моделей?

8. Які моделі називають лінійними? Що називають порядком регресійної моделі?

9. У яких випадках нелінійні однофакторні моделі можна звести до лінійних? Навести приклади відповідних перетворень.

10. Як використовують критерій Фішера для перевірки адекватності регресійних моделей?

11. Як визначають довірчі інтервали для коефіцієнтів однофакторних регресійних моделей?

12. Загальний вигляд поліноміальної регресійної моделі.

13. Загальний алгоритм визначення порядку і параметрів поліноміальних регресійних моделей.

Лабораторна робота № 5

Тема: Мультиколінеарність та її вплив на оцінки параметрів моделі

Основні відомості

Приклад 1. Для дослідження залежності між продуктивністю праці (X1), віком (X2) і виробничим стажем (X3) була проведена вибірка із 100 робітників тієї самої спеціальності. Обчислені парні коефіцієнти кореляції виявилися значущими і склали: r_{12}=0,20; r_{13}=0,41; r_{23}=0,82. Обчислити часткові коефіцієнти кореляції та оцінити їх значущість на рівні α=0,05.

Розв'язання

Часткові коефіцієнти кореляції обчислюють за формулою

$$r_{ijk} = \frac{r_{ij} - r_{ik} r_{jk}}{\sqrt{\left(1 - r_{ik}^2\right)\left(1 - r_{jk}^2\right)}}, \; r_{123} = \frac{0{,}2^2 - 0{,}41 \cdot 0{,}82}{\sqrt{(1 - 0{,}41)^2 (1 - 0{,}82)^2}} = -0{,}26 .$$

Аналогічно отримаємо r_{132}=0,44; r_{231}=0,83.

Оцінимо значущість r_{123}. Значення статистики t-критерію за формулою при n'=n-p+ 2= 100-3+2=99 (за абсолютною величиною)

$$|t_1| = \frac{|-0,26|\sqrt{99-2}}{\sqrt{1-(-0,26)^2}} = 2,65$$

перевищує табличне $t_{0,95;97}$=1,99, отже, частковий коефіцієнт кореляції r_{123} є значущим. Аналогічно встановлюється значущість інших часткових коефіцієнтів кореляції.

Порівнюючи часткові коефіцієнти кореляції r_{ijk} з відповідними парними коефіцієнтами, бачимо, що за рахунок «очищення зв'язку» найбільшій зміні піддався коефіцієнт кореляції між продуктивністю праці (X1) і віком (X2) робітників (змінилося не тільки його значення, але і знак r_{12}=0,20; r_{123}=-0,26, причому обидва ці коефіцієнти є значущими).

Отже, між продуктивністю праці (X1) і віком (X2) робітників існує прямий кореляційний зв'язок (r12=0,20). Якщо ж усунути (елімінувати) вплив змінної «виробничий стаж» (X3), то в чистому вигляді продуктивність праці (X1) знаходиться у зворотному за напрямом (і слабкому за тіснотою) зв'язку з віком робітників (X2) (r123=-0,26). Це цілком зрозуміло, якщо розглядати вік тільки як показник працездатності організму на певному етапі його життєдіяльності. Так само можуть бути інтерпретовані й інші часткові коефіцієнти кореляції.

Завдання.

Задача 1. Є наступні дані про споживання певного продукту Y (умовн. од.) залежно від рівня урбанізації (частки міського населення) X1, відносного освітнього рівня X2 і відносного заробітку X3 для дев'яти географічних районів (див. табл. 1).

Таблиця 1

i (номер району)	xi1	xi2	xi3	yi
1	42,2	11,2	31,9	167,1
2	48,6	10,6	13,2	174,4
3	42,6	10,6	28,7	160,8
4	39,0	10,4	26,1	162,0
5	34,7	9,3	30,1	140,8
6	44,5	10,8	8,5	174,6

7	39,1	10,7	24,3	163,7
8	40,1	10,0	18,6	174,5
9	45,9	12,0	20,4	185,7
Середні	41,85	10,62	24,42	167,07

Стандартні відхилення $s_{x1} = 4{,}176$; $s_{x2} = 0{,}7463$; $s_{x3} = 7{,}928$; s_y=12,645. Кореляційна матриця має вигляд, наведений у таблиці 2.

