ВІСНИК НАЦІОНАЛЬНОГО УНІВЕРСИТЕТУ
"ЛЬВІВСЬКА ПОЛІТЕХНІКА"
"Фізико-математичні науки"
№ 518, 2004, с. 51–59

JOURNAL OF LVIV POLYTECHNIC
NATIONAL UNIVERSITY
"Physical and mathematical sciences"
№ 518, 2004, p. 51–59

РАЦІОНАЛЬНИЙ АЛГОРИТМ ПСЕВДООБЕРНЕННЯ ДЛЯ МАЙЖЕ ВИРОДЖЕНИХ МАТРИЦЬ ІЗ ЗАСТОСУВАННЯМ ДО ЗАДАЧІ ЛІНІЙНОГО ПРОГРАМУВАННЯ

О. Рибицька

Національний університет "Львівська політехніка"
вул. С.Бандери 12, 79013, Львів, Україна

(Отримано 27 жовтня 2004 р.)

Вивчається дискретна система рівнянь з майже виродженою матрицею. На основі теорії параметризованих неперервних дробів і апроксимацій Паде здійснена регуляризація її розв'язку. Запропонований алгоритм дозволяє успішно розв'язувати клас систем лінійних рівнянь з матрицею невизначеного рангу. Результати ілюсртуються тестовими прикладами, зокрема класичним – задачею лінійного програмування.

Ключові слова: майже вироджена матриця, псевдообернення, регуляризація, неперервний дріб, задача лінійного програмування.

2000 MSC: 65J20

УДК: 51-96:62-50

Вступ

У практичних задачах найчастіше потрібно враховувати велику кількість можливо суперечливих вимог. Такі задачі можна розв'язати тільки шляхом вибору деякого компромісу – всі вимоги можуть бути задоволені не повністю, а лише частково.

Сьогодні існує струнка теорія, яка вирішує такого роду проблему з позицій псевдообернення. Але ця теорія недостатньо розвинута, оскільки існуючі нині методи псевдообернення залежать від того, які значення приписуються рангу вихідної матриці.

У запропонованій роботі зроблена спроба підвищення точності та вибору більш вдалого алгоритму стосовно звуження класу матриць із невизначеним рангом. Це, переважно здійснюється на засадах використання теорії параметризованих неперервних дробів та апроксимант Паде. Побудована процедура оцінок, які є нечутливими до наявності у моделі " засмічуючих " спостережень.

I. Випадок точного задання вхідної інформації та за точного виконання обчислень

Розглянемо систему лінійних рівнянь

$$Hx = b \tag{1}$$

із матрицею H розмірів $m \times n$, b – довільний стовпець розміру m. На параметри m, n і ранг матриці H не накладатимемо жодних умов.

Очевидно, система (1) може бути сумісною, або несумісною. Але вже система

$$H^T Hx = H^T b, \tag{2}$$

де символ " T " означає транспонування, обов'язково є сумісною [1, 2]. Систему (2) називатимемо *нормальною системою.*

Якщо розв'язок нормальної системи (2) не єдиний, то виникає задача про вибір якогось одного із розв'язків, і найчастіше вибирають розв'язок із мінімальною евклідовою нормою.

Означення 1. Нормальним псевдорозв'язком системи (1) називають стовпець із мінімальною нормою серед стовпців, що надають мінімальну за евклідовою нормою нев'язку при підставленні цього стовпця в систему.

Існує таке твердження [1].

Твердження 1. Кожна система (1) має один і тільки один нормальний псевдорозв'язок.

Крім (1) розглянемо системи лінійних рівнянь

$$Hx = \mathrm{I}_i, \tag{3}$$

де I_i – i-й стовпець одиничної матриці I розмірів $m \times m$.

Означення 2. Псевдооберненою матрицею до матриці H називатимемо матрицю H^+, стовпці якої – це псевдорозв'язки систем лінійних рівнянь (3).

Матриця H^+ має ті ж розміри, що і матриця H^{T}.

Кожна матриця H має одну і тільки одну псевдообернену [1], яка для випадку невиродженої матриці H перетворюється в H^{-1}. Крім того,

$$H^+ = (H^T H)^{-1} H^T$$

або

$$H^+ = H^T (HH^T)^{-1}$$

для випадків, коли відповідно стовпці або рядки матриці H лінійно незалежні [1].

Псевдорозв'язок x^+ системи (1) можна записати як [4]

$$x^+ = \mathrm{H}^+ b. \tag{4}$$

Псевдообернену мотивацію можна подати і у вигляді граничного співвідношення

$$H^+ = \lim_{\lambda \to 0} (H^T H + \lambda^2 I)^{-1} H^T.$$

Розглянемо систему лінійних рівнянь із невиродженою матрицею

$$(H^T H + \lambda^2 I) x_\lambda = H^T b \tag{5}$$

і запишемо її розв'язок у вигляді

$$x_\lambda = (H^T H + \lambda^2 I)^{-1} H^T b. \tag{6}$$

Здійснивши тут граничний перехід при $\lambda \to 0$, матимемо

$$x^+ = \lim_{\lambda \to 0} (H^T H + \lambda^2 I)^{-1} H^T b. \tag{7}$$

Останній запис має суттєве теоретичне значення. Річ у тому, що X^+ не є неперервною функцією стосовно елементів матриці H. Запис (7) вказує на те, що система (1) може бути включена у сім'ю систем (5) із параметром λ так, щоб розв'язок системи неперервно залежав від цього параметра.

