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

$$\boldsymbol{B}_1=\begin{pmatrix} \lambda_1-\lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda_1-\lambda & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & \lambda_1-\lambda \end{pmatrix}.$$

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

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

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 $\lambda-\lambda_1$, $\lambda-\lambda_2$, ... 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

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

где $t_1+t_2+\dots+t_k=n-1$, $\Delta_{t_i}^{(i)}$ — миноры порядка t_i матрицы $\boldsymbol{B}_i$. Так как сумма порядков миноров $\Delta_{t_1}^{(1)}$, ... 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. 200). 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, $D_{n-p}(\lambda)$, $D_{n-p-1}(\lambda)$, ..., $D_1(\lambda)$ вовсе не содержат $\lambda-\lambda_1$.

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

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

При этом, начиная с $D_{n-p}(\lambda)$, множитель $(\lambda-\lambda_1)\ldots$ заменяется единицей, начиная с $D_{n-q}(\lambda)$, множитель $(\lambda-\lambda_2)\ldots$ 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 $D_i(\lambda)$ имеют вид

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

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*

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

Мы видим, что задание последовательности инвариантных множителей $E_n(\lambda)$, $E_{n-1}(\lambda)$, ... 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 $E_n(\lambda)$, $E_{n-1}(\lambda)$, ...

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 уравнения на $e^{\lambda x}$. Мы 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

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

приводится 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.

---

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

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. *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 экви-*

валентную $A(\lambda)$ 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, указанным 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, ..., 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:

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

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 коэффици-

ент 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:

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

Замечание. 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.

Упражнение. 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:

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

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

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 $A-\lambda E$.

Теорема 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 $D_{r+1}(\lambda)=\ldots$ $\ldots=D_n(\lambda)=0$, то элементы нормальной диагональной формы* (5) *определяются по формулам*

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

а

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

Доказательство. 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 $D_1(\lambda)$, $D_2(\lambda)$, ..., $D_n(\lambda)$.*

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 $\operatorname{Det} P(\lambda)=\text{const}\neq 0$. 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

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

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

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

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

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

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

---

*) Этот пункт можно пропустить, так как он содержит другое, не зависимое 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 $A_0 + \lambda A_1$.

доказательство того, что всякая матрица подобна матрице, имеющей нормальную жорданову форму.

Доказательству предпошлем лемму:

Лемма. *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

$$\begin{aligned} K(\lambda)=(B-\lambda E)P_1(\lambda)(A-\lambda E)Q(\lambda)+\\ +P_0(A-\lambda E)Q_1(\lambda)(B-\lambda E). \end{aligned} \tag{15}$$

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

получим:

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

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

$$B-\lambda E=P_0(A-\lambda E)Q_0, \quad (17)$$

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

$$P_0Q_0=E,$$

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 подобно A.

Как было указано на стр. 216 (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.

ГЛАВА IV

ПОНЯТИЕ О ТЕНЗОРАХ

§ 23. Сопряженное (двойственное) пространство

1. Определение сопряженного пространства. Пусть R — линейное пространство. Одновременно с R часто рассматривают другое, тесно связанное с ним пространство, так называемое сопряженное пространство. Для того чтобы сформулировать определение сопряженного пространства, вернемся к понятию линейной функции, введенному нами в п. 1 § 4.

Линейной функцией мы назвали функцию $f(x)$, $x \in R$, удовлетворяющую следующим условиям:

1° $f(x+y)=f(x)+f(y)$,

2° $f(\lambda x)=\lambda f(x)$.

Пусть $e_1, e_2, \ldots, e_n$ — базис в n-мерном пространстве R. Если

$$x=\xi^1 e_1+\xi^2 e_2+\ldots+\xi^n e_n$$

— вектор из R, то линейная функция в R может быть записана в виде (см. § 4)

$$\begin{aligned} f(x)&=f(\xi^1 e_1+\xi^2 e_2+\ldots+\xi^n e_n)= \\ &=a_1\xi^1+a_2\xi^2+\ldots+a_n\xi^n, \end{aligned} \tag{1}$$

где коэффициенты $a_1, a_2, \ldots, a_n$, определяющие линейную функцию, вычисляются по формулам

$$a_1=f(e_1),\ a_2=f(e_2), \ldots,\ a_n=f(e_n). \tag{2}$$

Как это ясно из формулы (1), *при заданном базисе* $e_1, e_2, \ldots, e_n$ *всяким* n *числам* $a_1, a_2, \ldots, a_n$ *отвечает линейная функция, притом только одна.*

Пусть f и g — линейные функции. Их суммой называется функция h, ставящая в соответствие каждому вектору x число $f(x)+g(x)$. Произведением линейной функции f на число α называется функция, ставящая в соответствие каждому вектору x число $\alpha f(x)$.

Очевидно, что сумма линейных функций и произведение линейной функции на число есть снова линейная функция. При этом, если линейная функция f задается числами $a_1, a_2, \ldots, a_n$, а g — числами $b_1, b_2, \ldots, b_n$, то $f+g$ задается числами $a_1+b_1, a_2+b_2, \ldots, a_n+b_n$, а αf — числами $\alpha a_1, \alpha a_2, \ldots, \alpha a_n$.

Таким образом, множество заданных в R линейных функций образует линейное пространство.

Определение 1. *Пусть R есть n-мерное пространство. Пространством R', сопряженным к R, мы назовем линейное пространство, векторами которого являются линейные функции, заданные в R. Сумма в R' определяется как сумма линейных функций, а произведение вектора из R' на число — как произведение линейной функции на число.*

Так как при заданном базисе $e_1, e_2, \ldots, e_n$ в пространстве R каждая линейная функция однозначно задается системой n чисел $a_1, a_2, \ldots, a_n$, причем сумме функций отвечает сумма чисел, произведению функции на α произведение чисел a_i на α, то ясно, что R' изоморфно пространству, в котором вектор определен как совокупность n чисел.

Значит, *пространство R', сопряженное к n-мерному пространству R, также n-мерно.*

Если пространства R и R' рассматривают одновременно, то векторы из R называются *контравариантными*, а векторы из R' *ковариантными*. В дальнейшем символы $x, y, \ldots$ будут означать элементы из R, т. е. контравариантные векторы, а $f, g, \ldots$ — элементы из R', т. е. ковариантные векторы.

2. Биортогональные (взаимные) базисы. В дальнейшем мы будем значение линейной функции f в точке x обозначать через (f, x). Таким образом каждой паре $f \in R'$ и $x \in R$ отнесено число (f, x), причем

1° $(f, x_1+x_2)=(f, x_1)+(f, x_2)$,

2° $(f, \lambda x)=\lambda(f, x)$,

3° $(\lambda f, x) = \lambda (f, x)$,
4° $(f_1 + f_2, x) = (f_1, x) + (f_2, x)$.

Первое и второе из этих соотношений — это записанные в новых обозначениях равенства

$$f(x_1 + x_2) = f(x_1) + f(x_2) \text{ и } f(\lambda x) = \lambda f(x),$$

являющиеся определением линейной функции, а третье и четвертое — определения произведения линейной функции на число и суммы линейных функций. Соотношения 1° — 4° напоминают по внешнему виду аксиомы 2° и 3° скалярного произведения (§ 2). Надо лишь подчеркнуть, что в то время, как скалярное произведение есть число, отнесенное паре векторов одного и того же (евклидова) пространства, (f, x) есть число, отнесенное паре векторов, один из которых принадлежит аффинному пространству R, а другой — аффинному пространству R'.

Векторы $x \in R$ и $f \in R'$ мы назовем *ортогональными*, если

$$(f, x) = 0.$$

Таким образом, хотя в аффинном пространстве R (в отличие от евклидова) нет понятия ортогональности двух векторов $x, y \in R$, можно говорить об ортогональности векторов из R к векторам из R'.

Определение 2. *Пусть $e_1, e_2, \ldots, e_n$ — базис в R, а $f^1, f^2, \ldots, f^n$ — базис в R'. Мы назовем эти базисы биортогональными (взаимными), если*

$$(f^i, e_k) = \begin{cases} 1 \text{ при } i = k \\ 0 \text{ при } i \neq k \end{cases} \quad (i, k = 1, 2, \ldots, n). \tag{3}$$

Введем символ δ^i_k, положив

$$\delta^i_k = \begin{cases} 1 \text{ при } i = k \\ 0 \text{ при } i \neq k \end{cases} \quad (i, k = 1, 2, \ldots, n).$$

Тогда

$$(f^i, e_k) = \delta^i_k.$$

Если $e_1, e_2, \ldots, e_n$ — базис в R, то (f, e_k) являются числами a_k, определяющими линейную функцию $f \in R'$ [см. формулу (2)], так как (f, e_k) есть другая форма записи выражения $f(e_k)$.

