

## РЕГУЛЯРИЗАЦІЯ ПОЛІНОМІАЛЬНИХ ФУНКЦІЙ БЕЗ'Є-БЕРНШТЕЙНА ТА ЇЇ ЗАСТОСУВАННЯ В ЗАДАЧІ ІДЕНТИФІКАЦІЇ НЕЛІНІЙНИХ СИСТЕМ

О.Ю. МИТНИК

Розглядається задача визначення оптимального розподілу базових контрольних точок, на основі якого обернене відображення Кастельжо забезпечує найкраще наближення поліноміальних функцій Без'є–Бернштейна кривими та поверхнями Без'є. Вводяться критерії оптимальності системи базових контрольних точок, і для них наводяться відповідні оцінки. Запропоновано еволюційний метод побудови оптимального оберненого відображення Кастельжо в задачі ідентифікації нелінійних динамічних систем за допомогою нечітких нейронних моделей.

### ВСТУП

Технологія створення векторних зображень за допомогою кривих винайдена на початку 60-х років П'єром Без'є, інженером фірми Renault, і Полем де Кастельжо (Кастельє), інженером компанії Citroen. Оскільки вони обидва працювали над створенням креслень нових автомобілів, всі їх нароби до певного часу не оприлюднювалися. Робота Поля де Кастельжо була завершена першою, але так жодного разу і не публікувалася. П'єру Без'є пощастило більше. Його іменем названо нову технологію, хоча математичні вирази, відповідно до яких виконуються перетворення Без'є, винайдені, в основному, паном де Кастельжо [1].

Математична теорія барицентричних координат і алгоритм Кастельжо, які широко використовуються при побудові кривих та поверхонь Без'є в області автоматизованого проектування трьохвимірних систем, застосовуються і при побудові нечітких нейронних моделей для нелінійних динамічних систем.

### НЕЧІТКА ЛОГІКА В НЕЙРОННИХ СИСТЕМАХ

Оскільки нечітка логіка узагальнює поняття характеристичної функції булевої логіки, то за допомогою лінгвістичного терму  $A$ , що визначається нечіткою множиною з функцією належності  $\mu_A(x) \in [0, 1]$ , можна представити нечітке твердження щодо  $x$ .

Класична нечітка логіка оперує з логічними функціями **and**, **or**, **if(.) then(.)** над нечіткими змінними. Але якщо замість звичайних операторів **min** та **max** застосовувати алгебраїчні оператори добутку та суми, то нечіткі моделі стають більш гнучкими і забезпечують відповідність між нейронними системами і нечіткою логікою, наприклад, якщо функції належності задаються у вигляді радіальних базисних функцій або  $B$ -сплайнів [2]. Такі моде-

лі також полегшують формальне сприйняття нечіткої системи і відокремлюються від лінгвістичного апарату. В цьому випадку важливий такий результат [3].

**Теорема.** Якщо алгебраїчні оператори використовуються для реалізації функцій нечіткої логіки, дійсні значення входів представляються через функції належності, для дефазифікації виходу використовується метод середнього значення, то вихід узагальненої нечіткої нейронної системи буде мати вигляд

$$y(\mathbf{x}) = \sum_{i=1}^p \omega_i \mu_{A^i}(\mathbf{x}), \quad \mathbf{x} \in \mathfrak{R}^n, \quad (1)$$

де  $\mu_{A^i}(\mathbf{x})$  — функції належності вектору входу  $\mathbf{x}$ , які задовольняють  $\sum_{j=1}^p \mu_{A^j}(\mathbf{x}) = 1$ ; вагові коефіцієнти  $\omega_i = \sum_{j=1}^p c_{ij} y_j^c$ , де  $y_j^c$  — центр функції належності виходу  $\mu_{B^j}(y)$ ;  $c_{ij}$  — вага правила: «якщо  $\mathbf{x} \in A^i$ , то  $y \in B^j$ », при цьому  $\sum_{j=1}^p c_{ij} = 1$ ,  $c_{ij} \in [0, 1]$ ;  $A^i, B^j$  — лінгвістичні терми, визначені відповідно на вхідних та вихідних змінних.

### Висновок з теореми

Теорема показує, що нечітка нейронна нелінійна система може бути представлена у вигляді лінійної у параметрах структури, яка дозволяє застосовувати швидкі та прості процедури навчання. При цьому постає питання про вибір функцій належності.

### ПОЛНОМИ БЕРНШТЕЙНА ЯК ФУНКЦІЇ НАЛЕЖНОСТІ

Розглянемо алгоритм побудови нечітких нейронних моделей для нелінійних динамічних систем, запропонований К. Харрісом [3, 4]. Основа цього підходу полягає у використанні поліноміальних функцій Без'є–Бернштейна однієї та двох змінних як базисних для представлення виходу нелінійної системи у вигляді розкладу Габора–Колмогорова

$$f(\mathbf{x}) = f_0 + \sum_{k=1}^n f_k(x_k) + \sum_{p=1}^n \sum_{q=p+1}^{n-1} f_{pq}(x_p, x_q) + e(\mathbf{x}), \quad (2)$$

де вектор  $\mathbf{x} = [x_1, \dots, x_n]^T \in X^n \subset \mathfrak{R}^n$  — вхід нелінійної системи.

