УДК 51–7:519.87

ПОБУДОВА РЕГРЕСІЙНИХ ЗАЛЕЖНОСТЕЙ ІЗ ВИКОРИСТАННЯМ ОРТОГОНАЛЬНИХ ПЕРЕТВОРЕНЬ ТА ТЕОРІЇ МОМЕНТІВ

Ярослав П'ЯНИЛО[1], Олег ХУДОЛІЙ[2], Андрій ВЛАСОВ[3], Андрій ДЕМІЧКОВСЬКИЙ[3]

[1] *Центр математичного моделювання Інституту прикладних проблем механіки і математики ім. Я. С. Підстригача НАН України, Львів, Україна*
[2] *Харківський національний педагогічний університет імені Г. С. Сковороди, Україна*
[3] *Львівський державний університет фізичної культури імені Івана Боберського, Україна*

Досліджено питання побудови регресійних та емпіричних залежностей для вивчення процесів, описаних часовими рядами дискретних даних.

Об'єктом дослідження є табульовані та експериментальні дані природних процесів.

Мета – використання ортогональних базисів та статистично-ймовірностних методів для оброблення цифрової інформації та побудови емпіричних залежностей між параметрами процесів, які вивчають.

Одним з основних постулатів класичного регресійного аналізу є припущення, що найкращі оцінки параметрів можна одержати використовуючи метод найменших квадратів. На практиці оцінки, одержані за допомогою цього методу, часто бувають недостатньо точними і містять великі похибки. Причиною цього може бути структура регресійної моделі. Найчастіше задачу побудови регресійної моделі формулюють так. Необхідно знайти функцію заданого класу, для якої функціонал [2]:

$$F(\alpha) = \sum_{i=1}^{n} \left[z_i (\alpha, X) - y_i \right]^2 \to \min . \tag{1}$$

У виразі (1) $z_i(\alpha, X)$ – значення функції, що апроксимує залежність, в *i*-й точці, y_i – відповідне значення емпіричної залежності, α – вектор параметрів, які треба знайти, X – вектор незалежних змінних. Одержану функцію $z_i(\alpha, X)$ називають (середньоквадратичною) регресійною моделлю. Метод її пошуку, який базується на застосуванні критерію (1), називають методом найменших квадратів.

Спектральні методи побудови апроксимувальної функції. Нехай многочлени $u_n(x)$ ортогональні на проміжку $[a,b]$ з ваговою функцію $\omega(x)$ і функція $\varphi(x)$ представлено ортогональним рядом за даними многочленами [1]:

$$\varphi(x) = \sum_{n=0}^{\infty} \frac{\varphi_n}{r_n} u_n(x), \quad \varphi_n = \int_a^b \omega(x) \varphi(x) u_n(x) dx ,$$

де r_n – нормувальний множник. Відомо [2,4], що $N+1$-й ортогональний многочлен має $N+1$ дійсний корінь, який належить проміжку ортогональності. Для обчислення узагальнених спектрів φ_n наявна оптимальна в класі L_2 квадратурна формула:

$$\varphi_n \approx \sum_{j=0}^{N} \rho_j^2 \, u_n(x_j) \, \varphi(x_j) , \tag{2}$$

де $x_j \; j = \overline{1, \; N+1}$ – корені многочлена $\boldsymbol{u_{N+1}(x)}$, тобто $u_{N+1}(x_j) = 0$ та

$$\rho_j^{-2} = \sum_{i=0}^{N} u_i^2 (x_j).$$

Достатньо поширеним ортогональним базисом є базис многочленів Якобі $P_n^{(\alpha,\beta)}(x)$, де $\alpha > -1$, $\beta > -1$ – вільні параметри, ***n*** – порядок многочлена, **$\omega(x) = (1-x)^{\alpha}(1+x)^{\beta}$** – вагова функція для многочленів Якобі. Подавати функції рядом Фур'є – Якобі можна декількома способами:

$$1)\ \varphi(x) = \sum_{n=0}^{\infty} \frac{\varphi_n}{r_n} P_n^{(\alpha,\beta)}(x);$$

$$2)\ \varphi(x) = \omega(x) \sum_{n=0}^{\infty} \frac{\varphi_n}{r_n} P_n^{(\alpha,\beta)}(x);$$

$$3)\ \varphi(x) = (1-x)^{\alpha} \sum_{n=0}^{\infty} \frac{\varphi_n}{r_n} P_n^{(\alpha,\beta)}(x);$$

$$4)\ \varphi(x) = (1+x)^{\beta} \sum_{n=0}^{\infty} \frac{\varphi_n}{r_n} P_n^{(\alpha,\beta)}(x).$$

Вибором вигляду подання функції, запропонованим вище способом, можна врахувати поведінку функції на одному або обидвох проміжках визначеності. Це дасть можливість пришвидшити збіжність ряду, що сприятиме зменшенню нагромадженню похибки під час сумування.