Введено позначення $H^T \cdot H = A$ і $\lambda^2 = s$. Тоді згідно з (4) – (7) та результатами [3–5] матимемо

$$H^+ = \lim_{\lambda \to 0} \frac{s^{r-1} B_0 + s^{r-2} B_1 + \dots + B_{r-1}}{s^r + d_1 s^{r-1} + \dots + d_r} H^T \tag{8}$$

і

$$x^+ = \lim_{\lambda \to 0} \frac{s^{r-1} B_0 + s^{r-2} B_1 + \dots + B_{r-1}}{s^r + d_1 s^{r-1} + \dots + d_r} H^T b, \tag{9}$$

де $B_0 = I$

$$\begin{array}{ll}
d_1 = \operatorname{tr} A & B_1 = d_1 I - B_0 A \\
d_2 = \dfrac{1}{2} \operatorname{tr}(B_1 A) & B_2 = d_2 I - B_1 A \\
\dots\dots\dots & \dots\dots\dots \\
d_k = \dfrac{1}{k} \operatorname{tr}(B_{k-1} A) & B_k = d_k I - B_{k-1} A \\
\dots\dots\dots & \dots\dots\dots \\
d_{n-1} = \dfrac{1}{n-1} \operatorname{tr}(B_{n-2} A) & B_{n-1} = d_{n-1} I - B_{n-2} A \\
d_n = \dfrac{1}{n} \operatorname{tr}(B_{n-1} A) & 0 = d_n I - B_{n-1} A
\end{array} . \tag{10}$$

Символом $\operatorname{tr} A$ позначено слід матриці A.

Якщо у співвідношеннях (8) і (9) здійснити граничний перехід, то матимемо

$$H^+ = \frac{B_{r-1}}{d_r} H^T \tag{11}$$

і

$$x^+ = \frac{B_{r-1}}{d_r} H^T b, \tag{12}$$

Зауважимо, що якщо $d_{r+1} = 0$, тоді $d_{r+2} = d_{r+3} = \dots = d_n = 0$. Позначимо r – номер крайнього відмінного від нуля елемента $d_r (r \le n,\ B_r A = O)$. Слід наголосити, шо якщо відомий ранг матриціH, то $r = \operatorname{Rang} H$.

Приклад 1. [6]. Розглянемо систему (1) із

$$H = \begin{pmatrix} 14 & 35 & -7 & -63 \\ -10 & -25 & 5 & 45 \\ 26 & 65 & -13 & -117 \end{pmatrix}$$

$$\text{і} \quad b = \begin{pmatrix} 777 \\ -555 \\ 1443 \end{pmatrix}.$$

Обчислення $\operatorname{Rang} H = 1$,

$$A = \begin{pmatrix} 972 & 2430 & -496 & -4374 \\ 2430 & 6075 & -1215 & -10935 \\ -496 & -1215 & 243 & 2187 \\ -4374 & -10395 & 2187 & 19683 \end{pmatrix}$$

і

$$H^T b = \begin{pmatrix} 53946 \\ 134865 \\ -26973 \\ -242757 \end{pmatrix}$$

дозволяють, на засадах (12), записати

$$x^+ = \frac{B_0}{d_1} H^T b, \quad \text{або} \quad x^+ = \frac{H^T b}{d_1},$$

де

$$d_1 = trA = 972 + 6075 + 243 + 19683 = 26973$$

Отже,

$$x^+ = \frac{1}{26973} \begin{pmatrix} 53946 \\ 134865 \\ -26973 \\ -242758 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \\ -1 \\ 9 \end{pmatrix}.$$

У цьому прикладі

$$B_1 A = (269731 \cdot I - A) A = 26973 A - A^2 = O.$$

Такий результат не випадковий. Він дозволяє визначити H^+ і x^+ без попереднього підрахунку рангу матриці H.

Надалі нам буде важливим і покомпонентне подання x_λ (6), у вигляді скінченних j-дробів [7]. Згідно з [8, 9] кожен елемент стовпця невідомих x_λ можна обчислити за формулою

$$x_\lambda^{(k)} = \frac{\beta_1^{(k)}}{s+\beta_2^{(k)}} - \frac{\beta_2^{(k)}\beta_3^{(k)}}{s+(\beta_3^{(k)}+\beta_4^{(k)})} - \cdots - \frac{\beta_{2r-2}^{(k)}\beta_{2r-1}^{(k)}}{s+(\beta_{2r-1}^{(k)}+\beta_{2r}^{(k)})}, \quad k=\overline{1,n}, \tag{13}$$

де $x_\lambda = (x_\lambda^{(1)}, x_\lambda^{(2)}, ..., x_\lambda^{(n)})^T$.