Для формування базисних поліноміальних функцій Без'є–Бернштейна однієї та двох змінних використовуються лінійні комбінації відповідних поліномів Бернштейна від векторів барицентричних координат  $s, \mathbf{u}$ :

$$f_k(x_k) = \sum_{j=0}^d \omega_j B_j^{(d)}[s(x_k)], \quad (3)$$

$$f_{pq}(x_{pq}) = \sum_{i+j+k=d} \omega_{ijk} B_{ijk}^{(d)}[\mathbf{u}(x_{pq})], \quad (4)$$

де  $\mathbf{x}_{pq} = [x_p, x_q]^T \in X^2 \subset \mathfrak{R}^2$ . При цьому кожна функція Без'є–Бернштейна відповідає виходу нечіткої нейронної підсистеми (1), а поліноми Бернштейна

на виступають у ролі функцій належності для представлення її входів. Таким чином, вся нелінійна динамічна система (2) розглядається як суперпозиція таких нейронних підсистем.

Барицентричні координати, в свою чергу, обчислюються за допомогою оберненого алгоритму Кастельжо, який реалізує відображення

$$\Psi_k(\Lambda_k): x_k \rightarrow s_k \in [0, 1] \text{ та}$$

$$\Psi_{pq}(\Lambda_{pq}): \mathbf{x}_{pq} \rightarrow \mathbf{u}_{pq} = [u_p, u_q]^T \in \Delta \{u_p \geq 0, u_q \geq 0, u_p + u_q \leq 1\}.$$

Відображення будуються на основі наперед заданих множин базових контрольних точок  $\Lambda_k, \Lambda_{pq}$  із використанням ітераційного чисельного методу зворотного розповсюдження помилки або ж за допомогою швидкого прямого аналітичного алгоритму [5]. Після обчислення барицентричних координат знаходяться оцінки вагових коефіцієнтів  $\omega_j, \omega_{ijk}$  відповідних рівнянь регресії (2)–(4) двохетапним методом найменших квадратів (МНК) або ортогональним методом найменших квадратів (ОМНК) [6]. Тобто відбувається навчання нейронних підсистем.

Зауважимо також, що задачі (3) і (4) можна інтерпретувати як задачі наближення поліноміальних функцій однієї та двох змінних відповідно кривими та поверхнями Без'є [7]. Надалі, для простоти викладення, не порушуючи загальності, розглянемо задачу наближення поліноміальної функції однієї змінної кривою Без'є, тобто задачу (3).

### ПОСТАНОВКА ЗАДАЧІ РЕГУЛЯРИЗАЦІЇ ОБЕРНЕНОГО ВІДОБРАЖЕННЯ КАСТЕЛЬЖО

Залишається визначити систему базових контрольних точок (СБКТ)  $\Lambda^*$ , яка б дозволила найкращим чином побудувати наближення поліноміальної функції однієї змінної кривою Без'є, тобто задовольняла певним критеріям оптимальності. Тож під регуляризацією оберненого відображення Кастельжо будемо розуміти визначення саме такої оптимальної СБКТ, що забезпечує найкраще наближення поліноміальної функції кривою Без'є за певним критерієм.

Введемо означення.

*A-оптимальна* СБКТ така, що мінімізує слід коваріаційної матриці оцінок вагових коефіцієнтів

$$\Lambda_{A\text{-опт}}^* : \text{tr}(\text{cov} \hat{\Omega}) \rightarrow \min, \quad (5)$$

де  $\Omega = [\omega_0, \dots, \omega_d]^T$  — вектор вагових коефіцієнтів. Мінімізація сліду коваріаційної матриці відповідає мінімізації середньої дисперсії оцінок коефіцієнтів, що має зміст суми квадратів головних піввісей еліпсоїду розсіювання оцінок [7, 8].

*D-оптимальна* СБКТ така, що мінімізує значення визначника відповідної коваріаційної матриці

$$\Lambda_{D\text{-опт}}^* : \det(\text{cov} \hat{\Omega}) \rightarrow \min. \quad (6)$$

Такий визначник має зміст об'єму еліпсоїду розсіювання оцінок, тобто узагальненої дисперсії оцінок невідомих вагових коефіцієнтів рівняння регресії.

*G-оптимальна* СБКТ така, що мінімізує максимальну дисперсію спрогнозованих значень регресії

$$\Lambda_{G\text{-опт}}^* : \max_{x \in X} \{\hat{D}(x)\} \rightarrow \min, \quad (7)$$

де дисперсія спрогнозованих значень  $\hat{D}(x) = \mathbf{b}(x) \text{cov} \hat{\mathbf{\Omega}} \mathbf{b}^T(x)$ ; вектор  $\mathbf{b}(x) = [B_0^{(d)}(s(x|\Lambda)), \dots, B_d^{(d)}(s(x|\Lambda))]^T$ ,  $s(x|\Lambda)$  — барицентричні координати, визначені за допомогою оберненого відображення Кастельжо на основі СБКТ  $\Lambda$  для  $x \in X \subset \mathfrak{R}$ .

