ГЛАВА III

КАНОНИЧЕСКИЙ ВИД ПРОИЗВОЛЬНЫХ ЛИНЕЙНЫХ ПРЕОБРАЗОВАНИЙ

§ 18. Нормальная форма линейного преобразования

В главе II мы познакомились с различными классами линейных преобразований n-мерного пространства, имеющих n линейно независимых собственных векторов. Мы знаем, что в базисе, состоящем из собственных векторов такого преобразования, его матрица имеет особенно простой вид, так называемую диагональную форму.

Однако число линейно независимых собственных векторов у линейного преобразования может быть меньше, чем n *). Пример линейного преобразования с недостаточным числом собственных векторов мы приведем несколько позже (см. также § 10, п. 1, пример 3). Такое преобразование заведомо не может быть приведено к диагональной форме, так как базис, в котором матрица преобразования диагональна, состоит из соб-

*) Напомним, что если все корни характеристического многочлена различны, то преобразование имеет n линейно независимых собственных векторов. Поэтому, для того чтобы число собственных векторов было меньше чем n, необходимо наличие кратных корней у характеристического многочлена. Таким образом, этот случай является в некотором смысле исключительным.

ственных векторов. Возникает вопрос: каков простейший вид матрицы такого линейного преобразования?

В этой главе мы для произвольного преобразования укажем базис, в котором его матрица имеет сравнительно простой вид (так называемая *жорданова нормальная форма*). В случае, когда число линейно независимых собственных векторов преобразования равно размерности пространства, эта нормальная форма совпадает с диагональной. Мы сформулируем сейчас окончательный результат, который докажем в § 19.

Пусть задано произвольное линейное преобразование A в комплексном пространстве n измерений. Предположим, что у A имеется k ($k \leqslant n$) линейно независимых собственных векторов

$$e_1, f_1, \ldots, h_1,$$

соответствующих собственным значениям $\lambda_1, \lambda_2, \ldots$ $\ldots, \lambda_k$. Тогда существует базис, состоящий из k групп векторов *):

$$e_1, \ldots, e_p;\ f_1, \ldots, f_q;\ \ldots;\ h_1, \ldots, h_s, \tag{1}$$

в котором преобразование A имеет следующий вид:

$$\begin{aligned}
&Ae_1 = \lambda_1 e_1, \quad Ae_2 = e_1 + \lambda_1 e_2, \quad \ldots, \quad Ae_p = e_{p-1} + \lambda_1 e_p;\\
&Af_1 = \lambda_2 f_1, \quad Af_2 = f_1 + \lambda_2 f_2, \quad \ldots, \quad Af_q = f_{q-1} + \lambda_2 f_q;\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&Ah_1 = \lambda_k h_1, \quad Ah_2 = h_1 + \lambda_k h_2, \quad \ldots, \quad Ah_s = h_{s-1} + \lambda_k h_s.
\end{aligned} \tag{2}$$

Мы видим, что базисные векторы каждой группы переходят при нашем преобразовании в линейную комбинацию векторов той же группы. Отсюда следует, что каждая группа базисных векторов порождает подпространство, инвариантное относительно

*) Ясно, что $p+q+\ldots+s = n$. Если же $k = n$, то каждая группа состоит из одного вектора, а именно собственного вектора.

преобразования A. Рассмотрим несколько подробнее преобразование, задаваемое формулами (2).

В подпространстве, порожденном каждой группой, есть собственный вектор; например, в подпространстве, порожденном векторами $e_1, e_2, \ldots, e_p$, таким собственным вектором является e_1.

Вектор e_2 называют иногда присоединенным собственным вектором первого порядка. Это значит, что Ae_2 пропорционально e_2 с точностью до собственного вектора, как это видно из равенства

$$Ae_2 = \lambda_1 e_2 + e_1.$$

Аналогично $e_3, e_4, \ldots$ называют присоединенными векторами второго, третьего и т. д. порядков.

Каждый из них является «как бы собственным», т. е. собственным с точностью до присоединенного вектора низшего порядка

$$Ae_k = \lambda_1 e_k + e_{k-1}.$$

Таким образом, базис каждого инвариантного подпространства состоит из одного собственного вектора и такого же количества присоединенных, которое нужно добавить, чтобы получить базис данного подпространства.

Покажем, что в каждом из этих подпространств имеется, с точностью до множителя, *лишь один* собственный вектор.

Действительно, рассмотрим, например, подпространство, порожденное векторами $e_1, e_2, \ldots, e_p$. Допустим, что некоторый вектор из этого подпространства, т. е. некоторая линейная комбинация вида

$$c_1 e_1 + c_2 e_2 + \ldots + c_p e_p,$$

где не все c_k равны нулю, является собственным вектором, т. е.

$$A(c_1 e_1 + c_2 e_2 + \ldots + c_p e_p) = \lambda(c_1 e_1 + c_2 e_2 + \ldots + c_p e_p).$$

Подставляя вместо левой части ее выражение по формулам (2), получаем равенство

$$c_1\lambda_1 e_1 + c_2(e_1 + \lambda_1 e_2) + \ldots + c_p(e_{p-1} + \lambda_1 e_p) = \\ = \lambda c_1 e_1 + \lambda c_2 e_2 + \ldots + \lambda c_p e_p.$$

Отсюда, приравнивая коэффициенты при каждом из базисных векторов, имеем систему уравнений для нахождения величин $\lambda, c_1, c_2, \ldots, c_p$:

$$\begin{aligned} c_1\lambda_1 + c_2 &= \lambda c_1, \\ c_2\lambda_1 + c_3 &= \lambda c_2, \\ &\ldots\ldots\ldots \\ c_{p-1}\lambda_1 + c_p &= \lambda c_{p-1}, \\ c_p\lambda_1 \qquad &= \lambda c_p. \end{aligned}$$

Покажем прежде всего, что $\lambda = \lambda_1$. Действительно, если бы $\lambda \neq \lambda_1$, то из последнего равенства мы имели бы $c_p = 0$ и затем из остальных равенств $c_{p-1} = c_{p-2} = c_2 = c_1 = 0$. Итак, $\lambda = \lambda_1$; тогда из первого уравнения имеем $c_2 = 0$, из второго $c_3 = 0$ и т. д. до $c_p = 0$. Значит, собственный вектор равен $c_1 e_1$ и, следовательно, с точностью до множителя совпадает с первым вектором соответствующей группы.

Выпишем матрицу преобразования (2). Так как векторы каждой группы преобразуются в линейные комбинации векторов той же группы, то в первых p столбцах матрицы преобразования могут быть отличны от нуля лишь элементы первых p строк, в следующих q столбцах могут быть отличны от нуля лишь элементы, стоящие в строках с теми же номерами, что и у этих столбцов, и т. д. Таким образом, в данном базисе матрица преобразования будет состоять из k клеток, расположенных по главной диагонали, а все элементы, не принадлежащие ни одной из этих клеток, будут равны нулю.

Для того чтобы понять, что стоит в каждой клетке матрицы преобразования A, достаточно еще раз написать, как преобразуются векторы одной группы. Мы имеем:

$$\begin{aligned}
Ae_1 &= \lambda_1 e_1, \\
Ae_2 &= \quad e_1 + \lambda_1 e_2, \\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
Ae_{p-1} &= \qquad\qquad\qquad e_{p-2} + \lambda_1 e_{p-1}, \\
Ae_p &= \qquad\qquad\qquad\qquad\qquad e_{p-1} + \lambda_1 e_p.
\end{aligned}$$

Вспоминая, как строится матрица, отвечающая данному преобразованию базиса, получаем, что клетка матрицы, соответствующая данной группе векторов, имеет вид

$$\boldsymbol{A}_1 = \begin{pmatrix} \lambda_1 & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_1 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_1 & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda_1 \end{pmatrix}. \tag{3}$$

Вся же матрица оказывается составленной из таких клеток порядков $p, q, \dots, s$ соответственно, т. е. имеет вид

$$\left(\begin{array}{ccccc|ccccc|c|ccccc}
\lambda_1 & 1 & 0 & \dots & 0 & & & & & & & & & & & \\
0 & \lambda_1 & 1 & \dots & 0 & & & & & & & & & & & \\
\dots & \dots & \dots & \dots & \dots & & & & & & & & & & & \\
0 & 0 & 0 & \dots & \lambda_1 & & & & & & & & & & & \\
 & & & & & \lambda_2 & 1 & 0 & \dots & 0 & & & & & & \\
 & & & & & 0 & \lambda_2 & 1 & \dots & 0 & & & & & & \\
 & & & & & \dots & \dots & \dots & \dots & \dots & & & & & & \\
 & & & & & 0 & 0 & 0 & \dots & \lambda_2 & & & & & & \\
 & & & & & & & & & & \ddots & & & & & \\
 & & & & & & & & & & & \lambda_k & 1 & 0 & \dots & 0 \\
 & & & & & & & & & & & 0 & \lambda_k & 1 & \dots & 0 \\
 & & & & & & & & & & & \dots & \dots & \dots & \dots & \dots \\
 & & & & & & & & & & & 0 & 0 & 0 & \dots & \lambda_k
\end{array}\right), \tag{4}$$

где все элементы вне клеток — нули.

Заметим также, что не все λ_i обязаны быть различными.

Упражнение. Найти все инвариантные подпространства преобразования с матрицей (3).

Хотя приведенная здесь нормальная форма выглядит сложнее, чем, например, диагональная матрица, однако и с ней можно достаточно просто производить алгебраические операции. Мы покажем, например, как вычислить многочлен от матрицы (4). Матрица (4) имеет вид

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_1 & & & \\ & \boldsymbol{A}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{A}_k \end{pmatrix},$$

где $\boldsymbol{A}_i$ — отдельные клетки, а все невыписанные элементы — нули. Тогда

$$\boldsymbol{A}^2 = \begin{pmatrix} \boldsymbol{A}_1^2 & & & \\ & \boldsymbol{A}_2^2 & & \\ & & \ddots & \\ & & & \boldsymbol{A}_k^2 \end{pmatrix}, \ \ldots, \ \boldsymbol{A}^m = \begin{pmatrix} \boldsymbol{A}_1^m & & & \\ & \boldsymbol{A}_2^m & & \\ & & \ddots & \\ & & & \boldsymbol{A}_k^m \end{pmatrix},$$

т. е. для того, чтобы возвести в некоторую степень матрицу $\boldsymbol{A}$, достаточно уметь возвести в эту степень каждую из клеток. Пусть теперь $P(t) = a_0 + a_1 t + \ldots + a_m t^m$ — произвольный многочлен. Тогда легко видеть, что

$$P(\boldsymbol{A}) = \begin{pmatrix} P(\boldsymbol{A}_1) & & & \\ & P(\boldsymbol{A}_2) & & \\ & & \ddots & \\ & & & P(\boldsymbol{A}_k) \end{pmatrix}.$$

Покажем теперь, как вычислить $P(\boldsymbol{A}_1)$, т. е. многочлен от одной клетки нормальной формы матрицы (3). Для этого запишем матрицу (3) в виде

$$\boldsymbol{A}_1 = \lambda_1 \boldsymbol{E} + \boldsymbol{I},$$

где $\boldsymbol{E}$ — единичная матрица порядка p, а матрица $\boldsymbol{I}$ имеет вид

$$\boldsymbol{I} = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \multicolumn{6}{c}{\dotfill} \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & 0 & \ldots & 0 & 0 \end{pmatrix}.$$

Заметим, что матрицы $\boldsymbol{I}^2, \boldsymbol{I}^3, \ldots, \boldsymbol{I}^{p-1}$ имеют следующий вид *):

$$\boldsymbol{I}^2 = \begin{pmatrix} 0 & 0 & 1 & . & \ldots & 0 \\ 0 & 0 & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & . & \ldots & 0 \\ 0 & 0 & 0 & . & \ldots & 0 \end{pmatrix}, \ \ldots, \ \boldsymbol{I}^{p-1} = \begin{pmatrix} 0 & 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & 0 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 0 & \ldots & 0 & 0 \end{pmatrix},$$

а

$$\boldsymbol{I}^p = \boldsymbol{I}^{p+1} = \ldots = 0.$$

Теперь нетрудно вычислить произвольный многочлен от матрицы (3). Действительно, многочлен $P(t)$ можно по формуле Тейлора представить в виде

$$P(t) = P(\lambda_1) + (t-\lambda_1)P'(\lambda_1) + \frac{(t-\lambda_1)^2}{2!}P''(\lambda_1) + \ldots \\ \ldots + \frac{(t-\lambda_1)^n}{n!}P^{(n)}(\lambda_1),$$

где n — степень многочлена. Подставляя вместо t матрицу $\boldsymbol{A}_1$, имеем:

$$P(\boldsymbol{A}_1) = P(\lambda_1)\boldsymbol{E} + (\boldsymbol{A}_1 - \lambda_1\boldsymbol{E})P'(\lambda_1) + \frac{(\boldsymbol{A}_1 - \lambda_1\boldsymbol{E})^2}{2!}P''(\lambda_1) + \ldots \\ \ldots + \frac{(\boldsymbol{A}_1 - \lambda_1\boldsymbol{E})^n}{n!}P^{(n)}(\lambda_1).$$

Но $\boldsymbol{A}_1 - \lambda_1\boldsymbol{E} = \boldsymbol{I}$, следовательно,

$$P(\boldsymbol{A}_1) = P(\lambda_1)\boldsymbol{E} + P'(\lambda_1)\boldsymbol{I} + \frac{P''(\lambda_1)}{2!}\boldsymbol{I}^2 + \ldots + \frac{P^{(n)}(\lambda_1)}{n!}\boldsymbol{I}^n.$$

Подставляя вместо $\boldsymbol{I}, \boldsymbol{I}^2, \ldots, \boldsymbol{I}^{p-1}$ их выражения и учитывая, что $\boldsymbol{I}^p = \boldsymbol{I}^{p+1} = \ldots = 0$, получаем окончательный вид матрицы $P(\boldsymbol{A}_1)$:

$$P(\boldsymbol{A}_1) = \begin{pmatrix} P(\lambda_1) & \frac{P'(\lambda_1)}{1!} & \frac{P''(\lambda_1)}{2!} & \ldots & \frac{P^{(p-1)}(\lambda_1)}{(p-1)!} \\ 0 & P(\lambda_1) & \frac{P'(\lambda_1)}{1!} & \ldots & \frac{P^{(p-2)}(\lambda_1)}{(p-2)!} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & P(\lambda_1) \end{pmatrix}.$$

---

*) Проще всего это сосчитать так. Мы имеем $Ie_1 = 0$, $Ie_2 = e_1, \ \ldots, \ Ie_p = e_{p-1}$. Следовательно, $I^2e_1 = 0$, $I^2e_2 = 0$, $I^2e_3 = e_1, \ldots, I^2e_p = e_{p-2}$. Аналогично, $I^3e_1 = I^3e_2 = I^3e_3 = 0$, $I^3e_4 = e_1, \ldots, I^3e_p = e_{p-3}$.

