

УДК 519.876.2:338.27(045)

Казік А. В., Стрижак В. О.
Національний авіаційний університет, Київ

ПОРІВНЯЛЬНИЙ АНАЛІЗ ВИКОРИСТАННЯ МЕТОДУ ГРУПОВОГО ВРАХУВАННЯ АРГУМЕНТІВ ТА МОДЕЛЕЙ РЕГРЕСІЙНОГО ТИПУ ДЛЯ МОДЕЛЮВАННЯ ЕКОНОМІЧНИХ ПРОЦЕСІВ (НА ПРИКЛАДІ МОДЕЛЮВАННЯ ПРОЦЕСУ ІНФЛЯЦІЇ)

В даній статті розглянуто сутність методу групового урахування аргументів (МГУА), і зроблено порівняльний аналіз прогнозуючих властивостей регресійних моделей та моделей методу групового врахування аргументів на прикладі задачі моделювання процесу інфляції.

Однією з найважливіших задач, які стоять перед кібернетикою, являється моделювання складних систем на основі спостереження їх взаємодії з навколишнім світом. Моделювання необхідно для того, щоб дізнатися структуру і функції складного об'єкту (задача ідентифікації) і визначити відповідні засоби активного впливу на нього (задача управління) або, якщо ми не маємо таких засобів в повній мірі, щоб дізнатися, чого можна чекати від даного об'єкту в майбутньому (задача прогнозування або екстраполяції).

Проблеми моделювання складних економічних систем можуть бути вирішені за допомогою різних дедуктивних логіко-математичних, індуктивних переборних чи регресійних методів. Дедуктивні та імітаційні методи мають переваги у випадку простих задач моделювання, якщо відома теорія об'єкта, який моделюється, і тому можлива побудова моделі, виходячи з фізично обґрунтованих принципів, застосовуючи знання людини щодо процесів в об'єкті. Але ці методи не в змозі дати задовільний результат для складних систем.

Індуктивний переборний МГУА має переваги при дослідженні складних об'єктів, які не мають певної теорії, зокрема для об'єктів з розмитими характеристиками. Алгоритми МГУА знаходять єдину оптимальну для кожної вибірки модель за допомогою повного перебору всіх можливих моделей-кандидатів і оцінюють її за зовнішнім критерієм на незалежній вибірці даних. Дослідження показують, що МГУА є найкращим для вирішення задач ідентифікації та прогнозування [1].

Ці алгоритми давно і ефективно використовуються для моделей однократного і багаторазового прогнозування економічних процесів. МГУА вирішує проблему обробки вибірок спостережень. Метою є отримання математичної моделі об'єкту (задача ідентифікації та розпізнавання образів) чи опису процесів, що відбудуться

для об'єкта у майбутньому (задача прогнозування). МГУА вирішує за допомогою процедури перебору, багатовимірну проблему оптимізації моделі:

$$\tilde{g} = \arg \min_{g \in G} CR(g), \quad CR(g) = f(P, S, \xi^2, T, V), \quad (1)$$

де: G – множина моделей, що розглядаються, CR – зовнішній критерій якості моделі g з цієї множини; P – кількість множин змінних; S – складність моделі; ξ^2 – дисперсія завад; T – число трансформацій вибірки даних; V – кількість видів референтної функції. Для базової референтної функції кожна множина змінних відповідає певній структурі моделі $P = S$. Задача трансформується до більш простої одновимірної:

$$CR(g) = f(S),$$

коли $\xi^2 = \text{const}$, $T = \text{const}$ та $V = \text{const}$.

В основі лежить переборна процедура, тобто послідовна перевірка моделей, що вибираються з множини моделей-кандидатів у відповідності до вибраного критерію. Більшість алгоритмів МГУА використовують поліноміальну базисну функцію. Загальний зв'язок між вхідними та вихідними змінними може бути виражений у вигляді функціонального ряду Вольтерра, дискретним аналогом якого є поліном Колмогорова-Габор:

$$y = a_0 + \sum_{i=1}^M a_i x_i + \sum_{i=1}^M \sum_{j=1}^M a_{ij} x_i x_j + \sum_{i=1}^M \sum_{j=1}^M \sum_{k=1}^M a_{ijk} x_i x_j x_k,$$

де $X(x_1, x_2, \dots, x_M)$ – вхідний вектор змінних;

$A(a_1, a_2, \dots, a_M)$ – вектор коефіцієнтів чи ваг.

