УДК 004.89 https://doi.org/10.31713/vt1202223

Кінда В. В., аспірант (Рівненський державний гуманітарний університет, м. Рівне), **Турбал Ю. В., д.т.н., професор** (Національний університет водного господарства та природокористування, м. Рівне)

ЗАСТОСУВАННЯ НЕЧІТКОЇ ЛОГІКИ ДЛЯ РОБОТИ ІЗ КОРОТКОСТРОКОВИМИ ЧАСОВИМИ РЯДАМИ

Розглянуто основні компоненти нечітких множин, принцип та етапи їх побудови. Проведено опис основних функцій належності, правила їх застосування. Розглянуто метод побудови нечіткої множини для часових рядів із використанням його характеристики. Наведено приклад визначення універсальної множини для часових рядів, побудови трикутних функцій належності на інтервалах універсальної множин.

Ключові слова: **часовий ряд; нечітка множина; функція належності; універсальна множина; прогноз.**

Актуальність теми. Зі збільшенням обчислювальних потужностей дедалі більшої популярності набирають методи глибокого навчання. Еволюція архітектури штучних нейронних мереж пройшла свій розвиток від звичайного перцептрону Розенблата зі стандартною функцією активації до багатошарових складних ієрархічних архітектур, що дозволяють працювати не лише із економетричними процесами, але й даними аудіо, фото та відеоформатів. Такі методи та моделі показують добрий результат на великих наборах даних (big data), які дають можливість вирішувати задачі широкого спектру.

В нашому випадку дослідження малих об'ємів даних створює певні складнощі та практично унеможливлює використання вищезгаданих моделей. Саме тому питання прогнозування на малих об'ємах даних являє собою високий інтерес у науковій сфері.

Аналіз досліджень. За останні роки проведено велику кількість наукових досліджень в напрямі короткострокових прогнозів інтелектуальними методами. Прогнозування відіграє провідну роль у процесах прийняття управлінських рішень. Теорія нечітких множин застосовувалася до проблем прогнозування в різних галузях науки. Однією із перших робіт над нечіткими часовими рядами є дослідження Song and Chisson (1993) [1], яке знайшло продовження у роботі Fuzzy time series and its models. Fuzzy sets and systems [2]. Простота, мала обчислювальна потужність та швидкість одержання результату спонукає до необхідності в подальшому дослідженні нечітких множин для точкового прогнозу часових рядів.

Метою статті є дослідження теорії нечітких множин та їх застосування для короткострокового прогнозу економетричних часових рядів на базі обраного критерію.

Викладення основного матеріалу. Теорія нечітких множин (fuzzy sets) та нечітка логіка (fuzzy logic) є узагальненнями класичної теорії множин та класичної формальної логіки. Дані поняття були вперше запропоновані американським вченим Лотфі Заде (Lotfi Zadeh) у 1965 р. та глибоко розвинені наприкінці 80-х Бартоломеєм Коско – теорема FAT (Fuzzy Approximation Theorem) [1].

Під універсальною множиною ***U*** розуміємо множину всіх елементів, на базі якої можна побудувати нечітку множину. Із теорії множин нам відомо, що кожна множина задає деяку семантику, що об'єднує його елементи в одне ціле.

Терм – якісний опис значення параметра.

Лінгвістична змінна – сукупність всіх термів, за допомогою яких можна описати значення одного параметру.

Нехай *C* – підмножина ***U*** ($C \subset U$) характеризується функцією належності (MF – Membership Function), яка переводить кожен елемент із *C* в числовий інтервал.

$$MF_c : U \to [0,1]$$

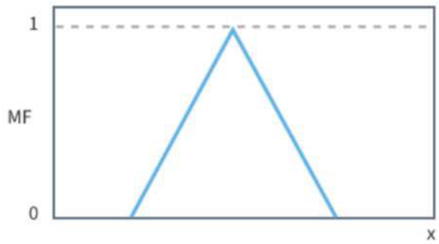
Значення $MF_c(x) = 0$ означає відсутність належності до множини, 1 – повну належність.

Існують різні варіанти задання функції належності. Найбільш поширеними є трикутна (рис. 1, а), трапецеїдальна (рис. 1, б) та Гаусівська (рис. 2) функції належності [2].

Трикутна MF визначається трійкою чисел *(a, b, c)*.

$$MF(x) = \begin{cases} 1 - \dfrac{b-x}{b-a}, a \le x \le b \\ 1 - \dfrac{x-b}{c-b}, b < x \le c \\ 0, x \notin [a,c] \end{cases}$$

При $(b-a) = (c-b)$ трикутна функція належності може задаватись двома значеннями із трійки.