Мы видим, таким образом, что, для того чтобы вычислить многочлен от одной клетки нормальной формы порядка p, достаточно знать значение этого многочлена и его производных до порядка $p-1$ в точке λ_1, где λ_1 — собственное значение, отвечающее клетке. Отсюда следует, что если матрица $\boldsymbol{A}$ имеет нормальную форму (4) с клетками порядков $p, q, \ldots, s$, то для вычисления матрицы $P(\boldsymbol{A}_1)$ достаточно знать значения $P(t)$ в точках $t = \lambda_1, \lambda_2, \ldots, \lambda_k$ с производными до порядков $p-1, q-1, \ldots, s-1$ соответственно.

Мы докажем следующую теорему.

Теорема. *Пусть в комплексном n-мерном пространстве задано линейное преобразование A. Тогда можно найти базис, в котором матрица линейного преобразования имеет нормальную форму. Другими словами, можно найти базис, в котором линейное преобразование имеет вид* (2).

Два независимых доказательства сформулированной теоремы будут даны в § 19 и 20. Кроме того, важная теория инвариантных множителей и λ-матриц дает нам третье независимое доказательство этого результата.

§ 19. Приведение произвольного преобразования к нормальной форме

Мы уже упоминали в § 18, что в случае, когда у преобразования A не хватает линейно независимых собственных векторов (т. е. когда их число меньше размерности пространства), базис приходится дополнять за счет так называемых присоединенных векторов (их точное определение будет дано чуть позже). В этом параграфе дается способ построения базиса, в котором матрица преобразования A имеет жорданову нормальную форму. Этот базис мы непосредственно наберем из собственных и присоединенных векторов, и такой способ выбора является, в некотором смысле, наиболее естественным *).

Перед этим параграфом мы рекомендуем читателю перечитать п. 4 § 9 и разобрать приведенные там примеры.

1. Собственные и присоединенные векторы линейного преобразования. Пусть λ_0 — некоторое собственное значение преобразования A. Мы уже имели раньше такое определение.

Определение 1. *Вектор $x \neq 0$ называется собственным вектором преобразования A, отвечающим собственному значению λ_0, если*

$$Ax = \lambda_0 x, \quad \textit{т. е.} \quad (A - \lambda_0 E)x = 0. \tag{1}$$

Рассмотрим совокупность всех векторов, удовлетворяющих условию (1) при фиксированном λ_0. Ясно, что совокупность этих векторов является подпространством пространства R.

Мы обозначим его $N_{\lambda_0}^{(1)}$. Легко видеть, что $N_{\lambda_0}^{(1)}$ инвариантно относительно преобразования A (проверьте!).

Заметим, что подпространство $N_{\lambda_0}^{(1)}$ состоит из всех собственных векторов преобразования A, отвечающих собственному значению λ_0, к которым добавлен еще *нулевой вектор.*

Определение 2. *Вектор x называется присоединенным вектором 1-го порядка преобразования A, отвечающим собственному значению λ_0, если вектор*

$$y = (A - \lambda_0 E)x$$

является собственным вектором преобразования A.

Пусть λ_0 — собственное значение преобразования A.

*) См. также И. В. Проскуряков, Сборник задач по линейной алгебре, где имеется аналогичное доказательство.

Рассмотрим подпространство, состоящее из всех векторов x, для которых выполнено условие

$$(A - \lambda_0 E)^2 x = 0, \tag{2}$$

т. е. ядро преобразования $(A - \lambda_0 E)^2$. Обозначим это подпространство $N_{\lambda_0}^{(2)}$; $N_{\lambda_0}^{(2)}$ является инвариантным подпространством пространства R. В самом деле, пусть, $x \in N_{\lambda_0}^{(2)}$, т. е. $(A - \lambda_0 E)^2 x = 0$. Нам надо доказать, что и вектор $Ax \in N_{\lambda_0}^{(2)}$, т. е. что $(A-\lambda_0 E)^2 Ax = 0$. Но преобразование A перестановочно с $(A-\lambda_0 E)^2$, т. е.

$$(A - \lambda_0 E)^2 Ax = A(A - \lambda_0 E)^2 x = 0.$$

Рассмотрим несколько более подробно структуру пространства $N_{\lambda_0}^{(2)}$. В нем есть векторы двух типов.

Если $x \in N_{\lambda_0}^{(1)}$, т. е. $(A - \lambda_0 E)x = 0$, то подавно и $(A - \lambda_0 E)^2 x = 0$, т. е. $x \in N_{\lambda_0}^{(2)}$. Таким образом, $N_{\lambda_0}^{(1)}$ целиком содержится в $N_{\lambda_0}^{(2)}$. Если $x \in N_{\lambda_0}^{(2)}$, но $x \notin N_{\lambda_0}^{(1)}$, т. е.

$$(A - \lambda_0 E)x \neq 0,$$
$$(A - \lambda_0 E)^2 x = 0,$$

то x — присоединенный вектор 1-го порядка. Действительно, в этом случае $y = (A - \lambda_0 E)x$ есть собственный вектор.

Таким образом, подпространство $N_{\lambda_0}^{(2)}$ получается, если к подпространству $N_{\lambda_0}^{(1)}$ добавить присоединенные векторы 1-го порядка.

Аналогично вводим подпространство $N_{\lambda_0}^{(k)}$, состоящее из всех векторов x, для которых

$$(A - \lambda_0 E)^k x = 0. \tag{3}$$

Это подпространство инвариантно относительно преобразования A. Ясно, что подпространство $N_{\lambda_0}^{(k)}$ содержит предыдущее подпространство $N_{\lambda_0}^{(k-1)}$.

Определение 3. *Вектор x называется присоединенным вектором k-го порядка, если вектор*

$$y = (A - \lambda_0 E)x$$

есть присоединенный вектор порядка $k-1$.

По индукции можно показать, что если x — присоединенный вектор k-го порядка, то

$$(A - \lambda_0 E)^k x \neq 0,$$
$$(A - \lambda_0 E)^{k+1} x = 0.$$

Другими словами, присоединенным вектором k-го порядка называется вектор, принадлежащий $N_{\lambda_0}^{(k+1)}$ и не принадлежащий $N_{\lambda_0}^{(k)}$.

Пример. Пусть R — пространство многочленов степени $\leqslant n-1$ и преобразование A — дифференцирование:

$$AP(t) = \frac{d}{dt}P(t).$$

Легко видеть, что $\lambda = 0$ есть собственное значение. Соответствующий ему собственный вектор $P(t) = \text{const}$. Найдем для этого преобразования пространства $N_0^{(k)}$. По определению $N_0^{(k)}$ состоит из всех многочленов $P(t)$, для которых $A^k P(t) = 0$, т. е.

$$\frac{d^k}{dt^k}P(t) = 0.$$

Это будут все многочлены, степень которых не превышает $k-1$. Присоединенными векторами k-го порядка будут многочлены, степень которых в точности равна $k-1$.

В этом примере размерность каждого из подпространств равна $N_0^{(k)}$ и она растет от 1 до n вместе с ростом k. Подпространство $N_0^{(n)}$ уже совпадает со всем пространством R, и если мы захотим определить $N_0^{(n+1)}, N_0^{(n+2)}$ и т. д., то все эти подпространства будут совпадать с $N_0^{(n)}$.

Легко видеть также, что в этом примере $AN_0^{(k+1)} = N_0^{(k)}$. Это следует из того, что каждый многочлен степени k есть производная от многочлена степени $k+1$.

Упражнение. Показать, что для любого линейного преобразования A имеет место включение

$$(A-\lambda_0 E)N_{\lambda_0}^{(k+1)} \subset N_{\lambda_0}^{(k)}.$$

Пусть A — линейное преобразование, а λ_0 — его собственное значение. Покажем, что подпространства $N_{\lambda_0}^{(1)}, N_{\lambda_0}^{(2)}, \ldots$ сначала строго возрастают с ростом индекса, а затем, начиная с некоторого номера $p \leqslant n$, этот рост прекращается, т. е.

$$N_{\lambda_0}^{(p)} = N_{\lambda_0}^{(p+1)} = \ldots$$

(см. приведенный в этом пункте пример).

Мы уже показали, что каждое подпространство $N_{\lambda_0}^{(k)}$ содержит $N_{\lambda_0}^{(k-1)}$, т. е. что с увеличением номера подпространства $N_{\lambda_0}^{(k)}$, а значит, и их размерности, могут только увеличиваться.

Так как наше пространство конечномерно, то для какого-то $p \leqslant n$ мы впервые получим, что $N_{\lambda_0}^{(p)} = N_{\lambda_0}^{(p+1)}$ (см. упражнение на стр. 25).

Докажем, что в этом случае $N_{\lambda_0}^{(p+1)} = N_{\lambda_0}^{(p+2)} = \ldots$, т. е. что дальнейшего возрастания подпространств происходить не будет.

Действительно, предположим противное, а именно, что $N_{\lambda_0}^{(p+1)} = N_{\lambda_0}^{(p)}$, но для некоторого $i > 0$ подпространство $N_{\lambda_0}^{(p+i+1)}$ строго больше, чем $N_{\lambda_0}^{(p+i)}$. Тогда существует вектор x такой, что

$$x \in N_{\lambda_0}^{(p+i+1)},\ x \notin N_{\lambda_0}^{(p+i)}.$$

Это значит, что

$$(A - \lambda_0 E)^{p+i+1} x = 0, \quad \text{но} \quad (A - \lambda_0 E)^{p+1} x \neq 0. \tag{4}$$

Обозначим через y вектор $y = (A - \lambda_0 E)^i x$. Тогда первое из равенств (4) означает, что $y \in N_{\lambda_0}^{(p+1)}$, а второе, что $y \notin N_{\lambda_0}^{(p)}$, что невозможно, так как подпространства $N_{\lambda_0}^{(p+1)}$ и $N_{\lambda_0}^{(p)}$ по предположению совпадают.

Итак, пусть λ_0 — некоторое собственное значение преобразования A. Основным результатом этого пункта является построение инвариантного подпространства $N_{\lambda_0}^{(p)}$, состоящего из всех собственных и присоединенных векторов, отвечающих этому собственному значению.

Кроме того, в п. 3 нам понадобится более детальная структура $N_{\lambda_0}^{(p)}$. А именно, обозначая через $N_{\lambda_0}^{(k)}$ подпространство, состоящее из присоединенных векторов порядка $\leqslant k-1$, мы получили возрастающую цепочку инвариантных подпространств

$$0 \subset N_{\lambda_0}^{(1)} \subset N_{\lambda_0}^{(2)} \subset \ldots \subset N_{\lambda_0}^{(p)}. \tag{5}$$

Все члены этой цепочки различны. Подпространство $N_{\lambda_0}^{(k)}$ состоит при этом из всех векторов x, для которых

$$(A - \lambda_0 E)^k x = 0,$$

т. е. это есть ядро преобразования $(A - \lambda_0 E)^k$.

Преобразование $A - \lambda_0 E$ переводит каждое из подпространств цепочки (5) в предшествующее.

2. Выделение подпространства, в котором преобразование A имеет только одно собственное значение. Пусть λ_1 — некоторое собственное значение преобразования A. В этом пункте мы покажем, что пространство R можно разложить в прямую сумму двух инвариантных подпространств, в первом из которых преобразование A имеет лишь одно собственное значение λ_1, а во втором у преобразования A уже нет собственного значения λ_1.

Не ограничивая общности, можно считать, что $\lambda_1 = 0$.

Действительно, пусть $\lambda_1 \neq 0$. Рассмотрим преобразование $B = A - \lambda_1 E$; оно уже имеет собственное значение, равное нулю *). Очевидно также, что инвариантные подпространства преобразований A и B совпадают.

Итак, впредь мы будем считать, что преобразование имеет собственное значение $\lambda = 0$. Докажем наше утверждение сначала для частного случая, когда *в пространстве нет присоединенных векторов, отвечающих этому собственному значению, а есть только собственные векторы* **).

Нам нужно построить два инвариантных подпространства, прямая сумма которых равна R. В качестве первого из них, в котором $\lambda = 0$ есть единственное собственное значение, можно взять совокупность N_0 всех

---

*) В самом деле, если λ_1 — собственное значение преобразования A, т. е. $Af = \lambda_1 f$ $(f \neq 0)$, то $Bf = (A - \lambda_1 E)f = 0$, т. е. f — собственный вектор B, отвечающий собственному значению $\lambda = 0$.

**) Хотя потом наше утверждение будет независимо доказано для общего случая, рассмотрение этого частного случая полезно, так как, во-первых, на нем более выпукло видна основная идея доказательства и, во-вторых, становится очевидной необходимость введения подпространств, отличных от введенных здесь N_0 и M.

собственных векторов, отвечающих собственному значению $\lambda = 0$ или, другими словами, ядро преобразования A.

В качестве второго подпространства возьмем образ M пространства R при преобразовании A, т. е. совокупность векторов $y = Ax$, где x пробегает все пространство R. Легко видеть, что каждое из этих подпространств инвариантно (это доказано в п. 4 § 9).

Докажем, что они дают разложение пространства в прямую сумму. Так как сумма размерностей ядра и образа для любого преобразования A равна n (см. п. 4 § 9), то достаточно доказать, что пересечение этих подпространств равно нулю.

Предположим, что это не так, т. е. пусть существует вектор $y \neq 0$ такой, что $y \in M$ и $y \in N_0$. Так как $y \in M$, то он имеет вид

$$y = Ax, \tag{6}$$

где x — некоторый вектор из R. Так как $y \in N_0$, то

$$Ay = 0, \quad \text{где} \quad y \neq 0. \tag{7}$$

Равенство (7) означает, что y есть собственный вектор преобразования A, отвечающий собственному значению $\lambda = 0$, а равенство (6) при этом означает, что x есть присоединенный вектор первого порядка, отвечающий тому же собственному значению. Мы же предположили, что у преобразования A нет присоединенных векторов, отвечающих собственному значению $\lambda = 0$.

Таким образом, доказано, что подпространства M и N_0 не имеют общих векторов, кроме нулевого.

Вспоминая, что сумма размерностей образа и ядра равна n, мы получаем отсюда, что пространство R разложимо в прямую сумму инвариантных подпространств M и N_0:

$$R = M \oplus N_0.$$

З а м е ч а н и е. Из приведенного выше доказательства видно, что образ и ядро имеют пересечение, отличное от нуля в том и только том случае, когда преобразование A имеет присоединенные векторы, отвечающие собственному значению $\lambda = 0$.

Разобранный частный случай дает нам идею того, как проводить доказательство в общем случае, когда A имеет также и присоединенные векторы, отвечающие собственному значению $\lambda = 0$. Подпространство N_0 при этом оказывается слишком узким, и его естественно расширить за счет добавления всех присоединенных векторов, отвечающих собственному значению $\lambda = 0$. Второе же подпространство M оказывается при этом слишком большим *).

