

Методика побудови математичних моделей динамічних процесів

Основи методики побудови моделей часових рядів запропоновані Боксом і Дженкінсом в роботі [1]. Модифікована авторами методика побудови математичної моделі динамічного процесу, складається з таких кроків [2]:

1. Аналіз процесу, для якого будується модель. Виконується на основі спеціальних літературних джерел, експертних оцінок протікання процесу, візуального дослідження вимірів вхідних і вихідних змінних, представлених часовими рядами або часовими перерізами, а також на основі іншої доступної інформації. Аналіз процесу спрямовується на розв'язання таких задач: визначення вимірності процесу; встановлення логічних зв'язків між змінними та можливостей їх спільного математичного опису; визначення кількості зовнішніх збурень та їх типу; встановлення можливості декомпозиції процесу на окремі простіші підпроцеси; розмежування рівнів процесів, що мають ієрархічну структуру; виявлення існуючих моделей процесу та досвіду його теоретичного чи експериментального дослідження, їх недоліків і переваг. Ігнорування цього етапу призводить до неможливості побудови моделі високого ступеня адекватності процесу та її придатності для розв'язання згаданих вище задач.

2. Попередня обробка експериментальних даних. Цей процес складається, як правило, з таких операцій: нормування та візуальна перевірка даних і, при необхідності, їх корегування, заповнення пропусків та зменшення викидів (аномальних значень); бутстреп аналіз наявних даних з метою збільшення об'ємів вибірок; заміна некоректних вимірів інтерпольованими або усередненими даними; формування перших або різниць вищих порядків; ортогональні перетворення та цифрова фільтрація даних з метою вилучення шумових складових.

3. Аналіз наявності нелінійностей. Наявність нелінійностей визначається за допомогою формальних тестів на основі множини статистичних критеріїв та візуального дослідження. Досвідченому фахівцю з моделювання візуальний аналіз дає можливість оперативно виявити наявність ділянок з лінійним або нелінійним трендом та значні викиди (імпульсів), які можуть суттєво впливати на якість моделі.

У випадку наявності більше трьох вибірок спостережень для одного і того ж процесу, можна рекомендувати простий тест для встановлення факту наявності не лінійності. Якщо статистика \hat{F} , яка розраховується як відношення (відхилення середніх значень від прямої регресії) до (відхилення значень $y(k)$ від вибірових середніх), з $v_1 = m - 2$ та $v_2 = n - m$ ступенями свободи не менше рівня значимості, то гіпотезу щодо лінійності необхідно відхилити, де n - загальне число вимірів, m - число вибірок даних.

Наявність нелінійності можна встановити також за допомогою вибірових *нелінійних кореляційних функцій* (НКФ), що розраховані за вибірками експериментальних даних.

Наявність нелінійного детермінованого тренду в процесі можна визначити шляхом оцінювання рівняння, що є поліномом порядку m відносно часу: $y(k) = a_0 + c_1k + c_2k^2 + \dots + c_mk^m$. Якщо хоча б один із коефіцієнтів c_i , $i = 1, \dots, m$ є статистично значущим, то гіпотеза стосовно відсутності тренду відхиляється. Наявність нелінійності стосовно регресора $x(k)$ можна встановити за допомогою відповідного полінома: $y(k) = a_0 + c_1k + c_2k^2 + \dots + c_mk^m + \varepsilon(k)$. Автоматично оцінює структуру математичної моделі *метод групового врахування аргументів* (МГВА), запропонований академіком Івахненком О. Г. Цей метод вже багаторазово застосовували до опису широкого класу процесів з метою оцінювання прогнозів та керування.

Загалом, задача встановлення наявності та визначення типу нелінійності залишається, на сьогодні, предметом активних досліджень.

4. Вибір (формування) структури моделей-кандидатів. Структура моделі включає *порядок моделі* (найвищий порядок рівнянь, що його утворюють); *вимірність* (число рівнянь моделі); *час запізнення* по входу (лаг) та його оцінка; *можливі нелінійності* та їх тип; *зовнішні збурення* та їх тип (детерміновані або випадкові; адитивні та мультиплікативні; імпульсні

та неперервні). Для формування структури моделі виконують такі дії: (1) обчислити та виконати аналіз кореляційної матриці для часових рядів залежної та незалежних змінних з метою визначення тих екзогенних змінних, які необхідно включити в модель; (2) обчислити автокореляційну (АКФ) та часткову автокореляційну функції (ЧАКФ) залежної змінної з метою визначення оцінки порядку авторегресійної частини моделі; (3) оцінити характеристики інших елементів структури математичної моделі.