Из этого замечания следует утверждение:

если $e_1, e_2 \ldots, e_n$ — произвольный базис в R, то в R' существует, и притом только один, базис $f^1, f^2, \ldots, f^n$ такой, что базисы $e_1 e_2, \ldots, e_n$ и $f^1, f^2, \ldots, f^n$ биортогональны (взаимны).

Действительно, из равенства (3) имеем

$$(f^1, e_1) = 1, \ (f^1, e_2) = 0, \ldots, (f^1, e_n) = 0.$$

Таким образом, здесь заданы числа $a_1 = 1$, $a_2 = 0, \ldots, a_n = 0$. Так как по всяким числам a_i можно построить единственную линейную функцию, то f^1 определено, и при этом однозначно. Аналогично определяется f^2 равенствами

$$(f^2, e_1) = 0, \ (f^2, e_2) = 1, \ldots, (f^2, e_n) = 0$$

и т. д. Построенные векторы $f^1, f^2, \ldots, f^n$ из R' (линейные функции) линейно независимы, так как отвечающие каждому из них системы чисел $a_1, a_2, \ldots, a_n$ линейно независимы между собой. Мы построили, таким образом, базис, биортогональный базису $e_1, e_2, \ldots, e_n$, и доказали его единственность.

В дальнейшем мы будем пользоваться принятыми в тензорном исчислении обозначениями, а именно, если в некотором выражении один и тот же индекс стоит один раз вверху, а другой раз внизу, то это означает, что по этому индексу производится суммирование (от 1 до n). Сам знак суммирования Σ мы при этом будем опускать.

Например, $\xi^i \eta_i$ означает $\xi^1 \eta_1 + \xi^2 \eta_2 + \ldots + \xi^n \eta_n$.

Имея в R и R' биортогональные базисы, легко вычислять координаты любого вектора. Пусть e_i и f^k — биортогональные базисы. Найдем координаты ξ^i вектора $x \in R$ в базисе e_i. Мы имеем

$$x = \xi^i e_i.$$

Отсюда

$$(f^k, x) = (f^k, \xi^i e_i) = \xi^i (f^k, e_i) = \xi^i \delta_i^k = \xi^k.$$

Следовательно, *координаты ξ^k вектора x в базисе $e_1, e_2, \ldots, e_n$ вычисляются по формулам*

$$\xi^k = (f^k, x),$$

где f^k — базис, взаимный с базисом e_i.

Аналогично получаем, что *координаты* η_i *вектора* f *в базисе* f^k *вычисляются по формулам*

$$\eta_i = (f, e_i).$$

Пусть $e_1, e_2, \ldots, e_n$ и $f^1, f^2, \ldots, f^n$ — два взаимных (биортогональных) базиса. Выразим величину (f, x) через координаты векторов f и x в базисах $e_1, e_2, \ldots, e_n$ и $f^1, f^2, \ldots, f^n$ соответственно. Пусть

$$x = \xi^1 e_1 + \xi^2 e_2 + \ldots + \xi^n e_n \quad \text{и} \quad f = \eta_1 f^1 + \eta_2 f^2 + \ldots + \eta_n f^n;$$

тогда

$$(f, x) = (\eta_1 f^1 + \eta_2 f^2 + \ldots + \eta_n f^n, \ \xi^1 e_1 + \xi^2 e_2 + \ldots + \xi^n e_n) = \\ = (f^i, e_k)\, \eta_i \xi^k = \delta^i_k \eta_i \xi^k = \eta_i \xi^i.$$

Итак, *если* $e_1, e_2, \ldots, e_n$ — *базис в* R, $f^1, f^2, \ldots, f^n$ — *взаимный с ним базис в* R', *то*

$$(f, x) = \eta_1 \xi^1 + \eta_2 \xi^2 + \ldots + \eta_n \xi^n, \tag{4}$$

где $\xi^1, \xi^2, \ldots, \xi^n$ — *координаты вектора* $x \in R$ *в базисе* $e_1, e_2, \ldots, e_n$, *а* $\eta_1, \eta_2, \ldots, \eta_n$ — *координаты вектора* $f \in R'$ *в базисе* $f^1, f^2, \ldots, f^n$.

Замечание. Если $e_1, e_2, \ldots, e_n$ и $f^1, f^2, \ldots, f^n$ — произвольные базисы в R и R' соответственно, то

$$(f, x) = a^i_k \eta_i \xi^k,$$

где $a^i_k = (f^i, e_k)$.

Мы видим, что во взаимных базисах значение (f, x) записывается особенно просто.

Итак, мы построили соответствие, относящее каждому линейному пространству R другое пространство, а именно сопряженное пространство R'. Мы можем теперь установить соответствие и между линейными преобразованиями пространств.

Пусть R_1, R_2 — два линейных пространства и R_1', R_2' — пространства, им сопряженные. Каждому линейному преобразованию A пространства R_1 в R_2 мы поставим в соответствие линейное преобразование A' пространства R_2' в R_1', которое определим следующим образом.

Пусть $f_2 \in R_2'$, $x_1 \in R_1$. Рассмотрим (f_2, Ax_1); при фиксированном f_2 это линейная функция от x_1, т. е. может быть записана в виде $(f_2, Ax_1) = (f_1, x_1)$, где $f_1 \in R_1'$. Положим по определению $f_1 = A' f_2$. Получаемое преобразование A' называется *сопряженным* к A. Итак, если A — линейное преобразование пространства R_1 в R_2, то

сопряженное ему преобразование есть линейное преобразование A' пространства R_2' в R_1', задаваемое тождеством

$$(A'f_2,\ x_1)=(f_2,\ Ax_1).$$

Установим одно важное свойство операции перехода к сопряженному преобразованию. Пусть A — линейное преобразование пространства R_1 в R_2, B — линейное преобразование пространства R_2 в R_3. Обозначим через BA композицию этих преобразований, т. е. линейное преобразование пространства R_1 в R_3 (по определению $BAx=B(Ax)$ для любого $x\in R_1$).

Покажем, что

$$(BA)'=A'B'.$$

В самом деле, согласно определению имеем:

$$((BA)'f,\ x)=(f,\ BAx) \text{ для любых } x\in R_1 \text{ и } f\in R_3'.$$

С другой стороны, $(A'B'f,\ x)=(B'f,\ Ax)=(f,\ BAx)$. Сопоставляя эти равенства, мы видим, что $(BA)'=A'B'$.

Упражнение. Доказать, что линейное преобразование, сопряженное к A', есть A.

3. Взаимозаменяемость R и R'. В предыдущем изложении R и R' играли различную роль. Мы покажем, что они совершенно равноправны, т. е. что все теоремы останутся справедливыми, если мы поменяем R и R' ролями.

Мы определили R' как совокупность линейных функций в R. Чтобы установить равноправность R и R', докажем, что всякая линейная функция $\varphi(f)$ в R' может быть записана в виде $(f,\ x_0)$, где x_0 — фиксированный вектор из R.

Пусть $e_1, e_2, \ldots, e_n$ — некоторый базис в R и $f^1, f^2, \ldots, f^n$ — взаимный с ним базис в R'. Линейная функция $\varphi(f)$ может быть записана в виде

$$\varphi(f)=a^1\eta_1+a^2\eta_2+\ldots+a^n\eta_n,$$

где $\eta_1, \eta_2, \ldots, \eta_n$ — координаты вектора f в базисе $f^1, f^2, \ldots, f^n$. Рассмотрим вектор x_0, имеющий в базисе $e_1, e_2, \ldots, e_n$ координаты $a^1, a^2, \ldots, a^n$. Тогда, как мы видели в п. 2,

$$(f,\ x_0)=a^1\eta_1+a^2\eta_2+\ldots+a^n\eta_n$$

и, следовательно,

$$\varphi(f)\equiv(f,\ x_0). \tag{5}$$

Эта формула устанавливает взаимно однозначное соответствие между линейными функциями φ, заданными в R', и векторами $x_0 \in R$.

Мы можем поэтому во всем изложении считать R пространством линейных функций над R', задавая эти линейные функции формулой (5). Этим установлено полное равноправие между R и R'.

Заметим, что при одновременном изучении пространства и сопряженного пространства мы употребляем лишь обычные для векторов операции сложения и умножения на число в каждом пространстве и операцию (f, x), связывающую элементы обоих пространств. Можно поэтому дать другое определение пары сопряженных пространств R и R', при котором их равноправие непосредственно видно. Это определение состоит в следующем: мы рассматриваем пару n-мерных пространств R и R' и каждой паре векторов $x \in R$, $f \in R'$ относим число (f, x), требуя при этом, чтобы выполнялись условия 1°—4° предыдущего пункта и условие

5° *Из* $(f, x)=0$ *для любого* x *следует* $f=0$ *и из* $(f, x)=0$ *для любого* f *следует* $x=0$.

