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 § 3 были введены операции сложения и умножения матриц, в § 5 — операция умножения матрицы на скаляр, а в § 7 — операция транспонирования матрицы. В тех же параграфах указан ряд свойств этих операций. Здесь мы приводим некоторые более глубокие свойства указанных операций, а в следующем параграфе рассмотрим одну новую операцию (обращение матрицы).

25.1. Определитель произведения матриц

Теорема 25.1

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:

$$D = \begin{pmatrix} A & O \\ -E & B \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & 0 & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 & 0 & \dots & 0 \\ -1 & 0 & \dots & 0 & b_{11} & b_{12} & \dots & b_{1n} \\ 0 & -1 & \dots & 0 & b_{21} & b_{22} & \dots & b_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & -1 & b_{n1} & b_{n2} & \dots & b_{nn} \end{pmatrix}.$$

В силу предложения 7.12

$$|D| = |A| \cdot |B|. \tag{1}$$

Определитель произведения матриц (2)

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:

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

С учетом 7-го свойства определителей и принципа равноправия строк и столбцов, получаем, что

$$|D'| = |D|. \tag{2}$$

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

$$D'' = \begin{pmatrix} C & A \\ O & -E \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & \ldots & c_{1n} & a_{11} & a_{12} & \ldots & a_{1n} \\ c_{21} & c_{22} & \ldots & c_{2n} & a_{21} & a_{22} & \ldots & a_{2n} \\ \hdotsfor{8} \\ c_{n1} & c_{n2} & \ldots & c_{nn} & a_{n1} & a_{n2} & \ldots & a_{nn} \\ 0 & 0 & \ldots & 0 & -1 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & 0 & -1 & \ldots & 0 \\ \hdotsfor{8} \\ 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & -1 \end{pmatrix}.$$

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

$$|D''| = (-1)^n \cdot |D'|. \qquad (3)$$

Определитель произведения матриц (4)

Матрица D'' является полураспавшейся матрицей с диагональными блоками C и $-E$. Предложения 7.11 и 7.12 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.

$$|C| = (-1)^n \cdot |D''|. \tag{4}$$

Объединяя сказанное выше, имеем:

$$\begin{aligned}
|C| &= (-1)^n \cdot |D''| && \text{в силу (4)}\\
&= (-1)^{2n} \cdot |D'| && \text{в силу (3)}\\
&= |D'| &&\\
&= |D| && \text{в силу (2)}\\
&= |A| \cdot |B|. && \text{в силу (1)}
\end{aligned}$$

Теорема доказана. $\square$

25.2. Ослабленный закон сокращения для матриц

Как уже отмечалось в § 10, в любом поле F выполнено следующее свойство, называемое *законом сокращения*: если $x, y, z \in F$, $xz = yz$ и $z \neq 0$, то $x = y$. В кольце матриц это свойство не выполняется. В самом деле, пусть A и B — матрицы размера $m \times n$, где $n > 1$, а X — матрица размера $n \times 1$, т. е. столбец длины n, в котором первый элемент равен 1, а остальные элементы равны 0. Тогда AX — первый столбец матрицы A, а BX — первый столбец матрицы B. Если матрицы A и B различны, но их первые столбцы совпадают, то $AX = BX$ и $X \neq O$, но $A \neq B$. Аналогично проверяется, что существуют матрицы A, B и X такие, что $XA = XB$ и $X \neq O$, но $A \neq B$ (в качестве A и B можно взять различные матрицы размера $m \times n$, где $m > 1$, у которых совпадают первые строки, а в качестве X — матрицу размера $1 \times m$, т. е. строку длины m, в которой первый элемент равен 1, а остальные элементы равны 0).

Ослабленный закон сокращения

Но справедлив следующий ослабленный аналог указанного выше свойства.

Предложение 25.1 (ослабленный закон сокращения для матриц)

Пусть A и B — матрицы размера $m \times n$ 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 размера $1 \times m$ 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 $n \times 1$, т. е. столбец длины n, в котором 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, а в качестве X_i брать матрицу размера $1 \times m$, т. е. строку длины m, в которой i-й элемент равен 1, а остальные элементы равны 0). $\square$

Отметим, что мы уже сталкивались со свойством такого типа при изучении скалярного произведения векторов (см. предложение 10.2).

25.3. Теорема Гамильтона–Кэли

Следующее понятие будет играть исключительно важную роль в дальнейшем.

Определение

Пусть A — квадратная матрица над полем F. Легко понять, что определитель матрицы $A - xE$, где E — единичная матрица того же порядка, что и A, является многочленом относительно неизвестного x. Этот многочлен называется *характеристическим многочленом* матрицы A и обозначается через $\chi_A(x)$.