Таблиця 2

	X1	X2	X3	Y
X1	1	0,684	-0,616	0,802
X2	0,684	-1	-0,173	0,770
X3	0,616	-0,173	1	-0,629
Y	0,802	0,770	-0,629	1

Використовуючи покрокову процедуру відбору найбільш інформативних пояснюючих змінних, визначити відповідну регресійну модель, виключивши при цьому мультиколінеарність. Оцінити значущість коефіцієнтів регресії одержаної моделі за t-критерієм.

Задача 2. Маємо дані про масу Y (у кілограмах) та вік X (у тижнях) 13 індичок, вирощених в областях А, В, С (див. табл.).

i	xi	yi	Область походження
1	28	12,3	A
2	20	8,9	A
3	32	15,1	A
4	22	10,4	A
5	29	13,1	B
6	27	12,4	B
7	28	13,2	B
8	26	11,8	B
9	21	11,5	C
10	27	14,2	C
11	29	15,4	C
12	23	13,1	C
13	25	13,8	C

Є підстава вважати, що на вагу індичок робить вплив не тільки їх вік, але й область походження. Необхідно:

а) знайти рівняння парної регресії Y_X та оцінити його значущість;

б) ввівши відповідні фіктивні змінні, знайти загальне рівняння множинної регресії за всіма пояснюючими змінними (включаючи фіктивні);

в) оцінити значущість загального рівняння множинної регресії за F-критерієм та значущість його коефіцієнтів за t-критерієм на рівні α=0,05;

г) простежити за зміною корегованого коефіцієнта детермінації при переході від парної до множинної регресії;

д) оцінити на рівні α=0,05 значущість відмінності між вільними членами рівнянь, що одержуються із загального рівняння множинної регресії Y для кожної області.

Задача 3. Під час побудови лінійної залежності витрат на одяг від доходу за вибіркою для 10 жінок одержані наступні суми квадратів та добутків спостережень:

$$\sum_1^{10} x_i = 110,\ \sum_1^{10} x_i^2 = 1540,\ \sum_1^{10} y_i = 60,\ \sum_1^{10} x_i y_i = 828,\ \sum_1^{10} y_i^2 = 448.$$

Аналогічні обчислення сум за вибіркою з 5 чоловіків дали:

$$\sum_1^{5} x_i = 35,\ \sum_1^{5} x_i^2 = 325,\ \sum_1^{5} y_i = 15,\ \sum_1^{5} x_i y_i = 140,\ \sum_1^{5} y_i^2 = 61.$$

За загальною (об'єднаною) вибіркою оцінено регресію з використанням фіктивної змінної Z (Z= 1 для чоловіка і Z= 0 для жінки), яка має вигляд:

$$\hat{y} = -0{,}06 + 0{,}438x + 0{,}46z + 0{,}072zx.$$

На рівні значущості α=0,05 перевірити гіпотезу, що функція споживання одна і та ж для чоловіків та жінок, якщо виконані всі передумови класичної нормальної лінійної регресії.

Задача 4. З метою дослідження впливу чинників X1 – середньомісячної кількості профілактичних налаштувань автоматичної лінії та X2 – середньомісячної кількості обривів нитки на чинник Y – середньомісячну характеристику якості тканини (у балах) за даними 37 підприємств легкої промисловості були обчислені парні коефіцієнти кореляції:

r_{y1}=0,105, r_{y2}=0,024 та r_{12}=0,996. Визначити часткові коефіцієнти кореляції r_{y12} та r_{y21} і оцінити їх значущість на 5%-му рівні.