Дріб (13) заміною $s = \dfrac{1}{\alpha}$ можна подати у вигляді правильного RITZ-дробу [8]

$$x_\lambda^{(k)} = \frac{\beta_1^{(k)}}{1} + \frac{\beta_2^{(k)}}{1} + \cdots + \frac{\beta_{2r}^{(k)}\alpha}{1}, \tag{14}$$

де $\alpha x_\alpha^{(k)} = x_\lambda^{(k)}$.

Здійснивши тепер в (13) граничний за $\lambda \to 0$ перехід, одержимо таке твердження [6, 8].

Теорема 1. *Якщо для всіх $k=\overline{1,n}$ підграничний вираз у (9) допускає j-дробове подання (13), то*

$$x_k^+ = \frac{\beta_1^{(k)}}{\beta_2^{(k)}} - \frac{\beta_2^{(k)}\beta_3^{(k)}}{\beta_3^{(k)}+\beta_4^{(k)}} - \cdots - \frac{\beta_{2r-2}^{(k)}\beta_{2r-1}^{(k)}}{\beta_{2r-1}^{(k)}+\beta_{2r}^{(k)}} \quad , \tag{15}$$

де x_k^+ – k-та компонента псевдорозв'язку системи (1).

Легко переконатись і в справедливості наступного твердження [6, 8].

Теорема 2. *Дріб (15) еквівалентний скінченній сумі*

$$x_k^+ = \sum_{i=1}^{r} \frac{\beta_1^{(k)}\beta_3^{(k)}...\beta_{2i-1}^{(k)}}{\beta_2^{(k)}\beta_4^{(k)}...\beta_{2i}^{(k)}} = \frac{\beta_1^{(k)}}{\beta_2^{(k)}}\left(1+\frac{\beta_3^{(k)}}{\beta_4^{(k)}}\left(1+...+\frac{\beta_{2r-3}^{(k)}}{\beta_{2r-2}^{(k)}}\left(1+\frac{\beta_{2r-1}^{(k)}}{\beta_{2r}^{(k)}}\right)...\right)\right).$$

Зауваження 1. Компоненти дробу (13) можна одержати також зі скінченного матричного ряду

$$\omega + \alpha A\omega + \alpha^2 A^2\omega + \alpha^3 A^3 \omega + ... + \alpha^{2r-1}A^{2r-1}\omega, \tag{16}$$

де $\alpha = -\dfrac{1}{s}$, $\omega = H^T b$, $A = H^T H$.

Для нього у роботі [8] виведені рекурентні співвідношення

$$\beta_{m+1}^{(k)} = -\frac{(-1)^m}{\beta_1^{(k)}...\beta_m^{(k)}}(A_k^m + Q_{m,1}^{(k)}A_k^{m-1} + ... + Q_{m,[m/2]}^{(k)}A_k^{m-[m/2]}), \tag{17}$$

де A_k^i – k-та компонента вектора $A^i\omega$, наприклад,

$$\beta_1^{(k)} = \omega^{(k)},\ \beta_2^{(k)} = -\frac{1}{\beta_1^{(k)}}A_k^1,\ \beta_3^{(k)} = \frac{(A_k^1)^2 - c^{(k)}A_k^2}{\omega^{(k)}A_k^1},\ \beta_4^{(k)} = \frac{\omega^{(k)}((A_k^2)^2 - A_k^1 A_k^3)}{A_k^1(\omega^{(k)}A_k^2 - (A_k^1)^2)}. \tag{18}$$

У (17)

$$Q_{m,j}^{(k)} = Q_{m-1,j}^{(k)} + \beta_2^{(k)}Q_{m-2,j-1}^{(k)},\ Q_{m,0}^{(k)} = 1,\ Q_{2,1}^{(k)} = \beta_2^{(k)},\ Q_{2m,m}^{(k)} = \beta_2^{(k)}\beta_4^{(k)}...\beta_{2m}^{(k)}.$$

Приклад 2. Нехай у системі (1)

$$H = \begin{pmatrix} 2 & -1 & 3 & 3 \\ 1 & 2 & 1 & 6 \\ -1 & 1 & 2 & 3 \\ 3 & 2 & 0 & 7 \\ 1 & -2 & 3 & 0 \\ 2 & 1 & -2 & 2 \\ -2 & 3 & 1 & 5 \\ 3 & 1 & -4 & 1 \end{pmatrix} \quad b = \begin{pmatrix} 3 \\ 22 \\ 20 \\ 30 \\ -2 \\ 14 \\ 5 \\ 15 \end{pmatrix}$$

Ця система несумісна із $\operatorname{Rang} H = 3$. Тут матриця

$$A=\begin{pmatrix} 33 & 2 & -10 & 27\\ 2 & 25 & -8 & 44\\ -10 & -8 & 44 & 18\\ 27 & 44 & 18 & 133 \end{pmatrix}, \quad \text{а} \quad H^Tb=\begin{pmatrix} 159\\ 169\\ -18\\ 479 \end{pmatrix}.$$

