УДК 519.6

В. В. Стоян

Псевдоінверсний підхід до розв'язання одного класу нелінійних алгебраїчних рівнянь

(Представлено членом-кореспондентом НАН України В. В. Скопецьким)

The problem of the root-mean-square inversion of the systems of nonlinear equations, whose left part is the Cartesian product of the set quantity of linear algebraic transformations, is solved. The conditions of exactness and unambiguity of the given solution are defined.

Методика псевдоінверсного обернення систем лінійних алгебраїчних рівнянь, започаткована в [1, 2] та розвинута в [3, 4], була успішно використана [5] при побудові та дослідженні на точність і однозначність середньоквадратичних наближень до розв'язків систем лінійних інтегральних та функціональних рівнянь. Поєднання отриманих при цьому математичних результатів із запропонованим у [6] підходом до моделювання динаміки розподілених просторово-часових процесів дозволило розвинути [7, 8] новий підхід до розв'язання задач динаміки лінійних систем із розподіленими параметрами, досліджуваних як в прямій, так і в оберненій постановках при неповноті інформації [9] про початково-крайовий стан. Для поширення методики [6–8] математичного моделювання розподілених просторово-часових процесів на нелінійні динамічні системи розглянемо можливості використання псевдоінверсних підходів до розв'язання нелінійних алгебраїчних систем спеціального вигляду.

1. Розглянемо систему алгебраїчних рівнянь

$$Ax \otimes Bx = a, \tag{1}$$

де $x \in \mathbb{R}^L$ — шуканий вектор; $A \in \mathbb{R}^{M\times L}$, $B \in \mathbb{R}^{M\times L}$, $a \in \mathbb{R}^M$ — задані матриці та вектор, а символом $\otimes$ позначена операція декартового добутку двох векторів.

Система (1), як і системи

$$Ax = a_1, \tag{2}$$

$$Bx = a_2, \tag{3}$$

де $a_1 \in \mathbb{R}^M$, $a_2 \in \mathbb{R}^M$, такі, що

$$a_1 \otimes a_2 = a, \tag{4}$$

може мати розв'язок (один або множину) або зовсім його не мати. В останньому випадку побудуємо

$$x = \arg\min_{\xi\in\mathbb{R}^L} \|A\xi \otimes B\xi - a\|^2. \tag{5}$$

Будемо виходити [8] з того, що

$$x_1 = \arg\min_{x\in\mathbb{R}^L} \|Ax - a_1\|^2 \in \Omega_1 = \{x_1 \colon x_1 = A^T P_1^+ a_1 + v_1 - A^T P_1^+ A v_1,\ \forall v_1 \in \mathbb{R}^L\}, \tag{6}$$

$$x_2 = \arg\min_{x\in\mathbb{R}^L} \|Bx - a_2\|^2 \in \Omega_2 = \{x_2\colon x_2 = B^T P_2^+ a_2 + v_2 - B^T P_2^+ B v_2,\ \forall v_2 \in \mathbb{R}^L\}, \tag{7}$$

$$\min_{x_1\in\Omega_1} \|Ax_1 - a_1\|^2 = a_1^T a_1 - a_1^T P_1 P_1^+ a_1 = \varepsilon_1^2, \tag{8}$$

$$\min_{x_2\in\Omega_2} \|Bx_2 - a_2\|^2 = a_2^T a_2 - a_2^T P_2 P_2^+ a_2 = \varepsilon_2^2, \tag{9}$$

$$A^+ a_1 = \arg\min_{x_1\in\Omega_1} \|x_1\|^2, \qquad B^+ a_2 = \arg\min_{x_2\in\Omega_2} \|x_2\|^2,$$

де знаком "+" позначена операція псевдообернення матриці, $P_1 = AA^T$, $P_2 = BB^T$, а $v_1 \equiv 0$, $v_2 \equiv 0$ при $\det A^T A > 0$ і $\det B^T B > 0$ відповідно.

При цьому

$$a_1 \otimes BA^+ a_1 = a, \tag{10}$$

$$a_2 \otimes AB^+ a_2 = a. \tag{11}$$

Розглянемо проблеми розв'язання рівнянь (10), (11), покладаючи

$$a = (c_1, c_2, \ldots, c_M)^T,$$

$$a_1 = (\alpha_1, \alpha_2, \ldots, \alpha_M)^T, \tag{12}$$

$$a_2 = (\beta_1, \beta_2, \ldots, \beta_M)^T \tag{13}$$

та враховуючи, що

$$\sum_{j=1}^{M} [BA^+]_{ij} \alpha_i \alpha_j = c_i \qquad (i = \overline{1, M}), \tag{14}$$

$$\sum_{j=1}^{M} [AB^+]_{ij} \beta_i \beta_j = c_i \qquad (i = \overline{1, M}), \tag{15}$$

де $[\cdot]_{ij}$ — ij-й елемент відповідної матриці.

