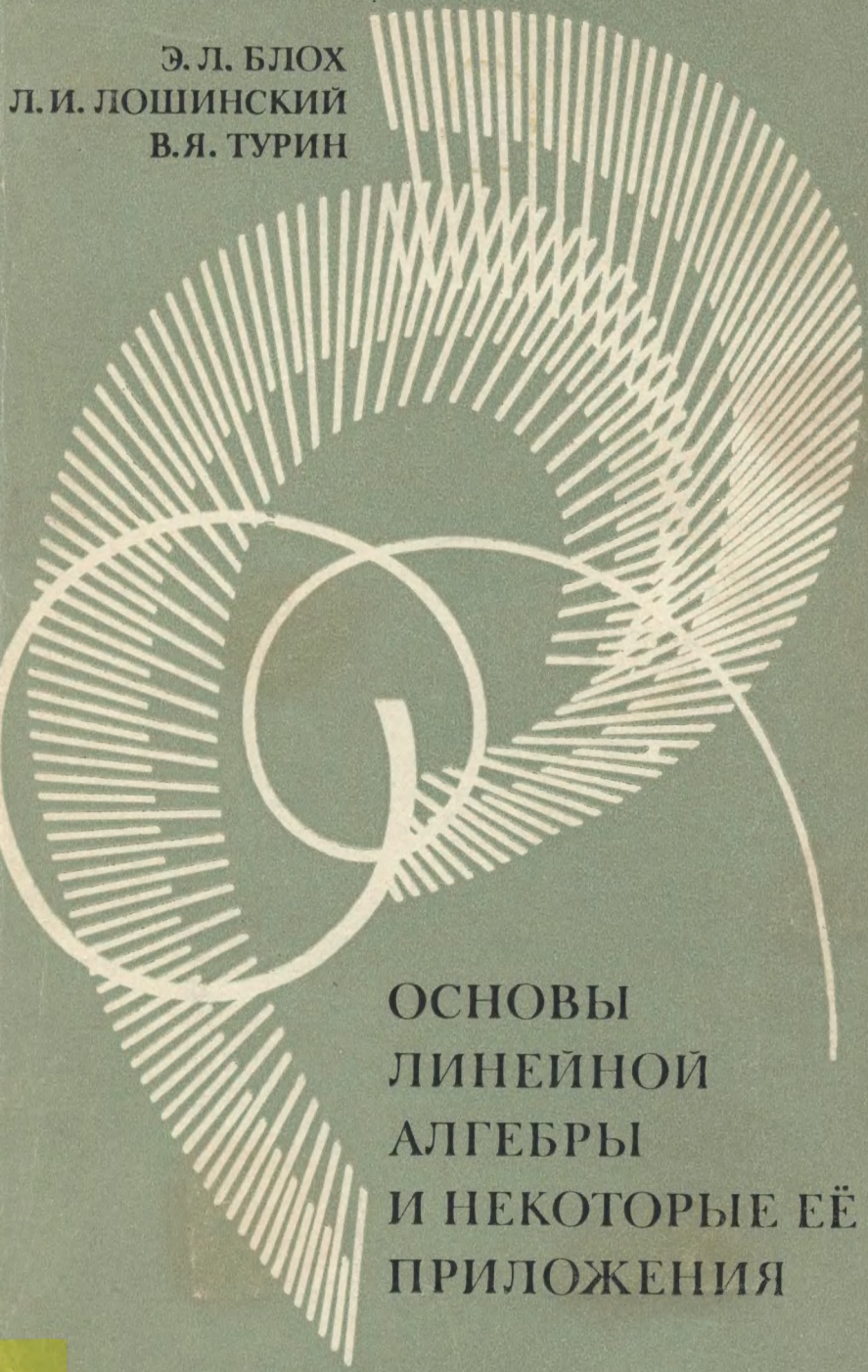
Э. Л. БЛОХ
Л. И. ЛОШИНСКИЙ
В. Я. ТУРИН



ОСНОВЫ ЛИНЕЙНОЙ АЛГЕБРЫ И НЕКОТОРЫЕ ЕЁ ПРИЛОЖЕНИЯ

Э. Л. БЛОХ,
И. ЛОШИНСКИЙ,
В. Я. ТУРИН

ОСНОВЫ ЛИНЕЙНОЙ АЛГЕБРЫ И НЕКОТОРЫЕ ЕЕ ПРИЛОЖЕНИЯ

Допущено Министерством высшего и среднего специального образования СССР в качестве учебного пособия для студентов втузов

Издательство
«ВЫСШАЯ ШКОЛА»
Москва — 1971

517.3
Б 70
512.8

УДК
5705

1 к
n8

Блох Э. Л., Лошинский Л. И., Турин В. Я.

Б70 Основы линейной алгебры и некоторые ее приложения. Учебное пособие. М., «Высшая школа», 1971.

256 стр. с илл.

В книге излагаются основы линейной алгебры: матрицы, линейные преобразования, системы линейных уравнений, линейное пространство, линейные операторы, евклидово пространство и квадратичные формы. Иллюстрируется применение методов линейной алгебры к некоторым вопросам анализа, теории линейных дифференциальных уравнений и цепей Маркова.

Имеется достаточное число поясняющих примеров. В конце каждой главы приводятся упражнения с ответами.

Книга предназначена для студентов технических вузов.

2-2-3 / 30-70 512.8

Рецензенты: кафедра высшей математики МЭИ и доц. В. В. Пашенков.

Эфраим Леонтьевич Блох,
Лев Ильич Лошинский,
Виллен Яковлевич Турин

ОСНОВЫ ЛИНЕЙНОЙ АЛГЕБРЫ И НЕКОТОРЫЕ ЕЕ ПРИЛОЖЕНИЯ

Редактор Д. А. Тальский Художественный редактор В. И. Пономаренко
Художник А. Е. Коленков Технический редактор Л. А. Муравьева
Корректор Л. З. Черникова

Сдано в набор 29/V 1970 г. Подп. к печати 25/XI 1970 г. Формат 84×108 1/32.
Объем 8 печ. л. 13,44 усл. п. л. Уч.-изд. л. 10,71. Изд. № ФМ—435.
Тираж 40 000 экз. Заказ 1229. Цена 30 коп.
План выпуска литературы издательства «Высшая школа» (вузы и техникумы) на 1970 г. Позиция № 30.
Москва, К-51, Неглинная ул., д. 29/14
Издательство «Высшая школа»

Ордена Трудового Красного Знамени Ленинградская типография № 1 «Печатный Двор» им. А. М. Горького Главполиграфпрома Комитета по печати при Совете Министров СССР, г. Ленинград, Гатчинская ул., 26.

НАУЧНАЯ БИБЛИОТЕКА
им. Горького
МГУ

3570-5-71

ПРЕДИСЛОВИЕ

В последние годы методы линейной алгебры находят все большее применение при рассмотрении многочисленных прикладных вопросов. Поэтому основные вопросы линейной алгебры (операции над матрицами, понятие о линейном и евклидовом пространствах, линейные операторы, квадратичные формы и т. д.) включены в программу по математике для технических вузов (см. гл. 1—7).

Кроме того, в книге излагаются некоторые приложения линейной алгебры не только к квадратичным формам, но и к другим разделам курса математики.

В главах 10—12 рассматриваются приложения методов линейной алгебры к теории линейных дифференциальных уравнений, к некоторым вопросам анализа и цепям Маркова.

Включение этих вопросов потребовало знакомства с дополнительными разделами линейной алгебры (главы 8 и 9), которые выходят за рамки действующих в настоящее время учебных программ.

Каждая глава иллюстрирована примерами, поясняющими применение основных теоретических результатов. В конце глав 1—10 предложены задачи для упражнений, часть из которых заимствована из известных задачников.

В книге не содержится теория определителей. Для удобства чтения во введении приводятся основные свой-

ства определителей, используемые при изложении материала.

Настоящее пособие написано на основе лекций, читанных авторами в течение ряда лет студентам и аспирантам Московского электротехнического института связи.

Авторы выражают искреннюю признательность рецензентам доцентам И. А. Брину и В. В. Пашенкову, а также научному редактору Д. А. Тальскому.

ВВЕДЕНИЕ

Определители n-го порядка

1. Перестановки. *Перестановкой* n-чисел 1, 2, ..., n называется любое расположение этих чисел в определенном порядке. Число всевозможных перестановок из n чисел равно $n!$ Два числа в перестановке образуют *инверсию*, если бо́льшее из них стоит впереди ме́ньшего.

Число пар в перестановке, образующих инверсию, называется *числом инверсий перестановки*. Перестановка называется четной, если число ее инверсий четное, и нечетной, если нечетное.

Например, в перестановке 5, 2, 3, 1, 6, 4 число инверсий равно $4+1+1+0+1+0=7$. Перестановка нечетная.

2. Определитель n-го порядка. Рассмотрим квадратную таблицу A, составленную из n^2 чисел

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

Определителем n-го порядка называется число, полученное из элементов данной таблицы по следующему правилу:

1) определитель n-го порядка равен алгебраической сумме $n!$ членов;

2) каждый член представляет собой произведение n элементов, взятых по одному из каждой строки и каждого столбца таблицы;

3) член берется со знаком плюс, если перестановки, образованные первыми и вторыми индексами элементов a_{ij}, входящими в произведение, одинаковой четности (либо

обе четные, либо обе нечетные) и со знаком минус в противоположном случае.

Определитель обозначается символом

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix},$$

или кратко

$$\det A = |A|.$$

Согласно этому определению, например, определитель второго порядка равен

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}\ a_{22} - a_{21}\, a_{12},$$

а определитель третьего порядка равен

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}\ a_{12}\ a_{13} + a_{12}\ a_{23}\ a_{31} + a_{21}\ a_{32}\ a_{13} - \\ - a_{13}\ a_{22}\ a_{31} - a_{11}\ a_{23}\ a_{32} - a_{33}\ a_{12}\ a_{21}.$$

Суммой нескольких строк определителя $|A|$ называется строка, каждый элемент которой равен сумме соответствующих элементов этих строк. *Произведением строки на число* называется строка, каждый элемент которой равен соответствующему элементу данной строки, умноженному на это число.

Линейной комбинацией нескольких строк определителя называется строка, равная сумме произведений этих строк на некоторые числа.

3. Минор и алгебраическое дополнение. Имеем определитель $|A| = |a_{ij}|$ n-го порядка ($n > 1$). Минором M_{ij} элемента a_{ij} определителя $|A|$ называется определитель $n-1$-го порядка, полученный из $|A|$ вычеркиванием i-й строки и j-го столбца, на пересечении которых стоит данный элемент a_{ij}. Так как в определителе $|A|$ n^2 элементов a_{ij}, то $|A|$ содержит n^2 различным образом составленных миноров. Алгебраическим дополнением элемента a_{ij}

называется число $A_{ij}=(-1)^{i+j}M_{ij}$. Например:

$$|A|=\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix};$$

$$M_{11}=\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}; \quad M_{21}=\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix}; \quad M_{32}=\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix};$$

$$A_{11}=(-1)^{1+1}M_{11}=M_{11}; \quad A_{21}=(-1)^{2+1}M_{21}=-M_{21};$$

$$A_{32}=(-1)^{3+2}M_{32}=-M_{32}.$$

4. Основные свойства определителей.

1. Если переменить местами строки и столбцы определителя, не меняя их порядка (иначе — транспонировав определитель), то определитель не меняется.

Следствие. Если известно какое-либо свойство определителя, относящееся к его строкам, то оно будет справедливо и по отношению к его столбцам. Поэтому в дальнейшем все свойства определителя будут даны по отношению к его строкам.

2. При перестановке двух строк определитель меняет знак.

3. Если все элементы какой-либо строки определителя равны нулю, то определитель равен нулю.

4. Общий множитель всех элементов строки определителя можно вынести за знак определителя.

5. Если все элементы некоторой строки определителя пропорциональны соответствующим элементам другой строки, то определитель равен нулю.

6. Если все элементы i-й строки определителя представлены в виде суммы двух слагаемых, то определитель равен сумме двух определителей, у которых все строки, кроме i-й, такие же, как и в данном определителе, а i-я строка в первом определителе состоит из первых слагаемых, во втором — из вторых слагаемых; это свойство распространяется на случай, когда каждый элемент i-й строки определителя есть сумма k слагаемых ($k\geqslant 2$).

7. Если к строке определителя прибавить линейную комбинацию нескольких других его строк (сумму произведений на произвольные числа), то определитель не меняется.

8. Если одна из строк определителя есть линейная комбинация других его строк, то определитель равен

нулю. Обратно, если определитель n-го порядка ($n \geqslant 2$) равен нулю, то одна из его строк есть линейная комбинация других строк.

9. Определитель $|A| = |a_{ij}|$ равен сумме произведений всех элементов какой-либо его строки на их алгебраические дополнения:

$$|A| = a_{i1}A_{i1} + a_{i2}A_{i2} + \ldots + a_{in}A_{in}; \quad i = 1, 2, \ldots n.$$

10. Сумма произведений всех элементов какой-либо строки определителя на алгебраические дополнения соответствующих элементов другой строки равна нулю:

$$a_{i1}A_{j1} + a_{i2}A_{j2} + \ldots + a_{in}A_{jn} = 0; \quad j \neq i.$$

Замечание. Последние два равенства можно объединить, введя символ Кронекера

$$\delta_{ij} = \begin{cases} 1, \text{ если } i = j \\ 0, \text{ если } i \neq j. \end{cases}$$

Тогда

$$a_{i1}A_{j1} + a_{i2}A_{j2} + \ldots + a_{in}A_{jn} = \delta_{ij}|A|.$$

5. Умножение определителей. Произведение определителей $|A| = |a_{ij}|$ и $|B| = |b_{ij}|$ одного и того же порядка n равно определителю $|C| = |c_{ij}|$ того же порядка, где

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj}.$$

Замечание. Элементы c_{ij} можно вычислять и по следующим формулам:

$$c_{ij} = a_{i1}b_{j1} + a_{i2}b_{j2} + \ldots + a_{in}b_{jn},$$
$$c_{ij} = a_{1i}b_{j1} + a_{2i}b_{j2} + \ldots + a_{ni}b_{jn},$$
$$c_{ij} = a_{1i}b_{1j} + a_{2i}b_{2j} + \ldots + a_{ni}b_{nj}.$$

6. Дифференцирование определителя.

Если элементы a_{ij} определителя $|A|$ — дифференцируемые функции x, то

$$|A|' = \begin{vmatrix} a'_{11} a'_{12} \ldots a'_{1n} \\ a_{21}\ a_{22} \ldots a_{2n} \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a_{n1}\ a_{n2} \ldots a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11}\ a_{12} \ldots a_{1n} \\ a'_{21}\ a'_{22} \ldots a'_{2n} \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a_{n1}\ a_{n2} \ldots a_{nn} \end{vmatrix} + \ldots + \begin{vmatrix} a_{11}\ a_{12} \ldots a_{1n} \\ a_{21}\ a_{22} \ldots a_{2n} \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a'_{n1}\ a'_{n2} \ldots a'_{nn} \end{vmatrix}$$

7. Формулы Крамера.

Имеем следующую систему n линейных уравнений с n неизвестными:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n &= b_n. \end{aligned}$$

Если определитель системы $|A| \neq 0$, то система имеет единственное решение, которое находится по формулам Крамера

$$x_j = \frac{|A_j|}{|A|}, \quad j = 1, 2, \ldots n,$$

где определитель $|A_j|$ получается из определителя системы заменой j-го столбца на столбец из свободных членов.

Глава I

МАТРИЦЫ

§ 1. ОСНОВНЫЕ ОПРЕДЕЛЕНИЯ

Матрицей размера $m \times n$ называется прямоугольная таблица чисел, расположенных в m строках и n столбцах

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix},$$

числа a_{ij} ($i = 1, 2, \dots m$; $j = 1, 2, \dots n$) называются *элементами матрицы* A. Первый индекс i указывает номер строки, а второй индекс j — номер столбца, на пересечении которых расположен элемент a_{ij}.

В дальнейшем мы матрицу A часто будем сокращенно обозначать символом

$$A = (a_{ij})_{m,n},$$

или еще проще

$$A = (a_{ij}).$$

Если $m = n$, то матрица A называется *квадратной матрицей* порядка n и сокращенно обозначается $A = (a_{ij})_n$; если $m \neq n$, то матрица называется *прямоугольной*. Отметим некоторые специальные виды матрицы.

1. *Матрица-строка* — прямоугольная матрица размера $1 \times n$. Так как в этом случае матрица содержит только одну строку, то достаточно отмечать элементы одним индексом

$$A = (a_1 a_2 \dots a_n).$$

2. *Матрица-столбец* — прямоугольная матрица размера $m \times 1$

$$A = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}.$$

3. Матрицу, состоящую из одного элемента, будем отождествлять с этим элементом

$$(a_{ij})_{1,1} = a_{11}.$$

4. *Нулевая матрица* — матрица, все элементы которой равны нулю. Будем обозначать нулевую матрицу 0.

5. *Единичная матрица* порядка n — квадратная матрица n-го порядка вида

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & 1 \end{pmatrix}, \quad \text{или сокращенно} \quad \begin{pmatrix} 1 & & & O \\ & 1 & & \\ & & \ddots & \\ O & & & 1 \end{pmatrix}.$$

Единичную матрицу будем обозначать символом E. Используя символ Кронекера (см. введение), можно записать

$$E = (\delta_{ij})_n.$$

6. *Диагональная матрица* n-го порядка — квадратная матрица n-го порядка, элементы которой $a_{ij} = \lambda_i \delta_{ij}$, т. е. матрица вида $(\lambda_i \delta_{ij})_n$.

В развернутой форме диагональная матрица запишется так:

$$\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}, \quad \text{или} \quad \begin{pmatrix} \lambda_1 & & & O \\ & \lambda_2 & & \\ & & \ddots & \\ O & & & \lambda_n \end{pmatrix}.$$

Одной из важнейших характеристик квадратной матрицы порядка n является ее определитель (или детерминант), который обозначается $|A|$ или $\det A$ и представляет собой определитель n-го порядка, составленный из тех же элементов (расположенных в том же порядке), что и матрица. Таким образом, если

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix},$$

то

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

Очевидно, что $\det E = 1$, а определитель диагональной матрицы равен произведению диагональных элементов. Если $\det A$ отличен от нуля, то матрица называется *невырожденной* (или *неособенной*). В противном случае (если $\det A = 0$) матрица A называется *вырожденной* (или *особенной*).

§ 2. ОПЕРАЦИИ НАД МАТРИЦАМИ

1. Равенство матриц. Две матрицы $A = (a_{ij})$ и $B = (b_{ij})$ называются равными (что обозначается $A = B$), если их размеры совпадают и соответствующие их элементы равны, т. е. при всех i и j

$$a_{ij} = b_{ij}.$$

Из определения равенства очевидно, что если $A = B$, то $B = A$. Если же $A = B$ и $B = C$, то $A = C$.

2. Сложение матриц. *Суммой двух матриц* $A = (a_{ij})$ и $B = (b_{ij})$ одинаковых размеров называется матрица $C = (c_{ij})$ (что обозначается $C = A + B$) тех же размеров, элементы которой определяются равенствами

$$c_{ij} = a_{ij} + b_{ij}.$$

Например,

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \quad B = \begin{pmatrix} -1 & 0 & 2 \\ 1 & 0 & -2 \end{pmatrix}, \quad C = A + B = \begin{pmatrix} 0 & 2 & 5 \\ 5 & 5 & 4 \end{pmatrix}.$$

Из определения операции сложения вытекают следующие свойства:

$$A + B = B + A;$$
$$(A + B) + C = A + (B + C).$$

Доказательства этих свойств предоставляются читателю.

3. Умножение матрицы на число. *Произведением матрицы* $A = (a_{ij})$ *на число* α называется матрица $B = (b_{ij})$ (что обозначается $B = \alpha A$ или $B = A\alpha$), элементы которой определяются равенствами

$$b_{ij} = \alpha a_{ij}.$$

Таким образом,

$$\alpha A = (\alpha a_{ij}).$$

Из этого определения вытекают очевидные следствия:

$$\alpha(A+B)=\alpha A+\alpha B;$$
$$(\alpha+\beta)A=\alpha A+\beta A;$$
$$(\alpha\beta)A=\alpha(\beta A),$$

где α и β — два произвольных числа, а A и B — две матрицы одного размера.

Из свойств сложения матриц и умножения матрицы на число следует, что

а) сумма матрицы A и нулевой матрицы тех же размеров равна матрице A, т. е.

$$A+0=A;$$

б) для матрицы A существует единственная матрица $-A=(-1)\cdot A$ такая, что $A+(-A)=0$;

в) существует однозначная операция, обратная сложению (которая называется *вычитанием*), т. е. для любых матриц A и B одинаковых размеров, существует единственная матрица C тех же размеров такая, что

$$B+C=A.$$

Матрица C обозначается

$$C=A-B$$

и называется *разностью* матриц A и B. Причем $C=A+(-B)$, где

$$-B=(-1)\cdot B.$$

4. Умножение матриц. Умножение матрицы $A=(a_{ij})$ на матрицу $B=(b_{ij})$ определяется только при условии, что число столбцов матрицы A равно числу строк матрицы B.

Если A имеет размеры $(m\times p)$, а B — размеры $(p\times n)$, то *произведением матрицы A на матрицу B* называется матрица $C=(c_{ij})$ (что обозначается $C=AB$) размеров $(m\times n)$, элементы которой c_{ij} определяются равенствами

$$c_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}+\ldots+a_{ip}b_{pj}=$$
$$=\sum_{k=1}^{p}a_{ik}b_{kj}\quad(i=1,\ 2,\ \ldots\ m;\ j=1,\ 2,\ \ldots\ n).$$

Таким образом, элемент матрицы $C=AB$, расположенный в i-й строке и j-м столбце ее, равен сумме произведений элементов i-й строки матрицы A на соответствующие элементы j-го столбца матрицы B.

Примеры.

1. Дано:

$$A=\begin{pmatrix} 2 & 3 & 1 \\ -1 & 0 & 1 \end{pmatrix}, \quad B=\begin{pmatrix} 2 & 1 & -1 \\ 1 & 3 & -2 \\ 0 & 2 & 1 \end{pmatrix}.$$ Найти AB.

Решение:

$$AB=\begin{pmatrix} 2\cdot 2+3\cdot 1+1\cdot 0 & 2\cdot 1+3\cdot 3+1\cdot 2 & 2(-1)+3(-2)+1\cdot 1 \\ -1\cdot 2+0\cdot 1+1\cdot 0 & -1\cdot 1+0\cdot 3+1\cdot 2 & -1(-1)+0(-2)+1\cdot 1 \end{pmatrix}=$$
$$=\begin{pmatrix} 8 & 13 & -7 \\ -2 & 1 & 2 \end{pmatrix}$$

2. Дано:

$$A=\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}; \quad B=\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$ Найти AB и BA.

Решение:

$$AB=\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad BA=\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Из определения умножения видно, что в общем случае

$$AB \neq BA,$$

т. е. произведение двух матриц не обладает переместительным свойством. Действительно, если произведение AB имеет смысл, то из этого не следует, что имеет смысл и произведение BA. Даже, если оба эти произведения имеют смысл, то из этого не следует их равенство (см. также предыдущий пример).

Если окажется, что $AB=BA$, то матрицы A и B называются *перестановочными*, в противном случае (если $AB \neq BA$) они называются *неперестановочными*.

Например, матрицы

$$A=\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \cos\alpha & \cos\alpha \end{pmatrix}, \quad B=\begin{pmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{pmatrix}$$

перестановочны, так как

$$AB=BA=\begin{pmatrix} \cos(\alpha+\beta) & -\sin(\alpha+\beta) \\ \sin(\alpha+\beta) & \cos(\alpha+\beta) \end{pmatrix}.$$

Заметим, что произведение матрицы-столбца на матрицу-строку всегда имеет смысл:

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} (b_1 b_2 \ldots b_n) = \begin{pmatrix} a_1 b_1 & a_1 b_2 & \ldots & a_1 b_n \\ a_2 b_1 & a_2 b_2 & \ldots & a_2 b_n \\ \cdot & \cdot & \cdot & \cdot \\ a_m b_1 & a_m b_2 & \ldots & a_m b_n \end{pmatrix}.$$

С другой стороны, умножение матрицы-строки на матрицу-столбец может быть выполнено только при условии, что они содержат одинаковое количество элементов:

$$(a_1 a_2 \ldots a_n) \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = (a_1 b_1 + a_2 b_2 + \ldots + a_n b_n) =$$
$$= a_1 b_1 + a_2 b_2 + a_n b_n.$$

Отметим основные свойства произведения матриц (считая, конечно, что все написанные произведения имеют смысл):

а) $A0 = 0A = 0$;

б) $AE = EA = A$;

в) $(A + B) C = AC + BC$;

г) $A (B + C) = AB + AC$;

д) $(AB) C = A (BC)$;

е) если A и B — квадратные матрицы одного порядка, то $\det (AB) = \det A \det B$.

Свойства а — г достаточно очевидны. Их доказательство представляется читателю.

Для доказательства свойства «д» положим, что матрицы $A = (a_{ij})$, $B = (b_{ij})$ и $C = (c_{ij})$ имеют размеры соответственно $(m \times p)$, $(p \times q)$, $(q \times n)$.

Обозначая $(AB) C = D = (d_{ij})$ и $A (BC) = L = (l_{ij})$, имеем

$$d_{ij} = \sum_{s=1}^{q} \left(\sum_{k=1}^{p} a_{ik} b_{ks} \right) c_{sj} = \sum_{s=1}^{q} \sum_{k=1}^{p} a_{ik} b_{ks} c_{sj},$$

$$l_{ij} = \sum_{k=1}^{p} a_{ik} \left(\sum_{s=1}^{q} b_{ks} c_{sj} \right) = \sum_{k=1}^{p} \sum_{s=1}^{q} a_{ik} b_{ks} c_{sj}.$$

Но порядок суммирования не влияет на величину приведенных сумм, т. е.

$$\sum_{s=1}^{q}\sum_{k=1}^{p} a_{ik}b_{ks}c_{sj} = \sum_{k=1}^{p}\sum_{s=1}^{q} a_{ik}b_{ks}c_{sj}.$$

Следовательно,

$$d_{ij} = l_{ij}.$$

Что касается свойства «е», то оно является следствием правила умножения определителей, указанного во введении.

В самом деле, элементы определителя $\det(AB)$, согласно правилу умножения матриц, составляются так же, как элементы определителя, равного произведению определителей $\det A$ и $\det B$. Но это значит, что

$$\det(AB) = \det A \det B.$$

5. Транспонированная матрица. *Транспонированием матрицы* называется замена строк этой матрицы ее столбцами с сохранением их номеров. Матрица, полученная таким образом из матрицы A, называется транспонированной по отношению к матрице A и обозначается A'.

Согласно определению, если $A = (a_{ij})_{m,n}$, то

$$A' = (a'_{ij})_{n,m}, \text{ где } a'_{ij} = a_{ji}.$$

Например, если

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 1 \\ 2 & 4 \end{pmatrix}, \quad \text{то } A' = \begin{pmatrix} 1 & 3 & 2 \\ 2 & 1 & 4 \end{pmatrix}.$$

Очевидно, что для матрицы-строки транспонированной будет матрица-столбец и наоборот.

Может оказаться, что квадратная матрица A совпадет со своей транспонированной матрицей, т. е. $A = A'$, или $a_{ij} = a_{ji}$. В этом случае матрица A называется *симметрической*.

Операция транспонирования обладает следующими свойствами:

а) $(A + B)' = A' + B'$;

б) $(\alpha A)' = \alpha A'$;

в) $(AB)' = B'A'$.

Первые два свойства очевидны. Приведем доказательство третьего свойства.

Обозначим $A=(a_{ij})_{m,\,p}$, $B=(b_{ij})_{p,\,n}$, тогда

$$B'A'=\left(\sum_{k=1}^{p} b'_{ik}a'_{kj}\right)=\left(\sum_{k=1}^{p} b_{ki}a_{jk}\right).$$

С другой стороны,

$$(AB)'=\left(\sum_{k=1}^{p} a_{ik}b_{kj}\right)'=\left(\sum_{k=1}^{p} a_{jk}b_{ki}\right).$$

Отсюда

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

6. Обратная матрица. Если A — квадратная матрица, то матрица B такая, что

$$AB=BA=E$$

называется *обратной* относительно A и обозначается A^{-1}. Таким образом, по определению $AA^{-1}=A^{-1}A=E$. Чтобы определить условия, при которых матрица A имеет обратную матрицу, докажем следующую теорему.

Теорема. *Для того чтобы матрица A имела обратную матрицу, необходимо и достаточно, чтобы она была невырожденной (т. е. $\det A\neq 0$).*

Необходимость непосредственно следует из соотношений

$$\det(AA^{-1})=\det E \text{ или } \det A\det A^{-1}=1.$$

Покажем, что условие $\det A\neq 0$ достаточно для существования обратной матрицы.

Необходимо показать, что найдется матрица B такая, что

$$AB=BA=E.$$

Обозначим $A=(a_{ij})$, $B=(b_{ij})$, где элементы b_{ij} необходимо определить так, чтобы выполнялось предыдущее равенство.

Так как $AB=\left(\sum_{k=1}^{n} a_{ik}b_{kj}\right)$, то из условия $AB=E$ получаем систему n^2 уравнений с n^2 неизвестными b_{ij}:

$$\sum_{k=1}^{n} a_{ik}b_{kj}=\delta_{ij};\qquad \begin{matrix} i=1,\ 2,\ \ldots\ n \\ j=1,\ 2,\ \ldots\ n \end{matrix}$$

НАУЧНАЯ БИБЛИОТЕКА
им. Горького
МГУ

При фиксированном j имеем

$$a_{11}b_{1j}+a_{12}b_{2j}+\ldots+a_{1n}b_{nj}=0,$$
$$a_{21}b_{1j}+a_{22}b_{2j}+\ldots+a_{2n}b_{nj}=0,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$a_{j1}b_{1j}+a_{j2}b_{2j}+\ldots+a_{jn}b_{nj}=1,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$a_{n1}b_{1j}+a_{n2}b_{2j}+\ldots+a_{nn}b_{nj}=0.$$

Так как определитель этой системы $|A|=\det A\neq 0$, то по формулам Крамера (см. Введение п. 7), получаем

$$b_{kj}=\frac{|A_k|}{|A|},$$

где определитель $|A_k|$ получается из определителя системы заменой k-го столбца на столбец из свободных членов (в рассматриваемом случае j-й элемент этого столбца равен единице, а остальные элементы равны нулю). Раскрывая определитель $|A_k|$ по элементам k-го столбца, получаем

$$b_{kj}=\frac{A_{jk}}{|A|},$$

где A_{jk} — алгебраическое дополнение элемента a_{jk}. Аналогично можно показать, что из равенства $BA=E$ также получаем $b_{kj}=\frac{A_{jk}}{|A|}$. Таким образом, обратная матрица $B=A^{-1}$ существует и определяется равенством

$$A^{-1}=\frac{1}{\det A}\begin{pmatrix} A_{11} & A_{21} & \ldots & A_{n1} \\ A_{12} & A_{22} & \ldots & A_{n2} \\ \cdot & \cdot & \cdot & \cdot \\ A_{n} & A_{2n} & \ldots & A_{nn} \end{pmatrix},$$

или

$$A^{-1}=\frac{1}{\det A}(A_{ij})'.$$

Из доказательства следует, что обратная матрица единственная и что она может быть определена только одним равенством $AB=E$ или $BA=E$.

Теорема дает нам также правило для нахождения обратной матрицы. Чтобы найти обратную матрицу A^{-1}, надо построить вспомогательную матрицу, составленную из алгебраических дополнений элементов a_{ij} матрицы A, транспонировать ее и умножить на число $\frac{1}{\det A}$.

Примеры. Найти обратные матрицы для матриц

1. $A=\begin{pmatrix}2 & -1\\ 3 & 1\end{pmatrix}$, $\det A=\begin{vmatrix}2 & -1\\ 3 & 1\end{vmatrix}=5.$

$A_{11}=1,\ A_{12}=-3,\ \ A_{21}=1,\ \ A_{22}=2.$

Вспомогательная матрица имеет вид $\begin{pmatrix}1 & -3\\ 1 & 2\end{pmatrix}$,

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

2. $A=\begin{pmatrix}1 & 0 & 0\\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{pmatrix};\quad \det A=1.$

$$A_{11}=1;\ A_{12}=0;\qquad A_{13}=0,$$
$$A_{21}=0;\ A_{22}=\frac{1}{\sqrt{2}};\qquad A_{23}=\frac{1}{\sqrt{2}},$$
$$A_{31}=0;\ A_{32}=-\frac{1}{\sqrt{2}};\quad A_{33}=\frac{1}{\sqrt{2}},$$

$$A^{-1}=\begin{pmatrix}1 & 0 & 0\\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}\\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{pmatrix}.$$

Укажем еще, что из определения обратной матрицы

$$A^{-1}A=AA^{-1}=E$$

следует, что матрицы A и A^{-1} — взаимно обратные, т. е.

$$(A^{-1})^{-1}=A.$$

Кроме того,

$$\det A^{-1}=\frac{1}{\det A}.$$

Отметим также следующее тождество:

$$(AB)^{-1} = B^{-1}A^{-1},$$

где A и B — невырожденные матрицы одного и того же порядка. Действительно,

$$(AB)\,B^{-1}A^{-1} = A\,(BB^{-1})\,A^{-1} = AEA^{-1} = AA^{-1} = E.$$

Аналогично

$$B^{-1}A^{-1}\,(AB) = E.$$

Таким образом, матрица $B^{-1}A^{-1}$ — обратная для матрицы AB.

7. Степени квадратной матрицы. Так как две любые квадратные матрицы одного и того же порядка можно перемножить, то всякую квадратную матрицу можно умножить саму на себя, т. е. найти матрицу AA. Эта матрица называется *квадратом матрицы* A и обозначается A^2. Аналогично A^2A называется *кубом матрицы* A и обозначается A^3. Наконец, $A^{n-1}A$ называется *n-й степенью матрицы* A и обозначается A^n. Исходная матрица A называется *первой степенью матрицы* A и может быть записана, как A^1. Нулевой степенью матрицы A называется единичная матрица E, т. е. $A^0 = E$. Целая отрицательная степень матрицы A (что обозначается A^{-n}, $n > 0$) определяется соотношением

$$A^{-n} = (A^{-1})^n;$$

отсюда следует

$$A^{-n} = (A^n)^{-1},$$

где A^{-1} — матрица обратная для матрицы A.

Степени матрицы обладают следующими очевидными свойствами:

а) $A^kA^m = A^{k+m}$;

б) $(A^k)^m = A^{km}$.

Пример

Дано:

$$A = \begin{pmatrix} 1 & 2 \\ 3 & -1 \end{pmatrix}. \quad \text{Найти } A^2 \text{ и } A^3.$$

Решение.

$$A^2=\begin{pmatrix}1 & 2\\ 3 & -1\end{pmatrix}\begin{pmatrix}1 & 2\\ 3 & -1\end{pmatrix}=\begin{pmatrix}7 & 0\\ 0 & 7\end{pmatrix},$$

$$A^3=\begin{pmatrix}7 & 0\\ 0 & 7\end{pmatrix}\begin{pmatrix}1 & 2\\ 3 & -1\end{pmatrix}=\begin{pmatrix}7 & 14\\ 21 & -7\end{pmatrix}.$$

§ 3. КЛЕТОЧНЫЕ МАТРИЦЫ

1. Разбиение матрицы на клетки. Имеем прямоугольную матрицу

$$A=\begin{pmatrix}a_{11} & a_{12} & \dots & a_{1n}\\ a_{21} & a_{22} & \dots & a_{2n}\\ \dots & \dots & \dots & \dots\\ a_{m1} & a_{m2} & \dots & a_{mn}\end{pmatrix}.$$

Разобьем горизонтальными и вертикальными прямыми матрицу A на клетки A^{sp}, где $s=1, 2, \dots \mu$; $p=1, 2, \dots \nu$. (Здесь введены верхние индексы, так как ранее мы обозначали символом A_{sp} алгебраическое дополнение элемента a_{sp}.)

Клетки A^{sp} являются матрицами размеров $m_s \times n_p$. Матрицу A можно сокращенно записать так

$$A=\begin{pmatrix}A^{11} & A^{12} & \dots & A^{1\nu}\\ A^{21} & A^{22} & \dots & A^{2\nu}\\ \dots & \dots & \dots & \dots\\ A^{\mu 1} & A^{\mu 2} & \dots & A^{\mu\nu}\end{pmatrix}, \text{ или } A=(A^{sp})_{\mu,\nu}.$$

Матрица A, записанная в таком виде, называется *клеточной* (*или блочной*), матрицы A^{sp} — *ее элементами*. Очевидно, что матрицу A можно разбивать на клетки различными способами. Например:

$$A=\begin{pmatrix}a_{11} & a_{12} & a_{13} & a_{14}\\ a_{21} & a_{22} & a_{23} & a_{24}\\ a_{31} & a_{32} & a_{33} & a_{34}\end{pmatrix}.$$

Разобьем ее на клетки

$$A=\left(\begin{array}{c:ccc} a_{11} & a_{12} & a_{13} & a_{14} \\ \hdashline a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{array}\right), \text{ или } A=\begin{pmatrix} A^{11} & A^{12} \\ A^{21} & A^{22} \end{pmatrix},$$

где

$$A^{11}=a_{11};\quad A^{12}=(a_{12}a_{13}a_{14});$$

$$A^{21}=\begin{pmatrix} a_{21} \\ a_{31} \end{pmatrix};\quad A^{22}=\begin{pmatrix} a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \end{pmatrix}.$$

Разбивая иначе, имеем

$$A=\left(\begin{array}{cc:c:c} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ \hdashline a_{31} & a_{32} & a_{33} & a_{34} \end{array}\right),$$

или

$$A=\begin{pmatrix} A^{11} & A^{12} & A^{13} \\ A^{21} & A^{22} & A^{23} \end{pmatrix},$$

где

$$A^{11}=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix};\quad A^{12}\begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix};\quad A^{13}=\begin{pmatrix} a_{14} \\ a_{24} \end{pmatrix};$$

$$A^{21}=(a_{31}\ a_{32});\quad A^{22}=a_{33};\quad A^{33}=a_{34}.$$

2. Квазидиагональная матрица. В частном случае, когда $\mu=\nu$ и $A^{sp}=0$ при $s\neq p$, клеточная матрица называется *квазидиагональной* (или *клеточно-диагональной*). В этом случае матрица A имеет вид:

$$A=\begin{pmatrix} A^{11} & & O \\ & \ddots & \\ O & & A^{\mu\mu} \end{pmatrix}.$$

Например:

матрица $A=\begin{pmatrix} 1 & 2 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ — квазидиагональная. Очевидно, если клетки A^{11}, A^{2a}, ..., $A^{\mu\mu}$ квадратные, то и матрица A — квадратная.

Примером квазидиагональной матрицы является так называемая Жорданова матрица I, которая имеет вид

$$I=\begin{pmatrix} I_1 & & & O \\ & I_2 & & \\ & & \ddots & \\ O & & & I_n \end{pmatrix},$$

где

$$I_i = \begin{pmatrix} \lambda_i & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_i & 1 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & \lambda_i & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda_i \end{pmatrix}.$$

3. Действия над клеточными матрицами. Клеточные матрицы удобны тем, что действия над ними производятся формально по тем же правилам, что и над обычными матрицами. Покажем это, предполагая, что соответствующие операции над матрицами выполнимы (см. § 2 этой главы).

Сложение матриц. Имеем матрицы $A = (a_{ij})_{m,n}$ и $B = (b_{ij})_{m,n}$ одинаковых размеров. Разобьем их одинаковым образом на клетки A^{sp} и B^{sp} (т. е. размеры клеток A^{sp} и B^{sp} одинаковы). Тогда очевидно

$$A + B = (A^{sp} + B^{sp})_{\mu,\nu} = (a_{ij} + b_{ij})_{m,n}.$$

Умножение матрицы на число. Если A — клеточная матрица и λ — число, то очевидно

$$\lambda A = (\lambda A^{sp})_{\mu,\nu} = (\lambda a_{ij})_{m,n}.$$

Умножение матриц. Имеем матрицы $A = (a_{ij})_{m,q}$ и $B = (b_{ij})_{q,n}$. В этом случае произведение AB определено. Разобьем матрицу A произвольным образом на клетки. Матрицу B разобьем на клетки так, чтобы их вертикальные размеры в столбце совпадали с соответствующими горизонтальными размерами клеток в строке матрицы A.

При таком разбиении произведение клеток $A^{sk}B^{kp}$ имеет смысл (читатель легко это может проверить). Поэтому

$$AB = \left(\sum_{k=1}^{\nu} A^{sk}B^{kp}\right) = \left(\sum_{k=1}^{q} a_{ik}b_{kj}\right).$$

В частности, если матрицы A и B квазидиагональные, т. е.

$$A = \begin{pmatrix} A^{11} & & O \\ & \ddots & \\ O & & A^{\nu\nu} \end{pmatrix}, \quad B = \begin{pmatrix} B^{11} & & O \\ & \ddots & \\ O & & B^{\nu\nu} \end{pmatrix},$$

то

$$AB = \begin{pmatrix} A^{11}B^{11} & & O \\ & \ddots & \\ O & & A^{\nu\nu}B^{\nu\nu} \end{pmatrix}.$$

Отсюда следует, что если $A=\begin{pmatrix} A^{11} & & 0 \\ & \ddots & \\ 0 & & A^{\nu\nu} \end{pmatrix}$, то

$$A^n=\begin{pmatrix} (A^{11})^n & & 0 \\ & \ddots & \\ 0 & & (A^{\nu\nu})^n \end{pmatrix}.$$

Пример.

Дано:

$$A=\begin{pmatrix} 5 & 2 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 0 & 0 & 8 & 3 \\ 0 & 0 & 5 & 2 \end{pmatrix}; \quad B=\begin{pmatrix} 1 & -2 & -1 & 0 & 0 \\ -2 & 5 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & -3 \\ 0 & 0 & 0 & -5 & 8 \end{pmatrix}.$$

Найти AB.

Решение.

Заменяя данные матрицы клеточными матрицами

$$A=\begin{pmatrix} A^{11} & 0 \\ 0 & A^{22} \end{pmatrix}; \quad B=\begin{pmatrix} B^{11} & 0 \\ 0 & B^{22} \end{pmatrix},$$

где

$$A^{11}=\begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}; \quad A^{22}=\begin{pmatrix} 8 & 3 \\ 5 & 2 \end{pmatrix};$$

$$B^{11}=\begin{pmatrix} 1 & -2 & -1 \\ -2 & 5 & 2 \end{pmatrix}; \quad B^{22}=\begin{pmatrix} 2 & -3 \\ -5 & 8 \end{pmatrix},$$

имеем

$$A\cdot B=\begin{pmatrix} A^{11}B^{11} & 0 \\ 0 & A^{22}B^{22} \end{pmatrix}$$

и так как

$$A^{11}B^{11}=\begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 1 & -2 & -1 \\ -2 & 5 & 2 \end{pmatrix}=\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

$$A^{22}B^{22}=\begin{pmatrix} 8 & 3 \\ 5 & 2 \end{pmatrix}\begin{pmatrix} 2 & -3 \\ -5 & 8 \end{pmatrix}=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

окончательно получаем

$$AB=\begin{pmatrix} 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Упражнения

1. Вычислить произведение матриц

а) $\begin{pmatrix} 5 & 8 & -4 \\ 6 & 9 & -5 \\ 4 & 7 & -3 \end{pmatrix}\begin{pmatrix} 3 & 2 & 5 \\ 4 & -1 & 3 \\ 9 & 6 & 5 \end{pmatrix}$,

б) $\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 & 5 & -3 \\ 1 & -1 & 0 & 1 \end{pmatrix}$,

в) $\begin{pmatrix} 4 & 3 \\ 7 & 5 \end{pmatrix}\begin{pmatrix} -28 & 93 \\ 38 & -126 \end{pmatrix}\begin{pmatrix} 7 & 3 \\ 2 & 1 \end{pmatrix}$.