Контрольні запитання

1. Поняття мультиколінеарності. У яких формах вона може проявлятися?
2. Які критерії використовують для визначення наявності або відсутності мультиколінеарності?
3. Охарактеризуйте методи, які використовують для усунення або зменшення мультиколінеарності.
4. Які наслідки включення до моделі мультиколінеарних факторів?
5. Як для оцінки мультиколінеарності використовують визначник матриці парних коефіцієнтів кореляції між факторами?
6. Поясніть алгоритм покрокових процедур відбору найінформативніших змінних.

Лабораторна робота № 6

Тема: Економетричні моделі динаміки

Основні відомості

Приклад 1. Часовий ряд попиту y_t наведено у таблиці. За даними таблиці визначити середнє значення, середнє квадратичне відхилення, коефіцієнти автокореляції (для лагів τ=1; 2) і частковий коефіцієнт автокореляції 1-го порядку.

Рік, t	1	2	3	4	5	6	7	8
Попит, y_t	213	171	291	309	317	362	351	361

Розв'язання

Знаходимо середнє значення часового ряду за формулою:

$$\bar{y}_t = \frac{\sum_{t=1}^{n} y_t}{n} = \frac{213+171+291+309+317+362+351+361}{8} = 296{,}88 \text{ (од.)}.$$

Дисперсію і середнє квадратичне відхилення обчислимо за співвідношеннями:

$$s_t^2 = \overline{y_t^2} - \overline{y_t}^2 = 92478{,}38 - 296{,}88^2 = 4343{,}61,$$

$$s_t = \sqrt{4343{,}61} = 65{,}31 \text{(од.)},$$

де

$$\overline{y_t^2} = \frac{\sum_{t=1}^{n} y_t^2}{n} = \frac{213^2 + 171^2 + 291^2 + 309^2 + 317^2 + 362^2 + 351^2 + 361^2}{8} = 92478{,}38.$$

Знайдемо коефіцієнт автокореляції r(τ) часового ряду (для лага τ=1), тобто коефіцієнт кореляції між послідовностями семи пар спостережень y_t та y_{t+1}(t= 1,2...,7). Спостереження наведені у таблиці.

y_t	213	171	291	309	317	362	351
y_{t+1}	171	291	309	317	362	351	361

Обчислюємо необхідні суми:

$$\sum_{t=1}^{7} y_t = 213 + 171 + 291 + 309 + 317 + 362 + 351 = 2014,$$

$$\sum_{t=1}^{7} y_t^2 = 213^2 + 171^2 + 291^2 + 309^2 + 317^2 + 362^2 + 351^2 = 609506,$$

$$\sum_{t=1}^{7} y_{t+\tau} = 171 + 291 + 309 + 317 + 362 + 351 + 361 = 2162,$$

$$\sum_{t=1}^{7} y_{t+\tau}^2 = 171^2 + 291^2 + 309^2 + 317^2 + 362^2 + 351^2 + 361^2 = 694458,$$

$$\sum_{t=1}^{7} y_t y_{t+\tau} = 213 \cdot 171 + 171 \cdot 291 + 291 \cdot 309 + 309 \cdot 317 + 317 \cdot 362 +$$
$$+ 362 \cdot 351 + 351 \cdot 361 = 642583.$$

Тепер обчислимо коефіцієнт автокореляції

$$r(1) = \frac{7 \cdot 642583 - 2014 \cdot 2162}{\sqrt{7 \cdot 609506 - 2014^2}\sqrt{7 \cdot 694458 - 2162^2}} = 0{,}725.$$

Коефіцієнт автокореляції r(2) для лага τ = 2 між членами ряду y_t і y_{t+2} (t=1,2...,6) за шістьма парами спостережень обчислюємо аналогічно: r(2) = 0,842.

Для визначення часткового коефіцієнта кореляції 1-го порядку $r_{част}(2) = r_{021}$ між членами ряду y_t і y_{t+2} при виключенні впливу y_{t+1} спочатку знайдемо (за аналогією з попереднім) коефіцієнт автокореляції r(1,2) між членами ряду y_{t+1} і y_{t+2}: r(1,2)=0,825, а потім обчислимо $r_{част}(2)$:

$$r_{част}(2) = r_{021} = \frac{0{,}842 - 0{,}725 \cdot 0{,}825}{\sqrt{1-0{,}725^2}\sqrt{1-0{,}825^2}} = 0{,}627\ .$$

Знання автокореляційних функцій r(τ) і $r_{част}(\tau)$ може надати істотну допомогу при підборі та ідентифікації моделі часового ряду, що аналізується, і статистичній оцінці його параметрів.

