Глава VII. Матрицы

§ 25. Умножение матриц. Матрицы и многочлены

Б.М.Верников

Уральский федеральный университет,
Институт естественных наук и математики,
кафедра алгебры и фундаментальной информатики

В § 4 были введены операции сложения и умножения матриц, в § 6 — операция умножения матрицы на скаляр, а в § 8 — операция транспонирования матрицы. В тех же параграфах указан ряд свойств этих операций. напомним те из них, которые связаны с умножнием матриц:

1) $(AB)C = A(BC)$,
2) $(A + B)C = AC + BC$,
3) $A(B + C) = AB + AC$,
4) $(tA)B = A(tB) = t(AB)$,
5) $t(sA) = (ts)A$,
6) $(AB)^\top = B^\top A^\top$.

Этот параграф посвящен более глубокому изучению операции умножения матриц.

Ослабленный закон сокращения

В любом поле F выполнено следующее свойство, называемое *законом сокращения*: если $x, y, z \in F$, $xz = yz$ и $z \neq 0$, то $x = y$ (чтобы убедиться в этом, надо умножить обе части равенства справа на z^{-1}). В кольце матриц это свойство не выполняется. Но справедлив следующий его ослабленный аналог.

Ослабленный закон сокращения для матриц

Пусть A и B — квадратные матрицы порядка n над полем F. Если для любой матрицы X размера $n \times 1$ над F выполнено равенство $AX = BX$, то $A = B$. Аналогичное заключение верно, если для любой матрицы Y порядка $1 \times n$ над F выполнено равенство $YA = YB$.

Доказательство. Докажем первое утверждение. Пусть $A = (a_{ij})$ и $B = (b_{ij})$. Для всякого $i = 1, 2, \ldots, n$ обозначим через A_i и B_i i-е столбцы матриц A и B соответственно, а через X_i — матрицу размера $n \times 1$, в которой i-й элемент равен 1, а все остальные элементы равны 0. Ясно, что $AX_i = A_i$ и $BX_i = B_i$. Учитывая, что $AX_i = BX_i$, получаем, что i-е столбцы матриц A и B совпадают. Поскольку это выполняется для любого $i = 1, 2, \ldots, n$, получаем, что $A = B$. Второе утверждение доказывается аналогично (надо только рассматривать не столбцы, а строки матриц A и B). □

Отметим, что мы уже сталкивались со свойством такого типа при изучении скалярного произведения векторов (см. § 11).

Наша следующая цель — доказать, что определитель произведения матриц равен произведению их определителей. Чтобы сделать это, нам потребуется рассмотреть один специальный вид матриц.

Определение

Квадратная матрица L порядка n называется *полураспавшейся*, если существуют квадратные матрицы A и B порядков p и q соответственно такие, что либо

$$L = \begin{pmatrix} A & N \\ O & B \end{pmatrix} \tag{1}$$

для нулевой матрицы O размера $q \times p$ и некоторой матрицы N размера $p \times q$, либо

$$L = \begin{pmatrix} A & O \\ N & B \end{pmatrix} \tag{2}$$

для нулевой матрицы O размера $p \times q$ и некоторой матрицы N размера $q \times p$. Матрицы A и B называются *диагональными блоками* полураспавшейся матрицы L.

Предложение об определителе полураспавшейся матрицы

Если L — полураспавшаяся матрица с диагональными блоками A и B, то $|L| = |A| \cdot |B|$.

Доказательство. Достаточно рассмотреть случай, когда матрица L имеет вид (1). В самом деле, если L имеет вид (2), то $L^\top$ — полураспавшаяся матрица вида (1) с диагональными блоками $A^\top$ и $B^\top$. Поэтому если для матриц вида (1) предложение уже доказано, то, используя 1-е свойство определителей (см. § 8), имеем $|L| = |L^\top| = |A^\top| \cdot |B^\top| = |A| \cdot |B|$. Пусть $A = (a_{ij})$, $B = (b_{ij})$ и $N = (n_{ij})$. Дальнейшие рассуждения проведем индукцией по p.

