УДК 512.83

И. Н. Молчанов, Е. Ф. Галба

Взвешенное псевдообращение комплексных матриц

Взвешенное псевдообращение матриц исследовалось в [1—4]. В статье рассматривается взвешенная псевдобратная матрица для комплексной матрицы. При доказательстве существования и единственности используется теорема Гамильтона—Кэли, что дает возможность представить взвешенную псевдообратную матрицу с помощью коэффициентов характеристического полинома некоторой квадратной матрицы.

Отметим, что применению теоремы Гамильтона—Кэли к вычислению псевдообратной матрицы Мура—Пенроуза [5—6] посвящена работа [7].

Обозначим через R_m и R_n соответственно m- и n- мерные векторные пространства над полем комплексных чисел со скалярными произведениями

$$(u, v)_B = (u, B\bar{v})_m, \quad (u, v)_C = (u, C\bar{v})_n \tag{1}$$

и нормами

$$\| u \|_B = (u, u)_B^{1/2}, \quad \| u \|_C = (u, u)_C^{1/2}, \tag{2}$$

где B и C — эрмитовы положительно-определенные матрицы соответственно порядков m и n,

$$(u, v)_m = \sum_{i=1}^{m} u_i v_i, \quad (u, v)_n = \sum_{i=1}^{n} u_i v_i,$$

а под $\bar{v}$ понимается вектор, компоненты которого комплексно сопряжены с компонентами вектора v.

Пусть A — произвольная прямоугольная комплексная матрица размера $m \times n$.

Теорема 1. *Матрица A^+ размера $n \times m$, удовлетворяющая уравнениям*

$$AA^+A = A; \tag{3}$$

$$A^+AA^+ = A^+; \tag{4}$$

$$(BAA^+)^* = BAA^+; \tag{5}$$

$$(CA^+A)^* = CA^+A, \tag{6}$$

существует и единственна. Она может быть представлена в виде:

$$A^+ = -\alpha_k^{-1} C^{-1} [(A^*BAC^{-1})^{k-1} + \alpha_1 (A^*BAC^{-1})^{k-2} + \ldots + \alpha_{k-1} E] A^*B, \tag{7}$$

где α_p, $p = 1, \ldots, n$, — коэффициенты характеристического полинома $f(\lambda) = \lambda^n + \alpha_1 \lambda^{n-1} + \ldots + \alpha_n = \det | \lambda E - A^*BAC^{-1} |$, а α_k — последний отличный от нуля коэффициент этого полинома.

Доказательство. Доказывать теорему будем по схеме, по которой доказана соответствующая теорема в [6] для псевдообратной матрицы Мура—Пенроуза. Сначала покажем, что уравнения (4), (5) эквивалентны одному уравнению

$$A^+B^{-1}A^{+^*}A^*B = A^+. \tag{8}$$

Уравнение (8) следует из (4), (5). Чтобы в этом убедиться умножим слева уравнение (5) на B^{-1}. Получим $B^{-1}A^{+*}A^*B = AA^+$. После подстановки этого уравнения в (4) получим (8). С другой стороны, из (8) следует, что $BAA^{+}B^{-1}A^{+*}A^{*}B = BAA^{+}B^{-1}(BAA^{+})^{*} = BAA^{+}$ — эрмитова матрица. Далее, подстановкой (5) в (8) получим (4).

Аналогично (3), (6) можно записать в виде одного уравнения

$$CA^{+}AC^{-1}A^{*} = A^{*}. \tag{9}$$

Таким образом, достаточно найти A^{+} из (8), (9). Такая матрица A^{+} существует, если существует матрица S, удовлетворяющая уравнению

$$SA^{*}BAC^{-1}A^{*} = A^{*}. \tag{10}$$

Действительно, пусть

$$A^{+} = C^{-1}SA^{*}B, \tag{11}$$

тогда из (10) следует, что A^{+} удовлетворяет (9). Покажем, что матрица A^{+}, определенная формулой (11), удовлетворяет (8). Из (3) имеем $A^{*}A^{+*}A^{*} = A^{*}$. Умножим это уравнение справа на B, а слева — на $C^{-1}S$, и полученное равенство перепишем в виде

$$C^{-1}SA^{*}BB^{-1}A^{+*}A^{*}B = C^{-1}SA^{*}B. \tag{12}$$

Из (12) следует, что A^{+}, определенная формулой (11), удовлетворяет (8).

