Глава IV. Матрицы

§ 14. Умножение матриц. Матрицы и многочлены

Б.М.Верников

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

Вступительные замечания

В § 4 были введены операции сложения и умножения матриц, в § 6 — 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).

14.1. Ослабленный закон сокращения для матриц

В любом поле F выполнено следующее свойство, называемое *закоеном 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 14.1 (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 в курсе аналитической геометрии доказано аналогичное свойство скалярного произведения векторов в трехмерном пространстве.

14.2. Многочлены, аннулирующие матрицу

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

Говорят, что многочлен f *аннулирует* матрицу A, если f — ненулевой многочлен и $f(A) = O$.

Чтобы привести пример многочлена, аннулирующего матрицу A, введем понятие, которое 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

$$\chi_A(x) = \begin{vmatrix} a_{11} - x & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - x & \dots & a_{2n} \\ \multicolumn{4}{c}{\dotfill} \\ a_{n1} & a_{n2} & \dots & a_{nn} - x \end{vmatrix}.$$

Ясно, что при вычислении определителя, стоящего в правой части этого равенства, появится лишь одно слагаемое, содержащее x^n, а 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*.

Теорема Гамильтона–Кэли (1)

Теорема 14.1 (теорема Гамильтона–Кэли)

Характеристический многочлен произвольной квадратной матрицы A над полем F аннулирует эту матрицу.

Для того, чтобы доказать эту теорему, 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 о присоединенной матрице

Замечание 14.1

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. □

В доказательстве теоремы Гамильтона–Кэли возникнет матрица, элементы которой являются многочленами. Заметим, что всякую такую матрицу можно переписать в виде многочлена, коэффициенты которого являются матрицами. Продемонстрируем это на конкретном примере. Рассмотрим матрицу

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

Ясно, что

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

Очевидно, что аналогичное равенство верно для любой матрицы над кольцом многочленов.

Доказательство теоремы Гамильтона–Кэли. 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. В силу сказанного на предыдущем слайде матрицу $B(x)$ можно переписать как многочлен степени $\leqslant n-1$ над кольцом $F^{n\times n}$. Иными словами, $B(x) = B_{n-1}x^{n-1} + \dots + B_1 x + B_0$, где $B_0, B_1, \dots, B_{n-1}$ — некоторые квадратные матрицы порядка n над полем F. Из замечания 14.1 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

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

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:

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

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. $\chi_A(A) = O$. $\square$

Минимальный многочлен матрицы (1)

Теорема Гамильтона–Кэли показывает, что для всякой квадратной матрицы порядка n существует аннулирующий ее многочлен степени n. Но некоторые матрицы могут аннулироваться многочленами меньшей степени. Например, единичная матрица произвольного порядка, очевидно, аннулируется линейным многочленом $f(x) = x - 1$.

Определение

Минимальным многочленом квадратной матрицы A называется унитарный многочлен, который аннулирует A и имеет наименьшую степень среди многочленов с этим свойством.

Лемма 14.1

Любая квадратная матрица имеет минимальный многочлен, и притом только один.

Доказательство. Теорема Гамильтона–Кэли показывает, что многочлен, аннулирующий данную квадратную матрицу, существует. Отсюда вытекает существование многочлена минимальной степени с этим свойством. Разделив последний многочлен на его старший коэффициент, получим минимальный многочлен. Докажем его единственность. Пусть f_1 и f_2 — минимальные многочлены одной и той же матрицы A.

Тогда $\deg f_1 = \deg f_2$ и $\ell c(f_1) = \ell c(f_2)$. Следовательно, $\ell m(f_1) = \ell m(f_2)$. Положим $g = f_1 - f_2$. Очевидно, что $\deg g < \deg f_1$ и $g(A) = f_1(A) - f_2(A) = O - O = O$. Если $g \neq o$, то разделив g на его старший коэффициент, получим унитарный многочлен, аннулирующий A, степень которого меньше степени f_1. Это противоречит минимальности многочлена f_1. Следовательно, $g = o$, и потому $f_1 = f_2$. $\square$

Следующее утверждение указывает важное свойство минимального многочлена.