База индукции. Если $p = 1$, то $A = (a_{11})$ и

$$|L| = \begin{vmatrix} a_{11} & n_{11} & n_{12} & \ldots & n_{1q} \\ 0 & b_{11} & b_{12} & \ldots & b_{1q} \\ 0 & b_{21} & b_{22} & \ldots & b_{2q} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & b_{q1} & b_{q2} & \ldots & b_{qq} \end{vmatrix}.$$

Разложив определитель из правой части этого равенства по первому столбцу, получим, что $|L| = a_{11} \cdot |B| = |A| \cdot |B|$.

Шаг индукции. Пусть $p > 1$. Минор матрицы A, соответствующий элементу a_{ij}, будем обозначать через M_{ij}. Разложив определитель матрицы L по первому столбцу и использовав предположение индукции, получим:

$$|L| = a_{11} \cdot \begin{vmatrix} a_{22} & \ldots & a_{2p} & n_{21} & \ldots & n_{2q} \\ a_{32} & \ldots & a_{3p} & n_{31} & \ldots & n_{3q} \\ \multicolumn{6}{c}{\dotfill} \\ a_{p2} & \ldots & a_{pp} & n_{r1} & \ldots & n_{rq} \\ 0 & \ldots & 0 & b_{11} & \ldots & b_{1q} \\ \multicolumn{6}{c}{\dotfill} \\ 0 & \ldots & 0 & b_{q1} & \ldots & b_{qq} \end{vmatrix} - a_{21} \cdot \begin{vmatrix} a_{12} & \ldots & a_{1p} & n_{11} & \ldots & n_{1q} \\ a_{32} & \ldots & a_{3p} & n_{31} & \ldots & n_{3q} \\ \multicolumn{6}{c}{\dotfill} \\ a_{p2} & \ldots & a_{pp} & n_{r1} & \ldots & n_{pq} \\ 0 & \ldots & 0 & b_{11} & \ldots & b_{1q} \\ \multicolumn{6}{c}{\dotfill} \\ 0 & \ldots & 0 & b_{q1} & \ldots & b_{qq} \end{vmatrix} +$$

$$+ \cdots + (-1)^{p+1} a_{p1} \cdot \begin{vmatrix} a_{12} & \ldots & a_{1p} & n_{11} & \ldots & n_{1q} \\ a_{22} & \ldots & a_{2p} & n_{21} & \ldots & n_{2q} \\ \multicolumn{6}{c}{\dotfill} \\ a_{p-1\,2} & \ldots & a_{p-1\,p} & n_{p-1\,1} & \ldots & n_{p-1\,q} \\ 0 & \ldots & 0 & b_{11} & \ldots & b_{1q} \\ \multicolumn{6}{c}{\dotfill} \\ 0 & \ldots & 0 & b_{q1} & \ldots & b_{qq} \end{vmatrix} =$$

$$= a_{11} M_{11} \cdot |B| - a_{21} M_{21} \cdot |B| + \cdots + (-1)^{p+1} a_{p1} M_{p1} \cdot |B| =$$
$$= a_{11} A_{11} \cdot |B| + a_{21} A_{21} \cdot |B| + \cdots + a_{p1} A_{p1} \cdot |B| =$$
$$= (a_{11} A_{11} + a_{21} A_{21} + \cdots + a_{p1} A_{p1}) \cdot |B| = |A| \cdot |B|. \qquad \square$$

Теорема об определителе произведения матриц

Если $A = (a_{ij})$ *и* $B = (b_{ij})$ *— квадратные матрицы одного и того же порядка, то* $|AB| = |A| \cdot |B|$.