*I-оптимальна* СБКТ така, що мінімізує узагальнену дисперсію спрогнозованих значень регресії [9]

$$\Lambda_{I\text{-опт}}^* : J = \int_X \hat{D}(x) d\mu(x) \rightarrow \min, \quad (8)$$

де  $\mu$  — рівномірна міра на просторі входу  $X$ .

Припускаючи, що похибки моделі (2) попарно некорельовані і мають однакову скінченну дисперсію, тобто  $e(\mathbf{x}) \sim N(0, \sigma^2)$ , коваріаційна матриця матиме вигляд  $\text{cov} \hat{\mathbf{\Omega}} = \sigma^2 (\mathbf{A}^T \mathbf{A})^{-1}$ , де матриця  $\mathbf{A} = \{B_j^{(d)}(s(x_i|\Lambda))\}_{ij}$ ,  $j=0, \dots, d$ ;  $i=1, \dots, N$ ,  $N$  — розмір вхідної вибірки. Надалі будемо вважати  $\sigma^2 = 1$ , що не обмежує загальності задач (5–8).

Застосування теорії побудови оптимальних планів у нашому випадку обмежується пасивною ідентифікацією, оскільки розглядаються нелінійні динамічні системи, в яких неможливо ставити експеримент і задавати значення входів, наприклад, економічні системи. Тож мова йде про оптимальний розподіл параметрів моделі, де в ролі параметрів виступають базові контрольні точки. Як правило, *A*, *D*-оптимальність використовується для побудови моделей, де важлива точність і якість оцінок коефіцієнтів регресії. У той же час *G*, *I*-оптимальність ефективна для задач пошуку екстремуму поверхні відгуку в заданій області. Важливо визначити, які з наведених оптимізаційних задач ефективніші для визначення структури моделі, що забезпечує найменшу похибку при прогнозуванні.

## ДЕЯКІ ОЦІНКИ ОПТИМАЛЬНОСТЕЙ

### Оцінка *I*-оптимальності

Функціонал (8) легко спрощується до вигляду  $J = \text{tr}(\mathbf{M} \text{cov} \hat{\mathbf{\Omega}})$ , де матриця моментів вхідного простору  $\mathbf{M} = \int_X \mathbf{b}^T(x) \mathbf{b}(x) d\mu(x)$ . Дійсно,

$$J = \text{tr}(\mathbf{J} \equiv J) = \int_X \text{tr}(\mathbf{b}(x) \text{cov} \hat{\mathbf{\Omega}} \mathbf{b}^T(x)) d\mu(x) =$$

$$= \int_{\mathbf{X}} \text{tr}(\mathbf{b}^T(x)\mathbf{b}(x)\text{cov}\hat{\mathbf{\Omega}})d\mu(x) = \text{tr}\left(\int_{\mathbf{X}} \mathbf{b}^T(x)\mathbf{b}(x)d\mu(x)\text{cov}\hat{\mathbf{\Omega}}\right),$$

оскільки для будь-яких квадратних матриць  $\mathbf{A}, \mathbf{B}$  виконується рівність  $\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$ .

Цей результат значно полегшує обчислення показника  $I$ -оптимальності, оскільки у багатьох випадках матрицю моментів можна визначити наперед. Крім того, очевидно, що у випадку ортогональності векторів  $\mathbf{b}(x)$  матриця моментів є одиничною матрицею і  $I$ -оптимальність зводиться до  $A$ -оптимальності.

**Теорема 1. Оцінка поліномів Бернштейна.** Будь-який поліном Бернштейна однієї змінної порядку  $d$  допускає оцінку зверху:

$$\forall s \in (0, 1), \forall j: B_j^{(d)}(s) \leq v(s) = \frac{1}{\sqrt{2\pi ds(1-s)}},$$

причому  $B_j^{(d)}(s) \approx v(s)$  при  $s = j/d$ .

**Доведення.** Поліноми Бернштейна однієї змінної порядку  $d$  мають вигляд

$$B_j^{(d)}(s) = \binom{d}{j} s^j (1-s)^{(d-j)},$$

і до них, як і для будь-яких біноміальних розподілів, можна застосувати нормальне наближення. Для довільного симетричного біноміального розподілу справедливі такі оцінки його членів, які перенумеровані у залежності від їх відстані до максимальної ймовірності [10]:

$$a_k = \binom{d}{m+k} s^{m+k} (1-s)^{(d-m-k)} \sim hr(kh), \quad k = -m, \dots, 0, \dots, d-m, \quad j = m+k,$$

де  $m = ds + \delta$ ,  $s - 1 < \delta \leq s$  — індекс максимальної ймовірності розподілу;

