члены $(\lambda-1)^2$, $\lambda-1$ и $\lambda+2$, а поэтому жордановой нормальной формой матрицы A служит матрица

$$J=\begin{pmatrix}1 & 1 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & -2\end{pmatrix}.$$

Если бы мы хотели найти ту невырожденную матрицу, которая трансформирует матрицу A в матрицу J, то должны были бы воспользоваться указаниями, сделанными в конце предшествующего параграфа.

На основании предшествующих результатов может быть доказано, наконец, следующее необходимое и достаточное условие приводимости матрицы к диагональному виду, условие, из которого немедленно вытекает достаточный критерий приводимости к диагональному виду, доказанный в § 33.

Матрица A порядка n с элементами из поля P тогда и только тогда приводится к диагональному виду, если все корни последнего инвариантного множителя $e_n(\lambda)$ ее характеристической матрицы лежат в поле P, причем среди этих корней нет кратных.

В самом деле, приводимость матрицы к диагональному виду равносильна приводимости к такому жорданову виду, все жордановы клетки которого имеют порядок 1. Иными словами, все элементарные делители матрицы A должны быть многочленами первой степени. Так как, однако, все инвариантные множители матрицы $A-\lambda E$ являются делителями многочлена $e_n(\lambda)$, то последнее условие равносильно тому, что все элементарные делители многочлена $e_n(\lambda)$ имеют степень 1, что и требовалось доказать.

§ 62. Минимальный многочлен

Пусть дана квадратная матрица A порядка n с элементами из поля P. Если

$$f(\lambda)=\alpha_0\lambda^k+\alpha_1\lambda^{k-1}+\ldots+\alpha_{k-1}\lambda+\alpha_k$$

— произвольный многочлен из кольца $P[\lambda]$, то матрица

$$f(A)=\alpha_0A^k+\alpha_1A^{k-1}+\ldots+\alpha_{k-1}A+\alpha_kE$$

будет называться *значением* многочлена $f(\lambda)$ при $\lambda=A$; обращаем внимание на то, что свободный член многочлена $f(\lambda)$ умножается при этом на нулевую степень матрицы A, т. е. на единичную матрицу E.

Легко проверяется, что если

$$f(\lambda)=\varphi(\lambda)+\psi(\lambda)$$

или

$$f(\lambda)=u(\lambda)\,v(\lambda),$$

то

$$f(A) = \varphi(A) + \psi(A)$$

и, соответственно,

$$f(A) = u(A)\,v(A).$$

Если многочлен $f(\lambda)$ аннулируется матрицей A, т. е.

$$f(A) = 0,$$

то матрицу A будем называть *матричным корнем* или, там, где это не может вызвать недоразумений, просто *корнем* многочлена $f(\lambda)$.

Всякая матрица A служит корнем некоторого ненулевого многочлена.

Мы знаем, в самом деле, что все квадратные матрицы порядка n составляют над полем P n^2-мерное векторное пространство. Отсюда следует, что система n^2+1 матриц

$$A^{n^2},\ A^{n^2-1},\ \ldots,\ A,\ E$$

линейно зависима над полем P, т. е. в P существуют такие элементы $\alpha_0, \alpha_1, \ldots, \alpha_{n^2}, \alpha_{n^2+1}$, не все равные нулю, что

$$\alpha_0 A^{n^2} + \alpha_1 A^{n^2-1} + \ldots + \alpha_{n^2} A + \alpha_{n^2+1} E = 0.$$

Таким образом, матрица A оказалась корнем ненулевого многочлена

$$\varphi(\lambda) = \alpha_0 \lambda^{n^2} + \alpha_1 \lambda^{n^2-1} + \ldots + \alpha_{n^2}\lambda + \alpha_{n^2+1},$$

степень которого не превосходит n^2.

Матрица A служит корнем и для некоторых таких многочленов, старшие коэффициенты которых равны единице — достаточно взять любой ненулевой многочлен, аннулируемый матрицей A, и разделить этот многочлен на его старший коэффициент. Многочлен наименьшей степени со старшим коэффициентом 1, аннулируемый матрицей A, называется *минимальным многочленом матрицы A*. Заметим, что *минимальный многочлен матрицы A определен однозначно*, так как разность двух таких многочленов имела бы меньшую степень, чем каждый из них, но также аннулировалась бы матрицей A.