Доказательство. Обозначим порядок матриц A и B через n, а матрицу AB — через C. Пусть $C = (c_{ij})$. Рассмотрим следующую полураспавшуюся матрицу с диагональными блоками A и B:

$$D = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} & 0 & 0 & \ldots & 0 \\ a_{21} & a_{22} & \ldots & a_{2n} & 0 & 0 & \ldots & 0 \\ \multicolumn{8}{c}{\dotfill} \\ a_{n1} & a_{n2} & \ldots & a_{nn} & 0 & 0 & \ldots & 0 \\ -1 & 0 & \ldots & 0 & b_{11} & b_{12} & \ldots & b_{1n} \\ 0 & -1 & \ldots & 0 & b_{21} & b_{22} & \ldots & b_{2n} \\ \multicolumn{8}{c}{\dotfill} \\ 0 & 0 & \ldots & -1 & b_{n1} & b_{n2} & \ldots & b_{nn} \end{pmatrix}.$$

В силу предложения об определителе полураспавшейся матрицы

$$|D| = |A| \cdot |B|. \qquad (3)$$

Для всякого $j = 1, 2, \ldots, n$ прибавим к $(n+j)$-му столбцу матрицы D ее первый столбец, умноженный на b_{1j}, второй, умноженный на b_{2j}, $\ldots$, и, наконец, n-й, умноженный на b_{nj}. Полученную матрицу обозначим через D'. Ясно, что первые n столбцов матрицы D' совпадают с соответствующими столбцами матрицы D. Для всякого $j = 1, 2, \ldots, n$ элемент, стоящий в i-й строке и $(n+j)$-м столбце матрицы D', равен $a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = c_{ij}$, если $1 \leqslant i \leqslant n$, и $-b_{ij} + b_{ij} = 0$, если $n+1 \leqslant i \leqslant 2n$. Таким образом, матрица D' имеет следующий вид:

$$D' = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} & c_{11} & c_{12} & \ldots & c_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} & c_{21} & c_{22} & \ldots & c_{2n} \\ \multicolumn{8}{c}{\dotfill} \\ a_{n1} & a_{n2} & \ldots & a_{nn} & c_{n1} & c_{n2} & \ldots & c_{nn} \\ -1 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 \\ 0 & -1 & \ldots & 0 & 0 & 0 & \ldots & 0 \\ \multicolumn{8}{c}{\dotfill} \\ 0 & 0 & \ldots & -1 & 0 & 0 & \ldots & 0 \end{pmatrix}.$$

С учетом 7-го свойства определителей и принципа равноправия строк и столбцов (см. § 8), получаем, что

$$|D'| = |D|. \qquad (4)$$

Поменяем в матрице D' местами сначала $(n+1)$-й столбец с первым, затем $(n+2)$-й столбец — со вторым, ..., наконец, последний столбец — с n-м. В результате мы получим матрицу

$$D'' = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1n} & a_{11} & a_{12} & \dots & a_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} & a_{21} & a_{22} & \dots & a_{2n} \\ \multicolumn{8}{c}{\dotfill} \\ c_{n1} & c_{n2} & \dots & c_{nn} & a_{n1} & a_{n2} & \dots & a_{nn} \\ 0 & 0 & \dots & 0 & -1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & -1 & \dots & 0 \\ \multicolumn{8}{c}{\dotfill} \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & -1 \end{pmatrix}.$$

Переходя от матрицы D' к матрице D'', мы сделали n перестановок столбцов. Применяя 4-е свойство определителей и принцип равноправия строк и столбцов (см. § 8), имеем

$$|D''| = (-1)^n \cdot |D'|. \tag{5}$$

Матрица D'' является полураспавшейся матрицей с диагональными блоками C и $-E$. Предложение об определителе треугольной матрицы (см. § 8) и предложение об определителе полураспавшейся матрицы показывают, что $|D''| = |C| \cdot (-1)^n$. Умножая обе части этого равенства на $(-1)^n$, имеем $(-1)^n \cdot |D''| = (-1)^{2n} \cdot |C| = |C|$, т. е.

$$|C| = (-1)^n \cdot |D''|. \qquad (6)$$

Из равенств (3)–(6) вытекает, что

$$|C| = (-1)^n \cdot |D''| = (-1)^{2n} \cdot |D'| = |D'| = |D| = |A| \cdot |B|.$$

Теорема доказана. $\square$

В дальнейшем нам понадобится следующее понятие.