Моменти в обробленні цифрової інформації та їх зв'язок із спектральними методами. Застосування ортогональних розкладів тісно пов'язане зі спектральни-ми моментами в статистичноймовірностних методах оброблення цифрової інформації. Таке поєднання доцільно використовувати для зменшення похибки у разі наближення до початкового значення аргументу функції розподілу, де вона виникає при використанні замість функції розподілу моментів для спрощення аналізу [2, 3]. Введення поняття спектральних моментів розподілу дасть змогу спростити розрахунки під час підсумовування декількох похибок за їх квантильного оцінцювання.

Початковим моментом *k*-го порядку випадкової величини *X* називається математичне сподівання *k*-го степеня *X*, тобто:

$$M_k(X) = \int_a^b x^k f(x)\,dx = a_k\,.$$

Центральним моментом *k*-го порядку називають математичне сподівання *k*-го степеня відхилення випадкової величини від її математичного сподівання:

$$\mu_k = \int_a^b (x - a_1)^k f(x)\, dx .$$

Центральні моменти використовують для кількісної характеристики розподілу випадкової величини. Коефіцієнтом асиметрії $\boldsymbol{C_s}$ називається відношення центрального моменту третього порядку до куба середнього квадратичного відхилення $C_s = \mu_3 \mu_2^{-3/2}$. Коефіцієнт ексцесу випадкових величин обчислюють за формулою $E_s = \mu_4 \mu_2^{-2} - 3$.

Використання ортогональних розкладів дає можливість розв'язати декілька задач оброблення цифрової інформації – апроксимацію сигналу та фільтрацію від адитивного шуму. Разом з тим, використовуючи зв'язок між коефіцієнтами ортогональних розкладів та статистичними параметрами, за відомими коефіцієнтами можна оцінити їх вірогідність та інші ймовірносно-статистичні параметри.

Теорема. Нехай функція $\varphi(x)$ розкладається в ряд за многочленами Якобі. Тоді наявна рівність:

$$\mu_n = \frac{1}{\eta_{nn}(\alpha,\beta)} \left(\varphi_n - \sum_{j=0}^{n-1} \eta_{jn}(\alpha,\beta) \mu_j \right).$$

Для побудови двовимірних регресійних та емпіричних залежностей, як правило, використовують поліноми різного степеня. Вони не завжди є ефективними. Здебільшого більш доцільно застосовувати дробово-раціональні вирази. Продемонструємо на побудові емпіричної формули для обчислення коефіцієнта стиску водню. Основою для цього є табульовані експериментальні дані. Результатом побудови є наступна формула, яка працює у таких проміжках: тиск $p \in [1,80]$ атм та температура $t \in [0,300]$ у градусах Цельсія.

За аналогією для газу метану залежність коефіцієнта стиску водню від температури та тиску будували у вигляді таких залежностей:

$$z = 1/(1+fp),\ z = (x_1 - x_2\, T/273)\, ((x_3\, p\, (T/273)^{y1} + x_4) ,$$

де $p(x)$ – вимірюється в атмосферах, а x_i, y_i – невідомі коефіцієнти, які знаходь на основі табличних даних. За отриманими результатами проведено числовий експеримент.

Висновок. У статті доведемо, що використання ортогональних базисів та методів статистично-ймовірностного оброблення інформації дає можливість розв'язувати складні задачі за умови значної невизначеності. Так, побудова кореляційних залежностей у базисах ортогональних многочленів дає можливість апроксимації та згладження експериментальних даних та на базі теорії моментів дає змогу знайти певні статистично-ймовірностні параметри процесів, що вивчають.

Список використаних джерел

1. П'янило Я. Д. Проекційно-ітераційні методи розв'язування прямих та обернених задач переносу / П'янило Я. Д. – Львів : Сплайн, 2011. – 248 с.
2. Авраменко В. І. Теорія ймовірностей і математична статистика: навч. посіб. / Авраменко В. І., Карімов І. К. – 2-ге вид., перероб. і доп. – Дніпродзержинськ: ДДТУ, 2013. – 245 с.
3. Системний підхід і математичне моделювання біологічних та природних об'єктів і процесів / А. Власов, А. Демічковський, О. Іващенко, А. Лопатьєв, М. Пітин, Я. П'янило, О. Худолій // Фізико-математичне моделювання та інформаційні технології. – 2016. – № 23. – С. 17–28.