2. Найти A^3, если $A=\begin{pmatrix} 1 & -2 \\ 3 & -4 \end{pmatrix}$.

3.

$A=\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$.

Найти:

а) A^2; б) A^3; в) A^n.

4. Найти A^5, если $A=\begin{pmatrix} 4 & -1 \\ 5 & -2 \end{pmatrix}$.

5.

$A=\begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$. Найти A^n.

6.

$A=T^{-1}BT$.

Показать, что $A^n=T^{-1}B^nT$.

7. Вычислить $\begin{pmatrix} 17 & -6 \\ 35 & -12 \end{pmatrix}^5$, если известно, что

$$\begin{pmatrix} 17 & -6 \\ 35 & -12 \end{pmatrix}=\begin{pmatrix} 2 & 3 \\ 5 & 7 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} -7 & 3 \\ 5 & -2 \end{pmatrix}.$$

8. Найти все матрицы второго порядка, квадрат которых равен:

а) нулевой матрице;

б) единичной матрице.

9. Найти обратные матрицы для матриц

а) $\begin{pmatrix} 3 & 4 \\ 5 & 7 \end{pmatrix}$,

б) $\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$,

в) $\begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{pmatrix}$.

Используя формулу $AA^{-1}=E$, проверить ответы.

10. Доказать, что если квадратные матрицы A и B перестановочны, т. е. $AB=BA$, то

а) $(AB)^n=A^nB^n$,

б) $(A+B)^n=A^n+nA^{n-1}B+\frac{n(n-1)}{2}A^{n-2}B+\ldots+B^n$.

Указание: применить метод полной математической индукции.

Глава 2

РЕШЕНИЕ СИСТЕМЫ ЛИНЕЙНЫХ УРАВНЕНИЙ

§ 1. ФОРМУЛЫ КРАМЕРА

1. Основные понятия. Системой из m линейных уравнений с n неизвестными называются соотношения вида

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = h_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = h_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = h_m. \end{cases} \tag{2.1}$$

Здесь x_i — неизвестные, подлежащие определению; a_{ij} — числа, называемые коэффициентами при неизвестных; h_i — числа, называемые свободными членами (или правыми частями) системы уравнений.

Решением системы уравнений (2.1) называется совокупность таких чисел $x_1 = \alpha_1$, $x_2 = \alpha_2$, ..., $x_n = \alpha_n$, которые обращают все уравнения системы (2.1) в тождества.

Если система линейных уравнений имеет хотя бы одно решение, то она называется *совместной*. В противном случае она называется *несовместной*.

Совместная система, имеющая единственное решение, называется *определенной*, а система, имеющая более одного решения, — *неопределенной*. Две системы линейных уравнений называются *эквивалентными*, если любое решение каждой из них является одновременно решением и другой системы. Две произвольные несовместные системы считаются эквивалентными.

2. Матричная запись системы. Составим из коэффициентов a_{ij} системы (2.1) матрицу размеров $(m \times n)$

$$A = (a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix},$$

которую назовем матрицей системы. Введем еще две

матрицы

$$X=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad H=\begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_m \end{pmatrix}.$$

Тогда систему (2.1) можно кратко представить в матричном виде

$$AX=H.$$

3. Решение системы с помощью обратной матрицы. Прежде чем исследовать общую систему линейных уравнений (при произвольных m и n), мы рассмотрим важнейший частный случай, когда число уравнений равно числу неизвестных ($m=n$). Итак, имеем

$$AX=H, \tag{2.2}$$

где $A=(a_{ij})$ — квадратная матрица n-го порядка. Покажем, что если матрица A невырожденная, то система имеет единственное решение. Так как по условию матрица A имеет обратную, то, умножая обе части равенства (2.2) слева на матрицу A^{-1}, получаем

$$A^{-1}AX=A^{-1}H.$$

Но $A^{-1}A=E$, а $EX=X$. Следовательно,

$$X=A^{-1}H. \tag{2.3}$$

Формулы Крамера. Запишем равенство (2.3) в развернутом виде

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}=\frac{1}{\det A}\begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix}\cdot\begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{pmatrix}.$$

Отсюда непосредственно следуют формулы Крамера, с которыми мы встречались во введении:

$$x_i=\frac{1}{\det A}\sum_{k=1}^{n} A_{ki}h_k=\frac{|A_i|}{|A|}, \quad i=1,\ 2,\ \dots,\ n.$$

Очевидно, решение системы по формуле (2.3) достаточно трудоемкое, так как нахождение обратной матрицы

при больших n затруднительно. Например, уже при $n=5$ необходимо вычислить 25 определителей 4-го порядка. Однако в приложениях часто возникает задача решить систему с неизменной левой частью при различных правых частях. Такая задача возникает при расчете конструкции, когда требуется найти нагрузки в узлах, при условии, что нагрузка на конструкцию меняется. Аналогичная задача возникает в электротехнике, когда требуется найти напряжение в некоторых элементах при различных входных напряжениях. В этом случае целесообразно вычислить обратную матрицу, после чего для различных H находить решения по формуле (2.3).

§ 2. МЕТОД ГАУССА

Простые по своей структуре формулы Крамера требуют проведения достаточно длинных и громоздких (при большом n) вычислений, связанных с определением алгебраических дополнений A_{ij}. Для практического решения системы (2.1) более удобным является способ (алгоритм) Гаусса, основанный на последовательном исключении неизвестных.

1. Элементарные преобразования системы. Имеем систему (2.1). Будем над уравнениями системы производить следующие преобразования:

1) умножение обеих частей одного из уравнений на произвольное число $\lambda \neq 0$;

2) прибавление к обеим частям одного из уравнений соответствующих частей любого другого уравнения, умноженного на произвольное число α;

3) перестановка любых двух уравнений.

Очевидно, что в результате каждого из этих преобразований, называемых элементарными, линейная система уравнений (2.1) перейдет в линейную систему уравнений, эквивалентную первоначальной системе (2.1).

2. Способ Гаусса. Имеем систему (2.1). Применяя элементарные преобразования, построим эквивалентную ей систему специального вида. Выберем в качестве первого уравнения одно из тех уравнений системы, где коэффициент при x_1 отличен от нуля. Не нарушая общности, предположим, что $a_{11} \neq 0$.

Итак, первым уравнением системы будет уравнение

$$a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = h_1, \quad \text{где } a_{11} \neq 0.$$

Умножим это уравнение на $-\frac{a_{i1}}{a_{11}}$, $(i=2, 3, \ldots, n)$ и прибавим его почленно к уравнениям системы (2.1) с номерами $i=2$, $i=3$, ..., $i=n$. После этой операции в уравнениях с номерами $i>1$ будет исключено неизвестное x_1. Может случиться, что вместе с x_1 будут исключены неизвестные x_2, x_3, ..., x_{k_2-1}, но найдется хоть одно уравнение, в котором сохранится неизвестное x_{k_2}. Одно из таких уравнений примем в качестве второго уравнения системы.

Таким образом, после первого шага второе уравнение будет иметь вид

$$a^{(1)}_{2,k_2}x_{k_2}+a^{(1)}_{2,k_2+1}x_{k_2+1}+\ldots+a^{(1)}_{2,n}x_n=h^{(1)}_2, \text{ где } a^{(1)}_{2,k_2}\neq 0.$$

Используем это уравнение для исключения неизвестного x_{k_2} из всех оставшихся уравнений (кроме 1-го уравнения). После 2-го шага в качестве третьего уравнения получим

$$a^{(2)}_{3,k_3}x_{k_3}+a^{(2)}_{3,k_3+1}x_{k_3+1}+\ldots+a^{(2)}_{3,n}x_n=h^{2}_{3,k_n}, \text{ где } a^{(2)}_{3,k_3}\neq 0.$$

Продолжая процесс, мы после r шагов приведем систему (2.1) к эквивалентной ей системе следующего вида:

$$\begin{cases} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=h_1, \\ a^{(1)}_{2,k_2}x_{k_2}+a^{(1)}_{2,k_2+1}x_{k_2+1}+\ldots+a^{(1)}_{2,n}x_n=h^{(1)}_2, \\ a^{(2)}_{3,k_3}x_{k_3}+a^{(2)}_{3,k_3+1}x_{k_3+1}+\ldots+a^{(2)}_{3,n}x_n=h^{(2)}_3, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a^{(r-1)}_{r,k_r}x_{k_r}+a^{(r-1)}_{r,k_r+1}x_{k_r+1}+\ldots+a^{(r-1)}_{r,n}x_n=h^{(r-1)}_r, \\ \qquad\qquad 0=h^{(r-1)}_{(r+1)}, \\ \qquad\qquad 0=h^{(r-1)}_{r+2}, \\ \qquad\qquad \cdots\cdots \\ \qquad\qquad 0=h^{(r-1)}_m, \end{cases}$$

где $a_{11}\neq 0$, $a_{2,k_2}\neq 0$, ..., $a_{r,k_r}\neq 0$. Очевидно, число шагов $r\leqslant m$. Если $r=m$, то все последние соотношения типа $0=h^{(r-1)}_j$ отсутствуют.

Что касается чисел $h^{(r-1)}_j$ при $j>r$, то они могут быть либо равны нулю, либо отличаться от нуля. Если хоть одно из этих чисел не равно нулю (пусть, например $h^{(r-1)}_{r+1}\neq 0$), то, умножая соотношение $0=h^{(r-1)}_{r+1}$ на $-\frac{h^{(r-1)}_j}{h^{(r-1)}_{r+1}}$

и складывая его с соответствующими оставшимися соотношениями этого типа, приведем их к виду $0=0$.

Итак, в самом общем случае система уравнений (2.1) приводится к эквивалентной системе специального вида (так называемой ступенчатой системе), в которой для удобства письма используем другие обозначения коэффициентов и правых частей

$$\begin{cases} b_{1,1}x_1+\ldots+b_{1,k_2}x_{k_2}+\ldots+b_{1,k_r}x_{k_r}+\ldots+b_{1,n}x_n=l_1, \\ \qquad\qquad b_{2,k_2}x_{k_2}+\ldots+b_{2,k_r}x_{k_r}+\ldots+b_{2,n}x_n=l_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \qquad\qquad\qquad\qquad b_{r,k_r}x_{k_r}+\ldots+b_{r,n}x_n=l_r, \\ \qquad\qquad\qquad\qquad\qquad\qquad 0=l_{r+1}, \\ \qquad\qquad\qquad\qquad\qquad\qquad 0=0, \\ \qquad\qquad\qquad\qquad\qquad\qquad \ldots \\ \qquad\qquad\qquad\qquad\qquad\qquad 0=0, \end{cases} \tag{2.4}$$

где $1<k_2<k_3<\ldots<k_r$; b_{11}, b_{2,k_2}, ..., b_{rk_r} отличны от нуля, а все остальные b_{ij} — произвольны; l_1, l_2, ..., $l_r l_{r+1}$ — произвольны.

3. **Совместность и несовместность системы.** Ступенчатая система (2.4) дает возможность сразу выяснить совместна или несовместна данная система уравнений (2.1). Кроме того, в случае совместности можно легко определить все решения системы.

Возможны два случая.

1. $l_{r+1}\neq 0$. В этом случае соотношение $0=l_{r+1}$ невозможно. Следовательно, система (2.4), а значит и эквивалентная ей система (2.1) несовместны.

Поэтому на практике вывод о несовместности системы можно сделать сразу же при появлении соотношения вида $0=h$, где $h\neq 0$.

2. $l_{r+1}=0$. Покажем, что в этом случае система (2.1) совместна.

Из уравнения системы (2.4) с номером r находим x_{k_r}, выражая его через x_{k_r+1}, x_{k_r+2}, ..., x_n, что всегда возможно, так как $b_{r,k_r}\neq 0$. Подставляя найденное значение x_{k_r} в предыдущее уравнение (что не изменяет коэффициента при x_{k_r-1}) находим x_{k_r-1}. Продолжая процесс, найдем последовательно x_{k_r-2}, ..., x_1.

В результате получим соотношения вида

$$\begin{cases} x_1 = \alpha_1 + \sum\limits_{(\nu)} \beta_{1\nu} x_\nu, \\ x_{k_2} = \alpha_{k_2} + \sum\limits_{(\nu)} \beta_{k_2,\nu} x_\nu, \\ \cdots\cdots\cdots\cdots \\ x_{k_r} = \alpha_{k_r} + \sum\limits_{(\nu)} \beta_{k_r,\nu} x_\nu, \end{cases} \tag{2.5}$$

где индекс ν принимает все значения $\nu \leqslant n$ за исключением значений $1, k_2, k_3, \ldots, k_r$.

Из соотношений (2.5) следует, что величинам x_ν, стоящим в правых частях равенств, можно придавать произвольные значения. Эти неизвестные x_ν ($\nu \neq 1, k_2, \ldots, k_r$) называются *свободными неизвестными*.

Неизвестные $x_1, x_{k_2}, \ldots, x_{k_r}$, которые однозначно определяются свободными неизвестными, называются *главными* или *основными*.

Любая совокупность значений свободных неизвестных и соответствующих им главных неизвестных будет решением системы (2.4), а значит и системы (2.1).

Итак, для совместности системы необходимо и достаточно, чтобы

$$l_{r+1} = 0.$$

4. Определенность и неопределенность системы. Пусть система (2.4) совместна (т. е. $l_{r+1} = 0$). Возможны два случая:

1) $r = n$ (что может быть только, когда $m \geqslant n$). В этом случае все неизвестные главные. Соотношения (2.5) принимают вид

$$x_1 = \alpha_1, \ x_2 = \alpha_2, \ \ldots, \ x_n = \alpha_n.$$

Значит, система (2.1) имеет единственное решение;

2) $r < n$. В этом случае имеются свободные неизвестные, которые можно задавать произвольно. Значит, система имеет бесконечное множество решений, общий вид которых определяется соотношениями (2.5).

Система (2.1) — неопределенная.

5. Общий вывод. Подводя итоги проведенному исследованию, приходим к следующему результату. Если после приведения системы (2.1) к ступенчатому виду (2.4)

1) $l_{r+1} \neq 0$, то система (2.1) — несовместна;

2) $l_{r+1} = 0$ и $r = n$, то система (2.1) имеет единственное решение;

3) $l_{r+1} = 0$ и $r < n$, то система (2.1) имеет бесконечное множество решений.

Пример.

Решить систему

$$\begin{cases} 2x_1 + x_2 - x_3 + x_4 = 3, \\ 4x_1 + 2x_2 + 5x_4 + x_5 = 5, \\ 2x_1 + x_2 - x_3 + 2x_4 + x_5 = 6, \\ 2x_1 + x_2 - x_3 + x_4 + 2x_5 = 5, \\ 4x_1 + 2x_2 - 2x_3 + 3x_4 + 3x_5 = 11. \end{cases}$$

Решение. Исключаем x_1

$$\begin{cases} 2x_1 + x_2 - x_3 + x_4 = 3, \\ 2x_3 + 3x_4 + x_5 = -1, \\ x_4 + x_5 = 3, \\ 2x_5 = 2, \\ x_4 + 3x_5 = 5. \end{cases}$$

Исключаем x_4 из последнего уравнения

$$\begin{cases} 2x_1 + x_2 - x_3 + x_4 = 3, \\ 2x_3 + 3x_4 + x_5 = -1, \\ x_4 + x_5 = 3, \\ x_5 = 1, \\ x_5 = 1. \end{cases}$$

Отсюда $x_5 = 1$; $x_4 = 2$; $x_3 = -4$; $x_1 = -\frac{3}{2} - \frac{1}{2}x_2$, где x_2 — свободное неизвестное.

§ 3. ЭКВИВАЛЕНТНЫЕ МАТРИЦЫ

В рассмотренном выше алгоритме Гаусса все операции производились непосредственно над уравнениями. Покажем, что все выкладки можно упростить, если рассмотреть соответствующие преобразования матрицы системы. Для этого исследуем изменение матрицы при элементарных преобразованиях системы.

1. Элементарные преобразования матрицы. Назовем *элементарными преобразованиями матрицы* следующие операции:

1. Умножение всех элементов некоторой строки матрицы на число $\lambda \neq 0$.

2. Прибавление к элементам некоторой строки матрицы соответствующих элементов другой строки, умноженных на произвольное число α.

3. Перемена местами строк матрицы.

Матрицы, полученные одна из другой при помощи конечного числа элементарных преобразований, называются *эквивалентными*, что обозначается так $A \backsim B$. Заметим, что элементарные преобразования обратимы. Это значит, что если матрица B получена из матрицы A с помощью конечного числа элементарных преобразований, то и матрица B в свою очередь может быть снова преобразована в матрицу A при помощи конечного числа элементарных преобразований. Для этого достаточно всю последовательность преобразований, выполняемых над матрицей A, применить в обратном порядке, заменяя при этом каждое из преобразований ему обратным, что означает следующее:

1. Если B получается из A умножением некоторой строки на число $\lambda \neq 0$, то, умножая соответствующую строку B на число $\frac{1}{\lambda}$, получим исходную матрицу A.

2. Если B получается из A прибавлением к i-й строке элементов j-й строки, умноженных на число α, то, добавляя к элементам i-й строки B элементы ее j-й строки, умноженные на $-\alpha$, получим A.

3. Если B получается из A перестановкой i-й и j-й строк, то, переставляя в B те же i-ю и j-ю строки, получим A.

Обратим внимание на то, что элементарное преобразование 3-го типа может быть осуществлено применением преобразований только первых двух типов.

Действительно, осуществить перестановку i-й и j-й строк можно при помощи следующих последовательно выполненных преобразований.

1. К элементам i-й строки прибавляем соответствующие элементы j-й строки.

2. Новую i-ю строку умножаем на -1 и прибавляем к j-й строке.

3. Преобразованную j-ю строку прибавляем к i-й строке.

4. Преобразованную j-ю строку умножаем на -1.

Сказанное проиллюстрируем схемой

$$\begin{pmatrix} i \\ j \end{pmatrix} \sim \begin{pmatrix} i+j \\ j \end{pmatrix} \sim \begin{pmatrix} i+j \\ -i \end{pmatrix} \sim \begin{pmatrix} j \\ -i \end{pmatrix} \sim \begin{pmatrix} j \\ i \end{pmatrix}.$$

2. Использование эквивалентных матриц в алгоритме Гаусса. Имеем систему (2.1). Рассмотрим матрицу системы

$$A = \begin{pmatrix} a_{11} a_{12} & \dots & a_{1n} \\ a_{21} a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots \\ a_{m1} a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

Составим матрицу B, получающуюся из A добавлением столбца из свободных членов системы (2.1),

$$B = \begin{pmatrix} a_{11} a_{12} & \dots & a_{1n} h_1 \\ a_{21} a_{22} & \dots & a_{2n} h_2 \\ \dots & \dots & \dots \\ a_{m1} a_{m2} & \dots & a_{mn} h_m \end{pmatrix}.$$

Матрица B называется *расширенной матрицей системы.*

Легко видеть, что каждому элементарному преобразованию системы (2.1) соответствует аналогичное элементарное преобразование над матрицей B. Справедливо и обратное. Поэтому вместо того, чтобы производить в алгоритме Гаусса элементарные операции над системой (2.1), нам достаточно проделать соответствующие элементарные преобразования матрицы B. Составим матрицу системы (2.4)

$$C = \begin{pmatrix} b_{11} & \dots & \dots & b_{1n} & l_1 \\ 0 & \dots & b_{2,k_2} \dots & b_{2n} & l_2 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & b_{r,k_r} \dots & b_{rn} & l_r \\ 0 & \dots & \dots & 0 & l_{r+1} \\ 0 & \dots & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & 0 & 0 \end{pmatrix}$$

Из сказанного следует, что $C \sim B$.

Пример.

Дано:

$$\begin{cases} 5x_1 - x_2 + 2x_3 + x_4 = 7, \\ 2x_1 + x_2 + 4x_3 - 2x_4 = 1, \\ x_1 - 3x_2 - 6x_3 + 5x_4 = 0. \end{cases}$$

Определить, является ли система совместной.

Решение. Составим расширенную матрицу и приведем ее к ступенчатому виду. Для удобства вычислений переставляем местами уравнения

$$\begin{pmatrix} 1 & -3 & -6 & 5 & 0 \\ 2 & 1 & 4 & -2 & 1 \\ 5 & -1 & 2 & 1 & 7 \end{pmatrix} \sim \begin{pmatrix} 1 & -3 & -6 & 5 & 0 \\ 0 & 7 & 16 & 12 & 1 \\ 0 & 14 & 32 & -24 & 7 \end{pmatrix} \sim \begin{pmatrix} 1 & -3 & -6 & 5 & 0 \\ 0 & 7 & 16 & -12 & 1 \\ 0 & 0 & 0 & 0 & 5 \end{pmatrix}.$$

Данная система несовместна.

§ 4. РАНГ МАТРИЦЫ

В алгоритме Гаусса мы последовательным применением элементарных преобразований переходим от расширенной матрицы B системы (2.1) к расширенной матрице C системы (2.4). Причем $B \sim C$. Эквивалентность этих матриц дает возможность сделать важный вывод о совместности системы (2.1). Для этого предварительно нам придется ввести одну характеристику матрицы, инвариантную относительно элементарных преобразований.

1. Минор матрицы. Дана матрица

$$\begin{pmatrix} a_{11} a_{12} & \dots & a_{1n} \\ a_{21} a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots \\ a_{m1} a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

Определитель k-го порядка

$$\begin{vmatrix} a_{i_1 j_1} a_{i_1 j_2} & \dots & a_{i_1 j_k} \\ a_{i_2 j_1} a_{i_2 j_2} & \dots & a_{i_2 j_k} \\ \dots & \dots & \dots \\ a_{i_k j_1} a_{i_k j_2} & \dots & a_{i_k j_k} \end{vmatrix},$$

составленный из элементов $a_{i_\nu j_\nu}$ матрицы, расположенных на пересечении некоторых ее k-строк $(i_1, i_2, \dots, i_k)$ и некоторых k столбцов $(j_1, j_2, \dots, j_k)$, называется *минором* k-го порядка этой матрицы.

Очевидно, $k \leqslant \min(m, n)$.

Например для матрицы $\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \\ d_1 & d_2 & d_3 \end{pmatrix}$

b_2 — минор 1-го порядка,

$\begin{vmatrix} b_1 & b_3 \\ d_1 & d_3 \end{vmatrix} \begin{vmatrix} a_1 & a_2 \\ c_1 & c_2 \end{vmatrix}$ — миноры 2-го порядка,

$\begin{vmatrix} b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \\ d_1 & d_2 & d_3 \end{vmatrix}$ — минор 3-го порядка.

Легко подсчитать, что, например, различных миноров 3-го порядка здесь будет 4.

2. Ранг матрицы. Если у матрицы все миноры порядка $k > r$ равны нулю, а среди миноров ее порядка r имеется хоть один, отличный от нуля, то число r называется *рангом матрицы*. Очевидно, что ранг невырожденной матрицы равен ее порядку. Вычисление ранга матрицы, непосредственно следующее из определения, является слишком громоздким. Используя элементарные преобразования матрицы, можно значительно проще вычислить ее ранг.

3. Неизменность ранга матрицы при элементарных преобразованиях ее.

Теорема. *Эквивалентные матрицы имеют один и тот же ранг.*

Для доказательства теоремы надо показать, что элементарные преобразования матрицы не меняют ее ранга. Для этого (см. § 3 п. 1) достаточно рассмотреть только преобразования первых двух типов. Пусть матрица A имеет ранг r. Это значит, что все ее миноры порядка, бо́льшего r, равны нулю, а среди миноров порядка r есть хоть один, отличный от нуля.

Если какую-либо строку матрицы A умножить на число $\lambda \neq 0$, то при этом любой ее минор окажется умноженным на λ (если он содержит элементы этой строки) или останется неизменным (в противном случае). Но это значит, что все миноры, равные нулю, останутся равными нулю, а все ненулевые миноры останутся ненулевыми.

Таким образом, применение 1-го элементарного преобразования не изменяет ранга матрицы.

Рассмотрим второе элементарное преобразование — прибавление к элементам строки матрицы соответствующих элементов другой строки, умноженных на число α. Очевидно, эта операция не изменит миноров эквивалентной матрицы B, которые не содержат измененную строку. Рассмотрим теперь произвольный минор ν_s^k порядка k матрицы B, содержащей измененную строку. На основании свойства определителей (см. введение п. 4) имеем

$$\nu_s^{(k)} = \mu_s^{(k)} + \alpha\mu_i^{(k)}, \text{ где } \mu_s^{(k)} \text{ и } \mu_i^{(k)} \text{ — миноры матрицы } A.$$

Отсюда следует, что если все миноры k-го порядка матрицы A были равны нулю, то и все миноры k-го порядка матрицы B также равны нулю. Следовательно, это преобразование не может привести к увеличению ранга матрицы. Итак, $r_B \leqslant r_A$. Как известно, элементарные преобразования обратимы, поэтому, совершая над матрицей B обратное преобразование, получим

$$r_A \leqslant r_B.$$

Таким образом, $r_B = r_A$. Из доказанного следует, что любое конечное число элементарных преобразований матрицы не изменяет ее ранга, что и требовалось доказать.

В заключение отметим, что преобразования 1, 2, 3 (§ 3 п. 1) можно ввести и для столбцов матрицы. Так как при транспонировании ранг матрицы не изменяется, то все, что было доказано для элементарных преобразований строк, сохраняет силу и для элементарных преобразований столбцов.

4. **Ранг ступенчатой матрицы.** Рассмотрим ступенчатую матрицу размеров $(m \times n)$

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & \cdots & a_{1n} \\ 0 & \cdots & a_{2,i_2} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & a_{r,i_r} & \cdots & a_{rn} \\ 0 & \cdots & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & & 0 & & 0 \end{pmatrix}$$

где $a_{11}, a_{2,i_2}, \ldots, a_{r,i_r}$ отличны от нуля. Минор матрицы порядка r

$$\begin{vmatrix} a_{11} & a_{1,i_2} & \ldots & a_{1,i_2} \\ 0 & a_{2,i_2} & \ldots & a_{2,i_r} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & a_{r,i_r} \end{vmatrix} = a_{11}\, a_{2,i_2} \ldots a_{r,i_r} \neq 0.$$

Матрица содержит только r ненулевых строк ($r \leqslant \min(m, n)$). Поэтому всякий минор ее порядка $k > r$ содержит хотя бы одну нулевую строку и равен нулю.

Следовательно, ранг ступенчатой матрицы равен r, т. е. числу ее ненулевых строк.

Простота вычисления ранга ступенчатой матрицы дает возможность сравнительно быстро подсчитать ранг произвольной матрицы. Для этого достаточно привести ее с помощью элементарных преобразований к ступенчатому виду и подсчитать число ненулевых строк.

Пример.

Дано:

$$A = \begin{pmatrix} 1 & 3 & 0 \\ 2 & 1 & -1 \\ 3 & 4 & -1 \end{pmatrix}.$$

Найти ранг.

Решение:

$$A \infty \begin{pmatrix} 1 & 3 & 0 \\ 0 & -5 & -1 \\ 0 & -5 & -1 \end{pmatrix} \infty \begin{pmatrix} 1 & 3 & 0 \\ 0 & -5 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

Следовательно, $r_A = 2$.

§ 5. КРИТЕРИЙ СОВМЕСТНОСТИ СИСТЕМЫ

Вернемся к системе (2.1). С помощью алгоритма Гаусса мы провели полное исследование этой системы и нашли ее решение (в случае совместности). Используем это для вывода условий совместности системы.

Теорема Кронекера — Капелли. *Для того чтобы система линейных уравнений* (2.1) *была совместна, необходимо и достаточно, чтобы ранги матрицы этой системы и ее расширенной матрицы были одинаковы.* Мы обоз-

начили A матрицу системы (2.1), B — ее расширенную матрицу, C — расширенную матрицу системы (2.4).

Составим еще матрицу системы (2.4)

$$D=\begin{vmatrix} b_{11} & \dots & \dots & \dots & b_{1,n} \\ 0 & \dots & b_{2,k_2} & \dots & b_{2,n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & b_{r,k_r} & b_{r,n} \\ 0 & \dots & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & 0 & 0 \end{vmatrix}$$

Система (2.4) совместна тогда и только тогда, когда $l_{r+1}=0$. Но в этом случае $r_C=r$ и поэтому $r_C=r_D$. Так как $C\sim B$, а $D\sim A$, то $r_A=r_B=r$. Теорема доказана.

Следствие. Пусть $r_A=r_B$, т. е. система (2.1) совместна.

Из выводов, полученных при рассмотрении алгоритма Гаусса, следует, что если $r=n$ (значит, при $m=n$ $\det A\neq 0$), то система имеет единственное решение (его можно получить по формулам Крамера).

Если же $r<n$, то система имеет бесконечное множество решений, причем из n неизвестных r являются главными, а $n-r$ — свободными.

§ 6. ОДНОРОДНАЯ СИСТЕМА ЛИНЕЙНЫХ УРАВНЕНИЙ

Рассмотрим частный случай системы (2.1), когда свободные члены ее равны нулю. Тогда получаем систему

$$\begin{cases} a_{11}x_1+a_{12}x_2+\dots+a_{1n}x_n=0, \\ a_{21}x_1+a_{22}x_2+\dots+a_{2n}x_n=0, \\ \dots\dots\dots\dots\dots\dots \\ a_{m1}x_1+a_{m2}x_2+\dots+a_{mn}x_n=0, \end{cases} \tag{2.6}$$

$$\text{или } AX=0,$$

которая называется *однородной*.

В этом случае $(h_1=h_2=\dots=h_n=0)$ расширенная матрица B отличается от матрицы системы A лишь наличием нулевого столбца. Поэтому $r_B=r_A$ и по теореме Кронекера — Капелли система совместна.

Таким образом, однородная система всегда совместна. Одно из ее решений очевидно: это так называемое нулевое, или тривиальное, решение

$$x_1 = x_2 = \ldots = x_n = 0.$$

Если $r_A = n$, то однородная система имеет единственное нулевое решение.

Если же $r_A < n$, то однородная система имеет бесчисленное множество решений. Таким образом, доказана **теорема**.

Для того чтобы однородная система имела решения, отличные от нулевого, необходимо и достаточно, чтобы ранг матрицы системы был меньше числа неизвестных.

Если число уравнений m совпадает с числом неизвестных, то система (2.6) имеет ненулевое решение тогда и только тогда, когда

$$\det A = 0.$$

Решения системы (2.6) обладают следующими свойствами:

1. Если $\alpha_1, \alpha_2, \ldots, \alpha_n$ — решения системы (2.6), то $\lambda\alpha_1, \lambda\alpha_2, \ldots, \lambda\alpha_n$, где λ — произвольное число, также решение системы.

2. Если $\alpha_1, \alpha_2, \ldots, \alpha_n$ и $\beta_1, \beta_2, \ldots, \beta_n$ — какие-либо два решения системы (2.6), то $\alpha_1 + \beta_1, \alpha_2 + \beta_2, \ldots, \alpha_n + \beta_n$ также будет ее решением. Справедливость этих свойств непосредственно следует из подстановки вместо x_i величин $\lambda\alpha_i$ и $\alpha_i + \beta_i$ в уравнения системы (2.6).

Заметим в заключение, что для системы (2.6) формулы (2.5) принимают вид:

$$\begin{cases} x_1 = \sum\limits_{k=r+1}^{n} \beta_{1,k} x_k, \\ x_2 = \sum\limits_{k=r+1}^{n} \beta_{2,k} x_k \\ x_r = \sum\limits_{k=r+1}^{n} \beta_{r,k} x_k \end{cases} \tag{2.7}$$

Здесь для удобства главные неизвестные мы обозначили

$$x_1, x_2, \ldots, x_r.$$

В матричной форме решение системы (2.6), учитывая формулы (2.7), имеет вид

$$X=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \\ x_{r+1} \\ \vdots \\ x_n \end{pmatrix}=\begin{pmatrix} \sum\limits_{k=r+1}^{n} \beta_{1,k} x_k \\ \sum\limits_{k=r+1}^{n} \beta_{2,k} x_k \\ \sum\limits_{k=r+1}^{n} \beta_{r,k} x_k \\ x_{r+1} \\ \vdots \\ x_n \end{pmatrix},$$

где неизвестным $x_{r+1}, \ldots, x_n$ можно придавать произвольные значения, или

$$X = x_{r+1}X_1 + x_{r+2}X_2 + \ldots + x_n X_{n-r} \ldots, \qquad (2.9)$$

где

$$X_1=\begin{pmatrix} \beta_{1,r+1} \\ \beta_{2,r+1} \\ \vdots \\ \beta_{r,r+1} \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}; X_2=\begin{pmatrix} \beta_{1,r+2} \\ \beta_{2,r+2} \\ \vdots \\ \beta_{r,r+2} \\ 0 \\ 1 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \ldots, X_{n-r}=\begin{pmatrix} \beta_{1,n} \\ \beta_{2,n} \\ \vdots \\ \beta_{r,n} \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Очевидно, $X_i\,(i=1,2,\ldots n-r)$ — частные решения системы (2.6) (так как при $x_{r+i}=1$ и остальных свободных неизвестных, равных нулю, $X=X_i$). Эти решения называются *фундаментальной системой решений уравнений* (2.6). Если фундаментальная система решений известна, то любое решение определяется по формуле (2.9).

Упражнения

1. Матрицу $\begin{pmatrix} 2 & 1 & -1 \\ 0 & -2 & 3 \\ 6 & 1 & 2 \end{pmatrix}$ элементарными преобразованиями привести к единичной.

2. Пусть A — матрица, содержащая m строк и n столбцов. Рассмотрим квадратную матрицу m-го порядка B_{ij}, у которой равны нулю все элементы, кроме элемента $b_{ij}=1$. Показать, что в матрице $\alpha B_{ij}A$ все строки, кроме i-й, состоят из нулей, i-я же строка ее представляет собой j-ю строку матрицы A, умноженную на α.

3. Показать, что элементарные преобразования над матрицей A могут быть осуществлены ее умножением слева на соответствующую матрицу C:

а) чтобы умножить i-ю строку матрицы A на число α, достаточно выбрать

$$C=E-(1-\alpha)B_{ii};$$

б) чтобы прибавить к i-й строке матрицы A ее j-ю строку, умноженную на α, достаточно выбрать

$$C=E+\alpha B_{ij};$$

в) чтобы переставить местами i-ю и j-ю строки матрицы A, достаточно выбрать

$$C=E-B_{ii}-B_{jj}+B_{ij}+B_{ji}$$

(заметим, что матрица C получается из E перестановкой i-й и j-й строк).

4. Показать, что любую невырожденную матрицу A можно элементарными преобразованиями привести к единичной матрице E.

Если те же преобразования в том же порядке произвести над матрицей E, то она приведется к матрице A^{-1}.

Решение.

Матрица A невырожденная. Поэтому элементарными преобразованиями она приводится к треугольной матрице, причем все ее диагональные элементы отличны от нуля. Треугольная матрица элементарными преобразованиями может быть приведена к диагональной форме. Последняя матрица умножением строк на соответствующие множители приводится к единичной (см. решение примера I).

Как мы видели (пример 3), элементарные преобразования над матрицей A можно выполнить умножением ее слева на соответствующие матрицы C.

Поэтому получаем

$$C_1C_2\ldots C_pA=E.$$

Значит,

$$C_1C_2\ldots C_p=A^{-1},$$

или

$$C_1C_2\ldots C_pE=A^{-1}.$$

5. Используя результаты примера 4, найти обратную матрицу для матриц:

$$\text{а)}\ \begin{pmatrix} 2 & 1 & -1 \\ 0 & -2 & 3 \\ 6 & 1 & 2 \end{pmatrix} \qquad \text{б)}\ \begin{pmatrix} 0 & 0 & 1 & -1 \\ 0 & 3 & 1 & 4 \\ 2 & 7 & 6 & -1 \\ 1 & 2 & 2 & -1 \end{pmatrix}.$$

6. Найти ранги следующих матриц:

$$\text{а)}\ \begin{pmatrix} 1 & 3 & 5 & -1 \\ 2 & -1 & -3 & 4 \\ 5 & 1 & -1 & 7 \\ 7 & 7 & 9 & 1 \end{pmatrix};$$

$$\text{б)}\ \begin{pmatrix} 4 & 3 & -5 & 2 & 3 \\ 8 & 6 & -7 & 4 & 2 \\ 4 & 3 & -8 & 2 & 7 \\ 4 & 3 & 1 & 2 & -5 \\ 8 & 6 & -1 & 4 & -6 \end{pmatrix}.$$

7. Решить систему методом Гаусса

$$\begin{cases} 2x_1 - 3x_2 + x_3 = -1, \\ x_1 + x_2 + x_3 = 6, \\ 3x_1 + x_2 - 2x_3 = -1. \end{cases}$$

8. Решить систему методом Гаусса

$$\begin{cases} 2x_1 - 3x_2 + 5x_3 - x_4 + x_5 = 5, \\ x_1 - 5x_2 + x_3 - 2x_4 + 3x_5 = 10, \\ 3x_1 + 4x_2 - 2x_3 + x_4 - x_5 = -12, \\ -x_1 + 3x_2 + 4x_3 + 2x_4 + 2x_5 = 10, \\ 4x_1 + x_2 - 2x_3 + x_4 + 4x_5 = 1. \end{cases}$$

9. Исследовать системы уравнений. В случае совместности системы найти ее решение.

$$\text{а)}\ \begin{cases} 2x_1 - 3x_2 + 5x_3 = 0, \\ 8x_1 + 2x_2 - x_3 = 21, \\ 2x_1 + 11x_2 - 16x_3 = 21. \end{cases}$$

$$\text{б)}\ \begin{cases} 2x_1 + x_2 + 3x_3 - x_4 + x_5 = 2, \\ x_1 - 2x_2 + x_3 + 3x_4 - x_5 = 1, \\ 3x_1 + 4x_2 + 5x_3 - 5x_4 + 3x_5 = 3, \\ x_1 + 3x_2 + 2x_3 - 4x_4 + 2x_5 = 1, \\ 5x_1 - 5x_2 + 6x_3 + 8x_4 - 2x_5 = 4. \end{cases}$$

в) $\begin{cases} 5x_1+3x_2+2x_3+6x_4+4x_5=-6, \\ 3x_1+5x_2+2x_3+5x_4+7x_5=-19, \\ 3x_1+9x_2+3x_3+10x_4+8x_5=-24, \\ 10x_1+2x_2+3x_3+7x_4+7x_5=-7. \end{cases}$

10. При каких λ система

$$\begin{cases} (5-\lambda)x-3y+2z=0, \\ 6x-(4+\lambda)y+4z=0, \\ 4x-4y+(5-\lambda)z=0 \end{cases}$$

имеет ненулевые решения?

11. Исследовать систему при различных значениях λ.

$$\begin{cases} \lambda x_1+x_2+x_3=1, \\ x_1+\lambda x_2+x_3=1, \\ x_1+x_2+\lambda x_3=1. \end{cases}$$

12. Найти общее решение и фундаментальную систему решений системы уравнений

$$\begin{cases} x_1+2x_2+4x_3-3x_4=0, \\ 3x_1+5x_2+6x_3-4x_4=0, \\ 4x_1+5x_2-2x_3+3x_4=0, \\ 3x_1+8x_2+24x_3-19x_4=0. \end{cases}$$

13. Имеем систему $AX=0$, где $A=\begin{pmatrix} a_{11}a_{12}\dots a_{1n} \\ a_{21}a_{22}\dots a_{2n} \\ \dots\dots \\ a_{n1}a_{n2}\dots a_{nn} \end{pmatrix}$, $X=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ и $\det A=0$. С помощью элементарных преобразований над столбцами матрицы A образуем нулевой столбец (пусть это будет j-й столбец). Те же элементарные преобразования произведем над матрицей E. Показать, что j-й столбец матрицы, полученной из E, образует решение (не нулевое) данной системы.

14. Используя результаты задачи 13, найти какое-нибудь не нулевое решение системы задачи 12.

Глава 3

ЛИНЕЙНЫЕ ПРЕОБРАЗОВАНИЯ И ЛИНЕЙНЫЕ ПРОСТРАНСТВА

§ 1. ЛИНЕЙНЫЕ ПРЕОБРАЗОВАНИЯ

1. Определение. Если некоторые величины y_1, $y_2, \ldots, y_m$ выражаются линейно и однородно через величины $x_1, x_2, \ldots, x_n$, т. е.

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n, \\ y_2 &= a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ y_m &= a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n, \end{aligned} \tag{3.1}$$

или сокращенно

$$y_i = \sum_{j=1}^{n} a_{ij}x_j, \quad (i = 1, 2, \ldots, m),$$

где a_{ij} — произвольные числа, то такое преобразование величин $x_1, x_2, \ldots, x_n$ в величины $y_1, y_2, \ldots, y_m$ называется *линейным преобразованием*.

Из коэффициентов линейного преобразования (3.1) можно составить матрицу A размеров $(m \times n)$

$$A = \begin{pmatrix} a_{11} \ldots a_{1n} \\ \cdot\ \cdot\ \cdot\ \cdot \\ a_{m1} \ldots a_{mn} \end{pmatrix},$$

которая однозначно определяет это линейное преобразование (и в свою очередь им определяется) и называется *матрицей линейного преобразования*.

Если ввести еще матрицы-столбцы

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \text{ и } Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix},$$

то линейное преобразование можно записать в матричной форме

$$Y = AX, \tag{3.2}$$

где AX — произведение матрицы A на матрицу-столбец X.

В качестве примера напомним формулы преобразования координат точки $M(x_1, x_2)$ плоскости при повороте системы декартовых координат на угол α (см. рис. 1).

Как известно, координаты (y_1, y_2) точки M в новой системе координат выражаются через координаты (x_1, x_2) в первоначальной системе координат в виде

$$y_1 = x_1 \cos\alpha + x_2 \sin\alpha,$$

$$y_2 = -x_1 \sin\alpha + x_2 \cos\alpha.$$

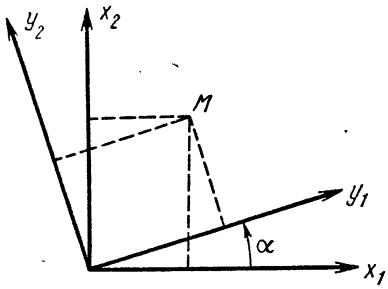


Рис. 1

Таким образом, преобразование координат точек плоскости при повороте осей на угол α является линейным преобразованием, матрица которого

$$A = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}.$$

Это линейное преобразование в матричной форме имеет вид

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

2. Свойства линейного преобразования. Линейные преобразования обладают двумя основными свойствами, которые следуют из соответствующих свойств матриц,

1) $A(X_1 + X_2) = AX_1 + AX_2$,

2) $A(\lambda X) = \lambda(AX)$.

3. Операции над линейными преобразованиями.