Коротко говоря, пара сопряженных пространств R и R' — это пара n-мерных пространств с введенной дополнительно операцией (f, x), удовлетворяющей перечисленным условиям.

Замечание. В п. 2 мы доказали, что для каждого базиса в R существует и притом единственный взаимный с ним базис в R'. Из равноправия между R и R' следует, что для всякого базиса в R' существует и притом единственный взаимный с ним базис в R.

4. Преобразования координат в R и R'. Если мы рассматриваем координаты векторов $x \in R$ в некотором базисе $e_1, e_2, \ldots, e_n$, то координаты векторов $f \in R'$ мы будем, как правило, рассматривать в базисе $f^1, f^2, \ldots, f^n$, взаимном к базису $e_1, e_2, \ldots, e_n$. Перейдем в R от базиса $e_1, e_2, \ldots, e_n$ к новому базису $e'_1, e'_2, \ldots, e'_n$, и пусть

$$e'_i = c_i^k e_k \tag{6}$$

— формулы этого перехода.

Обозначая через $f^1, f^2, \ldots, f^n$ базис, взаимный с базисом $e_1, e_2, \ldots, e_n$, а через $f'^1, f'^2, \ldots, f'^n$ — базис, взаимный с базисом $e'_1, e'_2, \ldots, e'_n$, найдем матрицу $\|b_i^k\|$ перехода от базиса f^i к базису f'^i.

Найдем сначала обратную ей матрицу $\| u_i^k \|$ перехода от $f'^1, f'^2, \ldots, f'^n$ к $f^1, f^2, \ldots, f^n$:

$$f^k = u_i^k f'^i. \tag{6'}$$

Для этого вычислим двумя способами выражение (f^k, e_i'):

$$(f^k, e_i') = (f^k, c_i^\alpha e_\alpha) = c_i^\alpha (f^k, e_\alpha) = c_i^k,$$
$$(f^k, e_i') = (u_i^k f'^i, e_i') = u_i^k.$$

Отсюда имеем $c_i^k = u_i^k$, т. е. матрица $\| u_i^k \|$ является транспонированной *) к матрице перехода (6). Следовательно, *матрица перехода*

$$f'^k = b_i^k f^i \tag{7}$$

от $f^1, f^2, \ldots, f^n$ *к* $f'^1, f'^2, \ldots, f'^n$ *равна матрице, транспонированной к матрице, обратной матрице* $\| c_i^k \|$ *перехода от* $e_1, e_2, \ldots, e_n$ *к* $e_1', e_2', \ldots, e_n'$.

Выясним теперь, как преобразуются координаты векторов в R и в R'. Пусть ξ^i — координаты вектора $x \in R$ в базисе $e_1, e_2, \ldots, e_n$ и ξ'^i — его координаты в новом базисе $e_1', e_2', \ldots, e_n'$.

Тогда

$$(f^i, x) = (f^i, \xi^1 e_1 + \xi^2 e_2 + \ldots + \xi^n e_n) = \xi^i$$

и

$$(f'^i, x) = (f'^i, \xi'^1 e_1' + \xi'^2 e_2' + \ldots + \xi'^n e_n') = \xi'^i.$$

Поэтому

$$\xi'^i = (f'^i, x) = (b_k^i f^k, x) = b_k^i (f^k, x) = b_k^i \xi^k.$$

Итак,

$$\xi'^i = b_k^i \xi^k, \tag{8}$$

т. е. координаты векторов в R преобразуются по тем же формулам, что и векторы взаимного базиса в R'. Аналогично, координаты векторов в R' преобразуются по тем

*) Мы говорим, что матрица $\| u_i^k \|$ является транспонированной к матрице перехода (6), так как суммирование в (6') производится по другому индексу.

же формулам, что и векторы взаимного базиса в R, т. е.

$$\eta'_i = c_i^k \eta_k. \tag{9}$$

Мы можем, таким образом, сформулировать следующее правило: *при переходе от старой системы координат к новой объекты, имеющие нижний индекс, преобразуются матрицей* $\| c_i^k \|$, *объекты, имеющие верхний индекс, преобразуются матрицей* $\| b_i^k \|$, *обратной к* $\| c_i^k \|$.

Тот факт, что матрица $\| b_i^k \|$ является обратной к матрице $\| c_i^k \|$, выражается соотношениями

$$c_i^\alpha b_\alpha^j = \delta_i^j, \quad b_i^\alpha c_\alpha^j = \delta_i^j.$$

5. Пространство, сопряженное к евклидову. Ограничимся для простоты евклидовым пространством над полем действительных чисел.

Лемма. *Пусть R есть n-мерное евклидово пространство. Тогда каждую линейную функцию в нем можно записать в виде*

$$f(x) = (x, y),$$

где y — фиксированный вектор, однозначно определяемый линейной функцией f. Обратно, каждый вектор y определяет линейную функцию $f(x) = (x, y)$.

Доказательство. Выберем в R некоторый ортогональный нормированный базис $e_1, e_2, \ldots, e_n$. Линейная функция $f(x)$ в этом базисе может быть записана в виде

$$f(x) = a_1\xi^1 + a_2\xi^2 + \ldots + a_n\xi^n.$$

Введем вектор y с координатами $a_1, a_2, \ldots, a_n$. Так как базис $e_1, e_2, \ldots, e_n$ — ортогональный, то

$$(x, y) = a_1\xi^1 + a_2\xi^2 + \ldots + a_n\xi^n.$$

Мы доказали, таким образом, существование такого вектора y, что для любого x имеет место равенство

$$f(x) = (x, y).$$

Докажем теперь, что такой вектор определяется однозначно. Пусть

$$f(x) = (x, y_1) \quad \text{и} \quad f(x) = (x, y_2).$$

Тогда

$$(x, y_1) = (x, y_2),$$

т. е.

$$(x,\ y_1 - y_2) = 0$$

для любого x. Следовательно, $y_1 - y_2 = 0$. Однозначность доказана.

Таким образом, в случае евклидова пространства мы можем каждый элемент f из R' заменить соответствующим элементом y из R и при этом вместо $(f,\ x)$ писать $(y,\ x)$. *Так как при одновременном изучении пространства и сопряженного пространства мы употребляем лишь обычные для векторов операции и операцию* $(f,\ x)$, *связывающую элементы* $f \in R'$ *и* $x \in R$, *то мы можем в случае евклидова пространства заменить* f *на* y, R' *на* R *и* $(f,\ x)$ *на* $(y,\ x)$, т. е. отождествить евклидово пространство с сопряженным к нему пространством R' *). Это выражают иногда и так: в евклидовом пространстве можно заменить ковариантные векторы контравариантными.

При таком отождествлении пространста R и сопряженного к нему пространства R' понятие ортогональности векторов $x \in R$ и $f \in R'$, введенное в пункте 2, переходит в обычное для евклидова пространства понятие ортогональности двух векторов из R.

Пусть $e_1,\ e_2,\ \ldots,\ e_n$ — произвольный базис в R, а $f^1,\ f^2,\ \ldots,\ f^n$ — взаимный с ним (биортогональный) базис в R'. Так как в случае евклидова пространства R и R' отождествлены, то мы можем считать векторы биортогонального к e_i базиса f^k также векторами из R.

Выясним, как найти в этом случае по базису $e_1, e_2, \ldots, e_n$ базис $f^1,\ f^2,\ \ldots,\ f^n$. Выразим сначала e_i через f^k:

$$e_i = g_{ik} f^k.$$

Нам нужно найти коэффициенты g_{ik}. Для этого умножим скалярно обе части равенства на e_α:

$$(e_i,\ e_\alpha) = g_{ik}\,(f^k,\ e_\alpha).$$

---

*) Если R — аффинное n-мерное пространство, то R' также n-мерно и, следовательно, R и R' изоморфны. Но если бы мы отождествили R и R', нам пришлось бы вместо $(f,\ x)$ писать $(y,\ x)$, где $y,\ x \in R$, т. е. мы тем самым ввели бы в R скалярное произведение.

Так как, в силу взаимности (биортогональности) базисов f^k и e_α,

$$(f^k, e_\alpha) = \delta^k_\alpha,$$

то

$$(e_i, e_\alpha) = g_{ik}\delta^k_\alpha = g_{i\alpha}.$$

Итак, если базис f^k биортогонален к базису e_i, то

$$e_i = g_{ik}f^k, \tag{10}$$

где матрица g_{ik} вычисляется по формуле

$$g_{ik} = (e_i, e_k).$$

Отсюда, разрешив соотношение (10) относительно f^i, имеем:

$$f^i = g^{ik}e_k, \tag{11}$$

где g^{ik} — матрица, обратная к g_{ik}, т. е.

$$g^{i\alpha}g_{\alpha k} = \delta^i_k.$$

Упражнение. Показать, что

$$g^{ik} = (f^i, f^k).$$

§ 24. Тензоры

1. Полилинейные функции. В первой главе мы изучили линейные и билинейные функции в n-мерном аффинном пространстве. Их естественным обобщением являются полилинейные функции, зависящие от произвольного числа векторов. При этом мы будем рассматривать функции, зависящие как от векторов из R, так и от векторов из R'.

