

Маняк Юрій Вікторович

студент

Навчально-науковий комплекс «Інститут прикладного системного аналізу» НТУУ «КПІ»

Маняк Юрий Викторович

студент

Учебно-научный комплекс «Институт прикладного системного анализа» НТУУ «КПИ»

Manyak Y.

student

ESC «Institute for applied systems analysis» NTUU «KPI»

МОДЕЛІ ТА ПРОГНОЗИ НЕСТАЦІОНАРНИХ ПРОЦЕСІВ В ЕКОНОМІЦІ ТА ФІНАНСАХ

МОДЕЛИ И ПРОГНОЗЫ НЕСТАЦИОНАРНЫХ ПРОЦЕССОВ В ЭКОНОМИКЕ И ФИНАНСАХ

MODELS AND FORECASTS OF NON-STATIONARY PROCESSES IN ECONOMY AND FINANCES

Анотація. Розглянуто основні методи регресійного аналізу, розроблено та реалізовано у вигляді програмного продукту архітектуру СППР для аналізу, побудови структури моделей, моделювання та прогнозування нестационарних економетричних параметрів.

Ключові слова: регресійний аналіз, структура моделі, кореляційний аналіз, нестационарні процеси, економетричні показники.

Аннотация. Рассмотрены основные методы регрессионного анализа, разработана и реализована в виде программного продукта архитектура СППР для анализа, построения структуры моделей, моделирования и прогнозирования нестационарных эконометрических параметров.

Ключевые слова: регрессионный анализ, структура модели, корреляционный анализ, нестационарные процессы, эконометрические показатели.

Summary. Regression analysis methods were surveyed, DSS architecture was developed and implemented as a software product for analysis, model structure inference, modeling and forecasting of non-stationary econometric parameters.

Key words: regression analysis, model structure, correlation analysis, non-stationary processes, econometric parameters.

Вступ

Оцінка економетричних показників для України є складною та неоднозначною задачею через її належність до групи держав з перехідною економікою. Крім цього, більшість популярних пакетів з побудови моделей для прогнозування часових рядів пропонують лише можливість побудови моделі з наперед заданою структурою, що значно уповільнює пошук найкращої моделі в кожному конкретному випадку. Тому доцільно розробити систему підтримки прийняття рішень, моделювання та прогнозування економетричних даних, що дозволяє автоматизувати процес пошуку структури моделей із застосуванням результатів кореляційного аналізу.

Мета дослідження

Розробка архітектури СППР, її реалізація у вигляді власного програмного продукту для прогнозування та моделювання фінансово-економічних процесів, представлених у вигляді часових рядів, апробація системи на реальних статистичних даних.

Об'єкт дослідження

Статистичні ряди даних, що описують фінансово-економічні процеси та потребують ефективної аналітичної обробки для виявлення практично корисних знань та взаємозв'язків між атрибутами, необхідних для прийняття рішень в економіці.

Предмет дослідження

Моделі регресійного аналізу з прогнозування часових рядів: AR, множинної регресії, ARMA, ARMAX, ARIMA, моделей у вигляді тренду та їх варіацій.

Методи дослідження

Спираються на теорію ймовірностей, математичну статистику, теорію моделювання часових рядів.

Постановка задачі

1. Провести огляд основних регресійних методів та підходів до задачі прогнозування.
2. Спроекувати архітектуру СППР, призначеної для аналізу, моделювання та прогнозування економічних процесів.
3. Реалізувати методи регресійного аналізу: AR, множинної регресії, ARMA, ARMAX, ARIMA, моделей у вигляді тренду та їх варіацій; представити дані методів у вигляді програмного модуля; організувати інтерфейс для програми, що використовує отримані модулі.
4. Застосувати розроблений продукт до реальних статистичних даних.
5. Вибрати моделі з найкращими прогнозуючими якостями та виконати порівняльний аналіз отриманих результатів.