а) **Сложение преобразований.** Рассмотрим два линейных преобразования величин $x_1, x_2, \ldots, x_n$ соответственно в величины $y_1, y_2, \ldots, y_m$ и $z_1, z_2, \ldots, z_m$, т. е.

$$y_i = \sum_{j=1}^{n} a_{ij} x_j,$$

$$z_i = \sum_{j=1}^{n} b_{ij} x_j,$$

или в матричной форме

$$Y = AX,$$
$$Z = BX,$$

где

$$A = (a_{ij}),\ B = (b_{ij}),\ X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},\ Y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix},\ Z = \begin{pmatrix} z_1 \\ \vdots \\ z_m \end{pmatrix}.$$

Тогда *суммой линейных преобразований* называется преобразование величин $x_1, \ldots, x_n$ в величины $u_1, \ldots, u_m$, определяемое соотношениями

$$u_i = y_i + z_i = \sum_{j=1}^{n} (a_{ij} + b_{ij})\, x_j,\ i = 1, 2, \ldots, m.$$

Полученное преобразование также является линейным, а его матрица

$$C = (c_{ij}) = (a_{ij} + b_{ij}) = (a_{ij}) + (b_{ij}) = A + B.$$

б) **Умножение преобразования на число** λ. Наряду с линейным преобразованием

$$y_i = \sum_{j=1}^{n} a_{ij} x_j;\quad i = 1, 2, \ldots, m,$$

или в матричной форме

$$Y = AX$$

рассмотрим преобразование

$$z_i = \sum_{j=1}^{n} \lambda a_{ij} x_j;\quad i = 1, 2, \ldots, m.$$

Это преобразование называется *произведением первоначального преобразования на число* λ. Оно является линейным, а его матрица

$$B = (b_{ij}) = (\lambda a_{ij}) = \lambda A.$$

в) **Произведение преобразований.** Рассмотрим теперь линейное преобразование величин $x_1, \ldots, x_n$ в величины $y_1, y_2, \ldots, y_p$ и последующее линейное преобразование

величин $y_1, \ldots, y_p$ в величины $z_1, z_2, \ldots, z_m$, т. е.

$$y_k = \sum_{j=1}^{n} a_{kj} x_j, \quad k = 1, 2, \ldots, p \tag{3.3}$$

и

$$z_i = \sum_{k=1}^{p} b_{ik} y_k, \quad i = 1, 2, \ldots, m, \tag{3.4}$$

или в матричной форме

$$Y = AX,$$
$$Z = BY.$$

Посмотрим теперь, как выражаются величины $z_1, z_2, \ldots$ $\ldots, z_m$ через величины $x_1, x_2, \ldots, x_n$. Для этого подставим y_k из выражения (3.3) в выражение (3.4).

Используя матричную форму записи линейного преобразования, получим

$$Z = BY = B(AX) = (BA)X = CX.$$

Отсюда видно, что преобразование переменных x_i в переменные z_i является линейным и имеет матрицу $C = BA$. В развернутом виде это преобразование можно записать так:

$$z_i = \sum_{j=1}^{n} c_{ij} x_j = \sum_{j=1}^{n} \left(\sum_{k=1}^{p} b_{ik} a_{kj} \right) x_j. \tag{3.5}$$

Преобразование (3.5) называется *произведением преобразования* (3.4) *на преобразование* (3.3).

Фактически именно эта особенность, последовательного применения двух линейных преобразований, является причиной данного ранее определения произведения матриц.

г) **Обратное преобразование.** Пусть

$$Y = AX, \tag{3.6}$$

где A — квадратная невырожденная матрица. Тогда можно однозначно выразить переменные $x_1, \ldots, x_n$ через $y_1, \ldots, y_n$. По формуле (2.3) получаем

$$X = A^{-1}Y.$$

Это преобразование, очевидно, является линейным, оно называется *обратным преобразованием для преобразования* (3.6).

Как видно, матрица обратного преобразования является обратной для матрицы A. Отсюда следует, что однозначное обратное преобразование существует только тогда, когда матрица A невырожденная. В этом случае и само преобразование называется невырожденным. Если же A — вырожденная матрица, то и преобразование называется вырожденным.

§ 2. ЛИНЕЙНЫЕ ПРОСТРАНСТВА

Во всем последующем изложении существенную роль будут играть множества элементов произвольной природы, для которых определены некоторые операции, подчиняющиеся определенным ограничениям (аксиомам).

Такие множества называются *«пространствами»*, а их элементы — *«точками»* или *«векторами»* этого пространства и в последнем случае обозначаются так же, как векторные величины в геометрии $\boldsymbol{x}$, $\boldsymbol{y}$, ...

С самого начала следует подчеркнуть, что «векторы» таких абстрактных пространств, как правило, ничего общего не имеют с обычными (изучаемыми в геометрии) векторами (в плоскости или пространстве). Элементами абстрактных пространств могут быть функции, матрицы, некоторые системы чисел и т. д., а в частном случае и обычные векторы. Сами абстрактные пространства отличаются друг от друга лишь числом и характером вводимых на них операций и аксиомами, определяющими эти операции.

1. Числовое поле. *Числовым полем* K называется множество чисел α, β, γ, ..., если для любых α и β из множества K числа $\alpha+\beta$, $\alpha-\beta$, $\alpha\beta$ и $\frac{\alpha}{\beta}$ (в последнем случае $\beta\neq 0$) также принадлежат этому множеству. Например, множество чисел вида $a+b\sqrt{2}$, где a и b — любые рациональные числа, образует поле. Действительно, для любой пары чисел $\alpha=a+b\sqrt{2}$ и $\beta=c+d\sqrt{2}$ получаем:

1) $\alpha\pm\beta=(a\pm c)+(b\pm d)\sqrt{2}$,

2) $\alpha\beta=(ac+2bd)+(ad+bc)\sqrt{2}$,

3) $\frac{\alpha}{\beta}=\frac{ac-2bd}{c^2-2d^2}+\frac{bc-ad}{c^2-2d^2}\sqrt{2}$,

т. е. числа того же вида.

Очевидно, что множество чисел того же вида $a+b\sqrt{2}$, где a и b целые числа, поля не образует.

Легко показать, что множества рациональных, вещественных и комплексных чисел образуют числовые поля.

2. Линейное пространство. Аксиомы линейного пространства. *Линейным пространством* над числовым полем K называется множество R элементов, которые будем называть *векторами* и обозначать $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{z}$, ..., если:

1) указан закон, согласно которому любой паре векторов $\boldsymbol{x}\in R$ и $\boldsymbol{y}\in R$* однозначно ставится в соответствие вектор $\boldsymbol{z}\in R$. Вектор $\boldsymbol{z}$ называется суммой векторов $\boldsymbol{x}$ и $\boldsymbol{y}$ и обозначается

$$\boldsymbol{z}=\boldsymbol{x}+\boldsymbol{y};$$

2) указан закон, согласно которому каждому числу λ из поля K и любому вектору $\boldsymbol{x}\in R$ однозначно ставится в соответствие вектор $\boldsymbol{z}\in R$. Вектор $\boldsymbol{z}$ называется произведением вектора $\boldsymbol{x}$ на число λ и обозначается $\boldsymbol{z}=\lambda\boldsymbol{x}$ (или $\boldsymbol{z}=\boldsymbol{x}\lambda$);

3) введенные в пп. 1 и 2 операции сложения векторов и умножения вектора на число удовлетворяют следующим условиям (аксиомам):

а) $\boldsymbol{x}+\boldsymbol{y}=\boldsymbol{y}+\boldsymbol{x}$;

б) $(\boldsymbol{x}+\boldsymbol{y})+\boldsymbol{z}=\boldsymbol{x}+(\boldsymbol{y}+\boldsymbol{z})$ для любых $\boldsymbol{x}$, $\boldsymbol{y}$ и $\boldsymbol{z}$ из R;

в) существует элемент $\boldsymbol{0}\in R$ (нулевой вектор), такой что $\boldsymbol{x}+\boldsymbol{0}=\boldsymbol{x}$ для любого $\boldsymbol{x}\in R$;

г) для каждого $\boldsymbol{x}\in R$ существует такой вектор $\boldsymbol{y}\in R$, что $\boldsymbol{x}+\boldsymbol{y}=\boldsymbol{0}$.

Вектор $\boldsymbol{y}$ называется противоположным вектору $\boldsymbol{x}$ и обозначается $-\boldsymbol{x}$;

д) $1\cdot\boldsymbol{x}=\boldsymbol{x}$ для любого $\boldsymbol{x}\in R$;

е) $\alpha(\beta\boldsymbol{x})=(\alpha\beta)\boldsymbol{x}$ для любого $\boldsymbol{x}\in R$ и любых чисел α и β из K;

ж) $(\alpha+\beta)\boldsymbol{x}=\alpha\boldsymbol{x}+\beta\boldsymbol{x}$ для любого $\boldsymbol{x}\in R$ и любых чисел α и β из K;

з) $\alpha(\boldsymbol{x}+\boldsymbol{y})=\alpha\boldsymbol{x}+\alpha\boldsymbol{y}$ для любых $\boldsymbol{x}$ и $\boldsymbol{y}$ из R и любого числа α из K.

Из этих аксиом следует:

1. В линейном пространстве существует единственный нулевой вектор.

---

* Запись $\boldsymbol{x}\in R$ означает, что $\boldsymbol{x}$ — элемент множества R.

2. В линейном пространстве каждый вектор имеет единственный противоположный вектор.

3. Для любого элемента $x \in R$ имеет место равенство

$$0x = 0.$$

Для любого вектора $x \in R$ противоположный вектор равен

$$-x = (-1)x.$$

Существование противоположного вектора определяет возможность введения для векторов линейного пространства операции вычитания, как операции обратной операции сложения.

Назовем *разностью векторов* x и y вектор z (который обозначим $z = x - y$), удовлетворяющий равенству

$$z + y = x.$$

Прибавляя к обеим частям этого равенства элемент $-y$ и учитывая, что

$$y + (-y) = 0,$$

получаем

$$z = x + (-y).$$

При изучении общих абстрактных линейных пространств в дальнейшем будем предполагать, что векторы этих пространств обладают только свойствами, вытекающими из приведенных при определении линейного пространства аксиом. Таким образом, для векторов общего (абстрактного) линейного пространства не введены такие привычные для нас из геометрии свойства, как длина вектора, угол между векторами, площадь параллелограмма или объем параллелепипеда и другие так называемые метрические свойства.

Наделение векторов линейного пространства различными метрическими свойствами (в тех случаях, когда они необходимы) естественно сужает общность характера линейного пространства и приводит к более частным случаям, с некоторыми из которых мы познакомимся в дальнейшем.

3. Примеры линейных пространств. Приведем следующие примеры некоторых конкретных линейных пространств:

1. Множество всех вещественных чисел с обычными операциями сложения и умножения.

Будет ли это множество линейным пространством над некоторым числовым полем? Рассмотрим частные случаи при специальном выборе числовых полей:

а) над полем рациональных чисел это множество будет линейным пространством;

б) над полем вещественных чисел это множество также будет линейным пространством;

в) над полем комплексных чисел это множество линейного пространства не образует (так как произведение действительного числа на комплексное число, вообще говоря, есть комплексное число).

2. Множество всех комплексных чисел образует линейное пространство над любым из рассмотренных в примере 1 полей.

3. Множество всех рациональных чисел образует линейное пространство над полем рациональных чисел. Над полем вещественных или комплексных чисел это множество линейного пространства не образует.

Наиболее важными как в теоретическом, так и в прикладном отношении являются линейные пространства над полем вещественных чисел (вещественные линейные пространства) и над полем комплексных чисел (комплексные линейные пространства). В настоящей книге всюду в дальнейшем рассматриваются только эти два типа линейных пространств. В тех случаях, когда тип линейного пространства либо очевиден, либо не существен для проводимых рассуждений или пояснений, характер выбранного поля не указывается. Во всех же случаях, когда он оказывается существенным, специально отмечается, является ли рассматриваемое линейное пространство вещественным или комплексным.

4. Рассмотрим множество всех свободных векторов в пространстве с обычными операциями сложения векторов и умножения вектора на вещественное число. Как известно из курса аналитической геометрии, эти операции удовлетворяют аксиомам а) — з). Поэтому это множество является линейным (вещественным) пространством, которое обозначим V_3. Аналогично множество всех векторов, параллельных одной плоскости, образует линейное пространство, которое обозначим V_2.

5. Рассмотрим множество элементов, каждый из которых является упорядоченной последовательностью из n-чисел, принадлежащих полю K. Элементы (векторы) этого

множества будем обозначать

$$x=(\alpha_1, \ldots, \alpha_n), \quad y=(\beta_1, \beta_2, \ldots, \beta_n), \ldots$$

Операции сложения векторов и умножение вектора на число введем по правилам

$$x+y=(\alpha_1+\beta_1, \alpha_2+\beta_2, \ldots, \alpha_n+\beta_n),$$
$$\lambda x=(\lambda\alpha_1, \lambda\alpha_2, \ldots, \lambda\alpha_n).$$

Введенные операции удовлетворяют всем аксиомам линейного пространства. Значит, это множество является линейным пространством, которое обозначим T_n. Очевидно, что нулевой вектор из T_n имеет вид

$$0=(0, 0, \ldots, 0).$$

6. Множество всех многочленов степени, не превосходящей n, с обычными для многочленов операциями сложения и умножения на число. В этом пространстве векторы x, y ... имеют вид

$$x=A_0t^n+A_1t^{n-1}+\ldots+A_n,$$

где $A_0, A_1, \ldots, A_n$ — произвольные числа; t — переменная.

7. Пространство $C(a, b)$. Множество всех вещественных непрерывных на отрезке $a \leqslant t \leqslant b$ функций с обычными для функций операциями сложения и умножения на число. Векторы x, y из $C(a, b)$, их сумма и произведение на α имеют вид:

$$x=f(t), \quad y=\varphi(t),$$
$$x+y=f(t)+\varphi(t), \quad \alpha x=\alpha\cdot f(t).$$

8. Множество всех функций вида

$$ae^t+be^{-t}, \quad -\infty<t<\infty,$$

где a и b — произвольные вещественные числа.

Легко проверить, что в примерах 4—8 выполняются все условия и аксиомы линейного пространства. Соответствующую проверку рекомендуется выполнить читателю.

4. Подпространство. Пусть множество R образует некоторое линейное пространство, тогда всякое подмножество (т. е. всякая часть) множества R, элементы которого в свою очередь образуют линейное пространство (с теми же самыми операциями сложения и умножения на число, что и в R), называется *подпространством линейного пространства* R.

Очевидно, единственный нулевой вектор $0 \in R$ и пространство R являются соответственно наименьшим и наибольшим подпространствами линейного пространства K.

Покажем, что для того, чтобы подмножество L множества R было подпространством линейного пространства R, необходимо и достаточно, чтобы выполнялись условия:

1) если $\boldsymbol{x} \in L$ и $\boldsymbol{y} \in L$, то $\boldsymbol{x}+\boldsymbol{y} \in L$;

2) если $\boldsymbol{x} \in L$ и λ — любое число, то $\lambda \boldsymbol{x} \in L$.

Необходимость следует из того, что эти условия должны выполняться для любого линейного пространства. Для доказательства достаточности надо показать, что при этом выполняются все 8 аксиом линейного пространства. Справедливость аксиом а), б), д), е), ж) и з) очевидна.

Докажем выполнимость аксиомы в). Так как по условию, если $\boldsymbol{x} \in L$, то $\alpha \boldsymbol{x} \in L$ при любом α, то, полагая, $\alpha=0$, видим, что

$$0\boldsymbol{x}=0 \in L.$$

Но $\boldsymbol{x}+0=\boldsymbol{x} \in L$ и, следовательно, аксиома в) верна.

Для доказательства аксиомы г) положим $\alpha=-1$.

Так как если $\boldsymbol{x} \in L$, то $(-1)\boldsymbol{x} \in L$, а $(-1)\boldsymbol{x}$ есть вектор, противоположный вектору $\boldsymbol{x}$, то подмножество L вместе с любым элементом $\boldsymbol{x}$ содержит и противоположный ему элемент, что и доказывает выполнимость аксиомы г).

Например, в пространстве V_3 векторы параллельные некоторой плоскости образуют подпространство. (Это подпространство является пространством V_2.)

§ 3. БАЗИС ЛИНЕЙНОГО ПРОСТРАНСТВА

Как известно, в пространстве V_3 каждый вектор можно единственным образом разложить по трем некомпланарным векторам, которые называются *базисом пространства* V_3. Рассмотрим вопрос о построении базиса в произвольном линейном пространстве R. Предварительно введем некоторые важные понятия.

1. Линейная зависимость векторов. Если $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ суть векторы линейного пространства R, а $\alpha_1, \alpha_2, \ldots, \alpha_n$ — произвольные числа из поля K, то выражение

$$\alpha_1 \boldsymbol{x}_1+\alpha_2 \boldsymbol{x}_2+\ldots+\alpha_n \boldsymbol{x}_n$$

называется *линейной комбинацией векторов* $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$.

Числа $\alpha_1, \alpha_2, \ldots, \alpha_n$ называются *коэффициентами этой линейной комбинации.*

Если линейная комбинация векторов $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ обращается в нуль только тогда, когда все коэффициенты $\alpha_1, \alpha_2, \ldots, \alpha_n$ равны нулю, то векторы $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ называются *линейно независимыми.* В противном случае, т. е. когда равенство

$$\alpha_1\boldsymbol{x}_1 + \alpha_2\boldsymbol{x}_2 + \ldots + \alpha_n\boldsymbol{x}_n = 0$$

возможно при условии, что хотя бы один из коэффициентов $\alpha_1, \ldots, \alpha_n$ отличен от нуля, векторы $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ называются *линейно зависимыми.*

Заметим, что любая совокупность векторов $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$, содержащая нулевой вектор, линейно зависима. В самом деле, пусть $\boldsymbol{x}_1 = 0$. Тогда, полагая $\alpha_1 = 1$; $\alpha_2 = \alpha_3 = \ldots = \alpha_n = 0$, получаем

$$1\boldsymbol{x}_1 + 0\boldsymbol{x}_2 + \ldots + 0\boldsymbol{x}_n = 10 + 0\boldsymbol{x}_2 + \ldots + 0\boldsymbol{x}_n = 0,$$

и так как $\alpha_1 \neq 0$, то $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ линейно зависимы.

Точно так же очевидно, что система векторов $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$, содержащая совокупность линейно зависимых векторов, также линейно зависима. Покажем теперь, что *векторы $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ линейно зависимы тогда и только тогда, когда хотя бы один из них можно представить как линейную комбинацию оставшихся.*

Пусть $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ линейно зависимы. Это значит, что существуют такие $\alpha_1, \alpha_2, \ldots, \alpha_n$, не все равные нулю, что

$$\alpha_1\boldsymbol{x}_1 + \alpha_2\boldsymbol{x}_2 + \ldots + \alpha_n\boldsymbol{x}_n = 0.$$

Примем для определенности, что $\alpha_n \neq 0$. Тогда очевидно

$$\boldsymbol{x}_n = -\frac{\alpha_1}{\alpha_n}\boldsymbol{x}_1 - \ldots - \frac{\alpha_{n-1}}{\alpha_n}\boldsymbol{x}_{n-1}.$$

Обратно, пусть один из векторов, например $\boldsymbol{x}_n$, можно представить в виде

$$\boldsymbol{x}_n = \beta_1\boldsymbol{x}_1 + \beta_2\boldsymbol{x}_2 + \ldots + \beta_{n-1}\boldsymbol{x}_{n-1}.$$

Тогда

$$-\beta_1 x_1 - \beta_2 x_2 - \ldots - \beta_{n-1} x_{n-1} + 1 x_n = 0,$$

т. е. векторы линейно зависимы.

2. **Размерность линейного пространства.** *Если в линейном пространстве R существует n линейно независимых векторов, а любые $n+1$ векторов этого пространства линейно зависимы, то линейное пространство называется n-мерным.* Число n называется *размерностью пространства.*

Например, пространства V_2 и V_3 соответственно двух- и трехмерные.

Если в линейном пространстве R существует линейно независимая система из любого числа векторов, то оно называется *бесконечномерным.* Например, пространство $C(a, b)$ — бесконечномерное.

В линейной алгебре изучаются только конечномерные пространства.

3. **Базис линейного пространства определение.** Система $\{e\}$ из n линейно независимых векторов n-мерного пространства

$$e_1, e_2, \ldots, e_n,$$

заданных в определенном порядке, называется *базисом этого пространства.*

Например, в пространстве V_3 базис образуют любые три некомпланарных вектора, заданных в определенном порядке, а в пространстве V_2 базис образуют любые два неколлинеарных вектора.

Теорема. *Любой вектор x, n-мерного пространства можно, и притом единственным образом, разложить по базису этого пространства $e_1, e_2, \ldots, e_n$.*

Действительно, векторы

$$x, e_1, e_2, \ldots, e_n$$

линейно зависимы (так как их число равно $n+1$) — значит можно составить выражения

$$\alpha_0 x + \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n = 0,$$

в котором хотя бы одно из чисел $\alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_n$ отлично от нуля.

В данном случае $\alpha_0 \neq 0$, так как в противном случае окажется, что базисные векторы линейно зависимы. По-

этому имеем

$$\boldsymbol{x} = -\frac{\alpha_1}{\alpha_0}\boldsymbol{e}_1 - \ldots - \frac{\alpha_n}{\alpha_0}\boldsymbol{e}_n = x_1\boldsymbol{e}_1 + \ldots + x_n\boldsymbol{e}_n. \quad (3.7)$$

Таким образом, любой вектор $\boldsymbol{x}$ можно представить в виде линейной комбинации базисных векторов.

Покажем, что это разложение единственно. Если предположить, что наряду с ним имеется разложение

$$\boldsymbol{x} = x_1^*\boldsymbol{e}_1 + x_2^*\boldsymbol{e}_2 + \ldots + x_n^*\boldsymbol{e}_n,$$

то, вычитая его из разложения (3.7), получим

$$0 = (x_1 - x_1^*)\,\boldsymbol{e} + (x_2 - x_2^*)\,\boldsymbol{e}_2 + \ldots + (x_n - x_n^*)\,\boldsymbol{e}_n.$$

Так как векторы $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ линейно независимы, то последнее равенство возможно только при условии

$$x_1 - x_1^* = 0, \ \ldots, \ x_n - x_n^* = 0,$$

т. е.

$$x_1 = x_1^*, \ \ldots, \ x_n = x_n^*,$$

что и доказывает единственность разложения (3.7).

Таким образом, если в n-мерном пространстве выбран некоторый базис, то пользуясь разложением (3.7), можно установить взаимно однозначное соответствие между векторами этого пространства и упорядоченными последовательностями из n чисел $(x_1, x_2, \ldots, x_n)$.

Числа $x_1, x_2, \ldots, x_n$ называются *координатами вектора* $\boldsymbol{x}$ в базисе $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$. В этом случае будем писать $\boldsymbol{x} = \{x_1, x_2, \ldots, x_n\}$ (или $\boldsymbol{x} = (x_1, x_2, \ldots x_n)$). Из теоремы следует, что два вектора $\boldsymbol{x} = x_1\boldsymbol{e}_1 + \ldots + x_n\boldsymbol{e}_n$ и $\boldsymbol{y} = y_1\boldsymbol{e}_1 + \ldots + y_n\boldsymbol{e}_n$ n-мерного линейного пространства, в котором задан некоторый базис, равны тогда и только тогда, когда их координаты в этом базисе равны, т. е. когда

$$x_1 = y_1, \quad x_2 = y_2, \ \ldots, \ x_n = y_n.$$

Очевидно, что если в пространстве выбрать другой базис, то тот же вектор $\boldsymbol{x}$ будет иметь другие координаты.

§ 4. ДЕЙСТВИЯ НАД ВЕКТОРАМИ В КООРДИНАТНОЙ ФОРМЕ

1. Линейные операции над векторами в координатной форме. Покажем, как осуществляются в координатной форме операции сложения и умножения на число векторов n-мерного линейного пространства, в котором выбран

некоторый базис $\{e\}$. Так как всякий вектор из R можно представить и притом единственным образом в виде линейной комбинации базисных векторов, т. е.

$$x = x_1e_1 + x_2e_2 + \ldots + x_ne_n,$$
$$y = y_1e_1 + y_2e_2 + \ldots + y_ne_n,$$

то на основании аксиом, которым удовлетворяют операции сложения и умножения на число (учитывая, что $e_i \in R$), имеем

$$x + y = (x_1e_1 + \ldots + x_ne_n) + (y_1e_1 + \ldots + y_ne_n) =$$
$$= (x_1 + y_1)\, e_1 + \ldots + (x_n + y_n)\, e_n$$
$$\lambda x = \lambda\,(x_1e_1 + \ldots + x_ne_n) = \lambda x_1e_1 + \ldots + \lambda x_ne_n.$$

Таким образом приходим к следующей **теореме.**

Если векторы n-мерного линейного пространства заданы своими координатами относительно некоторого базиса $\{e\}$, то при сложении векторов или умножения их на число λ координаты векторов соответственно складываются или умножаются на λ.

Таким образом:

$$x + y = \{x_1 + y_1,\ x_2 + y_2,\ \ldots,\ x_n + y_n\},$$
$$\lambda x = \{\lambda x_1,\ \lambda x_2,\ \ldots,\ \lambda x_n\}.$$

Поэтому, если

$$y = \alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_m x_m,$$

где

$$x_1 = \{x_{11},\ x_{12},\ \ldots,\ x_{1n}\},$$
$$x_2 = \{x_{21},\ x_{22},\ \ldots,\ x_{2n}\},$$
$$\cdots\cdots\cdots\cdots$$
$$x_m = \{x_{m1},\ x_{m2},\ \ldots,\ x_{mn}\},$$
$$y = \{y_1,\ y_2,\ \ldots,\ y_n\},$$

то

$$y_1 = \alpha_1 x_{11} + \alpha_2 x_{21} + \ldots + \alpha_m x_{m1},$$
$$y_2 = \alpha_1 x_{12} + \alpha_2 x_{22} + \ldots + \alpha_m x_{m2},$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$y_n = \alpha_1 x_{1n} + \alpha_2 x_{2n} + \ldots + \alpha_m x_{mn}.$$

2. Признак линейной зависимости и независимости векторов. Имеем систему векторов

$$x_1,\ x_2,\ \ldots,\ x_m \tag{3.8}$$

n-мерного пространства R.

Приравняем нулю их линейную комбинацию

$$\alpha_1 \boldsymbol{x}_1 + \alpha_2 \boldsymbol{x}_2 + \ldots + \alpha_m \boldsymbol{x}_m = 0. \tag{3.9}$$

В силу результатов предыдущего пункта (полагая $\boldsymbol{y} = 0$) получаем систему

$$\begin{cases} x_{11}\alpha_1 + x_{21}\alpha_2 + \ldots + x_{m1}\alpha_m = 0, \\ x_{12}\alpha_1 + x_{22}\alpha_2 + \ldots + x_{m2}\alpha_m = 0, \\ x_{1n}\alpha_1 + x_{2n}\alpha_2 + \ldots + x_{mn}\alpha_m = 0. \end{cases} \tag{3.10}$$

Матрицу этой системы

$$\begin{pmatrix} x_{11} & x_{21} & \ldots & x_{m1} \\ x_{12} & x_{22} & \ldots & x_{m2} \\ \cdot & \cdot & \cdot & \cdot \\ x_{1n} & x_{2n} & \ldots & x_{mn} \end{pmatrix},$$

столбцы которой составлены из координат векторов $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_m$, обозначим M_m и сокращенно запишем в виде

$$M_m = (X_1; X_2, \ldots, X_m), \text{ где } X_i = \begin{pmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{in} \end{pmatrix}; \; i = 1, 2, \ldots m.$$

Соотношение (3.9) эквивалентно системе (3.10). Поэтому векторы системы (3.8) *линейно зависимы тогда и только тогда, когда система (3.10) имеет ненулевое решение* $\alpha_1, \alpha_2, \ldots, \alpha_m$.

Иначе векторы системы (3.8) линейно независимы тогда и только тогда, когда ранг матрицы M_m равен m (т. е. числу векторов системы).

Пример.

Векторы

$$\begin{aligned} \boldsymbol{x}_1 &= 2\boldsymbol{e}_1 + \boldsymbol{e}_2 + \boldsymbol{e}_4, \\ \boldsymbol{x}_2 &= \boldsymbol{e}_1 - \boldsymbol{e}_2 + 2\boldsymbol{e}_3 - \boldsymbol{e}_4, \\ \boldsymbol{x}_3 &= \boldsymbol{e}_1 + 2\boldsymbol{e}_2 + \boldsymbol{e}_3 \end{aligned}$$

линейно независимы, так как ранг матрицы

$$M_3 = \begin{pmatrix} 2 & 1 & 1 \\ 1 & -1 & 2 \\ 0 & 2 & 1 \\ 1 & -1 & 0 \end{pmatrix}$$

равен 3.

Следствие 1. *Если любой вектор пространства R можно разложить по n линейно независимым векторам e_1, e_2, ..., e_n, то пространство является n-мерным.*

Для доказательства этого нам надо показать, что любые $n+1$ векторов x_1, x_2, ..., x_{n+1} пространства R линейно зависимы. По условию

$$x_i = \sum_{k=1}^{n} x_{ik} e_k.$$

Матрица M_{n+1} из коэффициентов разложения имеет теперь n строк и $n+1$ столбцов. Значит, ее ранг не превосходит n, т. е. векторы x_1, x_2, ..., x_{n+1} линейно зависимы.

Таким образом, размерность пространства можно определить числом базисных векторов, если понимать базис как совокупность линейно независимых векторов, по которым разлагается любой вектор пространства.

Следствие 2. Имеем базис пространства R: e_1, e_2, ..., e_n и систему из r линейно независимых векторов x_1, x_2, ..., x_r. Покажем, что *из базиса можно выбрать такие $n-r$ векторов, чтобы они вместе с векторами x_1, x_2, ..., x_r образовывали базис этого пространства.*

Мы знаем, что ранг матрицы $M_r = (X_1, X_2, \ldots, X_r)$ равен r. Не ограничивая общности, будем считать, что не равен нулю определитель, содержащий первые r строк.

Рассмотрим систему векторов $x_1, x_2, \ldots, x_r, e_{r+1}, \ldots, e_n$. Ранг матрицы

$$M_n = \begin{pmatrix} x_{11} & x_{21} & \ldots & x_{r1} & 0 & 0 & \ldots & 0 \\ x_{12} & x_{22} & \ldots & x_{r2} & 0 & 0 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ x_{1r} & x_{2r} & \ldots & x_{rr} & 0 & 0 & \ldots & 0 \\ x_{1,r+1} & x_{2,r+1} & \ldots & x_{r,r+1} & 1 & 0 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ x_{1n} & x_{2n} & \ldots & x_{rn} & 0 & 0 & \ldots & 1 \end{pmatrix}.$$

равен n, так как

$$\det M_n = \begin{vmatrix} x_{11} & x_{21} & \dots & x_{r1} \\ x_{12} & x_{22} & \dots & x_{r2} \\ \cdot & \cdot & \cdot & \cdot \\ x_{1r} & x_{2r} & \dots & x_{rr} \end{vmatrix} \neq 0.$$

Значит векторы $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_r, \boldsymbol{e}_{r+1}, \dots, \boldsymbol{e}_n$ линейно независимы и, следовательно, образуют базис пространства.

3. Ранг системы векторов. Максимальное число линейно независимых векторов системы (3.8) назовем ее *рангом*.

Теорема.

Ранг системы векторов (3.8) *равен рангу матрицы из координат этих векторов в некотором базисе.*

Пусть ранг системы векторов (3.8) равен r. Значит, в ней имеются r линейно независимых векторов. В силу результатов предыдущего пункта ранг матрицы M_r из координат этих векторов равен r. Но матрица M_r является частью матрицы M_m из координат всех векторов (3.8). Поэтому ранг матрицы M_m не меньше r. Но он не может быть больше r, так как в этом случае среди векторов системы (3.8) нашлись бы более чем r линейно независимых векторов. Итак, ранг матрицы M_m равен r.

Следствие. *Ранг матрицы равен максимальному числу ее линейно независимых столбцов* (*или строк*).

Пусть дана матрица

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix}.$$

Будем рассматривать ее как матрицу координат векторов

$$A_i = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{pmatrix}.$$

Будем считать, что столбцы матрицы линейно независимы (или зависимы), если независимы (или зависимы) соответствующие им векторы A_i. Тогда из теоремы следует, что ранг матрицы A равен рангу системы векто-

ров A_i, т. е. максимальному числу линейно независимых столбцов. Так как при транспонировании ранг матрицы не меняется, то аналогичное утверждение справедливо и для строк. Отсюда, в частности, следует, что определитель матрицы равен нулю тогда и только тогда, когда ее строки (или столбцы) линейно зависимы.

В заключение отметим, что повторяя выкладки п. 3 § 3 этой главы, легко показать, что если ранг системы векторов (3.8) равен r, то любой вектор этой системы можно однозначно разложить по ее r линейно независимым векторам.

Примеры.

1. Имеем пространство T_n. Рассмотрим векторы

$$e_1=(1,\ 0,\ \ldots\ 0),\quad e_2=(0,\ 1,\ \ldots\ 0),\quad e_n=(0,\ 0,\ \ldots\ 1),$$

покажем, что они линейно независимы.
Действительно, из равенства нулю линейной комбинации

$$\alpha_1 e_1+\alpha_2 e_2+\ldots+\alpha_n e_n=0$$

следует

$$(\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_n)=0=(0,\ 0,\ \ldots,\ 0),$$

т. е

$$\alpha_1=\alpha_2=\ldots=\alpha_n=0.$$

Далее, любой вектор $x=(x_1,\ x_2,\ \ldots,\ x_n)$ разлагается по векторам $e_1, e_2, \ldots, e_n$; $x=x_1e_1+\ldots+x_ne_n$.

Поэтому пространство T_n есть n-мерное, а векторы $e_1, e_2, \ldots, e_n$ образуют его базис.

2. Как мы знаем, пространство функций вида $f(t)=ae^t+be^{-t}$ является линейным. Его размерность равна двум. Это следует из того, что функции e^t и e^{-t} линейно независимы (равенство $\alpha_1e^t+\alpha_2e^{-t}=0$ справедливо лишь при $\alpha_1=\alpha_2=0$) и образуют базис.

3. Имеем векторы

$$x_1=\{1,\ 2,\ -1\},$$
$$x_2=\{2,\ 0,\ 3\},$$
$$x_3=\{\alpha,\ 1,\ 2\}.$$

Подобрать α так, чтобы эти векторы были линейно зависимы. Векторы линейно зависимы, если

$$\begin{vmatrix} 1 & 2 & \alpha \\ 2 & 0 & 1 \\ -1 & 3 & 2 \end{vmatrix}=0,\ \text{отсюда}\ \alpha=\frac{13}{6}.$$

4. Пространство решений линейной однородной системы. Рассмотрим множество всех решений некоторой

однородной системы линейных уравнений. Известно (гл. 2 § 6), что сумма решений этой системы и произведение любого ее решения на число также являются решениями этой системы и при этом выполняются все аксиомы линейного пространства. Поэтому множество решений однородной системы линейных уравнений является линейным пространством.

Найдем его размерность и базис.

Теорема. *Размерность пространства решений линейной однородной системы уравнений с n неизвестными равна $n-r$, где r — ранг матрицы системы.*

Действительно, как было показано в гл. 2 § 6, любое решение однородной системы можно представить в виде

$$X = x_{r+1}X_1 + x_{r+2}X_2 + \ldots + x_nX_{n-r}.$$

Поэтому для доказательства теоремы нам достаточно показать, что решения $X_1, X_2, \ldots, X_{n-r}$ линейно независимы. Составим матрицу $M_{n-r} = (X_1, X_2, \ldots, X_{n-r})$ или

$$M_{n-r} = \begin{pmatrix} \beta_{1,r+1} & \beta_{1,r+2} & \ldots & \beta_{1,n} \\ \beta_{2,r+1} & \beta_{2,r+2} & \ldots & \beta_{2,n} \\ \cdot & \cdot & \cdot & \cdot \\ \beta_{r,r+1} & \beta_{r,r+2} & \ldots & \beta_{r,n} \\ 1 & 0 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & 1 \end{pmatrix}.$$

Ранг этой матрицы равен $n-r$, так как минор матрицы, составленной из ее столбцов и последних $n-r$ строк

$$\begin{vmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & 1 \end{vmatrix} = 1.$$

Значит, решения $X_1, X_2, \ldots X_{n-r}$ линейно независимы и их можно принять за базис пространства решений.

Любой базис пространства решений называется *фундаментальной (или основной) системой решений.*

Упражнения

1*. Образуют ли линейное пространство многочлены:

а) степени n, б) степени, меньшей или равной n, в) степени, большей или равной n $(n \neq 0)$?

2. Доказать, что квадратные матрицы n-го порядка образуют линейное пространство.

3. Образуют ли линейное пространство все векторы плоскости, не параллельные некоторой оси?

4. Образуют ли линейное подпространство векторы плоскости, концы которых лежат в 1-й четверти (предполагаем, что векторы откладываются от начала координат)?

5. Перечислить все подпространства трехмерного пространства.

6. Образуют ли подпространство в пространстве R всевозможные линейные комбинации данных векторов $x_1, x_2, \ldots, x_n$ из R?

7. Будут ли линейно зависимыми векторы

а) $x_1 = (1, 2, 3)$, $x_2 = (3, 6, 7)$;

б) $x_1 = (5, 4, 3)$, $x_2 = (3, 3, 2)$, $x_3 = (8, 1, 3)$?

8. Можно ли в пространстве многочленов степени не выше второй выбрать за базис многочлены:

а) $1 + t^2$, $1 - t^2$, $1 + 2t$;

б) $1 + t^2$, $1 - t^2$, $1 + 2t^2$?

9. Имеем систему $AX = 0$, где

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} \quad \text{и} \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Пусть матрица A имеет ранг r, тогда с помощью элементарных преобразований над столбцами матрицы A можно образовать матрицу B с $n - r$ нулевыми столбцами. Проведем такие же преобразования над столбцами матрицы E, получим матрицу G.

Доказать, что столбцы матрицы G, соответствующие нулевым столбцам матрицы B, образуют фундаментальную систему решений данной системы уравнений.

10. Используя результаты задачи **9**, найти фундаментальную систему решений следующей системы:

$$\begin{cases} x_1 + x_2 - x_3 + x_4 = 0; \\ x_1 - x_2 + x_3 + x_4 = 0; \\ 3x_1 + x_2 - x_3 + 3x_4 = 0; \\ x_2 - x_3 = 0. \end{cases}$$

11. Найти размерности всех подпространств трехмерного пространства. Выбрать базис в каждом подпространстве.

---

* В этом примере и во всех последующих предполагается, что операции сложения и умножения на число элементов пространства вводятся обычным образом.

12. Найти размерность линейного пространства, образованного квадратными матрицами n-го порядка.

13. Показать, что векторы $e_1=(2, 1, -3)$, $e_2=(3, 2, -5)$, $e_3=(1, -1, 1)$ образуют базис трехмерного пространства. Найти координаты вектора $x=(6, 2, -7)$ в этом базисе.

14. Дана матрица A ранга r. Используя элементарные преобразования столбцов, выбрать такие r столбцов матрицы, чтобы построенная из них матрица имела ранг r.

15. Имеем систему из m векторов ранга r. Используя результаты задачи 14, выбрать из них r линейно независимых векторов.

16. Имеем систему векторов $x_1=(1, 1, 1)$, $x_2=(0, 0, 1)$, $x_3=(2, 2, 3)$, $x_4=(2, -1, 0)$, $x_5=(1, 2, 1)$.

Выбрать из них максимальное число линейно независимых векторов.

Глава 4

ЛИНЕЙНЫЙ ОПЕРАТОР

§ 1. ЛИНЕЙНЫЙ ОПЕРАТОР

1. Определение. Рассмотрим два линейных пространства R и R_1 над полем K. *Оператором* $\tilde{A}$ называется любое правило, согласно которому каждому вектору $x \in R$ ставится в соответствие единственный вектор $y \in R_1$, что символически записывается в виде

$$y=\tilde{A}x.$$

Вектор y называется *образом вектора* x, а вектор x — *прообразом вектора* y.

Очевидно, что оператор $\tilde{A}$ представляет собой обобщение известного из анализа определения функции на случай, когда областью задания функции является произвольное линейное пространство R, а область значений принадлежит пространству R_1.

Оператор $\tilde{A}$ называется *линейным*, если выполняются следующие два условия:

1) $\tilde{A}(x_1+x_2)=\tilde{A}x_1+\tilde{A}x_2$

для любых x_1 и x_2 из R (условие аддитивности);

2) $\tilde{A}(\lambda x)=\lambda\tilde{A}x$

для любого $x \in R$ и любого λ из поля K (условие однородности).

Полагая $\lambda=0$, имеем $\tilde{A}0=0$, т. е. линейный оператор преобразует нулевой вектор в нулевой.

В дальнейшем, если не будет оговорено противное, мы будем рассматривать только линейные операторы, поэтому слово «линейный» часто будем опускать.

Если $\tilde{A}$ — линейный оператор, то множество L всех образов

$$y = \tilde{A}x \qquad x \in R$$

образует подпространство линейного пространства R_1. Это подпространство называется *пространством образов*.

По определению оператора, L является подмножеством R_1 (так как $y \in R_1$). Остается лишь доказать (см. гл. 3 § 2 п. 4), что:

1) если $y_1 \in L$ и $y_2 \in L$, то $y_1 + y_2 \in L$;

2) если $y \in L$, то $\lambda y \in L$.

Так как $y_1 = \tilde{A}x_1$, $y_1 = \tilde{A}x_2$, то

$$y_1 + y_2 = \tilde{A}(x_1 + x_2) = \tilde{A}x_3 \in L.$$

Точно так же

$$\lambda y = \lambda \tilde{A}x = \tilde{A}(\lambda x) \in L,$$

что и доказывает наше утверждение. Так как L либо совпадает с R_1, либо является его частью, то преобразование R в L с помощью оператора $\tilde{A}$, что будем обозначать $L = \tilde{A}R$, называется *отображением пространства* R в R_1.

Очевидно, размерность пространства образов L не превышает размерности пространства R.

Действительно, оператор $\tilde{A}$ переводит линейную комбинацию векторов из R в линейную комбинацию векторов из L с теми же коэффициентами. Следовательно, максимальное число линейно независимых векторов из L не превышает максимального числа линейно независимых векторов из R.

2. Матрица линейного оператора. Пусть в n-мерном линейном пространстве R задан оператор A, который переводит вектор $x \in R$ в вектор y m-мерного пространства R_1, т. е.

$$y = \tilde{A}x. \tag{4.1}$$

Выберем базисы $e_1, e_2, \ldots, e_n$ и $g_1, g_2, \ldots, g_m$ в пространствах R и R_1.

Тогда

$$\begin{aligned} x &= x_1e_1 + x_2e_2 + \ldots + x_ne_n, \\ y &= y_1g_1 + y_2g_2 + \ldots + y_mg_m. \end{aligned} \tag{4.2}$$

Поставим перед собой задачу найти выражение координат образа $\boldsymbol{y}$ через координаты прообраза $\boldsymbol{x}$. Из (4.1) и (4.2) имеем

$$\boldsymbol{y}=y_1\boldsymbol{g}_1+y_2\boldsymbol{g}_2+\ldots+y_m\boldsymbol{g}_m=\tilde{A}(x_1\boldsymbol{e}_1+x_2\boldsymbol{e}_2+\ldots+x_n\boldsymbol{e}_n)=$$
$$=x_1\tilde{A}\boldsymbol{e}_1+x_2\tilde{A}\boldsymbol{e}_n+\ldots+x_n\tilde{A}\boldsymbol{e}_n. \qquad (4.3)$$

Отсюда видно, что для задания оператора достаточно задать только образы базисных векторов, так как если векторы $\tilde{A}\boldsymbol{e}_k$ известны, то известен и образ любого вектора $\boldsymbol{x}$.