Определение 1. *Полилинейной функцией*

$$l(x, y, \ldots; f, g, \ldots),$$

зависящей от p векторов $x, y, \ldots \in R$ и q векторов $f, g, \ldots \in R'$ (R' — пространство, сопряженное к R), называется функция, линейная относительно каждого из аргументов, когда остальные аргументы фиксированы.

Например, если зафиксированы все векторы, кроме первого, то

$$l(x' + x'', y, \ldots; f, g, \ldots) = \\ = l(x', y, \ldots; f, g, \ldots) + l(x'', y, \ldots; f, g, \ldots),$$

$$l(\lambda x, y, \ldots; f, g, \ldots) = \lambda l(x, y, \ldots; f, g, \ldots).$$

Аналогично

$$l(x, y, \ldots; f'+f'', g, \ldots) = \\ = l(x, y, \ldots; f', g, \ldots) + l(x, y, \ldots; f'', g, \ldots),$$
$$l(x, y, \ldots; \mu f, g, \ldots) = \mu l(x, y, \ldots; f, g, \ldots).$$

То же самое и для других аргументов.

Полилинейную функцию, зависящую от p векторов из R (контравариантных векторов) и q векторов из R' (ковариантных векторов) мы будем называть *полилинейной функцией типа* (p, q). Рассмотрим некоторые полилинейные функции.

Простейшие полилинейные функции — это функции типа (1, 0) и типа (0, 1).

Полилинейная функция типа (1, 0) — это линейная функция от одного вектора в пространстве R, т. е. вектор пространства R' (ковариантный вектор).

Аналогично, как это было показано в п. 3 предыдущего параграфа, полилинейная функция типа (0, 1) задает вектор из R (контравариантный вектор).

Полилинейные функции, зависящие от двух векторов (билинейные функции), бывают трех типов:

α) функции, зависящие от двух векторов из пространства R, — это введенные в § 4 билинейные функции в пространстве R;

β) функции, зависящие от двух векторов в пространстве R', — это билинейные функции в R';

γ) функции, зависящие от одного вектора из R и одного вектора из R'.

Функции третьего типа тесно связаны с линейными преобразованиями. Действительно, пусть

$$y = Ax$$

— линейное преобразование в R. Построим билинейную функцию

$$(f, Ax),$$

линейно зависящую от векторов $x \in R$ и $f \in R'$.

Мы можем, таким образом, каждому линейному преобразованию в R однозначно сопоставить билинейную функцию типа γ).

Как и в § 11 главы II, можно доказать и обратное, т. е. что каждой билинейной функции типа γ) отвечает линейное преобразование в R.

2. Выражения для полилинейной функции в данной системе координат. Переход от одной системы координат к другой. Выясним, как выражается полилинейная функция через координаты тех векторов, от которых она зависит. Для того чтобы не писать слишком длинных формул, проведем рассмотрение на случае полилинейной функции $l(x, y; f)$, зависящей от двух векторов из R и одного вектора из R' [функция типа (2, 1)].

Выберем в R некоторый базис $e_1, e_2, \ldots, e_n$, а в R' — взаимный с ним базис $f^1, f^2, \ldots, f^n$. Пусть

$$x = \xi^i e_i, \quad y = \eta^j e_j, \quad f = \zeta_k f^k.$$

Тогда

$$l(x, y; f) = l(\xi^i e_i, \eta^j e_j; \zeta_k f^k) = \xi^i \eta^j \zeta_k l(e_i, e_j; f^k).$$

Итак: при заданных в R и соответственно в R' базисах $e_1, e_2, \ldots, e_n$ и $f^1, f^2, \ldots, f^n$ полилинейная функция $l(x, y; f)$ записывается в виде

$$l(x, y; f) = a_{ij}^k \xi^i \eta^j \zeta_k,$$

где ξ^i, соответственно η^j, соответственно ζ_k — координаты вектора x, соответственно y, соответственно f. Числа a_{ij}^k, определяющие функцию $l(x, y; f)$, задаются формулой

$$a_{ij}^k = l(e_i, e_j; f^k)$$

и зависят, таким образом, от выбора базисов в R и R'.

Аналогичная формула имеет место для полилинейной функции общего вида:

$$l(x, y, \ldots; f, g, \ldots) = a_{ij\ldots}^{rs\ldots} \xi^i \eta^j \ldots \lambda_r \mu_s \ldots, \tag{1}$$

где числа $a_{ij\ldots}^{rs\ldots}$, определяющие полилинейную функцию, вычисляются по формулам

$$a_{ij\ldots}^{rs\ldots} = l(e_i, e_j, \ldots, f^r, f^s, \ldots). \tag{2}$$

Выясним теперь, как изменяется система чисел, определяющая полилинейную форму, при изменении базиса.

Пусть в R задан базис $e_1, e_2, \ldots, e_n$ и в R' — взаимный с ним базис $f^1, f^2, \ldots, f^n$. Перейдем к новому

базису $e'_1, e'_2, \ldots, e'_n$ в R и взаимному с ним базису $f'^1, f'^2, \ldots, f'^n$ в R'.

Пусть переход от базиса $e_1, e_2, \ldots, e_n$ к базису $e'_1, e'_2, \ldots, e'_n$ задается формулами

$$e'_\alpha = c^\beta_\alpha e_\beta. \tag{3}$$

Тогда переход от базиса $f^1, f^2, \ldots, f^n$ к базису $f'^1, f'^2, \ldots, f'^n$ задается формулами

$$f'^\beta = b^\beta_\alpha f^\alpha, \tag{4}$$

где $\| b^\beta_\alpha \|$ — матрица, транспонированная к матрице, обратной к $\| c^\beta_\alpha \|$.

Формула (3) показывает, что числа c^β_α при фиксированном α являются координатами вектора e'_α в базисе $e_1, e_2, \ldots, e_n$. Аналогично числа b^β_α при фиксированном β являются координатами вектора f'^β в базисе $f^1, f^2, \ldots, f^n$.

Найдем систему чисел $a'^{rs\ldots}_{ij\ldots}$, определяющих нашу полилинейную функцию в базисах $e'_1, e'_2, \ldots, e'_n$ и $f'^1, f'^2, \ldots, f'^n$. Мы знаем, что

$$a'^{rs\ldots}_{ij\ldots} = l\,(e'_i, e'_j, \ldots; f'^r, f'^s, \ldots).$$

Поэтому, чтобы найти $a'^{rs\ldots}_{ij\ldots}$ мы должны в формулу (1) вместо $\xi^i, \eta^j, \ldots; \lambda_r, \mu_s, \ldots$ подставить координаты векторов $e'_i, e'_j, \ldots; f'^r, f'^s, \ldots$, т. е. числа $c^\alpha_i, c^\beta_j, \ldots; b^r_\sigma, b^s_\tau, \ldots$ Мы получаем таким образом:

$$a'^{rs\ldots}_{ij\ldots} = c^\alpha_i c^\beta_j \ldots b^r_\sigma b^s_\tau \ldots a^{\sigma\tau\ldots}_{\alpha\beta\ldots}.$$

Итак, *система чисел $a^{rs\ldots}_{ij\ldots}$, определяющих полилинейную функцию $l\,(x, y, \ldots; f, g, \ldots)$ во взаимных базисах $e_1, e_2, \ldots, e_n$ и $f^1, f^2, \ldots, f^n$, при переходе к новым взаимным базисам $e'_1, e'_2, \ldots, e'_n$ и $f'^1, f'^2, \ldots, f'^n$ преобразуется по формулам*

$$a'^{rs\ldots}_{ij\ldots} = c^\alpha_i c^\beta_j \ldots b^r_\sigma b^s_\tau \ldots a^{\sigma\tau\ldots}_{\alpha\beta\ldots}, \tag{5}$$

где $\| c^j_i \|$ — матрица, определяющая преобразование базиса $e_1, e_2, \ldots, e_n$, а $\| b^j_i \|$ — матрица, определяющая преобразование взаимного с ним базиса $f^1, f^2, \ldots, f^n$.

Это можно выразить следующей фразой: *на нижние индексы системы чисел* $a_{ij\ldots}^{rs\ldots}$ *действует матрица* $\|c_i^j\|$, *на верхние индексы — матрица* $\|b_i^j\|$ (ср. § 23, п. 4, где рассмотрены формулы для преобразования координат ковариантного и контравариантного векторов).