Згідно з (10) маємо

$$d_1=33+25+44+133=235, \quad B_1=d_1I-A=\begin{pmatrix} 202 & -2 & 10 & -27\\ -2 & 210 & 8 & -44\\ 10 & 8 & 191 & -18\\ -27 & -44 & -18 & 102 \end{pmatrix},$$

$$d_2=\frac{1}{2}\operatorname{tr}B_1A=\frac{1}{2}\operatorname{tr}\begin{pmatrix} 5833 & -914 & -2050 & 1955\\ -914 & 3246 & -2100 & 2478\\ -2050 & -2100 & 7916 & 1666\\ 1955 & 3478 & 1666 & 10577 \end{pmatrix}=13786.$$

$$B_2=d_2I-B_1A=\begin{pmatrix} 7953 & 914 & 2050 & -1955\\ 914 & 10540 & 2100 & -3478\\ 2050 & 2100 & 5870 & -1666\\ -1955 & -3478 & -1666 & 3209 \end{pmatrix},$$

$$d_3=\frac{1}{3}\operatorname{tr}\, B_2A=222824, \quad d_4=0, \quad B_3H^Tb=(0000)^T.$$

Отже,

$$x_\lambda=\frac{s^2I+sB_1+B_2}{s^3+235s^2+13786s+222824}\begin{pmatrix} 159\\ 169\\ -18\\ 479 \end{pmatrix}$$

і, наприклад,

$$x_\lambda^{(1)}=\frac{159s^2+18657s+445648}{s^3+235s^2+13786s+222824}=\cfrac{159}{s+117,6404-\cfrac{2823,043s+106956,63}{s^2+117,3396s+2802,8176}}=$$

$$=\cfrac{159}{s+117,6604-\cfrac{2823,0403}{s+79,4506-\cfrac{207,3272}{s+37,8870}}}.$$

З цього виразу випливає, що

$$\beta_1^{(1)}=159; \quad \beta_2^{(1)}=117,6604; \quad \beta_2^{(1)}\beta_3^{(1)}=2823,0430; \quad \beta_3^{(1)}+\beta_4^{(1)}=79,4506;$$

$$\beta_4^{(1)}\beta_5^{(1)}=207,3272; \quad \beta_5^{(1)}+\beta_6^{(1)}=37,8870.$$

Із цих співвідношень визначаємо

$$\beta_3^{(1)}=23,9931; \quad \beta_4^{(1)}=55,4575; \quad \beta_5^{(1)}=3,7385; \quad \beta_6^{(1)}=34,1485.$$

Принагідно зауважимо, що в цьому прикладі

$$\beta_2^{(1)}\beta_4^{(1)}\beta_6^{(1)}=222824=d_3 \quad \text{і} \quad \frac{\beta_1^{(1)}}{\beta_2^{(1)}}\left(1+\frac{\beta_3^{(1)}}{\beta_4^{(1)}}\left(1+\frac{\beta_5^{(1)}}{\beta_6^{(1)}}\right)\right)=\frac{445648.}{222824}=2=x_1^+.$$

II. Випадок неточної вхідної інформації

Якщо довільна практична задача зводиться до системи лінійних рівнянь, то коефіцієнти такої системи найчастіше відомі неточно. Крім того, запис вхідних даних та їх обчислень пов'язані з похибками заокруглень, вплив яких рівносильний деякому викривленню коефіцієнтів системи.

Уникнути вказаних викривлень ми не в змозі, але можемо, по-перше, оцінити одержану похибку і, по-друге, постаратись вибрати такий метод розв'язування системи, який би не збільшив неточність результату, вже закладеного в саму систему.

Збурення коефіцієнтів системи лінійних рівнянь може не тільки викривити її розв'язок, а мати і серйозніші наслідки. Врахування похибок заокруглень та збурень роблять розмитими межі множин матриць деякого фіксованого рангу. Розмитим стає також і поняття лінійної залежності системи векторів тощо.

Нехай замість (1) задана збурена система

$$\tilde{H}x = \tilde{b}, \tag{19}$$

для якої матриці $\tilde{H}$ і стовпець $\tilde{b}$ задовольняють нерівності

$$||\tilde{H} - H|| < \delta, \quad ||\tilde{b} - b|| < \delta \tag{20}$$

(розглядатимемо спектральну норму матриць).

Означення 1. Матрицю A називають майже виродженою, якщо малі зміни її елементів можуть перетворити її у вироджену матрицю.

Малість визначника матриці не є необхідною чи достатньою умовою її майже виродженості. Властивість матриці бути майже виродженою пов'язана із нормою її оберненої матриці та числом обумовленості. Крім того, може трапитись, що матриця $A = H^T H$ є майже вироджена, хоча матриця H цією властивістю не володіє. Матриця $A = H^T H$ є, взагалі кажучи, значно гірше обумовлена, ніж сама матриця H.

Надалі використовуватимемо такі понятійні засади.