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

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

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 $(-1)^n x^n$. В частности, *степень многочлена $\chi_A(x)$ равна порядку матрицы 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 $A^n = A^{n-1} \cdot A$.

Ясно, что 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 $f(A) = O$.

Для дальнейшего нам понадобится одно новое 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 25.1

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 $x_{ij} = a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn}$, где n — порядок матрицы A. В силу 8-го и 9-го свойств определителей 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. $\square$

Теорема 25.2 (теорема Гамильтона–Кэли)

Характеристический многочлен произвольной квадратной матрицы A над полем F 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 $\leqslant n-1$ над полем F. Обозначим многочлен, стоящий в i-й строке и j-м столбце матрицы $B(x)$ через $b_{ij}(x)$ (для любых $i, j = 1, 2, \dots, n$). Для всякого $k = 0, 1, \dots, n-1$ обозначим через B_k квадратную матрицу порядка n над полем F, в i-й строке и j-м столбце которой (для любых $i, j = 1, 2, \dots, n$) стоит коэффициент при x^k в многочлене $b_{ij}(x)$. Ясно, что $B(x) = B_{n-1}x^{n-1} + \dots + B_1 x + B_0$.

Из замечания 25.1 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

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

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

$$a_n E = -B_{n-1},\, a_{n-1}E = AB_{n-1} - B_{n-2}, \ldots, a_1 E = AB_1 - B_0,\, a_0 E = AB_0.$$

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

$$\begin{array}{rcl}
a_n A^n & = & -A^n B_{n-1}, \\
a_{n-1}A^{n-1} & = & A^n B_{n-1} - A^{n-1}B_{n-2}, \\
a_{n-2}A^{n-2} & = & A^{n-1}B_{n-2} - A^{n-2}B_{n-3}, \\
\multicolumn{3}{c}{\dotfill} \\
a_1 A & = & A^2 B_1 - AB_0, \\
a_0 E & = & AB_0.
\end{array}$$

Теорема Гамильтона–Кэли (3)

Складывая левые и правые части этих равенств, имеем

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

т. е. $\chi_A(A) = O$. $\square$

25.4. Матричные уравнения

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

$$AX = B, \tag{5}$$

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 (5) 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 (5), то матрицы X и B содержат одинаковое число столбцов. Как мы видели в § 5, если матрицы X и B содержат один столбец, то уравнение (5) есть просто другой способ записи системы линейных уравнений. Обозначим через k число столбцов в матрицах X и B. Для всякого $i = 1, 2, \ldots, k$ 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* (5) *равносильно совокупности k систем линейных уравнений вида*

$$AX_1 = B_1, AX_2 = B_2, \dots, AX_k = B_k. \tag{6}$$

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.

Матричное уравнение вида $AX = B$ (3)

Алгоритм 25.1

Пусть дано уравнение (5), в котором A — матрица размера $n \times m$, а B — матрица размера $n \times k$. Запишем матрицу $(A \mid B)$, т. е. матрицу размера $n \times (m + k)$, в которой в первых m столбцах стоит матрица A, а в последних k столбцах — матрица B. С помощью элементарных преобразований всей этой матрицы приведем ее левую часть (т. е. первые m 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.

Матричное уравнение вида $AX = B$ в случае невырожденной квадратной матрицы A

Особый интерес с точки зрения дальнейшего представляет уравнение (5), в котором матрица A является квадратной и невырожденной. В этом случае каждая из систем (6) является крамеровской. В силу теоремы Крамера все эти системы имеют единственное решение. Следовательно, и уравнение (5) имеет единственное решение. Это позволяет использовать для его решения метод Гаусса–Жордана. Объединяя алгоритмы 6.1 и 25.1, получаем следующий алгоритм, который будет использован в дальнейшем для решения ряда важных задач.

Алгоритм 25.2

Пусть дано уравнение (5), 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 (5).

Матричные уравнения вида $XA = B$ и $AXB = C$

Рассмотрим теперь матричное уравнение вида $XA = B$, где A и B — известные матрицы, а X — неизвестная. Транспонируя обе части равенства $XA = B$ и используя равенство $(XA)^\top = A^\top X^\top$, получаем уравнение $A^\top X^\top = B^\top$, т. е. уравнение вида (5). 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). Используя равенство $(X^\top)^\top = X$, транспонируем матрицу $X^\top$ и найдем искомую матрицу X.

Рассмотрим, наконец, уравнение $AXB = C$, где матрицы A, B и C известны, а матрица X неизвестна. Положим $Y = XB$. Тогда наше уравнение принимает вид $AY = C$. Его можно решить, используя алгоритм 25.1. Обозначим решение этого уравнения через D. Остается решить уравнение $XB = D$, что можно сделать способом, указанным в предыдущем абзаце.