Итак, рассмотрим введенное в п. 1 инвариантное подпространство $N_0^{(p)}$, состоящее из всех собственных и присоединенных векторов преобразования A, отвечающих собственному значению $\lambda = 0$. Как мы помним, оно является ядром преобразования A^p, т. е. состоит из всех векторов x, для которых

$$A^p x = 0.$$

В качестве второго слагаемого прямой суммы мы возьмем подпространство $M^{(p)}$ — образ пространства R при том же преобразовании A^p.

Легко видеть, что $M^{(p)}$ также инвариантно относительно преобразования A. Действительно, если $y \in M^{(p)}$, т. е. $y = A^p x$, то

$$Ay = A^{p+1}x = A^p(Ax),$$

т. е. Ay также принадлежит $M^{(p)}$.

---

*) Что M «слишком велико», видно при этом не только из соображений размерности, но также и из того, что M пересекается даже с самим N_0, а не только с его расширением.

Теорема 1. *Пространство R можно разложить в прямую сумму инвариантных подпространств $N_0^{(p)}$ и $M^{(p)}$. При этом подпространство $N_0^{(p)}$ состоит только из собственных и присоединенных векторов, отвечающих собственному значению $\lambda = 0$, а в подпространстве $M^{(p)}$ преобразование A обратимо (т. е. $\lambda = 0$ не является собственным значением преобразования A в подпространстве $M^{(p)}$).*

Для доказательства первого утверждения нам, как и в рассмотренном выше частном случае, достаточно показать, что пересечение подпространств $N_0^{(p)}$ и $M^{(p)}$ равно нулю. Допустим противное, т. е. пусть существует вектор $y \neq 0$ такой, что $y \in M^{(p)}$ и $y \in N_0^{(p)}$. Так как $y \in M^{(p)}$, то

$$y = A^p x. \tag{8}$$

Далее, так как $y \in N_0^{(p)}$, то

$$A^p y = 0. \tag{9}$$

Но из равенств (8) и (9) следует, что существует такой вектор x, для которого

$$A^p x \neq 0$$

и в то же время

$$A^{2p} x = A^p y = 0.$$

Это значит, что x есть присоединенный вектор преобразования A с собственным значением $\lambda = 0$, не принадлежащий подпространству $N_0^{(p)}$, что невозможно, так как $N_0^{(p)}$ состоит из всех таких векторов.

Таким образом, мы доказали, что пересечение $N_0^{(p)}$ и $M^{(p)}$ равно нулю. Так как сумма размерностей этих подпространств равна n (это ядро и образ преобразования A^p), то отсюда следует, что пространство R

раскладывается в прямую сумму этих подпространств:

$$R = M^{(p)} \oplus N_0^{(p)}. \tag{10}$$

Докажем теперь второе утверждение теоремы, т. е. что в подпространстве $M^{(p)}$ преобразование A не имеет нулевого собственного значения. Действительно, если бы это было не так, то в $M^{(p)}$ существовал бы вектор $x \neq 0$ такой, что

$$A^p x = 0.$$

Но это равенство означает, что $x \in N_0^{(p)}$, т. е. является общим вектором $M^{(p)}$ и $N_0^{(p)}$, а мы доказали, что таким вектором может быть только нуль.

Теорема доказана полностью.

Теперь мы можем освободиться от предположения, что выделенное подпространство отвечает нулевому собственному значению, и считать установленным следующий факт.

Если λ_1 — некоторое собственное значение преобразования A, то пространство R можно разложить в прямую сумму инвариантных подпространств R_1 и $\tilde{R}$, в первом из которых преобразование A имеет только собственное значение λ_1, а во втором все собственные значения A отличны от λ_1.

Применяя полученный результат к преобразованию A в пространстве $\tilde{R}$ и к некоторому собственному значению λ_2 этого преобразования, мы «отщепим» инвариантное подпространство, отвечающее собственному значению λ_2. Продолжая этот процесс, пока не будут исчерпаны все собственные значения преобразования A, мы получим доказательство следующей теоремы:

Теорема 2. *Пусть преобразование A пространства R имеет k различных собственных значений $\lambda_1, \ldots, \lambda_k$. Тогда R можно разложить в прямую сумму*

k инвариантных подпространств $N_{\lambda_1}^{(p_1)}, \ldots, N_{\lambda_k}^{(p_k)}$:

$$R = N_{\lambda_1}^{(p_1)} \oplus \ldots \oplus N_{\lambda_k}^{(p_k)}. \tag{11}$$

Каждое из подпространств $N_{\lambda_i}^{(p_i)}$ состоит только из собственных и присоединенных векторов, отвечающих собственному значению λ_i.

Другими словами, для каждого i существует такое число p_i, что для всех $x \in N_{\lambda_i}^{(p_i)}$

$$(A - \lambda_i E)^{p_i} x = 0.$$

У нас осталась еще только одна, впрочем, не менее важная задача — выбрать в каждом из этих подпространств базис, в котором матрица преобразования имеет жорданову нормальную форму. Это будем сделано в следующем пункте.

3. Приведение к нормальной форме матрицы с одним собственным значением. В случае, если пространство состоит только из собственных векторов, базис в пространстве можно выбирать произвольно и матрица преобразования в этом базисе имеет диагональный вид.

В общем случае неосторожный выбор базиса может запутать картину.

Чтобы выбрать базис, в котором матрица преобразования имеет наиболее простой вид, мы будем тянуть цепочки собственных и присоединенных векторов, выбрав некоторый базис в подпространстве $N^{(p)}$ и последовательно применяя к векторам этого базиса преобразование A.

Введем предварительно некоторые понятия, удобные для дальнейшего.

Определение 4. *Векторы из пространства R называются относительно линейно независимыми над подпространством R_1, если никакая их линейная комбинация, отличная от нуля, не принадлежит R_1.*

Заметим, что всякие линейно зависимые векторы из R относительно линейно зависимы над любым подпространством.

О п р е д е л е н и е 5. *Базисом пространства R относительно подпространства R_1 называется такая система $e_1, \ldots, e_k$ линейно независимых векторов из R, которая после пополнения каким-нибудь базисом из R_1 образует базис во всем пространстве.*

Такой базис легко построить. Для этого достаточно выбрать какой-нибудь базис в R_1, дополнить его до базиса во всем пространстве и затем отбросить векторы исходного базиса из R_1. Число векторов в таком *относительном базисе* равно разности размерностей пространства и подпространства.

Всякую систему относительно линейно независимых векторов над R_1 можно дополнить до относительного базиса. Для этого нужно к выбранным векторам добавить какой-нибудь базис подпространства R_1. Получится некоторая система векторов из R, которые, как легко проверить, линейно независимы. Чтобы получить относительный базис, нужно дополнить эту систему до базиса во всем пространстве R, а затем отбросить базис подпространства.

Итак, пусть преобразование A в пространстве R имеет только одно собственное значение. Не ограничивая общности, можно предположить, что оно равно нулю.

Рассмотрим снова цепочку (5) подпространств, полученных в п. 1:

$$0 \subset N_0^{(1)} \subset \ldots \subset N_0^{(p)} = N_0^{(p+1)} = \ldots,$$

где подпространство $N_0^{(k)}$ есть ядро преобразования A^k. Так как преобразование A в пространстве R не имеет отличных от нуля собственных значений, то, очевидно, $N^{(p)}$ совпадает при этом со всем пространством R.

Выберем в максимальном из этих подпространств $N_0^{(p)}$ базис относительно содержащегося в нем подпространства $N_0^{(p-1)}$. Пусть векторы этого базиса будут

$$e_1, \ldots, e_q.$$

Очевидно, что это будут присоединенные векторы $(p-1)$-го порядка. Мы уже видели (см. упражнение на стр. 211), что $AN_0^{(p)} \subset N_0^{(p-1)}$. Поэтому векторы

$$Ae_1, \ldots, Ae_q$$

лежат в $N_0^{(p-1)}$. Покажем, что эти векторы линейно независимы в $N_0^{(p-1)}$ относительно лежащего в нем подпространства $N_0^{(p-2)}$. Действительно, пусть не все $\alpha_i = 0$ и

$$\alpha_1 Ae_1 + \ldots + \alpha_q Ae_q = A(\alpha_1 e_1 + \ldots + \alpha_q e_q) \in N_0^{(p-2)}.$$

Тогда вектор $x = \alpha_1 e_1 + \ldots + \alpha_q e_q \in N_0^{(p-1)}$, а это противоречит предположению, что векторы $e_1, \ldots, e_q$ линейно независимы над $N_0^{(p-1)}$.

Дополним векторы $Ae_1, \ldots, Ae_q$ до базиса в $N_0^{(p-1)}$ относительно $N_0^{(p-2)}$. Мы получим тогда $q+s$ векторов

$$Ae_1, \ldots, Ae_q, f_1, \ldots, f_s,$$

которые представляют собой максимальное число линейно независимых присоединенных векторов порядка $p-2$.

Снова применим к этим векторам преобразование A и полученную систему векторов из $N_0^{(p-2)}$ дополним, как и выше, до базиса в $N_0^{(p-2)}$ относительно $N_0^{(p-3)}$.

Продолжая этот процесс, мы дойдем до подпространства $N_0^{(1)}$ и выберем базис в этом пространстве, состоящий из максимального числа линейно независимых собственных векторов.

Расположим полученные векторы в следующую таблицу

$$\begin{array}{cccccccccc}
e_1 & \ldots & e_q & & & & & & & \\
Ae_1 & \ldots & Ae_q & f_1 & \ldots & f_s & & & & \\
A^2e_1 & \ldots & A^2e_q & Af_1 & \ldots & Af_s & & & & \\
\multicolumn{10}{c}{\dotfill} \\
\multicolumn{10}{c}{\dotfill} \\
A^{p-1}e_1 & \ldots & A^{p-1}e_q & A^{p-2}f_1 & \ldots & A^{p-2}f_s & \ldots & h_1 & \ldots & h_r
\end{array} \tag{12}$$

Векторы нижней строчки образуют базис в подпространстве $N_0^{(1)}$. Векторы двух нижних строчек образуют базис в $N_0^{(2)}$, так как это есть базис $N_0^{(2)}$ относительно $N_0^{(1)}$в соединении с базисом $N_0^{(1)}$. Векторы трех нижних строчек образуют базис в $N_0^{(3)}$ и т. д. Наконец все векторы таблицы образуют базис в $N_0^{(p)}$, т. е. во всем пространстве R.

Покажем, что в этом базисе матрица преобразования A имеет жорданову нормальную форму. Действительно, рассмотрим произвольный столбец таблицы (12), например, для определенности первый.

Обозначим для удобства $A^{p-1}e_1$ через $\overline{e}_1$, $A^{p-2}e_1$ — через $\overline{e}_2$ и т. д. и рассмотрим действие преобразования A на каждый из этих векторов. Так как $\overline{e}_1$ — собственный вектор, отвечающий нулевому собственному значению, то

$$A\overline{e}_1 = 0.$$

Дальше, по определению,

$$A\overline{e}_2 = AA^{p-2}e_1 = A^{p-1}e_1 = \overline{e}_1$$

и аналогично

$$\begin{aligned} &A\overline{e}_3 = \overline{e}_2, \\ &\ldots\ldots\ldots \\ &A\overline{e}_p = \overline{e}_{p-1}. \end{aligned}$$

Таким образом, преобразование A переводит векторы первого столбца снова в себя, т. е. подпространство R_1, натянутое на эти векторы, инвариантно относительно A. Матрица преобразования A в подпространстве R_1 в базисе $\overline{e}_1, \ldots, \overline{e}_p$ имеет вид

$$\begin{pmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & 0 & \ldots & 0 & 0 \end{pmatrix}, \tag{13}$$

т. е. это есть жорданова клетка, отвечающая собственному значению $\lambda = 0$. Аналогичное инвариантное подпространство отвечает каждому из столбцов таблицы (12), и размерность каждого такого подпространства равна числу векторов в соответствующем столбце. Так как матрица преобразования A в базисе, состоящем из векторов какого-либо столбца таблицы (12), имеет вид (13), то матрица преобразования во всем пространстве R в базисе, состоящем из всех векторов таблицы (12), состоит из жордановых клеток, число которых равно числу столбцов в этой таблице, а размер каждой клетки равен числу векторов соответствующего столбца.

Если вместо преобразования A рассмотреть преобразование $A + \lambda_1 E$, то, так как матрица преобразования $\lambda_1 E$ диагональна, мы получим тот же результат для преобразования пространства R, имеющего только одно собственное значение, равное произвольному числу λ_1. Соответствующие жордановы клетки матрицы преобразования $A + \lambda_1 E$ будут иметь вид:

$$\begin{pmatrix} \lambda_1 & 1 & 0 & \dots & 0 \\ 0 & \lambda_1 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_1 \end{pmatrix}. \tag{14}$$

Вспоминая теперь, что для произвольного преобразования A мы можем разложить пространство R в сумму инвариантных подпространств, в каждом из которых преобразование A имеет только одно собственное значение (см. формулу (11)), мы получаем отсюда полное доказательство теоремы § 18.

§ 20. Другое доказательство теоремы о приведении к нормальной форме

Это доказательство мы будем вести по индукции, именно предположим, что для линейного преобразования в пространстве n измерений такой базис существует, и докажем, что мы можем найти нужный базис в пространстве $n + 1$ измерений. Для доказательства теоремы нам понадобится следующая

Лемма. *У всякого линейного преобразования A в n-мерном комплексном пространстве R существует хотя бы одно $(n-1)$-мерное инвариантное подпространство R'.*

Доказательство. Рассмотрим преобразование A^*; у него, как и у всякого преобразования, есть собственный вектор e

$$A^* e = \lambda e.$$

Покажем, что $(n-1)$-мерное подпространство R', состоящее из векторов x, ортогональных *) e, т. е. для которых $(x, e) = 0$, инвариантно относительно A. Действительно, пусть $x \in R'$, т. е. $(x, e) = 0$. Тогда

$$(Ax, e) = (x, A^* e) = (x, \lambda e) = 0$$

и, значит, Ax также принадлежит R'. Инвариантность подпространства R' относительно преобразования A доказана.

Докажем теперь сформулированную выше основную теорему этого параграфа.

Пусть A — произвольное линейное преобразование в $(n+1)$-мерном пространстве R. Согласно лемме, в R существует n-мерное подпространство R', инвариантное относительно A. Так как в n-мерном пространстве мы предполагаем теорему доказанной, то в R' существует базис, в котором линейное преобразование имеет нормальную форму. Обозначим этот базис в R' через

$$e_1, e_2, \ldots, e_p; \; f_1, f_2, \ldots, f_q; \; \ldots; \; h_1, h_2, \ldots, h_s,$$

где $p + q + \ldots + s = n$. В этом базисе линейное преобра-

---

*) Мы здесь используем наличие в R скалярного произведения, т. е. считаем пространство R евклидовым. Незначительным изменением доказательства можно освободиться от этого предположения и на протяжении всей главы считать R аффинным пространством.