Позначимо через $F_{d,t}$ множину чисел, які записуються через d десяткових цифр із знаком " $+$ " або " $-$ ", причому t цифр розміщені після коми, тобто зображають дробову частину. Множина $F_{d,t}$ скінченна, хоча за великого d є дуже великою. Дійсне число можна наближено подати числом із $F_{d,t}$,якщо ціла частина менша за 10^{d-t} .

Якщо число наближене числом $\tilde{a}$ із $F_{d,t}$, то модуль абсолютної похибки $|a - \tilde{a}|$ не перевищує $\frac{1}{2}10^{-t}$.

При псевдооберненні найбільш незрозумілою є проблема визначення рангу матриці. Але тут допомогти може простіше питання – вияснити, чи є у розкладах (8), (9), (13) чи (14) зневажливо малі окремі доданки.

Для задачі (1), враховуючи оцінки і похибки вхідної інформації й заокруглень (19)–(20) та похибки, що внесені при обчисленні за схемою (10), чи (13), встановлюють деяке число ε як певну межу: якщо $d_r, d_{r-1}, \ldots, d_{r-p+1}$ чи $\beta_2^{(k)}\beta_4^{(k)}...\beta_{2(r-p+1)}^{(k)}$ менші ніж ε, то ранг матриці вважатимемо таким, що дорівнює $r - p$.

Отже, кількість обчислених рекурентно через (10) параметрів B_{k-1} і d_k, необхідних для побудови (11) чи (12), залежить від величини ε. У таких обчислювальних процесах маємо справу не із самою матрицею $\tilde{A} = \tilde{H}^T\tilde{H}$, а з іншою матрицею A_ε.

Ранг A_ε залежить від ε, причому $\mathrm{Rang}_{\,\varepsilon} \leq \mathrm{Rang}\,\widetilde{A}$, спектральне число обумовленості c задовольняє нерівності

$$c(A_\varepsilon) \leq \frac{\alpha_1}{\varepsilon} \leq c(A),$$

де α_1 – найбільше сингулярне число матриці A; $c(A) = ||A^{-1}||\ ||A||$.

Отже, зробивши ε більшим, ніж мінімальне сингулярне число матриці A, ми в змозі вибрати із $\widetilde{A}$ найвірогіднішу інформацію, якої, природно, замало. Цей ефект надалі буде використаний для виділення найсуттєвішої частини даних, записаних у матриці, а цьому, як буде показано нижче, допомагає теорія апроксимант Паде або її частковий випадок – теорія неперервних параметризованих дробів. Для цього ε задають настільки великим, щоб одержати матрицю A_ε прийнятного рангу, а саме: достатньо малого для того щоб, дані були оглядовими, і достатньо великого, щоб вони були значимими. Крім того, як буде показано нижче, вдалий вибір ε об'єднує два різні поняття: точність та стійкість.

Теорема 1.(Про регуляризацію).

Нехай для системи (19) виконуються умови (20). Тоді у розкладах (8)–(9) чи (13)–(14) для встановлення рангу матриці A_ε треба вибрати ε близьким до $\sqrt{\delta}$, де δ визначене в (20).

Як практично використати теорему регуляризації підкажуть такі модельні приклади: перший із майже нульовим визначником і другий з погано обумовленою матрицею.

Приклад 1. Розглянемо вироджену сумісну систему рівнянь

$$\left\{ \begin{array}{r} x_1 + \sqrt{2}x_2 = 1 \\ \sqrt{2}x_1 + 2x_2 = \sqrt{2} \end{array} \right. \tag{21}$$

і її наближений варіант

$$\left\{ \begin{array}{r} x_1 + 1,414x_2 = 1 \\ 1,414x_1 + 2x_2 = 1,41 \end{array} \right. . \tag{22}$$

Система (22) має єдиний розв'язок: $x_1 = 10,36$; $x_2 = -6,62$, що записаний згідно з точністю правої частини.

Інша наближена (21) система

$$\begin{cases} x_1 + 1,41x_2 = 1 \\ 1,41x_1 + 2x_2 = 1,41 \end{cases}$$

має також єдиний розв'язок: : $x_1 = 0,529$; $x_2 = 0,336$.

Як бачимо, тут стійкості немає. Це є наслідком того, що у разі неточного задання $\sqrt{2}$ система (21) перетворюється у невироджену систему з єдиним розв'язком, який не є близьким до нормального розв'язку $x^+ = \left(\frac{1}{3}, \frac{\sqrt{2}}{3}\right)$.