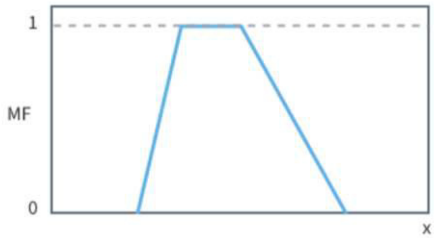


Рис. 1. а) вигляд трикутної функції належності, б) Вигляд трапецієвидної функції належності

Для задання трапецієвидної функції належності необхідна четвірка чисел *(a, b, c, d)*.

$$MF(x) = \begin{cases} 1 - \dfrac{b-x}{b-a}, a \le x \le b \\ 1, b < x \le c \\ 1 - \dfrac{x-c}{d-c}, c < x \le d \\ 0, x \notin [a, d] \end{cases}$$

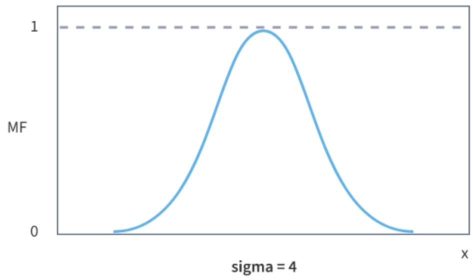


Рис. 2. Вигляд гаусівської функції належності

Функція належності гаусового типу описується формулою:

$MF(x) = exp[-\left(\frac{x-c}{\sigma}\right)^2]$ та оперує двома параметрами: *c* — позначає центр нечіткої множини, а параметр σ відповідає за крутість функції.

Нечітка множина задається наступним чином:

1. Формування універсальної множини;
2. Визначення сукупності термів, що визначають лінгвістичну змінну;
3. Для кожного терма *C* задається закон $MF_C(x)$, який визначає рівень відповідності елементів терму до універсальної множини ***U***.

$$C = \{\langle x, MF_c(x)\rangle | x \in U\}.$$

Більш детально про логічні операції над нечіткими множинами можна ознайомитись в роботах [3–5].

Алгоритм побудови прогнозу

При роботі із часовими рядами на малих вибірках можна досліджувати не самі значення ряду, а його різниці першого роду, тобто приріст функції. Фізичний зміст цих різниць являє собою швидкість процесу, оскільки чисельна зміна швидкості не може передати якісно інформацію в силу можливої різкої зміни значень величин. Для адекватної побудови будемо використовувати відсоткову зміну швидкості.

1. Нехай маємо часовий ряд виду $y_i, i \in [1, n]$

2. Побудуємо ряд відсоткової зміни швидкості наступним чином:

$$V_i = (\frac{y_i}{y_{i-1}} - 1)100\%, i \in [2, n].$$

3. Універсальну множину ***U*** задаємо наступним чином $U = [min(V_i), max(V_i)], i \in [2, n]$.

Автори [6] застосовують збільшення верхньої та нижньої межі на певну величину, що встановлюється емпіричним чином, задля забезпечення належності прогнозу цій множині.

4. Необхідно розділити ***U*** на інтервали, задати терми для лінгвістичної змінної. Кількість інтервалів є одним з найважливіших параметрів нечіткого часового ряду, який безпосередньо впливає на точність моделі.

Модель Stevenson and Porter передбачає розбиття універсальної множини на 7 інтервалів. В роботі [6] вказується, що оптимальна кількість інтервалів варіюється в залежності від величини вибірки на сягає переважно 5–15 інтервалів.

В роботі [7] зазначається про можливість поділу на субінтервали у відношенні 4,3,2 для інтервалів із найбільшими частотами.

5. Визначити терми лінгвістичної змінної із *Ci* у вигляді трикутних функцій приналежності, із заданням їх параметрів. $C = \{C_1, C_2 \dots C_m\}$, де *m* кількість інтервалів розбиття.

6. Провести дефазифікацію наступним чином:

$$v_j = \begin{cases} \dfrac{1+0.5}{\dfrac{1}{a_j}+\dfrac{0.5}{a_{j+1}}}, j = 1 \\ \dfrac{0.5+1+0.5}{\dfrac{0.5}{a_{j-1}}+\dfrac{1}{a_j}+\dfrac{0.5}{a_{j+1}}}, j \in [2, m-1] \\ \dfrac{0.5+1}{\dfrac{0.5}{a_{j-1}}+\dfrac{1}{a_j}}, j = m \end{cases}$$

a_j – центри трикутних функцій приналежності відповідних термів.