зование в n-мерном подпространстве R' имеет вид

$$
\begin{array}{lll}
\begin{aligned}
Ae_1 &= \qquad \lambda_1 e_1,\\
Ae_2 &= \quad e_1+\lambda_1 e_2,\\
&\ldots\ldots\ldots\\
Ae_p &= e_{p-1}+\lambda_1 e_p,
\end{aligned} & & \\
& \begin{aligned}
Af_1 &= \qquad \lambda_2 f_1,\\
Af_2 &= \quad f_1+\lambda_2 f_2,\\
&\ldots\ldots\ldots\\
Af_q &= f_{q-1}+\lambda_2 f_q,
\end{aligned} & \\
& & \ddots \\
& & \begin{aligned}
Ah_1 &= \qquad \lambda_k h_1,\\
Ah_2 &= \quad h_1+\lambda_k h_2,\\
&\ldots\ldots\ldots\\
Ah_s &= h_{s-1}+\lambda_k h_s.
\end{aligned}
\end{array}
$$

Дополним этот базис каким-нибудь вектором e, который вместе с $e_1, e_2, \ldots, e_p$; $f_1, f_2, \ldots, f_q$; $\ldots$; $h_1, h_2, \ldots \ldots, h_s$ составляет базис в R.

Применим к e преобразование A и разложим полученный вектор Ae по векторам базиса:

$$
Ae = \alpha_1 e_1 + \ldots + \alpha_p e_p + \beta_1 f_1 + \ldots + \beta_q f_q + \ldots \\
\ldots + \delta_1 h_1 + \ldots + \delta_s h_s + \tau e \ ^{*}).
$$

Мы можем считать, что $\tau = 0$. В самом деле, если в некотором базисе A имеет нормальную форму, то $A - \tau E$

---

*) Линейное преобразование A имеет в $(n+1)$-мерном пространстве R собственные значения $\lambda_1, \lambda_2, \ldots, \lambda_k$ и τ. Действительно, напишем матрицу преобразования A в базисе $e_1, e_2, \ldots \ldots, e_p; f_1, f_2, \ldots, f_q; \ldots; h_1, h_2, \ldots, h_s; e$. Она будет треугольной матрицей, в которой по диагонали стоят числа $\lambda_1, \lambda_2, \ldots, \lambda_k$ и τ.

Так как собственными значениями треугольной матрицы являются числа, стоящие по диагонали (см., например, § 10, п. 4), то $\lambda_1, \lambda_2, \ldots, \lambda_k$ и τ будут собственными значениями A в $(n+1)$-мерном пространстве. Таким образом, при переходе от инвариантного n-мерного к $(n+1)$-мерному пространству добавилось одно собственное значение: τ.

также имеет в этом базисе нормальную форму. Поэтому, если $\tau \neq 0$, то можно вместо преобразования A рассмотреть преобразование $A-\tau E$, причем A и $A-\tau E$, согласно сделанному замечанию, имеют нормальную форму в одном и том же базисе.

Мы полагаем, следовательно, что

$$Ae = \alpha_1 e_1 + \ldots + \alpha_p e_p + \beta_1 f_1 + \ldots \\ \ldots + \beta_q f_q + \ldots + \delta_1 h_1 + \ldots + \delta_s h_s. \quad (1)$$

Теперь постараемся заменить вектор e вектором e' так, чтобы после этой замены вектор Ae' стал возможно проще. Будем искать вектор e' в виде

$$e' = e - \varkappa_1 e_1 - \ldots - \varkappa_p e_p - \mu_1 f_1 - \ldots \\ \ldots - \mu_q f_q - \ldots - \omega_1 h_1 - \ldots - \omega_s h_s. \quad (2)$$

Мы имеем

$$Ae' = Ae - A(\varkappa_1 e_1 + \ldots + \varkappa_p e_p) - \\ - A(\mu_1 f_1 + \ldots + \mu_q f_q) - \ldots - A(\omega_1 h_1 + \ldots + \omega_s h_s),$$

или, пользуясь формулой (1),

$$Ae' = \alpha_1 e_1 + \ldots + \alpha_p e_p + \beta_1 f_1 + \ldots + \beta_q f_q + \ldots \\ \ldots + \delta_1 h_1 + \ldots + \delta_s h_s - A(\varkappa_1 e_1 + \ldots + \varkappa_p e_p) - \\ - A(\mu_1 f_1 + \ldots + \mu_q f_q) - \ldots - A(\omega_1 h_1 + \ldots + \omega_s h_s). \quad (3)$$

Коэффициенты $\varkappa_1, \ldots, \varkappa_p$; $\mu_1, \ldots, \mu_q$; $\ldots$; $\omega_1, \ldots, \omega_s$ мы можем выбрать произвольно. Подберем их так, чтобы в правой части формулы (3) осталось как можно меньше слагаемых.

Мы знаем, что каждой группе базисных векторов n-мерного подпространства R', в котором преобразование A имеет нормальную форму, отвечает свое собственное значение λ_1, λ_2 и т. д. Мы рассмотрим отдельно два случая, а именно разберем сначала случай, когда ни одно из этих собственных значений не равно нулю, а затем случай, когда это не так.

Разберем первый случай, когда $\lambda_1 \neq 0$, $\lambda_2 \neq 0$, $\ldots$, $\lambda_k \neq 0$. В этом случае мы покажем, что вектор e' можно выбрать так, чтобы $Ae' = 0$, т. е. подобрать $\varkappa_1, \ldots, \omega_s$ так, чтобы все слагаемые в правой части (3) сократились. Так как векторы каждой группы переходят при преобразовании в комбинацию векторов той же группы, то векторы различных групп можно уничтожать независимо друг от друга. Покажем, как подобрать коэффициенты $\varkappa_1, \varkappa_2, \ldots, \varkappa_p$, чтобы в правой части (3) сократились векторы $e_1, e_2, \ldots, e_p$. Члены, содержащие эти векторы, имеют вид

$$\begin{aligned}
&\alpha_1 e_1 + \ldots + \alpha_p e_p - A(\varkappa_1 e_1 + \ldots + \varkappa_p e_p) = \\
&\qquad = \alpha_1 e_1 + \ldots + \alpha_p e_p - \varkappa_1 \lambda_1 e_1 - \\
&\qquad\qquad - \varkappa_2(e_1 + \lambda_1 e_2) - \ldots - \varkappa_p(e_{p-1} + \lambda_1 e_p) = \\
&\qquad = (\alpha_1 - \varkappa_1\lambda_1 - \varkappa_2)e_1 + (\alpha_2 - \varkappa_2\lambda_1 - \varkappa_3)e_2 + \ldots \\
&\qquad\qquad \ldots + (\alpha_{p-1} - \varkappa_{p-1}\lambda_1 - \varkappa_p)e_{p-1} + (\alpha_p - \varkappa_p\lambda_1)e_p.
\end{aligned}$$

Приравнивая нулю коэффициент при e_p, определяем $\varkappa_p$, что возможно, так как $\lambda_1 \neq 0$, затем, приравнивая нулю коэффициент при e_{p-1}, определяем $\varkappa_{p-1}$ и так далее до $\varkappa_1$. Таким образом, мы уничтожили в (3) члены с $e_1, e_2, \ldots, e_p$. Аналогично вычисляем другие группы коэффициентов.

Мы получили, таким образом, вектор e', для которого

$$Ae' = 0.$$

Добавляя этот вектор к имеющемуся базису, получаем базис $e'; e_1, e_2, \ldots, e_p; f_1, f_2, \ldots, f_q; \ldots; h_1, h_2, \ldots, h_s$ в $(n+1)$-мерном пространстве, в котором преобразование имеет канонический вид. Вектор e' образует при этом отдельную группу с собственным значением, равным нулю (следовательно, с собственным значением τ, если бы мы не рассматривали вместо A преобразование $A - \tau E$).

Рассмотрим теперь второй случай, а именно пусть некоторым группам векторов базиса в n-мерном пространстве R' соответствуют собственные значения преобразования A, равные нулю. Тогда в правой части формулы (3) у нас будут слагаемые двух сортов — соответствующие группам с отличными от нуля собственными значениями и группам, для которых собственное значение равно нулю. С группами, у которых собственные значения отличны от нуля, мы можем поступить так же, как и в первом случае, т. е. подбором коэффициентов уничтожить векторы в правой части (3). Допустим, что после этой операции у нас останутся, например, три группы слагаемых $e_1, e_2, \ldots, e_p; f_1, f_2, \ldots, f_q$; $g_1, g_2, \ldots, g_r$ с собственным значением, равным нулю, т. е. что $\lambda_1 = \lambda_2 = \lambda_3 = 0$. Тогда

$$\begin{aligned} Ae' = \alpha_1 e_1 + \ldots + \alpha_p e_p + \beta_1 f_1 + \ldots + \beta_q f_q + \ldots + \gamma_1 g_1 + \ldots \\ \ldots + \gamma_r g_r - A(\varkappa_1 e_1 + \ldots + \varkappa_p e_p) - \\ - A(\mu_1 f_1 + \ldots + \mu_q f_q) - A(\nu_1 g_1 + \ldots + \nu_r g_r). \quad (4) \end{aligned}$$

Так как $\lambda_1 = \lambda_2 = \lambda_3 = 0$, то

$$\begin{aligned} &Ae_1 = 0, \quad Ae_2 = e_1, \quad \ldots, \quad Ae_p = e_{p-1}, \\ &Af_1 = 0, \quad Af_2 = f_1, \quad \ldots, \quad Af_q = f_{q-1}, \\ &Ag_1 = 0, \quad Ag_2 = g_1, \quad \ldots, \quad Ag_r = g_{r-1}. \end{aligned}$$

Поэтому линейная комбинация векторов $e_1, e_2, \ldots, e_p$, входящая в правую часть равенства (4), будет иметь вид

$$\alpha_1 e_1 + \ldots + \alpha_p e_p - \varkappa_2 e_1 - \ldots - \varkappa_p e_{p-1}.$$

Полагая $\varkappa_2 = \alpha_1, \ldots, \varkappa_p = \alpha_{p-1}$, мы уничтожим здесь все слагаемые кроме одного, равного $\alpha_p e_p$. Проделав ту же операцию в группах $f_1, \ldots, f_q$ и $g_1, \ldots, g_r$, мы получим вектор e', для которого

$$Ae' = \alpha_p e_p + \beta_q f_q + \gamma_r g_r.$$

Случайно может оказаться, что $\alpha_p = \beta_q = \gamma_r = 0$; тогда мы приходим к вектору e', для которого

$$Ae' = 0,$$

и тогда, как и в первом случае, наше преобразование уже в базисе $e'; e_1, e_2, \ldots, e_p; f_1, f_2, \ldots, f_q; \ldots; h_1, h_2, \ldots \ldots, h_s$ имеет нормальную форму. Вектор e' в этом случае образует новую клетку с собственным значением, равным нулю.

Пусть теперь хотя бы один из коэффициентов $\alpha_p, \beta_q, \gamma_r$ отличен от нуля. В этом случае, в отличие от рассмотренных ранее, нам придется для приведения к нормальной форме также изменить некоторые из векторов базиса, уже имеющегося в R'. Расположим числа p, q, r по их величине. Пусть, например, $p > q > r$. Тогда строим новую группу, начиная с e', следующим образом. Полагаем $e'_{p+1} = e'$, $e'_p = Ae'_{p+1}$, $e'_{p-1} = Ae'_p$, $\ldots$, $e'_1 = Ae'_2$. Мы имеем, следовательно,

$$\begin{aligned}
e'_{p+1} &= e', \\
e'_p &= Ae'_{p+1} = \alpha_p e_p + \beta_q f_q + \gamma_r g_r, \\
e'_{p-1} &= Ae'_p = \alpha_p e_{p-1} + \beta_q f_{q-1} + \gamma_r g_{r-1}, \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
e'_{p-r+1} &= Ae'_{p-r+2} = \alpha_p e_{p-r+1} + \beta_q f_{q-r+1} + \gamma_r g_1, \\
e'_{p-r} &= Ae'_{p-r+1} = \alpha_p e_{p-r} + \beta_q f_{q-r}, \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
e'_1 &= Ae'_2 = \alpha_p e_1.
\end{aligned}$$

Заменим теперь векторы $e', e_1, e_2, \ldots, e_p$ базиса векторами

$$e'_1, e'_2, \ldots, e'_p, e'_{p+1},$$

а остальные оставим без изменения. Мы получим тогда нормальную форму преобразования, причем размеры первой клетки увеличились на единицу. Теорема полностью доказана.

Мы видим, что в процессе построения нормальной формы нужно было различать два случая.

1. Случай, когда добавленное собственное значение τ (мы его полагали равным нулю) не совпадает ни с одним из прежних собственных значений $\lambda_1, \ldots, \lambda_k$. В этом случае добавлялась отдельная клетка первого порядка.

2. Случай, когда добавленное собственное значение совпадало с одним из уже имевшихся. В этом случае, вообще говоря, размер одной из имевшихся клеток увеличивался на 1. Если же коэффициенты α, β, γ равны нулю, то, как и в первом случае, добавлялась новая клетка.

§ 21. Инвариантные множители

В этом параграфе мы укажем способ, дающий возможность находить жорданову нормальную форму линейного преобразования. Из результатов этого параграфа будем также вытекать до сих пор еще не доказанная единственность этой формы.

Определение. *Матрицы $\boldsymbol{A}$ и $\boldsymbol{A}_1 = \boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}$, где $\boldsymbol{C}$ — произвольная невырожденная матрица, называются подобными.*

Если матрица $\boldsymbol{A}_1$ подобна матрице $\boldsymbol{A}_2$, то и обратно, $\boldsymbol{A}_2$ подобна $\boldsymbol{A}_1$. Действительно, пусть

$$\boldsymbol{A}_1 = \boldsymbol{C}^{-1}\boldsymbol{A}_2\boldsymbol{C}.$$

Тогда отсюда

$$\boldsymbol{A}_2 = \boldsymbol{C}\boldsymbol{A}_1\boldsymbol{C}^{-1},$$

т. е. если положить $\boldsymbol{C}^{-1} = \boldsymbol{C}_1$, имеем:

$$\boldsymbol{A}_2 = \boldsymbol{C}_1^{-1}\boldsymbol{A}_1\boldsymbol{C}_1$$

и, следовательно, $\boldsymbol{A}_2$ подобна $\boldsymbol{A}_1$.