$h = \frac{1}{\sqrt{ds(1-s)}}$ ,  $r(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$  — густина нормального розподілу. Точність такого наближення детально розглядається в роботі [10]. Очевидно, що

$$\forall s \in (0, 1), \forall j: B_j^{(d)}(s) = a_k \leq a_0 \sim hr(h \cdot 0) = \frac{1}{\sqrt{ds(1-s)}} \frac{1}{\sqrt{2\pi}} = v(s).$$

Це доводить оцінку зверху для поліномів Бернштейна. Крім того, оцінка максимальної ймовірності  $a_0 \sim v(s)$  є одночасно і оцінкою максимуму кожного поліному Бернштейна при  $s \rightarrow j/d$ .

$$\forall s \in [0, 1], \forall j: B_j^{(d)}(s) \leq B_j^{(d)}\left(s_j = \frac{j}{d}\right) \sim v(s_j) = \frac{1}{\sqrt{2\pi ds_j(1-s_j)}},$$

причому  $\forall k: B_j^{(d)}(s_j) \geq B_k^{(d)}(s_j)$  (рис. 1).

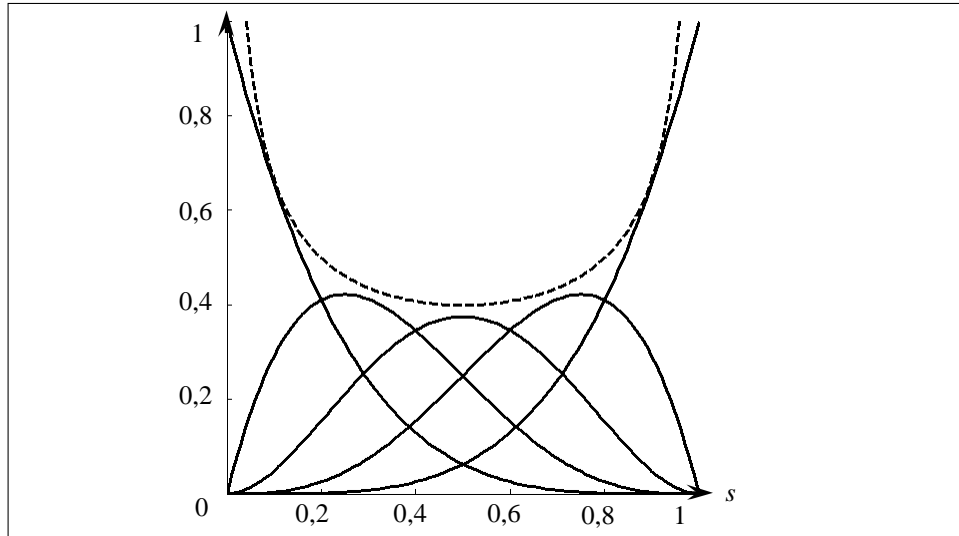


Рис. 1. Оцінка поліномів Бернштейна 4-го порядку зверху: —  $B_j^{(d)}(s)$ ; ----  $v(s)$

**Теорема 2. Оцінка сліду інформаційної матриці.** Слід інформаційної матриці від системи поліномів Бернштейна допускає наближення:

$$\text{tr}(\mathbf{A}^T \mathbf{A}) = \sum_{j=0}^d \sum_{i=1}^N [B_j^{(d)}(s(x_i | \Lambda))]^2 \approx \sum_{i=1}^N \frac{1}{2\sqrt{\pi d s_i (1-s_i)}}. \quad (9)$$

**Доведення.** Враховуючи наведені вище оцінки для поліномів Бернштейна,

$$\begin{aligned} \sum_{j=0}^d [B_j^{(d)}(s)]^2 &\sim \sum_{k=-m}^{d-m} [hr(kh)]^2 \stackrel{\langle k \rightarrow \xi \rangle}{\approx} \frac{1}{2\pi} h^2 \int_{-\infty}^{\infty} e^{-\xi^2 h^2} d\xi \stackrel{\langle x = \sqrt{2}\xi h \rangle}{=} \\ &= \frac{1}{2\pi} \frac{h}{\sqrt{2}} \int_{-\infty}^{\infty} e^{-x^2/2} dx = \frac{h}{2\sqrt{\pi}} = \frac{1}{2\sqrt{\pi d s_i (1-s_i)}}. \end{aligned}$$

Відповідно маємо для сліду інформаційної матриці таку оцінку (рис. 2):

$$\text{tr}(\mathbf{A}^T \mathbf{A}) = \sum_{i=1}^N \sum_{j=0}^d [B_j^{(d)}(s_i)]^2 \approx \sum_{i=1}^N \frac{1}{2\sqrt{\pi d s_i (1-s_i)}}.$$

### Висновки з теореми

1. Очевидно, що чим ближчі барицентричні координати до центру (до 0,5), тим менше значення сліду інформаційної матриці.

