Для разыскания матрицы R_2, трансформирующей A в B, найдем какую-либо цепочку элементарных преобразований, переводящую $A-\lambda E$ в $B-\lambda E$. Так,

$$A-\lambda E=\begin{pmatrix}-2-\lambda & 1\\ 0 & 3-\lambda\end{pmatrix}\sim\begin{pmatrix}-2-\lambda & 1\\ -16-8\lambda & 11-\lambda\end{pmatrix}\sim\begin{pmatrix}8+4\lambda & -4\\ -16-8\lambda & 11-\lambda\end{pmatrix}\sim$$

$$\sim\begin{pmatrix}40+4\lambda & -4\\ -104 & 11-\lambda\end{pmatrix}\sim\begin{pmatrix}-10-\lambda & -4\\ 26 & 11-\lambda\end{pmatrix}=B-\lambda E.$$

К столбцам относятся два последних преобразования: к первому столбцу прибавляется второй, умноженный на -8, а затем первый столбец умножается на $-\frac{1}{4}$. Произведение соответствующих элементарных матриц будет

$$V(\lambda)=\begin{pmatrix}1 & 0\\ -8 & 1\end{pmatrix}\begin{pmatrix}-\frac{1}{4} & 0\\ 0 & 1\end{pmatrix}=\begin{pmatrix}-\frac{1}{4} & 0\\ 2 & 1\end{pmatrix}.$$

Эта матрица не зависит от λ и поэтому она и будет искомой матрицей R_2.

Конечно, матрица, трансформирующая A в B, определяется далеко не однозначно. Такой будет, например, также матрица

$$\begin{pmatrix}3 & 1\\ 2 & 1\end{pmatrix}.$$

§ 61. Жорданова нормальная форма

Будем рассматривать сейчас квадратные матрицы порядка n с элементами из поля P. Будет выделен один специальный тип таких матриц, так называемые жордановы матрицы, и будет показано, что эти матрицы служат нормальной формой для весьма широкого класса матриц. Именно, *матрицы, все характеристические корни которых лежат в основном поле P (и только такие матрицы), подобны некоторым жордановым матрицам, т. е., как говорят, они приводятся к жордановой нормальной форме.* Отсюда будет следовать, если в качестве поля P взято поле комплексных чисел, что *всякая матрица с комплексными элементами приводится в поле комплексных чисел к жордановой нормальной форме.*

Введем необходимые определения. *Жордановой клеткой* порядка k, относящейся к числу λ_0, называется матрица порядка k, $1\leqslant k\leqslant n$, имеющая вид

$$\begin{pmatrix}\lambda_0 & 1 & & & \mathbf{0}\\ & \lambda_0 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1\\ \mathbf{0} & & & & \lambda_0\end{pmatrix}; \qquad (1)$$

иными словами, на ее главной диагонали стоит одно и то же число λ_0 из поля P; параллель, ближайшая к главной диагонали сверху, сплошь занята числом 1; все остальные элементы матрицы равны нулю. Так,

$$(\lambda_0), \quad \begin{pmatrix} \lambda_0 & 1 \\ 0 & \lambda_0 \end{pmatrix}, \quad \begin{pmatrix} \lambda_0 & 1 & 0 \\ 0 & \lambda_0 & 1 \\ 0 & 0 & \lambda_0 \end{pmatrix}$$

будут соответственно жордановыми клетками первого, второго и третьего порядков.

Жордановой матрицей порядка n называется матрица порядка n, имеющая вид

$$J = \begin{pmatrix} \boxed{J_1} & & & 0 \\ & \boxed{J_2} & & \\ & & \ddots & \\ 0 & & & \boxed{J_s} \end{pmatrix}; \tag{2}$$

здесь вдоль главной диагонали идут жордановы клетки $J_1, J_2, \ldots, J_s$ некоторых порядков, не обязательно различных, относящиеся к некоторым числам из поля P, также не обязательно различным; все места вне этих клеток заняты нулями. При этом $s \geqslant 1$, т. е. одна жорданова клетка порядка n принадлежит к числу жордановых матриц этого порядка, и, понятно, $s \leqslant n$.