Неважко бачити, що рівняння (14), (15) можна звести до вигляду

$$\overline{A}\alpha = a \tag{16}$$

та

$$\overline{B}\beta = a \tag{17}$$

відповідно, де

$$\alpha = \mathrm{col}(((\alpha_i \alpha_j),\ j = \overline{1, M}),\ i = \overline{1, M}), \tag{18}$$

$$\beta = \mathrm{col}(((\beta_i \beta_j),\ j = \overline{1, M}),\ i = \overline{1, M}), \tag{19}$$

$$\overline{A} = \mathrm{col}((\underbrace{0, \ldots, 0,}_{M(i-1)} \mathrm{str}([BA^+]_{ij},\ j = \overline{1, M}), \underbrace{0, \ldots, 0}_{M^2 - iM}),\ i = \overline{1, M}), \tag{20}$$

$$\overline{B} = \mathrm{col}((\underbrace{0,\ldots,0}_{M(i-1)}, \mathrm{str}([AB^+]_{ij},\ j=\overline{1,M}), \underbrace{0,\ldots,0}_{M^2-iM}),\ i=\overline{1,M}). \tag{21}$$

Звідси знаходимо

$$\Omega_\alpha = \{\alpha\colon \|\overline{A}\alpha - a\|^2 \to \min_\alpha\} = \{\alpha\colon \alpha = \overline{A}^+a + v_\alpha - \overline{A}^+\overline{A}v_\alpha, \forall v_\alpha \in \mathbb{R}^{M^2}\}, \tag{22}$$

$$\Omega_\beta = \{\beta\colon \|\overline{B}\beta - a\|^2 \to \min_\beta\} = \{\beta\colon \beta = \overline{B}^+a + v_\beta - \overline{B}^+\overline{B}v_\beta, \forall v_\beta \in \mathbb{R}^{M^2}\} \tag{23}$$

такі, що

$$\min_{\alpha\in\Omega_\alpha} \|a_1 \otimes BA^+a_1 - a\|^2 = \min_{\alpha\in\Omega_\alpha} \|\overline{A}\alpha - a\|^2 = a^Ta - a^T\overline{A}\overline{A}^+a = \delta_1^2,$$

$$\min_{\beta\in\Omega_\beta} \|a_2 \otimes AB^+a_2 - a\|^2 = \min_{\beta\in\Omega_\beta} \|\overline{A}\beta - a\|^2 = a^Ta - a^T\overline{B}\overline{B}^+\beta = \delta_2^2.$$

При $\det \overline{A}^T\overline{A} > 0$ та $\det \overline{B}^T\overline{B} > 0$ множини Ω_α та Ω_β будуть однозначними ($v_\alpha = v_\beta \equiv \equiv 0$), а

$$\overline{A}^+a = \arg\min_{\alpha\in\Omega_\alpha} \|\alpha\|^2, \tag{24}$$

$$\overline{B}^+a = \arg\min_{\beta\in\Omega_\beta} \|\beta\|^2. \tag{25}$$

Знайдені згідно з (22)–(25) псевдорозв'язки рівнянь (16), (17) із врахуванням визначення (18), (19) останніх дозволяють знайти компоненти α_i ($i = \overline{1,M}$) та β_i ($i = \overline{1,M}$) векторів a_1 та a_2. При цьому

$$\alpha_i^2 = q_{Ai}^Ta + [v_\alpha]_{(i-1)M+i} - q_{Ai}^T\overline{A}v_\alpha, \tag{26}$$

$$\beta_i^2 = q_{Bi}^Ta + [v_\beta]_{(i-1)M+i} - q_{Bi}^T\overline{B}v_\beta, \tag{27}$$

де q_{Ai}^T та q_{Bi}^T — $M(i-1)+i$-рядки матриць $\overline{A}^+$ та $\overline{B}^+$ відповідно. Використовуючи формулу Гревіля [2] обернення прямокутних матриць, розширених рядком, до визначених згідно з (20), (21) матриць $\overline{A}$ та $\overline{B}$, отримаємо такі співвідношення для знаходження q_{Ai}^T та q_{Bi}^T для $i = \overline{1,M}$:

$$(\overline{A}_i \vdots \overline{a}_i)^+ = \begin{pmatrix} Q_A \\ q_{Ai}^T \end{pmatrix}, \qquad (\overline{B}_i \vdots \overline{b}_i)^+ = \begin{pmatrix} Q_B \\ q_{Bi}^T \end{pmatrix}. \tag{28}$$