Всякий многочлен $f(\lambda)$, аннулируемый матрицей A, делится нацело на минимальный многочлен $m(\lambda)$ этой матрицы.

В самом деле, если

$$f(\lambda) = m(\lambda)\,q(\lambda) + r(\lambda),$$

где степень $r(\lambda)$ меньше степени $m(\lambda)$, то

$$f(A) = m(A)\,q(A) + r(A)$$

и из $f(A) = m(A) = 0$ следует $r(A) = 0$, что противоречит определению минимального многочлена.

Докажем теперь следующую теорему:

Минимальный многочлен матрицы A совпадает с последним инвариантным множителем $e_n(\lambda)$ характеристической матрицы $A-\lambda E$.

Доказательство. Сохраняя обозначения и используя результаты § 59, можно написать равенство

$$(-1)^n |A-\lambda E| = d_{n-1}(\lambda)\, e_n(\lambda). \tag{1}$$

Отсюда следует, в частности, что многочлены $e_n(\lambda)$ и $d_{n-1}(\lambda)$ не будут нулевыми. Обозначим, далее, через $B(\lambda)$ присоединенную матрицу к матрице $A-\lambda E$ (см. § 14),

$$B(\lambda) = (A-\lambda E)^*.$$

Как вытекает из § 14, равенство (3), справедливо равенство

$$(A-\lambda E)\, B(\lambda) = |A-\lambda E|\, E. \tag{2}$$

С другой стороны, так как элементами матрицы $B(\lambda)$ служат взятые со знаками плюс или минус миноры $(n-1)$-го порядка матрицы $A-\lambda E$ и только они, а многочлен $d_{n-1}(\lambda)$ есть общий наибольший делитель всёх этих миноров, то

$$B(\lambda) = d_{n-1}(\lambda)\, C(\lambda), \tag{3}$$

причем наибольший общий делитель элементов матрицы $C(\lambda)$ равен 1.

Из равенств (2), (3) и (1) вытекает равенство

$$(A-\lambda E)\, d_{n-1}(\lambda)\, C(\lambda) = (-1)^n d_{n-1}(\lambda)\, e_n(\lambda)\, E.$$

Это равенство можно сократить на ненулевой множитель $d_{n-1}(\lambda)$, как вытекает из следующего общего замечания: если $\varphi(\lambda)$ — ненулевой многочлен, $D(\lambda) = (d_{ij}(\lambda))$ — ненулевая λ-матрица, причем пусть $d_{st}(\lambda) \neq 0$, то в матрице $\varphi(\lambda)\, D(\lambda)$ на месте (s, t) будет стоять отличный от нуля элемент $\varphi(\lambda)\, d_{st}(\lambda)$. Таким образом,

$$(A-\lambda E)\, C(\lambda) = (-1)^n e_n(\lambda)\, E,$$

откуда

$$e_n(\lambda)\, E = (\lambda E - A)\, [(-1)^{n+1} C(\lambda)]. \tag{4}$$

Это равенство показывает, что остаток от «левого» деления λ-матрицы, стоящей слева, на двучлен $\lambda E - A$ равен нулю. Из леммы, доказанной в конце § 60, вытекает, однако, что этот остаток равен матрице $e_n(A)\, E = e_n(A)$. Действительно, матрица $e_n(\lambda)\, E$ может быть записана как матричный λ-многочлен, коэффициенты которого являются скалярными матрицами, т. е. перестановочны с матрицей A. Таким образом

$$e_n(A) = 0,$$

т. е. многочлен $e_n(\lambda)$ действительно аннулируется матрицей A.

Отсюда следует, что многочлен $e_n(\lambda)$ нацело делится на минимальный многочлен $m(\lambda)$ матрицы A,

$$e_n(\lambda) = m(\lambda)\, q(\lambda). \tag{5}$$

Ясно, что старший коэффициент многочлена $q(\lambda)$ равен единице.

