Записки научных
семинаров ПОМИ
Том 323, 2005 г.

Ю. А. Альпин, С. Н. Ильин

МАТРИЧНОЕ УРАВНЕНИЕ $AX - YB = C$ И СМЕЖНЫЕ ВОПРОСЫ

1. Введение и вспомогательные результаты

В 1952 г. В. Е. Рот опубликовал следующую теорему [1]: матричное уравнение $AX - YB = C$ имеет решение тогда и только тогда, когда матрицы $\begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$ и $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ эквивалентны. Другое доказательство теоремы Рота было дано в [2], оно помещено также в [3].

Основным результатом данной статьи является теорема 1, с помощью которой дается новое доказательство теоремы Рота, находится новый критерий разрешимости уравнения $AX - YB = C$ и выводится формула его частного решения, а также находится наименьший из порядков квадратных невырожденных матриц, включающих произвольно фиксированную прямоугольную матрицу в качестве подматрицы.

Наши рассуждения проводятся в терминах скелетных разложений матриц и псевдообратных матриц. Приведем необходимые сведения об этих объектах.

Пусть M – матрица ранга $n > 0$. Скелетным разложением матрицы M называется любое ее представление в виде произведения $M = LR$, в котором n – число столбцов L и число строк R. Ранги сомножителей скелетного разложения необходимо равны $n = \mathrm{rg}\, M$. Мы будем пользоваться следующими простыми свойствами скелетного разложения.

Предложение 1. *Матрица L (матрица R) может служить левым (правым) сомножителем скелетного разложения матрицы M тогда и только тогда, когда столбцы L (строки R) линейно независимы, а их линейная оболочка совпадает с линейной оболочкой столбцов (строк) матрицы M. Если известен левый (правый) сомножитель скелетного разложения, то правый (левый) сомножитель является единственным решением матричного уравнения $LZ = M$ ($ZR = M$).*

Справедливость этого предложения вытекает из определения скелетного разложения и стандартной теории систем линейных уравнений.

Предложение 2. *Пусть дано разложение $M = LR$, сомножители которого удовлетворяют условию* $\mathrm{rg}\ L = \mathrm{rg}\ R = n$, *где n – число столбцов L и число строк R. Тогда* $\mathrm{rg}\ M = n$, *то есть данное разложение матрицы M является скелетным.*

Доказательство. Чтобы проверить это утверждение, выпишем для произведения LR неравенства Сильвестра (см., напр., [4]):

$$\mathrm{rg}\ L + \mathrm{rg}\ R - n \leqslant \mathrm{rg}\ LR \leqslant \min(\mathrm{rg}\ L, \mathrm{rg}\ R).$$

Подставив в эти неравенства указанные в предложении значения рангов, получим $n \leqslant \mathrm{rg}\ LR \leqslant n$, то есть $\mathrm{rg}\ LR = n$.

В следующих предложениях собраны свойства псевдообратных матриц, полезные для дальнейшего изложения.

Предложение 3 (см., напр., [3]).*Если матричное уравнение $AX = C$ имеет решение, то одним из решений является матрица A^+C. Аналогично, если уравнение $YB = C$ имеет решение, то одним из решений является матрица CB^+.*

Предложение 4. *Если $M = LR$ – скелетное разложение, то*

$$M^+ = R^+L^+, \tag{1}$$

$$L = MR^+, \quad R = L^+M, \tag{2}$$

$$LL^+ = MM^+, \quad R^+R = M^+M. \tag{3}$$

Доказательство. Равенство (1) можно найти в [4], а (2) и (3) легко выводятся из элементарных свойств псевдообратных матриц, приведенных там же.

2. Основная теорема

Теорема 1. *Пусть блочная матрица* $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ *удовлетворяет условиям*

$$\mathrm{rg}\,(A, B) = \mathrm{rg}\ A + \mathrm{rg}\ B, \tag{4}$$

$$\mathrm{rg}\begin{pmatrix} A \\ C \end{pmatrix} = \mathrm{rg}\ A + \mathrm{rg}\ C. \tag{5}$$

Тогда следующие утверждения эквивалентны:
$\mathrm{rg}\ M = \mathrm{rg}\ A + \mathrm{rg}\ B + \mathrm{rg}\ C$;
уравнение $CX + YB = D$ *имеет решение.*

Доказательство состоит из трех частей.