2. При рівномірному розподілі барицентричної координати на  $[0, 1]$  можна перейти від граничних сум до визначеного інтегралу, який легко обчислити через тригонометричну заміну

$$\text{tr}(\mathbf{A}^T \mathbf{A}) \sim \frac{N}{2\sqrt{\pi d}} \int_0^1 \frac{ds}{\sqrt{s(1-s)}} = \frac{N}{2\sqrt{\pi d}} \pi = \frac{N}{2} \sqrt{\frac{\pi}{d}}.$$

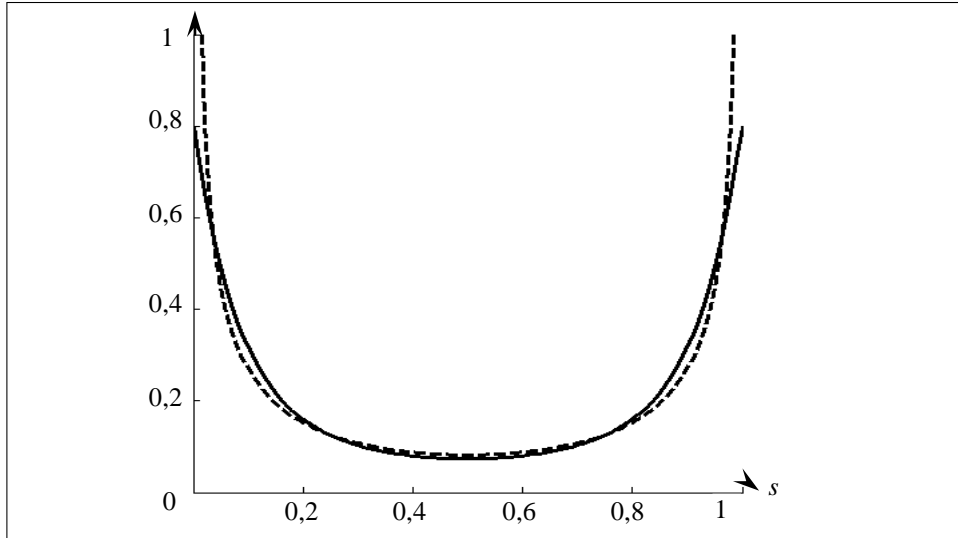


Рис. 2. Оцінка сліду інформаційної матриці ( $N = 1$ ): —  $\text{tr}(\mathbf{A}^T \mathbf{A})$ ; - - - оцінка  $\text{tr}(\mathbf{A}^T \mathbf{A})$

### Оцінка $A, D$ -оптимальності

Зазначимо, що коваріаційна матриця — це симетрична, позитивно визначена матриця, власні числа якої дійсні та невід’ємні, а їх сума дорівнює сліду матриці. За допомогою нерівностей між середнім арифметичним та середнім геометричним можна отримати оцінку знизу для сліду коваріаційної матриці

$$\begin{aligned} \text{tr}(\text{cov } \hat{\Omega}) &= \text{tr}(\mathbf{A}^T \mathbf{A})^{-1} = \sum_{i=0}^d \frac{1}{\lambda_i} \geq (d+1)^{d+1} \sqrt{\prod_{i=0}^d \frac{1}{\lambda_i}} \geq \\ &\geq (d+1)^2 \left[ \sum_{i=0}^d \lambda_i \right]^{-1} = (d+1)^2 [\text{tr}(\mathbf{A}^T \mathbf{A})]^{-1}, \end{aligned} \quad (10)$$

де  $\lambda_i^{-1}$  — власні числа коваріаційної матриці;  $\lambda_i$  — власні числа інформаційної матриці.

Аналогічно визначник коваріаційної матриці, який дорівнює добутку власних чисел, також допускає оцінку знизу

$$\begin{aligned} \det(\text{cov } \hat{\Omega}) &= \det(\mathbf{A}^T \mathbf{A})^{-1} = \prod_{i=0}^d \frac{1}{\lambda_i} \geq (d+1)^{d+1} \left[ \sum_{i=0}^d \lambda_i \right]^{-d-1} = \\ &= (d+1)^{d+1} [\text{tr}(\mathbf{A}^T \mathbf{A})]^{-d-1} \approx (d+1)^{d+1} \left[ \sum_{i=1}^N \frac{1}{2\sqrt{\pi d s_i (1-s_i)}} \right]^{-d-1}. \end{aligned} \quad (11)$$

Очевидно, що ті розподіли барицентричних координат, для яких в нерівностях (10), (11) мають місце рівності, тобто  $A, D$ -оптимальні розподіли, співпадають і можуть визначатися умовою [8]

$$\mathbf{A}^T \mathbf{A} \approx (d+1)^{-1} \sum_{i=1}^N \frac{1}{2\sqrt{\pi d s_i (1-s_i)}} \mathbf{I}_{d+1}.$$

Але легко показати, що такої матриці не існує, адже поліноми Бернштейна, з яких складається матриця  $\mathbf{A}$ , додатні. Тож можна зробити висновок про те, що оптимальність у нашому випадку це паритет двох критеріїв:

1. Кожна барицентрична координата знаходиться якомога ближче до центру області визначення.

2. Барицентричні координати розташовані якомога далі одна від одної з метою покращення обумовленості інформаційної матриці.