Використані для прогнозування моделі

До розгляду взято наступні регресійні моделі [1], на базі яких спроектовано архітектуру СППР:

1. Авторегресійна модель:

$$y(k) = a_0 + a_1 \cdot y(k-1) + a_2 \cdot y(k-2) + a_3 \cdot y(k-3) + \dots + a_p \cdot y(k-p) + \varepsilon(k),$$

де p – порядок авторегресійної частини.

2. Модель авторегресії з ковзним середнім:

$$y(k) = a_0 + a_1 \cdot y(k-1) + a_2 \cdot y(k-2) + a_3 \cdot y(k-3) + \dots + a_p \cdot y(k-p) + v(k) + b_1 \cdot v(k-1) + b_2 \cdot v(k-2) + b_3 \cdot v(k-3) + \dots + b_q \cdot v(k-q) + \varepsilon(k)$$

де q – порядок ковзного середнього.

3. Модель множинної регресії:

$$y(k) = a_0 + \sum_{s=1}^d c_s \cdot x_s + \varepsilon(k),$$

де d – кількість включених пояснюючих змінних.

4. Модель ARMAX:

$$y(k) = a_0 + \sum_{i=1}^p a_i \cdot y(k-i) + mv(k) + \sum_{j=1}^a b_j \cdot mv(k-j) + \sum_{s=1}^d c_s \cdot x_s + \varepsilon(k),$$

5. Модель у вигляді поліноміального тренду:

$$y(k) = a_0 + a_1 \cdot k + a_2 \cdot k^2 + \dots + a_m \cdot k^m + \varepsilon(k).$$

6. Модель авторегресії з інтегрованим ковзним середнім:

$$\hat{y}(k) = d^m y(k), \text{ де } dy(k) = y(k) - y(k-1)$$

$$\hat{y}(k) = a_0 + a_1 \cdot \hat{y}(k-1) + a_2 \cdot \hat{y}(k-2) + a_3 \cdot \hat{y}(k-3) + \dots + a_p \cdot \hat{y}(k-p) + \hat{v}(k) + b_1 \cdot \hat{v}(k-1) + b_2 \cdot \hat{v}(k-2) + b_3 \cdot \hat{v}(k-3) + \dots + b_q \cdot \hat{v}(k-q) + \varepsilon(k)$$

Архітектура розробленої СППР

Зі структурної схеми СППР, поданої на рис. 1, видно, що архітектура програми передбачає наявність таких основних блоків:

1. Блок аналізу структури моделі.
2. Блок попереднього аналізу.
3. Блок побудови моделей.
4. Блок прогнозування.
5. Блок обчислення оцінок.
6. Блок виведення результатів.