Компонентами вхідного вектора X можуть бути незалежні змінні, функціональні форми чи кінцеві різницеві члени. Інші нелінійні базисні функції, наприклад диференційні, логістичні, ймовірнісні або гармонійні також можуть

бути застосовані для побудови моделі. Метод дозволяє одночасно отримати оптимальну структуру моделі та залежність вихідних параметрів від вибраних найбільш значимих вхідних параметрів системи. Традиційно в алгоритмах МГУА для формування зовнішніх критеріїв виконуються послідовні (в порядку слідування даних) розбивання вибірки або розбивання «по дисперсії» – на подібні або неподібні частини вибірки.

Теорія МГВА вирішує проблеми: довгострокового прогнозування; короткострокового передбачення процесів та подій; ідентифікації фізичних залежностей при точних даних; апроксимації багатфакторних процесів; екстраполяції фізичних полів; кластеризації вибірок даних; розпізнавання образів у випадках неперервних та дискретних змінних; діагностики та розпізнавання ймовірністними переборними алгоритмами; нормативного прогнозування векторних процесів; безмодельного прогнозування за допомогою комплексування аналогів; самоорганізації подвійно-багаторядних нейромереж з активними нейронами.

Для того щоб переконатися, який з методів моделювання краще описує економічні процеси, розглянемо приклад моделювання інфляції за допомогою МНК та МГУА. В поданому прикладі використовується багаторядний алгоритм МГУА, який використовує на кожному рівні переборної процедури однаковий частковий опис (правило ітерації). Даний алгоритм має використовуватися для обробки великої кількості змінних (до 1000). Тобто сутність даного методу полягає в тому, що вихідна змінна повинна бути визначена наперед експертом, що відповідає використанню так званих явних шаблонів (рис.1). На кожному ряді, вибірка даних розширюється на F змінних, які вираховані за F кращими моделями (рис. 1).

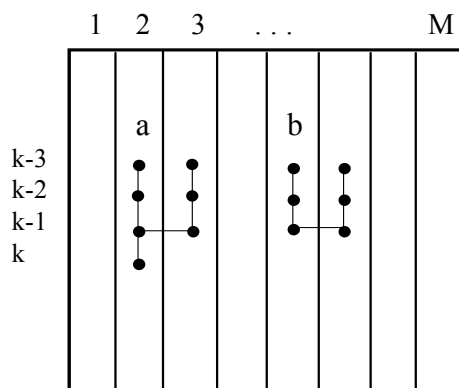


Рис. 1. Отримання умовних рівнянь у вибірці даних.

У Багаторядному Ітеративному алгоритмі правило ітерації залишається незмінним з одного ряду до наступного. Як показано на рис. 1, на першому ряді перебираються моделі, які можна отримати по інформації, яка міститься у будь-яких двох колонках змінних. На другому ряді – у чотирьох колонках, на третьому – у будь-яких восьми колонках і т.д. Правило зупинки таке: оптимальні моделі кожного ряду вибираються по мінімуму зовнішнього критерію.

Приклад. Для побудови моделі по даним, які взяті з [2] по описаному вище способу аналізу структур моделей, побудованих по багаторядному алгоритму МГУА, були вибрані найбільш вагомі аргументи, тобто:

x_1 – приватні накопичення (млн. грн.);

x_2 – число безробітних всього (відповідно до Муді);

x_3 – відсоткові ставки (відповідно до Муді);

x_4 – приватне споживання (млн. грн.)

x_5 – приватні доходи (млн. грн.)

x_6 – валовий національний продукт.

Вихідною величиною являється величина інфляції Y_2 .

По вибірці даних, яка приведена в [3] будувалися моделі залежності інфляції від поточних значень аргументів, причому три останні точки вибірки являли собою екзаменаційну послідовність, тобто для отримання моделі використовувалися лише перші 15 точок. Вибір такого варіанту розрахунків пов'язаний з різкими змінами характеру розвитку процесу. Ставилася така задача: чи можна, використовуючи дані, які відносяться до періоду монотонного розвитку інфляції, передбачити таку різку зміну? Іншими словами, слід було перевірити, чи закладено дані зміни в передісторію і чи можна їх виявити при моделюванні.