Определение

Пусть $A = (a_{ij})$ — квадратная матрица порядка > 1 над полем F. Матрица (A_{ij}), т. е. матрица, составленная из алгебраических дополнений к элементам матрицы A, называется *матрицей, присоединенной к* A, и обозначается через $A^{\vee}$.

Замечание о присоединенной матрице

Если A — произвольная квадратная матрица, то

$$A \cdot (A^{\vee})^{\top} = (A^{\vee})^{\top} \cdot A = |A| \cdot E.$$

Доказательство. Пусть $X = A \cdot (A^{\vee})^{\top}$. Положим $X = (x_{ij})$. Ясно, что $x_{ij} = a_{i1}A_{j1} + a_{i2}A_{j2} + \dots + a_{in}A_{jn}$. В силу формулы (7) из § 8 получаем, что

$$x_{ij} = \begin{cases} |A|, & \text{если } i = j, \\ 0, & \text{если } i \neq j. \end{cases}$$

Следовательно, $A \cdot (A^{\vee})^{\top} = |A| \cdot E$. Равенство $(A^{\vee})^{\top} \cdot A = |A| \cdot E$ проверяется точно так же. $\square$

Зафиксируем кольцо R и натуральное число n. Как уже отмечалось в § 4, множество $R^{n\times n}$ с операциями сложения и умножения матриц является ассоциативным кольцом с 1 (роль единицы играет единичная матрица порядка n). Рассмотрим кольцо $R^{n\times n}[x]$, т. е. кольцо многочленов над кольцом $R^{n\times n}$. Его элементами являются многочлены, коэффициентами которых являются квадратные матрицы порядка n над кольцом R, т. е. выражения вида $A_n x^n + A_{n-1}x^{n-1} + \cdots + A_0$, где $A_n, A_{n-1}, \ldots, A_0 \in R^{n\times n}$. Элементы кольца $R^{n\times n}[x]$ можно рассматривать как матрицы над кольцом $R[x]$. Другими словами, $R^{n\times n}[x] = \bigl(R[x]\bigr)^{n\times n}$. Например, в следующем равенстве слева от знака равенства стоит элемент кольца $\mathbb{R}^{2\times 2}[x]$, а справа от него — элемент кольца $\bigl(\mathbb{R}[x]\bigr)^{2\times 2}$:

$$\begin{pmatrix} 3 & -1 \\ 2 & 2 \end{pmatrix} x^2 + \begin{pmatrix} 5 & 1 \\ -1 & 0 \end{pmatrix} x + \begin{pmatrix} 1 & 2 \\ -2 & 4 \end{pmatrix} = \begin{pmatrix} 3x^2 + 5x + 1 & -x^2 + x + 2 \\ 2x^2 - x - 2 & 2x^2 + 4 \end{pmatrix}.$$

Следующее понятие будет играть исключительно важную роль в дальнейшем.

Определение

Пусть A — квадратная матрица над полем F. *Характеристическим многочленом* матрицы A называется многочлен $|A - xE|$, где E — единичная матрица того же порядка, что и A. Этот многочлен обозначается через $\chi_A(x)$.

Если $A = (a_{ij})$, а порядок A равен n, то

$$\chi_A(x) = \begin{vmatrix} a_{11} - x & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - x & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} - x \end{vmatrix}.$$

Ясно, что при вычислении определителя, стоящего в правой части этого равенства, появится лишь одно слагаемое, содержащее x^n, именно, — $(a_{11} - x)(a_{22} - x) \cdots (a_{nn} - x)$. Поэтому старший член многочлена $\chi_A(x)$ имеет вид $(-1)^n x^n$. Кроме того, очевидно, что свободный член этого многочлена есть $\chi_A(0) = |A|$.

Определение

Пусть A — квадратная матрица над полем F. Для всякого целого $n \geqslant 0$ определим по индукции матрицу A^n следующим образом: $A^0 = E$, где E — единичная матрица того же порядка, что и A, а если $n > 0$, то $A^n = A^{n-1} \cdot A$.