3. Определение тензора. Объекты, с которыми мы встречались на протяжении этой книги (векторы, линейные функции, линейные преобразования, билинейные функции и т. д.), определялись в каждом базисе своей системой чисел. Например, вектор определялся в каждом базисе системой n чисел — своими координатами. Линейная функция определяется в каждом базисе также системой n чисел — своими коэффициентами. Линейное преобразование определяется в каждом базисе системой n^2 чисел — матрицей линейного преобразования. Билинейная функция определяется в каждом базисе системой n^2 чисел — матрицей этой билинейной формы. При переходе от одной системы координат (базиса) к другой система чисел, определяющая данный объект, преобразуется определенным образом, причем закон преобразования различен для различных объектов. Например, как вектор из R, так и линейная функция в R задаются системой n чисел, однако при переходе к другому базису они преобразуются по-разному. Для полной характеристики встречающейся величины мы должны задать не только значения соответствующих чисел в какой-либо системе координат, но и закон преобразования соответствующей совокупности чисел при переходе к другой системе координат.

В пунктах 1 и 2 этого параграфа мы ввели понятие полилинейной функции, которая определяется в каждом данном базисе системой n^k чисел (2), преобразующихся при переходе к другому базису по формулам (5). В связи с ним вводится следующее определение, играющее важную роль во многих разделах физики, геометрии и алгебры.

Определение 2. *Если каждой системе координат в n-мерном аффинном пространстве отнесена система* n^{p+q} *чисел* $a_{ij\ldots}^{rs\ldots}$ (число нижних индексов обозначено через p, верхних — через q), *причем при переходе от одной системы координат к другой эти числа преобразуются*

по формуле

$$a'^{rs\ldots}_{ij\ldots} = c^{\alpha}_{i} c^{\beta}_{j} \ldots b^{r}_{\sigma} b^{s}_{\tau} \ldots a^{\sigma\tau\ldots}_{\alpha\beta\ldots}, \tag{6}$$

где $\| c^{j}_{i} \|$ — *матрица, задающая переход от одного базиса в* R *к другому, а* $\| b^{j}_{i} \|$ — *матрица, транспонированная к матрице, обратной к* $\| c^{j}_{i} \|$, *то мы говорим, что нам задан тензор. Этот тензор называется* p *раз ковариантным и* q *раз контравариантным. Число* $p+q$ *называется рангом (валентностью) тензора. Сами числа* $a^{rs\ldots}_{ij\ldots}$ *называются компонентами тензора.*

Так как система чисел, определяющих полилинейную функцию от p векторов из R и q векторов из R', при изменении базиса преобразуется как раз по формуле (6), то каждой такой полилинейной функции однозначно соответствует тензор ранга $p+q$, p раз ковариантный и q раз контравариантный. Обратно, каждому тензору однозначно отвечает полилинейная функция. В дальнейшем свойства тензоров и операции над ними мы будем изучать на «модели» полилинейных функций, хотя, конечно, полилинейные функции являются лишь одной из возможных реализаций тензоров.

Приведем некоторые примеры тензоров.

1. Скаляр. Если каждой системе координат отнесено одно и то же фиксированное число a, то его формально можно также считать тензором, а именно — тензором нулевого ранга. Тензор нулевого ранга называется скаляром.

2. Контравариантный вектор. Вектору из R в каждом базисе соответствует совокупность n его координат, которые при переходе к другому базису преобразуются по формулам

$$\eta'^{i} = b^{i}_{j} \eta^{j}$$

и, следовательно, представляют собой контравариантный тензор ранга 1.

3. Линейная функция (ковариантный вектор). Числа a_i, определяющие линейную функцию, преобразуются по формулам

$$a'_{i} = c^{j}_{i} a_{j}$$

и, следовательно, образуют ковариантный тензор ранга 1.

4. Билинейная функция. Пусть $A(x; y)$ — билинейная форма в пространстве R. Отнесем каждому базису матрицу данной билинейной формы в этом базисе. Мы получим при этом тензор ранга два, дважды ковариантный.

Аналогично, билинейная форма от векторов $x \in R$, $f \in R'$ определяет тензор ранга два, один раз ковариантный и один раз контравариантный, а билинейная форма от векторов f, $g \in R'$ определяет тензор, дважды контравариантный.

5. Линейные преобразования. Пусть A — линейное преобразование в пространстве R. Отнесем каждому базису матрицу $\| a_i^k \|$ преобразования A в этом базисе, т. е. положим

$$Ae_i = a_i^k e_k.$$

Покажем, что $\| a_i^k \|$ есть тензор ранга два, один раз ковариантный и один раз контравариантный. Действительно, пусть переход к новому базису задается формулой

$$e_i' = c_i^\alpha e_\alpha$$

и, следовательно, обратный переход — формулой

$$e_i = b_i^\alpha e_\alpha', \quad \text{где} \quad b_i^\alpha c_\alpha^k = \delta_i^k.$$

Тогда

$$Ae_i' = Ac_i^\alpha e_\alpha = c_i^\alpha Ae_\alpha = c_i^\alpha a_\alpha^\beta e_\beta = c_i^\alpha a_\alpha^\beta b_\beta^k e_k'.$$

Таким образом, матрица $\| a_i'^k \|$ преобразования A в базисе e_i' имеет вид

$$a_i'^k = a_\alpha^\beta c_i^\alpha b_\beta^k,$$

что и доказывает, что матрица линейного преобразования A есть тензор второго ранга, один раз ковариантный и один раз контравариантный.

В частности, единичному преобразованию E в каждом базисе соответствует единичная матрица, т. е. система чисел

$$\delta_i^k = \begin{cases} 1 & \text{при } i = k, \\ 0 & \text{при } i \neq k. \end{cases}$$

Таким образом, δ_i^k представляет собой простейший тензор ранга два, один раз ковариантный и один раз

контравариантный. Тензор δ_i^k интересен тем, что его компоненты в любой системе координат одни и те же.

Упражнение. Показать непосредственно, что если в каждой системе координат задать систему чисел

$$\delta_i^k = \begin{cases} 1 \text{ при } i=k, \\ 0 \text{ при } i \neq k, \end{cases}$$

то это будет тензор.

Докажем теперь два простых предложения о тензорах.

Пусть имеются два тензора одинакового типа. Тогда для равенства тензоров достаточно, чтобы их компоненты в каком-нибудь базисе были соответственно равны. Другими словами, из того, что компоненты этих двух тензоров равны в какой-либо системе координат, следует, что их компоненты соответственно равны в произвольной системе координат. Это предложение очевидно; действительно, так как оба тензора одинакового типа (т. е. имеют одно и то же число ковариантных и контравариантных индексов), то они преобразуются по одним и тем же формулам, и так как их компоненты в одной системе координат по предположению равны, то они равны и в любой другой системе координат. Заметим, что предположение, что оба тензора одинакового типа, является совершенно обязательным. Например, как билинейная форма, так и линейное преобразование определяются в данной системе координат матрицей. Однако из совпадения матриц линейного преобразования и билинейной формы в одной какой-либо системе координат не следует их совпадение в другой.

При заданных p и q мы можем построить тензор типа (p, q), компоненты которого в каком-нибудь одном базисе равны n^{p+q} наперед заданным числам. Докажем это.

Пусть в некотором базисе нам задана система чисел $a_{ij\ldots}^{rs\ldots}$. Этими числами задается полилинейная функция $l(x, y, \ldots; f, \ldots)$ по формуле (1) п. 2 этого параграфа, где ξ^i, соотв. η^j и т. д.,—координаты векторов x, соотв. y и т. д. в базисе e_i. Так как с полилинейной функцией однозначно связан тензор, то мы получили тем самым тензор, удовлетворяющий поставленным условиям.

4. Тензоры в евклидовом пространстве. Если R есть n-мерное евклидово пространство, то, как мы видели в п. 5 § 23, можно установить изоморфное соответствие

между R и R' так, что если $y \in R$ соответствует элементу $f \in R'$, то

$$(f, x) = (y, x)$$

для любого $x \in R$.

Если мы теперь в полилинейной функции, зависящей от p векторов $x, y, \ldots$ из R и q векторов $f, g, \ldots$ из R', заменим векторы из R' им соответствующими векторами $u, v, \ldots$ из R, то мы получим полилинейную функцию $l(x, y, \ldots; u, v, \ldots)$, зависящую от $p+q$ векторов из R.

Найдем коэффициенты функции $l(x, y, \ldots; u, v, \ldots)$ по коэффициентам функции $l(x, y, \ldots; f, g, \ldots)$.