Так как любой вектор $\tilde{A}\boldsymbol{e}_k$ может быть разложен по выбранному базису пространства R_1, то

$$\begin{aligned}
\tilde{A}\boldsymbol{e}_1&=a_{11}\boldsymbol{g}_1+a_{21}\boldsymbol{g}_2+\ldots+a_{m1}\boldsymbol{g}_m,\\
\tilde{A}\boldsymbol{e}_2&=a_{12}\boldsymbol{g}_1+a_{22}\boldsymbol{g}_2+\ldots+a_{m2}\boldsymbol{g}_m,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\\
\tilde{A}\boldsymbol{e}_n&=a_{1n}\boldsymbol{g}_1+a_{2n}\boldsymbol{g}_2+\ldots+a_{mn}\boldsymbol{g}_m.
\end{aligned} \qquad (4.4)$$

Составим матрицу (обозначим ее A') из коэффициентов этих разложений

$$A'=\begin{pmatrix} a_{11} & a_{21} & \ldots & a_{m1}\\ a_{12} & a_{22} & \ldots & a_{m2}\\ \ldots & \ldots & \ldots & \ldots\\ a_{1n} & a_{2n} & \ldots & a_{mn}\end{pmatrix}.$$

Очевидно, что если известны равенства (4.4), то известна матрица A' и наоборот; из (4.3) и (4.4) получаем

$$y_1\boldsymbol{g}_1+y_2\boldsymbol{g}_2+\ldots+y_m\boldsymbol{g}_m=(a_{11}x_1+a_{12}x_2+$$
$$+\ldots+a_{1n}x_n)\,\boldsymbol{g}_1+(a_{21}x_1+a_{22}x_1+\ldots+a_{2n}x_n)\,\boldsymbol{g}_2+$$
$$+\ldots+(a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n)\,\boldsymbol{g}_m.$$

Отсюда (в силу единственности разложения по базису) следует:

$$\begin{aligned}
y_1&=a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n;\\
y_2&=a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n;\\
&\ldots\ldots\ldots\ldots\ldots\ldots\\
y_m&=a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n.
\end{aligned} \qquad (4.5)$$

Формулы (4.5) решают поставленную нами задачу. Запишем их в матричном виде

$$Y=AX,$$

где

$$A=\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}; \quad X=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}; \quad Y=\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}.$$

Матрица A, получаемая транспонированием матрицы A', называется *матрицей линейного оператора* $\tilde{A}$; *k-й столбец этой матрицы составлен из последовательных координат вектора $\tilde{A}\boldsymbol{e}_k$ в базисе $\boldsymbol{g}_1, \boldsymbol{g}_2, \dots, \boldsymbol{g}_m$.*

3. Ранг оператора. Как следует из предыдущего, матрица линейного оператора, вообще говоря, зависит от выбора базисов. Однако, как будет показано, некоторые ее характеристики оказываются не зависящими от базисов и потому их величины характеризуют внутренние свойства линейного оператора. Докажем следующее утверждение.

Теорема. *Ранг матрицы линейного оператора равен размерности пространства образов.*

Пусть матрица A линейного оператора $\tilde{A}$ имеет ранг r. Следовательно, система образов базисных векторов $\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n$ имеет ранг r, значит, среди векторов $\tilde{A}\boldsymbol{e}_1, \tilde{A}\boldsymbol{e}_2, \dots, \tilde{A}\boldsymbol{e}_n$ имеются r линейно независимых векторов, пусть это $\tilde{A}\boldsymbol{e}_1, \tilde{A}\boldsymbol{e}_2, \dots, \tilde{A}\boldsymbol{e}_r$, по которым разлагаются остальные $n-r$ векторов. Так как любой вектор $\boldsymbol{y}$ пространства образов можно представить в виде

$$\boldsymbol{y}=x_1\tilde{A}\boldsymbol{e}_1+x_2\tilde{A}\boldsymbol{e}_2+\dots+x_n\tilde{A}\boldsymbol{e}_n$$

(см. п. 2), то вектор $\boldsymbol{y}$ можно разложить по векторам $\tilde{A}\boldsymbol{e}_1, \tilde{A}\boldsymbol{e}_2, \dots, \tilde{A}\boldsymbol{e}_r$. Значит, пространство образов r-мерное, а векторы $\tilde{A}\boldsymbol{e}_1, \tilde{A}\boldsymbol{e}_2, \dots, \tilde{A}\boldsymbol{e}_r$ образуют базис этого пространства.

Координаты этих векторов являются столбцами матрицы. Следовательно, за базис пространства образов можно принять любые r линейно независимых столбцов (столбцовых векторов) матрицы.

Таким образом, ранг матрицы оператора не зависит от выбора базисов и поэтому мы будем называть его *рангом оператора*. В частности, если ранг r оператора равен n (размерности пространства), то оператор переводит любую систему линейно независимых векторов в систему линейно независимых векторов, т. е. размерность

пространства образов L совпадает с размерностью пространства R. В этом случае оператор будем называть *невырожденным* (в противном случае, т. е. при $r < n$ — *вырожденным*).

Заметим, что оператор $\tilde{A}$ является вырожденным тогда и только тогда, когда он переводит какой-нибудь ненулевой вектор $\boldsymbol{x}$ в нулевой, так как из соотношения

$$\tilde{A}\boldsymbol{x} = x_1\tilde{A}\boldsymbol{e}_1 + x_2\tilde{A}\boldsymbol{e}_2 + \ldots + x_n\tilde{A}\boldsymbol{e}_n = 0$$

следует линейная зависимость образов базисных векторов. Например, оператор, преобразующий векторы пространства V_3 в векторы пространства V_2, имеет ранг, равный двум, и, следовательно, этот оператор вырожденный.

4. Отображение линейного пространства в себя. В дальнейшем будем рассматривать только наиболее важный случай, когда пространства R и R_1 совпадают. В этом случае говорят, что оператор $\tilde{A}$ задан в n-мерном пространстве R и отображает это пространство в себя, так как образы векторов из R принадлежат тому же пространству. В этом случае $m = n$ и базис $\{\boldsymbol{g}\}$ будем выбирать совпадающим с базисом $\{\boldsymbol{e}\}$.

Тогда уравнения (4.4) принимают вид:

$$\begin{aligned}
\tilde{A}\,\bar{\boldsymbol{e}}_1 &= a_{11}\,\bar{\boldsymbol{e}}_1 + a_{21}\,\bar{\boldsymbol{e}}_2 + \ldots + a_{n1}\,\bar{\boldsymbol{e}}_n \\
\tilde{A}\,\bar{\boldsymbol{e}}_2 &= a_{12}\,\bar{\boldsymbol{e}}_1 + a_{22}\,\bar{\boldsymbol{e}}_2 + \ldots + a_{n2}\,\bar{\boldsymbol{e}}_n \\
&\ldots\ldots\ldots\ldots\ldots\ldots \\
\tilde{A}\,\bar{\boldsymbol{e}}_n &= a_{1n}\,\bar{\boldsymbol{e}}_1 + a_{2n}\,\bar{\boldsymbol{e}}_2 + \ldots + a_{nn}\bar{\boldsymbol{e}}_n.
\end{aligned} \tag{4.6}$$

Теперь матрица оператора $\tilde{A}$ — квадратная и зависит только от выбора базиса $\{\boldsymbol{e}\}$. Формулы (4.5) принимают вид

$$\begin{aligned}
y_1 &= a_{11}\,x_1 + a_{12}\,x_2 + \ldots + a_{1n}\,x_n, \\
y_2 &= a_{21}\,x_1 + a_{22}\,x_2 + \ldots + a_{2n}\,x_n, \\
&\ldots\ldots\ldots\ldots\ldots\ldots \\
y_n &= a_{n1}\,x_1 + a_{n2}x_2 + \ldots + a_{nn}\,x_n,
\end{aligned} \tag{4.7}$$

или в матричной форме

$$Y = AX, \text{ где } X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}; \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

Таким образом, *всякому линейному оператору $\tilde{A}$ в n-мерном пространстве при выбранном базисе соответствует некоторая квадратная матрица A n-го порядка и линейное преобразование с этой матрицей.*

Справедливо и обратное утверждение. *Всякой матрице A n-го порядка при заданном базисе соответствует некоторый линейный оператор.*

Действительно, из формулы (4.7) видно, что в выбранном базисе вектору $\boldsymbol{x}$ соответствует некоторый вектор $\boldsymbol{y}$, т. е. $\boldsymbol{y}=\tilde{A}\boldsymbol{x}$. Из свойств линейных преобразований (см. гл. 3, § 1) следует, что оператор $\tilde{A}$ — линейный.

Таким образом, установлено взаимно однозначное соответствие между линейными операторами в n-мерном пространстве и матрицами n-го порядка, а, значит, и линейными преобразованиями. Особо отметим разницу между выражениями $\boldsymbol{y}=\tilde{A}\boldsymbol{x}$ и $Y=AX$. Первое из них означает символическую запись правила, согласно которому вектор $\boldsymbol{x}$ преобразуется в вектор $\boldsymbol{y}$ (независимо от выбора базиса), а второе устанавливает соответствие между координатами векторов $\boldsymbol{x}$ и $\boldsymbol{y}$ в выбранном базисе.

В заключение отметим, что если оператор $\tilde{A}$ невырожденный, т. е. $r=n$, то $\det A \neq 0$.

Таким образом, матрица невырожденного оператора невырожденная.

5. Примеры линейных операторов. Приведем несколько примеров линейных операторов в n-мерном линейном пространстве и соответствующих этим операторам матриц.

1. Если для каждого вектора $\boldsymbol{x} \in R$

$$\tilde{A}\boldsymbol{x}=0,$$

то оператор $\tilde{A}$ является линейным и называется *нулевым оператором*. Так как для любого базиса $\{e\}$

$$\tilde{A}e_i=0 \quad i=1, 2, \ldots, n,$$

то матрица A нулевого оператора $\tilde{A}$ в любом базисе является нулевой матрицей

$$A=0.$$

2. Если для каждого вектора $\boldsymbol{x} \in R$

$$\tilde{A}\boldsymbol{x}=\boldsymbol{x},$$

то оператор $\tilde{A}$ является линейным. Он называется *тождественным оператором* и обозначается $\tilde{E}$. Так как для

любого базиса $\{e\}$

$$\tilde{E}e_i = e_i \quad i = 1, 2, \ldots, n,$$

то матрица тождественного оператора в любом базисе является единичной матрицей

$$A = E = (\delta_{ij}).$$

3. Если для каждого вектора $x \in R$

$$\tilde{A}x = \lambda x, \ *$$

то оператор $\tilde{A}$ является линейным и называется *оператором подобия.* Так как для любого базиса $\{e\}$

$$\tilde{A}e_i = \lambda e_i \quad i = 1, 2, \ldots, n,$$

то матрица оператора подобия

$$A = \lambda E = (\lambda\delta_{ij}).$$

4. Если для каждого $x \in T_n$, т. е. для каждого вектора

$$x = (\alpha_1, \alpha_2, \ldots, \alpha_m, \alpha_{m+1}, \ldots, \alpha_n) \quad m < n,$$
$$\tilde{A}x = (\alpha_1, \alpha_2, \ldots, \alpha_m, 0, 0, \ldots, 0),$$

то оператор $\tilde{A}$ линейный и называется *оператором проектирования* (вектора в n-мерном пространстве на m-мерное подпространство этого пространства).

Матрица этого оператора существенно зависит от выбора базиса. Например, в базисе

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

матрица оператора проектирования принимает вид

$$A = \begin{array}{c} \quad\quad m \\ \quad\quad \vdots \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & \ldots & 1 & 0 & \ldots & 0 \\ 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & \ldots & 0 & 0 & \ldots & 0 \end{pmatrix} \end{array}.$$

* Вектор λx называется *коллинеарным вектору* x.

Этот результат следует из того, что в данном случае

$$\tilde{A}e_i = e_i, \quad \text{если} \quad i \leqslant m,$$
$$\tilde{A}e_i = 0, \quad \text{если} \quad i > m.$$

Очевидно, ранг оператора проектирования равен m.

5. Для иллюстрации того, как выбор базиса влияет на вид матрицы линейного оператора, рассмотрим линейное пространство всех функций вида

$$x = f(t) = ae^t + be^{-t}$$

и зададим оператор $\tilde{A}$ соотношением

$$\tilde{A}x = ae^t,$$

тогда в базисе $e_1 = e^t$, $e_2 = e^{-t}$ имеем

$$\tilde{A}e_1 = e_1 = 1e_1 + 0\,e_2,$$
$$\tilde{A}e_2 = 0 = 0 \cdot e_1 + 0 \cdot e_2$$

и, следовательно,

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

В то же время в базисе

$$e_1^* = \operatorname{ch} t = \frac{1}{2}e^t + \frac{1}{2}e^{-t},$$
$$e_2^* = \operatorname{sh} t = \frac{1}{2}e^t - \frac{1}{2}e^{-t},$$
$$\tilde{A}e_1^* = \frac{1}{2}e^t = \frac{1}{2}e_1^* + \frac{1}{2}e_2^*,$$
$$\tilde{A}e_2^* = \frac{1}{2}e^t = \frac{1}{2}e_1^* + \frac{1}{2}e_2^*$$

и, следовательно,

$$A^* = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

6. Рассмотрим пространство V_2 и используем присущие этому пространству геометрические свойства образующих его векторов, в частности, понятие длины и угла между векторами.

Пусть в пространстве V_2 каждому вектору $\boldsymbol{x}$ ставится в соответствие вектор

$$\boldsymbol{y}=\tilde{A}\boldsymbol{x},$$

представляющий собой зеркальное отображение вектора $\boldsymbol{x}$ относительно некоторой фиксированной прямой γ (рис. 2), проходящей через точку O, которая принимается за начало всех векторов $\boldsymbol{x}\in V_2$.

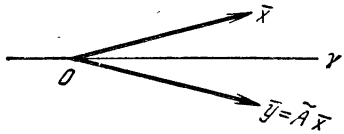


Рис. 2

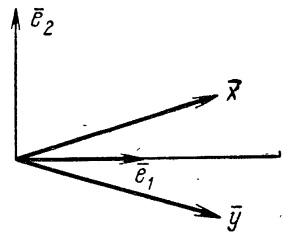


Рис. 3

Оператор $\tilde{A}$ в этом случае линейный и называется оператором зеркального отображения относительно прямой γ. Принимая в качестве базиса два единичных взаимно перпендикулярных вектора, один из которых направлен вдоль прямой γ (рис. 3), непосредственно находим, что

$$\tilde{A}\boldsymbol{e}_1=\boldsymbol{e}_1=1\cdot\boldsymbol{e}_1+0\cdot\boldsymbol{e}_2,$$
$$\tilde{A}\boldsymbol{e}_2=-\boldsymbol{e}_2=0\cdot\boldsymbol{e}_1-1\cdot\boldsymbol{e}_2$$

и, следовательно, в выбранном базисе матрица A имеет вид

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

7. В качестве примеров линейного оператора в бесконечномерных пространствах можно указать оператор интегрирования в пространстве $C(a, b)$, согласно которому каждому $\boldsymbol{x}=f(t)$ ставится в соответствие вектор $\boldsymbol{y}=\tilde{A}\boldsymbol{x}=\int\limits_a^t f(\tau)\,d\tau$ или оператор дифференцирования в пространстве всех бесконечно дифференцируемых на отрезке (a, b) функций, согласно которому вектору $\boldsymbol{x}=f(t)$ ставится в соответствие вектор

$$\boldsymbol{y}=\tilde{A}\boldsymbol{x}=f'(t).$$

§ 2. ДЕЙСТВИЯ С ЛИНЕЙНЫМИ ОПЕРАТОРАМИ

Подобно тому, как в гл. 2 были введены операции сложения и умножения линейных преобразований, можно ввести и аналогичные действия с линейными операторами. Заметим, что так как в n-мерном пространстве всякий линейный оператор порождает некоторое линейное преобразование с матрицей A (относительно выбранного базиса), то в конечномерных пространствах действия с линейными операторами аналогичны действиям с линейными преобразованиями.

1. Равенство операторов. Операторы $\tilde{A}$ и $\tilde{B}$ называются *равными*, что обозначается

$$\tilde{A}=\tilde{B},$$

если для любого $\boldsymbol{x}\in R$ справедливо равенство

$$\tilde{A}\boldsymbol{x}=\tilde{B}\boldsymbol{x}.$$

2. Сложение операторов. *Суммой* линейных операторов $\tilde{A}$ и $\tilde{B}$ называется оператор $\tilde{C}$, что обозначается

$$\tilde{C}=\tilde{A}+\tilde{B},$$

если для любого $\boldsymbol{x}\in R$ справедливо

$$\tilde{C}\boldsymbol{x}=(\tilde{A}+\tilde{B})\boldsymbol{x}=\tilde{A}\boldsymbol{x}+\tilde{B}\boldsymbol{x}.$$

Оператор $\tilde{C}$ — линейный, так как

$$\tilde{C}(\boldsymbol{x}+\boldsymbol{y})=\tilde{A}(\boldsymbol{x}+\boldsymbol{y})+\tilde{B}(\boldsymbol{x}+\boldsymbol{y})=\tilde{A}\boldsymbol{x}+\tilde{A}\boldsymbol{y}+$$
$$+\tilde{B}\boldsymbol{x}+\tilde{B}\boldsymbol{y}=(\tilde{A}\boldsymbol{x}+\tilde{B}\boldsymbol{x})+(\tilde{A}\boldsymbol{y}+\tilde{B}\boldsymbol{y})=\tilde{C}\boldsymbol{x}+\tilde{C}\boldsymbol{y}.$$
$$\tilde{C}(\lambda\boldsymbol{x})=\tilde{A}(\lambda\boldsymbol{x})+\tilde{B}(\lambda\boldsymbol{x})=\lambda\tilde{A}\boldsymbol{x}+\lambda\tilde{B}\boldsymbol{x}=$$
$$=\lambda(\tilde{A}\boldsymbol{x}+\tilde{B}\boldsymbol{x})=\lambda\tilde{C}\boldsymbol{x}.$$

3. Умножение оператора на число. *Произведением оператора $\tilde{A}$ на число* λ называется оператор $\tilde{B}$, что обозначается

$$\tilde{B}=\lambda\tilde{A},$$

если для любого $\boldsymbol{x}\in R$ справедливо равенство

$$\tilde{B}\boldsymbol{x}=\lambda(\tilde{A}\boldsymbol{x}).$$

Оператор $\tilde{B}$ является линейным, так как

$$\tilde{B}(x+y)=\lambda\tilde{A}(x+y)=\lambda(\tilde{A}x+\tilde{A}y)=\lambda\tilde{A}x+\lambda\tilde{A}y$$
$$\tilde{B}(\alpha x)=\lambda\tilde{A}(\alpha x)=\lambda(\alpha\tilde{A}x)=\lambda\alpha(\tilde{A}x)=\alpha(\lambda\tilde{A}x)=\alpha\tilde{B}x.$$

4. Умножение операторов. *Произведением линейных операторов* $\tilde{A}$ на $\tilde{B}$ называется оператор $\tilde{C}$, что обозначается

$$\tilde{C}=\tilde{A}\tilde{B},$$

если для любого $x\in R$ справедливо равенство

$$\tilde{C}x=\tilde{A}(\tilde{B}x),$$

т. е. сначала вектор x преобразуется в вектор $y=\tilde{B}x$, а затем — в вектор $z=\tilde{A}y$.

Оператор $\tilde{C}$ — линейный, так как

$$\tilde{C}(x+y)=\tilde{A}[\tilde{B}(x+y)]=\tilde{A}[\tilde{B}x+\tilde{B}y]=$$
$$=\tilde{A}(\tilde{B}x)+\tilde{A}(\tilde{B}y)=\tilde{C}x+\tilde{C}y$$

и

$$\tilde{C}(\lambda x)=\tilde{A}[\tilde{B}(\lambda x)]=\tilde{A}(\lambda\tilde{B}x)=\lambda\tilde{A}(\tilde{B}x)=\lambda\tilde{C}x.$$

Из теоремы о ранге оператора следует, что ранг произведения операторов $\tilde{A}\tilde{B}$ не больше ранга каждого сомножителя. В качестве упражнения предлагается читателю доказать, что введенные операции над линейными операторами обладают следующими свойствами:

1) $\tilde{A}+\tilde{B}=\tilde{B}+\tilde{A}$;
2) $(\tilde{A}+\tilde{B})+\tilde{C}=\tilde{A}+(\tilde{B}+\tilde{C})$;
3) $\tilde{A}+\tilde{0}=\tilde{A}$, где $\tilde{0}$ — нулевой оператор;
4) $\tilde{A}+(-1)\tilde{A}=\tilde{0}$;
5) $\lambda_1(\lambda_2\tilde{A})=(\lambda_1\lambda_2)\tilde{A}$;
6) $1\cdot\tilde{A}=\tilde{A}$;
7) $(\lambda_1+\lambda_2)\tilde{A}=\lambda_1\tilde{A}+\lambda_2\tilde{A}$;
8) $\lambda(\tilde{A}+\tilde{B})=\lambda\tilde{A}+\lambda\tilde{B}$;
9) $\lambda(\tilde{A}\tilde{B})=(\lambda\tilde{A})\tilde{B}$;
10) $(\tilde{A}\tilde{B})\tilde{C}=\tilde{A}(\tilde{B}\tilde{C})$;
11) $(\tilde{A}+\tilde{B})\tilde{C}=\tilde{A}\tilde{C}+\tilde{B}\tilde{C}$;
12) $\tilde{A}(\tilde{B}+\tilde{C})=\tilde{A}\tilde{B}+\tilde{A}\tilde{C}$.

В заключение заметим, что в общем случае

$$\tilde{A}\tilde{B} \neq \tilde{B}\tilde{A}.$$

5. Степень оператора. Если в произведении операторов $\tilde{A}\tilde{B}$ вместо $\tilde{B}$ подставить оператор $\tilde{A}$, то оператор $\tilde{A}\tilde{A}$ называется *квадратом оператора* A и обозначается $\tilde{A}^2$, т. е.

$$\tilde{A}^2 = \tilde{A}\tilde{A}.$$

Точно так же

$$\tilde{A}^3 = \tilde{A}^2\tilde{A}.$$

Продолжая этот процесс, получаем, что n-я степень оператора $\tilde{A}$, которая обозначается $\tilde{A}^n$, определяется равенством

$$\tilde{A}^n = \tilde{A}^{n-1}\tilde{A}.$$

Кроме того, по определению полагаем $\tilde{A}^0 = \tilde{E}$ — тождественный оператор.

6. Обратный оператор. Оператор $\tilde{B}$ называется *обратным* по отношению к оператору $\tilde{A}$, если $\tilde{A}\tilde{B} = \tilde{B}\tilde{A} = E$. В этом случае будем писать $\tilde{B} = \tilde{A}^{-1}$. Из определения следует, если $\boldsymbol{y} = \tilde{A}\boldsymbol{x}$, то $\boldsymbol{x} = \tilde{A}^{-1}\boldsymbol{y}$. Легко убедиться, что из линейности оператора $\tilde{A}$ следует линейность оператора $\tilde{A}^{-1}$. Так как ранг произведения $\tilde{A}\tilde{A}^{-1}$ равен рангу $\tilde{E}$, т. е. n, то ранги операторов $\tilde{A}$ и $\tilde{A}^{-1}$ равны n. Значит, оператор $\tilde{A}^{-1}$ — невырожденный.

Замечание. В силу эквивалентности оператора $\tilde{A}$ в конечномерном пространстве и некоторого линейного преобразования с матрицей A следует, что относительно одного и того же базиса:

1) если $\tilde{A} = \tilde{B}$, то $A = B$;
2) если $\tilde{C} = \tilde{A} + \tilde{B}$, то $C = A + B$;
3) если $\tilde{B} = \lambda\tilde{A}$, то $B = \lambda A$;
4) если $\tilde{C} = \tilde{A} \cdot \tilde{B}$, то $C = A \cdot B$;
5) если $\tilde{B} = \tilde{A}^n$, то $B = A^n$;
6) если $\tilde{B} = \tilde{A}^{-1}$, то $B = A^{-1}$.

Следовательно, обратный оператор существует тогда и только тогда, когда $\det A \neq 0$.

7. Операторы проектирования и зеркального отображения. Используя действия над операторами, можно дать

инвариантное определение (т. е. не зависящее от выбора базиса) некоторых операторов в n-мерном пространстве. Здесь мы рассмотрим два специальных вида операторов.

1. Оператор проектирования. Имеем пространство R и пусть $\boldsymbol{x} \in R$. Известно, что векторы $\boldsymbol{y} = \tilde{A}\boldsymbol{x}$ образуют подпространство $L = \tilde{A}R$. Если для любого $\boldsymbol{y} \in L$, $\tilde{A}\boldsymbol{y} = \boldsymbol{y}$, то будем называть вектор $\boldsymbol{y}$ проекцией вектора $\boldsymbol{x}$ на подпространство L с помощью оператора $\tilde{A}$. Оператор $\tilde{A}$, удовлетворяющий этим условиям, назовем *оператором проектирования*.

Введенное определение обобщает известное геометрическое представление о проектировании в пространствах V_2 и V_3: проекция проекции остается неизменной.

Из определения вытекает основное свойство оператора проектирования, а именно: для любого $\boldsymbol{x} \in R$

$$\tilde{A}^2\boldsymbol{x} = \tilde{A}\boldsymbol{y} = \boldsymbol{y} = \tilde{A}\boldsymbol{x}, \text{ т. е. } \tilde{A}^2 = \tilde{A}.$$

Отсюда получаем $A^2 = A$, где A — матрица оператора $\tilde{A}$ в некотором базисе.

Отметим, что введенные в примерах 4 и 5 операторы являются операторами проектирования, так как в обоих случаях $A^2 = A$, что легко проверить.

2. Оператор зеркального отображения. Аналогично определим оператор зеркального отображения векторов относительно некоторого подпространства L.

Если для любого $\boldsymbol{x} \in R$ $\boldsymbol{y} = \tilde{A}\boldsymbol{x}$, а $\boldsymbol{x} = \tilde{A}\boldsymbol{y}$ (т. е. $\tilde{A} = \tilde{A}^{-1}$ или $\tilde{A}^2 = \tilde{E}$), причем для векторов подпространства L

$$\tilde{A}\boldsymbol{x} = \boldsymbol{x},$$

то $\tilde{A}$ называется *оператором зеркального отображения относительно подпространства* L. Читатель легко убедится, что матрица оператора примера 6 (см. стр. 73) удовлетворяет равенству $A^2 = E$.

§ 3. ПРЕОБРАЗОВАНИЕ КООРДИНАТ

Ранее отмечалось, что координаты вектора зависят от выбранного базиса. Теперь определим, как связаны координаты одного и того же вектора в разных базисах.

Пусть в n-мерном пространстве выбраны два базиса:

$$\{e\} : \boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n \text{ — назовем его «}\textit{старым}\text{»}$$

и

$\{e^*\}: e_1^*, e_2^*, \ldots, e_n^*$, который назовем «*новым*».

Рассмотрим оператор $\tilde{T}$, который переводит базис $\{e\}$ в базис $\{e^*\}$, т. е.

$$e_k^* = \tilde{T} e_k. \tag{4.8}$$

Оператор $\tilde{T}$ — невырожденный, так как ранг системы векторов $\{e^*\}$ равен n.

Ему соответствует невырожденная матрица

$$T = (t_{ij}),$$

которая называется матрицей *перехода*.

Разложим теперь вектор x по старому и новому базисам

$$x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n,$$

$$x = x_1^* e_1^* + x_2^* e_2^* + \ldots + x_n^* e_n^*.$$

Учитывая равенство (4.8), имеем

$$x_1 e_1 + x_2 e_2 + \ldots + x_n e_n = x_1^* \tilde{T} e_1 + x_2^* \tilde{T} e_2 + \ldots + x_n^* \tilde{T} e_n.$$

Легко заметить, что написанное равенство совпадает с равенством (4.3), если в нем обозначить $y_i = x_i$, а $x_i = x_i^*$. Поэтому, проделав преобразование, как в § 1, п. 2, получим

$$X = TX^*. \tag{4.9}$$

Формула (4.9) выражает старые координаты вектора через новые.

Матрица T — невырожденная и поэтому из (4.9) имеем

$$X^* = T^{-1} X. \tag{4.10}$$

Формула (4.10) выражает новые координаты вектора через старые.

Примеры.

1. Рассмотрим двумерное пространство всех функций вида

$$f(t) = ae^t + be^{-t},$$

выбирая в качестве базисов $\{e\}$ и $\{e^*\}$ соответственно векторы

$$e_1 = e^t;\ e_2 = e^{-t};$$

$$e_1^* = \operatorname{ch} t = \frac{e^t + e^{-t}}{2};\quad e_2^* = \operatorname{sh} t = \frac{e^t - e^{-t}}{2}.$$

Тогда

$$e_1^* = \frac{1}{2} e_1 + \frac{1}{2} e_2,$$

$$e_2^* = \frac{1}{2} e_1 - \frac{1}{2} e_2$$

или

$$T = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}, \quad T^{-1} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Поэтому координаты вектора $f(t)$ в старом и новом базисе связаны соотношениями

$$\begin{pmatrix} a \\ b \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} a^* \\ b^* \end{pmatrix},$$

$$\begin{pmatrix} a^* \\ b^* \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix}.$$

2. Поворот системы координат. Возьмем на плоскости (пространство V_2) два единичных взаимно перпендикулярных вектора e_1, e_2. Они образуют базис. Построим новый базис e_1^*, e_2^*, который получается из старого поворотом на угол α.

Очевидно, имеем

$$e_1^* = \cos\alpha\, e_1 + \sin\alpha\, e_2,$$

$$e_2^* = -\sin\alpha\, e_1 + \cos\alpha\, e_2.$$

Значит,

$$T = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix},$$

$$T^{-1} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}.$$

Поэтому координаты вектора в старом и новом базисах связаны соотношениями

$$\begin{cases} x = x^* \cos\alpha - y^* \sin\alpha, \\ y = x^* \sin\alpha + y \cos\alpha, \end{cases} \qquad \begin{cases} x^* = x \cos\alpha + y \sin\alpha, \\ y^* = -x \sin\alpha + y \cos\alpha. \end{cases}$$

Таким образом, имеем обычные формулы поворота системы координат.

§ 4. ИЗМЕНЕНИЕ МАТРИЦЫ ЛИНЕЙНОГО ОПЕРАТОРА ПРИ ПЕРЕХОДЕ К НОВОМУ БАЗИСУ

Из примера 5 § 1 видно, что матрица оператора зависит от выбранного базиса.

Рассмотрим вопрос о зависимости матрицы оператора от базиса в общем случае.

1. Преобразование матрицы оператора. Пусть в n-мерном пространстве задан оператор

$$\boldsymbol{y}=\tilde{A}\boldsymbol{x}.$$

Выберем базис $\{\boldsymbol{e}\}$. В этом базисе оператору соответствует линейное преобразование $Y=AX$.

Введем новый базис $\{\boldsymbol{e}^*\}$. Пусть матрица перехода

$$T=(t_{ij}).$$

Тогда на основании (4.9) и (4.10) координаты векторов $\boldsymbol{x}$ и $\boldsymbol{y}$ в старом и новом базисах связаны соотношениями

$$X=TX^*, \quad Y=TY^*.$$

Подставляя X и Y в формулу $Y=AX$, получим

$$TY^*=ATX^*,$$

так как T — невырожденная матрица, то, умножая (слева) обе части на T^{-1}, находим

$$Y^*=T^{-1}ATX^*,$$

или, если обозначить $A^*=T^{-1}AT$, то

$$Y^*=A^*X^*.$$

Таким образом, мы получим линейное преобразование, соответствующее оператору $\tilde{A}$ в новом базисе. Это значит, что матрица A^* оператора $\tilde{A}$ относительно базиса $\{\boldsymbol{e}^*\}$ связана с матрицей A того же оператора относительно базиса $\{\boldsymbol{e}\}$ соотношением

$$A^*=T^{-1}AT.$$

Матрица перехода T называется при этом *преобразующей матрицей*. Очевидно, что

$$TA^*T^{-1}=A,$$

или

$$A=(T^{-1})^{-1}A^*T^{-1}.$$

2. Подобные матрицы. Матрицы A и $T^{-1}AT$ описывают действие одного и того же оператора $\tilde{A}$ в разных бази-

сах. Это обстоятельство указывает на внутреннее «родство» этих матриц. Квадратные матрицы A и $A^* = T^{-1}AT$ (где T — невырожденная матрица) называются *подобными*. Одним из важных свойств подобных матриц является равенство их определителей.

В самом деле,

$$|A^*| = |T^{-1}AT|.$$

Но определитель произведения матриц равен произведению их определителей, т. е.

$$|A^*| = |T^{-1}||A||T|,$$

и так как

$$|T^{-1}| = \frac{1}{|T|},$$

то отсюда следует, что

$$|A^*| = |A|.$$

Так как матрицы одного и того же линейного оператора в различных базисах являются подобными, то из равенства определителей подобных матриц в качестве следствия вытекает следующее утверждение:

Определитель матрицы линейного оператора не зависит от выбора базиса.

Заметим еще, что так как ранг матрицы оператора не зависит от выбранного базиса, то подобные матрицы имеют один и тот же ранг.

§ 5. ГЕОМЕТРИЧЕСКИЙ СМЫСЛ ОПРЕДЕЛИТЕЛЯ МАТРИЦЫ ОПЕРАТОРА В ПРОСТРАНСТВАХ V_3 И V_2.

Независимость определителя матрицы линейного оператора от выбора базиса указывает на то, что этот определитель связан с некоторыми внутренними свойствами оператора, не зависящими от такого привходящего фактора, как выбор базиса линейного пространства.

В пространствах V_2 и V_3 этому свойству определителя матрицы линейного оператора может быть дана чрезвычайно наглядная и простая геометрическая трактовка.

Однако при этом нам придется использовать ряд геометрических свойств пространств V_2 и V_3, которые не обязательно присущи любому линейному пространству (т. е. те, которые не вытекают из аксиом линейного простран-

ства). В частности, мы используем понятия площади параллелограмма, объема параллелепипеда и операции скалярного и векторного произведения векторов.

1. Изменение объема. Рассмотрим пространство V_3. Пусть оператор $\tilde{A}$ — невырожденный. Это означает, что он любые три некомпланарных вектора $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ переводит в три некомпланарных вектора $(\tilde{A}\boldsymbol{e}_1, \tilde{A}\boldsymbol{e}_2, \tilde{A}\boldsymbol{e}_3)$. Параллелепипед, построенный на векторах $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$, оператор $\tilde{A}$ в силу линейности преобразует в параллелепипед, построенный на их образах. Найдем соотношения между объемами этих параллелепипедов.

Как известно, ориентированный объем параллелепипеда, построенного на трех векторах, равен их смешанному произведению. Поэтому

$$V^{(1)} = \boldsymbol{e}_1\boldsymbol{e}_2\boldsymbol{e}_3,$$
$$V^{(2)} = \tilde{A}\boldsymbol{e}_1\tilde{A}\boldsymbol{e}_2\tilde{A}\boldsymbol{e}_3.$$

Выражая $\tilde{A}\boldsymbol{e}_1$, $\tilde{A}\boldsymbol{e}_2$ и $\tilde{A}\boldsymbol{e}_3$ по формулам (4.6), получаем

$$V^{(2)} = (a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{33} - \\ - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32})\cdot(\boldsymbol{e}_1\boldsymbol{e}_2\boldsymbol{e}_3), \text{ или}$$

$$V^{(2)} = V^{(1)}\det A,$$

где A — матрица оператора $\tilde{A}$. Отметим, что если $\det A > 0$, то оба параллелепипеда одинаково ориентированы, если же $\det A < 0$, то ориентация их разная.

Итак, *определитель матрицы оператора равен отношению ориентированных объемов соответствующих параллелепипедов в пространствах образов и прообразов. Иначе, он является коэффициентом искажения объема.*

2. Изменение площади. Пусть оператор $\tilde{A}$ переводит векторы $\boldsymbol{e}_1, \boldsymbol{e}_2$ пространства V_2 (неколлинеарные) в векторы $\tilde{A}\boldsymbol{e}_1$, $\tilde{A}\boldsymbol{e}_2$. Возьмем единичный вектор $\boldsymbol{e}_3$, перпендикулярный плоскости векторов $\boldsymbol{e}_1, \boldsymbol{e}_2$. Введем вспомогательный оператор $\tilde{A}_1$ (действующий в пространстве V_3) такой, что

$$\tilde{A}_1\boldsymbol{e}_1 = \tilde{A}\boldsymbol{e}_1,$$
$$\tilde{A}_1\boldsymbol{e}_2 = \tilde{A}\boldsymbol{e}_2,$$
$$\tilde{A}_1\boldsymbol{e}_3 = \boldsymbol{e}_3.$$

Тогда ориентированные объемы $V^{(1)}$ и $V^{(2)}$ численно равны ориентированным площадям параллелограммов $S^{(1)}$ и $S^{(2)}$,

построенных на векторах e_1, e_2 и их образах. Учитывая результат п. 1, получаем

$$S^{(2)} = S^{(1)} \det A_1.$$

Но легко видеть, что $\det A_1 = \det A$.

Значит, $S^{(2)} = \det A \cdot S^{(1)}$, т. е. $\det A$ является *коэффициентом искажения площади.*

Упражнения

1. $x = (x_1, x_2, x_3)$,

$$\tilde{A}x = (x_2 + x_3,\ 2x_1 + x_3,\ 3x_1 - x_2 + x_3).$$

Проверить, что оператор $\tilde{A}$ — линейный и найти его матрицу в базисе $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$.

2. Показать, что ортогональное проектирование трехмерного пространства на ось x есть линейный оператор. Найти его матрицу в базисе i, j, k.

3. Показать, что умножение квадратных матриц 2-го порядка на матрицу $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ есть линейный оператор. Найти его матрицу в базисе

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

4. Векторы $a_1 = (2, 3, 5)$, $a_2 = (0, 1, 2)$, $a_3 = (1, 0, 0)$ линейным оператором $\tilde{A}$ преобразуются соответственно в векторы $b_1 = (1, 1, 1)$, $b_2 = (1, 1, -1)$, $b_3 = (2, 1, 2)$.

Найти матрицу этого оператора в том же базисе, в каком указаны координаты векторов a и b.

5. Векторы e_1 и e_2 образуют базис пространства. Выберем за новый базис векторы

$$e_1^* = 2e_1 + 3e_2, \quad e_2^* = e_1 + e_2.$$

Найти формулы преобразования координат.

6. В пространстве многочленов степени не выше 2-ой выбраны два базиса $1, t, t^2$ и $1, t-1, \frac{(t-1)^2}{2}$.

Найти матрицу преобразования координат.

7. Вывести формулы преобразования координат при повороте прямоугольной системы координат в пространстве:

а) если известны углы между новыми осями и старыми осями:

	x'	y'	z
x	α	α_1	α_2
y	β	β_1	β_2
z	γ	γ_1	γ_2

б) если даны углы Эйлера:

θ — угол между осями z и z';

φ — угол между линией пересечения l плоскостей xOy и $x'Oy'$ и осью x;

ψ — угол между линией l и осью x'.

8. В базисе e_1, e_2, e_3, e_4 дана матрица оператора

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

Найти матрицу оператора в базисе из векторов:

а) e_1, e_2, e_3, e_4;

б) e_1, e_1+e_2, $e_1+e_2+e_3$, $e_1+e_2+e_3+e_4$.

9. Линейный оператор в базисе

$$e_1=(8,\ -6,\ 7);\ e_2=(-16,\ 7,\ -13);\ e_3=(9,\ -3,\ 7)$$

имеет матрицу

$$\begin{pmatrix} 1 & -18 & 15 \\ -1 & -22 & 15 \\ 1 & -25 & 22 \end{pmatrix}.$$

Найти его матрицу в базисе

$$e_1^*=(1,\ -2,\ 1),\ e_2^*=(3,\ -1,\ 2),\ e_3^*=(2,\ 1,\ 2).$$

10. Найти матрицу оператора дифференцирования в пространстве многочленов степени не выше 2-й в базисах:

а) $1,\ t,\ t^2$;

б) $1,\ t-1,\ \dfrac{(t-1)^2}{2}$;

в) показать, что матрицы оператора в первом базисе (A) и втором базисе (B) подобны.

11. В пространстве V_3 задан вектор $z=z_1\boldsymbol{i}+z_2\boldsymbol{j}+z_3\boldsymbol{k}$. Оператор $\tilde{A}$ определяется условием $\tilde{A}\boldsymbol{x}=\boldsymbol{z}\times\boldsymbol{x}$. Доказать, что оператор линейный и найти его матрицу в базисе $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$.

Глава 5

СОБСТВЕННЫЕ ВЕКТОРЫ И СОБСТВЕННЫЕ ЧИСЛА ЛИНЕЙНОГО ОПЕРАТОРА

§ 1. ХАРАКТЕРИСТИЧЕСКИЙ МНОГОЧЛЕН И ХАРАКТЕРИСТИЧЕСКОЕ УРАВНЕНИЕ МАТРИЦЫ

1. Характеристический многочлен. Рассмотрим квадратную матрицу

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

как было показано, все матрицы, подобные матрице A, т. е. все матрицы вида

$$A^*=T^{-1}AT,$$

где T — любая невырожденная матрица, обладают одним и тем же определителем

$$|A|=|A^*|.$$

Однако подобные матрицы обладают еще одной более общей для всех них характеристикой.

Наряду с матрицей A рассмотрим матрицу

$$A-\lambda E=\begin{pmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda \end{pmatrix},$$

которая образована из A заменой диагональных элементов a_{ii} элементами $a_{ii}-\lambda$, где λ — произвольное число. Определитель этой матрицы

$$|A-\lambda E|=\begin{vmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & & a_{nn}-\lambda \end{vmatrix}=\Delta(\lambda)$$

представляет собой многочлен степени n относительно λ (коэффициент при λ^n равен $(-1)^n$). Этот многочлен, кото-

рый мы обозначим $\Delta(\lambda)$, называется *характеристическим многочленом матрицы* A.

Алгебраическое уравнение n-й степени

$$\Delta(\lambda)=0$$

называется *характеристическим уравнением матрицы* A, а его корни — *характеристическими числами матрицы* A. Из основной теоремы алгебры следует, что *характеристическое уравнение имеет по меньшей мере один корень (действительный или комплексный)*.

2. Инвариантность характеристического многочлена. Покажем теперь, что все подобные матрицы имеют один и тот же характеристический многочлен, т. е.

$$|A^*-\lambda E|=|A-\lambda E|=\Delta(\lambda),$$

если

$$A^*=T^{-1}AT.$$

Для этого воспользуемся очевидным тождеством

$$E=T^{-1}ET.$$

Тогда, заменяя в матрице $A^*-\lambda E$ матрицы A^* и E — соответственно на $T^{-1}AT$ и $T^{-1}ET$, получаем

$$A^*-\lambda E=T^{-1}AT-\lambda T^{-1}ET=T^{-1}(A-\lambda E)T,$$

откуда следует, что матрицы $A-\lambda E$ и $A^*-\lambda E$ — подобные, а значит, их определители равны.

Таким образом, все подобные матрицы имеют один и тот же характеристический многочлен $\Delta(\lambda)$ и одно и то же характеристическое уравнение. Значит, характеристические числа матрицы оператора не зависят от выбора базиса. Поэтому их называют характеристическими числами оператора.

§ 2. СОБСТВЕННЫЙ ВЕКТОР И СОБСТВЕННОЕ ЧИСЛО ЛИНЕЙНОГО ОПЕРАТОРА

1. Собственный вектор. Пусть в линейном пространстве R над полем K задан линейный оператор $\tilde{A}$, тогда, если существует ненулевой вектор $\boldsymbol{x}\in R$ и число λ из K такие, что

$$\tilde{A}\boldsymbol{x}=\lambda\boldsymbol{x}, \tag{5.1}$$

то число λ называется *собственным числом* (или *собственным значением*) *линейного оператора*, а вектор $\boldsymbol{x}$ назы-

вается *собственным вектором оператора $\tilde{A}$ с собственным числом* λ. Это значит, что образом собственного вектора $\boldsymbol{x}$ является коллинеарный ему вектор $\lambda\boldsymbol{x}$.