Предложение 14.2

Минимальный многочлен произвольной квадратной матрицы A делит любой аннулирующий ее многочлен.

Доказательство. Пусть f — многочлен, аннулирующий матрицу A, а g — ее минимальный многочлен. Разделим f на g с остатком: $f = qg + r$, где $\deg r < \deg g$. Тогда $f(A) = q(A)g(A) + r(A)$. Учитывая, что $f(A) = g(A) = O$, получаем, что $r(A) = O$. Если $r \neq o$, то это противоречит минимальности многочлена g. Следовательно, $r = o$, и потому $g \mid f$. $\square$

Минимальный многочлен матрицы (3)

Из предложения 14.2 и теоремы Гамильтона–Кэли непосредственно вытекает

Следствие 14.1

Минимальный многочлен произвольной квадратной матрицы A делит ее характеристический многочлен. □

Следствие 14.1 подсказывает один из способов нахождения минимального многочлена матрицы A: надо разложить многочлен $\chi_A(x)$ на неприводимые множители и затем перебирать всевозможные произведения этих множителей (т. е. делители многочлена $\chi_A(x)$) в порядке возрастания степеней этих произведений до тех пор, пока не окажется, что очередное произведение аннулирует матрицу A — оно и будет минимальным многочленом этой матрицы.

Определим новый тип матриц, который будет играть важную роль в дальнейшем.

Определение

Квадратная матрица A называется *клеточно-диагональной*, если существуют квадратные матрицы $A_1, A_2, \ldots, A_m$ такие, что

$$A = \begin{pmatrix} A_1 & O & \ldots & O \\ O & A_2 & \ldots & O \\ \multicolumn{4}{c}{\dotfill} \\ O & O & \ldots & A_m \end{pmatrix}.$$

Матрицы $A_1, A_2, \ldots, A_m$ называются *диагональными клетками* или просто *клетками* матрицы A.

- Порядки клеток A_1, A_2, ..., A_m могут быть любыми.
- Всякая диагональная матрица является клеточно-диагональной матрицей, в которой все клетки имеют порядок 1.
- Всякая квадратная матрица A является клеточно-диагональной матрицей с одной клеткой $A_1 = A$. Поэтому интерес представляют те клеточно-диагональные матрицы, у которых число клеток как можно больше и/или сами клетки устроены как можно проще.
- Всякая клеточно-диагональная матрица с более чем одной диагональной клеткой является полураспавшейся.

Предложение 14.3

Пусть A — клеточно-диагональная матрица с клетками $A_1, A_2, \ldots, A_m$. Тогда минимальный многочлен матрицы A равен наименьшему общему кратному минимальных многочленов матриц $A_1, A_2, \ldots, A_m$.

Доказательство. Легко понять, что если f — произвольный многочлен над тем же полем, что и матрица A, то

$$f(A) = \begin{pmatrix} f(A_1) & O & \ldots & O \\ O & f(A_2) & \ldots & O \\ \ldots & \ldots & \ldots & \ldots \\ O & O & \ldots & f(A_m) \end{pmatrix}.$$

Предположим, что $f(A) = O$. Пусть $k \in \{1, 2, \ldots, m\}$. В силу сказанного, $f(A_k) = O$. Поэтому из предложения 14.2 вытекает, что минимальный многочлен матрицы A_k делит f. Таким образом, произвольный многочлен, аннулирующий матрицу A, является общим кратным минимальных многочленов матриц $A_1, A_2, \ldots, A_m$. Следовательно, чтобы получить минимальный многочлен матрицы A, надо взять унитарный многочлен наименьшей степени среди всех общих кратных этих многочленов, т. е. наименьшее общее кратное минимальных многочленов матриц $A_1, A_2, \ldots, A_m$. $\square$

Очевидно, что минимальный многочлен квадратной матрицы первого порядка $A = (\lambda)$ равен $x - \lambda$. Поэтому из предложения 14.3 вытекает

Следствие 14.2

Минимальный многочлен диагональной матрицы равен $(x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_m)$, где $\lambda_1, \lambda_2, \ldots, \lambda_m$ — совокупность всех попарно различных диагональных элементов этой матрицы. $\square$