Пусть $a_{ij\ldots}^{rs\ldots}$ — коэффициенты полилинейной функции $l(x, y, \ldots; f, g, \ldots)$, т. е.

$$a_{ij\ldots}^{rs\ldots} = l(e_i, e_j, \ldots; f^r, f^s, \ldots),$$

и пусть $b_{ij\ldots rs\ldots}$ — коэффициенты полилинейной функции $l(x, y, \ldots; u, v, \ldots)$, т. е.

$$b_{ij\ldots rs\ldots} = l(e_i, e_j, \ldots; e_r, e_s, \ldots).$$

Мы доказали в п. 5 § 23, что в евклидовом пространстве векторы e_k базиса, биортогонального f^i, выражаются через векторы базиса f^i по формулам

$$e_r = g_{r\alpha} f^\alpha,$$

где

$$g_{ik} = (e_i, e_k).$$

Подставляя вместо $e_r, \ldots$ их выражения, получаем

$$\begin{aligned} b_{ij\ldots rs\ldots} &= l(e_i, e_j \ldots; e_r, e_s, \ldots) = \\ &= l(e_i, e_j, \ldots; g_{\alpha r} f^\alpha, g_{\beta s} f^\beta, \ldots) = \\ &= g_{\alpha r} g_{\beta s} \ldots l(e_i, e_j, \ldots; f^\alpha, f^\beta, \ldots) = \\ &= g_{\alpha r} g_{\beta s} \ldots a_{ij\ldots}^{\alpha\beta\ldots}. \end{aligned}$$

Ввиду установленного соответствия между полилинейными функциями и тензорами мы можем сформулировать полученный результат для тензоров:

Если $a_{ij\ldots}^{rs\ldots}$ — тензор, построенный в евклидовом пространстве, p раз ковариантный и q раз контравариантный, то по нему можно построить новый тензор $b_{ij\ldots rs\ldots}$, являющийся $p+q$ раз ковариантным. Эта операция

называется операцией опускания индексов. Она определяется формулой

$$b_{ij\ldots rs\ldots} = g_{\alpha r} g_{\beta s} \ldots a_{ij\ldots}^{\alpha\beta\ldots}.$$

g_{ik} является дважды ковариантным тензором. Действительно, $g_{ik} = (e_i, e_k)$ представляют собой в данной системе координат коэффициенты некоторой билинейной формы, а именно скалярного произведения. Ввиду его связи со скалярным произведением (метрикой) пространства *тензор g_{ik} называется метрическим тензором.*

Совершенно аналогично операции опускания индексов можно ввести операцию поднимания индексов с помощью формулы

$$b^{ij\ldots rs\ldots} = g^{\alpha i} g^{\beta j} \ldots a_{\alpha\beta\ldots}^{rs\ldots},$$

где g^{ik} имеет смысл, указанный в § 23, п. 5.

Упражнение. Показать, что g^{ik} — дважды контравариантный тензор.

5. Операции над тензорами. Ввиду установленной связи между тензорами и полилинейными функциями мы будем определять операции над полилинейными функциями. Запись полученных результатов в произвольном базисе даст нам соответствующую операцию над тензорами.

Сложение тензоров. Пусть

$$l'(x, y, \ldots; f, g, \ldots), \quad l''(x, y, \ldots; f, g, \ldots)$$

— две полилинейные функции от одного и того же числа векторов из R и одного и того же числа векторов из R'. Определим их сумму $l(x, y, \ldots; f, g, \ldots)$ формулой

$$l(x, y, \ldots; f, g, \ldots) = \\ = l'(x, y, \ldots; f, g, \ldots) + l''(x, y, \ldots; f, g, \ldots).$$

Ясно, что эта сумма есть снова полилинейная функция от того же числа векторов из R и из R'. Сложение тензоров определяется поэтому формулой:

$$a_{ij\ldots}^{rs\ldots} = a'^{rs\ldots}_{ij\ldots} + a''^{rs\ldots}_{ij\ldots}.$$

Умножение тензоров. Пусть

$$l'(x, y, \ldots; f, g, \ldots) \text{ и } l''(z, \ldots; h, \ldots)$$

— две полилинейные функции, из которых первая зависит от p' векторов из R и q' векторов из R', а вторая — от p'' векторов из R и q'' векторов из R'. Определим функцию $l(x, y, \ldots, z, \ldots; f, g, \ldots, h, \ldots)$ формулой

$$l(x, y, \ldots, z, \ldots; f, g, \ldots, h, \ldots) = \\ = l'(x, y, \ldots; f, g, \ldots) l''(z, \ldots; h, \ldots).$$

Функция l называется произведением полилинейных функций l' и l''. Покажем, что l есть полилинейная функция, зависящая от $p'+p''$ векторов из R и $q'+q''$ векторов из R'. Действительно, при проверке того, что

$$l(x, y, \ldots, z, \ldots; f, g, \ldots, h, \ldots)$$

есть полилинейная функция, мы фиксируем, по очереди, все векторы, кроме одного; при этом ясно, что l есть линейная функция от вектора, оставшегося незафиксированным.

Выразим компоненты тензора, отвечающего произведению полилинейных функций l' и l'', через компоненты тензоров, отвечающих самим этим полилинейным функциям. Так как

$$a'^{rs}_{ij}{}^{\cdots}_{\cdots} = l'(e_i, e_j, \ldots; f^r, f^s, \ldots)$$

и

$$a''^{tu}_{kl}{}^{\cdots}_{\cdots} = l''(e_k, e_l, \ldots; f^t, f^u, \ldots),$$

то

$$a^{rstu}_{ijkl}{}^{\cdots}_{\cdots} = a^{rs}_{ij}{}^{\cdots}_{\cdots} a^{tu}_{kl}{}^{\cdots}_{\cdots}$$

Эта формула определяет, таким образом, произведение двух тензоров.

Свертка тензора. Пусть $l(x, y, \ldots; f, g, \ldots)$ — полилинейная функция, зависящая от p векторов $x, y, \ldots$ из R $(p \geqslant 1)$ и q векторов $f, g, \ldots$ из R' $(q \geqslant 1)$. Мы построим по ней полилинейную функцию, зависящую от $p-1$ векторов из R и $q-1$ векторов из R'. Выберем для этого какой-либо базис $e_1, e_2, \ldots, e_n$ в R и взаимный с ним базис $f^1, f^2, \ldots, f^n$ в R'. Будем теперь вместо x и f подставлять соответственно $e_1, f^1; e_2, f^2; \ldots; e_n, f^n$

и рассмотрим сумму *)

$$l'(y, \ldots; g, \ldots) = l(e_\alpha, y, \ldots; f^\alpha, g, \ldots). \qquad (7)$$

Ясно, что каждое слагаемое, а значит и вся сумма, есть полилинейная функция от $y, \ldots$ и $g, \ldots$ Покажем, что, хотя каждое слагаемое зависит от выбора базиса, построенная нами сумма от выбора базиса уже не зависит.

Перейдем для этого к другому базису $e'_1, e'_2, \ldots, e'_n$ и соответственно к взаимному с ним базису $f'^1, f'^2, \ldots, f'^n$. Так как мы не меняем при этом векторов $y, \ldots$ и $g, \ldots$, то мы можем их фиксировать и доказывать наше утверждение для билинейной формы $A(x; f)$. Итак, нам нужно доказать, что если $A(x; f)$ — билинейная форма, то

$$A(e_\alpha; f^\alpha) = A(e'_\alpha; f'^\alpha).$$

Если переход от базиса $e_1, e_2, \ldots, e_n$ к базису $e'_1, e'_2, \ldots$ $\ldots, e'_n$ задается формулой

$$e'_i = c_i^k e_k,$$

то переход от базиса $f'^1, f'^2, \ldots, f'^n$ к базису $f^1, f^2, \ldots, f^n$ задается формулой

$$f^k = c_i^k f'^i.$$

Поэтому

$$A(e'_\alpha; f'^\alpha) = A(c_\alpha^k e_k; f'^\alpha) = c_\alpha^k A(e_k; f'^\alpha) =$$
$$= A(e_k; c_\alpha^k f'^\alpha) = A(e_k; f^k),$$

т. е. $A(e_\alpha; f^\alpha)$ действительно не зависит от системы координат.

Найдем по коэффициентам формы $l(x, y, \ldots; f, g, \ldots)$ коэффициенты формы (7). Так как

$$a'^s_j \ldots = l'(e_j, \ldots; f^s, \ldots)$$

и

$$l'(e_j, \ldots; f^s, \ldots) = l(e_\alpha, e_j, \ldots; f^\alpha, f^s, \ldots),$$

то

$$a'^s_j \ldots = a^{\alpha s}_{\alpha j} \ldots. \qquad (8)$$

---

*) Напоминаем, что если в некотором выражении один и тот же индекс (в данном случае α) встречается вверху и внизу, то по нему производится суммирование.