Заметим, хотя это и не будет дальше использоваться, что строение жордановой матрицы можно было бы описать, не прибегая к понятию жордановой клетки. Очевидно, именно, что матрица J будет жордановой матрицей тогда и только тогда, если она имеет вид

$$\begin{pmatrix} \lambda_1 & \varepsilon_1 & & & 0 \\ & \lambda_2 & \varepsilon_2 & & \\ & & \ddots & \ddots & \\ & & & \ddots & \varepsilon_{n-1} \\ 0 & & & & \lambda_n \end{pmatrix},$$

где λ_i, $i = 1, 2, \ldots, n$,— произвольные числа из поля P, а каждое ε_j, $j = 1, 2, \ldots, n-1$, равно единице или нулю, причем, если $\varepsilon_j = 1$, то $\lambda_j = \lambda_{j+1}$.

Диагональные матрицы являются частным случаем жордановых матриц: это будут в точности те жордановы матрицы, у которых все жордановы клетки имеют порядок 1.

Нашей ближайшей целью является разыскание канонического вида для характеристической матрицы $J - \lambda E$

произвольной жордановой матрицы J порядка n. Найдем сначала канонический вид для характеристической матрицы

$$\begin{pmatrix} \lambda_0-\lambda & 1 & & 0 \\ & \lambda_0-\lambda & 1 & \\ & & \ddots & \ddots \\ & & & 1 \\ 0 & & & \lambda_0-\lambda \end{pmatrix} \tag{3}$$

одной жордановой клетки (1) порядка k. Вычисляя определитель этой матрицы и вспоминая, что старший коэффициент многочлена $d_k(\lambda)$ должен равняться 1, получаем, что

$$d_k(\lambda)=(\lambda-\lambda_0)^k.$$

С другой стороны, среди миноров $(k-1)$-го порядка матрицы (3) имеется минор, равный единице, а именно тот, который получается после вычеркивания первого столбца и последней строки этой матрицы. Поэтому

$$d_{k-1}(\lambda)=1.$$

Отсюда следует, что *каноническим видом для матрицы* (3) *служит следующая λ-матрица порядка k*:

$$\begin{pmatrix} 1 & & & 0 \\ & \ddots & & \\ & & 1 & \\ 0 & & & (\lambda-\lambda_0)^k \end{pmatrix}. \tag{4}$$

Докажем теперь следующую лемму:

Если многочлены $\varphi_1(\lambda), \varphi_2(\lambda), \ldots, \varphi_t(\lambda)$ из кольца $P[\lambda]$ попарно взаимно просты, то имеет место следующая эквивалентность:

$$\begin{pmatrix} \varphi_1(\lambda) & & & 0 \\ & \varphi_2(\lambda) & & \\ & & \ddots & \\ 0 & & & \varphi_t(\lambda) \end{pmatrix} \sim \begin{pmatrix} 1 & & & 0 \\ & \ddots & & \\ & & 1 & \\ 0 & & & \prod_{i=1}^{t}\varphi_i(\lambda) \end{pmatrix}.$$

Достаточно, очевидно, рассмотреть случай $t=2$. Так как многочлены $\varphi_1(\lambda)$ и $\varphi_2(\lambda)$ взаимно просты, то в кольце $P[\lambda]$ существуют такие многочлены $u_1(\lambda)$ и $u_2(\lambda)$, что

$$\varphi_1(\lambda)u_1(\lambda)+\varphi_2(\lambda)u_2(\lambda)=1.$$

Поэтому

$$\begin{pmatrix} \varphi_1(\lambda) & 0 \\ 0 & \varphi_2(\lambda) \end{pmatrix} \sim \begin{pmatrix} \varphi_1(\lambda) & \varphi_1(\lambda) u_1(\lambda) \\ 0 & \varphi_2(\lambda) \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} \varphi_1(\lambda) & \varphi_1(\lambda) u_1(\lambda) + \varphi_2(\lambda) u_2(\lambda) \\ 0 & \varphi_2(\lambda) \end{pmatrix} = \begin{pmatrix} \varphi_1(\lambda) & 1 \\ 0 & \varphi_2(\lambda) \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & \varphi_1(\lambda) \\ \varphi_2(\lambda) & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & \varphi_1(\lambda) \\ 0 & -\varphi_1(\lambda)\varphi_2(\lambda) \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & 0 \\ 0 & -\varphi_1(\lambda)\varphi_2(\lambda) \end{pmatrix} \sim \begin{pmatrix} 1 & 0 \\ 0 & \varphi_1(\lambda)\varphi_2(\lambda) \end{pmatrix},$$

что и требовалось доказать.