1. Вначале рассмотрим простой случай, когда B или C – нулевая подматрица. Если B – нулевая подматрица, то, с учетом (5), утверждение теоремы сводится к следующему:

$$\mathrm{rg}\begin{pmatrix} A & 0 \\ C & D \end{pmatrix} = \mathrm{rg}\begin{pmatrix} A \\ C \end{pmatrix} \quad \Leftrightarrow \text{ уравнение } CX = D \text{ имеет решение.}$$

Но оно несомненно верно. Аналогично, если C – нулевая подматрица, то, с учетом (4), утверждение теоремы равносильно следующему:

$$\mathrm{rg}\begin{pmatrix} A & B \\ 0 & D \end{pmatrix} = \mathrm{rg}\,(A, B) \ \Leftrightarrow \text{ уравнение } YB = D \text{ имеет решение.}$$

Но и оно очевидно верно.

2. Теперь пусть B и C ненулевые, но A – нулевая подматрица. Фиксируем скелетные разложения

$$B = B_1 B_2, \quad C = C_1 C_2 \tag{6}$$

и рассмотрим матрицы

$$M = \begin{pmatrix} 0 & B \\ C & D \end{pmatrix}, \quad M_1 = \begin{pmatrix} 0 & B_1 \\ C_1 & DB_2^+ \end{pmatrix}.$$

Нетрудно заметить, что $\mathrm{rg}\ M_1 = \mathrm{rg}\ B_1 + \mathrm{rg}\ C_1 = \mathrm{rg}\ B + \mathrm{rg}\ C$. Поскольку $\begin{pmatrix} B_1 \\ DB_2^+ \end{pmatrix} = \begin{pmatrix} B \\ D \end{pmatrix} B_2^+$, то, очевидно, линейная оболочка столбцов M_1 содержится в линейной оболочке столбцов M.

Предположим, что оболочки столбцов матриц M и M_1 совпадают, то есть выполняется равенство $\mathrm{rg}\ M = \mathrm{rg}\ B + \mathrm{rg}\ C$. Тогда по предложению 1 матрица M_1 может служить левым сомножителем скелетного разложения матрицы M, и уравнение $M_1 Z = M$ имеет, и притом единственное, решение. Разобьем матрицу Z на блоки так, чтобы было возможно поблочное умножение матриц. Тогда уравнение приобретает вид

$$\begin{pmatrix} 0 & B_1 \\ C_1 & DB_2^+ \end{pmatrix} \begin{pmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{pmatrix} = \begin{pmatrix} 0 & B \\ C & D \end{pmatrix}. \tag{7}$$

Это уравнение равносильно системе из четырех уравнений

$$\begin{cases} B_1 Z_{21} = 0, \\ B_1 Z_{22} = B, \\ C_1 Z_{11} + D B_2^{+} Z_{21} = C, \\ C_1 Z_{12} + D B_2^{+} Z_{22} = D. \end{cases}$$

Решим эту систему. Поскольку столбцы B_1 линейно независимы, то $Z_{21} = 0$. Принимая во внимание равенства (6) и предложение 1, видим, что второе уравнение имеет решение $Z_{22} = B_2$, а третье – решение $Z_{11} = C_2$, $Z_{21} = 0$. С учетом полученного ранее значения $Z_{22} = B_2$, четвертое уравнение принимает вид $C_1 Z_{12} + D B_2^{+} B_2 = D$, или

$$C_1 Z_{12} = D(I - B_2^{+} B_2). \tag{8}$$

Поскольку уравнение (7) имеет решение, то и уравнение (8) имеет (единственное) решение. Согласно предложению 3, оно находится по формуле $Z_{12} = C_1^{+} D(I - B_2^{+} B_2)$. Подставляя его в (8) и используя соотношения (3), получаем равенство

$$CC^{+} D(I - B^{+} B) + DB^{+} B = D.$$

Но это равенство означает, что уравнение $CX + YB = D$ имеет решение

$$X = C^{+} D(I - B^{+} B), \quad Y = DB^{+}.$$

Теперь пусть известно, что уравнение $CX + YB = D$ имеет решение. Если X_0, Y_0 – решение, то положим $P = C_2 X_0$, $Q = Y_0 B_1$. Нетрудно проверить, что тогда справедливо разложение

$$M = \begin{pmatrix} 0 & B \\ C & D \end{pmatrix} = \begin{pmatrix} 0 & B_1 \\ C_1 & Q \end{pmatrix} \begin{pmatrix} C_2 & P \\ 0 & B_2 \end{pmatrix}. \tag{9}$$