Легко также показать, что если две матрицы $\boldsymbol{A}_1$ и $\boldsymbol{A}_2$ подобны одной и той же матрице $\boldsymbol{A}$, то они подобны между собой. Действительно, пусть

$$\boldsymbol{A} = \boldsymbol{C}_1^{-1}\boldsymbol{A}_1\boldsymbol{C}_1, \quad \boldsymbol{A} = \boldsymbol{C}_2^{-1}\boldsymbol{A}_2\boldsymbol{C}_2.$$

Тогда $\boldsymbol{C}_1^{-1}\boldsymbol{A}_1\boldsymbol{C}_1 = \boldsymbol{C}_2^{-1}\boldsymbol{A}_2\boldsymbol{C}_2$, т. е.

$$\boldsymbol{A}_1 = \boldsymbol{C}_1\boldsymbol{C}_2^{-1}\boldsymbol{A}_2\boldsymbol{C}_2\boldsymbol{C}_1^{-1},$$

и если положить $\boldsymbol{C}_2\boldsymbol{C}_1^{-1} = \boldsymbol{C}$, то получим:

$$\boldsymbol{A}_1 = \boldsymbol{C}^{-1}\boldsymbol{A}_2\boldsymbol{C},$$

т. е. $\boldsymbol{A}_1$ и $\boldsymbol{A}_2$ подобны.

Пусть $\boldsymbol{A}$ — матрица преобразования A в некотором базисе. При переходе к другому базису матрица $\boldsymbol{A}$ заменяется подобной ей матрицей $\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}$, где $\boldsymbol{C}$ — матрица перехода от первого базиса ко второму (§ 9). Таким образом, подобные матрицы — это матрицы одного и того же линейного преобразования в различных базисах.

Наша задача — по матрице преобразования построить *инварианты* самого преобразования, т. е. выражения, зависящие лишь от самого преобразования A. Другими словами, нам нужно построить функции от элементов матрицы, совпадающие для подобных матриц.

Один такой инвариант установлен уже в § 10. Именно, там было доказано, что характеристический многочлен матрицы $\boldsymbol{A}$, т. е. определитель матрицы $\boldsymbol{A} - \lambda\boldsymbol{E}$:

$$D_n(\lambda) = |\boldsymbol{A} - \lambda\boldsymbol{E}|$$

не меняется при замене матрицы $\boldsymbol{A}$ подобной матрицей. Мы построим здесь ряд инвариантов, среди которых будет содержаться и характеристический многочлен; они будут *полной системой инвариантов*, в том смысле, что из их совпадения для двух матриц следует подобие этих матриц.

Пусть $\boldsymbol{A}$ — произвольная матрица n-го порядка. Миноры k-го порядка матрицы $\boldsymbol{A} - \lambda\boldsymbol{E}$ суть некоторые многочлены от λ. *Обозначим через $D_k(\lambda)$ их наибольший общий делитель* *). В частности, $D_n(\lambda)$ — определитель матрицы $\boldsymbol{A} - \lambda\boldsymbol{E}$, т. е. характеристический многочлен матрицы $\boldsymbol{A}$. В дальнейшем мы покажем, что все $D_k(\lambda)$ являются инвариантами.

Заметим, что $D_n(\lambda)$ делится на $D_{n-1}(\lambda)$. Действительно, по определению $D_{n-1}(\lambda)$ все миноры $(n-1)$-го порядка делятся на $D_{n-1}(\lambda)$. Разлагая определитель на сумму произведений элементов какой-нибудь строки на их алгебраические дополнения, мы получаем, что и определитель $D_n(\lambda)$ делится на $D_{n-1}(\lambda)$. Аналогично, $D_{n-1}(\lambda)$ делится на $D_{n-2}(\lambda)$ и т. д.

Упражнение. Найти $D_k(\lambda)$ $(k = 1, 2, 3)$ для матрицы

$$\begin{pmatrix} \lambda_0 & 1 & 0 \\ 0 & \lambda_0 & 1 \\ 0 & 0 & \lambda_0 \end{pmatrix}.$$

Ответ. $D_3(\lambda) = (\lambda - \lambda_0)^3$, $D_2(\lambda) = D_1(\lambda) = 1$.

Лемма 1. *Если $\boldsymbol{C}$ — произвольная невырожденная матрица, то общие наибольшие делители миноров k-го порядка матриц $\boldsymbol{A} - \lambda\boldsymbol{E}$ и $\boldsymbol{C}(\boldsymbol{A} - \lambda\boldsymbol{E})$ совпадают. Аналогичное утверждение имеет место и для $(\boldsymbol{A} - \lambda\boldsymbol{E})\boldsymbol{C}$.*

---

*) Наибольший общий делитель определен с точностью до числового множителя. Мы выбираем $D_k(\lambda)$ так, чтобы старший коэффициент был равен 1. В частности, если миноры k-го порядка взаимно просты, то $D_k(\lambda) = 1$.

Доказательство. Строки матрицы $\boldsymbol{C}(\boldsymbol{A}-\lambda\boldsymbol{E})$ являются линейными комбинациями строк матрицы $\boldsymbol{A}-\lambda\boldsymbol{E}$ с коэффициентами, являющимися элементами матрицы $\boldsymbol{C}$, т. е. не зависящими от λ. Действительно, обозначим через a_{ik} элементы матрицы $\boldsymbol{A}-\lambda\boldsymbol{E}$ и через a'_{ik} элементы матрицы $\boldsymbol{C}(\boldsymbol{A}-\lambda\boldsymbol{E})$. Тогда, например,

$$a'_{1k}=\sum_{j=1}^{n} c_{1j}a_{jk},$$

т. е. элементы первой строки матрицы $\boldsymbol{C}(\boldsymbol{A}-\lambda\boldsymbol{E})$ являются линейными комбинациями строк матрицы $\boldsymbol{A}-\lambda\boldsymbol{E}$ с коэффициентами c_{1j}. Аналогично показываем это и для других строк.

Поэтому минор матрицы $\boldsymbol{C}(\boldsymbol{A}-\lambda\boldsymbol{E})$ разлагается на сумму миноров матрицы $\boldsymbol{A}-\lambda\boldsymbol{E}$ с некоторыми численными коэффициентами. Следовательно, всякий делитель миноров k-го порядка матрицы $\boldsymbol{A}-\lambda\boldsymbol{E}$ будет также делителем миноров того же порядка матрицы $\boldsymbol{C}(\boldsymbol{A}-\lambda\boldsymbol{E})$. Так как от матрицы $\boldsymbol{C}(\boldsymbol{A}-\lambda\boldsymbol{E})$ мы можем перейти к матрице $\boldsymbol{A}-\lambda\boldsymbol{E}$ умножением на $\boldsymbol{C}^{-1}$, то и обратно, каждый делитель миноров k-го порядка матрицы $\boldsymbol{C}(\boldsymbol{A}-\lambda\boldsymbol{E})$ является делителем миноров k-го порядка матрицы $\boldsymbol{A}-\lambda\boldsymbol{E}$. Следовательно, у $\boldsymbol{A}-\lambda\boldsymbol{E}$ и $\boldsymbol{C}(\boldsymbol{A}-\lambda\boldsymbol{E})$ общие делители миноров k-го порядка совпадают.

Лемма 2. *У подобных матриц многочлены $D_k(\lambda)$ совпадают.*

Доказательство. Пусть $\boldsymbol{A}$ и $\boldsymbol{A}'=\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}$ — две подобные матрицы. Согласно предыдущей лемме, общие наибольшие делители миноров k-го порядка у $\boldsymbol{A}-\lambda\boldsymbol{E}$ и $(\boldsymbol{A}-\lambda\boldsymbol{E})\boldsymbol{C}$ совпадают. По той же лемме совпадают между собой общие наибольшие делители миноров k-го порядка у $\boldsymbol{C}^{-1}(\boldsymbol{A}-\lambda\boldsymbol{E})$ и $\boldsymbol{C}^{-1}(\boldsymbol{A}-\lambda\boldsymbol{E})\boldsymbol{C}=$ $=\boldsymbol{A}'-\lambda\boldsymbol{E}$. Следовательно, $D_k(\lambda)$ для $\boldsymbol{A}$ и $\boldsymbol{A}'$ равны между собой.

Так как при переходе от одного базиса к другому матрица линейного преобразования заменяется подобной, то из леммы 2 вытекает следующая

Теорема 1. *Пусть A — линейное преобразование. Тогда наибольший общий делитель $D_k(\lambda)$ миноров k-го порядка матрицы $\boldsymbol{A} - \lambda\boldsymbol{E}$, где $\boldsymbol{A}$ — матрица преобразования A в некотором базисе, не зависит от выбора базиса.*

Перейдем к вычислению многочленов $D_k(\lambda)$ для данного линейного преобразования A. В силу теоремы 1, при их вычислении можно пользоваться матрицей линейного преобразования в любом базисе. Выберем базис, в котором матрица линейного преобразования имеет жорданову нормальную форму. Нам нужно, следовательно, вычислить многочлены $D_k(\lambda)$ для матрицы $\boldsymbol{A}$, имеющей нормальную форму.

Найдем сначала все $D_k(\lambda)$ для матрицы n-го порядка вида

$$\begin{pmatrix} \lambda_0 & 1 & 0 & \dots & 0 \\ 0 & \lambda_0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & \lambda_0 \end{pmatrix}, \tag{1}$$

т. е. для одной «клетки» нормальной формы. Мы имеем $D_n(\lambda) = (\lambda - \lambda_0)^n$. Если в матрице (1) зачеркнуть первый столбец и последнюю строку, то получим матрицу $\boldsymbol{A}_1$, в которой по диагонали стоят единицы, а над диагональю нули. Поэтому $D_{n-1}(\lambda) = 1$. Вычеркивая далее в матрице $\boldsymbol{A}_1$ строки и столбцы с одинаковыми номерами, мы сможем доказать, что и $D_{n-2}(\lambda) = \dots$ $\dots = D_1(\lambda) = 1$. Окончательно мы имеем, что *для отдельной клетки* [*матрицы* (1)] *последовательность* $D_k(\lambda)$ *следующая*:

$$(\lambda - \lambda_0)^n,\ 1,\ 1,\ \dots,\ 1.$$

Далее, заметим: *пусть матрица* $\boldsymbol{B}$ *имеет вид*

$$\begin{pmatrix} \boldsymbol{B}_1 & 0 \\ 0 & \boldsymbol{B}_2 \end{pmatrix},$$

где $\boldsymbol{B}_1$ *и* $\boldsymbol{B}_2$ — *какие-либо матрицы порядков* n_1 *и* n_2. *Тогда отличные от нуля миноры* m-*го порядка матрицы* $\boldsymbol{B}$ *имеют вид*

$$\Delta_m = \Delta_{m_1}^{(1)} \Delta_{m_2}^{(2)}, \quad m_1 + m_2 = m,$$

где $\Delta_{m_1}^{(1)}$ — *миноры* m_1-*го порядка матрицы* $\boldsymbol{B}_1$, *а* $\Delta_{m_2}^{(2)}$ — *миноры* m_2-*го порядка матрицы* $\boldsymbol{B}_2$ *). Действительно, если выделить те из первых n_1 строк, которые входят в состав данного минора, и разложить по ним минор (воспользовавшись теоремой Лапласа), то этот минор будет либо равен нулю, либо иметь вид $\Delta_{m_1}^{(1)} \Delta_{m-m_1}^{(2)}$.

Найдем теперь многочлены $D_k(\lambda)$ для произвольной матрицы $\boldsymbol{A}$, имеющей жорданову нормальною форму. Мы предположим, что в матрице $\boldsymbol{A}$ имеется p клеток, отвечающих собственному значению λ_1, q клеток, отвечающих собственному значению λ_2, и т. д. Обозначим порядки клеток, отвечающих собственному значению λ_1, через $n_1, n_2, \ldots, n_p$ $(n_1 \geqslant n_2 \geqslant n_3 \geqslant \ldots \geqslant n_p)$.

Матрица $\boldsymbol{B} = \boldsymbol{A} - \lambda \boldsymbol{E}$ распадается на отдельные клетки $\boldsymbol{B}_i$, из которых, например, $\boldsymbol{B}_1$ имеет вид

$$\boldsymbol{B_1} = \begin{pmatrix} \lambda_1 - \lambda & 1 & 0 & \ldots & 0 \\ 0 & \lambda_1 - \lambda & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & \lambda_1 - \lambda \end{pmatrix}.$$

*) Отличный от нуля минор Δ_k k-го порядка матрицы $\boldsymbol{B}$ может, конечно, иметь вид $\Delta_k^{(1)}$, т. е. быть составлен из элементов $\boldsymbol{B}_1$. В этом случае мы его запишем формально в виде $\Delta_k = \Delta_k^{(1)} \Delta_0^{(2)}$, где положено $\Delta_0^{(2)} = 1$.

Вычислим сначала $D_n(\lambda)$, т. е. определитель матрицы $\boldsymbol{B}$. Он равен произведению определителей матриц $\boldsymbol{B}_1$, т. е.

$$D_n(\lambda) = (\lambda - \lambda_1)^{n_1+n_2+\ldots+n_p}(\lambda - \lambda_2)^{m_1+m_2+\ldots+m_q}\ldots$$

Перейдем теперь к вычислению $D_{n-1}(\lambda)$. Так как $D_{n-1}(\lambda)$ есть делитель многочлена $D_n(\lambda)$, то $D_{n-1}(\lambda)$ состоит из множителей $\lambda - \lambda_1$, $\lambda - \lambda_2$, $\ldots$. Вычислим, в какой степени в $D_{n-1}(\lambda)$ входит $\lambda - \lambda_1$. Для этого заметим, что произвольный отличный от нуля минор $(n-1)$-го порядка матрицы $\boldsymbol{B} = \boldsymbol{A} - \lambda\boldsymbol{E}$ имеет вид

$$\Delta_{n-1} = \Delta_{t_1}^{(1)}\Delta_{t_2}^{(2)}\ldots\Delta_{t_k}^{(k)},$$

где $t_1+t_2+\ldots+t_k = n-1$, $\Delta_{t_i}^{(i)}$ миноры порядка t_i матрицы $\boldsymbol{B}_i$. Так как сумма порядков миноров $\Delta_{t_i}^{(1)}, \ldots$ равна $n-1$, то один и только один из этих миноров имеет порядок на единицу ниже, чем порядок соответствующей матрицы $\boldsymbol{B}_i$, т. е. получается из соответствующей клетки матрицы $\boldsymbol{B}$ вычеркиванием одной строки и одного столбца. Мы видели, что в отдельной клетке мы можем вычеркиванием одной строки и одного столбца получить минор, равный единице (см. стр. 234). Поэтому мы можем подобрать Δ_{n-1} так, чтобы какой-нибудь один из миноров $\Delta_{t_i}^{(i)}$ стал равным единице, не меняя при этом остальных, равных определителям соответствующих клеток. Отсюда ясно, что, для того чтобы получить минор, содержащий $\lambda - \lambda_1$ в возможно более низкой степени, достаточно вычеркнуть строку и столбец в клетке, отвечающей λ_1 и имеющей наибольший порядок, а именно порядок n_1. Таким образом, наибольший общий делитель $D_{n-1}(\lambda)$ миноров $(n-1)$-го порядка содержит $\lambda - \lambda_1$ в степени $n_2 + n_3 + \ldots + n_p$.