Модель, отримана методом найменших квадратів (МНК) включає в собі 6 вихідних аргументів і має такий вигляд:

$$V_{2,t} = 0,1646 + 0,00015x_{1t} - 0,00702x_{2t} - 0,00996x_{3t} + 0,0143x_{4t} + 0,00031x_{5t} - 0,00056x_{6t} \quad (2)$$

де V_2 – модельна оцінка вихідної змінної V_2 , x_{it} – значення i -го аргументу в t -ій точці ($i = 1, \dots, 6$; $t = 1, \dots, 18$).

Модель має такі характеристики якості: СКВ = 0,0297; $R = 0,0854$; $S = 157,9\%$ і такі статистичні характеристики:

$$R^2 = 0,864; R_y^2 = 0,762; F = 8,475; m1 = 6; m2 = 8; F_{0,05,6,8} = 3,58.$$

Значення t -статистик $t_1 (i = 1, \dots, 6)$:

$$t_0 = 2,9155; t_1 = 0,296; t_2 = 1,619; t_3 = 1,005; t_4 = 0,685; t_5 = 1,801; t_6 = 2,927; t_{(0,05,15)} = 1,753.$$

На рис. 2 приведені оцінки МЗ інфляції (МНК), отримані по моделі (2), на рис. 3 представлено графік їх зміни. Видно, що модель не зовсім відповідає реальній ситуації, причому найбільші помилки відносяться до трьох екзаменаційних точок, коли тенденція розвитку інфляції різко змінилася. Таким чином, МНК не дозволяє побудувати модель, яка б відображала системні закономірності і могла б застосовуватися для отримання надійного прогнозу, не дивлячись на те, що по статистичних оцінках вона значима,

адже $F > F_{(0,05,6,8)}, t_0, t_5, t_6 > t_{(0,05,15)}$.

Модель інфляції, отримана при цих же умовах методом МГУА (М2) має такий вигляд:

$$V_{2,t} = 0,0003x_{1t} + 0,0109x_{3t} - 0,00046x_{4t} + 0,00019x_{5t} - 0,0001156x_{6t}, \quad (3)$$

Характеризується такими показниками якості: СКВ = 0,00873; R=0,0194; S=35,9 % і має такі статистичні характеристики:

$$R^2 = 0,718; R_y^2 = 0,604; F = 6,796; F_{0,05,4,10} = 3,48.$$

Значення t -статистик $t_i (i = 1, \dots, 6)$:

$$t_1 = 0,631; t_2 = 1,252; t_4 = 0,816; t_5 = 1,299; t_6 = 2,646; t_{(0,05,15)} = 1,753.$$

t	Y ₂	V ₂	
		М3	М4
1	0.0145	0.025	0.0125
2	0.0357	0.024	0.0163
3	0.0267	0.018	0.0107
4	0.0081	0.016	0.0187
5	0.0160	0.014	0.0205
6	0.0092	0.009	0.0174
7	0.0117	0.016	0.0147
8	0.0116	0.014	0.0152
9	0.0140	0.017	0.0192

t	Y ₂	V ₂	
		М3	М4
10	0.0175	0.020	0.0227
11	0.0283	0.021	0.0307
12	0.0287	0.031	0.0305
13	0.0419	0.040	0.0382
14	0.0536	0.051	0.0451
15	0.0593	0.063	0.0618
16	0.0430	0.090	0.0483
17	0.0326	0.118	0.0421
18	0.0622	0.132	0.0736