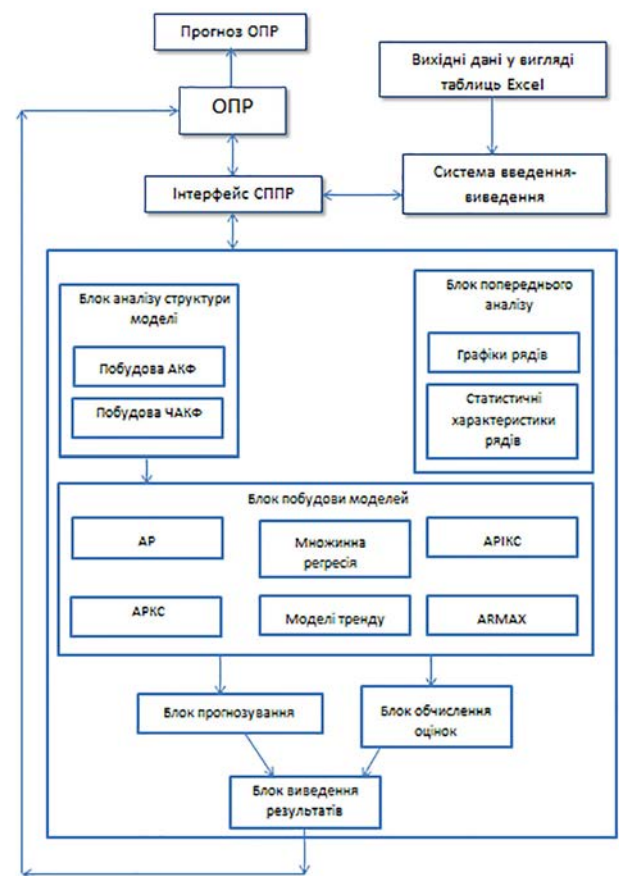


Рисунок 1. Блок-схема СППР для аналізу, моделювання та прогнозування економічних процесів

У програмному продукті реалізовано візуалізацію часових рядів у вигляді графіків, що дозволяє аналітику оцінити характер поведінки даних. В програмі здійснюється попереднє обчислення статистичних

характеристик аналізованого процесу. Для побудови структури моделі обчислюються значення АКФ та ЧАКФ для вихідних рядів. В результаті побудови моделі з'являється можливість прогнозування цільової змінної, використовуючи значення регресорів.

Коефіцієнти регресійних рівнянь знаходяться за допомогою методу МНК. З метою пришвидшення роботи методу використовується бібліотека MathNet.Numerics для платформи.NET, що дозволяє пришвидшити операції над матрицями з допомогою попереднього застосування до них різних типів розкладу.

Моделні значення на тестовій вибірці та прогноз на задану кількість кроків відображаються у вигляді графіків.

Програмний продукт містить достатню кількість налаштувань для керування процесом підбору структури моделі.

Опис програмного продукту

Вихідні дані завантажуються у вигляді сукупності часових рядів макроекономічних процесів. Дані ряди представляють цільову та пояснювальні змінні. Для зручності організації вхідних дані заносяться у таблицю Excel.

Для завантаження вихідних даних необхідно натиснути кнопку «Файл -> Завантажити дані» і обрати файл з таблицею вихідних даних.

Після завантаження даних про пояснювальні змінні, цільову змінну та їх опис можна переглянути у відповідній частині інтерфейсу.

Паралельно з відображенням даних відображаються графічне відображення ряду цільової змінної та набір характеристик цільового процесу.

Формат виведення результатів попереднього аналізу показано на рис. 2



Рисунок 2. Результати попереднього аналізу вихідного ряду цільової змінної

Використовуючи панель налаштування вибірки користувач має можливість виділити частину вибір-

ки, що використовується для побудови та тестування моделі.

Вибір моделі здійснюється шляхом зміни поточної вкладки на панелі налаштування моделей.

На цій же панелі можна змінити параметри структури моделі, такі як граничне значення ЧАКФ (кореляційної функції тощо).

Побудувати модель можна шляхом натискання кнопки «Побудова моделі».

Після побудови обраної моделі процесу виводиться модель у текстовому вигляді (див. рис. 3).

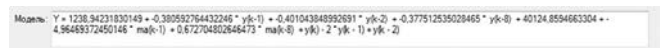


Рисунок 3. Панель відображення моделі

Побудова моделей ВВП України

Опишемо побудовані моделі та прогнози для ВВП України з використанням розробленого програмного продукту.

При побудові моделей розглянуто дані про об'єм ВВП України за часовому інтервалі від 1-го кварталу 2011 року до 1 кварталу 2016 року включно [4].