Перейдем теперь к рассмотрению характеристической матрицы

$$J - \lambda E = \begin{pmatrix} \boxed{J_1 - \lambda E_1} & & & \mathbf{0} \\ & \boxed{J_2 - \lambda E_2} & & \\ & & \ddots & \\ \mathbf{0} & & & \boxed{J_s - \lambda E_s} \end{pmatrix} \tag{5}$$

для жордановой матрицы J вида (2); здесь E_i, $i = 1, 2, \ldots, s$, есть единичная матрица того же порядка, что и клетка J_i. Пусть жордановы клетки матрицы J относятся к следующим различным числам: $\lambda_1, \lambda_2, \ldots, \lambda_t$, где $t \leqslant s$. Пусть, далее, к числу λ_i, $i = 1, 2, \ldots, t$, относится q_i жордановых клеток, $q_i \geqslant 1$, и пусть порядки этих клеток, расположенные в невозрастающем порядке, будут

$$k_{i1} \geqslant k_{i2} \geqslant \ldots \geqslant k_{iq_i}. \tag{6}$$

Отметим, хотя и не будем этим пользоваться, что

$$\sum_{i=1}^{t} q_i = s,$$

$$\sum_{i=1}^{t} \sum_{j=1}^{q_i} k_{ij} = n.$$

Применяя элементарные преобразования к тем строкам и столбцам матрицы (5), которые проходят через клетку $J_i - \lambda E_i$ этой матрицы, мы не будем затрагивать, очевидно, других диагональных клеток. Отсюда следует, что в матрице (5) можно при помощи элементарных преобразований заменить каждую клетку $J_i - \lambda E_i$, $i = 1, 2, \ldots, s$, соответствующей клеткой вида (4). Иными словами, *матрица $J - \lambda E$ эквивалентна диагональной матрице, на диагонали которой стоят, помимо некоторого числа единиц, также*

следующие многочлены, соответствующие всем жордановым клеткам матрицы J:

$$\left.\begin{array}{c}(\lambda-\lambda_1)^{k_{11}},\ (\lambda-\lambda_1)^{k_{12}},\ \ldots,\ (\lambda-\lambda_1)^{k_{1q_1}},\\ (\lambda-\lambda_2)^{k_{21}},\ (\lambda-\lambda_2)^{k_{22}},\ \ldots,\ (\lambda-\lambda_2)^{k_{2q_2}},\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ (\lambda-\lambda_t)^{k_{t1}},\ (\lambda-\lambda_t)^{k_{t2}},\ \ldots,\ (\lambda-\lambda_t)^{k_{tq_t}}.\end{array}\right\} \tag{7}$$

Мы не указываем при этом те места на диагонали, на которых стоят многочлены (7), так как в любой диагональной λ-матрице диагональные элементы можно произвольно переставлять при помощи перестановок строк и одноименных столбцов. Это замечание следует учитывать и в дальнейшем.

Пусть q — наибольшее среди чисел q_i, $i=1, 2, \ldots, t$. Обозначим через $e_{n-j+1}(\lambda)$ произведение многочленов, стоящих в j-м столбце таблицы (7), $j=1, 2, \ldots, q$, т. е.

$$e_{n-j+1}(\lambda)=\prod_{i=1}^{t}(\lambda-\lambda_i)^{k_{ij}}; \tag{8}$$

если при этом в j-м столбце имеются пустые места — для некоторых i может оказаться, что $q_i<j$, — то соответствующие множители в (8) считаем равными единице. Так как числа $\lambda_1, \lambda_2, \ldots, \lambda_t$ по условию различные, то степени линейных двучленов, стоящие в j-м столбце таблицы (7), попарно взаимно просты. Поэтому, на основании доказанной выше леммы, они при помощи элементарных преобразований могут быть заменены в рассматриваемой диагональной матрице их произведением $e_{n-j+1}(\lambda)$ и некоторым числом единиц.