Аналогично, среди миноров $(n-2)$-го порядка наинизшую степень $\lambda - \lambda_1$ содержит минор Δ_{n-2}, полученный вычеркиванием по строке и столбцу из клеток, соответствующих собственному значению λ_1 и имеющих порядки n_1 и n_2. Таким образом, $D_{n-2}(\lambda)$ содержит $\lambda - \lambda_1$ в степени $n_3 + n_4 + \ldots + n_p$ и т. д. Наконец, $D_{n-p}(\lambda), D_{n-p-1}(\lambda), \ldots, D_1(\lambda)$ вовсе не содержит $\lambda - \lambda_1$.

Совершенно так же мы выясняем, в каких степенях в $D_k(\lambda)$ входят множители $\lambda - \lambda_2$, $\lambda - \lambda_3$, ...

Итак, мы доказали следующее утверждение:

Пусть матрица преобразования A имеет жорданову нормальную форму, в которой имеется p клеток порядков $n_1, n_2, \ldots, n_p$ $(n_1 \geqslant n_2 \geqslant \ldots \geqslant n_p)$, отвечающих собственному значению λ_1, q клеток порядков $m_1, m_2, \ldots, m_q$ $(m_1 \geqslant m_2 \geqslant \ldots \geqslant m_q)$, отвечающих собственному значению λ_2, и т. д.; тогда

$$
\begin{aligned}
&D_n(\lambda) = (\lambda-\lambda_1)^{n_1+n_2+n_3+\ldots+n_p}(\lambda-\lambda_2)^{m_1+m_2+m_3+\ldots+m_q}\ldots,\\
&D_{n-1}(\lambda) = (\lambda-\lambda_1)^{n_2+n_3+\ldots+n_p}(\lambda-\lambda_2)^{m_2+m_3+\ldots+m_q}\ldots,\\
&D_{n-2}(\lambda) = (\lambda-\lambda_1)^{n_3+\ldots+n_p}(\lambda-\lambda_2)^{m_3+\ldots+m_q}\ldots,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots
\end{aligned}
$$

При этом, начиная с $D_{n-p}(\lambda)$ множитель $(\lambda - \lambda_1)^{\cdots}$ заменяется единицей, начиная с $D_{n-q}(\lambda)$, множитель $(\lambda - \lambda_2)^{\cdots}$ заменяется единицей и т. д.

Рассмотрим важный пример. Пусть собственному значению λ_1 отвечает лишь одна клетка, порядок которой равен n_1, собственному значению λ_2 — только одна клетка порядка m_1, λ_3 — одна клетка порядка k_1 и т. д. (т. е. собственные значения, отвечающие различным клеткам, различны). Тогда $D_i(\lambda)$ имеет вид

$$
\begin{aligned}
&D_n(\lambda) = (\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{m_1}(\lambda-\lambda_3)^{k_1}\ldots,\\
&D_{n-1}(\lambda) = 1,\\
&D_{n-2}(\lambda) = 1,\\
&\ldots\ldots\ldots
\end{aligned}
$$

Указанный выше общий вид для $D_k(\lambda)$ показывает, что вместо многочленов $D_k(\lambda)$ удобнее ввести их отношения

$$E_k(\lambda) = \frac{D_k(\lambda)}{D_{k-1}(\lambda)}.$$

Многочлены $E_k(\lambda)$ называются *инвариантными множителями.* Таким образом, если *матрица* $\boldsymbol{A}$ *имеет жорданову нормальную форму, в которой имеется* p *«клеток» порядков* $n_1, n_2, \ldots, n_p$ $(n_1 \geqslant n_2 \geqslant \ldots \geqslant n_p)$, *отвечающих собственному значению* λ_1, q *«клеток» порядков* $m_1, m_2, \ldots, m_q$ $(m_1 \geqslant m_2 \geqslant \ldots \geqslant m_q)$, *отвечающих собственному значению* λ_2 *и т. д., то инвариантные множители* $E_k(\lambda)$ *имеют вид*

$$\begin{aligned}
&E_n(\lambda) = (\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{m_1}\ldots,\\
&E_{n-1}(\lambda) = (\lambda - \lambda_1)^{n_2}(\lambda - \lambda_2)^{m_2}\ldots,\\
&E_{n-2}(\lambda) = (\lambda - \lambda_1)^{n_3}(\lambda - \lambda_2)^{m_3}\ldots,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots
\end{aligned}$$

Мы видим, что задание последовательности инвариантных множителей $E_n(\lambda), E_{n-1}(\lambda), \ldots$ полностью определяет жорданову нормальную форму матрицы $\boldsymbol{A}$; собственные значения λ_i получаются как корни уравнения $E_n(\lambda) = 0$. Размеры же $n_1, n_2, \ldots, n_p$ клеток, отвечающих данному собственному значению λ_1, равны степеням, с которыми двучлен $\lambda - \lambda_1$ содержится соответственно в $E_n(\lambda), E_{n-1}(\lambda), \ldots$

Мы теперь в состоянии сформулировать необходимые и достаточные условия существования базиса, в котором матрица линейного преобразования диагональна.

Для того чтобы существовал базис, в котором матрица преобразования диагональна, необходимо и достаточно, чтобы инвариантные множители этой матрицы имели лишь простые корни.

Действительно, мы видели, что кратности корней $\lambda_1, \lambda_2, \ldots$ инвариантных множителей определяют порядки клеток в жордановой нормальной форме. Простота корней инвариантных множителей означает, таким образом, что эти клетки первого порядка, т. е. что жорданова нормальная форма матрицы сводится к диагональной.

Теорема 2. *Для того чтобы две матрицы были подобны, необходимо и достаточно, чтобы их инвариантные множители совпадали.*

Доказательство. Мы доказали (лемма 2), что у подобных матриц совпадают многочлены $D_k(\lambda)$. Следовательно, совпадают и инвариантные множители $E_k(\lambda)$, являющиеся их отношениями.

Обратно, пусть инвариантные множители матриц $\boldsymbol{A}$ и $\boldsymbol{B}$ совпадают. Мы знаем, что каждая матрица подобна некоторой матрице, имеющей жорданову нормальную форму. Так как инвариантные множители у $\boldsymbol{A}$ и $\boldsymbol{B}$ совпадают, то их нормальные жордановы формы тоже совпадают. Таким образом, матрицы $\boldsymbol{A}$ и $\boldsymbol{B}$ подобны одной и той же матрице и, значит, $\boldsymbol{A}$ подобна $\boldsymbol{B}$.

Теорема 3. *Нормальная форма линейного преобразования однозначно определяется самим линейным преобразованием.*

Доказательство. Матрицы преобразования A в различных базисах подобны. Так как подобные матрицы имеют одинаковые инвариантные множители, а инвариантными множителями однозначно определяется нормальная форма, то нормальная форма данного линейного преобразования определена однозначно, и теорема доказана.

Мы в состоянии теперь найти жорданову нормальную форму матрицы линейного преобразования. Для этого достаточно взять матрицу линейного преобразования в каком-нибудь базисе и найти инвариантные множители матрицы $\boldsymbol{A}$. Разложив инвариантные множители в произведение вида $(\lambda-\lambda_1)^n(\lambda-\lambda_2)^m\ldots$, мы будем знать как собственные значения, так и отвечающие им порядки клеток.

§ 22. λ-матрицы

1. λ-матрицей (полиномиальной матрицей) называется матрица, элементами которой являются многочлены относительно некоторой буквы λ. Степенью λ-матрицы называется наивысшая из степеней многочленов, входящих в состав матрицы. Ясно, что λ-матрица степени n может быть представлена в виде

$$A_0\lambda^n + A_1\lambda^{n-1} + \ldots + A_n,$$

где A_k — матрицы, уже не зависящие от λ *). Частный случай λ-матриц нам уже встречался неоднократно, а именно матрицы вида $A-\lambda E$. Результаты, которые мы получим в этом параграфе, для случая λ-матриц вида $A-\lambda E$ содержат как частный случай многие из результатов, полученных в предыдущих параграфах этой главы.

λ-матрицы встречаются во всех вопросах математики. Так, например, решение системы однородных линейных дифференциальных уравнений первого порядка с постоянными коэффициентами

$$\frac{dy_i}{dx} = \sum_{k=1}^{n} a_{ik}y_k \quad (i=1,2,\ldots,n) \tag{1}$$

ищется обычно в виде

$$y_k = c_k e^{\lambda x}, \tag{2}$$

где λ и c_k — некоторые постоянные. Для их определения подставим функции (2) в систему и сократим уравнение на $e^{\lambda x}$. Мы

*) В этом параграфе мы будем матрицы обозначать светлыми буквами.

получим систему линейных уравнений

$$\lambda c_i = \sum_{k=1}^{n} a_{ik} c_k,$$

матрица которой есть $A - \lambda E$, где A — матрица из коэффициентов системы (1). Таким образом, изучение системы дифференциальных уравнений (1) тесно связано с λ-матрицей первой степени относительно λ: $A - \lambda E$.

Аналогично, исследование системы уравнений порядка выше первого приводит к исследованию λ-матриц высших степеней. Например, исследование системы уравнений

$$\sum_{k=1}^{n} a_{ik} \frac{d^2 y_k}{dx^2} + \sum_{k=1}^{n} b_{ik} \frac{dy_k}{dx} + \sum_{k=1}^{n} c_{ik} y_k = 0$$

приводит к исследованию λ-матрицы $A\lambda^2 + B\lambda + C$, где $A = \|a_{ik}\|$, $B = \|b_{ik}\|$, $C = \|c_{ik}\|$.

Мы рассмотрим сейчас вопрос о каноническом виде λ-матриц относительно так называемых элементарных преобразований.

Элементарными преобразованиями λ-матриц называются преобразования следующих типов.

1° Перестановка между собой двух каких-либо строк или столбцов матрицы.

2° Прибавление к строке какой-либо другой строки, умноженной на некоторый многочлен $\varphi(\lambda)$, и, аналогично, прибавление к столбцу другого столбца, умноженного на некоторый многочлен.

3° Умножение строки или столбца на некоторое число, отличное от нуля.

Определение. *Две λ-матрицы называются эквивалентными, если одна может быть получена из другой некоторой последовательностью элементарных преобразований.*

Обратное к каждому элементарному преобразованию есть снова элементарное преобразование. Это легко проверяется для каждого из трех типов элементарных преобразований. Так, если λ-матрица $B(\lambda)$

получается из λ-матрицы $A(\lambda)$ перестановкой строк, то обратной перестановкой строк мы можем из $B(\lambda)$ получить $A(\lambda)$. Если $B(\lambda)$ получается из $A(\lambda)$ прибавлением к k-й строке i-й, умноженной на $\varphi(\lambda)$, то, обратно, $A(\lambda)$ можно получить из $B(\lambda)$ прибавлением к k-й строке i-й, умноженной на $-\varphi(\lambda)$.

Из сделанного замечания следует, что если λ-матрица $K(\lambda)$ эквивалентна $L(\lambda)$, то и обратно, $L(\lambda)$ эквивалентна $K(\lambda)$. В самом деле, пусть из $K(\lambda)$ применением некоторой последовательности элементарных преобразований получается $L(\lambda)$. Тогда, применяя к $L(\lambda)$ в обратном порядке обратные преобразования, мы придем к $K(\lambda)$.

Если две λ-матрицы $K_1(\lambda)$ и $K_2(\lambda)$ эквивалентны некоторой матрице $K(\lambda)$, то они эквивалентны между собой. Действительно, если сначала провести последовательность элементарных преобразований, переводящих $K_1(\lambda)$ в $K(\lambda)$, а затем элементарные преобразования, переводящие $K(\lambda)$ в $K_2(\lambda)$, то мы переведем $K_1(\lambda)$ в $K_2(\lambda)$, т.е. $K_1(\lambda)$ эквивалентна $K_2(\lambda)$.

Основной результат п. 1 этого параграфа состоит в доказательстве теоремы о том, что всякую λ-матрицу можно элементарными преобразованиями привести к диагональному виду. Доказательству этого предложения предпошлем лемму:

Лемма 1. *Если элемент $a_{11}(\lambda)$ в λ-матрице $A(\lambda)$ не равен нулю и если не все элементы $a_{ik}(\lambda)$ матрицы $A(\lambda)$ делятся на многочлен $a_{11}(\lambda)$, то можно подобрать эквивалентную $A(\lambda)$ λ-матрицу $B(\lambda)$, для которой элемент $b_{11}(\lambda)$ также не равен нулю и имеет степень более низкую, чем $a_{11}(\lambda)$.*

Доказательство. Предположим сначала, что не делящийся на $a_{11}(\lambda)$ элемент матрицы $A(\lambda)$ находится в первой строке. Пусть, например, $a_{1k}(\lambda)$ не делится на $a_{11}(\lambda)$. Тогда $a_{1k}(\lambda)$ можно представить в виде

$$a_{1k}(\lambda) = a_{11}(\lambda)\varphi(\lambda) + b(\lambda),$$

где $\varphi(\lambda)$ — частное, $b(\lambda) \neq 0$ — остаток от деления $a_{1k}(\lambda)$ на $a_{11}(\lambda)$ и, следовательно, степень $b(\lambda)$ ниже, чем степень $a_{11}(\lambda)$. Вычтем из k-го столбца первый, умноженный на $\varphi(\lambda)$. Получим матрицу, где вместо $a_{1k}(\lambda)$ стоит теперь многочлен $b(\lambda)$, имеющий более низкую степень, чем $a_{11}(\lambda)$. Переставляя теперь k-й столбец с первым, мы переведем $b(\lambda)$ в левый верхний угол.

Рассмотрим теперь случай, когда все элементы первой строки и первого столбца делятся на $a_{11}(\lambda)$, а некоторый элемент $a_{ik}(\lambda)$ не делится на $a_{11}(\lambda)$. Этот случай мы сведем к предыдущему следующим образом: $a_{i1}(\lambda)$ делится на $a_{11}(\lambda)$, т. е. имеет вид $a_{i1}(\lambda) = \varphi(\lambda)a_{11}(\lambda)$. Вычтем из i-й строки первую, умноженную на $\varphi(\lambda)$. Тогда $a_{i1}(\lambda)$ заменится нулем, элемент $a_{ik}(\lambda)$ заменится элементом $a'_{ik}(\lambda) = a_{ik}(\lambda) - \varphi(\lambda)a_{1k}(\lambda)$, который по-прежнему не делится на $a_{11}(\lambda)$ (так как $a_{1k}(\lambda)$ по предположению делится на $a_{11}(\lambda)$). Прибавим теперь i-ю строку к первой. Так как на первом месте в i-й строке теперь стоит нуль, то $a_{11}(\lambda)$ не изменится, а на k-м месте в первой строке теперь будет стоять $a_{1k}(\lambda) + a'_{ik}(\lambda) = a_{1k}(\lambda)(1 - \varphi(\lambda)) + a_{ik}(\lambda)$ и, следовательно, в первой строке имеется элемент, не делящийся на $a_{11}(\lambda)$. Мы свели этот случай к рассмотренному выше и, следовательно, лемма доказана.