Тут $Q_A \in \mathbb{R}^{(M^2-1)\times M}$, $Q_B \in \mathbb{R}^{(M^2-1)\times M}$, $\overline{A}_i$ та $\overline{B}_i$ — матриці $\overline{A}$ та $\overline{B}$ без $M(i-1)+i$-х стовпців, $\overline{a}_i$ та $\overline{b}_i$ — ці стовпці.

Виходячи із загальності розв'язків (6), (7) та (22), (23) рівнянь (2), (3) та (16), (17) з урахуванням точностей ε_i^2, δ_i^2 цих розв'язків, розв'язок задачі визначимо співвідношеннями (6), (12), (26), якщо $\delta_1^2 + \varepsilon_1^2 < \delta_2^2 + \varepsilon_2^2$ або (7), (13), (27), якщо це не так.

2. Узагальнимо отримані вище результати на задачу середньоквадратичного обернення системи алгебраїчних рівнянь вигляду

$$A_1x \otimes A_2x \otimes \cdots \otimes A_nx = a, \tag{29}$$

де, як і раніше, $\otimes$ — операція декартового добутку двох векторів; $x \in \mathbb{R}^L$ — шуканий вектор; $A_i \in \mathbb{R}^{M\times L}$ $(i=\overline{1,n})$ та $a \in \mathbb{R}^M$ — задані матриці та вектор.

За аналогією з (2)–(4) робимо висновок, що знаходження

$$x = \arg\min_{\xi\in\mathbb{R}^L} \|A_1\xi \otimes A_2\xi \otimes \cdots \otimes A_n\xi - a\|^2 \tag{30}$$

еквівалентне середньоквадратичному оберненню рівняння

$$a_1 \otimes a_2 \otimes \cdots \otimes a_n = a, \tag{31}$$

де

$$a_i = A_i x \qquad (i=\overline{1,n}). \tag{32}$$

Розв'язок i-го рівняння системи (32) x_i такий, щоб

$$x_i = \arg\min_{x\in\mathbb{R}^L} \|A_i x - a_i\|^2,$$

визначимо співвідношенням

$$x_i \in \Omega_i = \{x_i \colon x_i = P_i^+ A_i^T a_i + v_i - P_i^+ P_i v_i,\ \forall\, v_i \in \mathbb{R}^L\},$$

де плюсом, як і раніше, позначена операція псевдообернення, $P_i = A_i^T A_i$, $v_i \equiv 0$ при $\det P_i >$ > 0, а

$$\min_{x_i\in\Omega_i} \|A_i x_i - a_i\|^2 = a_i^T a_i - a_i^T A_i P_i^+ A_i^T a_i = \varepsilon_i^2.$$

Для знаходження векторів

$$a_i = (a_1^{(i)}, a_2^{(i)}, \ldots, a_M^{(i)})^T \qquad (i=\overline{1,n})$$

через компоненти вектора

$$a = (a_1^{(0)}, a_2^{(0)}, \ldots, a_M^{(0)})^T$$

будемо виходити з того, що рівняння (31) може мати одне з таких n представлень:

$$[A_1 A_i^+ a_i]_j [A_2 A_i^+ a_i]_j \cdots [A_{i-1} A_i^+ a_i]_j a_j^{(i)} [A_{i+1} A_i^+ a_i]_j \cdots [A_n A_i^+ a_i]_j = a_j^{(0)}$$
$$(j=\overline{1,M},\ i=\overline{1,n}),$$

або, що еквівалентне,

$$\overline{A}_i \alpha_i = a \qquad (i=\overline{1,n}). \tag{33}$$

Тут $[\cdot]_j$ — j-й елемент вектора $[\cdot]$,

$$\overline{A}_i = \mathrm{col}\left((\underbrace{0,\ldots 0,}_{M^{n-1}(j-1)}\ A_j^{(i)},\ \underbrace{0,\ldots 0}_{M^n - jM^{n-1}}), j=\overline{1,M} \right),$$