Проделав это для $j=1, 2, \ldots, q$, мы получим, что

$$J-\lambda E \sim \left\{\begin{array}{ccccccc} 1 & & & & & & 0\\ & \cdot & & & & & \\ & & \cdot & & & & \\ & & & \cdot & & & \\ & & & & 1 & & \\ & & & & & e_{n-q+1}(\lambda) & \\ & & & & & & \cdot \\ & & & & & & & \cdot \\ & & & & & & & & e_{n-1}(\lambda)\\ 0 & & & & & & & & e_n(\lambda)\end{array}\right\}. \tag{9}$$

Это и будет искомый канонический вид матрицы $J-\lambda E$. Действительно, старшие коэффициенты всех многочленов, стоящих в (9) на главной диагонали, равны единице и каждый из этих многочленов нацело делится на предыдущий ввиду условия (6).

Пример. Пусть

$$J=\begin{pmatrix} \boxed{\begin{matrix}2&1&0\\0&2&1\\0&0&2\end{matrix}} & & & & & \mathbf{0} \\ & \boxed{2} & & & & \\ & & \boxed{\begin{matrix}5&1\\0&5\end{matrix}} & & & \\ & & & \boxed{\begin{matrix}5&1\\0&5\end{matrix}} & & \\ \mathbf{0} & & & & \boxed{2} & \end{pmatrix}.$$

Для этой жордановой матрицы 9-го порядка таблица многочленов (7) имеет вид

$$(\lambda-2)^3,\quad \lambda-2,\quad \lambda-2,$$

$$(\lambda-5)^2,\quad (\lambda-5)^2.$$

Поэтому инвариантными множителями матрицы J будут многочлены

$$e_9(\lambda)=(\lambda-2)^3(\lambda-5)^2,$$

$$e_8(\lambda)=(\lambda-2)(\lambda-5)^2,$$

$$e_7(\lambda)=\lambda-2,$$

в то время как $e_6(\lambda)=\ldots=e_1(\lambda)=1$.

Теперь, когда мы научились по виду данной жордановой матрицы J сразу писать канонический вид ее характеристической матрицы, может быть доказана следующая теорема:

Две жордановы матрицы тогда и только тогда подобны, если они состоят из одних и тех же жордановых клеток, т. е. отличаются, быть может, лишь расположением этих клеток вдоль главной диагонали.

В самом деле, таблица многочленов (7) полностью определялась набором жордановых клеток жордановой матрицы J и в ней никак не отражалось расположение жордановых клеток вдоль главной диагонали этой матрицы. Отсюда следует, что если жордановы матрицы J и J' обладают одним и тем же набором жордановых клеток, то им соответствует одна и та же таблица многочленов (7), а поэтому одни и те же многочлены (8). Таким образом, характеристические матрицы $J-\lambda E$ и $J'-\lambda E$ обладают одинаковыми инвариантными множителями, т. е. эквивалентны, а поэтому сами матрицы J и J' подобны.

Обратно, если жордановы матрицы J и J' подобны, то их характеристические матрицы обладают одинаковыми инвариантными множителями. Пусть многочлены (8) для $j=1, 2, \ldots, q$ будут те из этих инвариантных множителей, которые отличны от единицы. Однако по многочленам (8) восстанавливается таблица многочленов (7). Именно, многочлены (8) разлагаются в произведение степеней линейных множителей, так как этим свойством обладают, как уже доказано, инвариантные множители характеристической матрицы для любой жордановой матрицы. Таблица (7) как раз и состоит из всех тех максимальных степеней линейных множителей, на которые разлагаются многочлены (8). Наконец, по таблице (7) восстанавливаются жордановы клетки исходных жордановых матриц: каждому многочлену $(\lambda-\lambda_i)^{k_{ij}}$ из таблицы (7) соответствует жорданова клетка порядка k_{ij}, относящаяся к числу λ_i. Этим доказано, что матрицы J и J' состоят из одних и тех же жордановых клеток и отличаются, быть может, лишь их расположением.