Теперь покажем, что существует такая матрица S, которая удовлетворяет уравнению (10). Для этого используем теорему Гамильтона—Кэли, согласно которой любая квадратная матрица удовлетворяет своему характеристическому уравнению. Так как $A^{*}BAC^{-1}$— квадратная матрица порядка n, то имеет место равенство

$$(A^{*}BAC^{-1})^{n} + \alpha_1 (A^{*}BAC^{-1})^{n-1} + \dots + \alpha_{n-1}A^{*}BAC^{-1} + \alpha_n E = 0. \tag{13}$$

Если $\alpha_n \neq 0$, то матрица $A^{*}BAC^{-1}$ имеет обратную. И тогда матрица $S = (A^{*}BAC^{-1})^{-1}$ удовлетворяет (10). Пусть $\alpha_n = 0$, а среди коэффициентов α_p, $p = 1, \dots, n-1$, α_k будет последний отличный от нуля коэффициент полинома $f(\lambda) = \det|\lambda E - A^{*}BAC^{-1}|$ и пусть

$$S = -\alpha_k^{-1}[(A^{*}BAC^{-1})^{k-1} + \alpha_1 (A^{*}BAC^{-1})^{k-2} + \dots + \alpha_{k-1}E], \tag{14}$$

тогда (13) дает

$$S(A^{*}BAC^{-1})^{n-k+1} = (A^{*}BAC^{-1})^{n-k}. \tag{15}$$

Пусть L и H эрмитовы положительно-определенные матрицы порядков m и n соответственно, $\widetilde{L} = L^{1/2}$, $\widetilde{H} = H^{1/2}$; V, F, P — произвольные комплексные $(m \times n)$-матрицы; U, T — произвольные комплексные $(n \times m)$-матрицы. Тогда

$$VP^{*}\widetilde{L}\widetilde{L}P = FP^{*}\widetilde{L}\widetilde{L}P \Rightarrow VP^{*}\widetilde{L} = FP^{*}\widetilde{L}, \tag{16}$$

так как

$$(VP^{*}LP - FP^{*}LP)(V-F)^{*} = (VP^{*}\widetilde{L} - FP^{*}\widetilde{L})(VP^{*}\widetilde{L} - FP^{*}\widetilde{L})^{*}$$

и аналогично

$$UP\widetilde{H}\widetilde{H}P^{*} = TP\widetilde{H}\widetilde{H}P^{*} \Rightarrow UP\widetilde{H} = TP\widetilde{H}. \tag{17}$$

Используя (16), (17), из (15) получим (10), чем и устанавливается существование матрицы A^{+}.

Из (11), (14) следует представление матрицы A^{+} формулой (7).

Чтобы показать, что A^{+} единственна, предположим, что X удовлетворяет (8) и (9), которое перепишем в виде

$$C^{-1}A^{*} = XAC^{-1}A^{*}, \tag{18}$$

а Y удовлетворяет соотношениям

$$Y = C^{-1}A^{*}Y^{*}CY; \tag{19}$$

$$A^{*}B = A^{*}BAY. \tag{20}$$

Равенство (19) получено подстановкой (6) в (4). после умножения (6) слева на C^{-1}, а справа — на Y. Равенство (20) получено из (3), (5).

Учитывая (8), (18), (19), (20), получим

$$X = XB^{-1}X^*A^*B = XB^{-1}X^*A^*BAY = XAY = XAC^{-1}A^*Y^*CY = C^{-1}A^*Y^*CY = Y.$$

Теорема 1 доказана.

Матрица A^+, удовлетворяющая уравнениям (3)—(6), называется взвешенной псевдообратной в отличие от матрицы Мура—Пенроуза, определяемой уравнениями

$$AA^{\neq}A = A,\ A^{\neq}AA^{\neq} = A^{\neq},\ (AA^{\neq})^* = AA^{\neq},\ (A^{\neq}A)^* = A^{\neq}A.$$

Положительно-определенные эрмитовы матрицы B и C называются весами. Если в (5), (6) положить $B = C = E$, то получим псевдообратную матрицу Мура—Пенроуза.

Пусть

$$Ax = b,\ x \in R_n,\ b \in R_m, \tag{21}$$

— система линейных алгебраических уравнений с произвольной прямоугольной комплексной $(m \times n)$-матрицей A. Обозначим через D $(n \times m)$-матрицу, удовлетворяющую уравнениям (3), (5), т. е.

$$ADA = A,\ (BAD)^* = BAD \tag{22}$$

Лемма 1. *Псевдорешение $x_1 = Db$ удовлетворяет условию*

$$\| Ax_1 - b \|_B = \min_{x_1 \in R_n}. \tag{23}$$

Доказательство. Условие (23) равносильно неравенству

$$\| ADb - b \|_B \leqslant \| Ax - b \|_B,\ \forall x \in R_n,\ b \in R_m.$$

Так как $Ax - b = ADb - b + A(x - Db)$, то для выполнения (23) должно выполняться

$$\| ADb - b \|_B \leqslant \| ADb - b + Aw \|_B,$$

где $w = x - Db$.

