ОБОБЩЕНИЕ ФОРМУЛЫ ШЕРМАНА—МОРРИСОНА И БЛИЗКИЕ ВОПРОСЫ

В.Н.Малоземов
(Доклад на семинаре 29.12.98)

1^o. **ТЕОРЕМА 1.** *Пусть A — обратимая $(n\times n)$-матрица, X, Y — $(n\times r)$-матрицы, где $r\in 1:n$, и E_r — единичная матрица порядка r. Матрица*

$$D = A - XY^T \tag{1}$$

обратима тогда и только тогда, когда обратима матрица r-го порядка $P = E_r - Y^T A^{-1} X$. При этом

$$D^{-1} = A^{-1} + A^{-1} X P^{-1} Y^T A^{-1}. \tag{2}$$

Доказательство. Возьмем решение u_0 однородной системы $Du = \mathbf{0}$. Согласно (1) имеем $Au_0 = XY^T u_0$, так что

$$u_0 = A^{-1} X Y^T u_0. \tag{3}$$

Умножив на Y^T слева, получим

$$(E_r - Y^T A^{-1} X) Y^T u_0 = \mathbf{0}, \text{ или } PY^T u_0 = \mathbf{0}.$$

Если матрица P обратима, то $Y^T u_0 = \mathbf{0}$, откуда в силу (3) следует, что $u_0 = \mathbf{0}$. Установлено, что обратимость матрицы P гарантирует обратимость матрицы D.

Наоборот, пусть обратимой будет матрица D. Покажем, что и матрица P обратима. Возьмем решение v_0 однородной системы $Pv = \mathbf{0}$. Согласно определению матрицы P имеем

$$v_0 = Y^T A^{-1} X v_0.$$

Положим $u_0 = A^{-1} X v_0$. Тогда $Au_0 = Xv_0$, $v_0 = Y^T u_0$ и

$$Du_0 = Au_0 - XY^T u_0 = \mathbf{0}.$$

В силу обратимости матрицы D получаем $u_0 = \mathbf{0}$. Поскольку $v_0 = Y^T u_0$, то и $v_0 = \mathbf{0}$. Таким образом, из обратимости матрицы D следует обратимость матрицы P.

Выведем формулу (2). Согласно (1)

$$DA^{-1}X = X - X(Y^T A^{-1} X) = X(E_r - Y^T A^{-1} X) = XP,$$

$$DA^{-1}XP^{-1}Y^T = XY^T = A - D.$$

Умножив последнее равенство слева на D^{-1}, получим

$$A^{-1}XP^{-1}Y^T = D^{-1}A - E_n,$$

или

$$D^{-1}A = E_n + A^{-1}XP^{-1}Y^T. \tag{4}$$

Остается умножить (4) на A^{-1} справа. Теорема доказана.

При $r = 1$ приходим к следующему утверждению.

Пусть x, y — n-мерные вектор-столбцы и A — обратимая матрица порядка n. Матрица $D = A - xy^T$ обратима тогда и только тогда, когда число $\alpha = 1 - \langle y, A^{-1}x\rangle$ отлично от нуля. При этом

$$D^{-1} = A^{-1} + \alpha^{-1}A^{-1}xy^T A^{-1}. \tag{5}$$

Формула (5) называется *формулой Шермана—Моррисона.* Она эффективно используется при обращении матриц.

ПРИМЕР. В обратимой матрице n-го порядка A заменим p-й столбец a_p на вектор b. Новая матрица D примет вид

$$D = A - (a_p - b)e_p^T,$$

где e_p — p-й орт. Когда матрица D будет обратимой?

В данном случае

$$\alpha = 1 - \langle e_p, A^{-1}(a_p - b)\rangle = \langle e_p, A^{-1}b\rangle.$$

Обозначим $z = A^{-1}b$. Тогда $\alpha = z[p]$. Критерием обратимости матрицы D является условие $z[p] \neq 0$. При этом

$$D^{-1} = A^{-1} + (z[p])^{-1}A^{-1}(a_p - b)e_p^T A^{-1} =$$

$$= A^{-1} + (z[p])^{-1}(e_p - z)e_p^T A^{-1}. \tag{6}$$

Обозначим через A_p^{-1} p-ю строку матрицы A^{-1} и перепишем матричное равенство (6) в виде равенства строк. Учитывая, что $e_p^T A^{-1} = A_p^{-1}$, получаем

$$D_p^{-1} = A_p^{-1} + (z[p])^{-1}(1 - z[p])A_p^{-1} = (z[p])^{-1}A_p^{-1}$$

и при $k \neq p$

$$D_k^{-1} = A_k^{-1} - z[k]D_p^{-1}.$$

Это известные формулы пересчета обратной матрицы. Строка D_p^{-1} называется *рабочей строкой.*