Из этой теоремы следует, в частности, *что жорданова матрица, подобная диагональной матрице, сама диагональна и что две диагональные матрицы тогда и только тогда подобны, если получаются друг из друга перестановкой чисел, стоящих на главной диагонали.*

Приведение матрицы к жордановой нормальной форме. Если матрица A с элементами из поля P *приводится к жордановой нормальной форме,* т. е. подобна жордановой матрице, то, как следует из доказанной выше теоремы, *жорданова нормальная форма определяется для матрицы A однозначно с точностью до расположения жордановых клеток на главной диагонали.* Условие для того, чтобы матрица A допускала такое приведение, указывается в следующей теореме, доказательство которой дает одновременно практический способ для разыскания жордановой матрицы, подобной матрице A, если такая жорданова матрица существует. При этом заметим, что приводимость в поле P означает, что все элементы трансформирующей матрицы содержатся в поле P.

Матрица A с элементами из поля P тогда и только тогда приводится в поле P к жордановой нормальной форме, если все характеристические корни матрицы A лежат в самом основном поле P.

В самом деле, если матрица A подобна жордановой матрице J, то эти две матрицы обладают одними и теми же характеристическими корнями. Характеристические корни матрицы J находятся, однако, без всяких затруднений: так как определитель матрицы $J-\lambda E$ равен произведению ее элементов, стоящих на главной диагонали, то многочлен $|J-\lambda E|$ разлагается над полем P на линейные множители и его корнями служат числа, стоящие на главной диагонали матрицы J, и только они.

Обратно, пусть все характеристические корни матрицы A лежат в самом поле P. Если отличные от 1 инвариантные множители матрицы $A-\lambda E$ будут

$$e_{n-q+1}(\lambda),\ \ldots,\ e_{n-1}(\lambda),\ e_n(\lambda), \tag{10}$$

то

$$|A-\lambda E|=(-1)^n e_{n-q+1}(\lambda)\ldots e_{n-1}(\lambda)\,e_n(\lambda).$$

Действительно, определители матрицы $A-\lambda E$ и ее канонической матрицы могут отличаться друг от друга лишь постоянным множителем, который на самом деле равен $(-1)^n$, так как именно таков старший коэффициент характеристического многочлена $|A-\lambda E|$. Таким образом, среди многочленов (10) нет равных нулю, сумма степеней этих многочленов равна n и все они разлагаются над полем P на линейные множители — последнее ввиду того, что, по условию, многочлен $|A-\lambda E|$ обладает таким разложением.

Пусть (8) будут разложения многочленов (10) в произведения степеней линейных множителей. Назовем *элементарными делителями многочлена* e_{n-j+1}, $j=1, 2, \ldots, q$, отличные от единицы степени различных линейных двучленов, входящие в его разложение (8), т. е.

$$(\lambda-\lambda_1)^{k_{1j}},\ (\lambda-\lambda_2)^{k_{2j}},\ \ldots,\ (\lambda-\lambda_t)^{k_{tj}}.$$

Элементарные делители всех многочленов (10) назовем *элементарными делителями матрицы* A и выпишем их в виде таблицы (7).

Возьмем теперь жорданову матрицу J порядка n, составленную из жордановых клеток, определяемых следующим образом: каждому элементарному делителю $(\lambda-\lambda_i)^{k_{ij}}$ матрицы A ставим в соответствие жорданову клетку порядка k_{ij}, относящуюся к числу λ_i. Очевидно, что отличными от 1 инвариантными множителями матрицы $J-\lambda E$ будут многочлены (10) и только они. Поэтому матрицы $A-\lambda E$ и $J-\lambda E$ эквивалентны и, следовательно, матрица A подобна жордановой матрице J.

Пример. Пусть дана матрица

$$A=\begin{pmatrix} -16 & -17 & 87 & -108 \\ 8 & 9 & -42 & 54 \\ -3 & -3 & 16 & -18 \\ -1 & -1 & 6 & -8 \end{pmatrix}.$$

Приводя обычным способом матрицу $A-\lambda E$ к каноническому виду, получим, что отличными от единицы инвариантными множителями этой матрицы будут многочлены

$$e_4(\lambda)=(\lambda-1)^2(\lambda+2),$$
$$e_3(\lambda)=\lambda-1.$$

Мы видим, что матрица A приводится к жордановой нормальной форме даже в поле рациональных чисел. Ее элементарными делителями являются много-

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

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