В дальнейшем мы будем пользоваться также следующим замечанием: если все элементы λ-матрицы $B(\lambda)$ делятся на некоторый многочлен $E(\lambda)$, то после элементарных преобразований над матрицей $B(\lambda)$ мы снова получим матрицу, элементы которой делятся на $E(\lambda)$.

Перейдем теперь к приведению λ-матрицы к диагональному виду.

Мы можем считать, что $a_{11}(\lambda) \neq 0$, так как, если в матрице есть хоть один элемент, отличный от нуля, то перестановками строк и столбцов его можно перевести на это место. Если не все элементы матрицы делятся на $a_{11}(\lambda)$, то мы можем способом, указанном в лемме, заменить матрицу эквивалентной, в которой элемент, стоящий в левом верхнем углу, имеет более низкую степень и по-прежнему отличен от нуля. Если не все элементы делятся на него, то мы можем опять понизить степень этого элемента и т. д. Процесс закончится, когда мы придем к матрице $B(\lambda)$, в которой все элементы делятся на $b_{11}(\lambda)$.

Так как элементы $b_{12}(\lambda), \ldots, b_{1n}(\lambda)$ первой строки делятся на $b_{11}(\lambda)$, то, вычитая из второго, третьего и т. д. столбца первый, умноженный на соответственно подобранные многочлены от λ, мы можем обратить в нуль 2-й, 3-й, $\ldots$, n-й элементы первой строки. Аналогично обратим в нуль все элементы, начиная со второго, в первом столбце. Так как в матрице $B(\lambda)$ все элементы делились на $b_{11}(\lambda)$, то в полученной матрице все элементы также делятся на $b_{11}(\lambda)$. Разделим все элементы первой строки на старший коэффициент многочлена $b_{11}(\lambda)$. На первом месте получится многочлен со старшим коэффициентом 1, который мы обозначим через $E_1(\lambda)$, а на остальных местах будут по-прежнему нули.

Мы пришли, таким образом, к матрице следующего вида:

$$\begin{pmatrix} E_1(\lambda) & 0 & 0 & \ldots & 0 \\ 0 & c_{22}(\lambda) & c_{23}(\lambda) & \ldots & c_{2n}(\lambda) \\ 0 & c_{32}(\lambda) & c_{33}(\lambda) & \ldots & c_{3n}(\lambda) \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & c_{n2}(\lambda) & c_{n3}(\lambda) & \ldots & c_{nn}(\lambda) \end{pmatrix}, \tag{3}$$

все элементы которой делятся на $E_1(\lambda)$.

Мы можем теперь повторить с матрицей $(n-1)$-го порядка $\|c_{ik}(\lambda)\|$ те же операции, что с матрицей n-го порядка. Заметим, что всякое элементарное преобразование матрицы $\|c_{ik}\|$ есть в то же время элементарное преобразование матрицы (3), так как в первой строке и столбце все элементы, кроме $E_1(\lambda)$, равны нулю.

Таким образом, мы обратим в нуль все элементы второй строки и второго столбца, кроме диагонального. Полученный диагональный элемент (старший коэффициент которого также считаем равным единице) обозначим $E_2(\lambda)$. Все элементы $c_{ik}(\lambda)$ делятся на $E_1(\lambda)$. Поэтому все дальнейшие элементарные преобразования всегда приводят нас к элементам, делящимся на $E_1(\lambda)$. В частности, $E_2(\lambda)$ делится на $E_1(\lambda)$.

Мы пришли, таким образом, к матрице, у которой в первых двух строках и столбцах все элементы, кроме диагональных, равны нулю, а по диагонали стоят $E_1(\lambda)$ и $E_2(\lambda)$, причем $E_2(\lambda)$ делится на $E_1(\lambda)$. Мы сможем продолжать этот процесс далее, пока не приведем всю матрицу к диагональному виду. Может, конечно, оказаться, что мы закончим процесс раньше, придя к матрице, состоящей сплошь из нулей.

Итак, доказана следующая

Теорема 1. *Всякая λ-матрица может быть элементарными преобразованиями приведена к виду*

$$\begin{pmatrix} E_1(\lambda) & 0 & 0 & \dots & 0 \\ 0 & E_2(\lambda) & 0 & \dots & 0 \\ 0 & 0 & E_3(\lambda) & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & E_n(\lambda) \end{pmatrix}, \tag{4}$$

где многочлены $E_k(\lambda)$, стоящие по диагонали, имеют старшие коэффициенты, равные единице, многочлен $E_2(\lambda)$ делится на $E_1(\lambda)$, $E_3(\lambda)$ делится на $E_2(\lambda)$,

$E_4(\lambda)$ на $E_3(\lambda)$ и т. д. Этот вид называется нормальной диагональной формой λ-матрицы.

Конечно, некоторое число последних многочленов $E_k(\lambda)$ в матрице (4) может оказаться равным нулю:

$$E_{r+1}(\lambda) = E_{r+2}(\lambda) = \ldots = 0.$$

З а м е ч а н и е. Мы привели матрицу $A(\lambda)$ к нормальному диагональному виду, в котором каждый диагональный элемент делится на предшествующий. Если поставить себе целью приведение матрицы к какому-нибудь диагональному виду, отбросив требование делимости, то задача решается проще.

Действительно, для того чтобы обратить в нуль все элементы первой строки и первого столбца кроме $a_{11}(\lambda)$, достаточно, чтобы эти элементы (а не все элементы матрицы) делились на $a_{11}(\lambda)$. Как видно из доказательства леммы, для того чтобы этого достигнуть, требуется гораздо меньшее число элементарных преобразований, чем для приведения к нормальной диагональной форме. Обратив в нуль все элементы первой строки и первого столбца, кроме диагонального, мы можем проделать то же самое с оставшейся матрицей $(n-1)$-го порядка и т. д., пока матрица не будет приведена к диагональному виду. Этим путем можно привести матрицу к различным диагональным формам, т. е. диагональная форма не определена однозначно. В следующем пункте этого параграфа мы покажем, что *нормальная* диагональная форма данной λ-матрицы определяется уже однозначно.

У п р а ж н е н и е. Привести λ-матрицу

$$\begin{pmatrix} \lambda-\lambda_1 & 0 \\ 0 & \lambda-\lambda_2 \end{pmatrix}, \quad \lambda_1 \neq \lambda_2,$$

к нормальной диагональной форме.

Ответ.

$$\begin{pmatrix} 1 & 0 \\ 0 & (\lambda-\lambda_1)(\lambda-\lambda_2) \end{pmatrix}.$$

2. В этом пункте мы докажем, что нормальная диагональная форма данной матрицы определена однозначно. Для этого мы построим систему многочленов, связанных с данной λ-матрицей, которые не меняются при элементарных преобразованиях и которыми, как мы увидим, нормальная диагональная форма λ-матрицы вполне определяется.

Пусть дана произвольная λ-матрица. Наибольший общий делитель всех миноров k-го порядка данной λ-матрицы обозначим через $D_k(\lambda)$. Так как наибольший общий делитель определен с точностью до постоянного множителя, то будем считать, что старший коэффициент у $D_k(\lambda)$ равен единице. В частности, если общий наибольший делитель миноров k-го порядка равен постоянной, то мы полагаем $D_k(\lambda) = 1$.

Докажем, что элементарные преобразования не меняют многочленов $D_k(\lambda)$, т.е. что у эквивалентных λ-матриц многочлены $D_k(\lambda)$ совпадают.

Для элементарных преобразований вида 1°, переставляющих строки или столбцы, это очевидно, так как при них каждый минор k-го порядка либо вовсе не меняется, либо меняет знак, либо заменяется другим минором k-го порядка, что, конечно, не меняет общего наибольшего делителя всех таких миноров. Аналогично, не меняют $D_k(\lambda)$ элементарные преобразования вида 3°, так как при этих преобразованиях миноры самое большее умножаются на постоянное. Рассмотрим теперь элементарное преобразование вида 2°, например, прибавим к i-му столбцу j-й, умноженный на $\varphi(\lambda)$. При этом минор k-го порядка вовсе не изменится, если он содержит и i-й и j-й столбцы либо если он не содержит ни одного из них.

В случае же, если минор содержит i-й столбец и не содержит j-го столбца, то его можно представить как комбинацию двух миноров, которые имелись у исходной матрицы. Таким образом, наибольший общий делитель миноров k-го порядка и в этом случае не изменится.

Если все миноры порядка k, а следовательно, и более высоких порядков, матрицы $A(\lambda)$ равны нулю, то мы будем считать $D_k(\lambda) \equiv D_{k+1}(\lambda) \equiv \ldots \equiv D_n(\lambda) = 0$. Заметим, что из совпадения у всех эквивалентных матриц многочленов $D_k(\lambda)$ следует, что эквивалентные матрицы имеют один и тот же ранг.

Найдем многочлены $D_k(\lambda)$ для матрицы, приведенной к нормальной диагональной форме:

$$\begin{pmatrix} E_1(\lambda) & 0 & \ldots & 0 \\ 0 & E_2(\lambda) & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & E_n(\lambda) \end{pmatrix}. \tag{5}$$

Заметим, что у диагональной матрицы отличны от нуля только главные миноры, т. е. миноры, в которые входят строки и столбцы с одинаковыми номерами. Эти миноры имеют вид $E_{i_1}(\lambda)E_{i_2}(\lambda)\ldots E_{i_k}(\lambda)$.

Так как $E_2(\lambda)$ делится на $E_1(\lambda)$, $E_3(\lambda)$ делится на $E_2(\lambda)$ и т. д., то наибольший общий делитель миноров первого порядка $D_1(\lambda)$ равен $E_1(\lambda)$. Так как все многочлены $E_k(\lambda)$ делятся на $E_1(\lambda)$, а все многочлены, кроме $E_1(\lambda)$, делятся на $E_2(\lambda)$, то произведение $E_i(\lambda)E_j(\lambda)$ $(i < j)$ всегда делится на минор $E_1(\lambda)E_2(\lambda)$. Таким образом, $D_2(\lambda) = E_1(\lambda)E_2(\lambda)$. Так как, кроме того, все $E_k(\lambda)$, кроме $E_1(\lambda)$ и $E_2(\lambda)$, делятся на $E_3(\lambda)$, то $E_i(\lambda)E_j(\lambda)E_k(\lambda)$ $(i < j < k)$ делится на минор $E_1(\lambda)E_2(\lambda)E_3(\lambda)$ и, следовательно, $D_3(\lambda) = E_1(\lambda)E_2(\lambda)E_3(\lambda)$.

Таким же образом для матрицы (4)

$$D_k(\lambda) = E_1(\lambda)E_2(\lambda)\ldots E_k(\lambda) \quad (k = 1, 2, \ldots, n). \tag{6}$$

Очевидно, что если, начиная с некоторого r, $E_{r+1}(\lambda) = E_{r+2}(\lambda) = \ldots = E_n(\lambda) = 0$, то $D_{r+1}(\lambda) = D_{r+2}(\lambda) = \ldots = D_n(\lambda) = 0$.

Отсюда получается, что для λ-матрицы, имеющей нормальную диагональную форму (5), диагональные элементы $E_k(\lambda)$ вычисляются по формулам

$$E_k(\lambda) = \frac{D_k(\lambda)}{D_{k-1}(\lambda)}.$$

При этом, если $D_{r+1}(\lambda) = \ldots = D_n(\lambda) = 0$, то надо положить $E_{r+1}(\lambda) = \ldots = E_n(\lambda) = 0$.

Многочлены $E_k(\lambda)$ *называются инвариантными множителями.* В § 20 мы уже определили их для матриц вида $A - \lambda E$.

Т е о р е м а 2. *Нормальная диагональная форма данной λ-матрицы $A(\lambda)$ определяется по ней однозначно. Если $D_k(\lambda)$ $(k = 2, 3, \ldots, r)$ — наибольший общий делитель миноров k-го порядка матрицы $A(\lambda)$, а $D_{r+1}(\lambda) = \ldots = D_n(\lambda) = 0$, то элементы нормальной диагональной формы* (5) *определяются по формулам*

$$E_k(\lambda) = \frac{D_k(\lambda)}{D_{k-1}(\lambda)} \quad (k = 1, 2, \ldots, r),$$

а

$$E_{r+1}(\lambda) = E_{r+2}(\lambda) = \ldots = E_n(\lambda) = 0.$$

Д о к а з а т е л ь с т в о. Мы показали, что при элементарных преобразованиях многочлены $D_k(\lambda)$ не меняются. Поэтому, если матрица $A(\lambda)$ эквивалентна диагональной нормальной матрице (5), то $D_k(\lambda)$ у них совпадают. Так как для матрицы (5) мы получили, что

$$D_k(\lambda) = E_1(\lambda) \ldots E_k(\lambda) \quad (k = 1, 2, \ldots, r;\ r \leqslant n)$$

и что $D_{r+1}(\lambda) = D_{r+2}(\lambda) = \ldots = D_n(\lambda) = 0$, то теорема доказана.

Следствие. *Для того чтобы две λ-матрицы $A(\lambda)$ и $B(\lambda)$ были эквивалентны, необходимо и достаточно, чтобы для них совпадали многочлены $D_1(\lambda)$, $D_2(\lambda), \ldots, D_n(\lambda)$.*

Действительно, если многочлены $D_k(\lambda)$ у $A(\lambda)$ и $B(\lambda)$ совпадают, то эти матрицы эквивалентны одной и той же нормальной диагональной λ-матрице и, следовательно, эквивалентны между собой.

3. Назовем λ-матрицу $P(\lambda)$ обратимой, если матрица $[P(\lambda)]^{-1}$ также есть λ-матрица. Если $\operatorname{Det} P(\lambda)$ равен постоянной, отличной от нуля, то $P(\lambda)$ обратима. Действительно, элементы обратной матрицы равны минорам $(n-1)$-го порядка, деленным на $\operatorname{Det} P(\lambda)$, т.е. в нашем случае они будут многочленами от λ и, значит $[P(\lambda)]^{-1}$, будет λ-матрицей.

Обратно, если $P(\lambda)$ обратима, то $\operatorname{Det} P(\lambda) = \text{const} \neq \neq 0$. В самом деле, пусть $[P(\lambda)]^{-1} = P_1(\lambda)$. Тогда $\operatorname{Det} P(\lambda) \operatorname{Det} P_1(\lambda) = 1$, а произведение двух многочленов может быть тождественно равно единице лишь в том случае, если многочлены суть отличные от нуля постоянные. Таким образом, мы показали, что *λ-матрица обратима тогда и только тогда, когда ее определитель есть постоянная, отличная от нуля.*

Все обратимые матрицы эквивалентны единичной матрице. В самом деле, определитель обратимой матрицы равен постоянной, отличной от нуля, и значит, $D_n(\lambda) = 1$. Так как $D_n(\lambda)$ делится на $D_k(\lambda)$, то и $D_k(\lambda) = 1$ $(k = 1, 2, \ldots, n)$. Поэтому все инвариантные множители $E_k(\lambda)$ обратимой матрицы равны 1, и нормальная диагональная форма для них будет совпадать с единичной матрицей.