Пусть $u = ADb - b$, $v = Aw$, тогда $\| u \|_B \leqslant \| u + v \|_B$. Учитывая определение нормы $\| \cdot \|_B$ в (2) получаем $\| u \|_B^2 \leqslant (u + v, B(\overline{u} + \overline{v}))_m = \| u \|_B^2 + \| v \|_B^2 + (u, B\overline{v})_m + (v, B\overline{u})_m$, откуда

$$0 \leqslant \| v \|_B^2 + (\overline{u}, Bv)_m + (\overline{\overline{u}, Bv})_m. \tag{24}$$

Так как $\| v \|_B^2 \geqslant 0$ и $(\overline{u}, Bv)_m = (ADb - b)^* BAw = b^* (D^*A^*BA - BA) w$ зависит от произвольных b и w, то неравенство (24) равносильно требованию $(\overline{u}, Bv)_m = 0$ или

$$D^*A^*BA = BA. \tag{25}$$

Равенство (25) и два равенства (22) эквивалентны. Действительно, так как из первого равенства в (22) $BADA = BA$, то после подстановки второго равенства из (22) в это равенство получаем $D^*A^*BA = BA$. Обратно, после умножения (25) справа на D получаем, что

$$BAD = D^*A^*BAD = (AD)^* BAD,$$

т. е. является эрмитовой матрицей. Кроме того, из (25) и второго равенства из (22) имеем $BA = D^*A^*BA = (BAD)^*A = BADA$, откуда $A = ADA$. Лемма 1 доказана.

Согласно лемме 1 псевдорешение $x_1 = Db$ минимизирует невязку для системы (21). Но это псевдорешение в общем случае неединственно. Общий вид псевдорешения, удовлетворяющего (23), устанавливает такая лемма.

Лемма 2. *Общее псевдорешение, удовлетворяющее (23), определяется по формуле*

$$x = Db + (E - DA)\, y, \tag{26}$$

где E — единичная матрица, y — произвольный вектор.

Доказательство. Во-первых, x удовлетворяет условию (23) так как согласно (22) $A(E - DA) = 0$; во-вторых, каждое решение $x_1 = D_1 b$ удовлетворяющее (23), можно представить формулой (26), взяв $y = (E - DA)^{-1}(D_1 b - Db)$. Но такое представление y возможно, если матрица $E - DA$ имеет обратную. Покажем, что матрица $E - DA$ имеет обратную, если система $Ax = b$ имеет множество псевдорешений, и не имеет обратной, если система имеет единственное псевдорешение.

В силу равенства $ADA = A$, имеем $(E - DA)^k = (E - DA)$, $k = 1, 2, \ldots$ и согласно теореме Гамильтона—Кэли

$$(E - DA) + \alpha_1 (E - DA) + \ldots + \alpha_{n-1}(E - DA) + \alpha_n E = 0. \tag{27}$$

Если $\alpha_n \neq 0$, то матрица $E - DA$ имеет обратную, если же $\alpha_n = 0$, то из (27) следует $E = DA$.

Пусть $\alpha_n = 0$. Покажем, что в этом случае система $Ax = b$ имеет единственное псевдорешение. Так как ранг единичной матрицы равен n, то из равенства $E = DA$ следует, что и ранг матрицы A равен n ($r_{DA} \leqslant \min\{r_D, r_A\}$). Теперь покажем, что при $r_A = n$ несовместная система $Ax = b$ имеет единственное псевдорешение, удовлетворяющее (23). Пусть эта система имеет еще псевдорешение $x_1 = D_1 b$. Тогда для D_1 должно выполняться (25) т. е. $D_1^* A^* BA = BA$ и, следовательно, $D_1^* A^* BA - D^* A^* BA = 0$ или

$$(D_1^* - D^*)\, A^* BA = 0. \tag{28}$$

Так как матрица A по предположению имеет максимально возможный ранг, то матрицы A^*A и AA^* являются неособенными и, следовательно, имеют обратные. Учитывая это, из (28) получим $D_1 = D$.

Таким образом, если ранг матрицы A равен n, то существует единственное псевдорешение. В противном случае $E - DA$ имеет обратную и общее псевдорешение определяется формулой (26). Лемма 2 доказана.