Приклад 2. За даними прикладу 1 знайти рівняння невипадкової складової (тренду) для часового ряду y_t, вважаючи тренд лінійним.

Розв'язання

Відповідно до методу найменших квадратів параметри лінійної залежності визначають із системи нормальних рівнянь. Враховуючи, що значення t утворюють натуральний ряд чисел від 1 до n, суми $\sum_{t=1}^{n} t$ та $\sum_{t=1}^{n} t^2$ можна виразити як кількість членів ряду n за відомими формулами:

$$\sum_{t=1}^{n} t = \frac{n(n+1)}{2} = \frac{8\cdot 9}{2} = 36, \quad \sum_{t=1}^{n} t^2 = \frac{n(n+1)(2n+1)}{6} = \frac{8\cdot 9\cdot 17}{6} = 204\ .$$

Далі

$$\sum_{t=1}^{8} y_t = 213 + 171 + 291 + 309 + 317 + 362 + 351 + 361 = 2375\ ,$$

$$\sum_{t=1}^{8} y_t^2 = 213^2 + 171^2 + 291^2 + 309^2 + 317^2 + 362^2 + 351^2 + 361^2 = 739827\ ,$$

$$\sum_{t=1}^{8} y_t t = 213\cdot 1 + 171\cdot 2 + 291\cdot 3 + 309\cdot 4 + 317\cdot 5 +$$

$$+ 362\cdot 6 + 351\cdot 7 + 361\cdot 8 = 11766\ .$$

Система нормальних рівнянь має вигляд:

$$\begin{cases} 8b_0 + 36b_1 = 2375\ , \\ 36b_0 + 204b_1 = 11766\ . \end{cases}$$

Розв'язавши систему, отримаємо b_0=181,32 та b_1=25,679. Отже рівняння тренду запишеться у вигляді:

$\hat{y}_t = 181{,}32 + 25{,}679t$, тобто попит щорічно збільшується у середньому на 25,7 одиниць. Перевіримо значущість отриманого рівняння тренду за F-критерієм на 5%-му рівні значущості. Обчислимо дисперсії. Дисперсія зумовлена регресією:

$$Q_R = \sum_{t=1}^{n}\left(\hat{y}_t - \bar{y}_t\right)^2 = \sum b_1^2\left(t - \bar{t}\right)^2 = b_1^2\left(\sum_{t=1}^{n} t^2 - \frac{\left(\sum_{t=1}^{n} t\right)^2}{n}\right) =$$

$$= 25{,}679^2\left(204 - \frac{36^2}{8}\right) = 27695{,}3.$$

Дисперсія загальна:

$$Q = \sum_{t=1}^{n}\left(y_t - \bar{y}_t\right)^2 = \sum y_t^2 - \frac{\left(\sum_{t=1}^{n} y_t\right)^2}{n} = 739827 - \frac{2375^2}{8} = 34748{,}9 .$$

Дисперсія залишкова:

Qзал=Q-QR=34748,9-27695,3=7053,6.

Обчислимо значення статистики

$$F = \frac{Q_R(n-2)}{Q_{зал}} = \frac{27695{,}3 \cdot 6}{7053{,}6} = 23{,}56 .$$

Оскільки $F > F_{0,05;1;6}$, рівняння тренду є значущим.

Завдання.

Задача 1. Дані про врожайність озимої пшениці у (ц/га) за 10 років наведено в таблиці.

t	1	2	3	4	5	6	7	8	9	10
y_t	16,3	20,2	17,1	7,7	15,3	16,3	19,9	14,4	18,7	20,7

Знайти середнє значення, середнє квадратичне відхилення та коефіцієнти автокореляції для лагів τ = 1;2 часового ряду.