Ранги сомножителей в (9) равны числу столбцов первого сомножителя и числу строк второго сомножителя, то есть оба равны $\mathrm{rg}\ B + \mathrm{rg}\ C$. По предложению 2 это значит, что разложение (9) скелетное, причем

$$\mathrm{rg}\ M = \mathrm{rg} \begin{pmatrix} 0 & B \\ C & D \end{pmatrix} = \mathrm{rg}\ B + \mathrm{rg}\ C.$$

3. Рассмотрим оставшийся случай, когда блоки A, B, C ненулевые. Добавим к списку (6) скелетное разложение $A = A_1 A_2$.

Наше рассуждение будет следовать схеме предыдущего пункта, но теперь матрица M_1 расширяется и приобретает вид

$$M_1 = \begin{pmatrix} A_1 & 0 & B_1 \\ 0 & C_1 & DB_2^+ \end{pmatrix}.$$

Блочное строение матрицы M_1 и условие (4) обеспечивают линейную независимость ее столбцов. Стало быть $\operatorname{rg} M_1 = \operatorname{rg} A + \operatorname{rg} B + \operatorname{rg} C$. Докажем, что линейная оболочка столбцов M_1 содержится в линейной оболочке столбцов M. Рассмотрим равенство

$$\begin{pmatrix} A \\ C \end{pmatrix} = \begin{pmatrix} A_1 & 0 \\ 0 & C_1 \end{pmatrix} \begin{pmatrix} A_2 \\ C_2 \end{pmatrix}.$$

Оно является скелетным разложением, поскольку, согласно условию (5), строки матрицы $\begin{pmatrix} A_2 \\ C_2 \end{pmatrix}$ линейно независимы, а их линейная оболочка очевидным образом совпадает с оболочкой строк $\begin{pmatrix} A \\ C \end{pmatrix}$. Но тогда, по предложению 1, оболочка столбцов $\begin{pmatrix} A_1 & 0 \\ 0 & C_1 \end{pmatrix}$ совпадает с оболочкой столбцов $\begin{pmatrix} A \\ C \end{pmatrix}$, и, как теперь ясно, оболочка столбцов M_1 содержится в оболочке столбцов M.

Предположим, что линейные оболочки столбцов M_1 и M совпадают, а значит выполняется равенство $\operatorname{rg} M_1 = \operatorname{rg} M = \operatorname{rg} A + \operatorname{rg} B + \operatorname{rg} C$. Тогда уравнение $M_1 Z = M$, или, в блочной записи,

$$\begin{pmatrix} A_1 & 0 & B_1 \\ 0 & C_1 & DB_2^+ \end{pmatrix} \begin{pmatrix} Z_{01} & Z_{02} \\ Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

имеет (единственное) решение. Равносильная система уравнений в блоках такова:

$$\begin{cases} A_1 Z_{01} + B_1 Z_{21} = A, \\ A_1 Z_{02} + B_1 Z_{22} = B, \\ C_1 Z_{11} + DB_2^+ Z_{21} = C, \\ C_1 Z_{12} + DB_2^+ Z_{22} = D. \end{cases}$$

Несложный анализ, подобный проведенному в пункте 2, дает следующие значения блоков:

$$Z_{01} = A_2, \quad Z_{21} = 0, \quad Z_{02} = 0, \quad Z_{22} = B_2, \quad Z_{11} = C_2.$$

С учетом последнего равенства, четвертое уравнение системы имеет вид $C_1Z_{12} + DB_2^+B_2 = D$. Разрешимость этого уравнения влечет разрешимость уравнения $CX+YB=D$ точно так же, как в пункте 2.