Так как $m(A) = 0$, то снова, ввиду той же леммы из § 60, остаток от левого деления λ-матрицы $m(\lambda)E$ на двучлен $\lambda E - A$ равен нулю, т. е.

$$m(\lambda)\, E = (\lambda E - A)\, Q(\lambda). \tag{6}$$

Равенства (5), (4) и (6) приводят к равенству

$$(\lambda E - A)\,[(-1)^{n+1} C(\lambda)] = (\lambda E - A)\,[Q(\lambda)\, q(\lambda)].$$

Обе части этого равенства можно сократить на общий множитель $\lambda E - A$, так как старший коэффициент E этого матричного λ-многочлена является невырожденной матрицей. Таким образом,

$$C(\lambda) = (-1)^{n+1} Q(\lambda)\, q(\lambda).$$

Мы помним, однако, что наибольший общий делитель элементов матрицы $C(\lambda)$ равен 1. Поэтому многочлен $q(\lambda)$ должен иметь нулевую степень, а так как его старший коэффициент равен 1, то $q(\lambda) = 1$. Таким образом, ввиду (5),

$$e_n(\lambda) = m(\lambda),$$

что и требовалось доказать.

Так как, ввиду (1), характеристический многочлен матрицы A нацело делится на многочлен $e_n(\lambda)$, то из доказанной сейчас теоремы вытекает следующая

Теорема Гамильтона—Кэли. *Всякая матрица является корнем своего характеристического многочлена.*

Минимальный многочлен линейного преобразования. Докажем сначала следующее утверждение:

Если матрицы A и B подобны и если многочлен $f(\lambda)$ аннулируется матрицей A, то он аннулируется и матрицей B.

Действительно, пусть

$$B = C^{-1} A C.$$

Если

$$f(\lambda) = \alpha_0 \lambda^k + \alpha_1 \lambda^{k-1} + \ldots + \alpha_{k-1}\lambda + \alpha_k,$$

то

$$\alpha_0 A^k + \alpha_1 A^{k-1} + \ldots + \alpha_{k-1} A + \alpha_k E = 0.$$

Трансформируя обе части этого равенства матрицей C, получаем:

$$C^{-1}(\alpha_0 A^k + \alpha_1 A^{k-1} + \ldots + \alpha_{k-1} A + \alpha_k E)\, C =$$
$$= \alpha_0 (C^{-1} AC)^k + \alpha_1 (C^{-1} AC)^{k-1} + \ldots + \alpha_{k-1} (C^{-1} AC) + \alpha_k E =$$
$$= \alpha_0 B^k + \alpha_1 B^{k-1} + \ldots + \alpha_{k-1} B + \alpha_k E = 0,$$

т. е. $f(B) = 0$.

Отсюда следует, что *подобные матрицы обладают одним и тем же минимальным многочленом.*

Пусть теперь φ будет линейное преобразование n-мерного линейного пространства над полем P. Матрицы, задающие это преобразование в разных базах пространства, подобны между собой. Общий минимальный многочлен этих матриц называется *минимальным многочленом линейного преобразования* φ.

Используя операции над линейными преобразованиями, введенные в § 32, можно ввести понятие *значения* многочлена

$$f(\lambda) = \alpha_0 \lambda^k + \alpha_1 \lambda^{k-1} + \ldots + \alpha_{k-1} \lambda + \alpha_k$$

из кольца $P[\lambda]$ при λ, равном линейному преобразованию φ: это будет линейное преобразование

$$f(\varphi) = \alpha_0 \varphi^k + \alpha_1 \varphi^{k-1} + \ldots + \alpha_{k-1} \varphi + \alpha_k \varepsilon,$$

где ε — тождественное преобразование.

Мы скажем, далее, что многочлен $f(\lambda)$ *аннулируется* линейным преобразованием φ, если

$$f(\varphi) = \omega,$$

где ω — нулевое преобразование.

Учитывая связь между операциями над линейными преобразованиями и над матрицами, читатель без труда докажет, что *минимальный многочлен линейного преобразования* φ *является тем однозначно определенным многочленом наименьшей степени со старшим коэффициентом* 1, *который аннулируется преобразованием* φ. После этого результаты, полученные выше, в частности теорема Гамильтона—Кэли, могут быть переформулированы на языке линейных преобразований.