Теорема 3. *Для того чтобы λ-матрицы $A(\lambda)$ и $B(\lambda)$ были эквивалентны между собой, необходимо* *и достаточно, чтобы существовали обратимые λ-матрицы $P(\lambda)$ и $Q(\lambda)$ такие, что*

$$A(\lambda) = P(\lambda)B(\lambda)Q(\lambda). \tag{7}$$

Доказательство. Докажем сначала, что если матрицы $A(\lambda)$ и $B(\lambda)$ эквивалентны, то можно подобрать обратимые матрицы $P(\lambda)$ и $Q(\lambda)$ так, чтобы выполнялось равенство (7). Для этого заметим, что каждое элементарное преобразование λ-матрицы $A(\lambda)$ можно осуществить, умножая $A(\lambda)$ слева или справа на некоторую обратимую λ-матрицу — матрицу этого элементарного преобразования.

Покажем это для всех трех типов элементарных преобразований. Пусть дана λ-матрица

$$A(\lambda) = \begin{pmatrix} a_{11}(\lambda) & a_{12}(\lambda) & \dots & a_{1n}(\lambda) \\ a_{21}(\lambda) & a_{22}(\lambda) & \dots & a_{2n}(\lambda) \\ \dots & \dots & \dots & \dots \\ a_{n1}(\lambda) & a_{n2}(\lambda) & \dots & a_{nn}(\lambda) \end{pmatrix}.$$

Чтобы поменять местами, например, первый и второй столбцы (соответственно строки) этой матрицы, надо умножить $A(\lambda)$ справа (соответственно слева) на матрицу

$$\begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}, \tag{8}$$

полученную из единичной перестановкой тех же столбцов (или, что все равно, строк).

Чтобы умножить второй столбец (соответственно строку) матрицы $A(\lambda)$ на число α, нужно умножить

$A(\lambda)$ справа (соответственно слева) на матрицу

$$\begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & \alpha & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix}, \tag{9}$$

полученную из единичной также умножением на α второго столбца (или, что все равно, второй строки).

Наконец, чтобы прибавить к первому столбцу $A(\lambda)$ второй, умноженный на $\varphi(\lambda)$, надо умножить $A(\lambda)$ справа на матрицу

$$\begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ \varphi(\lambda) & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix}, \tag{10}$$

полученную с помощью той же операции из единичной, а чтобы прибавить к первой строке вторую, умноженную на $\varphi(\lambda)$, нужно умножить $A(\lambda)$ слева на матрицу

$$\begin{pmatrix} 1 & \varphi(\lambda) & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix}, \tag{11}$$

которая также получается из единичной с помощью соответствующего элементарного преобразования.

Мы видим, таким образом, что матрицы элементарных преобразований — это матрицы, полученные одним элементарным преобразованием из E, причем, чтобы произвести элементарное преобразование над столбцами, $A(\lambda)$ надо умножать на матрицу преобразования справа, а чтобы преобразовать строки, $A(\lambda)$ надо умножать на соответствующую матрицу слева.

Можно сосчитать определитель каждой из приведенных матриц (8)–(11) и, таким образом, проверить, что он равен отличной от нуля постоянной; следовательно, все эти матрицы обратимы. Так как определитель произведения матриц равен произведению определителей, то и произведение матриц элементарных преобразований есть обратимая матрица.

Так как мы предположили, что $A(\lambda)$ и $B(\lambda)$ эквивалентны, то $A(\lambda)$ можно получить, применяя к $B(\lambda)$ некоторую цепочку элементарных преобразований. Каждое элементарное преобразование можно осуществить, умножая $B(\lambda)$ на обратимую λ-матрицу; следовательно, весь переход от $B(\lambda)$ к $A(\lambda)$ можно получить, умножая $B(\lambda)$ последовательно на некоторую совокупность обратимых λ-матриц слева и аналогично на некоторую совокупность справа. Так как произведение обратимых матриц также есть обратимая матрица, то первая часть теоремы тем самым доказана.

Отсюда следует, что всякая обратимая матрица есть произведение матриц элементарных преобразований. Действительно, всякая обратимая матрица $Q(\lambda)$ эквивалентна единичной матрице и поэтому может быть представлена в виде

$$Q(\lambda) = P_1(\lambda) E P_2(\lambda),$$

где $P_1(\lambda)$ и $P_2(\lambda)$ — произведения матриц элементарных преобразований. Но это значит, что и сама $Q(\lambda) = P_1(\lambda)P_2(\lambda)$ есть произведение матриц элементарных преобразований.

Этим замечанием можно воспользоваться для доказательства второй половины теоремы. Действительно, пусть дано, что

$$A(\lambda) = P(\lambda)B(\lambda)Q(\lambda),$$

где $P(\lambda)$ и $Q(\lambda)$ обратимы. Но, согласно только что сделанному замечанию, умножение слева на $P(\lambda)$ и справа на $Q(\lambda)$ эквивалентно некоторой совокупности элементарных преобразований, произведенных над $B(\lambda)$. Таким образом, $A(\lambda)$ эквивалентна $B(\lambda)$, что и требовалось доказать.

4. *) В этом пункте мы будем заниматься λ-матрицами вида $A-\lambda E$, где A — постоянная матрица. Основной вопрос, который будет решен, это вопрос об эквивалентности λ-матриц первой степени $A-\lambda E$ и $B-\lambda E$ **).

Легко видеть, что если матрицы A и B подобны, т. е. существует такая невырожденная постоянная матрица C, что $B = C^{-1}AC$, то λ-матрицы $A-\lambda E$ и $B-\lambda E$ эквивалентны. Действительно, если

$$B = C^{-1}AC,$$

то

$$B - \lambda E = C^{-1}(A - \lambda E)C.$$

Так как постоянная невырожденная матрица есть частный случай обратимой λ-матрицы, то, по теореме 3, из этого равенства следует эквивалентность $A - \lambda E$ и $B - \lambda E$.

Мы покажем позднее и обратное, что из эквивалентности λ-матриц $A-\lambda E$ и $B-\lambda E$ следует подобие матриц A и B. Отсюда мы получим, в частности, новое доказательство того, что всякая матрица подобна матрице, имеющей нормальную жорданову форму.

---

*) Этот пункт можно пропустить, так как он содержит другое, независимое от §§ 19 и 20 доказательство того, что всякую матрицу можно привести к жордановой форме.

**) Произвольная λ-матрица первой степени $A_0 + \lambda A_1$, у которой $\operatorname{Det} A_1 \neq 0$, эквивалентна некоторой матрице вида $A - \lambda E$. Действительно, в этом случае $A_0 + \lambda A_1 = -A_1(-A_1^{-1}A_0 - \lambda E)$ и, обозначая $-A_1^{-1}A_0$ через A, имеем $A_0 + \lambda A_1 = -A_1(A - \lambda E)$, откуда по теореме 3 следует эквивалентность матриц $A - \lambda E$ и $A_0 + \lambda A_1$.

Доказательству предпошлем лемму:

Л е м м а 2. *Произвольную λ-матрицу*

$$P(\lambda) = P_0\lambda^n + P_1\lambda^{n-1} + \ldots + P_n$$

можно разделить слева на матрицу вида $A - \lambda E$ (где A — любая постоянная матрица), т. е. можно найти такие матрицы $S(\lambda)$ и R (R постоянна), что,

$$P(\lambda) = (A - \lambda E)S(\lambda) + R.$$

Процесс деления, с помощью которого доказывается лемма, отличается от обычного деления многочленов только тем, что при умножении нельзя изменять порядок сомножителей.

Пусть

$$P(\lambda) = P_0\lambda^n + P_1\lambda^{n-1} + \ldots + P_n,$$

где P_k — постоянные матрицы.

Легко видеть, что λ-матрица

$$P(\lambda) + (A - \lambda E)P_0\lambda^{n-1}$$

будет иметь степень не выше $n - 1$.

Если

$$P(\lambda) + (A - \lambda E)P_0\lambda^{n-1} = P'_0\lambda^{n-1} + P'_1\lambda^{n-2} + \ldots + P'_{n-1},$$

то аналогично многочлен

$$P(\lambda) + (A - \lambda E)P_0\lambda^{n-1} + (A - \lambda E)P'_0\lambda^{n-2}$$

есть многочлен степени не выше $n-2$. Продолжая этот процесс, мы придем к многочлену

$$P(\lambda) + (A - \lambda E)(P_0\lambda^{n-1} + P'_0\lambda^{n-2} + \ldots)$$

степени не выше нулевой, т. е. не зависящему от λ. Обозначив полученную постоянную матрицу через R, мы получим

$$P(\lambda) = (A - \lambda E)[-P_0\lambda^{n-1} - P'_0\lambda^{n-2} - \ldots] + R.$$

Если теперь обозначить многочлен в квадратных скобках через $S(\lambda)$, то мы будем иметь

$$P(\lambda) = (A - \lambda E)S(\lambda) + R,$$

т. е. лемма доказана.

Аналогично доказывается возможность деления справа, т. е. существование матриц $S_1(\lambda)$ и R_1 таких, что

$$P(\lambda) = S_1(\lambda)(A - \lambda E) + R_1.$$

Заметим кстати, что здесь, как и в обычной теореме Безу, можно утверждать, что

$$R = R_1 = P(A).$$

Теорема 4. *Для того чтобы λ-матрицы $A - \lambda E$ и $B - \lambda E$ были эквивалентны, необходимо и достаточно, чтобы матрицы A и B были подобны.*

Доказательство. Достаточность была доказана в начале этого пункта. Докажем необходимость. Нам надо доказать, что если λ-матрицы $A - \lambda E$ и $B - \lambda E$ эквивалентны, то матрицы A и B подобны. По теореме 3 существуют такие обратимые λ-матрицы $P(\lambda)$ и $Q(\lambda)$, что

$$B - \lambda E = P(\lambda)(A - \lambda E)Q(\lambda). \tag{12}$$

Покажем сначала, что в равенстве (12) $P(\lambda)$ и $Q(\lambda)$ можно заменить постоянными матрицами.

С этой целью разделим $P(\lambda)$ на $B - \lambda E$ слева, $Q(\lambda)$ — справа. Мы получим равенства

$$\begin{aligned} P(\lambda) &= (B - \lambda E)P_1(\lambda) + P_0, \\ Q(\lambda) &= Q_1(\lambda)(B - \lambda E) + Q_0, \end{aligned} \tag{13}$$

где P_0 и Q_0 — постоянные матрицы.

Подставим в формулу (12) выражение для $P(\lambda)$ и произведем умножение. Мы получим:

$$B - \lambda E = (B - \lambda E)P_1(\lambda)(A - \lambda E)Q(\lambda) + P_0(A - \lambda E)Q(\lambda).$$

Во второе слагаемое подставим выражение для $Q(\lambda)$, произведем умножение и перенесем слагаемое $P_0(A-\lambda E)Q_0$ в левую часть равенства. Мы получим:

$$B-\lambda E-P_0(A-\lambda E)Q_0=K(\lambda), \tag{14}$$

где

$$K(\lambda)=(B-\lambda E)P_1(\lambda)(A-\lambda E)Q(\lambda)+ \\ +P_0(A-\lambda E)Q(\lambda)(B-\lambda E). \tag{15}$$

Из равенства (13) следует, что $P_0 = P(\lambda)-(B-\lambda E)P_1(\lambda)$. Заменив этим выражением P_0 во втором слагаемом, получим:

$$K(\lambda)=(B-\lambda E)P_1(\lambda)(A-\lambda E)Q(\lambda)+ \\ +P(\lambda)(A-\lambda E)Q_1(\lambda)(B-\lambda E)- \\ -(B-\lambda E)P_1(\lambda)(A-\lambda E)Q_1(\lambda)(B-\lambda E). \tag{16}$$

Но из равенства (12) мы имеем

$$(A-\lambda E)Q(\lambda)=P^{-1}(\lambda)(B-\lambda E),$$
$$P(\lambda)(A-\lambda E)=(B-\lambda E)Q^{-1}(\lambda).$$

Пользуясь этими равенствами, мы можем ввести множитель $B-\lambda E$ в конец первого и начало второго слагаемого в выражении для $K(\lambda)$, после чего получим окончательно

$$K(\lambda)=(B-\lambda E)[P_1(\lambda)P^{-1}(\lambda)+Q^{-1}(\lambda)Q_1(\lambda)- \\ -P_1(\lambda)(A-\lambda E)Q_1(\lambda)](B-\lambda E).$$

Докажем теперь, что $K(\lambda)=0$. Выражение в квадратных скобках, в силу обратимости $P(\lambda)$ и $Q(\lambda)$, есть многочлен относительно λ. Докажем, что он равен нулю. Предположим, что этот многочлен отличен от нуля и имеет степень m. Нетрудно убедиться тогда, что $K(\lambda)$ имеет степень $m+2$ и так как $m \geqslant 0$, является многочленом не ниже второй степени. Но из равенства (14) следует, что $K(\lambda)$ не выше первой степени. Следовательно, выражение в квадратных скобках, а значит, и $K(\lambda) = 0$.

Мы получили таким образом, что

$$B - \lambda E = P_0(A - \lambda E)Q_0, \tag{17}$$

где P_0 и Q_0 — постоянные матрицы, т. е. в равенстве (12) можно матрицы $P(\lambda), Q(\lambda)$ заменить постоянными матрицами.

Сравнивая коэффициенты при первой степени λ в обеих частях равенства (17), мы получаем

$$P_0 Q_0 = E,$$

откуда следует невырожденность каждой из матриц P_0 и Q_0 и равенство

$$P_0 = Q_0^{-1}.$$

Сравнение свободных членов дает

$$B = P_0 A Q_0 = Q_0^{-1} A Q_0,$$

т. е. B и A подобны. Теорема доказана.

Так как условием эквивалентности $A - \lambda E$ и $B - \lambda E$ служит совпадение их инвариантных множителей, то из доказанной теоремы следует, что, *для того чтобы матрицы A и B были подобны, необходимо и достаточно, чтобы инвариантные множители у $A - \lambda E$ и $B - \lambda E$ совпадали между собой.* Покажем теперь, что *всякая матрица A подобна матрице, имеющей жорданову нормальную форму.*

Для этого рассмотрим матрицу $A - \lambda E$ и найдем ее инвариантные множители. По этим инвариантным множителям построим, как было указано в § 21, матрицу B, имеющую жорданову нормальную форму. Тогда $B - \lambda E$ имеет те же инвариантные множители, что и $A - \lambda E$, и, значит, B подобна A.

Как было указано на стр. 254 (сноска), изложенное в п. 4 является другим, заменяющим §§ 19 и 20, доказательством того, что всякая матрица подобна матрице, имеющей жорданову нормальную форму. С другой стороны, конечно, содержание п. 4 может быть непосредственно выведено из содержания §§ 19 или 20 и 21.