Тензор $a_j^{\prime s}{}_{\cdot\cdot\cdot}$, *полученный из* $a_{ij}^{rs}{}_{\cdot\cdot\cdot}$ *по формуле* (8), *называется сверткой тензора* $a_{ij}^{rs}{}_{\cdot\cdot\cdot}$.

Ясно, что свертку мы можем провести не обязательно по первому верхнему и первому нижнему индексам. Обязательно лишь, чтобы суммирование производилось по одному ковариантному и одному контравариантному индексу. Если бы мы суммировали, например, по двум нижним индексам, то полученная система чисел не образовывала бы тензора (так как при переходе от одной системы координат к другой эти числа не преобразовывались бы по предписанному тензору закону преобразования).

Заметим, что в случае свертки тензора ранга два мы получаем тензор нулевого ранга (скаляр), т. е. число, не зависящее от системы координат.

Рассмотренная нами в п. 4 операция опускания индексов есть не что иное, как свертка произведения данного тензора и метрического тензора g_{ik} (взятого сомножителем соответствующее число раз). Аналогично, поднимание индексов есть свертка произведения данного тензора и тензора g^{ik}.

Приведем еще пример. Пусть a_{ij}^k — тензор ранга три, а b_l^m — тензор ранга два. Их произведение есть тензор $c_{ijl}^{km} = a_{ij}^k b_l^m$ ранга пять. Если теперь свернуть этот тензор, например, по индексам i и m, то мы получим тензор ранга три. Если мы полученный тензор еще раз свернем, например, по индексам j и k, то мы получим тензор ранга один (вектор).

Пусть a_l^i и b_k^l — два тензора ранга два. Умножением и свертыванием можно построить по ним новый тензор ранга два:

$$c_i^l = a_i^\alpha b_\alpha^l.$$

Если тензоры a_i^l и b_k^l трактовать как матрицы линейных преобразований, то полученный тензор есть матрица произведения этих преобразований.

Мы можем также построить по данному тензору ранга два a_l^i ряд инвариантов (т. е. чисел, не зависящих от системы координат, — скаляров), а именно:

$$a_\alpha^\alpha,\ a_\alpha^\beta a_\beta^\alpha,\ \ldots$$

Введенные нами операции над тензорами дают нам возможность по данным тензорам строить ряд новых, инвариантно связанных с ними, тензоров.

Приведем некоторые примеры.

Операцией умножения мы можем из векторов построить тензоры сколь угодно высокого ранга. Пусть, например, ξ^i — координаты контравариантного, а η_j — координаты ковариантного вектора. Тогда $\xi^i\eta_j$ есть тензор ранга два. Аналогично, взяв бо́льшее число векторов, можно получить тензоры более высокого ранга. Заметим, что не всякий тензор можно получить умножением векторов. Можно, однако, доказать, что всякий тензор может быть получен из векторов (тензоров ранга один) операциями сложения и умножения.

Целым рациональным инвариантом от данной системы тензоров называется многочлен от компонент тензора, который не меняется при замене компонент тензоров в какой-нибудь системе координат их компонентами в другой системе координат.

Имеет место следующая теорема, которую мы не будем доказывать:

Если задана некоторая система тензоров, то всякий целый рациональный инвариант, построенный по данным тензорам, можно получить из них операциями перемножения тензоров, сложения и умножения на числа и полного свертывания (т. е. свертывания по всем индексам).

6. Симметрические и знакопеременные (антисимметрические) тензоры.

Определение. *Тензор называется симметрическим по данным индексам, если при любой перестановке этих индексов компоненты тензора не меняются* *). Например, симметричность тензора по первым двум индексам означает, что имеет место равенство

$$a_{ik\ldots}^{st\ldots} = a_{ki\ldots}^{st\ldots}.$$

Если $l(x, y, \ldots; f, g, \ldots)$ — соответствующая тензору $a_{ik\ldots}^{st\ldots}$ полилинейная форма

$$l(x, y, \ldots; f, g, \ldots) = a_{ik\ldots}^{st\ldots}\xi^i\eta^k \ldots \lambda_s\mu_t \ldots \qquad (9)$$

то симметричность тензора по некоторой группе индексов, как это непосредственно видно из формулы (9), эквива-

*) Само собой разумеется, что речь идет лишь об индексах одной и той же группы (верхней или нижней).

лентна симметричности полилинейной формы по соответствующей группе векторов. Так как для симметричности полилинейной формы по некоторой группе векторов достаточно, чтобы $a_{ik\ldots}^{st\ldots}$ были симметричны по соответствующим индексам лишь в одной какой-нибудь системе координат, то отсюда следует, что *если компоненты тензора симметричны в одной системе координат, то такая же симметрия будет иметь место и в любой другой системе координат.*

Определение. *Знакопеременным (антисимметрическим) называется тензор, который меняет знак при перемене любых двух индексов местами.*

При этом предполагается, конечно, что у этого тензора все индексы одинакового характера, т. е. либо все ковариантные, либо все контравариантные.

Из определения знакопеременного тензора непосредственно следует, что при любой перестановке индексов компоненты тензора не меняются, если перестановка четная, и меняют знак, если перестановка нечетная. Знакопеременным тензорам соответствуют знакопеременные полилинейные функции.

Полилинейная функция $l(x, y, \ldots)$, зависящая от p векторов $x, y, \ldots$ из R, называется знакопеременной, если при перестановке любой пары из векторов $x, y, \ldots$ знак функции меняется.

Для проверки знакопеременности полилинейной функции достаточно проверить знакопеременность компонент соответствующего ей тензора в какой-либо одной системе координат, как это непосредственно следует из формулы (9). С другой стороны, из знакопеременности полилинейной функции следует знакопеременность соответствующего ей тензора (в любой системе координат). Следовательно, если компоненты тензора знакопеременны в какой-либо одной системе координат, то это же имеет место и в любой другой системе координат, и, значит, тензор является знакопеременным (антисимметрическим).

Выясним число независимых компонент антисимметрического тензора. Пусть, например, a_{ik} есть знакопеременный тензор ранга 2. Тогда $a_{ik} = -a_{ki}$ и, следовательно, число различных компонент равно $\frac{n(n-1)}{2}$. Аналогично,

для знакопеременного тензора a_{ijk} число различных компонент равно $\frac{n(n-1)(n-2)}{3!}$, так как компоненты с одинаковыми индексами равны нулю, а компоненты, отличающиеся лишь порядком индексов, определяются одна через другую.

Аналогично, число независимых компонент знакопеременного тензора с k индексами ($k \leqslant n$) равно C_n^k. (Отличных от нуля знакопеременных тензоров с числом индексов больше, чем n, не существует, так как у знакопеременного тензора компоненты хотя бы с двумя одинаковыми индексами равны нулю, а если число индексов превышает n, то у каждой компоненты совпадают, по крайней мере, два индекса.)

Рассмотрим более подробно знакопеременный тензор с n индексами. Так как все группы по n различных индексов, принимающих значение от 1 до n, отличаются лишь порядком, то у такого тензора есть лишь одна независимая компонента и он имеет, таким образом, следующий вид.

Пусть $i_1, i_2, \ldots, i_n$ — некоторая перестановка чисел $1, 2, \ldots, n$. Положим $a_{12\ldots n} = a$. Тогда

$$a_{i_1 i_2 \ldots i_n} = \pm a, \tag{10}$$

где $+$ отвечает четной подстановке, а $-$ нечетной.

Упражнение. Показать, что при переходе к другой системе координат число $a_{12\ldots n} = a$ умножится на определитель матрицы перехода.

Напишем полилинейную функцию, соответствующую знакопеременному тензору с n индексами. В силу формулы (10) она имеет вид:

$$l(x, y, \ldots, z) = a_{i_1 i_2 \ldots i_n} \xi^{i_1} \eta^{i_2} \ldots \zeta^{i_n} = a \begin{vmatrix} \xi_1 & \xi_2 & \ldots & \xi_n \\ \eta_1 & \eta_2 & \ldots & \eta_n \\ \cdot & \cdot & \cdot & \cdot \\ \zeta_1 & \zeta_2 & \ldots & \zeta_n \end{vmatrix}.$$

Мы доказали, таким образом, что определитель из координат векторов есть, с точностью до множителя, единственная знакопеременная полилинейная функция от n векторов в n-мерном линейном пространстве.

Операция симметрирования. Мы можем по всякому тензору построить новый тензор, симметричный по некоторой наперед заданной группе индексов. Эта операция называется симметрированием и состоит в следующем.