5. Вибір методу (методів) оцінювання параметрів математичних моделей вибраних структур і оцінювання векторів їх параметрів. Найчастіше це метод найменших квадратів (МНК), метод максимальної правдоподібності (ММП) та їх модифікації (рекурсивні та нелінійні). Коефіцієнти моделей оцінюють, використовуючи принцип економії або збереження – *кількість коефіцієнтів, що оцінюються, не повинна перевищувати їх необхідне число* (“необхідність” можна визначити, наприклад, як необхідність збереження в моделі основних статистичних характеристик процесу). Також необхідно пам’ятати, що поведінку процесу необхідно *апроксимувати* за допомогою рівнянь, а не старатися описати її до найменших дрібниць. Модель, що оцінюється, повинна задовольняти принципу інверсії, тобто щоб за допомогою отриманого рівняння можна було б згенерувати початковий ряд, на основі якого оцінювались коефіцієнти. Тобто, модель повинна співпадати з досліджуванним процесом за основними характеристиками: середнє, дисперсія та коваріація.

6. Вибір кращої з оцінених моделей-кандидатів. Вибір здійснюється за допомогою множини статистичних критеріїв адекватності (якості) моделі. Аналізується якість моделей і виконується перевірка оцінених кандидатів на адекватність процесу, що складається з таких кроків:

- Візуальне дослідження графіка похибок моделі. На графіку не повинно бути значних викидів та довгих інтервалів, на яких похибка приймає великі значення (тобто довгих інтервалів суттєвої неадекватності).
- Аналіз наявності кореляції між значеннями похибок. Похибки моделі повинні бути некорельованими. Для аналізу обчислюється АКФ та ЧАКФ для ряду похибок. Ступінь корельованості визначається за допомогою Q -статистики (наприклад, Q -статистика вважається несуттєвою до рівня 10%) та статистики Дарбіна-Уотсона (DW) [2].
- Перевірка кількості ітерацій оцінювання. Для лінійної моделі 2-3 порядку оцінки параметрів повинні збігатися до усталених значень після 30-40 ітерацій алгоритму оцінювання. При значно більшій кількості ітерацій процес може бути нестационарним.
- Перевірка статистичної значущості оцінок параметрів моделі за *t-статистикою Стюдента* [2].
- Обчислення *коефіцієнту множинної детермінації R^2* , який показує рівень інформативності моделі по відношенню до інформативності вибірки даних, за якою вона була оцінена.
- Перевірка, що *сума квадратів похибок* для вибраної моделі повинна бути мінімальною у порівнянні з усіма іншими моделями.
- Розрахунок *інформаційних критерію Акайке та критерію Байеса-Шварца* [2], за якими вибирають ту модель, для якої критерії приймають найменші значення.
- Розрахунок *F-статистики Фішера* [2] дозволяє визначити адекватність моделі в цілому.

7. Висновки. Коректне застосування запропонованої методики забезпечує побудову адекватної математичної моделі процесу, якщо експериментальні дані відповідають вимогам представництва та інформативності. Перша вимога означає, що вибірка даних повинна охоплювати достатньо довгий проміжок часу, щоб повністю відображати поведінку того режиму функціонування процесу, для яких будуватиметься модель. Вимога інформативності означає, що вибірка повинна містити в собі об’єм інформації, достатній для оцінювання коефіцієнтів моделі. Розглянута методика побудови математичних моделей орієнтована на побудову моделей фінансово-економічних, соціальних та екологічних процесів, для яких, як правило, набагато складніше поставити експеримент та отримати інформативні експериментальні дані в доста-

тньому об'ємі. В наведеному вигляді ця методика також може бути успішно застосована до побудови моделей динаміки технічних систем та технологічних процесів.

Література. 1. Бокс Дж., Дженкінс Г. Анализ временных рядов. – Москва: Мир, 1974. – 406 с.
2. Половцев О.В. Системний підхід до моделювання, прогнозування та управління фінансово-економічними процесами / О.В. Половцев, П.І. Бідюк, Л.О. Коршевніюк, І.І. Семенчев. – Донецьк: "Східний видавничий дім", 2009. – 286 с.

Відомості про авторів

Коршевнюк Лев Олександрович

кандидат технічних наук, докторант Інституту прикладного системного аналізу Національного технічного університету України «Київський політехнічний інститут», доцент НТУУ «КПІ»

Терентьєв Олександр Миколайович

кандидат технічних наук, науковий співробітник Інституту прикладного системного аналізу Національного технічного університету України «Київський політехнічний інститут»

Бідюк Петро Іванович

доктор технічних наук, професор Інституту прикладного системного аналізу Національного технічного університету України «Київський політехнічний інститут»