Теперь пусть известно, что уравнение $CX+YB=D$ имеет решение X_0, Y_0. Тогда верно разложение

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A_1 & 0 & B_1 \\ 0 & C_1 & Q \end{pmatrix} \begin{pmatrix} A_2 & 0 \\ C_2 & P \\ 0 & B_2 \end{pmatrix}, \tag{10}$$

где $P = C_2X_0$, $Q = Y_0B_1$. Ранги сомножителей в (10), благодаря их блочному строению и условиям (4), (5), равны числу столбцов первого сомножителя и числу строк второго сомножителя, то есть равны $\mathrm{rg}\ A + \mathrm{rg}\ B + \mathrm{rg}\ C$. Согласно предложению 2, разложение (10) скелетное, и

$$\mathrm{rg}\ M = \mathrm{rg}\begin{pmatrix} 0 & B \\ C & D \end{pmatrix} = \mathrm{rg}\ A + \mathrm{rg}\ B + \mathrm{rg}\ C.$$

Теорема доказана.

3. Об уравнении $CX + YB = D$

Предположим, что наш интерес состоит в решении матричного уравнения $CX+YB=D$. Тогда из доказательства теоремы 1 можно извлечь следующий результат.

Теорема 2. *Матричное уравнение $CX+YB=D$ имеет решение тогда и только тогда, когда матрицы $\begin{pmatrix} 0 & B \\ C & D \end{pmatrix}$ и $\begin{pmatrix} 0 & B \\ C & 0 \end{pmatrix}$ одного ранга. Если уравнение $CX+YB=D$ разрешимо, то решением являются матрицы*

$$X = C^+D(I - B^+B), \quad Y = DB^+. \tag{11}$$

Доказательство. Матрица $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$, где $A = 0$, очевидно удовлетворяет условиям (4), (5) теоремы 1. Применяя теорему 1 к этой матрице и учитывая, что $\mathrm{rg}\ B + \mathrm{rg}\ C = \mathrm{rg}\begin{pmatrix} 0 & B \\ C & 0 \end{pmatrix}$, получаем первое утверждение теоремы 2.

Формулы (11) выведены в доказательстве теоремы 1 в предположении, что B, C – ненулевые блоки. Но они верны и тогда, когда $B = 0$ или $C = 0$. Действительно, в первом из этих случаев для уравнения $CX = D$ формулы (11) дают $X = C^+D$, $Y = 0$, а во втором случае для уравнения $YB = D$ они дают $X = 0$, $Y = DB^+$, и доказательство завершается с использованием предложения 3.

Отметим, что первое утверждение теоремы 2 совпадает (с точностью до обозначений) с теоремой Рота. Формулы (11), выражающие частное решение в явном виде, по-видимому, не были известны. Согласно этим формулам, уравнение $CX + YB = D$ разрешимо тогда и только тогда, когда

$$CC^+D(I - B^+B) + DB^+B = D. \tag{12}$$

Придадим критерию разрешимости более симметричный вид.

Теорема 3. *Уравнение $CX+YB = D$ имеет решение тогда и только тогда, когда*

$$(I - CC^+)D(I - B^+B) = 0. \tag{13}$$

Доказательство. Эквивалентность равенств (12) и (13) легко устанавливается непосредственной проверкой.

Добавим, в заключение этой части, что открыв скобки в левой части (12) и перегруппировав члены, можно получить равенство $CC^+D + (I - CC^+)DB^+B = D$. Таким образом, в случае разрешимости исследуемого уравнения матрицы $X = C^+D$ и $Y = (I - CC^+)DB^+$, как и матрицы (11), являются его решением.

4. Невырожденное окаймление наименьшего порядка

В качестве еще одного приложения теоремы 1 решим следующую задачу. Дана $m \times n$-матрица A. Каков наименьший порядок квадратной невырожденной матрицы, содержащей A в качестве подматрицы? Без ограничения общности мы можем искать окаймляющую матрицу M в виде

$$M = \begin{array}{cc} & \begin{array}{cc} n & l \end{array} \\ \begin{array}{c} m \\ k \end{array} & \begin{pmatrix} A & B \\ C & D \end{pmatrix} \end{array}. \tag{14}$$

В случае $m = n$ ответ $2m - \mathrm{rg}\, A$ получен в [5] с помощью формулы Фробениуса для обращения блочной матрицы. В общем случае ответ дает следующая

Теорема 4. *Наименьший из порядков квадратных невырожденных матриц, содержащих A в качестве подматрицы, равен $m+n-\operatorname{rg} A$. Матрица M, определенная в (14), где $m+k=n+l=m+n-\operatorname{rg} A$, невырождена тогда и только тогда, когда $\operatorname{rg}(A,B)=m$ и $\operatorname{rg}\begin{pmatrix} A \\ C \end{pmatrix} = n$.*