$$\begin{aligned} A_j^{(i)} = \mathrm{str}((\ldots(((\ldots([A_1A_i^+]_{jk_1}[A_2A_i^+]_{jk_2}\ldots[A_{i-1}A_i^+]_{jk_{i-1}}[A_{i+1}A_i^+]_{jk_{i+1}}\ldots[A_nA_i^+]_{jk_n},\\ k_n=\overline{1,M}),\ldots),k_{i+1}=\overline{1,M}),k_{i-1}=\overline{1,M}),\ldots),k_1=\overline{1,M}),\end{aligned} \tag{34}$$

$$\alpha_i = \mathrm{col}((\ldots((a_{k_1}^{(i)} a_{k_2}^{(i)} \ldots a_{k_n}^{(i)}, k_n=\overline{1,M}), k_{n-1}=\overline{1,M}),\ldots), k_1=\overline{1,M}).$$

З (33) знаходимо, що

$$\alpha_i = \arg\min_{\alpha\in\mathbb{R}^{M^n}} \|\overline{A}_i\alpha - a\|^2 \in \Omega_i = \{\alpha_i\colon \alpha_i = \overline{A}_i^+ a + v_i - \overline{A}_i^+\overline{A}_i v_i,\ \forall v_i \in \mathbb{R}^{M^n}\},$$

де $v_i \equiv 0$ при $\det \overline{A}_i^T\overline{A}_i > 0$. Звідси з урахуванням (34) маємо

$$(a_j^{(i)})^n = q_{A_ij}^T a + [v_i]_{j^*} - q_{A_ij}^T\overline{A}_i v_i \qquad (j = \overline{1,M}), \tag{35}$$

де $q_{A_ij}^T$ — рядок матриці $\overline{A}_i^+$ з номером $j^* = ((j-1)(M^{n-1} + M^{n-2} + \cdots + 1) + 1)$. Цей рядок, за аналогією з (28), визначимо співвідношенням

$$(\overline{A}_{ij} \vdots \overline{a}_j)^+ = \begin{pmatrix} Q_{ij} \\ q_{A_ij}^T \end{pmatrix},$$

в якому $\overline{A}_{ij}$ — матриця $\overline{A}_i$ без j^*-го стовпця, а $\overline{a}_j$ — цей стовпець. При цьому

$$\delta_i^2 = \arg\min_{\alpha_i\in\Omega_i} \|\overline{A}_i\alpha_i - a\|^2 = a^T a - a^T\overline{A}_i\overline{P}_i^+\overline{A}_i^T a \qquad (i = \overline{1,n}),$$

де

$$\overline{P}_i = \overline{A}_i^T\overline{A}_i.$$

Розв'язком, або визначеним згідно з (30) псевдорозв'язком, рівняння (29) будемо вважати $x = x_{i_0}$, де $i_0 = \arg\min_{i\in\{1,\dots,n\}} (\varepsilon_i^2 + \delta_i^2)$.

1. *Гантмахер А. Ф.* Теория матриц. – Москва: Наука, 1967. – 287 с.
2. *Альберт А.* Регрессия, псевдоинверсия, рекуррентное оценивание. – Москва: Наука, 1977. – 305 с.
3. *Кириченко Н. Ф.* Псевдообращение матриц и их рекуррентность в задачах моделирования и управления // Пробл. управления и информатики. – 1995. – № 1. – С. 114–127.
4. *Кириченко Н. Ф.* Рекуррентность операций псевдообращения в задачах идентификации и синтеза матриц // Кибернетика и вычислительная техника. – 1994. – № 104. – С. 17–21.
5. *Кириченко Н. Ф., Стоян В. А.* Аналитическое представление матричных и интегральных линейных преобразований // Кибернетика и системный анализ. – 1998. – № 3. – С. 90–104.
6. *Стоян В. А.* Об одном подходе к исследованию начально-краевых задач матфизики // Пробл. управления и информатики. – 1998. – № 1. – С. 79–86.
7. *Скопецький В. В., Стоян В. А., Кривонос Ю. Г.* Математичне моделювання прямих та обернених задач динаміки систем з розподіленими параметрами. – Київ: Наук. думка, 2001. – 361 с.
8. *Стоян В. А.* Моделювання та ідентифікація динаміки систем із розподіленими параметрами. – Київ: ВПЦ "Київський університет", 2004. – 187 с.
9. *Скопецкий В. В., Стоян В. А.* О некоторых новых результатах по решению проблем моделирования и управления динамикой систем с распределенными параметрами // Пробл. управления и информатики. – 2002. – № 3. – С. 73–84.

Київський національний університет
ім. Тараса Шевченка

Надійшло до редакції 17.09.2007