Поскольку, как уже отмечалось выше, произведение двух квадратных матриц одного и того же порядка всегда определено и является квадратной матрицей того же порядка, матрица A^n при любом n определена и является квадратной матрицей того же порядка, что и A.

Определение

Пусть $f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_0$ — многочлен над полем F, а A — квадратная матрица над F. *Значением многочлена f от матрицы A* называется матрица $f(A) = a_n A^n + a_{n-1} A^{n-1} + \dots + a_0 E$.

Ясно, что $f(A)$ — квадратная матрица над F того же порядка, что и A.

Определение

Пусть F — поле, $A \in F^{n \times n}$ и $f \in F[x]$. Говорят, что многочлен f *аннулирует* матрицу A, если $f(A) = O$.

Теорема Гамильтона–Кэли

Характеристический многочлен произвольной квадратной матрицы A аннулирует эту матрицу.

Доказательство. Пусть $\chi_A(x) = a_n x^n + a_{n-1}x^{n-1} + \dots + a_0$. Положим $B(x) = \left((A - xE)^{\vee}\right)^{\top}$. Элементами матрицы $B(x)$ являются алгебраические дополнения к элементам матрицы $A - xE$, т. е., с точностью до знака, миноры $(n-1)$-го порядка этой матрицы. Ясно, что эти миноры суть многочлены степени $\leqslant n-1$. Следовательно, матрицу $B(x)$ можно записать в виде $B_{n-1}x^{n-1} + \dots + B_1 x + B_0$, где $B_0, B_1, \dots . B_{n-1} \in F^{n\times n}$. Из замечания о присоединенной матрице вытекает, что $(A - xE)B(x) = |A - xE| \cdot E = \chi_A(x) \cdot E$. Следовательно,

$$(A - xE)(B_{n-1}x^{n-1} + \dots + B_1 x + B_0) = a_n x^n E + a_{n-1}x^{n-1}E + \dots + a_0 E.$$

Раскрывая скобки в левой части этого равенства и приводя подобные, имеем

$$\begin{aligned} &- B_{n-1}x^n + (AB_{n-1} - B_{n-2})x^{n-1} + \dots + (AB_1 - B_0)x + AB_0 = \\ &= a_n x^n E + a_{n-1}x^{n-1}E + \dots + a_0 E. \end{aligned}$$

Приравнивая коэффициенты при одинаковых степенях x в левой и правой частях последнего равенства, имеем

$$a_nE = -B_{n-1},\ a_{n-1}E = AB_{n-1} - B_{n-2}, \ldots, a_1E = AB_1 - B_0,\ a_0E = AB_0.$$

Умножая слева первое из этих равенств на A^n, второе – на A^{n-1}, ..., предпоследнее — на A, последнее — на E, получим следующий набор равенств:

$$\begin{array}{rcl} a_nA^n & = & -A^nB_{n-1}, \\ a_{n-1}A^{n-1} & = & A^nB_{n-1} - A^{n-1}B_{n-2}, \\ a_{n-2}A^{n-2} & = & A^{n-1}B_{n-2} - A^{n-2}B_{n-3}, \\ \multicolumn{3}{c}{\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots} \\ a_1A & = & A^2B_1 - AB_0, \\ a_0E & = & AB_0. \end{array}$$

Сумма левых частей этих равенств равна

$$a_nA^n + a_{n-1}A^{n-1} + \cdots + a_1A + a_0E = \chi_A(A),$$

а сумма их правых частей равна O. Следовательно, $\chi_A(A) = O$. $\square$

Матричным уравнением называется уравнение, в котором неизвестным является матрица. Здесь мы подробно рассмотрим матричное уравнение вида

$$AX = B, \tag{7}$$

где A и B — известные матрицы, а X — неизвестная. Из определения произведения матриц видно, что число строк в матрице AX равно числу строк в матрице A. Следовательно, если число строк в матрицах A и B различно, то уравнение (7) решений не имеет. Поэтому всюду в дальнейшем, говоря об этом уравнении, мы будем считать, что *матрицы A и B имеют одинаковое число строк*.