В результаті виконання програми отримано наступні моделі процесу:

АР:

$$y(k) = 310,0493 + 1,1114 \cdot y(k-1).$$

Множинна регресія:

$$y(k) = -25263,58 + 22,0301 \cdot X_1 + 2,56 \cdot 10^{-4} \cdot X_2.$$

АРКС:

$$y(k) = 310,049 + 1,111 \cdot y(k-1) + 573,375 - 0,618 \cdot ma(k-1)$$

Тренд 1-го порядку:

$$y(k) = 1570,332 + 2960,162 \cdot k.$$

Тренд 2-го порядку:

$$y(k) = 333,1323 + 45,21 \cdot k + 12,521 \cdot k^2.$$

АРИКС:

$$y(k) = 10,575 - 0,3621 \cdot \hat{y}(k-1) + 173,901 - 5,9 \cdot 10^{-4} \cdot ma(k-1) + y(k) - 2 \cdot y(k-1) + y(k-2)$$

ARMAX:

$$y(k) = 93,1141 + 0,4333 \cdot y(k-1) + 172,1121 - 0,0855 \cdot ma(k-1) - 89,4331 + 0,0711 \cdot X_1 + 9,0725 \cdot 10^{-7} \cdot X_2$$

Значення якісних статистик для навчальної та тестової вибірок наведено в табл. 1

Таблиця 1

Значення якісних статистик для навчальної та тестової вибірок при побудові моделі ВВП

| Метод | RMSE | | MAPE | | Коефіцієнт детермінації R ² | |
|--------------------|--------------------------|------------------------|-----------------------------|---------------------------|--|-------------------|
| | RMSE по навчальним даним | RMSE по тестовим даним | MAPE по навчальним даним, % | MAPE по тестовим даним, % | По навчальним даним | По тестовим даним |
| АР | 724,68 | 4874,21 | 3,33 | 2,17 | 0,997 | 0,842 |
| Множинна регресія | 1243,52 | 5232,19 | 3,72 | 7,68 | 0,996 | 0,385 |
| АРКС | 984,32 | 3421,42 | 3,20 | 2,71 | 0,997 | 0,842 |
| Тренд 1-го порядку | 3293,12 | 14076,23 | 14,33 | 19,26 | 0,945 | 0,713 |
| Тренд 2-го порядку | 606,19 | 516,29 | 3,01 | 0,67 | 0,998 | 0,862 |
| АРІКС | 634,66 | 1781,38 | 3,24 | 2,44 | 0,990 | 0,832 |
| ARMAX | 939,26 | 2244,45 | 3,30 | 3,20 | 0,999 | 0,842 |

Найкращою моделлю виявилась модель тренду 2-го порядку.

Побудова моделей індексу інфляції України

В даному підрозділі побудовано моделі та прогнози для індексу інфляції України з використанням розробленого програмного продукту.

При побудові моделей розглянуто дані про значення індексу інфляції України за часовому інтервалу від 1-го кварталу 2011 року до 1 кварталу 2016 року включно.

В результаті виконання програми отримано наступні моделі процесу:

АР:

$$y(k) = 131,4034 + 0,4052 \cdot y(k-2) - 0,3697 \cdot y(k-5) - 0,3431 \cdot y(k-8)$$

Множинна регресія:

$$y(k) = 104,1055 - 0,0004 \cdot X_1 - 2,143 \cdot 10^{-10} \cdot X_2 - 0,2811 \cdot X_3 - 0,0153 \cdot X_4$$

АРКС:

$$y(k) = 77,0521 + 0,4075 \cdot y(k-2) - 0,1788 \cdot y(k-8) + 1,0066$$

Тренд 1-го порядку:

$$y(k) = 102,4473 - 0,2033 \cdot k$$

Тренд 2-го порядку

$$y(k) = 102,3921 - 0,1842 \cdot k - 0,0011 \cdot k^2$$

АРІКС:

$$y(k) = 2,3271 - 0,3855 \cdot \hat{y}(k-2) - 0,4069 \cdot \hat{y}(k-8) + 0,726 + y(k) - 2 \cdot y(k-1) + y(k-2)$$

ARMAX:

$$y(k) = 23,2154 + 0,2221 \cdot y(k-2) + 0,0463 \cdot y(k-8) + 0,4018 - 1,372 \cdot 10^{-5} \cdot X_1 - 7,589 \cdot 10^{-13} \cdot X_2 - 9,9 \cdot 10^{-4} \cdot X_3 - 5,633 \cdot 10^{-5} \cdot X_4$$

Значення статистик для навчальної та тестової вибірок наведено в табл. 2

Найкращою моделлю виявилась модель АРІКС.