Отметим некоторые свойства собственных векторов оператора $\tilde{A}$.

1. *Каждому собственному вектору соответствует единственное собственное число.*

Предположим обратное: пусть $\boldsymbol{x}$ — собственный вектор оператора $\tilde{A}$, а λ и λ_1, соответствующие вектору $\boldsymbol{x}$, — обственные числа. Это значит, что

$$\tilde{A}\boldsymbol{x}=\lambda\boldsymbol{x},$$
$$\tilde{A}\boldsymbol{x}=\lambda_1\boldsymbol{x}.$$

Но отсюда следует, что

$$\lambda\boldsymbol{x}-\lambda_1\boldsymbol{x}=0,$$

или

$$(\lambda-\lambda_1)\boldsymbol{x}=0.$$

Так как по условию $\boldsymbol{x}\neq 0$, то $\lambda=\lambda_1$.

2. *Если $\boldsymbol{x}_1$ и $\boldsymbol{x}_2$ — собственные векторы оператора $\tilde{A}$ с одним и тем же собственным числом λ, то их сумма $\boldsymbol{x}_1+\boldsymbol{x}_2$ также является собственным вектором оператора A с собственным числом λ.*

Действительно, так как

$$\tilde{A}\boldsymbol{x}_1=\lambda\boldsymbol{x}_1 \text{ и } \tilde{A}\boldsymbol{x}_2=\lambda\boldsymbol{x}_2,$$

то

$$\tilde{A}(\boldsymbol{x}_1+\boldsymbol{x}_2)=\tilde{A}\boldsymbol{x}_1+\tilde{A}\boldsymbol{x}_2=\lambda\boldsymbol{x}_1+\lambda\boldsymbol{x}_2=\lambda(\boldsymbol{x}_1+\boldsymbol{x}_2).$$

3. *Если $\boldsymbol{x}$ — собственный вектор оператора $\tilde{A}$ с собственным числом λ, то любой вектор $\alpha\boldsymbol{x}$, коллинеарный вектору $\boldsymbol{x}$, также является собственным вектором оператора $\tilde{A}$ с тем же самым собственным числом λ.*

Действительно,

$$\tilde{A}(\alpha\boldsymbol{x})=\alpha\tilde{A}\boldsymbol{x}=\alpha(\lambda\boldsymbol{x})=\lambda(\alpha\boldsymbol{x}).$$

Таким образом, каждому собственному числу λ соответствует бесчисленное множество (коллинеарных) собственных векторов. Отметим следующее: множество собственных векторов оператора $\tilde{A}$, соответствующих одному и тому же собственному числу, образует пространство, которое является подпространством пространства R (это следует из свойств 2 и 3).

2. Существование собственного вектора. Теорема. *В комплексном линейном пространстве R каждый линейный оператор $\tilde{A}$ имеет по меньшей мере один собственный вектор.*

Пусть $\tilde{A}$ — линейный оператор, заданный в n-мерном пространстве R, а $\boldsymbol{x}$ — собственный вектор этого оператора с собственным числом λ, т. е. $\tilde{A}\boldsymbol{x}=\lambda\boldsymbol{x}$. Выбирая произвольный базис $\{\boldsymbol{e}\}=(\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n)$, обозначим координаты вектора $\boldsymbol{x}$ в этом базисе через $x_1, x_2, \ldots, x_n$. Тогда, если $A=(a_{ij})$ — матрица оператора $\tilde{A}$ в базисе $\{\boldsymbol{e}\}$, то, записывая соотношение (5.1) в матричной форме, получаем

$$AX=\lambda X=\lambda EX,$$

или

$$(A-\lambda E)X=0,$$

где

$$X=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Переходя к координатной форме записи, имеем

$$\begin{aligned}
&(a_{11}-\lambda)x_1+a_{12}x_2+\ldots+a_{1n}x_n=0,\\
&a_{21}x_1+(a_{22}-\lambda)x_2+\ldots+a_{2n}x_n=0,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&a_{n1}x_1+a_{n2}x_2+\ldots+(a_{nn}-\lambda)x_n=0.
\end{aligned}$$

Для отыскания собственного вектора необходимо найти ненулевые решения этой системы уравнений, которые существуют тогда и только тогда, когда определитель системы

$$\Delta(\lambda)=\begin{vmatrix} a_{11}-\lambda & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22}-\lambda & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn}-\lambda \end{vmatrix}=|A-\lambda E|$$

равен нулю, т. е. когда $\Delta(\lambda)=0$. Значит, собственное число линейного оператора является его характеристическим числом, которое всегда существует. Подставляя это число в систему (5.2), найдем ненулевое решение этой системы, которое определит искомый собственный вектор.

Таким образом, для нахождения собственных векторов надо составить характеристический многочлен $\Delta(\lambda) = |A - \lambda E|$ и найти все его корни λ_i, которые будут собственными числами оператора. После этого каждое из найденных собственных чисел λ_i следует подставить вместо λ в систему (5.2) и найти все ее линейно независимые решения, которые и определяют собственные векторы, соответствующие собственному числу λ_i.

Число таких решений равно $n - r_i$, где r_i — ранг матрицы $A - \lambda_i E$. Отсюда следует, что размерность пространства собственных векторов, соответствующих одному и тому же собственному числу λ_i, равна $n - r_i$.

Замечание. Очевидно, что если оператор $\tilde{A}$ задан в вещественном линейном пространстве R, то теорема не всегда справедлива. (Так как в этом случае характеристический многочлен может не иметь ни одного вещественного корня.) Например, оператор поворота вектора на угол $\alpha \neq k\pi$ в пространстве V_2 не имеет ни одного собственного вектора. В то же время, если $\alpha = k\pi$, то любой вектор собственный с собственным числом $\lambda = (-1)^k$.

Пример.

Пусть

$$A = \begin{pmatrix} 7 & -12 & 6 \\ 10 & -19 & 10 \\ 12 & -24 & 13 \end{pmatrix}.$$ Найти собственные векторы.

Решение. Характеристическое уравнение

$$\Delta(\lambda) = \begin{vmatrix} 7-\lambda & -12 & 6 \\ 10 & -19-\lambda & 10 \\ 12 & -24 & 13-\lambda \end{vmatrix} = 0$$

имеет корни $\lambda_1 = \lambda_2 = 1$; $\lambda_3 = -1$.

Составим систему (5.2)

$$\begin{cases} (7-\lambda)x_1 - 12x_2 + 6x_3 = 0, \\ 10x_1 - (19+\lambda)x_2 + 10x_3 = 0, \\ 12x_1 - 24x_2 + (13-\lambda)x_3 = 0. \end{cases}$$

Подставляем $\lambda = 1$. Имеем

$$\begin{cases} 6x_1 - 12x_2 + 6x_3 = 0, \\ 10x_1 - 20x_2 + 10x_3 = 0, \\ 12x_1 - 24x_2 + 12x_3 = 0. \end{cases}$$

Ранг матрицы $(A - \lambda_1 E)$ этой системы $r_1 = 1$, т. е. система эквивалентна одному уравнению, которое, как легко видеть, имеет вид

$x_1 - 2x_2 + x_3 = 0$. Значит, система имеет два линейно независимых решения, например:

$$x_1 = (4,\ 5,\ 6),$$
$$x_2 = (3,\ 5,\ 7).$$

Подставим теперь в систему $\lambda_3 = -1$, имеем

$$\begin{cases} 8x_1 - 12x_2 + 6x_3 = 0, \\ 10x_1 - 18x_2 + 10x_3 = 0, \\ 12x_1 - 24x_2 + 14x_3 = 0. \end{cases}$$

Ранг матрицы $(A - \lambda_3 E)$ этой системы $r_3 = 2$. Значит, система эквивалентна двум уравнениям и ее частное решение имеет вид

$$x_3 = (3,\ 5,\ 6).$$

Таким образом, x_1, x_2, x_3 — собственные векторы оператора $\tilde{A}$. Отметим, что все собственные векторы, соответствующие собственному числу $\lambda = 1$, определяются равенством

$$x = \alpha x_1 + \beta x_2,$$

а собственному числу $\lambda = -1$ — равенством

$$x = \gamma x_3,$$

где α, β, γ — произвольные числа.

3. Другой метод нахождения собственных векторов. Пусть оператор $\tilde{A}$ в некотором базисе пространства R имеет матрицу $A = (a_{ij})$.

Как мы знаем, столбцы A составлены из координат образов базисных векторов. Но базисные векторы имеют координаты

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

Поэтому можно считать, что столбцы A — это образы соответствующих столбцов E.

Схематично это соответствие можно представить так:

$$E = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 1 \\ & \cdot & \cdot\ \cdot\ \cdot & \\ 0 & 0 & \ldots & 1 \end{pmatrix}.$$

$$\downarrow \quad \downarrow \qquad \downarrow$$

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ & \cdot & \cdot\ \cdot\ \cdot & \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}.$$

Очевидно, что если среди столбцов A имеется нулевой столбец, то соответствующий столбец E состоит из координат собственного вектора оператора $\tilde{A}$ с собственным числом $\lambda=0$.

Произведем теперь над столбцами матрицы E элементарные операции (см. гл. 2, § 3, п. 1). В результате получим матрицу E_1, эквивалентную матрице E. Те же элементарные операции произведем и над столбцами матрицы A. Получим матрицу A_1, эквивалентную A. В силу линейности оператора соответствие между столбцами матриц E_1 и A_1 сохраняется, т. е. столбцы матрицы A_1 — образы соответствующих столбцов матрицы E_1.

Допустим теперь, что в результате произведенных операций среди столбцов A_1 появился нулевой столбец. Тогда, очевидно, соответствующий столбец E_1 состоит из координат собственного вектора оператора $\tilde{A}$ с собственным значением $\lambda=0$. Поэтому для нахождения собственного вектора с $\lambda=0$ нам достаточно с помощью элементарных преобразований над столбцами матриц A и E получить матрицы A_1 и E_1, где A_1 содержит нулевой столбец.

Практически указанные элементарные операции удобно производить над столбцами матрицы $\begin{pmatrix} E \\ A \end{pmatrix}$.

Пример. Дано:

$$A=\begin{pmatrix} 8 & -12 & 6 \\ 10 & -18 & 10 \\ 12 & -24 & 14 \end{pmatrix}.$$

Найти собственный вектор, соответствующий собственному числу $\lambda=0$.

Решение. Рассмотрим матрицу

$$\begin{pmatrix} E \\ A \end{pmatrix}=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 8 & -12 & 6 \\ 10 & -18 & 10 \\ 12 & -24 & 14 \end{pmatrix}\sim\begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 2 & 1 \\ 0 & 0 & 6 \\ -10 & 2 & 10 \\ -20 & 4 & 14 \end{pmatrix}\sim\begin{pmatrix} 3 & 0 & 0 \\ 5 & 1 & 0 \\ 6 & 2 & 1 \\ 0 & 0 & 6 \\ 0 & 2 & 10 \\ 0 & 4 & 14 \end{pmatrix}.$$

Сначала умножили последний столбец на -4 и прибавили к первому, умноженному на 3; затем последний столбец умножили на 2 и прибавили ко второму. После этого второй столбец умножили на 5 и прибавили к первому. Значит, вектор (3, 5, 6) — собственный с собственным числом $\lambda=0$.

Отметим, что этим методом мы можем найти все линейно независимые собственные векторы, соответствующие собственному числу $\lambda = 0$.

Действительно, пусть ранг матрицы A есть r. Это означает, что оператор $\tilde{A}$ имеет $n - r$ линейно независимых собственных векторов, соответствующих собственному числу $\lambda = 0$. Но одновременно это означает (см. гл. 2, § 4, п. 4), что с помощью элементарных преобразований мы можем получить матрицу A_1, у которой $n - r$ столбцов будут нулевыми. Соответствующие им столбцы матрицы E_1 и будут искомыми собственными векторами (их линейная независимость следует из того, что матрица E, а потому и матрица E_1 — невырожденная). Мы рассмотрели метод нахождения собственных векторов с $\lambda = 0$. От этого ограничения легко избавиться. Очевидно, если $\lambda \neq 0$, то вместо элементарных преобразований над матрицей A надо производить элементарные преобразования над матрицей $A - \lambda E$.

Пример. Дано:

$$A = \begin{pmatrix} 7 & -12 & 6 \\ 10 & -19 & 10 \\ 12 & -24 & 13 \end{pmatrix}.$$

Найти линейно независимые собственные векторы.

Решение. Характеристическое уравнение

$$\Delta(\lambda) = \begin{vmatrix} 7-\lambda & -12 & 6 \\ 10 & -19-\lambda & 10 \\ 12 & -24 & 13-\lambda \end{vmatrix} = 0.$$

Его корни: $\lambda_1 = \lambda_2 = 1$; $\lambda_3 = -1$.

Ищем собственные векторы, соответствующие $\lambda = 1$. Имеем

$$\begin{pmatrix} E \\ A-E \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 6 & -12 & 6 \\ 10 & -20 & 10 \\ 12 & -24 & 12 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 6 & 0 & 0 \\ 10 & 0 & 0 \\ 12 & 0 & 0 \end{pmatrix}.$$

Значит, имеем два линейно независимых собственных вектора (2, 1, 0) и (—1, 0, 1). Найдем собственный вектор, соответствующий $\lambda = -1$. Имеем

$$\begin{pmatrix} E \\ A+E \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 8 & -12 & 6 \\ -10 & -18 & 10 \\ -12 & -24 & 14 \end{pmatrix} \sim \begin{pmatrix} 3 & 0 & 0 \\ 5 & 1 & 0 \\ 6 & 2 & 1 \\ 0 & 0 & 6 \\ 0 & 2 & 10 \\ 0 & 4 & 14 \end{pmatrix}.$$

Итак, вектор (3, 5, 6) — собственный.

Замечание. Указанный метод нахождения собственных векторов может быть использован для решения однородной системы $AX=0$, так как координаты собственных векторов с собственным значением $\lambda=0$ (по определению) удовлетворяют этой системе (см. также задачу 13 гл. 2).

4. Свойства оператора $\tilde{A}-\mu\tilde{E}$. В дальнейшем большую роль будет играть оператор $\tilde{A}-\mu\tilde{E}$. Изучим его свойства. Пусть $\boldsymbol{x}\in R$ и $\tilde{A}\boldsymbol{x}=\boldsymbol{y}$. Тогда $(\tilde{A}-\mu\tilde{E})\boldsymbol{x}=\boldsymbol{y}-\mu\boldsymbol{x}$. Из последнего соотношения следует, что если $\boldsymbol{x}$ — собственный вектор оператора $\tilde{A}$ с собственным числом λ, то

$$(\tilde{A}-\mu\tilde{E})\boldsymbol{x}=(\lambda-\mu)\boldsymbol{x}.$$

Значит, операторы $\tilde{A}$ и $\tilde{A}-\mu\tilde{E}$ имеют одни и те же собственные векторы, собственные числа которых отличаются на μ.

В частности, если $\mu=\lambda$, то $(\tilde{A}-\lambda\tilde{E})\boldsymbol{x}=0$, и, следовательно, оператор $\tilde{A}-\lambda\tilde{E}$ — вырожденный и аннулирует собственный вектор, соответствующий собственному числу λ.

Аналогично, если $\boldsymbol{x}$ — собственный вектор оператора $\tilde{A}$, соответствующий собственному числу λ, то

$$(\tilde{A}-\lambda_1\tilde{E})(\tilde{A}-\lambda_2\tilde{E})\boldsymbol{x}=(\tilde{A}-\lambda_1\tilde{E})(\lambda-\lambda_2)\boldsymbol{x}=$$
$$=(\lambda-\lambda_1)(\lambda-\lambda_2)\boldsymbol{x}.$$

Значит вектор $\boldsymbol{x}$ является собственным вектором оператора $(\tilde{A}-\lambda_1\tilde{E})(\tilde{A}-\lambda_2\tilde{E})$ с собственным числом $(\lambda-\lambda_1)(\lambda-\lambda_2)$. В частности, если λ совпадает с одним из чисел λ_1 или λ_2, то оператор $(\tilde{A}-\lambda_1\tilde{E})(\tilde{A}-\lambda_2\tilde{E})$ аннулирует этот вектор.

Методом математической индукции нетрудно показать, что вектор $\boldsymbol{x}$ будет собственным вектором оператора

$$\tilde{B}_{st}=(\tilde{A}-\lambda_1\tilde{E})\ldots(\tilde{A}-\lambda_{s-1}\tilde{E})(\tilde{A}-\lambda_{s+1}\tilde{E})\ldots(\tilde{A}-\lambda_t\tilde{E})$$

с собственным числом $(\lambda-\lambda_1)\ldots(\lambda-\lambda_{s-1})(\lambda-\lambda_{s+1})\ldots$ $\ldots(\lambda-\lambda_t)$, т. е.

$$\tilde{B}_{st}\boldsymbol{x}=(\lambda-\lambda_1)\ldots(\lambda-\lambda_{s-1})(\lambda-\lambda_{s+1})\ldots(\lambda-\lambda_t)\boldsymbol{x}.$$

Если λ совпадает с одним из чисел $\lambda_1,\ldots,\lambda_{s-1},\lambda_{s+1},\ldots,\lambda_t$, то очевидно $\tilde{B}_{st}\boldsymbol{x}=0$.

В заключение докажем следующее утверждение.

Теорема. *Собственные векторы $\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots\boldsymbol{x}_m$ оператора $\tilde{A}$ с попарно различными собственными числами $\lambda_1,\lambda_2,\ldots\lambda_m$, ($\lambda_i\neq\lambda_j$, если $i\neq j$) линейно независимы.*

Доказательство. Предположим противное. Пусть

$$\alpha_1\boldsymbol{x}_1+\alpha_2\boldsymbol{x}_2+\ldots+\alpha_m\boldsymbol{x}_m=0$$

и хотя бы одно из чисел, например α_s, не равно нулю. Тогда

$$\tilde{B}_{sm}(\alpha_1\boldsymbol{x}_1+\alpha_2\boldsymbol{x}_2+\ldots+\alpha_m\boldsymbol{x}_m)=0,$$

или

$$\alpha_s(\lambda_s-\lambda_1)\ldots(\lambda_s-\lambda_{s-1})(\lambda_s-\lambda_{s+1})\ldots(\lambda_s-\lambda_m)=0.$$

Отсюда $\alpha_s=0$, что противоречит предположению.

5. Матрица оператора в базисе, содержащем собственные векторы. Как мы видели (§ 2, п. 2), в пространстве R оператор $\tilde{A}$ имеет $n-r_i$ линейно независимых собственных векторов $\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_{n-r_i}$ с собственным числом λ_i. Образуем базис пространства из этих векторов и векторов $\boldsymbol{y}_1,\boldsymbol{y}_2,\ldots\boldsymbol{y}_{r_i}$ (это всегда можно сделать. См. гл. 3, § 4, п. 2). В этом базисе $\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots\boldsymbol{x}_{n-r_i},\boldsymbol{y}_1,\boldsymbol{y}_2,\ldots\boldsymbol{y}_{r_i}$, в силу соотношений

$$\tilde{A}\boldsymbol{x}_j=\lambda_i\boldsymbol{x}_j;\quad j=1,2,\ldots n-r_i;$$

$$\tilde{A}\boldsymbol{y}_k=a_{1k}\boldsymbol{x}_1+\ldots+a_{n-r_i,k}\boldsymbol{x}_{n-r_i}+b_{1k}\boldsymbol{y}_1+\ldots+b_{r_i,k}\boldsymbol{y}_{r_i},$$

оператор $\tilde{A}$ имеет матрицу

$$A=\left(\begin{array}{cccc|cccc}
\lambda_i & 0 & \ldots & 0 & a_{11} & a_{12} & \ldots & a_{1r_i}\\
0 & \lambda_i & \ldots & 0 & a_{21} & a_{22} & \ldots & a_{2r_i}\\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\
0 & 0 & \ldots & \lambda_i & a_{n-r_i,1} & a_{n-r_i,2} & \ldots & a_{n-r_i,r_i}\\
\hline
0 & 0 & \ldots & 0 & b_{11} & b_{12} & \ldots & b_{1r_i}\\
0 & 0 & \ldots & 0 & b_{21} & b_{22} & \ldots & b_{2r_i}\\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\
0 & 0 & \ldots & 0 & b_{r_i,1} & b_{r_i,2} & \ldots & b_{r_i,r_i}
\end{array}\right),$$

или сокращенно

$$A=\begin{pmatrix} I_i & A_i \\ 0 & B_i \end{pmatrix}.$$

Проделав очевидные выкладки, легко получить характеристический многочлен

$$\Delta(\lambda)=|A-\lambda E|=(\lambda_i-\lambda)^{n-r_i}|B_i-\lambda E|.$$

Отсюда, в частности, следует, что *кратность k_i собственного числа λ_i не меньше, чем $n-r_i$, т. е. $0<n-r_i\leqslant k_i$*. В заключение отметим, что включение в базис *собственных векторов упрощает вид матрицы оператора*, а именно: в $n-r_i$ столбцах матрицы A все элементы, кроме диагональных, равны нулю.

6. Оператор $\tilde{A}$ в пространстве $L=(\tilde{A}-\lambda_i\tilde{E})R$. Пространство L является подпространством размерности r_i пространства R.

Из теорем о ранге оператора (гл. 4, § 1, п. 3) следует, что из образов базисных векторов пространства R можно построить базис пространства L. Выберем в R базис, указанный в п. 5. Тогда из соотношений п. 4 следует, что $(\tilde{A}-\lambda_i\tilde{E})\,\boldsymbol{x}_j=0$ и поэтому за базис L можно выбрать векторы $(\tilde{A}-\lambda_i\tilde{E})\,\boldsymbol{y}_k=\boldsymbol{g}_k$, $k=1,2,\ldots r_i$. Найдем матрицу оператора $\tilde{A}$ в этом базисе. Для этого достаточно найти образы базисных векторов $\boldsymbol{g}_k$. Имеем $\tilde{A}\boldsymbol{g}_k=\tilde{A}\,(\tilde{A}-\lambda_i\tilde{E})\,\boldsymbol{y}_k=(\tilde{A}-\lambda_i\tilde{E})\,\tilde{A}\boldsymbol{y}_k$ (мы воспользовались тем, что операторы $\tilde{A}$ и $\tilde{A}-\lambda_i\tilde{E}$, как легко видеть, перестановочны).

Подставляя в это соотношение $\tilde{A}\boldsymbol{y}_k$ из п. 5, получаем

$$\tilde{A}\boldsymbol{g}_k=b_{1k}\boldsymbol{g}_1+b_{2k}\boldsymbol{g}_2+\ldots+b_{r_i,k}\boldsymbol{g}_{r_i}.$$

Таким образом, в выбранном базисе пространства L оператор $\tilde{A}$ имеет матрицу

$$A_L=\begin{pmatrix} b_{11} & b_{12} & \ldots & b_{1,r_i} \\ b_{21} & b_{22} & \ldots & b_{2,r_i} \\ \cdot & \cdot & \cdot & \cdot \\ b_{r_i,1} & b_{r_i,2} & \ldots & b_{r_i,r_i} \end{pmatrix},$$

которая является клеткой B_i матрицы A п. 5.

Отсюда следует, что характеристические многочлены оператора $\tilde{A}$ в пространствах R и L связаны соотно-

шением

$$\Delta_L(\lambda) = |A_L - \lambda E| = \frac{\Delta(\lambda)}{(\lambda_i - \lambda)^{n-r_i}}.$$

Поэтому *кратность корня λ_i характеристического многочлена оператора $\tilde{A}$ при переходе от пространства R к пространству L уменьшается на $n - r_i$* и становится равной $k_i + r_i - n$, кратность же остальных корней остается неизменной.

§ 3. ОПЕРАТОР ПРОСТОЙ СТРУКТУРЫ

Если линейный оператор $\tilde{A}$ в n-мерном пространстве R имеет n линейно независимых собственных векторов, то он называется оператором простой структуры.

1. Матрица оператора простой структуры. Докажем следующее утверждение.

Теорема. *Оператор $\tilde{A}$ простой структуры однозначно определен, если заданы его n линейно независимых собственных векторов и соответствующие им собственные числа.*

Действительно, выбирая собственные векторы оператора в качестве базиса в пространстве R, получим

$$\begin{aligned} \tilde{A}e_1 &= \lambda_1 e_1, \\ \tilde{A}e_2 &= \lambda_2 e_2, \\ &\dots\dots \\ \tilde{A}e_n &= \lambda_n e_n. \end{aligned}$$

А это значит, что матрица I линейного оператора в выбранном базисе имеет вид

$$I = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}.$$

Следовательно, оператор $\tilde{A}$ однозначно определен, а его матрица относительно базиса из собственных векторов является диагональной. Заметим, что среди чисел λ_i могут быть и одинаковые.

Если через A обозначить матрицу оператора $\tilde{A}$ в произвольном базисе $e_1^*, e_2^*, \dots e_n^*$, то

$$A = T^{-1}\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} T,$$

где T — матрица перехода от базиса $\{e\}$ из собственных векторов к базису $\{e^*\}$, т. е. матрица оператора простой структуры подобна диагональной матрице. Очевидно, справедливо и обратное утверждение. *Любая матрица, подобная диагональной, является матрицей некоторого оператора простой структуры.*

Поэтому, если оператор $\tilde{A}$ имеет в некотором базисе $\{e\}$ матрицу A, то в базисе $\{e^*\}$ из собственных векторов он имеет матрицу

$$I = T^{-1}AT.$$

Здесь T — матрица перехода от базиса $\{e\}$ к базису из собственных векторов $\{e^*\}$.

Пример.

Пусть,

$$A = \begin{pmatrix} 7 & -12 & 6 \\ 10 & -19 & 10 \\ 12 & -24 & 13 \end{pmatrix}.$$

Как мы видели (пример п. 2, § 2), собственные векторы матрицы A таковы:

$$x_1 = (4,\ 5,\ 6);\quad x_2 = (3,\ 5,\ 7);\quad x_3 = (3,\ 5,\ 6).$$

Эти векторы линейно независимы, так как

$$\begin{vmatrix} 4 & 3 & 3 \\ 5 & 5 & 5 \\ 6 & 7 & 6 \end{vmatrix} \neq 0.$$

Значит, оператор $\tilde{A}$ имеет простую структуру.

Выберем x_1, x_2, x_3 за базис пространства. В этом базисе матрица имеет диагональный вид

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

В этом можно убедиться непосредственно. В данном случае

$$T = \begin{pmatrix} 4 & 3 & 3 \\ 5 & 5 & 5 \\ 6 & 7 & 6 \end{pmatrix} \quad \text{и} \quad T^{-1}AT = I.$$

2. Достаточный признак простой структуры оператора. *Если все корни характеристического уравнения различны, то оператор имеет простую структуру.*

Действительно, в этом случае оператор имеет n по-

парно различных собственных чисел и, следовательно, n линейно независимых собственных векторов.

3. Геометрический смысл оператора простой структуры. Действие оператора простой структуры можно описать следующим образом. В пространстве R имеется n таких «направлений», что каждый из n линейно независимых векторов, имеющих одно из этих «направлений», преобразуется оператором в вектор, ему коллинеарный.

Произвольный вектор $\boldsymbol{x}$ преобразуется по формуле

$$\tilde{A}\boldsymbol{x} = \tilde{A}(x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + \ldots + x_n\boldsymbol{e}_n) = \\ = \lambda_1 x_1 \boldsymbol{e}_1 + \lambda_2 x_2 \boldsymbol{e}_2 + \ldots + \lambda_n x_n \boldsymbol{e}_n.$$

Отсюда видно, что действие оператора приводит к умножению i-й компоненты вектора $\boldsymbol{x}$ на λ_i.

4. Критерий простоты структуры оператора. Найдем теперь необходимые и достаточные условия, при которых оператор имеет простую структуру.

Теорема. *Для того чтобы оператор $\tilde{A}$ имел простую структуру, необходимо и достаточно, чтобы для каждого корня λ_i характеристического уравнения кратности k_i ранг r_i матрицы $(A - \lambda_i E)$ был равен $n - k_i$.*

Необходимость. Если матрица A имеет простую структуру, то

$$A = T^{-1} I T.$$

Тогда

$$A - \lambda_i E = T^{-1}(I - \lambda_i E)\, T.$$

Следовательно, матрицы $A - \lambda_i E$ и $I - \lambda_i E$ подобны, а потому имеют одинаковый ранг.

Ранг же матрицы

$$I - \lambda_i E = \begin{pmatrix} \lambda_1 - \lambda_i & 0 & \ldots & 0 \\ 0 & \lambda_2 - \lambda_i & \ldots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \ldots & \lambda_n - \lambda_i \end{pmatrix}$$

равен числу отличных от нуля ее элементов, т. е. числу корней характеристического уравнения, не равных λ_i, т. е. равен $n - k_i$.

Достаточность. Пусть $\lambda_1, \lambda_2, \ldots \lambda_\nu$ — различные собственные числа оператора $\tilde{A}$. Как мы видели (см. § 2, п. 2), собственные векторы с собственным числом λ_i образуют подпространство размерности $n - r_i$ пространства

R, так как по условию $n - r_i = k_i$, то оператор $\tilde{A}$ имеет k_i линейно независимых собственных векторов с собственным числом λ_i $(i = 1, 2, \ldots, \nu)$.

Таким образом, мы имеем всего $k_1 + k_2 + \ldots + k_\nu = n$ собственных векторов $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$. Покажем, что они линейно независимы.

Предположим противное. Пусть $\sum_{j=1}^{n} \alpha_j \boldsymbol{x}_j = 0$, где, например, $\alpha_1 \neq 0$. Тогда из соотношения п. 4, § 2, имеем

$$\tilde{B}_{1\nu} \sum_{j=1}^{n} \alpha_j \boldsymbol{x}_j = \sum_{j=1}^{k_1} \alpha_j (\lambda_1 - \lambda_2) \ldots (\lambda_1 - \lambda_\nu) \boldsymbol{x}_j = 0.$$

Отсюда $\sum_{j=1}^{k_1} \alpha_j \boldsymbol{x}_j = 0$, $\alpha_1 \neq 0$, что противоречит линейной независимости векторов, отвечающих собственному числу λ_1.

Пример.

Возвратимся к примеру (см. § 2, п. 2). Мы имели $\lambda_1 = \lambda_2 = 1$, $\lambda_3 = -1$. Значит, $k_1 = 2$, $k_3 = 1$. Далее, $r_1 = 1$, $r_3 = 2$. Так как $n = 3$, то, очевидно, $n - k_i = r_i$, т. е. оператор имеет простую структуру (в примере § 3, п. 1 мы это установили непосредственно).

5. Построение базиса из собственных векторов оператора простой структуры. Из определения оператора простой структуры следует, что можно построить базис пространства R из собственных векторов этого оператора.

Координаты собственных векторов в некотором базисе можно найти, решая системы уравнений

$$(A - \lambda_i E) X = 0,$$

где λ_i $(i = 1, 2, \ldots \nu)$ — различные собственные значения оператора. В найденном базисе из собственных векторов оператор имеет диагональную матрицу.

Мы укажем еще один способ построения базиса из собственных векторов, который не требует решения систем уравнений. Этот способ в дальнейшем будет применен к построению базиса, в котором оператор, не имеющий простой структуры, имеет в некотором смысле простейшую матрицу. Предлагаемый способ построения базиса из собственных векторов состоит в нахождении подпространства пространства R, целиком состоящего из собственных векторов оператора, соответствующих собственному значению λ_i.

Рассмотрим подпространство $(\tilde{A}-\lambda_j\tilde{E})R$ $(j\neq i)$, размерность которого r_j. В силу того, что оператор имеет простую структуру $n-r_j=k_j$, это означает, что в подпространстве $(\tilde{A}-\lambda_j\tilde{E})R$ оператор не имеет собственных векторов, отвечающих собственному значению λ_j (на основании результатов п. 6, § 2 характеристическое число имеет кратность $r_j+k_j-n=0$.) Рассуждая аналогично, приходим к заключению, что подпространство

$$(\tilde{A}-\lambda_1\tilde{E})\ldots(\tilde{A}-\lambda_{i-1}\tilde{E})(\tilde{A}-\lambda_{i+1}\tilde{E})\ldots(\tilde{A}-\lambda_\nu\tilde{E})R=\tilde{B}_{i\nu}R,$$

размерность которого равна

$$n-k_1-k_2-\ldots-k_{i-1}-k_{i+1}-\ldots-k_\nu=k_i,$$

не содержит собственных векторов с собственными значениями $\lambda_1, \ldots, \lambda_{i-1}, \lambda_{i+1}, \ldots, \lambda_\nu$, а следовательно, оно состоит только из собственных векторов оператора $\tilde{A}$, отвечающих собственному значению λ_i. Действительно в силу соотношения $n-r_i=k_i$ в этом подпространстве можно выбрать базис только из собственных векторов, отвечающих λ_i, следовательно, любой ненулевой вектор этого подпространства собственный.

Таким образом, используя операторы $\tilde{B}_{1\nu}, \tilde{B}_{2\nu}, \ldots, \tilde{B}_{\nu\nu}$ мы производим расслоение пространства R на подпространства $\tilde{B}_{1\nu}R, \tilde{B}_{2\nu}R, \ldots, \tilde{B}_{\nu\nu}R$ (размерности $k_1, k_2, \ldots, k_\nu$), каждое из которых состоит только из собственных векторов оператора $\tilde{A}$ отвечающих одному собственному числу. Выбирая базис в каждом подпространстве $\tilde{B}_{i\nu}R$, получим базис всего пространства (линейная независимость построенных векторов доказывается так же, как в п. 4).

Таким образом, для отыскания всех k_i линейно независимых собственных векторов оператора $\tilde{A}$ простой структуры, отвечающих собственному значению λ_i, достаточно построить любой базис пространства $\tilde{B}_{i\nu}R$. Для этого выберем произвольный базис пространства R, составленный из векторов $e_1^*, e_2^*, \ldots, e_n^*$, координаты которых относительно исходного базиса $e_1, e_2, \ldots, e_n$ (т. е. базиса, в котором задана матрица A оператора) соот-

ветственно равны

$$e_1^* = \{t_{11}, t_{12}, \ldots, t_{1n},\}$$
$$e_2^* = \{t_{21}, t_{22}, \ldots, t_{2n},\}$$
$$\cdots\cdots\cdots$$
$$e_n^* = \{t_{n1}, t_{n2}, \ldots, t_{nn}\}.$$

Этот произвольный базис будем задавать матрицей

$$T = \begin{pmatrix} t_{11} & t_{21} & \ldots & t_{n1} \\ t_{12} & t_{22} & \ldots & t_{n2} \\ \cdot & \cdot & \cdot & \cdot \\ t_{1n} & t_{2n} & \ldots & t_{nn} \end{pmatrix}$$

и обозначать в дальнейшем

$$e_i^* = (t_{i1}, t_{i2}, \ldots t_{in}).$$

Найдем теперь образы векторов e_1^*, e_2^*, ... e_n^* в пространстве $\tilde{B}_{i\nu}R$:

$$\tilde{B}_{i\nu}e_1^*, \tilde{B}_{i\nu}e_2^*, \ldots \tilde{B}_{i\nu}e_n^*,$$

среди которых всегда найдется k_i линейно независимых, которые и примем в качестве базиса пространства $\tilde{B}_{i\nu}R$ (гл. 3, § 1, п. 3). Заметим, что каждый из векторов $\tilde{B}_{i\nu}e_j^*$ можно представить как столбцовую матрицу, являющуюся j-столбцом матрицы $B_{i\nu}T$. Так как T — произвольная невырожденная матрица, то мы можем положить $T = E$ или $e_i^* = e_i$. В этом случае матрица $B_{i\nu}T$ принимает вид $B_{i\nu}E = B_{i\nu}$, а это значит, что в качестве базиса пространства $B_{i\nu}R$ можно выбрать (в матричной форме) любые k_i — линейно независимых столбцов матрицы $B_{i\nu}T$:

$$B_{i\nu}T = (A - \lambda_1 E) \ldots (A - \lambda_{i-1}E)(A - \lambda_{i+1}E) \ldots (A - \lambda_\nu E)\, T.$$

Из предыдущих рассуждений следует утверждение.

Теорема. *Для того чтобы оператор имел простую структуру, необходимо и достаточно, чтобы*

$$(\tilde{A} - \lambda_1\tilde{E})(\tilde{A} - \lambda_2\tilde{E}) \ldots (\tilde{A} - \lambda_\nu\tilde{E})\, R = (\tilde{A} - \lambda_i\tilde{E})\, \tilde{B}_{i\nu}R = 0.$$

Необходимость непосредственно вытекает из предыдущих рассуждений. Действительно, пространство $B_{i\nu}R$ состоит только из собственных векторов оператора $\tilde{A}$, отвечающих собственному числу λ_i, а поэтому оператор $\tilde{A} - \lambda_i\tilde{E}$ аннулирует это пространство.

Достаточность. Из соотношения $(\tilde{A}-\lambda_i\tilde{E})\tilde{B}_{i\nu}R=0$ следует, что $\tilde{B}_{i\nu}R$ состоит только из собственных векторов оператора, отвечающих λ_i. Размерность этого пространства равна кратности k_i корня λ_i, так как операторы $\tilde{A}-\lambda_j\tilde{E}$ $(j\neq i)$, входящие в $\tilde{B}_{i\nu}$, не могут изменить кратности корня λ_i. Строя базис в каждом подпространстве $\tilde{B}_{i\nu}R$ $(i=1, 2, \ldots \nu)$, мы получим базис всего пространства (линейная независимость построенных векторов следует из предыдущих рассуждений, а их количество равно $k_1+k_2+\ldots+k_\nu=n$).

Пример.

Пусть матрица A оператора $\tilde{A}$ имеет вид (см. пример § 2, п. 2)

$$A=\begin{pmatrix} 7 & -12 & 6 \\ 10 & -19 & 10 \\ 12 & -24 & 13 \end{pmatrix}.$$

Характеристическое уравнение

$$|A-\lambda E|=\begin{vmatrix} 7-\lambda & -12 & 6 \\ 10 & -19-\lambda & 10 \\ 12 & -24 & 13-\lambda \end{vmatrix}=0$$

имеет корни

$$\lambda_1=\lambda_2=1;\quad \lambda_3=-1.$$

Так как для единственного кратного корня $\lambda_1=1$ (кратности $k_1=2$) матрица

$$A-\lambda_1E=\begin{pmatrix} 6 & -12 & 6 \\ 10 & -20 & 10 \\ 12 & 24 & 12 \end{pmatrix}$$

имеет ранг $r_1=1$ и $r_1=n-k_1$, то A является оператором простой структуры.

Рассмотрим теперь матрицы

$$A-\lambda_1E=A-E \text{ и } A-\lambda_3E=A+E.$$

Первая из них (ранга 1) имеет один линейно независимый столбец, определяющий собственный вектор x_3, соответствующий собственному значению λ_3. В качестве x_3 можно принять, например,

$$x_3=(6,\ 10,\ 12).$$

Вторая матрица

$$A+E=\begin{pmatrix} 8 & -12 & 6 \\ 10 & -18 & 10 \\ 12 & -24 & 14 \end{pmatrix}$$

ранга $r_3=2$ содержит два линейно независимых столбца, определяющих собственные векторы x_1 и x_2, отвечающие собственному значению λ_1. В качестве x_1 и x_2 можно принять $x_1=(8, 10, 12)$ и $x_2=(6, 10, 14)$.

Найденные собственные векторы лишь множителями отличаются от собственных векторов примера из § 2, п. 2.

6. Условие отсутствия простой структуры оператора. В дальнейшем понадобится условие, при выполнении которого оператор $\tilde{A}$ не является оператором простой структуры. Это условие сформулируем в следующем виде.

Теорема. *Чтобы оператор $\tilde{A}$ не имел простой структуры, необходимо и достаточно, чтобы существовали вектор y_i и собственный вектор x_i (с собственным числом λ_i), удовлетворяющие условию*

$$(\tilde{A}-\lambda_i\tilde{E})\,y_i=x_i.$$

Необходимость. Пусть $(\tilde{A}-\lambda_i\tilde{E})\,\tilde{B}_{i\nu}R\neq 0$, т. е. оператор не имеет простой структуры. Тогда в подпространстве $L=(\tilde{A}-\lambda_i\tilde{E})\,\tilde{B}_{i\nu}R$ найдется хоть один собственный вектор оператора $\tilde{A}$. Обозначим его x_i и пусть y_i — вектор подпространства $\tilde{B}_{i\nu}R$, перешедший в x_i под действием оператора $(\tilde{A}-\lambda_i\tilde{E})$. Тогда $(\tilde{A}-\lambda_i\tilde{E})\,y_i=x_i$.

Достаточность. Пусть $(\tilde{A}-\lambda_i\tilde{E})\,y_i=x_i$, тогда

$$\tilde{B}_{i\nu}(\tilde{A}-\lambda_i\tilde{E})\,y_i=\tilde{B}_{i\nu}x_i.$$

На основании результатов § 2, п. 4 имеем

$$\tilde{B}_{i\nu}x_i=(\lambda_i-\lambda_1)\ldots(\lambda_i-\lambda_{i-1})(\lambda_i-\lambda_{i+1})\ldots(\lambda_i-\lambda_\nu)\,x_i\neq 0.$$

Поэтому

$$(\tilde{A}-\lambda_i\tilde{E})\,\tilde{B}_{i\nu}R\neq 0.$$

Упражнения

1. Найти собственные векторы матриц

$$\text{а)}\ \begin{pmatrix} -1 & 3 & -1 \\ -3 & 5 & -1 \\ -3 & 3 & 1 \end{pmatrix}, \qquad \text{б)}\ \begin{pmatrix} 2 & -1 & 2 \\ 5 & -3 & 3 \\ -1 & 0 & -2 \end{pmatrix},$$

$$\text{в)}\ \begin{pmatrix} -\frac{1}{3} & \frac{4}{3} & -\frac{2}{3} \\ \frac{4}{3} & -\frac{1}{3} & \frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & \frac{2}{3} \end{pmatrix}$$

2. Какие матрицы примера 1 имеют простую структуру?

3. Привести матрицы примера 1 (а и в) к простейшему виду.

4. Привести матрицу

$$\begin{pmatrix} 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$$

к простейшему виду и найти преобрузующую матрицу.

5. Привести к диагональному виду матрицу

$$\begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & 2 & 0 & \dots & 0 \\ 1 & 2 & 3 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & 2 & 3 & \dots & n \end{pmatrix}.$$

6. Выяснить, являются ли следующие матрицы подобными диагональным матрицам в полях рациональных, вещественных и комплексных чисел:

$$\text{а)}\ \begin{pmatrix} 8 & 15 & -36 \\ 8 & 21 & -46 \\ 5 & 12 & -27 \end{pmatrix}; \qquad \text{б)}\ \begin{pmatrix} 4 & 7 & -5 \\ -4 & 5 & 0 \\ 1 & 9 & -4 \end{pmatrix};$$

$$\text{в)}\ \begin{pmatrix} 5 & 2 & -3 \\ 4 & 5 & -4 \\ 6 & 4 & -4 \end{pmatrix}; \qquad \text{г)}\ \begin{pmatrix} 4 & 2 & -5 \\ 6 & 4 & -9 \\ 5 & 3 & -7 \end{pmatrix}.$$

Глава 6

ПРОСТРАНСТВА СО СКАЛЯРНЫМ ПРОИЗВЕДЕНИЕМ

§ 1. ЕВКЛИДОВО ПРОСТРАНСТВО

До сих пор мы рассматривали только такие свойства линейных операторов, которые не зависят от метрических свойств пространства R и полностью определяются лишь аксиомами линейного пространства. Это обстоятельство определяет чрезвычайно большую общность всех полученных результатов, которые, естественно, справедливы для любых линейных пространств. Однако для получения новых результатов необходимо несколько ограничить класс рассматриваемых пространств и ценой некоторой потери общности получить возможность более глубокого изучения свойств линейных операторов. Иллюстрацией к сказанному может служить тот факт, что уже в простейших примерах линейных операторов (см. пример 6 в § I главы 4) и при трактовке геометрического смысла определителя матрицы линейного оператора в пространствах V_2 и V_3 мы существенно использовали метрические свойства этих пространств (считая эти свойства очевидными и известными из геометрии).