Розв'яжемо тепер (22) за схемою (9)–(10) із використанням теореми регуляризації. Для неї $d_1 = 3$, а

$$B_1 = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} - \begin{pmatrix} 1 & 1,414 \\ 1,414 & 2 \end{pmatrix} =$$

$$= \begin{pmatrix} 2 & -1,414 \\ -1,414 & 1 \end{pmatrix}$$

і

$$B_1A = \begin{pmatrix} 1-(1,414)^2 & 0 \\ 0 & 2-(1,414)^2 \end{pmatrix} =$$

$$= \begin{pmatrix} 0,000604 & 0 \\ 0 & 0,000604 \end{pmatrix}.$$

Оскільки в нашому випадку $\delta = \frac{1}{2}10^{-2}$ і $\sqrt{\delta} = \frac{1}{10\sqrt{2}}$, то треба вважати (із точністю $\sqrt{\delta}$) матрицю B_1A нульовою, а це означає, що наближеним значенням нормального розв'язку задачі (22) є

$$x^+ \approx \frac{B_0}{d_1}\begin{pmatrix} 1 \\ 1,41 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 1 \\ 1,41 \end{pmatrix} \approx \begin{pmatrix} 0,33 \\ 0,47 \end{pmatrix}.$$

Приклад 2. [13]. Нехай матриця $A = H^TH$ і вектор H^Tb для (1) мають відповідно вигляд

$$A = \begin{pmatrix} 7 & 8 & 9 \\ 8 & 9 & 10 \\ 9 & 10 & 8 \end{pmatrix}, \quad H^Tb = \begin{pmatrix} 24 \\ 27 \\ 27 \end{pmatrix}.$$

Матриця A погано обумовлена, оскільки її нормалізований визначник

$$|\bar{A}| = \frac{\det A}{A_1A_2A_3} = \frac{3}{\sqrt{194 \cdot 245^2}} \approx 8,79 \cdot 10^{-4},$$

де $A_i = \sqrt{\sum_{j=1}^{3} a_{ij}^2}$, $i = 1,2,3$, а a_{ij} – елементи матриці A.

За таких точних значень параметрів система (2) має єдиний розв'язок $x = (1,1,1)^T$.

Якщо тепер незначно змінити вектор $g = H^Tb$, наприклад, вважати його таким:

$$\tilde{g} = (24,2;\ 27,0;\ 27,2)^T,$$

то одержимо розв'язок $\tilde{x} = (-0,933;\ 2,866;\ 0,866)^T$.

Такий результат є наслідком того, що матриця прикладу майже вироджена. Для того, щоб і тут знайти стійкий розв'язок, потрібно використати теорему про регуляризацію.

За схемою (16) для наближеного $\tilde{g}$ маємо

$$\begin{pmatrix} 24,2 \\ 27,0 \\ 27,2 \end{pmatrix} + \alpha\begin{pmatrix} 630,2 \\ 708,6 \\ 705,4 \end{pmatrix} + \alpha^2\begin{pmatrix} 16428,8 \\ 18437 \\ 18401 \end{pmatrix} + \ldots$$

Нехай коефіцієнти вектора $\tilde{g}$ обчислюється із точністю $\delta = 4 \cdot 10^{-2}$. Отже, тут $\sqrt{\delta} \approx 0,2$ і тому згідно з теоремою про регуляризацію надалі у рахунках всі параметри, що є меншими, або рівними, ніж 0,2, вважатимемо такими, що дорівнюють нулю. Наприклад, для координати $\tilde{x}_1$ вектора-розв'язку $\tilde{x} = (\tilde{x}_1, \tilde{x}_2, \tilde{x}_3)^T$ маємо

$$\beta_1^{(1)} = 24,2; \quad \beta_2^{(1)} = -\frac{1}{\beta_1^{(1)}}A_1^1 = -\frac{1}{24,2}630,2;$$

$$\beta_3^{(1)} = \frac{(630,2)^2 - 24,2 \cdot 16428,8}{24,2 \cdot 630,2} \approx 0.$$

Тому,

$$\tilde{x}_1 = -\frac{\beta_1}{\beta_2} = \frac{(24,2)^2}{630,2} \approx 0,9.$$

Аналогічно, $\tilde{x}_2 = \tilde{x}_3 \approx 1,0$.

Отже, для прикладу ранг матриці A_ε потрібно вибрати таким, що дорівнює одиниці. Тоді обчислювальний процес стабілізується, а результуючі дані стають оглядовими і значимими.

Підсумовуючи вищенаведене, зазначимо, що вказаний у роботі метод регуляризації не потребує громіздких обчислень для знаходження параметра регуляризації (див., наприклад, метод регуляризації Тіхонова-Філіпса [12]). Він споріднений з ітераційними методами регуляризації, в яких для знаходження нормального розв'язку потрібно вчасно призупинити ітераційний процес, тобто знайти потрібний індекс ітерації. Варто тут наголосити і на тому, що вдале використання неперервного дробу часто приводить до бажаних результатів. Це, наприклад, проблема доведення стійкості многочлена [7], побудова апроксимант Паде [14] тощо. У роботі використана ще одна дуже важлива властивість неперервного дробу: здатність зосереджувати основну результуючу інформацію на перших поверхах дробу.