Висновки

В статті розглянуто питання аналізу, моделювання та прогнозування нестационарних економетричних процесів за допомогою регресійних моделей та побудови структури моделі з врахуванням кореляції екзогенних та ендогенних змінних.

Спроектовано архітектуру СППР для аналізу, моделювання та прогнозування економетричних процесів, що автоматизує процес підбору структури моделі, та реалізовано методи регресійного аналізу: АР, множинної регресії, АРМА, АРМАХ, АРІМА, моделей у вигляді тренду та їх варіацій.

Описані методи представлено у вигляді програмного модуля, організовано інтерфейс для програми, що використовує отримані модулі.

Програму апробовано на статистичних даних ВВП та коефіцієнта інфляції України (з використанням

Таблиця 2

Значення статистик для навчальної та тестової вибірок при побудові моделі ВВП

| Метод | RMSE | | MAPE | | Коефіцієнт детермінації R ² | |
|--------------------|--------------------------|---------------------|-----------------------------|---------------------------|--|-------------------|
| | RMSE по навчальним даним | RMSE тестових даних | MAPE по навчальним даним, % | MAPE по тестових даних, % | По навчальним даним | По тестовим даним |
| АР | 0,944 | 0,951 | 0,450 | 0,866 | 0,753 | 0,721 |
| Множинна регресія | 0,954 | 1,24 | 0,774 | 1,158 | 0,819 | 0,337 |
| АРКС | 1,252 | 0,992 | 1,012 | 0,863 | 0,812 | 0,743 |
| Тренд 1-го порядку | 0,996 | 1,977 | 0,810 | 1,888 | 0,653 | 0,531 |
| Тренд 2-го порядку | 0,996 | 1,067 | 0,812 | 1,977 | 0,894 | 0,912 |
| АРІКС | 0,771 | 0,899 | 0,621 | 0,763 | 0,953 | 0,993 |
| ARMAX | 1,052 | 1,298 | 1,929 | 2,167 | 0,883 | 0,721 |

інших показників у якості пояснювальних змінних); проведено аналіз та відбір найкращих моделей для побудови прогнозу.

При моделюванні ВВП України найкращою виявилась модель тренду 2-го порядку

$$y(k) = 333,1323 + 45,21 \cdot k + 12,521 \cdot k^2.$$

При моделюванні індексу інфляції найкращий результат отримано для моделі АРІКС

$$y(k) = 2,3271 - 0,3855 \cdot \hat{y}(k-2) - 0,4069 \cdot \hat{y}(k-8) + 0,726 + y(k) - 2 \cdot y(k-1) + y(k-2)$$

Література

1. Бідюк П. І. Аналіз часових рядів. / П. І. Бідюк, В. Д. Романенко, О. Л. Тимошук – Київ: «Політехніка», НТУУ «КПІ», 2013. – 600 с.
2. Бідюк П. І., Половцев О. В. Аналіз та моделювання економічних процесів перехідного періоду. – Київ: НТУУ «КПІ», 1999. – 230 с.
3. Бокс Д. Аналіз временных рядов прогноз и управление [Електронний ресурс]/ Бокс Д. – Режим доступу: <http://www.rusnauka.com//Economics/16434.doc.html>.
4. Державна служба статистики [Електронний ресурс] – Режим доступу: <http://www.ukrstat.gov.ua/>
5. Лукашин Ю. П. Адаптивные методы краткосрочного прогнозирования. – М.: Финансы и статистика, 2003. – 414 с.
6. Методичні вказівки до виконання практичних робіт з проектування інформаційних комп'ютерних систем [Текст] / Укл. Бідюк П. І. – К.: НТУУ «КПІ», 2008. – 87 с.