

АНАЛІЗ ТА ПРОГНОЗУВАННЯ РІЧНИХ ТЕМПІВ ІНФЛЯЦІЇ

Вяліна А.В.

e-mail: anzhelika.vialina@nure.ua

Науковий керівник – канд. фіз.-мат. наук, доц. Ламтюгова С.М.

Харківський національний університет радіоелектроніки,
кафедра прикладної математики

м. Харків, Україна

The paper proposes an algorithm for studying inflation processes in Asian countries based on a combination of cluster analysis methods and statistical modeling of time series. The algorithm includes the following stages: data preprocessing, visual analysis of inflation dynamics using a time-series plot and a box-plot, clustering of countries with similar inflation patterns using the k-means method and forecasting annual inflation rates for the countries belonging to one of the clusters.

Інфляція є одним із ключових макроекономічних показників, що характеризує стан економічного розвитку країни. Дослідження динаміки інфляційних процесів дає змогу виявити закономірності зміни цін, порівняти економічну стабільність різних країн та побудувати прогнози майбутніх значень інфляції.

Нехай y_t – значення річного темпу інфляції в момент часу t ($t = \overline{1, T}$) для обраної країни. Метою дослідження є побудова прогнозу річного темпу інфляції. Для опису динаміки досліджуваного показника та здійснення прогнозування використовується модель авторегресії інтегрованого ковзного середнього $ARIMA(p, d, q)$, що застосовується до диференційованого часового ряду [1]:

$$y_t = c + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t,$$

де ϕ_i , $i = 1, \dots, p$, – коефіцієнти авторегресії; θ_j , $j = 1, \dots, q$, – коефіцієнти ковзного середнього; ε_t – випадкова похибка (білий шум).

Розглянемо основні етапи дослідження інфляційних процесів.

На першому етапі виконуються аналіз та попередня обробка даних. Цей етап включає збір даних, пошук статистичних зв'язків, обробку пропущених значень та нормування даних. Збір даних передбачає отримання загальної інформації про набір даних, а також обчислення описової статистики. Для аналізу розподілу значень інфляції будуються гістограми, які дають змогу зробити висновки про можливий тип розподілу. Пошук статистичних зв'язків між країнами здійснюється шляхом побудови матриці кореляції. Особлива увага приділяється якості даних, оскільки багато алгоритмів машинного навчання є чутливими до наявності пропущених значень. У зв'язку з цим набір даних перевіряється на наявність пропущених або некоректних значень, які заповнюються середнім значенням або медіаною. Для

забезпечення коректної роботи алгоритмів дані мають бути приведені до єдиної шкали. У даному дослідженні застосовується стандартизація, яка забезпечує приведення ознак до порівнюваного масштабу.

На другому етапі будується графік динаміки річних темпів інфляції для всіх країн за період 1980–2024 років, що дає змогу візуально оцінити зміну інфляційних процесів у різних країнах та визначити періоди економічної нестабільності. Для порівняння статистичних характеристик показника між країнами використовується діаграма розмаху. Її аналіз дає змогу оцінити варіативність темпів інфляції, проаналізувати медіану та міжквартильний розмах, а також виявити можливі викиди й країни з високою волатильністю інфляції.

На третьому етапі проводиться кластеризація країн із подібною динамікою інфляції за допомогою методу k-середніх (k-means) [2]. Додатково застосовується ієрархічна кластеризація, на основі якої будується дендрограма, що дає змогу проаналізувати структуру кластерів та оцінити ступінь подібності інфляційної динаміки країн на різних рівнях агрегації. Для визначення оптимальної кількості кластерів застосовуються метод ліктя або метод силуєту.

На четвертому етапі обирається один із отриманих кластерів, для країн якого здійснюється прогнозування інфляції. Часові ряди поділяються на навчальну та тестову вибірки. Побудова ARIMA-моделі включає наступні кроки: перевірку часового ряду на стаціонарність, ідентифікацію структури моделі, оцінювання її параметрів, діагностику адекватності моделі та використання побудованої моделі для прогнозування майбутніх значень часового ряду [1]. Для перевірки стаціонарності можуть застосовуватися статистичні тести, зокрема тест Дікі-Фуллера. У разі необхідності ряд перетворюється у стаціонарний шляхом усунення трендової компоненти та сезонності. Параметри моделі оцінюються методом максимальної правдоподібності. Під час прогнозування також обчислюються довірчі інтервали, які характеризують можливий діапазон майбутніх значень інфляції.

На завершальному етапі проводиться оцінювання точності прогнозної моделі. Для цього використовуються статистичні метрики, зокрема середня абсолютна помилка (MAE), середньоквадратична помилка (RMSE) та середня абсолютна відносна помилка (MAPE). Порівняння прогнозних і фактичних значень показника дає змогу оцінити адекватність побудованої моделі та визначити ступінь її відповідності реальній динаміці інфляційних процесів.

Список використаної літератури:

1. Box G. Jenkins G. Time Series Analysis: forecasting and control, rev. ed. Oakland, California : Holden-Day, 1976. 575 p.
2. Alpaydin E. Introduction to Machine Learning. 3rd ed. Cambridge : MIT Press, 2014. 613 p.