7. Зробити прогноз використовуючи $\hat{y}_i = y_{i-1}(1 + \frac{v_i}{100})$.

Висновки та перспективи подальших досліджень

Розглянутий метод моделювання часових рядів дозволяє отримати точковий прогноз шляхом аналізу приростів часового ряду за допомогою нечіткого оцінювання на основі нечітких множин.

Подальші дослідження в цьому напрямі будуть спрямовані на розробку єдиної загальної методології інтелектуального розбиття універсальної множини на інтервали та принцип отримання кількості цих інтервалів для прогнозу із найменшою похибкою відповідного критерію. Удосконалення таких методів дасть можливість запровадити кращі алгоритми прогнозування в системах підтримки прийняття рішень.

1. Q. Song and B. S. Chissom. Forecasting enrollments with fuzzy time series – part I, Fuzzy Sets and Systems 54. 1993. **2.** Q. Song and B. S. Chissom. Forecasting enrollments with fuzzy time series – part II, Fuzzy Sets and Systems 62. 1994. **3.** S. M. Chen. Forecasting enrollments based on fuzzy time series, Fuzzy Sets and Systems 81. 1996. **4.** C. H. Cheng, J. R. Chang and C. A. Yeh. Entropy-based and trapezoid fuzzification fuzzy time series approaches for forecasting IT project cost. *Technological Forecasting & Social Change*. 2006. 73. **5.** S. R. Sing. A robust method of forecasting based on fuzzy time series. *Applied Mathematics and Computation.* 2007. 188. **6.** T. A. Jilani, S. M. A. Burney, C. Ardil. Fuzzy Metric Approach for Fuzzy Time Series Forecasting based on Frequency Density Based Partitioning. *Proceedings of World Academy of Science, Engineering and Technology*. 2007. Vol. 23, Pp. 333–338. **7.** T. A. Jilani, S. M. A. Burney, C. Ardil. Multivariate high order fuzzy time series forecasting for car road accidents. *International Journal of Computational Intelligence*. 2007. Vol. 4, No. 1. Pp. 15–20.

REFERENCES:

1. Q. Song and B. S. Chissom. Forecasting enrollments with fuzzy time series – part I, Fuzzy Sets and Systems 54. 1993. **2.** Q. Song and B. S. Chissom. Forecasting enrollments with fuzzy time series – part II, Fuzzy Sets and Systems 62. 1994. **3.** S. M. Chen. Forecasting enrollments based on fuzzy time series, Fuzzy Sets and Systems 81. 1996. **4.** C. H. Cheng, J. R. Chang and C. A. Yeh. Entropy-based and trapezoid fuzzification fuzzy time series approaches for forecasting IT project cost. *Technological Forecasting & Social Change*. 2006. 73. **5.** S. R. Sing. A robust method of forecasting based on fuzzy time series. *Applied Mathematics and Computation.* 2007. 188. **6.** T. A. Jilani, S. M. A. Burney, C. Ardil. Fuzzy Metric Approach for Fuzzy Time Series Forecasting based on Frequency Density Based Partitioning. *Proceedings of World Academy of Science, Engineering and Technology*. 2007. Vol. 23, Pp. 333–338. **7.** T. A. Jilani, S. M. A. Burney, C. Ardil. Multivariate high order fuzzy time series forecasting for car road accidents. *International Journal of Computational Intelligence*. 2007. Vol. 4, No. 1. Pp. 15–20.

Kinda V. V., Post-graduate Student (Rivne State University of the Humanities, Rivne), **Turbal Yu. V., Doctor of Engineering, Professor** (National University of Water and Environmental Engineering, Rivne)

APPLICATION OF FUZZY LOGIC FOR SHORT-TERM TIME SERIES

The study of econometric time series forecasting on small volumes of data is one of the key forecasting tasks. That is why a detailed study of the possibility of using fuzzy systems in this class of problems is expedient. This article considers the main components of fuzzy sets, the main principles and stages of their construction.

A description of the membership functions, the rules of their application, and the methods of specifying a fuzzy set is carried out. The method of constructing a fuzzy set for time series using its characteristics is considered. An example of defining a universal set for time series, constructing triangular membership functions on intervals of a universal set is given.

The algorithm for constructing a forecast using the abovementioned method was considered and its implementation in the high-level python programming environment was made.

Further research in this direction will be aimed at the development of a single general methodology for the intelligent division of the universal set into intervals and the principle of obtaining the number of these intervals for forecasting with the smallest error of the corresponding criterion. Improvement of such methods will make it possible to introduce better forecasting algorithms in decision support systems.

***Keywords*: time series; fuzzy set; membership function; universal set; forecast.**