Оскільки вказаний у роботі метод регуляризації суттєво відрізняється від найуживанішого на сьогодні Тіхоновського методу регуляризації некоректних задач, розглянемо дещо детальніше останній. Метод Тіхонова–Філіпса містить деякі суттєві вади. Поперше, згідно з таким підходом змінюється сам оператор системи рівнянь, що стає коректним, тобто стійким до збурень. Отже, тут не знаходиться розв'язок вихідного рівняння і не ставиться питання про побудову регуляризуючого алгоритму його розв'язування. По-друге, під час тіхоновської регуляризації немаловажне значення має вдалий вибір параметра регуляризації. Здебільшого його відшукання є складнішою задачею за первісну, тому практики найчастіше вибирають його емпірично. По-третє, саме знаходження розв'язку збуреного рівняння збігається досить повільно. Крім того, за такої регуляризації одержують не прямі, а зміщені оцінки.

III. Регуляризуючий алгоритм задачі лінійного програмування

Будь-яку задачу лінійного програмування (ЛП) можна записати в матричній формі так:

максимізувати функцію

$$(c^T, x^T), \tag{23}$$

за умов

$$Ax = b, \quad x \geq 0. \tag{24}$$

Тут: x – n-вимірний шуканий вектор-стовпець; x^T – його транспонента; c^T – n-вимірний вектор-рядок; b – m-вимірний вектор-стовпець; A – прямокутна матриця розмірів $m \times n$.

В основі теорії ЛП лежить теорема Куна-Таккера про необхідні та достатні умови існування оптимальної точки $x*$, яка б задовольняла умови (23) і (24). Умови Куна-Таккера мають вигляд

$$Ax = b \tag{25}$$

$$A^T\lambda - \nu = c \tag{26}$$

$$x_j\nu_j = 0, \ j = \overline{1, n}. \tag{27}$$

Отже, для існування розв'язку x^* задачі (23)–(24), треба і досить, щоб віднайшлися такі вектори $\nu^* \geq 0$ і λ^*, щоб система векторів x^*, λ^*, ν^* задовольняла умови (25)–(27).

Легко переконатися в тому, що довільний $2n+m$-вимірний вектор (x, λ, ν), що є розв'язком системи (25)–(26) за умов $x \geq 0$ і $\nu \geq 0$, буде її базисним розв'язком (крайньою точкою).

Для задачі (23)–(24) пошук точки Куна-Таккера зводиться до розв'язання двох СЛАР (25) і (26). У першій системі маємо n невідомих $x = (x_1, x_2, \ldots, x_n)$. У другій $m + n$ невідомих $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_m)$ і $\nu = (\nu_1, \nu_2, \ldots, \nu_n)$.

Обчислювальний пошук оптимальної точки (x^*, λ^*, ν^*) полягає ось у чому. Передусім розв'язуємо систему (25) за схемою (13), (15). Одержимо нормальний розв'язок, серед компонент якого, можливо, є і від'ємні числа. Замінимо останні нулями і повернемось знову, за тим же алгоритмом до відшукання розв'язків скороченої системи. Отже, буде знайдений нормальний невід'ємний розв'язок системи (25). Аналогічно обчислюється стосовно невід'ємних значень ν нормальний розв'язок системи (26). Такий розв'язок (x, λ, ν) за умов $x \geq 0$ і $\nu \geq 0$ буде бажаним розв'язком задачі (25)–(26).

Якщо для цього розв'язку для певного j умова (27) не виконується, то приймаємо в (25) $x_j = 0$ і продовжуємо весь процес (25)–(27) заново до повного виконання умов Куна-Таккера.

Приклад 1. [11]. Нехай в (23)–(24)

$$A = \begin{pmatrix} 2 & -2 & 1 & 1 & 0,5 \\ 1 & 1 & 1 & 0 & 0,5 \\ 3 & -1 & 2 & 1 & 1 \end{pmatrix}, \qquad b = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}, \quad c^T = (-1, -1, 0, 0, 0) \tag{28}$$

Стандартна програма розв'язання цієї задачі, побудована за сімплекс-методом, видала як розв'язок вектор $(0, 0, 1, 1, 0)^T$.

Це насправді розв'язок задачі (28), хоча він не володіє мінімальною нормою.

Збуримо тепер в (28) вектор b, прийнявши $b_3 = 3,01$ (замість 3), залишивши без змін всі інші дані цього вектора. Тепер ця ж сама стандартна програма не видає жодного розв'язку і повідомляє про несумісність обмежень. Таку ж відповідь одержимо і за інших збурень (навіть дуже малих).

З погляду класичного підходу ці результати цілком природні і кажуть про достатньо високу якість стандартної програми. Якщо ж розглянути з погляду регуляризації, то ці ж самі результати вказують на непридатність сімплекс-методу як регуляризуючого алгоритму.