Зауважимо, що наведені вище оцінки знизу для  $A$ ,  $D$ -оптимальностей не слід використовувати як критерії мінімізації самих оптимальностей, тобто мінімізація сліду чи визначника коваріаційної матриці не еквівалентна максимізації сліду інформаційної матриці.

Зазначимо також, що для неперервних планів справедливою є теорема еквівалентності [7], згідно якої неперервний  $D$ -оптимальний план є також і  $G$ -оптимальним.

## ЗАСТОСУВАННЯ ГЕНЕТИЧНОГО АЛГОРИТМУ ДЛЯ ВИЗНАЧЕННЯ ОПТИМАЛЬНОГО РОЗПОДІЛУ БАЗОВИХ КОНТРОЛЬНИХ ТОЧОК

Відомі приклади успішного застосування генетичних алгоритмів (ГА) у задачах оптимального планування експерименту [11], а також у деяких оптимізаційних задачах [12]. Для розв'язку оптимізаційних задач (5)–(8) пропонується використовувати ГА, а саме його реалізацію з популяціями, які не перекриваються [13]. Зазначимо, що в цьому випадку задачу (1), (4) можна віднести до класу задач навчання нейронних систем еволюційними методами.

Об'єкти ГА в нашому випадку мають такий зміст. В якості *гену* виступає базова контрольна точка. Набір *алелей* характеризує множину можливих значень контрольних точок. Кожна *особина* (*хромосома*) складається з набору генів, який відповідає системі контрольних точок. *Мірою придатності* кожної особини, яка виражає той чи інший критерій оптимальності, виступає цільова функція вигляду

$$F_{A-\text{опт}}(\Lambda) \sim -\text{tr}(\text{cov} \hat{\Omega}), \quad F_{D-\text{опт}}(\Lambda) \sim -\det(\text{cov} \hat{\Omega}),$$

$$F_{G-\text{опт}}(\Lambda) \sim -\max \{\hat{D}(x)\}, \quad F_{I-\text{опт}}(\Lambda) \sim -\text{tr}(\mathbf{M} \text{cov} \hat{\Omega}).$$

В результаті роботи ГА на тестових моделях отримані оптимальні розподіли базових контрольних точок для різних розподілів вектору входу. При цьому точність апроксимації і короткострокового прогнозування виходу системи за допомогою моделі К. Харріса зросла в декілька разів, оскільки наведені вище оптимізації зменшують також і значення середньоквадратичної похибки.

Окремою темою для досліджень в цьому напрямку є питання про збіжність ГА. Хоч ГА іноді знаходить гарні розв'язки для складних задач, проте теоретичних результатів щодо збіжних властивостей і умов збіжності дуже мало. Одна із спроб дослідити ГА на збіжність — створення відповідної мо-



делі ланцюга Маркова та ймовірнісний аналіз збіжності найпридатнішої особини до глобального екстремуму [14].

Для задач побудови довгострокового прогнозу проведення таких оптимізацій, як  $A, D, I, G$ , недостатньо, оскільки середньоквадратична похибка, що використовується при побудові регресії на всіх точках вибірки, є внутрішнім критерієм системи. Для однозначного вибору моделі (параметрів моделі), як правило, використовується принцип зовнішнього доповнення і, відповідно, зовнішні критерії [15]. У нашому випадку в якості зовнішнього доповнення використаємо додаткову перевірочну вибірку для такого критерію оптимальності:

Оптимальна за виходом ( $O$ -оптимальна) СБКТ така, що мінімізує штрафну функцію

$$\Lambda_{O\text{-опт}}^* : V_{N_p}(\Omega, D_{N_p}) \rightarrow \min, \quad (12)$$

яка може бути визначена в різний спосіб:

$$V_{N_p}(\Omega, D_{N_p}) = \left\langle E[e_f(x)], E[e_f^2(x)], \max_x |e_f(x)| \right\rangle, \quad x \in D_{N_p} \subset X,$$

де похибка виходу моделі  $e_f(x) = f(x) - \hat{f}(x)$  є одночасно і похибкою наближення поліноміальної функції Без'є–Бернштейна однієї змінної кривою Без'є;  $D_{N_p}$  — перевірочна вибірка даних. При  $V_{N_p}(\Omega, D_{N_p}) = E[e_f^2(x)]$  оптимізаційна задача (12) виражає відомий критерій регулярності. Цільова функція ГА для знаходження СБКТ оптимальної за виходом має вигляд

$$F_{O\text{-опт}}(\Lambda) \sim -V_{N_p}(\Omega, D_{N_p}).$$

Мінімальну похибку прогнозу моделі К. Харріса було отримано саме із застосуванням оптимальної за виходом СБКТ, яка забезпечувала оптимальне нечітке представлення входу.