Пусть задан некоторый тензор, например $a_{i_1 i_2 \dots i_n}$; симметрирование его, например, по первым k индексам, состоит в построении тензора

$$a_{(i_1 i_2 \dots i_k) i_{k+1} \dots} = \frac{1}{k!} \sum a_{j_1 j_2 \dots j_k i_{k+1} \dots},$$

где сумма распространяется по всем перестановкам $j_1, j_2, \dots, j_k$ индексов $i_1, i_2, \dots, i_k$. Например,

$$a_{(i_1, i_2)} = \frac{1}{2} (a_{i_1 i_2} + a_{i_2 i_1}).$$

Операция симметрирования тензора по группе из k индексов $i_1, i_2, \dots, i_k$ обозначается следующим образом:

$$a_{j_1 j_2 \dots j_l (i_1 i_2 \dots i_k) \dots}.$$

Операция альтернирования вводится аналогично операции симметрирования и дает возможность по данному тензору построить тензор, знакопеременный по данной группе индексов. Она определяется следующим образом:

$$a_{[i_1 i_2 \dots i_k] i_{k+1} \dots} = \frac{1}{k!} \sum \pm a_{j_1 j_2 \dots j_k i_{k+1} \dots},$$

где сумма распространяется по всем перестановкам $j_1, j_2, \dots, j_k$ индексов $i_1, i_2, \dots, i_k$, а знак определяется четностью или нечетностью этой перестановки. Например,

$$a_{[i_1 i_2]} = \frac{1}{2} (a_{i_1 i_2} - a_{i_2 i_1}).$$

Операция альтернирования обозначается скобками [ ]; в них заключаются те индексы, по которым тензор альтернируется.

По всяким k векторам $\xi^{i_1}, \eta^{i_2}, \dots, \zeta^{i_k}$ можно построить антисимметрический тензор

$$a^{i_1 i_2 \dots i_k} = \xi^{[i_1} \eta^{i_2} \dots \zeta^{i_k]}, \tag{11}$$

где через $\xi^{[i_1} \eta^{i_2} \dots \zeta^{i_k]}$ обозначен тензор, полученный альтернированием тензора $\xi^{i_1} \eta^{i_2} \dots \zeta^{i_k}$. Как нетрудно

усмотреть из написанной формулы, компонентами этого тензора являются миноры k-го порядка следующей матрицы из n столбцов:

$$\begin{pmatrix} \xi^1 & \xi^2 & \dots & \xi^n \\ \eta^1 & \eta^2 & \dots & \eta^n \\ \cdot & \cdot & \cdot & \cdot \\ \zeta^1 & \zeta^2 & \dots & \zeta^n \end{pmatrix}.$$

Построенный тензор (11) обладает тем свойством, что если к какому-либо из векторов ξ^{i_1}, η^{i_2}, ... добавить линейную комбинацию остальных, то тензор $a^{i_1 \dots i_k}$ от этого не изменится.

Рассмотрим k-мерное подпространство n-мерного пространства R. Поставим вопрос о том, чтобы охарактеризовать это k-мерное подпространство системой чисел, т. е. ввести координаты подпространства.

k-мерное подпространство порождается k линейно независимыми векторами ξ^{i_1}, η^{i_2}, ..., ζ^{i_k}. При этом разные системы из k векторов могут породить одно и то же подпространство. Однако нетрудно показать, и мы предоставляем это читателю, что если две системы векторов порождают одно и то же подпространство, то построенные по каждой из них тензоры

$$a^{i_1 i_2 \dots i_k} = \xi^{[i_1} \eta^{i_2} \dots \zeta^{i_k]}$$

совпадают с точностью до множителя.

Таким образом, тензор $a^{i_1 i_2 \dots i_k}$, построенный по векторам ξ^{i_1}, η^{i_2}, ..., ζ^{i_k}, порождающим некоторое подпространство, определяет это подпространство.

§ 25. Тензорное произведение

1. Тензорное произведение $R \otimes R$. В первой главе мы изучали билинейные функции в аффинном пространстве R. Здесь мы покажем, что билинейные функции можно трактовать и как линейные функции в некотором новом пространстве. Это пространство, играющее очень важную роль, называется тензорным произведением R и R (по-другому, тензорным квадратом R) и обозначается $R \otimes R$ или $\overset{2}{\otimes} R$. Дадим его определение.

Рассмотрим всевозможные упорядоченные пары x, y элементов из R. Каждую такую пару будем называть *тензорным произведением* x и y и обозначать $x \otimes y$. Образуем формальные конечные суммы таких пар:

$$X = x_1 \otimes y_1 + \ldots + x_k \otimes y_k. \tag{1}$$

При этом формальные суммы, отличающиеся только порядком слагаемых, мы не будем различать между собой. Запись (1) означает, таким образом, только то, что нам задано множество k пар $x_1, y_1; \ldots; x_k, y_k$.

Введем для выражений вида (1) операции сложения и умножения на число. Сумму двух таких выражений определим как результат формального дописывания к первому выражению второго:

$$(x_1 \otimes y_1 + \ldots + x_k \otimes y_k) + (x_{k+1} \otimes y_{k+1} + \ldots + x_{k+l} \otimes y_{k+l}) = \\ = x_1 \otimes y_1 + \ldots + x_{k+l} \otimes y_{k+l}. \tag{2}$$

Произведение на число λ определим так:

$$\lambda (x_1 \otimes y_1 + \ldots + x_k \otimes y_k) = (\lambda x_1) \otimes y_1 + \ldots + (\lambda x_k) \otimes y_k. \tag{3}$$

Элемент $0 \otimes 0$ будем называть *нулем* тензорного произведения и обозначать коротко через 0.

Мы должны еще объяснить, какие выражения вида (1) считаются равными. Будем предполагать, что

1) $(x_1 + x_2) \otimes y - x_1 \otimes y - x_2 \otimes y = 0$ *);
2) $x \otimes (y_1 + y_2) - x \otimes y_1 - x \otimes y_2 = 0$;
3) $(\lambda x) \otimes y - x \otimes (\lambda y) = 0$.

Кроме того, *приравняем нулю любое выражение, получающееся из выражений* 1), 2) *и* 3) *сложением и умножением на число.*

Теперь *два выражения* X *и* X' *вида* (1) *будем считать равными, если их можно превратить в одинаковые, прибавляя к* X *и* X' *выражения, равные нулю, т. е. если существуют такие выражения* $Z = 0$ *и* $Z' = 0$, *что* $X + Z$ *совпадает с* $X' + Z'$. Очевидно, что введенное отношение равенства рефлексивно (т. е. $X = X$) и симметрично (т. е. из $X = Y$ следует $Y = X$); легко проверить, что оно

*) Более подробно, это означает, что $(x_1 + x_2) \otimes y + (-x_1) \otimes y + (-x_2) \otimes y = 0$.

обладает также и свойством транзитивности (т. е. из $X=Y$ и $Y=Z$ следует $X=Z$).

В результате мы получаем пространство, элементы которого — классы равных между собой выражений вида (1), а сложение и умножение на число определены по формулам (2) и (3).

Заметим, что операции сложения и умножения на число мы ввели до того, как было определено отношение равенства двух выражений (1). Поэтому нам нужно было бы убедиться в корректности определений этих операций; именно нужно показать, что сумма выражений (1) и произведение на число не меняются при замене этих выражений на равные. Эта простая проверка предоставляется читателю.

Нетрудно убедиться, что построенное пространство является линейным пространством. Покажем, например, что для каждого X существует элемент, ему противоположный. Это достаточно проверить для элементов вида $X=x\otimes y$. Из условия 3) при $\lambda=1$ получаем $x\otimes y+(-x)\otimes y=0$, т. е. $Y=(-x)\otimes y$ является элементом, противоположным X.

Построенное линейное пространство называется *тензорным произведением* R на R и обозначается через $R\otimes R$.

Итак, *мы определили тензорное произведение $R\otimes R$ как линейное пространство, элементами которого являются формальные выражения вида $x_1\otimes y_1+\ldots+x_k\otimes y_k$, где x_i, y_i — элементы из R. Точнее, элементами пространства $R\otimes R$ являются классы равных между собой выражений вида $x_1\otimes y_1+\ldots+x_k\otimes y_k$* (условие равенства дано выше). *Сложение в $R\otimes R$ и умножение на число определяются по формулам* (2) *и* (3).

Отметим, что из условий 1)—3) следует:

$$(x_1+x_2)\otimes y=x_1\otimes y+x_2\otimes y;$$
$$x\otimes(y_1+y_2)=x\otimes y_1+x\otimes y_2;$$
$$\lambda(x\otimes y)=(\lambda x)\otimes y=x\otimes(\lambda y).$$

Поэтому в тензорном произведении можно раскрывать скобки по обычному правилу:

$$(\lambda_1x_1+\ldots+\lambda_mx_m)\otimes(\mu_1y_1+\ldots+\mu_ny_n)=$$
$$=\sum_{i=1}^{m}\sum_{j=1}^{n}\lambda_i\mu_j(x_i\otimes y_j).$$