Рис. 2. Оцінка процесу інфляції

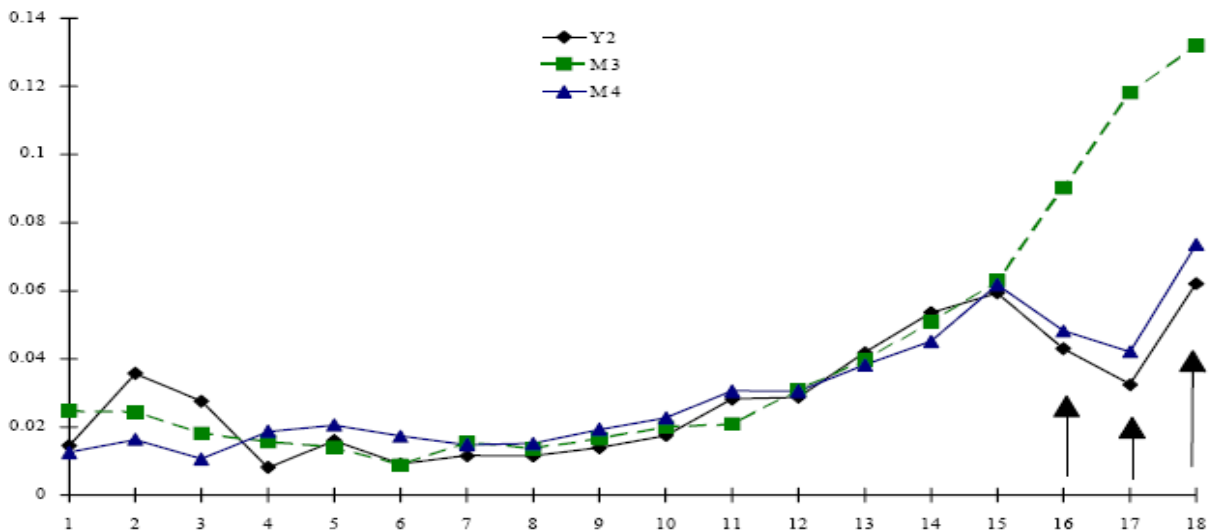


Рис. 3. Порівняння моделей інфляції, отриманих за МНК (М3) і за МГУА (М4) (стрілками позначені точки екзаменаційної послідовності).

На рис. 2 приведено оцінки М4 інфляції (МГУА), отримані по моделі 3, якість цієї моделі наглядно проілюстрована на рисунку 2, з якого видно, що вона чітко відображає зміни тенденцій процесу, що не досить очевидно з рис. 1, тобто з передісторії (до 16 точки) процесу.

Слід відмітити, що покращення прогнозу відбулося за рахунок спрощення прогнозованої моделі (в даному випадку за рахунок виключення з неї аргументу x_2), що характерно для використання в МГУА (ефект виключення «зайвих», не інформованих факторів).

Отже, регресійні моделі, навіть якщо вони по статистичних характеристиках являються значимими, недостатньо точно відображають прогнозовані тенденції. Натомість моделі, побудовані за алгоритмом МГУА, за своїми прогнозуючими властивостями значно ефективніші, ніж регресійні моделі в силу того, що за цими алгоритмами автоматично (за рахунок використання зовнішнього доповнення) відбираються аргументи (фактори), які є найбільш інформативними та значущими для моделювання даного об'єкту.

Саме тому доцільно звернути увагу на відмінності алгоритмів МГВА від інших алгоритмів структурної ідентифікації та селекції кращої регресії полягає у властивостях:

використання зовнішнього критерію, що базується на поділі вибірки даних та є адекватним до задачі побудови прогнозуючих моделей, за зменшенням потреб до об'єму первісної інформації; значно більшою різноманітністю генераторів структур: використання як у регресійних алгоритмах шляхів повного чи зменшеного перебору варіантів структур та застосування оригінальних багаторядних ітераційних процедур;

♦ більшим ступенем автоматизації – достатньо лише ввести первісні дані та вказати зовнішній критерій;

♦ автоматичною адаптацією складності оптимальної моделі та зовнішніх критеріїв до рівня завад у системі чи порушень – ефект завадостійкості обумовлює робастність підходу;

♦ запровадження принципу некінцевих рішень у процес поступового ускладнення моделей.

Список літератури

1. *Ивахненко А. Г., Мюллер Й. А.* Самоорганизация прогнозирующих моделей. – К.: Техніка, 1985. – 223 с.
2. *Бюлетень економічної кон'юнктури України.* – Київ: НДІ статистики Мінстату України. 1997. – випуск № 3. – 134 с.
3. *Степанко В.С., Копна Ю.В.* Опыт применения системы АСТРИД для моделирования экономических процессов по статистическим данным // Кибернетика и выч. Техника, 1999. – Вып. 117. – С. 23-29.

*Науковий керівник – Лецинський О.Л.,
канд. фіз.-мат. наук, доц.*