## ПОЛІНОМІАЛЬНІ ФУНКЦІЇ БЕЗ'Є–БЕРНШТЕЙНА ЯК ЧАСТКОВІ ПЕРЕДСТАВЛЕННЯ

Метод групового врахування аргументів (МГВА) надає певну свободу у виборі опорних функцій, які визначають часткове представлення для базових моделей двох змінних. В класичному багаторядному МГВА використовуються лінійні та квадратичні представлення. Пропонується використовувати поліноміальні функції Без'є–Бернштейна (ПФББ) двох змінних в якості таких опорних функцій, при цьому відбір моделей проводити не тільки за структурою, а й за параметрами, в ролі яких виступає СБКТ для побудови найкращого наближення поліноміальних функцій поверхнею Без'є, яка буде поверхнею відгуку для базової моделі.

Результати показали, що прогноз, складений МГВА з опорними поліноміальними функціями Без'є–Бернштейна, побудованими на основі  $O$ -оптимальної СБКТ, за показником середньоквадратичної похибки (СКП) кращий від прогнозу як класичного багаторядного МГВА, так і методу

К. Харріса, навіть із застосуванням регуляризації оберненого відображення Кастельжо.

## РЕЗУЛЬТАТИ МОДЕЛЮВАННЯ ТА ЇХ АНАЛІЗ

Розглянемо модельний приклад [4]. Нехай нелінійна динамічна система задана у вигляді

$$y(t) = \frac{y^2(t-1) + u_s(t-1)}{1 + y^2(t-1) + y(t-1)u_s(t-2)} + 0,3e(t-1) + e(t),$$

де вхід системи  $u_s(t)$  — рівномірно розподілений сигнал на  $[-1, 1]$ , і шумова складова  $e(t) \sim N(0, 0,1^2)$ . Згенеровано вибірку із 40 основних і 10 перевірочних точок. Нелінійна система моделюється у вигляді

$$y(t) = f(\mathbf{x}(t)) + ce(t-1) + e(t),$$

де  $\mathbf{x}(t) = [y(t-1), u_s(t-1), u_s(t-2)]^T$ .

При побудові моделей отримані результати наведені у таблиці та на рис. 3.

Порівняльний аналіз методів

Метод		Слід матриці коваріації	Визначник матриці коваріації	СКП прогнозу, %
МГВА	Класичний	11919,82	755401	2,37
	з опорними ПФББ O-опт.	149140,66	$3,42 \cdot 10^{10}$	1,37
Алгоритм Харріса	Класичний (нерегуляризований)	253,56	156899,38	3,31
	A-опт.	97,5	3317,78	3,26
	D-опт.	100,68	3464,36	3,41
	O-опт.	173,45	35721,34	2,99

Зазначимо, що дані за класичним багаторядним МГВА приведені по останньому (другому) ряду, а за методом Харріса — по першій частковій підсистемі. Як і очікувалось, A-оптимальність дає найменше значення сліду матриці коваріації і, крім цього, показує менше значення визначника матриці коваріації ніж D-оптимальність. Очевидно, це пов'язано із гіршою збіжністю ГА на D-оптимальності. O-оптимальність стабільно забезпечує найменше значення СКП. Відповідні розподіли СБКТ показані на рис. 3. Як видно з рисунку, A-, D-оптимізації дають більш густий розподіл, і контрольні точки розміщуються щільніше одна до одної, що призводить, відповідно, до збільшення міжбарицентричних відстаней. Тож перевага A-, D-оптимальностей полягає у зниженні чутливості побудованих моделей.

Параметри ГА емпірично були підібрані такими:

- розмір популяції — 30 осіб;

- рівень мутацій — 1%;
- рівень схрещування — 60%;
- кількість поколінь — 400.

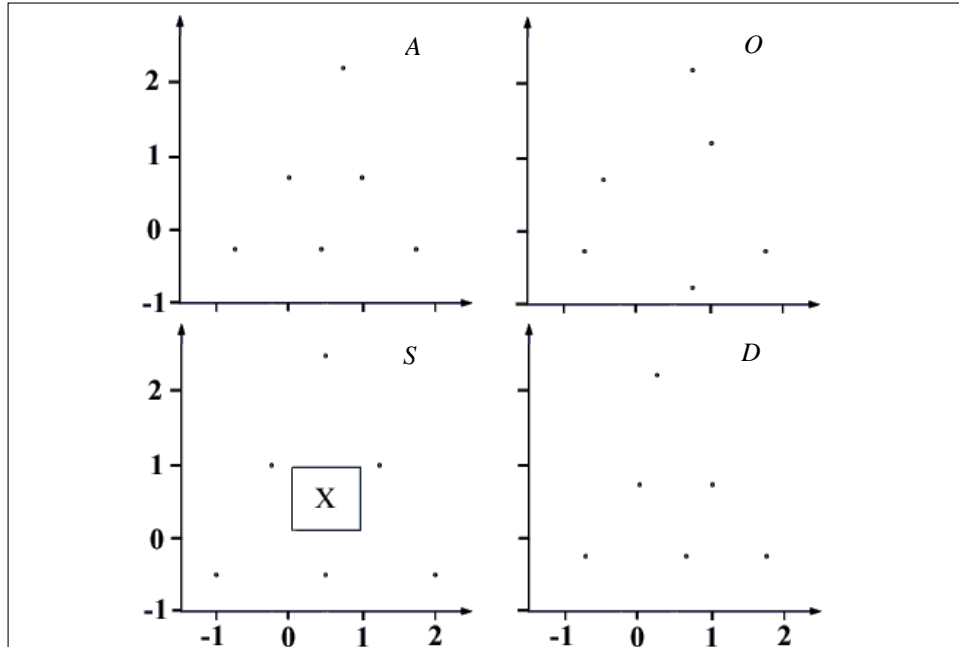


Рис. 3. Розподіли базових контрольних точок: *A*, *O*, *D* — відповідні оптимальності; *S* — нерегуляризований розподіл; *X* — область визначення нормованого входу