2^o. Рассмотрим вопрос об обратимости блочной матрицы G вида

$$G = \begin{pmatrix} A & X \\ Y^T & H \end{pmatrix}. \tag{7}$$

ТЕОРЕМА 2. *Предположим, что матрица A имеет обратную. Критерием обратимости матрицы G является обратимость матрицы r-го порядка $F = H - Y^T A^{-1} X$. При этом*

$$G^{-1} = \begin{pmatrix} A^{-1} + A^{-1}XF^{-1}Y^TA^{-1} & -A^{-1}XF^{-1} \\ -F^{-1}Y^TA^{-1} & F^{-1} \end{pmatrix}. \tag{8}$$

Доказательство. Возьмем решение (u_0, v_0) однородной системы

$$\begin{aligned} Au + Xv &= \mathbf{0}, \\ Y^Tu + Hv &= \mathbf{0}. \end{aligned} \tag{9}$$

Имеем $u_0 = -A^{-1}Xv_0,\quad Hv_0 - Y^TA^{-1}Xv_0 = \mathbf{0}$. Последнее равенство можно переписать в виде $Fv_0 = \mathbf{0}$. Если матрица F обратима, то $v_0 = \mathbf{0}$, а тогда и $u_0 = \mathbf{0}$. Установлено, что из обратимости матрицы F следует обратимость матрицы G.

Предположим теперь, что матрица G обратима, и проверим обратимость матрицы F. Возьмем решение v_0 однородной системы $Fv = \mathbf{0}$. Согласно определению матрицы F имеем

$$Hv_0 - Y^TA^{-1}Xv_0 = \mathbf{0}.$$

Обозначим $u_0 = -A^{-1}Xv_0$. В этом случае

$$Au_0 + Xv_0 = \mathbf{0}, \quad Hv_0 + Y^T u_0 = \mathbf{0},$$

т. е. пара (u_0, v_0) удовлетворяет однородной системе (9). В силу обратимости матрицы G получаем $u_0 = \mathbf{0}$, $v_0 = \mathbf{0}$. Условие $v_0 = \mathbf{0}$ гарантирует обратимость матрицы F.

Выведем формулу (8). Пусть

$$G^{-1} = \begin{pmatrix} K & B \\ C^T & L \end{pmatrix}. \tag{10}$$

Тогда

$$\begin{pmatrix} A & X \\ Y^T & H \end{pmatrix} \begin{pmatrix} K & B \\ C^T & L \end{pmatrix} = \begin{pmatrix} E_n & \mathbf{0} \\ \mathbf{0} & E_r \end{pmatrix}.$$

Распишем это равенство подробно:

$$\begin{array}{cc} AK + XC^T = E_n, & Y^T K + HC^T = \mathbf{0}, \\ AB + XL = \mathbf{0}, & Y^T B + HL = E_r. \end{array}$$

Получаем

$$\begin{array}{rr} K = A^{-1} - A^{-1}XC^T, & Y^T A^{-1} - Y^T A^{-1} X C^T + HC^T = \mathbf{0}, \\ B = -A^{-1}XL, & -Y^T A^{-1} X L + HL = E_r. \end{array}$$

Поскольку $Y^T A^{-1} + FC^T = \mathbf{0}$, то

$$C^T = -F^{-1}Y^T A^{-1}, \quad K = A^{-1} + A^{-1}XF^{-1}Y^T A^{-1}.$$

Далее $FL = E_r$, так что

$$L = F^{-1}, \quad B = -A^{-1}XF^{-1}.$$

Теорема доказана.

Формула (8) называется *формулой Фробениуса.* Она используется в методах окаймления.

3^o. Поставим обратную задачу. Пусть матрица G вида (7) имеет обратную матрицу G^{-1} вида (10). Когда обратима матрица A?