В общем случае для того, чтобы наделить линейное пространство метрическими свойствами, надо к аксиомам линейного пространства добавить дополнительные аксиомы, определяющие соответствующие метрические свойства пространства. Мы не будем рассматривать общих метрических пространств, а ограничимся лишь важнейшим частным случаем последних. Этим частным случаем являются пространства со скалярным произведением.

1. Скалярное произведение. *Комплексное линейное пространство* R называется пространством со скалярным произведением, если задан закон, согласно которому любым двум векторам $\boldsymbol{x}$ и $\boldsymbol{y}$ из R единственным образом сопоставляется комплексное число, которое называется *скалярным произведением векторов* $\boldsymbol{x}$ и $\boldsymbol{y}$, обозначается $(\boldsymbol{x}, \boldsymbol{y})$ и удовлетворяет следующим условиям:

1) $(\boldsymbol{x}, \boldsymbol{y}) = \overline{(\boldsymbol{y}, \boldsymbol{x})}$,

где число $\overline{(y, x)}$, комплексно сопряженное с числом (y, x);

2) $(x+y, z)=(x, z)+(y, z)$;

3) $(\lambda x, y)=\lambda(x, y)$,

где λ — произвольное комплексное число.

Из введенных аксиом имеем

$$(x, y+z)=\overline{(y+z, x)}=\overline{(y, x)}+\overline{(z, x)}=$$
$$=(x, y)+(x, z),$$

$$(x, \lambda y)=\overline{(\lambda y, x)}=\bar{\lambda}\overline{(y, x)}=\bar{\lambda}(x, y),$$
$$(x, x)=\overline{(x, x)},$$

откуда следует, что скалярное произведение линейно относительно каждого из сомножителей. Наконец, из аксиомы 3 вытекает, что при $x=0$ $(x, y)=0$, так как если $x=0$, то его можно представить в виде $0\cdot z$, где $z\in R$, и тогда

$$(x, y)=(0\cdot z, y)=0(z, y)=0.$$

Таким образом, чтобы линейное пространство было пространством со скалярным произведением, надо к 8 аксиомам линейного пространства добавить 3 аксиомы скалярного произведения. Из всех возможных пространств со скалярным произведением мы в дальнейшем будем изучать наиболее важный частный случай, который получается при введении дополнительного условия

4) $$(x, x)>0$$

при любом $x\neq 0$. В этом случае пространство называется *унитарным*.

Вещественное унитарное пространство называют *евклидовым* и обозначают $\mathscr{E}$. В этом случае первая аксиома принимает вид

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

Введенные аксиомы позволяют обобщить на векторы произвольной природы из R характерные свойства скалярного произведения векторов из пространств V_2 и V_3.

2. Примеры скалярного произведения в различных пространствах

1. В пространствах V_2 и V_3 скалярное произведение вводится обычным образом:

$$(\boldsymbol{x}, \boldsymbol{y}) = |\boldsymbol{x}| \cdot |\boldsymbol{y}| \cdot \cos(\widehat{\boldsymbol{x}, \boldsymbol{y}}),$$

где $|\boldsymbol{x}|$ и $|\boldsymbol{y}|$ — длины векторов $\boldsymbol{x}$ и $\boldsymbol{y}$;

$\cos(\widehat{\boldsymbol{x}, \boldsymbol{y}})$ — косинус угла между векторами $\boldsymbol{x}$ и $\boldsymbol{y}$.

2. В пространстве T_n можно ввести скалярное произведение равенством

$$(\boldsymbol{x}, \boldsymbol{y}) = \alpha_1\bar{\beta}_1 + \alpha_2\bar{\beta}_2 + \ldots + \alpha_n\bar{\beta}_n,$$

где

$$\boldsymbol{x} = (\alpha_1, \alpha_2, \ldots, \alpha_n);$$
$$\boldsymbol{y} = (\beta_1, \beta_2, \ldots, \beta_n).$$

3. В пространстве $C(a, b)$ можно ввести скалярное произведение равенством

$$(\boldsymbol{x}, \boldsymbol{y}) = \int_a^b f(t)\,\varphi(t)\,dt,$$

где

$$\boldsymbol{x} = f(t);$$
$$\boldsymbol{y} = \varphi(t).$$

Легко видеть, что рассмотренные скалярные произведения удовлетворяют всем аксиомам унитарного пространства.

Следует подчеркнуть, что в одном и том же пространстве можно ввести скалярное произведение различным образом (лишь бы удовлетворялись аксиомы скалярного произведения). Например, в пространстве $C(a, b)$ можно ввести скалярное произведение с помощью соотношения

$$(\boldsymbol{x}, \boldsymbol{y}) = \int_a^b \psi^2(t)\,f(t)\,\varphi(t)\,dt,$$

где $\psi(t)$ — произвольная непрерывная на (a, b) ненулевая функция.

3. Скалярное произведение в координатной форме. Поставим вопрос о наиболее общем виде скалярного произведения в произвольном конечномерном унитарном пространстве.

Пусть $\boldsymbol{x}$ и $\boldsymbol{y}$ — векторы n-мерного линейного пространства. Выберем в этом пространстве базис $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$. Тогда

$$\boldsymbol{x} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + \ldots + x_n\boldsymbol{e}_n,$$
$$\boldsymbol{y} = y_1\boldsymbol{e}_1 + y_2\boldsymbol{e}_2 + \ldots + y_n\boldsymbol{e}_n.$$

На основании аксиом 2) и 3) скалярного произведения можно записать

$$(\boldsymbol{x}, \boldsymbol{y}) = (x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + \ldots + x_n\boldsymbol{e}_n, y_1\boldsymbol{e}_1 + y_2\boldsymbol{e}_2 + \ldots + y_n\boldsymbol{e}_n = \\ = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i\bar{y}_j(\boldsymbol{e}_i, \boldsymbol{e}_j).$$

Обозначим $(\boldsymbol{e}_i, \boldsymbol{e}_j) = \alpha_{ij}$, где в силу аксиомы (4)

$$\alpha_{ii} \neq 0.$$

Тогда

$$(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_{ij}x_i\bar{y}_j.$$

Полученное выражение представляет собой общий вид скалярного произведения в конечномерном унитарном пространстве.

Задавая различным образом скалярные произведения $(\boldsymbol{e}_i, \boldsymbol{e}_j)$ базовых векторов, получаем различные формы скалярного произведения. Очевидно, в евклидовом пространстве

$$(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_{ij}x_iy_j.$$

§ 2. МЕТРИЧЕСКИЕ СВОЙСТВА ВЕКТОРОВ

Введем метрические свойства, подобные тем, которые присущи векторам в пространствах V_2 и V_3.

1. Длина вектора. Рассмотрим *длину*, или *норму*, *вектора*, которая обозначается $|\boldsymbol{x}|$ и определяется выражением

$$|\boldsymbol{x}| = \sqrt{(\boldsymbol{x}, \boldsymbol{x})}.$$

В этом определении учитывается, что $(\boldsymbol{x}, \boldsymbol{x}) \geqslant 0$ и берется арифметическое значение корня. В пространствах V_2 и V_3 норма $|\boldsymbol{x}|$ совпадает с обычной длиной вектора $\boldsymbol{x}$. В про-

странствах T_n и $C(a, b)$ соответственно имеем

$$|x| = \sqrt{|\alpha_1|^2 + |\alpha_2|^2 + \ldots + |\alpha_n|^2},$$

$$|x| = \left\{\int_a^b f^2(t)\,dt\right\}^{1/2}.$$

Из определения нормы вектора вытекает:

1) $|x| > 0$ при $x \neq 0$ и $|x| = 0$ только, если $x = 0$;

2) $|\lambda x| = |\lambda|\,|x|$.

Вектор x, норма (длина) которого равна 1 (т. е. $|x| = 1$), называется *нормированным вектором*. Очевидно, всякий ненулевой вектор x можно пронормировать. Для этого надо умножить его на число $\lambda = \frac{1}{|x|}$. Полученный вектор $y = \frac{1}{|x|}x$ будет нормированным.

2. Угол между векторами. Рассмотрим евклидово пространство. Косинус угла между векторами x и y определим соотношением

$$\cos\varphi = \frac{(x, y)}{|x| \cdot |y|}.$$

Для того чтобы введенное понятие угла имело смысл, надо, чтобы для любых векторов x и y из $\mathscr{E}$ выполнялось условие

$$\frac{|(x, y)|}{|x| \cdot |y|} \leqslant 1 \quad \text{или} \quad |(x, y)| \leqslant |x| \cdot |y|.$$

Покажем, что это условие всегда выполняется. Действительно, рассмотрим скалярное произведение вектора $\lambda x - y$ самого на себя.

Тогда в силу аксиомы (4)

$$(\lambda x - y, \lambda x - y) \geqslant 0.$$

Отсюда

$$\lambda^2 (x, x) - 2\lambda (x, y) + (y, y) \geqslant 0.$$

Этот трехчлен относительно λ не отрицателен. Поэтому его дискриминант меньше или равен нулю, т. е.

$$(x, y)^2 \leqslant (x, x) \cdot (y, y) = |x|^2 \cdot |y|^2, \quad |(x, y)| \leqslant |x| \cdot |y|.$$

Это неравенство называется *неравенством Коши—Буняковского*.

Введенное определение косинуса угла между векторами обобщает известное из геометрии определение для пространств V_2 и V_3.

Для вещественного пространства T_n и пространства $C(a, b)$ соответственно имеем

$$\cos\varphi = \frac{\alpha_1\beta_1 + \alpha_2\beta_2 + \ldots + \alpha_n\beta_n}{\sqrt{\alpha_1^2 + \ldots + \alpha_n^2} \cdot \sqrt{\beta_1^2 + \ldots + \beta_n^2}},$$

$$\cos\varphi = \frac{\int\limits_a^b f(t) \cdot \varphi(t)\, dt}{\left\{\int\limits_a^b f^2(t)\, dt \cdot \int\limits_a^b \varphi^2(t)\, dt\right\}^{1/2}}.$$

3. Ортогональные векторы. Если для двух векторов $\boldsymbol{x}$ и $\boldsymbol{y}$ из унитарного пространства

$$(\boldsymbol{x}, \boldsymbol{y}) = 0,$$

то векторы $\boldsymbol{x}$ и $\boldsymbol{y}$ называются ортогональными.

Отсюда, в частности, следует, что нулевой вектор $\boldsymbol{x} = 0$ ортогонален к всякому вектору $\boldsymbol{y}$.

В пространстве V_2 и V_3 условие ортогональности $\boldsymbol{x}$ и $\boldsymbol{y}$ означает, что либо один из векторов $\boldsymbol{x}$ или $\boldsymbol{y}$ нулевой, либо, если $\boldsymbol{x} \neq 0$ и $\boldsymbol{y} \neq 0$, то угол между ними равен $\frac{\pi}{2}$.

В пространствах T_n и $C(a, b)$ условие ортогональности соответственно имеет вид

$$\alpha_1\bar{\beta}_1 + \ldots + \alpha_n\bar{\beta}_n = 0,$$

$$\int\limits_a^b f(t)\, \bar{\varphi}(t)\, dt = 0.$$

Например, векторы $\boldsymbol{x} = \cos nt$, $\boldsymbol{y} = \sin mt$, где m и n — целые, ортогональны при $a = -\pi$ и $b = \pi$. Это очевидно, так как

$$\int\limits_{-\pi}^{\pi} \cos nt \sin mt\, dt = 0.$$

Если L — подпространство пространства R и вектор $\boldsymbol{x} \in R$ ортогонален любому вектору $\boldsymbol{y} \in L$, то говорят, что вектор $\boldsymbol{x}$ ортогонален к подпространству L. Это определение сохраняется и в случае, когда L — произвольное множество векторов из R (необязательно подпространство пространства R).

4. Расстояние между векторами (точками). Определим расстояние между векторами $\boldsymbol{x}$ и $\boldsymbol{y}$ (обозначим его $\rho(\boldsymbol{x}, \boldsymbol{y})$) как длину вектора $\boldsymbol{x}-\boldsymbol{y}$. Для простоты ограничимся рассмотрением евклидова пространства

$$\rho(\boldsymbol{x}, \boldsymbol{y})=|\boldsymbol{x}-\boldsymbol{y}|=\sqrt{(\boldsymbol{x}-\boldsymbol{y}, \boldsymbol{x}-\boldsymbol{y})}=$$
$$=\sqrt{(\boldsymbol{x}, \boldsymbol{x})-2(\boldsymbol{x}, \boldsymbol{y})+(\boldsymbol{y}, \boldsymbol{y})}.$$

В пространствах V_2 и V_3 $\rho(\boldsymbol{x}, \boldsymbol{y})$ равно обычному расстоянию между концами векторов $\boldsymbol{x}$ и $\boldsymbol{y}$.

В пространствах T_n (вещественном) и $C(a, b)$ соответственно имеем

$$\rho(\boldsymbol{x}, \boldsymbol{y})=\sqrt{(\alpha_1-\beta_1)^2+\ldots+(\alpha_n-\beta_n)^2},$$
$$\rho(\boldsymbol{x}, \boldsymbol{y})=\left\{\int_a^b [f(t)-\varphi(t)]^2\,dt\right\}^{1/2}.$$

Из определения расстояния следует, что

1) $\rho(\boldsymbol{x}, \boldsymbol{y})=\rho(\boldsymbol{y}, \boldsymbol{x})$;

2) $\rho(\boldsymbol{x}, \boldsymbol{y})>0$ при $\boldsymbol{x}\neq\boldsymbol{y}$;

3) $\rho(\boldsymbol{x}, \boldsymbol{y})=0$ только при $\boldsymbol{x}=\boldsymbol{y}$;

4) $\rho(\boldsymbol{x}, \boldsymbol{z})\leqslant\rho(\boldsymbol{x}, \boldsymbol{y})+\rho(\boldsymbol{y}, \boldsymbol{z})$ для любых $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$ из $\mathscr{E}$.

Первые три из этих соотношений очевидны. Докажем последнее. Используя неравенство Коши—Буняковского, имеем

$$|\boldsymbol{x}+\boldsymbol{y}|^2=(\boldsymbol{x}+\boldsymbol{y}, \boldsymbol{x}+\boldsymbol{y})=|\boldsymbol{x}|^2+2(\boldsymbol{x}, \boldsymbol{y})+|\boldsymbol{y}|^2\leqslant$$
$$\leqslant|\boldsymbol{x}|^2+2|\boldsymbol{x}||\boldsymbol{y}|+|\boldsymbol{y}|^2=(|\boldsymbol{x}|+|\boldsymbol{y}|)^2.$$

Отсюда

$$|\boldsymbol{x}+\boldsymbol{y}|\leqslant|\boldsymbol{x}|+|\boldsymbol{y}|.$$

Это неравенство называется *неравенством треугольника*, так как в пространстве V_2 оно означает, что длина стороны треугольника не превосходит суммы длин двух других его сторон.

На основании неравенства треугольника получаем

$$|\boldsymbol{x}-\boldsymbol{z}|=|\boldsymbol{x}-\boldsymbol{y}+\boldsymbol{y}-\boldsymbol{z}|\leqslant|\boldsymbol{x}-\boldsymbol{y}|+|\boldsymbol{y}-\boldsymbol{z}|,$$

т. е.

$$\rho(\boldsymbol{x}, \boldsymbol{z})\leqslant\rho(\boldsymbol{x}, \boldsymbol{y})+\rho(\boldsymbol{y}, \boldsymbol{z}).$$

Заметим, что если векторы $\boldsymbol{x}$ и $\boldsymbol{y}$ ортогональны, то

$$|\boldsymbol{x}+\boldsymbol{y}|^2=|\boldsymbol{x}|^2+2(\boldsymbol{x},\ \boldsymbol{y})+|\boldsymbol{y}|^2=|\boldsymbol{x}|^2+|\boldsymbol{y}|^2.$$

Таким образом, для ортогональных векторов справедлива теорема Пифагора.

Теорема сохраняет силу для любого конечного числа попарно ортогональных векторов:

$$|\boldsymbol{x}_1+\boldsymbol{x}_2+\ldots+\boldsymbol{x}_n|^2=|\boldsymbol{x}_1|^2+|\boldsymbol{x}_2|^2+\ldots+|\boldsymbol{x}_n|^2.$$

§ 3. ОРТОНОРМИРОВАННЫЙ БАЗИС

Покажем, что в любом n-мерном унитарном пространстве существует базис, состоящий из попарно ортогональных векторов.

1. Линейная независимость ортогональной системы векторов. Назовем совокупность попарно ортогональных векторов *ортогональной системой векторов.*

Теорема. *Ортогональная система ненулевых векторов линейно независима.*

Пусть $\boldsymbol{x}_1,\ \boldsymbol{x}_2,\ \ldots,\ \boldsymbol{x}_n$ — ненулевые векторы, удовлетворяющие условиям

$$(\boldsymbol{x}_i,\ \boldsymbol{x}_j)=0,\ \text{если}\ i\neq j.$$

Предположим, что $\boldsymbol{x}_1,\ \boldsymbol{x}_2,\ \ldots,\ \boldsymbol{x}_n$ линейно зависимы. Следовательно, существуют такие $\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_n$, не все равные нулю, что

$$\alpha_1\boldsymbol{x}_1+\alpha_2\boldsymbol{x}_2+\ldots+\alpha_n\boldsymbol{x}_n=0.$$

Пусть для определенности $\alpha_1\neq 0$. Тогда, умножая это равенство скалярно на $\boldsymbol{x}_1$, получим

$$\alpha_1(\boldsymbol{x}_1,\ \boldsymbol{x}_1)+\alpha_2(\boldsymbol{x}_2,\ \boldsymbol{x}_1)+\ldots+\alpha_n(\boldsymbol{x}_n,\ \boldsymbol{x}_1)=0$$

и в силу условия

$$(\boldsymbol{x}_i,\ \boldsymbol{x}_j)=0\qquad \text{при}\ i\neq j$$

имеем $\alpha_1(\boldsymbol{x}_1,\ \boldsymbol{x}_1)=0$.

Но $(\boldsymbol{x}_1,\ \boldsymbol{x}_1)\neq 0$, так как $\boldsymbol{x}_1\neq 0$, значит $\alpha_1=0$, что противоречит условию и, следовательно, доказывает теорему.

2. Построение ортонормированного базиса. Рассмотрим n-мерное унитарное пространство. Так как это пространство линейное, то в нем существует некоторый базис, т. е. в нем всегда можно (и не единственным образом)

выделить n линейно независимых векторов $e_1, e_2, \ldots, e_n$. Тогда обозначим $e_1 = e_1^*$ и построим вектор $e_2^* = e_2 + \alpha_{11} e_1^*$ ($e_2^* \neq 0$ в силу линейной независимости e_1 и e_2). Коэффициент α_{11} выберем так, чтобы векторы e_1^* и e_2^* были ортогональны, т. е.

$$(e_2^*, e_1^*) = 0.$$

Но $(e_2^*, e_1^*) = (e_2, e_1^*) + \alpha_{11}(e_2^*, e_1^*)$ и, так как $e_1^* = e_1$, получаем

$$\alpha_{11} = -\frac{(e_2, e_1)}{(e_1, e_2)}.$$

Далее строим вектор

$$e_3^* = e_3 + \alpha_{21} e_1^* + \alpha_{22} e_2^*$$

($e_3^* \neq 0$, так как e_1^* и e_3 линейно независимы) и выбираем числа α_{21} и α_{22} так, чтобы вектор e_3^* был ортогонален векторам e_1^* и e_2^*. Но это значит, что должны выполняться условия

$$(e_3, e_1^*) + \alpha_{21}(e_1^*, e_1^*) + \alpha_{22} \cdot 0 = 0,$$
$$(e_3, e_2^*) + \alpha_{21} \cdot 0 + \alpha_{22}(e_2^*, e_2^*) = 0,$$

откуда

$$\alpha_{21} = -\frac{(e_3, e_1^*)}{(e_1^*, e_1^*)},$$
$$\alpha_{22} = -\frac{(e_3, e_2^*)}{(e_2^*, e_2^*)}.$$

Если же таким образом построены $n-1$ попарно ортогональных векторов $e_1^*, e_2^*, \ldots, e_{n-1}^*$, то вектор e_n^* выберем в виде

$$e_n^* = e_n + \alpha_{n-1,1} e_1^* + \alpha_{n-1,2} e_2^* + \ldots + \alpha_{n-1,n-1} e_{n-1}^*$$

($e_n^* \neq 0$, так как $e_1^*, e_n^*, \ldots, e_{n-1}^*$ и e_n линейно независимы) и потребуем, чтобы он был ортогонален векторам e_1^*, $e_2^*, \ldots, e_{n-1}^*$. Но для этого должны выполняться условия

$$(e_n^*, e_1^*) = (e_n, e_1^*) + \alpha_{n-1,1}(e_1^*, e_1^*) \ldots = 0,$$
$$(e_n^*, e_2^*) = (e_n, e_2^*) + \alpha_{n-1,2}(e_2^*, e_2^*) = 0,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$(e_n^*, e_{n-1}^*) = (e_n, e_{n-1}^*) + \ldots + \alpha_{n-1,n-1}(e_{n-1}^* e_{n-1}^*) = 0$$

и значит,

$$\alpha_{n-1,1}=\frac{(e_n, e_1^*)}{(e_1^*, e_1^*)}, \quad \alpha_{n-1,2}=-\frac{(e_n, e_2^*)}{(e_2^*, e_2^*)}, \ldots,$$

$$\alpha_{n-1,n-1}=-\frac{e_n, e_{n-1}^*}{e_{n-1}^*, e_{n-1}^*}.$$

Таким образом, по произвольному базису $e_1, e_2, \ldots, e_n$ всегда можно построить n попарно ортогональных векторов $e_1^*, e_2^*, \ldots, e_n^*$. Эти векторы в силу доказанной в п. 1 теореме линейно независимы и, следовательно, могут быть приняты в качестве базиса нашего пространства. Такой базис назовем *ортогональным*. Отметим, что описанный выше процесс построения ортогонального базиса называется *ортогонолизацией системы векторов*.

Если каждый вектор ортогонального базиса $e_1^*, \ldots, e_n^*$ пронормировать, то получим n линейно независимых нормированных ортогональных векторов:

$$\frac{e_1^*}{|e_1^*|}, \frac{e_2^*}{|e_2^*|}, \ldots, \frac{e_n^*}{|e_n^*|},$$

которые также можно принять в качестве базиса нашего пространства.

Такой базис называется *ортонормированным*. Таким образом, приходим к следующему утверждению.

Теорема. *В каждом n-мерном унитарном пространстве существует ортонормированный базис.*

3. Координатное выражение скалярного произведения в ортонормированном базисе. Пусть $e_1, e_2, \ldots, e_n$ — произвольный ортонормированный базис в n-мерном унитарном пространстве, а x — произвольный вектор этого пространства. Это значит, что

$$x=x_1e_1+x_2e_2+\ldots+x_ne_n,$$

причем

$$|e_i|=1; \quad i=1, 2, \ldots, n$$

$$|e_i, e_j|=0 \quad \text{при} \quad i\neq j.$$

Умножим выражение для x скалярно на e_i. Тогда

$$(x, e_i)=x_1(e_1, e_i)+\ldots+x_i(e_i, e_i)\ldots+x_n(e_n, e_i)$$

и, следовательно,

$$x_i=(x, e_i).$$

Таким образом, *координаты любого вектора* $\boldsymbol{x}$ *относительно ортонормированного базиса равны скалярным произведениям вектора* $\boldsymbol{x}$ *на соответствующие базисные векторы.* Далее, если

$$\boldsymbol{x} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + \ldots + x_n\boldsymbol{e}_n,$$
$$\boldsymbol{y} = y_1\boldsymbol{e}_1 + y_2\boldsymbol{e}_2 + \ldots + y_n\boldsymbol{e}_n,$$

то

$$(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i\bar{y}_j \cdot (\boldsymbol{e}_i\boldsymbol{e}_j). \tag{6.1}$$

Отсюда

$$(\boldsymbol{x}, \boldsymbol{y}) = x_1\bar{y}_1 + x_2\bar{y}_2 + \ldots + x_n\bar{y}_n. \tag{6.2}$$

Отметим, что из равенства (6.1) очевидно обратное: если для любых векторов $\boldsymbol{x}$ и $\boldsymbol{y}$ $(\boldsymbol{x}, \boldsymbol{y}) = x_1\bar{y}_1 + x_2\bar{y}_2 + \ldots + x_n\bar{y}_n$, то соответствующий базис ортонормированный.

В пространстве $\mathscr{E}$ формула (6.2) принимает вид

$$(\boldsymbol{x}, \boldsymbol{y}) = x_1y_1 + x_2y_2 + \ldots + x_ny_n.$$

Заметим, что в пространствах V_2 и V_3 в качестве ортонормированного базиса можно выбрать любые два и соответственно три единичных взаимно ортогональных вектора. В пространстве T_n в качестве ортонормированного базиса можно выбрать векторы

$$\boldsymbol{e}_1 = (1, 0, \ldots 0),\ \boldsymbol{e}_2 = (0, 1, \ldots 0),\ \ldots,\ \boldsymbol{e}_n = (0, 0, \ldots 1).$$

4. Матричная форма записи скалярного произведения в евклидовом пространстве. Рассмотрим две столбцовые матрицы:

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix},$$

элементами которой являются координаты векторов $\boldsymbol{x}$ и $\boldsymbol{y}$ в ортонормированном базисе.

Так как $(\boldsymbol{x}, \boldsymbol{y}) = x_1y_1 + x_2y_2 + \ldots + x_ny_n$, а матрица

$$X'Y = Y'X = (x_1y_1 + x_2y_2 + \ldots + x_ny_n) =$$
$$= x_1y_1 + x_2y_2 + \ldots + x_ny_n, \text{ то}$$
$$(\boldsymbol{x}, \boldsymbol{y}) = X'Y = Y'X.$$

§ 4. САМОСОПРЯЖЕННЫЙ ОПЕРАТОР

1. **Определение.** Пусть $\tilde{A}$ и $\tilde{B}$ — линейные операторы, действующие в унитарном пространстве.

Если для любых двух векторов x и y

$(\tilde{A}x, y) = (x, \tilde{B}y)$, то операторы $\tilde{A}$ и $\tilde{B}$ называются *сопряженными*.

Если сопряженные операторы совпадают, т. е. $\tilde{A} = \tilde{B}$, то оператор $\tilde{A}$ называется *самосопряженным*.

В этом случае имеем $(\tilde{A}x, y) = (x, \tilde{A}y)$ и, в частности, $(\tilde{A}x, x) = (x, \tilde{A}x)$.

2. Свойства самосопряженного оператора.

Теорема 1. *Собственные числа самосопряженного оператора вещественны.*

Пусть λ — собственное число самосопряженного оператора $\tilde{A}$ и x — соответствующий собственный вектор, т. е.

$$\tilde{A}x = \lambda x, \quad \text{где} \quad x \neq 0.$$

Так как

$$(\tilde{A}x, x) = (x, \tilde{A}x), \quad \text{то} \quad (\lambda x, x) = (x, \lambda x),$$

или

$$\lambda (x, x) = \bar{\lambda} (x, x).$$

Но $(x, x) \neq 0$. Значит, $\lambda = \bar{\lambda}$, т. е. λ — вещественное число.

Теорема 2. *Собственные векторы самосопряженного оператора $\tilde{A}$, соответствующие различным его собственным числам, ортогональны.*

Пусть λ_1 и λ_2 — различные собственные числа (самосопряженного) оператора $\tilde{A}$, а x_1 и x_2 — соответствующие им собственные векторы.

Так как $\tilde{A}x_1 = \lambda_1 x$, $\tilde{A}x_2 = \lambda_2 x_2$ и $(\tilde{A}x_1, x_2) = (x_1, \tilde{A}x_2)$, то $\lambda_1 (x_1, x_2) = \lambda_2 (x_1, x_2)$ и так как $\lambda_1 \neq \lambda_2$, то $(x_1, x_2) = 0$. Значит векторы x_1 и x_2 — ортогональны.

Теорема 3. *Самосопряженный оператор имеет простую структуру.*

Предположим противное. Пусть самосопряженный оператор не имеет простой структуры. В этом случае (см. гл. 5, § 3, п. 6) существует вектор y такой, что $(\tilde{A} - \lambda \tilde{E}) y = x$, где x — собственный вектор, соответствующий собственному числу λ. Отсюда $\tilde{A}y = \lambda y + x$.

Умножим это равенство скалярно на вектор x

$$(x, \tilde{A}y) = \lambda(x, y) + (x, x).$$

Но $\tilde{A}$ — самосопряженный оператор, поэтому

$$(\tilde{A}x, y) = \lambda(x, y) + (x, x).$$

Так как $\tilde{A}x = \lambda x$, то

$$\lambda(x, y) = \lambda(x, y) + (x, x).$$

Отсюда $(x, x) = 0$, что невозможно, так как x — собственный вектор и поэтому $x \neq 0$, что и требовалось доказать.

§ 5. СИММЕТРИЧНЫЙ ОПЕРАТОР

Самосопряженный оператор $\tilde{A}$, действующий в евклидовом пространстве, называется *симметричным оператором*. Из определения следует, что для симметричного оператора справедливы все теоремы, доказанные в предыдущем параграфе.

1. Матрица симметричного оператора. *Чтобы оператор $\tilde{A}$ был симметричным, необходимо и достаточно, чтобы в ортонормированном базисе его матрица была симметрической.*

Пусть $\tilde{A}$ — симметричный оператор. Тогда при любых x и y справедливо равенство

$$(\tilde{A}x, y) = (x, \tilde{A}y).$$

Обозначим через A матрицу оператора $\tilde{A}$. Тогда

$$(\tilde{A}x, y) = (AX)'Y = X'A'Y,$$
$$(x, \tilde{A}y) = X'(AY) = X'AY.$$

Поэтому $X'A'Y = X'AY$, где X и Y — произвольные столбцовые матрицы. Последнее равенство справедливо тогда и только тогда, когда $A' = A$, т. е. A — симметрическая матрица.

2. Приведение вещественной симметрической матрицы к диагональному виду. Как мы знаем, симметричный оператор $\tilde{A}$ имеет простую структуру, поэтому симметрическая матрица $\tilde{A}$ преобразованием подобия может быть приведена к диагональному виду, т. е. найдется такая матрица T, что $T^{-1}AT = (\lambda_i \delta_{ij})$.

§ 6. ОРТОГОНАЛЬНЫЕ ПРЕОБРАЗОВАНИЯ

Выясним теперь, каковы специфические свойства матрицы перехода от одного ортонормированного базиса к другому ортонормированному базису в евклидовом пространстве.

1. Ортогональные матрицы. Матрица T с вещественными элементами называется *ортогональной*, если $T' = T^{-1}$, т. е.

$$TT' = T'T = E.$$

Из определения следует, что ортогональная матрица всегда невырожденная. Так как $\det(TT') = \det T' \det T' = \det E = 1$, а $\det T' = \det T$, то $\det T = \pm 1$.

Основные свойства ортогональной матрицы

1. Матрица, обратная ортогональной, также ортогональна.

Пусть T — ортогональная матрица, т. е. $T' = T^{-1}$. Отсюда $(T^{-1})' = (T')' = T = (T^{-1})^{-1}$, т. е. $(T^{-1})' = (T^{-1})^{-1}$.

Значит, матрица T^{-1} — ортогональная.

Кроме того, так как $T' = T^{-1}$, то матрица T' — ортогональная.

2. Матрица $T = (t_{ij})$ — ортогональная тогда и только тогда, когда ее элементы удовлетворяют равенствам

$$\sum_{k=1}^{n} t_{ki} t_{kj} = \delta_{ij}.$$

Действительно,

$$T'T = \left(\sum_{k=1}^{n} t'_{ik} t_{kj}\right) = \left(\sum_{k=1}^{n} t_{ki} t_{kj}\right).$$

Отсюда $T'T$ может быть равно E тогда и только тогда, когда

$$\sum_{k=1}^{n} t_{ki} t_{kj} = \delta_{ij},$$

Или иначе

$$\sum_{k=1}^{n} t_{ki} t_{kj} = 0 \text{ при } i \neq j.$$

$$\sum_{k=1}^{n} t_{ki}^2 = 1.$$

Так как матрица T' также ортогональна, то

$$\sum_{k=1}^{n} t_{ik}t_{jk} = \delta_{ij}.$$

2. Ортогональное преобразование и его свойства. Линейное преобразование $Y = TX$ с ортогональной матрицей называется *ортогональным*. Так как $\det T = \pm 1$, то ортогональное преобразование всегда невырожденное. Докажем **теорему.**

Ортогональное пргобразование координат не изменяет скалярного произведения векторов.

Рассмотрим линейный оператор $\tilde{T}$, соответствующий матрице T, и два произвольных вектора $\boldsymbol{x}$ и $\boldsymbol{y}$. Их образы обозначим $\boldsymbol{x}_1$ и $\boldsymbol{y}_1$, т. е. $\boldsymbol{x}_1 = \tilde{T}\boldsymbol{x}$; $\boldsymbol{y}_1 = \tilde{T}\boldsymbol{y}$.
Тогда

$$(\boldsymbol{x}, \boldsymbol{y}) = X'Y,$$

$$(\boldsymbol{x}_1, \boldsymbol{y}_1) = (TX)'TY = X'T'TY = X'EY = X'Y.$$

Поэтому

$$(\boldsymbol{x}_1, \boldsymbol{y}_1) = (\tilde{T}\boldsymbol{x}, \tilde{T}\boldsymbol{y}) = (\boldsymbol{x}, \boldsymbol{y}).$$

Следствие 1. *Ортогональное преобразование не меняет длин векторов и углов между векторами.*

Следствие 2. *Ортогональное преобразование переводит ортонормированный базис в ортонормированный.*

Следствие 3. *Матрица T перехода от одного ортонормированного базиса $\{e\}$ к другому ортонормированному базису $\{e^*\}$ является ортогональной.*

Действительно, элементы столбцов матрицы T являются координатами векторов e_i^* в базисе $\{e\}$; так как векторы e_i^* ортонормированы, то элементы матрицы T удовлетворяют условиям

$$\sum_{k=1}^{n} t_{ki}t_{kj} = \delta_{ij}.$$

Следовательно, матрица T — ортогональная.

Следствие 4. *Покажем, что матрицу T, приводящую симметрическую матрицу A к диагональному виду, можно выбрать ортогональной.* Для этого нам достаточно показать, что из собственных векторов матрицы A можно построить ортонормированный базис.

Пусть собственное число λ матрицы A имеет кратность k. Так как оператор $\tilde{A}$ имеет простую структуру, то числу λ соответствует k линейно независимых собственных векторов. Эти векторы образуют базис подпространства из собственных векторов размерности k. Выберем в этом подпространстве ортонормированный базис.

На основании теоремы 2 из § 4 собственные векторы, соответствующие другим собственным значениям, ортогональны этому подпространству.

Поступая аналогично с векторами, соответствующими другим собственным числам (учитывая их кратность), мы получим ортонормированный базис всего пространства.

Таким образом, найдется такая ортогональная матрица T, что

$$T'AT = (\lambda_i \delta_{ij}).$$

Упражнения

1. Проверить, что векторы $x_1 = (1, 1, 1, 2)$ и $x_2 = (1, 2, 3, -3)$ ортогональны. Дополнить их до ортогонального базиса четырехмерного пространства.

2. Применяя ортогонализацию, построить ортогональный базис пространства, представляющего собой линейные комбинации следующих векторов

$$(2, 1, 3, -1), \quad (7, 4, 3, -3), \quad (1, 1, -6, 0), \quad (5, 7, 7, 8).$$

3. Матрица оператора в некотором ортонормированном базисе

$$A = \begin{pmatrix} 11 & 2 & -8 \\ 2 & 2 & 10 \\ -8 & 10 & 5 \end{pmatrix}.$$

а) Найти ортонормированный базис из собственных векторов матрицы A.

б) Найти матрицу оператора $\tilde{A}$ в собственном ортонормированном базисе.

4. Доказать, что полиномы Лежандра

$$P_0 = 1, \quad P_k = \frac{1}{2^k k!} \frac{d^k}{dx^k} [(x^2 - 1)^k], \; k = 1, 2, 3$$

образуют ортогональный базис пространства многочленов степени не выше 3-й на отрезке $-1 \leqslant x \leqslant 1$, если скалярное произведение определить равенством $(f, g) = \int_{-1}^{1} f(x) g(x) dx$, где f и g — многочлены.

5. При каких условиях диагональная матрица будет ортогональной?

6. Показать, что матрица $\begin{pmatrix} \cos 2 & -\sin 2 \\ \sin 2 & \cos 2 \end{pmatrix}$ ортогональна.

7. Квадратная матрица $A=(a_{ij})$ порядка n с комплексными элементами называется *унитарной*, если $A^{-1}=\bar{A}'$. Доказать, что

а) $\det A=\pm 1$,

б) столбцы матрицы A попарно ортогональны.

8. При каких условиях диагональная матрица является унитарной?

9. Доказать, что собственные значения ортогональной матрицы равны по модулю единице.

10. Оператор $\tilde{A}$ называется *кососимметричным*, если для любых векторов x и y справедливо равенство

$$(x, \tilde{A}y)=-(\tilde{A}x, y).$$

Доказать, что матрица этого оператора в ортонормированном базисе кососимметрическая (т. е. $A=-A'$) и собственные числа нулевые или чисто мнимые.

Глава 7

КВАДРАТИЧНЫЕ ФОРМЫ

§ 1. КВАДРАТИЧНЫЕ ФОРМЫ И ПРИВЕДЕНИЕ ИХ К КАНОНИЧЕСКОМУ ВИДУ

В предыдущих главах нам часто приходилось рассматривать однородный многочлен первой степени от n переменных

$$a_1x_1+a_2x_2+a_3x_3+\ldots+a_nx_n,$$

который называется линейной формой этих переменных. Непосредственным обобщением линейной формы n переменных являются так называемые *квадратичные формы*, которые тесно связаны с изучением поверхностей второго порядка в n мерном евклидовом пространстве.

Общее уравнение таких поверхностей в декартовых координатах содержит сумму квадратичной формы, линейной формы и некоторой постоянной, причем преобразованиям (определенного типа) переменных и связанным с ними преобразованиям квадратичной (и линейной) формы соответствуют преобразования поворота и параллельного переноса системы координат, что используется для исследования поверхностей второго порядка и приведения уравнений этих поверхностей к каноническому (простейшему) виду.

1. Квадратичная форма и ее матричная запись. *Квадратичной формой переменных* $x_1, x_2, \ldots, x_n$ называется одно-

родный многочлен второй степени относительно этих переменных, т. е.

$$L(x_1, x_2, \ldots x_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j, \tag{7.1}$$

где $a_{ij} = a_{ji}$.

Например, квадратичная форма при $n = 2$ имеет вид

$$L(x, y) = Ax^2 + 2Bxy + Cy^2.$$

Здесь

$$a_{11} = A, \quad a_{12} = a_{21} = B, \quad a_{22} = C.$$

В принятых обозначениях имеем:

$$L(x_1, x_2) = \sum_{i=1}^{2} \sum_{j=1}^{2} a_{ij} x_i x_j = a_{11} x_1^2 + a_{12} x_1 x_2 + a_{21} x_2 x_1 + a_{12} x_2^2.$$

Это можно записать еще так:

$$L(x_1, x_2) = x_1 (a_{11} x_1 + a_{12} x_2) + x_2 (a_{21} x_1 + a_{22} x_2).$$

Если ввести матрицы

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

то

$$L(x_1, x_2) = X'AX.$$

Последняя формула справедлива и в общем случае. Действительно, введем матрицу из коэффициентов формы

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}.$$

Очевидно, что если дана форма (7.1), то матрица A (симметрическая) известна. Обратно, если дана матрица A, то ей соответствует вполне определенная (с точностью до обозначения переменных) квадратичная форма. В силу этого A называют *матрицей квадратичной формы*.

Форму (7.1) можно записать в матричном виде

$$L(x_1, x_2, \ldots x_n) = X'AX, \qquad (7.2)$$

где $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$.

В дальнейшем мы будем изучать вещественные квадратичные формы (a_{ij} — вещественные числа).

2. Преобразование квадратичной формы. Произведем линейное преобразование переменных формы (7.1). Положим $X = TY$, где T — квадратная матрица порядка n, а матрица

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

Тогда, так как $X' = Y'T'$, то из равенства (7.2) получаем $L = X'AX = Y'T'ATY$, или $L = Y'(T'AT)Y$. Обозначим $B = T'AT$. Тогда

$$L = Y'BY.$$

Отметим, что матрица B — симметрическая. Действительно,

$$B' = T'A'T'' = T'A'T = T'AT = B, \quad \text{так как } A' = A.$$

Итак, квадратичная форма (7.1) с матрицей A после линейного преобразования переменных переходит в квадратичную форму от новых переменных с матрицей $B = T'AT$.

Если T — ортогональная матрица, то $T' = T^{-1}$ и $B = T^{-1}AT$, т. е. матрица B подобна матрице A.

3. Приведение квадратичной формы к каноническому виду. Рассмотрим вещественную квадратичную форму $L = X'AX$. После линейного преобразования получаем квадратичную форму от новых переменных с матрицей $B = T'AT$. Так как матрица A — вещественная и симметрическая, то она имеет простую структуру и может быть ортогональным преобразованием приведена к диаго-

нальному виду. Значит, существует такая ортогональная матрица T, что $B = T^{-1}AT = (\lambda_i \delta_{ij})$, где λ_i — собственные значения (вещественные!) матрицы A.

Преобразованная форма имеет вид:

$$L = \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \delta_{ij} y_i y_j,$$

или

$$L = \sum_{i=1}^{n} \lambda_i y_i^2, \tag{7.3}$$

где число отличных от нуля коэффициентов λ_i равно рангу матрицы A.

Пример.

Привести к каноническому виду форму

$$L = x_1^2 - 2x_2^2 + x_3^2 + 4x_1x_2 - 8x_1x_3 - 4x_2x_3.$$

Решение. Матрица A квадратичной формы имеет вид

$$A = \begin{pmatrix} 1 & 2 & -4 \\ 2 & -2 & -2 \\ -4 & -2 & 1 \end{pmatrix}.$$

Ее характеристическое уравнение

$$\Delta(\lambda) = |A - \lambda E| = \begin{vmatrix} 1-\lambda & 2 & -4 \\ 2 & -2-\lambda & -2 \\ -4 & -2 & 1-\lambda \end{vmatrix} = 0.$$

Отсюда

$$(\lambda - 6)(\lambda + 3)^2 = 0.$$

Следовательно,

$$\lambda_1 = 6, \quad \lambda_2 = -3, \quad \lambda_3 = -3$$

и канонический вид квадратичной формы будет

$$L = 6y_1^2 - 3y_2^2 - 3y_3^2.$$

§ 2. НЕОДНОРОДНЫЙ МНОГОЧЛЕН ВТОРОЙ СТЕПЕНИ

Как известно, в аналитической геометрии ставится и решается задача приведения общего уравнения линии второго порядка на плоскости или поверхности второго порядка к каноническому виду с помощью преобразования декартовой системы координат (без изменения масштаба). Мы рассмотрим задачу приведения общего уравнения второй степени (содержащего n переменных) к кано-

ническому виду. Предварительно решим задачу о преобразовании многочлена второй степени.