Следствие. *Если ранг матрицы равен n, т. е. числу неизвестных, то несовместная система $Ax = b$ имеет единственное псевдорешение, удовлетворяющее (23), которое определяется по формуле $x = Db$. В этом случае псевдообратная матрица определяется уравнениями (22).*

Замечание. Если ранг матрицы A равен n, то $E - DA = 0$, так что формула (26), определяющая общее псевдорешение, верна и для этого случая.

Теорема 2. *Псевдорешение для системы (21) $x = A^+ b$, где матрица A^+ определена уравнениями (3)—(6), удовлетворяет условиям*

$$\| Ax - b \|_B = \min_{x \in R_n}, \quad \| x \|_C = \min_{x \in R_n}. \tag{29}$$

Доказательство. В лемме 1 установлено, что псевдорешение $x_1 = Db$ с матрицей D, удовлетворяющей (3), (5), минимизирует невязку для системы (21) в смысле нормы $\| \cdot \|_B$, определенной в (2). В лемме 2 установлено, что общее псевдорешение, удовлетворяющее (23), определяется по формуле (26). Для доказательства теоремы осталось показать, что если матрица D кроме условий (22), удовлетворяет условиям (4), (6), т. е.

$$DAD = D, \quad (CDA)^* = CDA, \tag{30}$$

то $x = Db$ есть псевдорешение, удовлетворяющее, кроме первого, и второму условию из (29). Для этого должно выполняться неравенство

$$\| Db \|_C \leqslant \| Db + (E - DA)\, y \|_C, \quad \forall b, \; y \in R_n. \tag{31}$$

Пусть $u = Db$, $v = (E - DA)y$, тогда $\|u\|_C \leqslant \|u + v\|_C$ и, учитывая определение нормы $\|\cdot\|_C$ в (2), получим

$$\|u\|_C^2 \leqslant (u + v, C(\bar{u} + \bar{v}))_n = \|u\|_C^2 + \|v\|_C^2 + (u, C\bar{v})_n + (v, C\bar{u})_n,$$

откуда

$$0 \leqslant \|v\|_C^2 + (\bar{u}, Cv)_n + \overline{(\bar{u}, Cv)}_n. \qquad (32)$$

Так как $\|v\|_C^2 \geqslant 0$ и

$$(\bar{u}, Cv)_n = b^* D^* C (E - DA) y = b^* (D^* C - D^* CDA) y$$

зависит от произвольных b и y, то неравенство (32) равнозначно требованию $(u, Cv)_n = 0$ или

$$D^* C = D^* CDA. \qquad (33)$$

Равенство (33) и два равенства (30) эквивалентны. Действительно, так как из первого равенства в (30) следует $D^*C = D^*A^*D^*C$, по после подстановки второго равенства из (30) в это равенство получаем $D^*C = D^*CDA$. Обратно, после умножения (33) слева на A^* получим, что $A^*D^*C = (CDA)^* = A^*D^*CDA = (DA)^*CDA$ — эрмитова матрица. Кроме того, из (33) и второго уравнения (30) $D^*C = D^*CDA = D^*A^*D^*C$, т. е. $D^* = D^*A^*D^*$ или $D = DAD$. Следовательно, если $D = A^+$, то $x = Db$ удовлетворяет (29), т. е. условию теоремы.

Таким образом, псевдорешение $x = A^+b$, где A^+— взвешенная псевдообратная матрица, удовлетворяющая (3)—(6), минимизирует невязку системы (21) в смысле нормы $\|\cdot\|_B$ и является минимальным в смысле нормы $\|\cdot\|_C$.

1. *Chipman J. S.* On least squares with insufficient observation.— J. Amer. Statist. Assoc., 1964, 59, N 308, p. 1078—1111.
2. *Ward J. F., Boullion T. L., Lewis T. O.* A note on the oblique matrix pseudoinverse.— SIAM J. Appl. Math., 1971, 20, N 2, p. 173—175.
3. *Milne R. D.* An oblique matrix pseudoinverse.— Ibid., 1968 16, N 5, p. 931—944.
4. *Ward J. F., Boullion T. L., Lewis T. O.* Weighted pseudoinverses with singular weights. — Ibid., 1971, 21, N 3, p. 480—482.
5. *Moore E. H.* Abstract.— Bull. Amer. Math. Soc., 1920, 26, p. 394—395.
6. *Penrose R.* A generalized inverse for matrices.— Proc. Cambridge Phil. Soc., 1955, 51, N 3, p. 406—413.
7. *Decell H. P.* An application of the Cayley — Hamilton theorem to generalized matrix inversion.— SIAM Rev., 1965, 7, N 4, p. 526—528.

Институт кибернетики
АН УССР

Поступила в редакцию
11.06.81