ТЕОРЕМА 3. *Для обратимости матрицы A необходима и достаточна обратимость матрицы r-го порядка L. При этом*

$$A^{-1} = K - BL^{-1}C^T. \tag{11}$$

Доказательство. Пусть L — обратимая матрица. Из равенства

$$AB + XL = \mathbf{0} \tag{12}$$

следует, что $X = -ABL^{-1}$. Подставляя это выражение в равенство $AK + XC^T = E_n$, получаем

$$A(K - BL^{-1}C^T) = E_n.$$

Последнее гарантирует и обратимость матрицы A, и справедливость формулы (11).

Допустим теперь, что L — вырожденная матрица. Тогда найдется ненулевой вектор u_0, такой, что $Lu_0 = \mathbf{0}$. Умножив обе части (12) на u_0, получим

$$ABu_0 = \mathbf{0}. \tag{13}$$

Из равенства $Y^TB+HL = E_r$ следует, что $Y^TBu_0 = u_0$. В частности, вектор Bu_0 отличен от нуля. Учитывая (13), заключаем, что A — особая матрица. Таким образом, если A — невырожденная матрица, то необходимо невырожденной будет и L.

Теорема доказана.

Формула (11) установлена В. А. Даугавет.

4^o. Пусть A — вырожденная симметричная $(n\times n)$-матрица и b — n-мерный вектор-столбец. Составим матрицу

$$S = \begin{pmatrix} A & b \\ b^T & 0 \end{pmatrix}.$$

ТЕОРЕМА 4. *Для того чтобы существовала обратная матрица S^{-1}, необходимо и достаточно, чтобы выполнялись следующие условия:*

$$\mathrm{rank}(A) = n - 1, \quad b \notin \mathrm{Im}(A). \tag{14}$$

Доказательство. Необходимость. По условию

$$\mathrm{rank}(S) = n + 1.$$

Поскольку матрица A — вырожденная, то $\mathrm{rank}(A) \le n-1$. Если $\mathrm{rank}(A) \le n-2$, то ранг матрицы (A, b) не превосходит $n-1$, а тогда $\mathrm{rank}(S) \le n$. Это противоречит тому, что $\mathrm{rank}(S) = n+1$. Значит, $\mathrm{rank}(A) = n-1$.

Если $b \in \mathrm{Im}(A)$, т. е. $b = Ay_0$, то ранг матрицы (A, b) равен $n-1$. Снова получаем, что $\mathrm{rank}(S) \le n$, что противоречит условию $\mathrm{rank}(S) = n+1$.

Достаточность. Покажем, что при выполнении условий (14) однородная система

$$\begin{array}{llcl} Au & +bt & = & \mathbf{0}, \\ b^T u & & = & 0 \end{array}$$

имеет только нулевое решение. Возьмем любое решение (u_0, t_0) этой системы. Тогда $t_0 = 0$, ибо иначе $A(-u_0/t_0) = b$, что противоречит условию $b \notin \mathrm{Im}(A)$. Имеем

$$Au_0 = \mathbf{0}, \quad b^T u_0 = 0. \tag{15}$$

Обозначим $\mathcal{L} = \{\, u \,|\, Au = \mathbf{0} \,\}$. Как известно [2],

$$\dim(\mathcal{L}) = n - \mathrm{rank}(A).$$

В нашем случае $\mathrm{rank}(A) = n-1$, так что $\dim(\mathcal{L}) = 1$. Возьмем ненулевое решение u_* системы $Au = \mathbf{0}$. Поскольку $u_0 \in \mathcal{L}$, то $u_0 = \lambda_0 u_*$. Проверим, что $\lambda_0 = 0$. Этим завершится доказательство теоремы.

Допустим противное: $\lambda_0 \ne 0$. Согласно (15)

$$0 = b^T u_0 = \lambda_0 b^T u_*,$$

откуда следует, что $b^T u_* = 0$. В силу одномерности подпространства $\mathcal{L}$ получаем $b \in \mathcal{L}^\perp$. Но, как известно [2],

$$\mathcal{L}^\perp = \{\, x = A^T y \,|\, y \in \mathbb{R}^n \,\}.$$

Значит, $b = A^T y_0$. Матрица A — симметричная, так что $b = Ay_0$, а это противоречит условию $b \notin \mathrm{Im}(A)$. Теорема доказана.

Теорема 4 сформулирована Ф. Монако.

ЛИТЕРАТУРА

1. Broyden C. G. Basic Matrices. An Introduction to Matrix Theory and Practice. London, 1975.

2. Малоземов В. Н. Линейная алгебра без определителей. Квадратичная функция. СПб., 1997.