Из определения произведения матриц видно также, что число столбцов в матрице AX равно числу столбцов в матрице X. Следовательно, если X — решение уравнения (7), то матрицы X и B содержат одинаковое число столбцов. Как мы видели в § 6, если матрицы X и B содержат один столбец, то уравнение (7) есть просто другой способ записи системы линейных уравнений. Обозначим через k число столбцов в матрицах X и B. Для всякого $i = 1, 2, \dots, k$ обозначим i-й столбец матрицы X через X_i, а i-й столбец матрицы B — через B_i. Из определения произведения матриц вытекает, что i-й столбец матрицы AX равен AX_i.

Поэтому

!! *в общем случае уравнение (7) равносильно совокупности k систем линейных уравнений вида*

$$AX_1 = B_1, AX_2 = B_2, \ldots, AX_k = B_k. \tag{8}$$

Для того чтобы решить каждую из этих систем методом Гаусса, надо записать расширенную матрицу системы и с помощью элементарных преобразований привести ее к ступенчатому виду. Если при этом окажется, что хотя бы одна из систем несовместна, то и исходное матричное уравнение не имеет решений. Если же все системы совместны, то, решив каждую из них, мы найдем все столбцы матрицы X, а значит и саму эту матрицу. Но у всех решаемых систем основная матрица одна и та же — матрица A. Это позволяет решать все системы одновременно, действуя по алгоритму, который приведен на следующем слайде.

Алгоритм решения матричного уравнения вида $AX = B$

Пусть дано уравнение (7), в котором A — матрица размера $n \times k$, а B — матрица размера $n \times m$. Запишем матрицу $(A \,|\, B)$, т. е. матрицу размера $n \times (k + m)$, в которой в первых k столбцах стоит матрица A, а в последних m столбцах — матрица B. С помощью элементарных преобразований всей этой матрицы приведем ее левую часть (т. е. первые k столбцов) к ступенчатому виду. После этого для всякого $i = 1, 2, \ldots, k$ можно найти i-й столбец матрицы X, решив систему линейных уравнений вида $A'X_i = B'_i$, где A' — левая часть полученной матрицы, а B'_i — i-й столбец правой части полученной матрицы. Если при этом окажется, что хотя бы одна из этих систем несовместна, то уравнение $AX = B$ решений не имеет.

Особый интерес с точки зрения дальнейшего представляет уравнение (7), в котором матрица A является квадратной и невырожденной. В этом случае каждая из систем (8) является крамеровской. В силу теоремы Крамера (см. § 9) все эти системы имеют единственное решение. Следовательно, и уравнение (7) имеет единственное решение. Это позволяет использовать для его решения метод Гаусса–Жордана в том виде, в каком он был изложен в конце § 7. Объединяя приведенный там алгоритм нахождения решения системы, имеющей единственное решение, и алгоритм с предыдущего слайда, получаем изложенный на следующем слайде алгоритм, который будет использован в дальнейшем для решения ряда важных задач.

Алгоритм решения матричного уравнения вида $AX = B$ в случае невырожденной квадратной матрицы A

Пусть дано уравнение (7), в котором A — невырожденная квадратная матрица порядка n, а B — матрица размера $n \times k$. Запишем матрицу $(A \mid B)$ размера $n \times (n + k)$ и с помощью элементарных преобразований всей этой матрицы приведем ее левую часть (т. е. первые n столбцов) к единичному виду. В правой части (т. е. в последних k столбцах) полученной матрицы будет записана матрица X, являющаяся единственным решением уравнения (7).

Все сказанное выше об уравнении (7) можно использовать для того, чтобы решить матричное уравнение вида $XA = B$, где A и B — известные матрицы, а X — неизвестная. Транспонируя обе части равенства $XA = B$ и используя свойство 7) произведения матриц, получаем уравнение $A^\top X^\top = B^\top$, т. е. уравнение вида (7). Решив его одним из двух описанных выше способов, мы найдем матрицу $X^\top$ (естественно, второй способ можно применять лишь в случае, когда A — невырожденная квадратная матрица). Транспонировав матрицу $X^\top$, найдем искомую матрицу X.