Доказательство. В случае, когда $\operatorname{rg} A = \min(m,n)$, наименьший порядок искомой невырожденной матрицы, очевидно, равен $\max(m,n) = m+n-\min(m,n)$. Таким образом, можно считать, что $\operatorname{rg} A < \min(m,n)$. Очевидно, что $\operatorname{rg}(A,B) \leqslant \operatorname{rg} A + l$; поэтому если $l < m - \operatorname{rg} A$, то $\operatorname{rg}(A,B) < m$, и, ввиду линейной зависимости строк подматрицы (A,B), матрица M вырождена. Аналогично, если $k < n - \operatorname{rg} A$, то столбцы подматрицы $\begin{pmatrix} A \\ C \end{pmatrix}$ линейно зависимы, что влечет вырожденность M.

Пусть теперь $l = m - \operatorname{rg} A$ и $k = n - \operatorname{rg} A$. Заметим, что при указанных значениях параметров k и l матрица M – квадратная, и ее порядок равен $m+n-\operatorname{rg} A$. Ясно, что для невырожденности M необходимо выполнение равенств $\operatorname{rg}(A,B) = m$ и $\operatorname{rg}\begin{pmatrix} A \\ C \end{pmatrix} = n$ и что блоки B и C, удовлетворяющие требуемым условиям, существуют. Тогда

$$m = \operatorname{rg}(A,B) \leqslant \operatorname{rg} A + \operatorname{rg} B \leqslant \operatorname{rg} A + l = m,$$

$$n = \operatorname{rg}\begin{pmatrix} A \\ C \end{pmatrix} \leqslant \operatorname{rg} A + \operatorname{rg} C \leqslant \operatorname{rg} A + k = n.$$

Следовательно,

$$\operatorname{rg}(A,B) = \operatorname{rg} A + \operatorname{rg} B, \quad \operatorname{rg} B = l,$$

$$\operatorname{rg}\begin{pmatrix} A \\ C \end{pmatrix} = \operatorname{rg} A + \operatorname{rg} C, \quad \operatorname{rg} C = k.$$

Как видно, матрица M удовлетворяет условиям (4), (5) теоремы 1. Матрица B ранга l содержит l линейно независимых строк длины l, поэтому для любой $k \times l$-матрицы D уравнение $YB = D$ имеет решение. Так как последнее уравнение – частный случай уравнения $CX + YB = D$, то применяя теорему 1, получаем

$$\operatorname{rg} M = \operatorname{rg} A + \operatorname{rg} B + \operatorname{rg} C = \operatorname{rg} A + (m - \operatorname{rg} A) + (n - \operatorname{rg} A) = m + n - \operatorname{rg} A. \tag{15}$$

Мы доказали, что ранг окаймляющей матрицы M равен ее порядку, что и означает невырожденность. Теорема доказана.

Литература

1. W. E. Roth, *The equations $AX - YB = C$ and $AX - XB = C$ in matrices.* — Proc. Amer. Math. Soc., **3** (1953), 392–396.
2. H. Flanders, H. K. Wimmer, *On matrix equations $AX - XB = C$ and $AX - YB = C$.* — SIAM J. Appl. Math., **32** (1977), 707–710.
3. В. В. Прасолов, *Задачи и теоремы линейной алгебры.* — Наука, Физматлит, М. (1996).
4. Ф. Р. Гантмахер, *Теория матриц.* — Наука, М. (1967).
5. В. Н. Малоземов, М. Ф. Монако, А. А. Петров, *Формулы Фробениуса, Шермана–Моррисона и близкие вопросы.* — Ж. вычисл. матем. матем. физ., **42**, No. 10 (2002), 1459–1465.

Al'pin Yu. A., Il'yin S. N. The matrix equation $AX - YB = C$ and related problems.

The main result of the paper is a theorem, using which a new proof of Roth's theorem is obtained, a new solvability criterion for the matrix equation $AX - YB = C$ is proved, a formula for a particular solution of the latter is derived, and the least of the orders of nonsingular matrices containing a given rectangular matrix as a submatrix is determined.

Казанский
государственный университет
E-mail: Sergey.Ilyin@ksu.ru

Поступило 1 июня 2005 г.