1. Классификация многочленов. *Неоднородным многочленом второй степени* называется выражение вида

$$P(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j + 2 \sum_{i=1}^{n} b_i x_i + c, \qquad (7.4)$$

где все коэффициенты — вещественные числа и $a_{ij} = a_{ji}$. Запишем многочлен в матричной форме

$$P = X'AX + 2X'B + c, \qquad (7.5)$$

где

$$A = (a_{ij}); \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}; \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

Рассмотрим матрицу

$$D = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} & b_n \end{pmatrix}.$$

Если ранги матриц A и D равны, то многочлен P будем называть *центральным*. В противном случае — *нецентральным*. Очевидно, если A — неособенная матрица ($\det A \neq 0$), то многочлен P — центральный.

2. Приведение неоднородного многочлена 2-й степени к каноническому виду. Имеем многочлен (7.4). Рассмотрим два возможных случая:

а) Пусть многочлен — центральный. Подвергнем его преобразованию

$$X = Y + V,$$

где

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}; \quad V = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix};$$

v_i — некоторые числа, которые мы в дальнейшем выберем специальным образом.

Внося X в выражение (7.5), получаем

$$P=(Y'+V')A(Y+V)+2(Y'+V')B+c$$

или после преобразований

$$P=Y'AY+2Y'(AV+B)+V'(AV+B)+V'B+c.$$

При выводе этого соотношения учитывалось, что $(V'AY)'=$ $=V'AY$, так как $V'AY$ — одноэлементная матрица.

Выберем теперь V так, чтобы выражение $AV+B$ обратилось в нуль. Это можно сделать, так как система уравнений $AV+B=0$ совместна (ранги матрицы системы и расширенной матрицы системы по условию равны).

Итак, если выберем матрицу V решением уравнения $AV+B=0$, то многочлен примет вид

$$P=Y'AY+V'B+c.$$

Квадратичную форму $Y'AY$ с матрицей A приведем к каноническому виду при помощи ортогонального преобразования переменных $Y=TZ$, где $T=(t_{ij})$ — ортогональная матрица.

Получаем канонический вид центрального многочлена

$$P=\sum_{i=1}^{n}\lambda_i z_i^2+c_1, \quad \text{где} \quad c_1=V'B+c,$$

причем, если $\det A=0$, то некоторые коэффициенты λ_i равны нулю.

Следует отметить, что если $\det A\neq 0$, то свободный член c_1 можно определить, не решая предварительно системы

$$AV+B=0$$

относительно V. Так как

$$c_1=V'B+c,$$

где V определяется из предыдущего условия, то c_1 можно найти из системы $n+1$ уравнений с $n+1$ неизвестными:

$$AV=-B,$$
$$V'B-c_1=-c,$$

или в развернутой форме

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 \\ b_1 & b_2 & \dots & b_n & -1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \\ c_1 \end{pmatrix} = \begin{pmatrix} -b_1 \\ -b_2 \\ \vdots \\ -b_n \\ -c \end{pmatrix}.$$

Отсюда

$$c_1 = \frac{\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} & -b_1 \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} & -b_n \\ b_1 & b_2 & \dots & b_n & -c \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} & 0 \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 \\ b_1 & b_2 & \dots & b_n & -1 \end{vmatrix}} = \frac{1}{\det A} \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} & b_n \\ b_1 & b_2 & \dots & b_n & c \end{vmatrix}.$$

Итак, если $\det A \neq 0$, то канонический вид многочлена таков:

$$P = \sum_{i=1}^{n} \lambda_i z_i^2 + \frac{K}{\det A},$$

где

$$K = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} & b_n \\ b_1 & b_2 & \dots & b_n & c \end{vmatrix} \quad \text{и} \quad \lambda_i \neq 0 \quad i = 1, 2, \dots n.$$

б) Пусть теперь многочлен — нецентральный. Подвергнем переменные x_i ортогональному преобразованию $X = TY$, где $T = (t_{ij})$, ортогональная матрица такая, что квадратичная форма XAX преобразуется к каноническому виду. Преобразованный многочлен будет иметь вид

$$P = \sum_{i=1}^{m} \lambda_i y_i^2 + 2 \sum_{i=1}^{n} b_i^* y_i + c,$$

где $m < n$ (по условию $\det A = 0$ и некоторые λ_i равны нулю), $\lambda_i \neq 0$, $i = 1, 2, \ldots m$, а

$$\begin{pmatrix} b_1^* \\ b_2^* \\ \vdots \\ b_n^* \end{pmatrix} = T'B,$$

причем хотя бы одно $b_i^* \neq 0$, так как в противном случае многочлен был бы центральным. Так как

$$\lambda_i y_i^2 + 2b_i^* y_i = \lambda_i \left(y_i + \frac{b_i^*}{\lambda_i}\right)^2 - \frac{b_i^{*2}}{\lambda_i}; \quad i = 1, 2, \ldots m,$$

то обозначая $y_i + \frac{b_i^*}{\lambda_i} = z_i$, получим

$$P = \sum_{i=1}^{m} \lambda_i z_i^2 + 2\sum_{i=m+1}^{n} b_i^* y_i + c_1,$$

где $c_1 = c - \sum_{i=1}^{m} \frac{b_i^{*2}}{\lambda_i}$.

Произведем теперь еще ортогональное преобразование, не меняющее переменных z_i, вида

$$u_i = z_i, \quad i = 1, 2, \ldots m;$$

$$u_j = \sum_{i=m+1}^{n} q_{ij} y_i, \quad j = m+1, m+2, \ldots n.$$

В частности, выберем это преобразование так:

$$u_i = z_i; \quad i = 1, 2, \ldots m$$

$$u_{m+1} = \frac{1}{\sqrt{b_{m+1}^{*2} + \ldots + b_n^{*2}}} \sum_{i=m+1}^{n} b_i^* y_i$$

$$u_j = y_j; \quad j = m+2, m+3, \ldots n,$$

где хотя бы одно из b_i^* при $i \geqslant m+1$ отлично от нуля. Обозначив $\lambda_{m+1} = \sqrt{b_{m+1}^{*2} + \ldots + b_n^{*2}}$, имеем

$$P = \sum_{i=1}^{m} \lambda_i u_i^2 + 2\lambda_{m+1} u_{m+1} + c_1.$$

Наконец, полагая $u_i = v_i$, где $i \neq m+1$,

$$u_{m+1} + \frac{c_1}{2\lambda_{m+1}} = v_{m+1},$$

находим канонический вид нецентрального многочлена

$$P = \sum_{i=1}^{m} \lambda_i v_i^2 + 2\lambda_{m+1} v_{m+1}.$$

Здесь

$$\lambda_i \neq 0, \quad i = 1, 2, \ldots m; \quad m < n.$$

§ 3. УПРОЩЕНИЕ УРАВНЕНИЙ ЛИНИЙ И ПОВЕРХНОСТЕЙ ВТОРОГО ПОРЯДКА

1. Общее уравнение второй степени. *Общим уравнением второй степени* называется уравнение вида $P(x_1, x_2, \ldots x_n) = 0$, где левая часть есть многочлен второй степени. Этот многочлен имеет вид (7.4)

$$P = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j + 2 \sum_{i=1}^{n} b_i x_i + c$$

(здесь все коэффициенты действительны $a_{ij} = a_{ji}$), или в матричной форме

$$P = X'AX + 2X'B + c.$$

Учитывая результаты § 2, мы с помощью линейных преобразований можем привести общее уравнение второй степени к одному из следующих видов:

1. $\sum_{i=1}^{n} \lambda_i z_i^2 + c_1 = 0$, если многочлен P — центральный, причем некоторые λ_i могут быть равны нулю.

2. $\sum_{i=1}^{m} \lambda_i z_i^2 + 2\lambda_{m+1} z_{m+1} = 0$, если многочлен — нецентральный, причем $\lambda_i \neq 0$, $i = 1, 2, \ldots, m, m+1$ и $m < n$. Рассмотрим два частных случая.

2. Линии второго порядка. Положим в общем уравнении $n = 2$. Получим

$$\sum_{i=1}^{2} \sum_{j=1}^{2} a_{ij} x_i x_j + 2 \sum_{i=1}^{2} b_i x_i + c = 0.$$

Это уравнение есть уравнение линии второго порядка на плоскости в декартовой системе координат. С помощью преобразований вида $X=Y+V$ (параллельный перенос осей координат) и $X=TZ$ (поворот осей координат) уравнение приводится к каноническому виду. Изменяя обозначения переменных и коэффициентов, получаем:

1. $\lambda_1 x^2+\lambda_2 y^2+c=0$, если многочлен — центральный. Причем хоть одно из λ_1 и λ_2 не равно нулю. Имеем центральную линию.

2. $\nu_1 y^2+\nu_2 x=0$, если многочлен нецентральный, причем ν_1 и ν_2 не равны нулю. Имеем нецентральную линию.

Придавая коэффициентам различные значения и вводя стандартные обозначения, получаем следующие геометрические образы:

1) $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ — эллипс;

2) $\frac{x^2}{a^2}-\frac{y^2}{b^2}=1$ — гипербола;

3) $\frac{x^2}{a^2}+\frac{y^2}{b^2}=-1$ — мнимый эллипс;

4) $\frac{x^2}{a^2}+\frac{y^2}{b^2}=0$ — точка;

5) $\frac{x^2}{a^2}-\frac{y^2}{b^2}=0$ — две пересекающиеся прямые;

6) $x^2=a^2$ — две параллельные линии;
7) $x^2=-a^2$ — две мнимые прямые;
8) $x^2=0$ — две слившиеся прямые;
9) $y^2=2px$ — парабола.

3. Поверхности второго порядка. Положив в общем уравнении второй степени $n=3$, получим уравнение поверхности второго порядка в декартовой системе координат

$$\sum_{i=1}^{3}\sum_{j=1}^{3} a_{ij}x_i x_j+2\sum_{i=1}^{3} b_i x_i+c=0.$$

Аналогично предыдущему пункту в результате соответствующих преобразований получаем следующие канонические виды уравнений:

1. Центральная поверхность

$$\lambda_1 x^2+\lambda_2 y^2+\lambda_3 z^2+c=0,$$

где хоть одно из λ_1, λ_2, λ_3 не нуль.

2. Нецентральная поверхность

$$\nu_1 x^2 + \nu_2 y^2 + \nu_3 z = 0,$$

где хоть одно из ν_1 и ν_2 не нуль и $\nu_3 \neq 0$. Эти уравнения определяют следующие геометрические образы:

1) $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ — эллипсоид;

2) $\frac{x^2}{a^2} - \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1$ — однополостный гиперболоид;

3) $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1$ — двухполостный гиперболоид;

4) $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = -1$ — мнимый эллипсоид;

5) $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 0$ — мнимый конус;

6) $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0$ — конус;

7) $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ — эллиптический цилиндр;

8) $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ — гиперболический цилиндр;

9) $\frac{x^2}{a^2} + \frac{y^2}{b^2} = -1$ — мнимый эллиптический цилиндр;

10) $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 0$ — две мнимые пересекающиеся плоскости;

11) $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0$ — две пересекающиеся плоскости;

12) $x^2 = a^2$ — две параллельные плоскости;

13) $x^2 = -a^2$ — две мнимые параллельные плоскости;

14) $x^2 = 0$ — две слившиеся плоскости;

15) $\frac{x^2}{p} + \frac{y^2}{q} = 2z$ — эллиптический параболоид

$$(pq > 0);$$

16) $\frac{x^2}{p} - \frac{y^2}{q} = 2z$ — гиперболический параболоид

$$(pq > 0);$$

17) $y^2 = 2px$ — параболический цилиндр.

Примеры.

1. Привести к простейшему виду уравнение линии

$$5x^2+4xy+8y^2-32x-56y+80=0.$$

Решение. Найдем характеристические корни матрицы $A=\begin{pmatrix}5 & 2\\ 2 & 8\end{pmatrix}$

$$\begin{vmatrix}5-\lambda & 2\\ 2 & 8-\lambda\end{vmatrix}=0;\ \lambda^2-13\lambda+36=0,\ \text{отсюда}\ \lambda_1=4;\ \lambda_2=9.$$

Так как $\det A=36\neq 0$, то легко составить преобразованное уравнение. Для этого (см. п. 2 § 2) найдем K

$$K=\begin{vmatrix}5 & 2 & -16\\ 2 & 8 & -28\\ -16 & -28 & 80\end{vmatrix}=-16\cdot 81.$$

Значит, $c_1=-36$ и канонический вид уравнения будет

$$4x_1^2+9y_1^2=36,\quad \text{или}\quad \frac{x_1^2}{9}+\frac{y_1^2}{4}=1.$$

Дополним условие задачи требованием определить расположение данной линии относительно исходной системы координат. В этом случае задача усложняется. Надо найти угол поворота осей координат и новое начало координат.

Сначала найдем направление новых осей координат. У нас $\lambda_1=4$, $\lambda_2=9$. Координаты соответствующих собственных векторов находятся из систем

$$(5-\lambda)\,u_1+2u_2=0,$$
$$2u_1+(8-\lambda)\,u_2=0.$$

При $\lambda_1=4$ имеем

$$u_1+2u_2=0,$$
$$2u_1+4u_2=0.$$

Отсюда $u_1=2$; $u_2=-1$.

При $\lambda_2=9$ имеем

$$-4u_1+2u_2=0,$$
$$2u_1-u_2=0.$$

Отсюда $u_1=1$; $u_2=2$.

Нормированные собственные векторы имеют вид

$$e_1=\frac{2}{\sqrt{5}}i-\frac{1}{\sqrt{5}}j,$$
$$e_2=\frac{1}{\sqrt{5}}i+\frac{2}{\sqrt{5}}j.$$

Новое начало координат (так как $\det A\neq 0$) легко найти методом, указанным в п. 2 § 2.

Второй способ, годный для центральных и нецентральных кривых. Составим матрицу линейного преобразования

$$T=\begin{pmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix}.$$

Значит, формулы преобразования координат имеют вид

$$x=\frac{2}{\sqrt{5}}x_1+\frac{1}{\sqrt{5}}y_1,$$

$$y=-\frac{1}{\sqrt{5}}x_1+\frac{2}{\sqrt{5}}y_1.$$

Внесем x_1 и y_1 в левую часть данного уравнения. Так как трехчлен $5x^2+4xy+8y^2$ в новой системе будет иметь вид $\lambda_1x_1^2+\lambda_2y_1^2=$ $=4x_1^2+9y_1^2$, то надо преобразовать лишь члены, содержащие x и y в первой степени

$$-32x-56y=\frac{-64x_1-32y_1}{\sqrt{5}}+\frac{56x_1-112y_1}{\sqrt{5}}=-\frac{8}{\sqrt{5}}x_1-\frac{144}{\sqrt{5}}y_1.$$

Следовательно, в новой системе координат уравнение имеет вид

$$4x_1^2+9y_1^2-\frac{8}{\sqrt{5}}x_1-\frac{144}{\sqrt{5}}y_1+80=0$$

или

$$4\left(x_1-\frac{1}{\sqrt{5}}\right)^2+9\left(y_1-\frac{8}{\sqrt{5}}\right)^2-36=0.$$

Обозначая $x_1-\frac{1}{\sqrt{5}}=x_2$; $y_1-\frac{8}{\sqrt{5}}=y_2$, имеем $4x_2^2+9y_2^2=36$.

Окончательно $\frac{x_2^2}{9}+\frac{y_2^2}{4}=1$.

Очевидно, направление повернутых осей координат x_1 и y_1 определяется векторами e_1 и e_2. Затем параллельным переносом осей x_1 и y_1 в точку с координатами $\left(\frac{1}{\sqrt{5}}, \frac{8}{\sqrt{5}}\right)$ в системе x_1, y_1 получаем систему координат x_2, y_2.

2. Определить канонический вид и расположение поверхности

$$2x^2+y^2+2z^2-2xy+2yz+4x-2y=0$$

$$A=\begin{pmatrix} 2 & -1 & 0 \\ -1 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix}; \quad \begin{vmatrix} 2-\lambda & -1 & 0 \\ -1 & 1-\lambda & 1 \\ 0 & 1 & 2-\lambda \end{vmatrix}=0,$$

отсюда $\lambda_1=3$; $\lambda_2=2$; $\lambda_3=0$.

Находим направление новых осей

$$(2-\lambda)u_1-u_2=0,$$
$$-u_1+(1-\lambda)u_2+u_3=0,$$
$$u_2+(2-\lambda)u_3=0,$$

при $\lambda=3$ $u_1=1$, $u_2=-1$, $u_3=-1$;
при $\lambda=2$ $u_1=1$, $u_2=0$, $u_3=1$;
при $\lambda=0$ $u_1=1$, $u_2=2$, $u_3=-1$.

Нормированные собственные векторы имеют вид

$$e_1=\frac{1}{\sqrt{3}}i-\frac{1}{\sqrt{3}}j-\frac{1}{\sqrt{3}}k,$$
$$e_2=\frac{1}{\sqrt{2}}i+\frac{1}{\sqrt{2}}k,$$
$$e_3=\frac{1}{\sqrt{6}}i+\frac{1}{\sqrt{6}}j-\frac{1}{\sqrt{6}}k,$$
$$T=\begin{pmatrix}\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}}\\ -\frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}}\\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}}\end{pmatrix}.$$

Следовательно,

$$x=\frac{1}{\sqrt{3}}x_1+\frac{1}{\sqrt{2}}y_1+\frac{1}{\sqrt{6}}z_1,$$
$$y=-\frac{1}{\sqrt{3}}x_1+\frac{2}{\sqrt{6}}z_1,$$
$$z=-\frac{1}{\sqrt{3}}x_1+\frac{1}{\sqrt{2}}y_1-\frac{1}{\sqrt{6}}z_1,$$
$$4x-2y=\frac{6x_1}{\sqrt{3}}+\frac{4y_1}{\sqrt{2}}.$$

Преобразованное (после поворота осей) уравнение принимает вид

$$3x_1^2+2y_1^2+\frac{6x_1}{\sqrt{3}}+\frac{4y_1}{\sqrt{2}}=0,$$

или

$$3\left(x_1+\frac{1}{\sqrt{3}}\right)^2+2\left(y_1+\frac{1}{\sqrt{2}}\right)^2=2.$$

Обозначим

$$x_1+\frac{1}{\sqrt{3}}=x_2,\quad y_1+\frac{1}{\sqrt{2}}=y_2.$$

Имеем $3x_2^2+2y_2^2=2$, или $\frac{x_2^2}{\frac{2}{3}}+\frac{y_2^2}{1}=1$ — эллиптический цилиндр.

§ 4. ЗНАКООПРЕДЕЛЕННЫЕ КВАДРАТИЧНЫЕ ФОРМЫ

Во многих вопросах, связанных с применением квадратичных форм, особую роль играет знак формы при различных значениях переменных.

1. Положительно и отрицательно определенные формы. Действительная квадратичная форма $L = X'AX$ называется *положительно определенной*, если для любых действительных значений переменных $x_1, x_2, \ldots x_n$, не всех равных нулю, $L > 0$. Если при этих условиях $L < 0$, то форма называется *отрицательно опредгленной*.

В обоих этих случаях форма L называется *знакоопределенной*.

Если L не изменяет знака, но при некоторых значениях переменных (которые не все равны нулю) обращается в нуль, то форма L называется *знакопостоянной*.

Например, очевидно, что форма $L = 2x^2 + 3y^2$ — положительно определенная, а форма $L = -x^2 - y^2$ — отрицательно определенная. Несколько сложнее дать ответ в следующем случае $L = 2xy - x^2 - 3y^2$. После элементарного преобразования получаем $L = -(x-y)^2 - 2y^2$. Значит, форма — отрицательно определенная.

Очевидно, что при большом числе переменных ответить на вопрос, будет ли квадратичная форма знакоопределенной, сложно. Докажем теорему, которая поможет дать ответ на поставленный вопрос.

Теорема. *Для того чтобы квадратичная форма $L = X'AX$ была знакоопределенной, необходимо и достаточно, чтобы характеристические числа матрицы A имели одинаковые знаки (плюс для положительно определенной и минус для отрицательно определенной формы).*

Очевидно, что если L — положительно определенная форма, то форма $-L$ — отрицательно определенная. Поэтому нам достаточно рассмотреть лишь случай положительно определенной формы.

Приведем форму $L = X'AX$ к каноническому виду $L = \sum_{i=1}^{n} \lambda_i y_i^2$ ортогональным преобразованием $X = TY$. Так как матрица T — ортогональная, а, значит, — невырожденная, то $Y = T^{-1}X$. Поэтому, если $X \neq 0$, то $Y \neq 0$ и, наоборот.

Доказательство необходимого условия.

Пусть при $X \neq 0$ $L > 0$. Значит, и при $Y \neq 0$ $L > 0$. Предположим, что хоть одно λ_i, например, $\lambda_1 \leqslant 0$. Тогда, полагая $y_1 = 1$; $y_2 = y_3 = \ldots = y_n = 0$, имеем $L = \lambda_1 \leqslant 0$, что противоречит условию.

Доказательство достаточного условия.

Пусть все $\lambda_i > 0$, $i = 1, 2, \ldots n$. Очевидно, что $L = \sum_{i=1}^{n} \lambda_i y_i^2 > 0$, если $Y \neq 0$ (значит, и $X \neq 0$), что и требовалось доказать.

Отметим, что практическое применение результатов этой теоремы затруднительно, так как для нахождения характеристических чисел матрицы A необходимо решить уравнение n-й степени. Приведем доказательство более удобного для приложений признака.

2. Критерий Сильвестра. *Для того чтобы квадратичная форма $L = X'AX$ была положительно определенной, необходимо и достаточно, чтобы выполнялись неравенства*

$$\det A_1 = a_{11} > 0,$$

$$\det A_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \ldots, \det A_n = \det A = \begin{vmatrix} a_{11} \ldots a_{1n} \\ a_{21} \ldots a_{2n} \\ \cdot \quad \cdot \quad \cdot \\ a_{n1} \ldots a_{nn} \end{vmatrix} > 0.$$

Докажем теорему методом полной математической индукции. При $n = 1$ $L = a_{11}x_1^2$. В этом случае критерий очевиден, так как $a_{11}x_1^2 > 0$, тогда и только тогда, когда $a_{11} > 0$. Предположим, что критерий справедлив для формы $n-1$-го порядка. Покажем, что он будет справедлив и для формы n-го порядка. Докажем необходимость условий. Имеем форму n-го порядка $L = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}x_ix_j$. Пусть $L > 0$. Положим $x_n = 0$. Получим форму $n-1$-го порядка $L_1 = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}x_ix_j$, для которой по предположению индукции выполняются неравенства

$$\det A_1 > 0, \quad \det A_2 > 0, \ldots \det A_{n-1} > 0.$$

Остается показать, что $\det A > 0$. Для этого форму L ортогональным преобразованием приведем к каноническому

виду. Учитывая неизменность $\det A$ при этом преобразовании и предыдущую теорему, получаем

$$\det A = \begin{vmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & \lambda_n \end{vmatrix} = \lambda_1\lambda_2\dots\lambda_n > 0.$$

Докажем теперь достаточность условий. Допустим, что

$$\det A_1 > 0,\ \det A_2 > 0, \dots, \det A_n = \det A > 0.$$

Форму $L = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$ представим в виде

$$L = L_1 + 2\sum_{i=1}^{n} a_{in} x_i x_n + a_{nn} x_n^2,$$

где

$$L_1 = \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} a_{ij} x_i x_j.$$

При произвольном, но фиксированном x_n, мы можем рассматривать L как неоднородный многочлен второй степени от переменных $x_1, x_2, \dots x_{n-1}$. Так как определитель матрицы квадратичной формы L_1 отличен от нуля (по условию $\det A_{n-1} > 0$), то этот многочлен — центральный. Поэтому мы можем преобразовать его к виду

$$L = \sum_{i=1}^{n-1} \lambda_i^* y_i^2 + \frac{K}{\det A_{n-1}} x_n^2,$$

где λ_i^* — характеристические числа матрицы A_{n-1}, а

$$K = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \det A.$$

По условию все $\lambda_i^* > 0$ (по предположению индукции форма L_1 — положительно определенная) и $\det A > 0$. Поэтому $L > 0$, что и требовалось доказать.

Следствие. *Чтобы квадратичная форма L была отрицательно определенной, необходимо и достаточно выполнение неравенств*

$$\det A_1 < 0,\ \det A_2 > 0,\ \det A_3 < 0,\ \ldots\ (-1)^n \det A_n > 0.$$

Это очевидно, так как в этом случае форма $-L$ — положительно определенная.

Примеры.

1. $L = -x^2 + 2xy - 3y^2,\quad A = \begin{pmatrix} -1 & 1 \\ 1 & -3 \end{pmatrix},$

$$\det A_1 = -1 < 0,\quad \det A_2 = \begin{vmatrix} -1 & 1 \\ 1 & -3 \end{vmatrix} = 2 > 0.$$

Форма L — отрицательно определенная.

2. $L = ax^2 + 2bxy + cy^2$.

Аналогично предыдущему получаем: при $ac - b^2 > 0$ форма — знакоопределенная, причем при $a > 0$ она положительно определенная, а при $a < 0$ — отрицательно определенная.

Упражнения

1. Привести к каноническому виду квадратичную форму

$$3x_2^2 + 3x_3^2 + 4x_1x_2 + 4x_1x_3 - 2x_2x_3.$$

2. Привести к каноническому виду квадратичную форму

$$6x_1^2 + 5x_2^2 + 7x_3^2 - 4x_1x_2 + 4x_2x_3$$

и найти преобразующую матрицу.

3. Показать, что квадратичная форма

$$7x_1^2 + 7x_2^2 + 7x_3^2 + 2x_1x_2 + 2x_1x_3 + 2x_2x_3$$

положительно определенная.

4. Привести к каноническому виду уравнения следующих поверхностей:

а) $6x^2 - 2y^2 + 6z^2 + 4xz + 8x - 4y - 8z + 1 = 0$;

б) $4x^2 + 3y^2 + 2z^2 + 4xy - 4yz + 4x - 2y - 4z - 3 = 0$;

в) $2x^2 + 2y^2 + 3z^2 + 4xy + 2xz + 2yz - 4x + 6y - 2z + 3 = 0$.

5. При каких значениях параметра λ квадратичная форма

$$5x_1^2 + x_3^2 + \lambda x_3^2 + 4x_1x_2 - 2x_1x_3 - 2x_2x_3$$

будет положительно определенной.

Глава 8

ЖОРДАНОВА НОРМАЛЬНАЯ ФОРМА МАТРИЦ

§ 1. ВСПОМОГАТЕЛЬНЫЙ БАЗИС СПЕЦИАЛЬНОГО ВИДА

1. Постановка задачи. Ранее было показано, что для того, чтобы матрица A n-го порядка была подобна диагональной матрице, необходимо и достаточно, чтобы соответствующий ей оператор $\tilde{A}$ был оператором простой структуры, т. е. чтобы он имел n линейно независимых собственных векторов.

В противном случае, т. е. когда число m линейно независимых векторов оператора $\tilde{A}$ меньше чем n, то в любом базисе число столбцов матрицы A, все элементы которых, кроме диагонального, равны нулю, не больше чем m. Для того чтобы число указанных столбцов было наибольшим, т. е. равнялось m, надо, чтобы все линейно независимые собственные векторы были включены в базис. Однако, как бы мы не выбирали недостающие $n-m$ базисных векторов, соответствующие этим векторам столбцы матрицы A будут содержать более одного отличного от нуля элемента. Естественно считать, что матрица, в которой m столбцов содержат лишь один (диагональный) ненулевой элемент, а остальные $n-m$ столбцов содержат минимальное число элементов, отличных от нуля, является простейшей (после диагональной) матрицей, подобной матрице, не имеющей простой структуры. Подробному выявлению вида такой матрицы и способу построения соответствующего ей базиса и посвящается настоящая глава. Итак, пусть оператор $\tilde{A}$ не имеет простой структуры, это значит, что хотя бы для одного характеристического числа λ_i $(i=1, 2\ldots\nu)$ ранг r_i матрицы $A-\lambda_i E$ отличен от $n-k_i$, где k_i кратность корня λ_i характеристического многочлена $\Delta(\lambda)$ матрицы A. Поэтому $n-k_i<r_i$ (гл. 5, § 2, п. 5) и, следовательно, система

$$(A-\lambda_i E)X=0$$

имеет только $m_i=n-r_i<k_i$ линейно независимых решений, а это значит, что оператор $\tilde{A}$ имеет соответственно $m_i<k_i$ линейно независимых собственных векторов

с характеристическим числом λ_i. Для построения базиса надо найти (для каждого λ_i) недостающие $k_i - m_i = k_i + r_i - n$ векторов таких, которые вместе со всеми $m = m_1 + m_2 + \ldots + m_\nu$ линейно независимыми собственными векторами оператора образуют линейно независимую систему.

2. Построение базиса специального вида. Пусть ранг оператора $\tilde{A} - \lambda\tilde{E}$ в пространстве R равен r. Это значит, что оператор $\tilde{A}$ имеет $n - r$ линейно независимых собственных векторов $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{n-r}$, соответствующих собственному числу λ.

Рассмотрим подпространство $L = (\tilde{A} - \lambda\tilde{E})R$. Из теоремы о ранге следует, что размерность его равна r. Выберем в L базис $\boldsymbol{g}_1, \boldsymbol{g}_2, \ldots, \boldsymbol{g}_r$. Обозначим прообразы этих векторов $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_r$, т. е.

$$(\tilde{A} - \lambda\tilde{E})\boldsymbol{y}_i = \boldsymbol{g}_i, \quad i = 1, 2, \ldots, r.$$

Лемма. *Векторы* $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{n-r}, \boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_r$ *образуют базис пространства* R.

Количество этих векторов совпадает с размерностью пространства R. Поэтому нам достаточно доказать, что векторы $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{n-r}, \boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_r$ линейно независимы.

Предположим, что это не так, т. е.

$$\sum_{i=1}^{n-r} \alpha_i \boldsymbol{x}_i + \sum_{i=1}^{r} \beta_i \boldsymbol{y}_i = 0,$$

где хотя бы один из коэффициентов отличен от нуля. Отсюда

$$(\tilde{A} - \lambda\tilde{E}) \sum_{i=1}^{n-r} \alpha_i \boldsymbol{x}_i + (\tilde{A} - \lambda\tilde{E}) \sum_{i=i}^{r} \beta_i \boldsymbol{y}_i = 0,$$

или

$$\sum_{i=1}^{r} \beta_i \boldsymbol{g}_i = 0.$$

Так как векторы $\boldsymbol{g}_i$ образуют базис L, то $\beta_i = 0$, $i = 1, 2, \ldots, r$. Значит, $\sum_{i=1}^{n-r} \alpha_i \boldsymbol{x}_i = 0$, что противоречит линейной независимости векторов $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{n-r}$.

Итак, базис пространства R можно построить из собственных векторов, оператора $\tilde{A}$, соответствующих

числу λ, и прообразов базисных векторов подпространства $L=(\tilde{A}-\lambda\tilde{E})R$.

Используем лемму для построения базиса специального вида.

Пусть в подпространстве $L=(\tilde{A}-\lambda\tilde{E})R$ размерности r имеется m линейно независимых собственных векторов оператора $\tilde{A}$, соответствующих числу λ.

Обозначим их

$$\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_m, \text{ где } m \leqslant r \text{ и } m \leqslant n-r$$

и дополним их векторами

$$\boldsymbol{g}_1, \boldsymbol{g}_2, \ldots, \boldsymbol{g}_{r-m}$$

до базиса подпространства L, а также обозначим $\boldsymbol{y}_i$ прообразы векторов $\boldsymbol{x}_i$ и $\boldsymbol{z}_i$ прообразы векторов $\boldsymbol{g}_i$, т. е.

$$(\tilde{A}-\lambda\tilde{E})\boldsymbol{y}_i=\boldsymbol{x}_i\ (i=1, 2, \ldots, m),\ (\tilde{A}-\lambda\tilde{E})\boldsymbol{z}_i=\boldsymbol{g}_i$$
$$i=1, 2, \ldots, r-m.$$

Из доказанной леммы следует **теорема.** *Можно построить базис пространства R из векторов*

$$\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_m, \boldsymbol{z}_1, \boldsymbol{z}_2 \ldots, \boldsymbol{z}_{r-m}, \boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{n-r},$$

где $\boldsymbol{x}_{m+1}, \boldsymbol{x}_{m+r}, \ldots, \boldsymbol{x}_{n-r}$ — *недостающие собственные векторы оператора A, соответствующие числу λ, т. е. такие, что векторы*

$$\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{n-r} —$$

линейно независимы.

Содержание доказанного утверждения иллюстрируется схемой, приведенной на рис. 4.

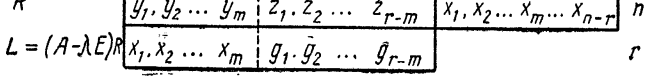


Рис. 4

§ 2. ЖОРДАНОВ БАЗИС В ЧАСТНОМ СЛУЧАЕ

1. Жорданов базис. Рассмотрим частный случай, когда оператор $\tilde{A}$ имеет единственное собственное число λ и не имеет простой структуры. В этом случае матрица $A-\lambda E$ имеет ранг r, не равный нулю (при $r=0$ оператор

имеет простую структуру), а значит в пространстве R имеется $n-r$ линейно независимых собственных векторов оператора $\tilde{A}$, соответствующих числу λ.

Рассмотрим подпространство $L_1=(\tilde{A}-\lambda\tilde{E})R$. Его размерность равна $r\neq 0$. Следовательно, в L_1, как известно (см. гл. 5, § 2, п. 2), найдется хотя бы один собственный вектор оператора $\tilde{A}$.

Поэтому размерность подпространства $L_2=(\tilde{A}-\lambda\tilde{E})L_1=(\tilde{A}-\lambda\tilde{E})^2R$ будет меньше размерности L_1.

Если L_2 — ненулевое пространство, то в нем найдется хотя бы один собственный вектор оператора $\tilde{A}$. В этом случае построим подпространство $L_3=(\tilde{A}-\lambda\tilde{E})L_2=(\tilde{A}-\lambda\tilde{E})^3R$, размерность которого меньше размерности L_2. Продолжая процесс, мы получим ненулевое подпространство L_k такое, что

$$L_{k+1}=(\tilde{A}-\lambda\tilde{E})L_k=(\tilde{A}-\lambda\tilde{E})^{k+1}R=0.$$

Это означает, что пространство L_k содержит только собственные векторы оператора $\tilde{A}$, значит, оператор $\tilde{A}$ в L_k имеет простую структуру.

Выберем в L_k базис. Используя предыдущую теорему, мы можем дополнить выбранный базис L_k до базиса L_{k-1}, затем базис L_{k-1} дополнить до базиса L_{k-2} и т. д. В итоге получим базис всего пространства R. Перейдем к его построению. Пусть p_1 — размерность L_k и векторы $x_1, x_2, \ldots, x_{p_1}$ образуют его базис, а пространство L_{k-1} содержит $p_2\geqslant p_1$ линейно независимых собственных векторов оператора $\tilde{A}$. Тогда векторы

$$x_1,\ x_2,\ldots,\ x_{p_1};\ y_1 y_2,\ldots,\ y_{p_1};\ x_{p_1+1},\ldots,\ x_{p_2},$$

где

$$(\tilde{A}-\lambda\tilde{E})y_i=x_i\quad (i=1,\ 2,\ldots,\ p_1);$$
$$(\tilde{A}-\lambda\tilde{E})x_i=0\quad (i=1,\ 2,\ldots,\ p_2)$$

образуют базис L_{k-1}, размерность которого равна p_1+p_2. Аналогично векторы

$$x_1, x_2,\ldots, x_{p_1};\ y_1, y_2,\ldots,\ y_{p_1};\ z_1,\ z_2,\ldots, z_{p_1};\ x_{p_1+1},\ldots, x_{p_2};$$
$$y_{p_1+1},\ldots,\ y_{p_2};\ x_{p_2+1},\ldots,\ x_{p_3},$$

где

$$(\tilde{A}-\lambda\tilde{E})\,z_i=y_i,\quad i=1,\ldots,p_1,$$
$$(\tilde{A}-\lambda\tilde{E})\,y_i=x_i,\quad i=1,\ldots,p_2,$$
$$(\tilde{A}-\lambda\tilde{E})\,x_i=0,\quad i=1,\ldots,p_3$$

($p_3\geqslant p_2$ — число линейно независимых собственных векторов оператора $\tilde{A}$, содержащихся в пространстве L_{k-2}) образуют базис L_{k-2}, размерность которого равна $p_1+p_2+p_3$. Продолжая процесс, получим базис пространства R. Этот базис называется *жордановым базисом*.

Для последующего нам будет полезно разбить полученный базис на отдельные группы векторов.

2. Жорданова цепочка векторов. Рассмотрим векторы x_i $(i=1, 2,\ldots, p_1)$ из L_k. Как мы видели, они являются образами векторов y_i $(i=1, 2,\ldots, p_1)$ пространства L_{k-1}, которые в свою очередь являются образами векторов z_i $(i=1, 2,\ldots, p_1)$ из L_{k-2} и т. д.

Совокупность этих векторов $x_i, y_i, z_i,\ldots$ при фиксированном i назовем *жордановой цепочкой векторов* длины $k+1$. Таким образом, мы получили p_1 жордановых цепочек, каждая из которых состоит из $k+1$ вектора.

Рассмотрим теперь векторы x_i $(i=p_1+1,\ldots, p_2)$ из L_{k-1}, каждому из этих векторов также соответствует цепочка из k векторов:

$$x_i,\ y_i,\ z_i,\ldots,\ v_i\ \ (i=p_1+1,\ldots, p_2).$$

Таким образом, мы получаем дополнительно p_2-p_1 жордановых цепочек длины k. Рассуждая аналогичным образом, получим далее p_3-p_2 жордановых цепочек длины $k-1$, p_4-p_3 цепочек длины $k-2$ и т. д. Наконец, $p_{k+1}-p_k$ цепочек длины 1, т. е. цепочек, состоящих из одного собственного вектора оператора $\tilde{A}$. Здесь p_ν — число линейно независимых собственных векторов оператора $\tilde{A}$, содержащихся в пространстве $L_{k+1-\nu}$, размерность которого равна

$$p_1+p_2+\ldots+p_\nu\ \ (\nu=1, 2,\ldots, k+1),$$
$$p_{k+1}=n-r \text{ и } L_0=R.$$

Таким образом, общее число жордановых цепочек равно

$$p_1+(p_2-p_1)+\ldots+(p_{k+1}-p_k)=p_{k+1}=n-r,$$

т. е. числу линейно независимых собственных векторов оператора $\tilde{A}$, а их суммарная длина (т. е. число векторов, содержащихся во всех цепочках) равна

$$p_1(k+1)+(p_2-p_1)k+\ldots+(p_{k+1}-p_k)1=$$
$$=p_1+p_2+\ldots+p_{k+1}=n,$$

т. е. размерности пространства R.

Построенный таким образом базис пространства R схематично изображен на рис. 5, на котором в столбцах

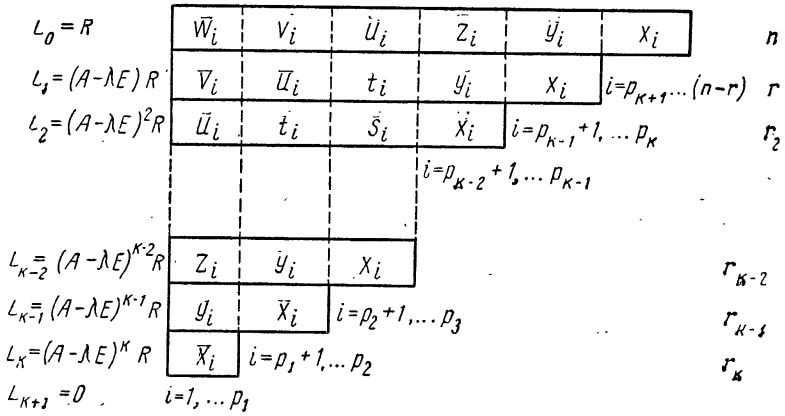


Рис. 5

показаны жордановы цепочки, а каждая строчка указывает, из какого подпространства выбираются векторы этих цепочек.

Отметим еще раз соотношения между векторами одной жордановой цепочки. Если обозначить

$$\boldsymbol{x},\ \boldsymbol{y},\ \boldsymbol{z},\ldots,\ \boldsymbol{u},\ \boldsymbol{v},\ \boldsymbol{w}$$

векторы одной цепочки длины s, то имеем

$$\begin{aligned}
&(\tilde{A}-\lambda\tilde{E})\,\boldsymbol{x}=0,\\
&(\tilde{A}-\lambda\tilde{E})\,\boldsymbol{y}=\boldsymbol{x},\\
&(\tilde{A}-\lambda\tilde{E})\,\boldsymbol{z}=\boldsymbol{y},\\
&\cdots\cdots\cdots\\
&(\tilde{A}-\lambda\tilde{E})\,\boldsymbol{v}=\boldsymbol{u},\\
&(\tilde{A}-\lambda\tilde{E})\,\boldsymbol{w}=\boldsymbol{v}.
\end{aligned}\qquad(8.1)$$

3. Жорданова матрица. Итак, мы построили жорданов базис, который состоит из $n-r$ жордановых цепочек, суммарная длина которых равна n. Покажем теперь, что в этом базисе оператор A имеет простейшую матрицу.

Соотношения (8.1) запишем в следующем виде:

$$\begin{aligned}
\tilde{A}\boldsymbol{x} &= \lambda\boldsymbol{x},\\
\tilde{A}\boldsymbol{y} &= \lambda\boldsymbol{y}+\boldsymbol{x},\\
\tilde{A}\boldsymbol{z} &= \lambda\boldsymbol{z}+\boldsymbol{y},\\
&\dots\dots\dots\\
\tilde{A}\boldsymbol{v} &= \lambda\boldsymbol{v}+\boldsymbol{u},\\
\tilde{A}\boldsymbol{w} &= \lambda\boldsymbol{w}+\boldsymbol{v}.
\end{aligned}$$

Напомним, что по определению матрицы линейного оператора столбцы ее составлены из координат образов базисных векторов. Поэтому, если векторы цепочки $\boldsymbol{x}$, $\boldsymbol{y}$, ..., $\boldsymbol{w}$ мы включим в базис, то в нем (в силу предыдущих соотношений) этой цепочке будет соответствовать в матрице оператора прямоугольная клетка из s столбцов и n строк вида

$$\begin{pmatrix}
 & & & & & & \mathrm{O} & & & & & \\
\lambda & 1 & 0 & \dots & & & 0 & 0\\
0 & \lambda & 1 & \dots & & & 0 & 0\\
\dots & \dots & \dots & \dots & & & & \\
0 & 0 & 0 & \dots & & & \lambda & 1\\
0 & 0 & 0 & \dots & & & 0 & \lambda\\
 & & & & & & \mathrm{O} & & & & &
\end{pmatrix}.$$

Поэтому в построенном нами жордановом базисе матрица оператора будет иметь вид

$$I=\begin{pmatrix}
I_{n_1} & 0 & \dots & 0\\
0 & I_{n_2} & \dots & 0\\
\dots & \dots & \dots & \dots\\
0 & 0 & \dots & I_{n_m}
\end{pmatrix},$$

где

$$I_{n_j}=\begin{pmatrix}\lambda & 1 & \dots & 0 & 0\\ 0 & \lambda & \dots & 0 & 0\\ \cdot & \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & \dots & \lambda & 1\\ 0 & 0 & \dots & 0 & \lambda\end{pmatrix}-$$

матрица, называемая жордановой клеткой порядка n_j; m — число жордановых цепочек, а $n_1+n_2+\ldots+n_m=n$. Матрица I называется *жордановой матрицей*, или *жордановой нормальной формой матрицы*. Очевидно, в частности, если оператор $\tilde{A}$ имеет простую структуру, то $I_{n_j}=I_1=\lambda$ и жорданова матрица вырождается в диагональную.