## ВИСНОВКИ

За допомогою регуляризації оберненого відображення Кастельжо і, відповідно, поліноміальних функцій Без'є–Бернштейна можна досягти певного зниження абсолютної і середньоквадратичної похибок при застосуванні алгоритму К. Харріса для задач наближення та прогнозування. Але основною перевагою є зниження чутливості побудованих моделей і вирішення питання про розміщення базових контрольних точок, яке до цього часу залишалося відкритим. Тепер можна дати деякі рекомендації по застосуванню алгоритму Харріса. Зокрема, при побудові систем реального часу, де питання швидкодії алгоритму стоїть на першому місці, доцільно використовувати удосконалений алгоритм, де аналітичність обчислення барицентричних координат забезпечує максимальну швидкодію [5]. Якщо ж системи, що проектуються, не обмежені в часі обробки інформації, але критичні по точності і чутливості моделей, то ефективною є розроблена в цій статті регуляризація на основі різного роду оптимальностей (*O*-оптимальність для точності, *A(D)*-оптимальності для чутливості).

Для оцінки ефективності представленої теорії наведено порівняння із відомим класичним багаторядним методом групового врахування аргументів як еталоном в області прогнозування. Крім цього, представлена модифікація МГВА з опорними поліноміальними функціями Без'є–Бернштейна, що потребує більш детального дослідження. Перспективним напрямком є також дослідження ГА і його збіжності для розглянутого класу задач.

## ЛІТЕРАТУРА

1. *О'Квин Донни*. Допечатная подготовка. Руководство дизайнера / Пер. с англ. — М.: Издательский дом «Вильямс», 2002. — 592 с.
2. *Harris C.J., Hong X., Gan Q.* Adaptive modeling, estimation and fusion from data. — Berlin: Springer, 2002. — 323 p.
3. *Brown M., Harris C.J.* Neurofuzzy adaptive modelling and control. — Hemel Hempstead: Prentice Hall, 1994. — 508 p.
4. *Harris C.J., Hong X.* Neurofuzzy network model construction using Bezier Bernstein polynomial functions // IEE Proc. D Control Theory and Applications. — 2000. — **147**, № 3. — P. 337–343.
5. *Митник О.Ю., Бідюк П.І.* Обернене відображення Кагельжо в нечітких нейронних моделях // Системні дослідження та інформаційні технології. — 2004. — № 2. — С. 24–34.
6. *Chen S., Billings S.A., Luo W.* Orthogonal least squares methods and their application to non-linear system identification // Int. J. Control. — 1989. — **50**, № 5. — P. 1873–1896.
7. *Круг Г.К., Сосулин Ю.А., Фатуев В.А.* Планирование эксперимента в задачах идентификации и экстраполяции. — М.: Наука, 1977. — 207 с.
8. *Малютов М.Б., Заиграев А.Ю.* Современные задачи оптимального планирования регрессионных экспериментов. — Киев: Выща шк., 1989. — 64 с.
9. *Hardin R.H., Sloane N. J. A.* A new approach to the construction of optimal designs // Journal of statistical planning and inference. — 1993. — **37**. — P. 339–369.
10. *Феллер В.* Введение в теорию вероятностей и ее приложения. В 2-х т. — Т.1: Пер. с англ. — М.: Мир, 1984. — 528 с.
11. *Genetic algorithms can improve the construction of d-optimal experimental designs / J. Poland, A. Mitterer, K. Knödler, A. Zell // Advances In Fuzzy Systems and Evolutionary Computation. — 2001. — P. 227–231.*
12. *Бідюк П.І., Митник О.Ю.* Застосування генетичного алгоритму в задачі оцінювання вмісту хлорофілу в рослинності // Наук. вісті НТУУ «КПІ». — 2004. — № 4. — С. 65–70.
13. *Rogers A., Prugel-Bennett A.* Genetic Drift in Genetic Algorithm Selection Schemes // IEEE Transactions on Evolutionary Computation. — 1999. — **3**, № 4. — P. 298–303.
14. *Eiben A.E., Aarts E.H.L., Van Hee K.M.* Global convergence of genetic algorithms: a markov chain analysis. — Berlin: Springer, 1991. — P. 4–12.
15. *Ивахненко А.Г.* Индуктивный метод самоорганизации сложных систем. — Киев: Наук. думка, 1981. — 296 с.

Надійшла 20.01.2005