Задача 2. За даними таблиці із задачі 1 знайти рівняння тренду часового ряду y_t, вважаючи що він лінійний. Перевірити значущість тренду часового ряду на рівні 0,05.

Задача 3. У наступній таблиці подані дані, що відображають динаміку зростання доходів на душу населення y_t (грош. од.) за восьмирічний період.

t	1	2	3	4	5	6	7	8
y_t	1133	1222	1354	1389	1342	1377	1491	1684

Вважаючи, що тренд лінійний і умови класичної моделі виконані, знайти рівняння тренду, оцінити його значущість на рівні 0,05 та дати точковий і з надійністю 0,95 інтервальний прогнози середнього та індивідуального значень доходів на дев'ятий рік.

Контрольні запитання

1. Наведіть визначення часового ряду.
2. Чим відрізняються часові ряди від просторових вибірок?
3. Які фактори впливають на рівні часового ряду?
4. Які компоненти містить реальний часовий ряд?
5. Поясніть адитивну і мультиплікативну моделі часового ряду.
6. Тренд часового ряду, види тренду.
7. Як визначають значущість тренду?
8. Перелічіть основні етапи аналізу часових рядів.
9. Перелічіть найпоширеніші методи аналізу часових рядів.
10. Автокореляція рівнів часового ряду та її вимірювання.
11. Перелічіть властивості коефіцієнта автокореляції.
12. Автокореляційна функція часового ряду і корелограма.
13. Призначення аналітичного вирівнювання часового ряду.
14. Які функції найчастіше застосовують для побудови трендів?
15. Який метод дозволяє визначити параметри тренду?
16. Методи згладжування часового ряду.
17. В чому полягає метод ковзних середніх?
18. Як провадиться усунення сезонних коливань за методом ковзної середньої?
19. Яка мета спектрального аналізу часового ряду?
20. Поясніть суть зв'язного аналізу часових рядів.

Національний університет
водного господарства
та природокористування

Література

1. Ачкасов А. Є., Воронков О. О. Конспект лекцій з курсу «Економіко-математичне моделювання». Харків: ХНАМГ, 2011. 204 с.
2. Бахрушин В.Є. Методи аналізу даних: навч. посібник. Запоріжжя: КПУ, 2011. 268 с.
3. Вітлінський В. В. Моделювання економіки: Навч. посібник. К.: КНЕУ, 2005. 408 с.

4. Воронков О. О., Медведєв І. А. Методичні вказівки для практичних занять та самостійної роботи з навч. дисципліни «Економіко-математичне моделювання». Харків: ХНУМГ, 2016. 69 с.

5. Красін М. А. Завдання для самостійної роботи студентів з навчальної дисципліни "Моделі економічної динаміки". Рівне: РДГУ, 2014. 29 с.

6. Лук'янова В. В. Комп'ютерний аналіз даних: посібник. К.: Академія, 2003. 344 с.

7. Ляшенко І. М., Коробова М. В., Столяр А. М. Основи математичного моделювання економічних, екологічних та соціальних процесів. Навч. посіб. Тернопіль: Навчальна книга-Богдан, 2006. 304 с.

8. Іванків К. С., Щербатий М. В. Математичне моделювання біологічних та еколого-економічних процесів Львів: ЛНУ ім. І. Франка, 2005. 154 с.

9. Пономаренко К. А. Основи економічної кібернетики. К.: КНТЕУ, 2002. 432 с.

10. Шарапов О. Д., Дербенцев В. Д., Семьонов Д. Є. Економічна кібернетика: Навч. посібник. К.: КНЕУ, 2003. 154с.

11. Елисеева И. И., Курышева С. В., Костеева Т. В. Эконометрика: учебник. М.: Финансы и статистика, 2007. 576 с.

12. Наконечный С. И., Терещенко Т. П. Эконометрия: учебник. К.: КНЕУ, 2001.

13. Салманов О. Н. Математическая экономика с применением MathCAD и Excel. СПб.: БХВ-Петербург, 2003.