Розв'яжемо тепер цю задачу запропонованим у роботі алгоритмом. У прикладі $x = (x_1, x_2, x_3, x_4, x_5)^T$, $\lambda = (\lambda_1, \lambda_2, \lambda_3)$ і $\nu = (\nu_1, \nu_2, \nu_3, \nu_4, \nu_5)^T$q. Початкова система (25) із коефіцієнтами (28) має нормальний розв'язок (точність 10^{-4})

$$x = (0,5748;\ -0,0861;\ 0,4094;\ 0,1654;\ 0,2047)^T.$$

Прийнявши тепер у системі (25) $x_2 = 0$, її скорочений варіант встановлює невід'ємний нормальний розв'язок

$$x^+ = (0,6429;\ 0;\ 0,2857;\ 0,3571;\ 0,1429)^T.$$

Розглянемо систему (26)

$$\begin{pmatrix} 2 & 1 & 3 \\ 1 & 1 & 2 \\ 1 & 0 & 1 \\ 0,5 & 0,5 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} - \begin{pmatrix} \nu_1 \\ \nu_3 \\ \nu_4 \\ \nu_5 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Її нормальний розв'язок має від'ємні значення ν_3, ν_4 і ν_5. Якщо прийняти їх за нулі, матимемо систему

$$\begin{aligned}2\lambda_1+\lambda_2+3\lambda_3-\nu_1&=-1\\ \lambda_1+\lambda_2+2\lambda_3&=0\\ \lambda_1+\lambda_3&=0\\ 0,5\lambda_1+0,5\lambda_2+\lambda_3&=0,\end{aligned}$$

для нормального розв'язку якої $\nu_1=1$. Тобто не виконується умова (27): $x_1\cdot\nu_1=0,6429\cdot 1\neq 0$.

Тому слід в системі (25) взяти $x_1=0$ (а раніше нами було прийнято, що і $x_2=0$). Тому нормальний розв'язок прикладу визначається через нормальний додатний розв'язок системи

$$\begin{aligned}x_3+x_4+0,5x_5&=2\\ x_3+0,5x_5&=1\\ 2x_3+x_4+x_5&=3\end{aligned}$$

Він такий: $x^+_{opt}=(0;0;\ 0,8;\ 1;\ 0,4)^T$.

Здійснивши для задачі (28) аналогічні розрахунки із збуреними даними $a_{11}=2,001$ і $b_3=3,01$, а всі інші параметри не змінено, одержимо результат

$$x^+_{opt}=(0;\ 0;\ 0,8027;\ 1;\ 0,4013)^T.$$

Порівнявши цей результат із запропонованим у роботі [11] методом ітеративної регуляризації, що обчислює за 19000 ітерацій значення

$$x^+_{opt}=(0;\ 0;\ 0,8052;\ 0,9916;\ 0,4026),$$

переконуємось у високій ефективності дробово-раціонального методу регуляризації.

Література

[1] Алберт А. Регрессия, псевдоинверсия и рекуррентное оценивание. – М.: Наука, 1977. – 224 с.

[2] Беллман Р. Введение в теорию матриц. – М.: Наука, 1969.– 180 с.

[3] Фадеев Д.К., Фадеева В.Н. Вычислительные методы линейной алгебры. – М.: Физматгиз, 1960.

[4] Беклемишев Д.В. Дополнительные главы линейной алгебры. – М.: Наука, 1983. – 336 с.

[5] Заде Л., Дезоер Ч. Теория линейных систем. – М.: Наука, 1970. – 704 с.

[6] Сявавко М., Рибицька О. Математичне моделювання за умов невизначеності. – Львів: Українські технології, 2000. – 320 с.

[7] Джоунс У., Трон В. Непрерывные дроби. – М.: Мир, 1985. – 414 с.

[8] Сявавко М.С. J-дробова регуляризація лінійних некоректних рівнянь // Укр.мат.журн. – 1996.– 48, № 8. – С. 1130–1143.

[9] Сявавко М.С. Узагальнення формули Крамера // Укр.мат.журн. – 2001. – 53, № 3. – С. 662–679.

[10] Тихонов А.Н., Арсенин В.Я. Методы решения некорректных задач.– М.: Наука, 1986. – 287 с.

[11] Бакушинский А.Б., Гончарский А.В. Итеративные методы решения некорректных задач. – М.: Наука, 1989. – 128 с.

[12] Наттерер Ф. Математические аспекты компьютерной томографии. – М.: Мир, 1990. – 288 с.

[13] Большаков В.Д., Маркузе Ю.И. Голубев В.В. Уравнение геодезических построений. Справочное пособие. – М.: Недра, 1989. – 413 с.

[14] Бейкер Дж., Грейвс-Моррис. Аппроксимации Паде. – М.: Мир, 1986. – 502 с.

THE RATIONAL ALGORITHM OF PSEUDOINVERTION FOR ALMOST DEGENERATE MATRIXES AND ITS APPLICATION TO THE LINEAR PROGRAMMING PROBLEMS

O. Rybytska

Lviv Polytechnic National University
12, S. Bandery Str., Lviv, 79013, Ukraine

The discrete system of linear equations with almost degenerate matrix is investigated. The regularization of such system solution is realized on the base of parameterized continuous fractions theory and Pade Approximants. The proposed algorithm makes it possible to solve linear equations systems with matrix having indefinite rank. The theoretical results are illustrated by a number of test examples including linear programming task.

Keywords: almost degenerate matrix, pseudoinversion, regularisation, continuous fraction, linear programming task.

2000 MSC: 65J20

UDK: 51-96:62-50