4. Построение жордановой матрицы. Как было показано выше, структура жордановой матрицы в рассматриваемом случае определена, если известны порядки n_j ($j=1, 2, \ldots, m$) жордановых клеток. Возникает вопрос, нельзя ли установить эти порядки, не находя предварительно жорданового базиса? На этот вопрос можно дать утвердительный ответ. Из рассуждений п. 2 ясно, что длины жордановых цепочек векторов однозначно связаны с размерностями вспомогательных пространств $L_1, L_2, \ldots, L_k$, а, следовательно (см. гл. 4), с рангами матриц $A-\lambda E$, $(A-\lambda E)^2, \ldots, (A-\lambda E)^k$. Найдем явные формулы для определения длин жордановых цепочек. Предварительно обозначим r_ν ранг матрицы $(A-\lambda E)^\nu$, а q_h — количество цепочек длины h ($h=1, 2, \ldots, k-1$).

В п. 2 было показано, что имеется p_1 цепочек максимальной длины $k+1$, где p_1 — размерность пространства L_k, т. е. $p_1=r_k$ и $q_{k+1}=r_k$. Жордановых цепочек длины k будет $q_k=p_2-p_1$, где p_2 максимальное число линейно независимых собственных векторов в пространстве L_{k-1}, т. е. $p_2=r_{k-1}-r_k$. Таким образом, $q_k=r_{k-1}-2r_k$. Так как число жордановых цепочек длины h равно

$p_{k+2-h}-p_{k+1-h}$, а $r_{\nu-1}=p_1+p_2+\ldots+p_{k+2-\nu}=r_\nu+p_{k+2-\nu}$

и, значит,

$p_{k+2-\nu}=(r_{\nu-1}-r_\nu)$, то

$$q_h=(r_{h-1}-r_h)-(r_h-r_{h-1})=r_{h-1}-2r_h+r_{h+1},$$

где $h=1, 2, \ldots, k-1$; $r_0=n$; $r_1=r$; $r_{k+1}=r_{k+2}=0$.

Таким образом, мы определили количество всех жордановых клеток длины h, а тем самым количество всех жордановых клеток, составляющих жорданову матрицу и их порядки n_j.

5. **Единственность жордановой матрицы.** Докажем, что для данной матрицы существует единственная совокупность жордановых клеток, расположение которых по диагонали жордановой матрицы может быть произвольным. Иначе: докажем, что существует единственная (с точностью до порядка расположения клеток) жорданова матрица, подобная данной.

В самом деле, жорданова матрица определена, если известны величины q_h, но $q_h = r_{h-1} - 2r_h + r_{h+1}$ и так как ранги r_ν не зависят от выбранного базиса, то и q_h не зависит от базиса.

Значит, если оператор $\tilde{A}$ задан, то q_h, а поэтому и матрицы I_{n_j} однозначно определены.

Пример.

Дана матрица оператора

$$A = \begin{pmatrix} 3 & 2 & -3 \\ 4 & 10 & -12 \\ 3 & 6 & -7 \end{pmatrix}.$$ Найти ее жорданову матрицу.

Решение. Характеристическое уравнение

$$\Delta(\lambda) = |A - \lambda E| = \begin{vmatrix} 3-\lambda & 2 & -3 \\ 4 & 10-\lambda & -12 \\ 3 & 6 & -7-\lambda \end{vmatrix} = 0$$

имеет единственный корень $\lambda = 2$ кратности 3.

Так как ранг r_1 матрицы

$$A - 2E = \begin{pmatrix} 1 & 2 & -3 \\ 4 & 8 & -12 \\ 3 & 6 & -9 \end{pmatrix}$$

равен 1, то оператор не имеет простой структуры.

Так как $(A - 2E)^2 = 0$, то

$$r_2 = 0, \text{ т. е. } k + 1 = 2 \text{ и } k = 1.$$

Значит, $q_2 = r_1 = 1$, $q_1 = r_0 - 2r_1 = 3 - 2 = 1$. Итак, имеем одну клетку порядка 2 и одну клетку порядка 1. Поэтому

$$I = \left(\begin{array}{cc:c} 2 & 1 & 0 \\ 0 & 2 & 0 \\ \hdashline 0 & 0 & 2 \end{array}\right).$$

6. Построение жорданова базиса. Укажем следующий практический прием для построения жорданова базиса, предполагая выполненными все предыдущие условия (оператор $\tilde{A}$ не имеет простой структуры в R и имеет единственное собственное число λ кратности n).

Подпространство $L_1 = (A - \lambda\tilde{E})R$, построенное в п. 2, имеет размерность $r \neq 0$. Его базис образуют r линейно независимых столбцевых векторов матрицы $A - \lambda E$ (см. гл. 5, § 2, п. 6). Аналогично базис $L_2 = (\tilde{A} - \lambda\tilde{E})L_1 = (\tilde{A} - \lambda\tilde{E})^2 R$ можно построить из столбцевых векторов матрицы $(A - \lambda E)^2$ и т. д. Наконец, базис $L_k = (\tilde{A} - \lambda\tilde{E})R$ можно построить из столбцевых векторов матрицы $(A - \lambda E)^k$. Так как пространство L_{k+1} — нулевое, то $(A - \lambda E)^{k+1} = 0$. Поэтому за векторы $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots \boldsymbol{x}_p$, образующие базис L_k, можно взять любые p_1 линейно независимые столбцевые векторы матрицы $(A - \lambda E)^k$. Теперь построим базис L_{k-1}. Очевидно, векторы $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots \boldsymbol{y}_p$ — это столбцевые векторы матрицы $(A - \lambda E)^{k-1}$ с теми же номерами, что и выбранные столбцы матрицы $(A - \lambda E)^k$. Недостающие собственные векторы $\boldsymbol{x}_{p_1+1}, \ldots, \boldsymbol{x}_{p_2}$ найдем методом, рассмотренным в гл. 5, § 2, п. 3. Аналогично строим базис пространства L_{k-2} и т. д.

Таким образом, для построения жорданова базиса можно поступать следующим образом:

1. Построить совокупность матриц

$$E,\ A - \lambda E,\ (A - \lambda E)^2,\ \ldots,\ (A - \lambda E)^k \neq 0,\ (A - \lambda E)^{k+1} = 0.$$

2. Выберем p_1 линейно независимых столбцов матрицы $(A - \lambda E)^k$ и все столбцы с теми же номерами всех предшествующих матриц. Таким образом, получим p_1 жордановых цепочек длины $k + 1$.

3. С помощью элементарных преобразований остальные $n - p_1$ столбцов матрицы $(A - \lambda E)^k$ сделаем нулевыми. Те же преобразования проделаем под всеми предшествующими матрицами. Получим последовательность матриц

$$E_1,\ (A - \lambda E)_1,\ \ldots,\ (A - \lambda E)_1^{k-1},\ (A - \lambda E)_1^k.$$

После этого столбцами матрицы $(A - \lambda E)_1^{k-1}$, соответствующими нулевым столбцам матрицы $(A - \lambda E)_1^k$, будут либо собственные векторы, либо нулевые векторы.

4. Выберем $p_2 - p_1$ столбцевых векторов матрицы $(A - \lambda E)_1^{k-1}$ (линейно независимых с выбранными раньше) и все столбцы с теми же номерами матриц

$$E_1,\ (A - \lambda E)_1,\ \ldots,\ (A - \lambda E)_1^{k-2}.$$

Мы получим $p_2 - p_1$ жордановых цепочек длины k.

Таким же образом, продолжая процесс, можно построить жорданов базис всего пространства.

Примеры.

1. Найдем жорданов базис оператора предыдущего примера

$$A = \begin{pmatrix} 3 & 2 & -3 \\ 4 & 10 & -12 \\ 3 & 6 & -7 \end{pmatrix} \text{ и } \lambda_1 = \lambda_2 = \lambda_3 = 2.$$

Составляем матрицы

$$E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},\quad A - 2E = \begin{pmatrix} 1 & 2 & -3 \\ 4 & 8 & -12 \\ 3 & 6 & -9 \end{pmatrix},\quad (A - 2E)^2 = 0.$$

В данном случае $p_1 = 1$. В качестве собственного вектора x_1 можно взять любой столбец матрицы $A - 2E$. Выберем первый, т. е. $x_1 = (1, 4, 3)$. Тогда первый столбец матрицы E дает вектор $y = (1, 0, 0)$. Для отыскания недостающего собственного вектора проделаем элементарные преобразования над столбцами матрицы $\begin{pmatrix} E \\ A - 2E \end{pmatrix}$

$$\begin{pmatrix} E \\ A - 2E \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 2 & -3 \\ 4 & 8 & -12 \\ 3 & 6 & -9 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 4 & 0 & 0 \\ 3 & 0 & 0 \end{pmatrix}.$$

Значит, векторы $x_2 = (-2, 1, 0)$ и $x_3 = (3, 0, 1)$ — собственные. Один из них, например x_2, включим в базис.

В данном случае безразлично какой из векторов x_2 или x_3 включать в базис, так как очевидно, что пары x_1, x_2 и x_1, x_3 линейно независимы. В общем случае, когда $p_1 > 1$ и $p_2 - p_1 > 1$ полезно после получения $n - p_1$ нулевых столбцов матрицы $(A - \lambda E)_1^k$ продолжить элементарные преобразования этих столбцов (которые оставят их нулевыми) до тех пор пока в соответствующих столбцах матрицы $(A - \lambda E)_1^{k-1}$ не появится p_1 столбцов, совпадающих с ненулевыми столбцами матрицы $(A - \lambda E)_1^k$. Тогда для получения недостающих $p_2 - p_1$ собственных векторов достаточно

выбрать отличные от $x_1, x_2, \ldots x_{p_1}$ ненулевые столбцы матрицы $(A-\lambda E)_1^{k-1}$, соответствующие нулевым столбцам матрицы $(A-\lambda E)_1^k$. В нашем случае это равносильно, например, прибавлению второго столбца, умноженного на 4, к третьему столбцу, предварительно умноженному на 3

$$\begin{pmatrix} 1 & -2 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 4 & 0 & 0 \\ 3 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & 1 \\ 0 & 1 & 4 \\ 0 & 0 & 3 \\ 1 & 0 & 0 \\ 4 & 0 & 0 \\ 3 & 0 & 0 \end{pmatrix},$$

после чего ясно, что в качестве x_2 надо выбрать вектор $x = (-2, 1, 0)$. При больших p_1 и $p_2 - p_1$ можно воспользоваться общим приемом, указанным в задаче главы 3.

Таким образом, мы получили жорданов базис, состоящий из двух цепочек x_1, y_1 и x_2. В этом базисе матрица оператора $\tilde{A}$ имеет вид

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

Заметим, что иногда нет необходимости производить элементарные операции.

2. Пусть

$$A = \begin{pmatrix} 4 & 5 & -2 \\ -2 & -2 & 1 \\ -1 & -1 & 1 \end{pmatrix}.$$

В этом случае $\lambda_1 = \lambda_2 = \lambda_3 = 1$. Ранг матрицы $A - E$, как легко подсчитать, равен 2 и оператор $\tilde{A}$ имеет лишь один собственный вектор.

Далее:

$$E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}; \quad A - E = \begin{pmatrix} 3 & 5 & -2 \\ -2 & -3 & 1 \\ -1 & -1 & 0 \end{pmatrix};$$

$$(A-E)^2 = \begin{pmatrix} 1 & 2 & -1 \\ -1 & -2 & 1 \\ -1 & -2 & 1 \end{pmatrix}; \quad (A-E)^3 = 0.$$

Получаем сразу одну жорданову цепочку длины 3:

$$x=(1,\ -1,\ -1);\quad y=(3,\ -2,\ -1);\quad z=(1,\ 0,\ 0).$$

В этом базисе матрица оператора имеет вид

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

§ 3. ЖОРДАНОВ БАЗИС В ОБЩЕМ СЛУЧАЕ

1. Построение жорданова базиса. Переходя к общему случаю, предположим, что оператор $\tilde{A}$, не имеющий простой структуры, обладает ν различными собственными числами $\lambda_1, \lambda_2, \ldots, \lambda_\nu$. Построение жорданова базиса произведем, сведя задачу к предыдущей, когда оператор имел только одно собственное число. Для этого исключим из пространства R собственные векторы, отвечающие собственному значению λ_j кратности k_j. Это можно сделать при помощи оператора $\tilde{A}-\lambda_j\tilde{E}$. Действительно, как мы видели (гл. 5 § 2 п. 6), оператор $\tilde{A}$ в пространстве $(\tilde{A}-\lambda_j\tilde{E})R$ имеет собственное значение λ_j, но кратность его $k_j^{(1)}$ меньше кратности k_j на число $n-r_j$ линейно независимых собственных векторов оператора A, отвечающих λ_j (r_j — ранг матрицы $A-\lambda_jE$). Если в пространстве $(\tilde{A}-\lambda_j\tilde{E})R$ отсутствуют собственные векторы ($k_j^{(1)}=0$), то процесс исключения закончен. Если же $k_j^{(1)}>0$, то понизим размерность пространства, переходя к пространству $(\tilde{A}-\lambda_j\tilde{E})^2R$. Если обозначим через $k_j^{(2)}$ кратность собственного значения λ_j оператора $\tilde{A}$ в пространстве $(\tilde{A}-\lambda_j\tilde{E})^2R$, то очевидно $k_j^{(2)}<k_j^{(1)}$. При $k_j^{(2)}\neq 0$ переходим к пространству $(\tilde{A}-\lambda\tilde{E})^3R$ и т. д. При некотором $s_j>0$ получим пространство $(\tilde{A}-\lambda_j\tilde{E})^{s_j}R=\tilde{C}_jR$, в котором отсутствуют собственные векторы, отвечающие собственному значению λ_j. Показатель s_j можно определить из условия $k_j^{(s_j)}=0$, $k_j^{(s_j-1)}\neq 0$, т. е. как наименьшую степень оператора $(\tilde{A}-\lambda_j\tilde{E})^s$, аннулирующего собственные векторы, отвечающие λ_j.

По построению в пространстве $(\tilde{A}-\lambda_j\tilde{E})^{s_j}R$ отсутствуют собственные векторы, отвечающие собственному числу λ_j.

Поэтому размерности пространства $(\tilde{A}-\lambda_j\tilde{E})^{s_j}R$ и $(\tilde{A}-\lambda_j\tilde{E})^{s}R$, где $s>s_j$ совпадают. Но размерность пространства $(\tilde{A}-\lambda_j\tilde{E})^{s}R$ равна рангу матрицы $(A-\lambda_jE)^s$ (по теореме о ранге). Значит, степень s_j является наименьшим из чисел, удовлетворяющих условию $r_s=r_{s+1}$, где r_s — ранг матрицы $(A-\lambda_iE)^s=C_j$. Аналогично можно исключить из R собственные векторы, отвечающие другому собственному числу. Эти рассуждения показывают, что в пространстве

$$\tilde{C}_1\tilde{C}_2\ldots\tilde{C}_{i-1}\tilde{C}_{i+1}\ldots\tilde{C}_\nu=\tilde{D}_iR$$

отсутствуют собственные векторы с собственными значениями $\lambda_1, \lambda_2, \ldots, \lambda_{i-1}, \lambda_{i+1}, \ldots, \lambda_\nu$ и, таким образом, в пространстве $\tilde{D}_iR$ оператор A имеет единственное характеристическое число $\lambda=\lambda_i$. Согласно результатам гл. 5, § 2, п. 6 кратность характеристического числа равна k_i и совпадает с размерностью пространства $\tilde{D}_iR$ (ибо размерность $\tilde{D}_iR$ равна $n-k_1-k_2-\ldots-k_{i-1}-k_{i+1}-k_\nu=k_i$).

Переходим теперь к построению жорданова базиса в пространстве R.

Методами, предложенными в § 2, построим жорданов базис в каждом подпространстве $\tilde{D}_iR$ $(i=1, 2, \ldots \nu)$. Покажем, что совокупность всех $k_1+k_2+\ldots+k_\nu=n$ построенных векторов можно принять за базис пространства R. Для этого достаточно показать, что все построенные векторы линейно независимы. Обозначая эти векторы $e_1, e_2, \ldots, e_n$, приравняем нулю линейную комбинацию

$$\alpha_1e_1+\alpha_2e_2+\ldots+\alpha_ne_n=0 \tag{8.2}$$

в предположении, что хотя бы одно из чисел α_j отлично от нуля. Вектор e_j по построению принадлежит одному из подпространств $\tilde{D}_iR$ $(i=1, 2, \ldots \nu)$. Не ограничивая общности, будем считать, что он принадлежит пространству $\tilde{D}_jR$. В этом случае, применяя оператор $\tilde{D}_j$ к соотношению (8.2), получим равную нулю линейную комбинацию векторов пространства $\tilde{D}_jR$, которые по построению линейно независимы. Следовательно, $\alpha_j=0$, что противоречит предположению и, значит, векторы

$$e_1, e_2, \ldots, e_n$$

линейно независимы, т. е. образуют базис пространства R. Эти рассуждения указывают способ практического построения жорданова базиса. Выберем произвольный базис в пространстве R

$$e_1^*,\ e_2^*,\ \ldots,\ e_n^*.$$

Тогда базис пространства $\tilde{D}_i R$ будет составлять k_i линейно независимых векторов из

$$\tilde{D}_i e_1^*,\ \tilde{D}_i e_2^*,\ \ldots,\ \tilde{D}_i e_n^*.$$

Их координаты в базисе $\{e^*\}$ образуют столбцы матрицы D_i. Например, пусть матрица оператора A имеет вид

$$A=\begin{pmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 1 \end{pmatrix}.$$

Тогда

$$|A-\lambda E|=\begin{vmatrix} -\lambda & 1 & 0 \\ -4 & 4-\lambda & 0 \\ -2 & 1 & 1-\lambda \end{vmatrix}$$

и $\lambda_1=2$ кратности $k_1=2$,
$\lambda_2=1$ кратности $k_2=1$.

Ранг матрицы $A-\lambda_1 E=A-2E$ равен $r_1=2$ и, значит, оператор не имеет простой структуры ($n-r_1=1<k_1=2$). В данном случае операторы

$$\tilde{D}_1=(\tilde{A}-\lambda_2\tilde{E})=\tilde{C}_2,$$
$$\tilde{D}_2=(\tilde{A}-\lambda_1\tilde{E})^2=\tilde{C}_1,$$

имеют матрицы

$$D_1=\begin{pmatrix} -1 & 1 & 0 \\ -4 & 3 & 0 \\ -2 & 1 & 0 \end{pmatrix},\quad D_2=\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 2 & -1 & 1 \end{pmatrix}.$$

Пространство $\tilde{D}_2 R$ (в данном случае одномерное) состоит из собственных векторов, отвечающих $\lambda_2=1$. В данном случае любой столбец матрицы D_2 определяет собственный вектор. Положим $x_2=(0,\ 0,\ 1)$. Вектор x_2 образует базис пространства $\tilde{D}_2 R$. В пространстве $\tilde{D}_1 R$ (двумерном) оператор $\tilde{A}$ имеет одно собственное значение $\lambda_1=2$ кратности 2 и не имеет простой структуры. По-

строим жорданов базис в пространстве $\tilde{D}_1 R$. Согласно результатам § 2 пространство $(\tilde{A}-\lambda_1\tilde{E})\tilde{D}_1 R$ (одномерное) состоит из собственных векторов, отвечающих $\lambda_1=2$. Имеем

$$(A-\lambda_1 E)D_1=(A-\lambda_1 E)(A-\lambda_2 E)=\begin{pmatrix}-2 & 1 & 0\\ -4 & 2 & 0\\ 0 & 0 & 0\end{pmatrix}.$$

Положим $x_1=(1,\ 2,\ 0)$, но тогда вектор цепочки, перешедший в x_1, имеет вид

$$y_1=(1,\ 3,\ 1).$$

В итоге векторы x_1, y_1, x_2 образуют жорданов базис пространства R.

2. **Второй способ построения жорданова базиса.** Можно указать еще один способ построения жорданова базиса, основанный на использовании уравнений (8.1), которым удовлетворяют жордановы цепочки векторов. Изменяя несколько обозначения, запишем в координатной форме систему уравнений, определяющих координаты векторов одной из жордановых цепочек,

$$\begin{aligned}&(A-\lambda_i E)X_1=0,\\ &(A-\lambda_i E)X_2=X_1,\\ &\cdots\cdots\cdots\cdots\\ &(A-\lambda_i E)X_h=X_{h-1},\end{aligned}\tag{8.3}$$

где X_j — матрица-столбец из координат вектора x_j. Здесь только вектор x_1 является собственным вектором оператора $\tilde{A}$ (с собственным числом λ_1), а остальные векторы $x_2, x_3, \ldots, x_h$ принадлежат цепочке, определяемой этим собственным вектором.

Весьма существенным для применения этого метода, состоящего просто в решении системы (8.3), является то, что собственный вектор x_1, отвечающий характеристическому числу λ_i, надо выбрать таким, чтобы система (8.3) была совместна. Длина h цепочки определяется из условия, что система

$$(A-\lambda_i E)X_{h+1}=X_h$$

несовместна, если $X_1 X_2, \ldots, X_h$ — решение системы (8.3). Указанный прием построения жорданова базиса особенно

удобен тогда, когда вопрос о выборе x_1 не возникает, например, когда характеристическому числу λ_i соответствует единственный собственный вектор, т. е., когда ранг матрицы $A-\lambda_i E$ равен $n-1$.

Пример.

Пусть матрица оператора A имеет вид

$$A=\begin{pmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{pmatrix}.$$

Тогда

$$|A-\lambda E|=\begin{vmatrix} -\lambda & 1 & 0 \\ -4 & 4-\lambda & 0 \\ -2 & 1 & 2-\lambda \end{vmatrix}$$

и

$$\lambda_1=\lambda_2=\lambda_3=2.$$

Так как ранг матрицы $A-\lambda_1 E$ равен 1, то оператор не имеет простой структуры и обладает двумя собственными векторами, координаты которых x_1, x_2, x_3 удовлетворяют системе уравнений

$$\begin{cases} -2x_1+x_2=0, \\ -4x_1+2x_2=0, \\ -2x_1+x_2=0. \end{cases}$$

Эта система имеет два линейно независимых решения, например:

$$x_1=0 \qquad x_2=0 \qquad x_3=1$$

и

$$x_1=1 \qquad x_2=2 \qquad x_3=0.$$

Таким образом, линейно независимыми собственными векторами оператора A являются векторы

$$e_1=(0,\ 0,\ 1) \text{ и } e_2=(1,\ 2,\ 0).$$

Недостающий третий вектор e_3 должен удовлетворять соотношению

$$(\tilde{A}-\lambda_1\tilde{E})\,e_3=x,$$

где x — некоторая линейная комбинация собственных векторов e_1 и e_2 (т. е. тоже собственный вектор оператора A). Представляя x в виде

$$x=\gamma_1 e_1+\gamma_2 e_2=(\gamma_2,\ 2\gamma_2,\ \gamma_1),$$

получаем для определения координат y_1, y_2, y_3 вектора e_3 систему уравнений

$$\begin{aligned} -2y_1+y_2&=\gamma_2, \\ -4y_1+2y_2&=2\gamma_2, \\ -2y_1+y_2&=\gamma_1. \end{aligned}$$

Эта система совместна только, если

$$\gamma_1=\gamma_2.$$

Тогда, полагая $\gamma_1 = \gamma_2 = 1$ в качестве решения системы, можно выбрать

$$y_1 = 0, \quad y_2 = 1, \quad y_3 = 0$$

и, значит,

$$e_3 = (0, 1, 0).$$

Таким образом, в качестве жорданова базиса можно выбрать собственный вектор

$$x = e_1 + e_2 = (1, 2, 3),$$

связанный с ним вектор жордановской цепочки

$$e_3 = (0, 1, 0)$$

и любой из собственных векторов e_1 или e_2, например,

$$e_1 = (0, 0, 1).$$

Таким образом, окончательно жорданов базис (изменяя обозначения составляющих его векторов) в данном случае имеет вид

$$e_1 = (1, 2, 1)$$
$$e_2 = (0, 1, 0)$$
$$e_3 = (0, 0, 1).$$

3. Жорданова нормальная форма матрицы. Выясним, какой вид имеет матрица оператора $\tilde{A}$ в жордановом базисе.

Построение жорданова базиса для оператора $\tilde{A}$, имеющего ν различных собственных чисел $\lambda_1, \lambda_2, \ldots, \lambda_\nu$, мы свели к построению жорданова базиса для оператора, имеющего только одно собственное число. Поэтому базис будет состоять из жордановых цепочек, длины которых можно определить методами, изложенными в § 2.

Матрица оператора в жордановом базисе имеет вид

$$\begin{pmatrix} I_{n_1} & 0 & \ldots & 0 \\ 0 & I_{n_2} & \ldots & 0 \\ 0 & 0 & \ldots & I_{n_m} \end{pmatrix},$$

где I_{n_j} — жорданова клетка порядка n_j;
m — число жордановых цепочек ($m \geqslant \nu$);

$$n_1 + n_2 + \ldots + n_m = n.$$

Матрица I называется жордановой матрицей, или жордановой нормальной формой матрицы.

Так как у любого оператора $\tilde{A}$ существует жорданов базис, а в этом базисе матрица оператора является жордановой, то это значит, что любая матрица A подобна

некоторой жордановой матрице, т. е. для любой матрицы A существует переходная матрица T такая, что

$$T^{-1}AT = I.$$

Столбцами матрицы T являются координаты векторов жордановых цепочек, причем столбцы в матрице T располагаются в том же порядке, что и векторы соответствующих цепочек. Так, в приведенных ранее примерах для матрицы (из п. 6 § 2)

$$A = \begin{pmatrix} 4 & 5 & -2 \\ -2 & -2 & 1 \\ -1 & -1 & 1 \end{pmatrix}$$

матрица

$$T = \begin{pmatrix} 1 & 3 & 1 \\ -1 & -2 & 0 \\ -1 & -1 & 0 \end{pmatrix},$$

а жорданова матрица

$$I = T^{-1}AT = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

состоит из одной жордановой клетки, соответствующей характеристическому числу $\lambda = 1$.

Для матрицы (из п. 1)

$$A = \begin{pmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 1 \end{pmatrix}$$

матрица

$$T = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 3 & 0 \\ 0 & 1 & 1 \end{pmatrix},$$

а жорданова матрица

$$I = T^{-1}AT = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

состоит из двух жордановых клеток

$$I_1 = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix},$$

соответствующей $\lambda=2$ и $I_2=(1)$, соответствующей $\lambda=1$. Для матрицы (из п. 2)

$$A=\begin{pmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{pmatrix}$$

матрица

$$T=\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix},$$

а жорданова матрица

$$I=T^{-1}AT=\begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

состоит из двух жордановых клеток

$$I_2=\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$$

и

$$I_1=(2),$$

соответствующих единственному собственному числу $\lambda=2$.

Упражнения

1. Найти жорданову форму следующих матриц:

а) $\begin{pmatrix} 1 & -3 & 4 \\ 4 & -7 & 8 \\ 6 & -7 & 7 \end{pmatrix}$;

б) $\begin{pmatrix} 2 & 6 & -15 \\ 1 & 1 & -5 \\ 1 & 2 & -6 \end{pmatrix}$;

в) $\begin{pmatrix} 1 & -3 & 3 \\ -2 & -6 & 13 \\ -1 & -4 & 8 \end{pmatrix}$.

2. Привести матрицу

$$A=\begin{pmatrix} 12 & -6 & -2 \\ 18 & -9 & -3 \\ 18 & -9 & -3 \end{pmatrix}$$

к простейшему виду и найти преобразующую матрицу T.

Проверить справедливость равенства

$$I=T^{-1}AT.$$

3. Доказать:

для того, чтобы матрицы A и B были подобны, необходимо и достаточно, чтобы они были подобны одной и той же жордановой матрице (с точностью до расположения жордановых клеток).

4. Доказать:

для того, чтобы матрицы A и B были подобны, необходимо и достаточно, чтобы они имели одинаковые собственные значения одной и той же кратности и чтобы матрицы

$$(A-\lambda_i E)^s,\ (B-\lambda_i E)^s \qquad \begin{matrix} i=1,\ 2,\ \ldots,\ \mu \\ s=1,\ 2,\ \ldots,\ \mu \end{matrix}$$

имели одинаковые ранги.

5. Выяснить, используя результаты задачи 4, подобны ли следующие матрицы:

$$A=\begin{pmatrix} 4 & 6 & -15 \\ 1 & 3 & -5 \\ 1 & 2 & -4 \end{pmatrix}; \quad B=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

6. Доказать, что следующие матрицы подобны

$$A=\begin{pmatrix} 5 & -1 \\ 9 & -1 \end{pmatrix} \quad \text{и} \quad B=\begin{pmatrix} 38 & -81 \\ 16 & 34 \end{pmatrix}.$$

Пользуясь решением задачи 3, найти преобразующую матрицу T. Проверить справедливость равенства

$$A=T^{-1}BT.$$

7. Дано:

$$\begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} \qquad (\lambda \neq 0).$$

Найти жорданову форму матрицы I^2.

8. Дана матрица I порядка n

$$I=\begin{vmatrix} \lambda & 1 & 0 & \ldots & 0 \\ 0 & \lambda & 1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & & \lambda \end{vmatrix}. \text{ Найти } I^k.$$

9. Доказать, что если матрица A клеточно-диагональная

$$A=\begin{pmatrix} A_1 & & O \\ & A_2 & \\ & & \ddots \\ O & & & A_n \end{pmatrix},$$

то ее жорданова нормальная форма имеет вид

$$I=\begin{pmatrix} I_1 & & O \\ & I_2 & \\ & & \ddots \\ O & & & I_n \end{pmatrix},$$

где I_i — жорданова форма матрицы A_i.

Найти преобразующую матрицу.

Глава 9

ФУНКЦИИ ОТ МАТРИЦ

В основе определения функций от матриц будет лежать операция возведения матрицы в целую положительную степень. Так как эта операция определена только для квадратных матриц, то функции от матриц будут определены также только для квадратных матриц.

§ 1. ВОЗВЕДЕНИЕ КВАДРАТНОЙ МАТРИЦЫ В ЦЕЛУЮ ПОЛОЖИТЕЛЬНУЮ СТЕПЕНЬ

Если $A=(a_{ij})$ — квадратная матрица, то, как известно, s-й степенью матрицы A называется *матрица* A^s, получающаяся в результате s-кратного умножения матрицы A на себя, т. е.

$$A^s=\underbrace{AA\ldots A}_{s}.$$

Однако практическое вычисление A^s для произвольной матрицы, опирающееся на непосредственное s-кратное умножение матрицы A на себя, наталкивается на значительные чисто вычислительные трудности.

Существенное упрощение этой операции связано с приведением матрицы A к нормальной жордановой форме. Согласно результатам главы 8 для любой квадратной матрицы A существует невырожденная матрица T такая, что

$$T^{-1}AT=I,$$

где I — жорданова матрица. Отсюда следует, что

$$A=TIT^{-1},$$

а ее квадрат

$$A^2=A\cdot A=TIT^{-1}TIT^{-1}=TIIT^{-1}.$$

Следовательно,

$$A^2=TI^2T^{-1}.$$

Точно также находим, что

$$A^3=TI^3T^{-1}$$

и т. д., так что для любого целого положительного s

$$A^s = TI^sT^{-1}.$$

Рассмотрим возведение жордановой матрицы в степень. Пусть

$$I = \begin{pmatrix} I_{n_1} & 0 & \dots & 0 \\ 0 & I_{n_2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & & I_{n_m} \end{pmatrix}.$$

Так как эта матрица клеточно диагональная, то

$$I^s = \begin{pmatrix} I^s_{n_1} & 0 & \dots & 0 \\ 0 & I^s_{n_2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & I^s_{n_m} \end{pmatrix}.$$

(см. п. 3 § 3, гл. I).

В частности, если A — матрица простой структуры, то

$$I^s = (\lambda_i^s \delta_{ij}),$$

где δ_{ij} — символ Кронекера, причем среди λ_i могут быть и одинаковые числа.

Если же A не имеет простой структуры, то s-я степень жордановой клетки I_{n_i} определяется достаточно очевидными равенствами:

$$I^2_{n_i} = \begin{pmatrix} \lambda_i & 1 & 0 & \dots & 0 \\ 0 & \lambda_i & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & \lambda_i \end{pmatrix} \begin{pmatrix} \lambda_i & 1 & 0 & \dots & 0 \\ 0 & \lambda_i & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & \lambda_i \end{pmatrix} = \begin{pmatrix} \lambda_i^2 & 2\lambda_i & 1 & 0 & \dots & 0 \\ 0 & \lambda_i^2 & 2\lambda_i & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & \lambda_i^2 \end{pmatrix},$$

$$I^3_{n_i} = I^2_{n_i} = I_{n_i} = \begin{pmatrix} \lambda_i^3 & 3\lambda_i^2 & 3\lambda_i & 1 & 0 & \dots & 0 \\ 0 & \lambda_i^3 & 3\lambda_i^2 & 3\lambda_i & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & \dots & \lambda_i^3 \end{pmatrix}$$

и, наконец,

$$I_{n_i}^s = I_{n_i}^{s-1} I_{n_i} = \begin{pmatrix} \lambda_i^s & s\lambda_i^{s-1} & \frac{s(s-1)}{2}\lambda_i^{s-2} & \dots & 1 & 0 & \dots & 0 \\ 0 & \lambda_i^s & s\lambda_i^{s-1} & \dots & s\lambda_i & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & \lambda_i^s \end{pmatrix}.$$

Последнее легко доказывается методом индукции.

Пример.

Пусть матрица

$$A = \begin{pmatrix} 1-b & a \\ b & 1-a \end{pmatrix}$$

$$a \neq 0,\ b \neq 0 \text{ и } a+b \neq 0.$$

Тогда

$$\Delta(\lambda) = |A - \lambda E| = \begin{vmatrix} 1-b-\lambda & a \\ b & 1-a-\lambda \end{vmatrix} = (\lambda - 1)(\lambda - 1 + a + b)$$

и, следовательно,

$$\lambda_1 = 1,\ \lambda_2 = 1 - a - b \neq \lambda_1.$$

Этим характеристическим числам соответствуют линейно независимые собственные векторы

$$e_1 = (a,\ b),\ e_2 = (1,\ -1)$$

и, следовательно,

$$T = \begin{pmatrix} a & 1 \\ b & -1 \end{pmatrix}, \quad \text{а} \quad T^{-1} = \begin{pmatrix} \frac{1}{a+b} & \frac{1}{a+b} \\ \frac{b}{a+b} & -\frac{a}{a+b} \end{pmatrix}.$$

Так как матрица A имеет простую структуру, то

$$A^s = \begin{pmatrix} a & 1 \\ b & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & (1-a-b)^s \end{pmatrix} \begin{pmatrix} \frac{1}{a+b} & \frac{1}{a+b} \\ \frac{b}{a+b} & -\frac{a}{a+b} \end{pmatrix},$$

или окончательно

$$A^s = \begin{pmatrix} \frac{a+b(1-a-b)^s}{a+b} & \frac{a-a(1-a-b)^s}{a+b} \\ \frac{b-b(1-a-b)^s}{a+b} & \frac{b+a(1-a-b)^s}{a+b} \end{pmatrix}.$$

§ 2. МНОГОЧЛЕНЫ ОТ МАТРИЦЫ

Пусть A — квадратная матрица n-го порядка, а $P(t)=a_0t^s+a_1t^{s-1}+\ldots+a_s$ — многочлен скалярного аргумента t. Тогда матрицу

$$P(A)=a_0A^s+a_1A^{s-1}+\ldots+a_sE$$

называют многочленом от матрицы A (соответствующим многочлену $P(t)$).

Заменяя A на TIT^{-1} и используя результаты § 1, имеем

$$P(A)=T(a_0I^s+a_1I^{s-1}+\ldots+a_sE)T^{-1}=TP(I)T^{-1},$$

где, очевидно,

$$P(I)=\begin{pmatrix} P(I_{n_1}) & 0 & \ldots & 0 \\ 0 & P(I_{n_2}) & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & P(I_{n_m}) \end{pmatrix}.$$

В частности, если матрица A имеет простую структуру, то

$$P(A)=T\begin{pmatrix} P(\lambda_1) & 0 & \ldots & 0 \\ 0 & P(\lambda_2) & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & P(\lambda_n) \end{pmatrix}T^{-1}.$$

В общем же случае (для произвольной матрицы A) надо определить $P(I_{n_i})$. Так как

$$P(I_{n_i})=a_0\begin{pmatrix} \lambda_i^s & s\lambda_i^{s-1} & \ldots & 0 \\ 0 & \lambda_i^s & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & \lambda_i^s \end{pmatrix}+\ldots+a_s\begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & 1 \end{pmatrix}$$

то, проделав соответствующие выкладки, получим

$$P(I_{n_i})=\begin{pmatrix} P(\lambda_i), & P'(\lambda_i), & \frac{1}{2!}P''(\lambda_i) & \ldots & \frac{1}{(n_i-1)!}P^{(n_i-1)}(\lambda_i) \\ 0 & P(\lambda_i) & P'(\lambda_i) & \ldots & \frac{1}{(n_i-2)!}P^{(n_i-2)}(\lambda_i) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & P(\lambda_i) \end{pmatrix}.$$

Таким образом, элементы матрицы $P(I_{n_i})$ определяются значением многочлена $P(t)$ и его производных до (n_i-1)-го порядка включительно в точке $t=\lambda_i$, где λ_i — характеристическое число, отвечающее жордановой клетке I_{n_i}.

Отсюда непосредственно следует, что из равенства $P(A)=Q(A)$ не вытекает тождество $P(t)=Q(t)$, а лишь равенства

$$\begin{aligned} &P(\lambda_i)=Q(\lambda_i),\\ &P'(\lambda_i)=Q'(\lambda_i),\\ &\cdots\cdots\cdots\\ &P^{(n_i-1)}(\lambda_i)=Q^{(n_i-1)}(\lambda_i). \end{aligned}$$

В этом случае мы будем говорить, что многочлены $P(t)$ и $Q(t)$ совпадают на спектре матрицы A. Эти соотношения позволяют сделать заключение, что для выполнения равенства $P(A)=Q(A)$ необходимо и достаточно совпадение многочленов $P(t)$ и $Q(t)$ на спектре матрицы A.

Пример.

Пусть $P(t)=t^3+1$ и $A=\begin{pmatrix}1-b & a\\ b & 1-a\end{pmatrix}$.

Найти $P(A)$.

Решение. Используя результаты, полученные при рассмотрении примера в § 1 и учитывая, что

$$P(\lambda_1)=P(1)=2;\ P(\lambda_2)=P(1-a-b)=(1-a-b)^3+1,$$

получаем

$$P(A)=T\begin{pmatrix}2 & 0\\ 0 & (1-a-b)^3+1\end{pmatrix}T^{-1}.$$

Подставляя вместо T и T^{-1} их значения, находим

$$P(A)=\begin{pmatrix}\dfrac{a+b(1-a-b)^3}{a+b}+1 & \dfrac{a-a(1-a-b)^3}{a+b}\\ \dfrac{b-b(1-a-b)^3}{a+b} & \dfrac{b+a(1-a-b)^3}{a+b}+1\end{pmatrix}.$$

Заметим, что этот пример можно решить, непосредственно учитывая, что

$$P(A)=A^3+E,$$

и используя (для A^3) результат примера, приведенного в § 1.

§ 3. ФУНКЦИИ ОТ МАТРИЦЫ

Пусть $f(t)$ — функция скалярного аргумента t. Говорят, что эта функция определена на спектре матрицы A, если существуют

$$f(\lambda_i),\ f'(\lambda_i), \ldots, f^{(n_i-1)}(\lambda_i)\quad i=1,2,\ldots, m$$

(которые мы будем называть значениями функции $f(t)$ на спектре A) для всех характеристических чисел λ_i матрицы A, где n_i — порядок соответствующих этим характеристическим числам жордановых клеток.

В предыдущем параграфе показано, что многочлены от матрицы A совпадают, если совпадают их значения на спектре матрицы. Мы определим функцию от матрицы по такому же принципу: значения функции от матрицы $f(A)$ должны полностью определяться значениями функции $f(t)$ на спектре матрицы A.

Иными словами, по определению, если две (в общем случае различные) функции $f(t)$ и $g(t)$, определенные на спектре матрицы A, совпадают на нем, т. е.

$$f(\lambda_i)=g(\lambda_i),\ f'(\lambda_i)=g'(\lambda_i), \ldots, f^{(n_i-1)}(\lambda_i)=g^{(n_i-1)}(\lambda_i),$$

то матрицы $f(A)$ и $g(A)$ совпадают, т. е. $f(A)=g(A)$. Из этого определения следует, что если $f(t)$ — некоторая функция, а $P(t)$ — многочлен, совпадающий с ней на спектре матрицы A, то

$$f(A)=P(A).$$

Последнее соотношение можно использовать для нахождения функции от матрицы, ибо определение многочлена от матрицы было дано в § 2.

Следовательно, если функция $f(t)$ определена на спектре матрицы A, то в силу результатов § 2 функцию от матрицы A можно определять по формуле

$$f(A)=Tf(I)T^{-1},$$

где

$$A=TIT^{-1}$$

$$f(I)=\begin{pmatrix} f(I_{n_1}) & 0 & 0 \\ 0 & f(I_{n_2}) & 0 \\ \cdots & \cdots & \cdots \\ 0 & 0 & f(I_{n_m}) \end{pmatrix},$$

причем

$$f(I_{ni})=\begin{pmatrix} f(\lambda_i), & f'(\lambda_i) & \dots & \frac{1}{(n_i-1)!}f^{(n_i-1)}(\lambda_i) \\ 0 & f(\lambda_i) & \dots & \frac{1}{(n_i-2)!}f^{(n_i-2)}(\lambda_i) \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & f(\lambda_i) \end{pmatrix}.$$

В частности, если A — матрица простой структуры, то

$$f(A)=T\begin{pmatrix} f(\lambda_1) & 0 & \dots & 0 \\ 0 & f(\lambda_2) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & f(\lambda_n) \end{pmatrix}T^{-1}.$$

Пример.

Дана матрица

$$A=\begin{pmatrix} 1-b & a \\ b & 1-a \end{pmatrix}.$$

Определить e^A и $\ln A$.

Решение. Используя данное определение и результаты п. 2, получаем

$$e^A=\begin{pmatrix} a & 1 \\ b & -1 \end{pmatrix}\begin{pmatrix} e & 0 \\ 0 & e^{1-a-b} \end{pmatrix}\begin{pmatrix} \frac{1}{a+b} & \frac{1}{a+b} \\ \frac{b}{a+b} & -\frac{a}{a+b} \end{pmatrix}$$

или окончательно

$$e^A=\begin{pmatrix} \frac{ae+be^{1-a-b}}{a+b} & \frac{ae-ae^{1-a-b}}{a+b} \\ \frac{be-be^{1-a-b}}{a+b} & \frac{be+ae^{1-a-b}}{a+b} \end{pmatrix}.$$

Точно также

$$\ln A=\begin{pmatrix} a & 1 \\ b & -1 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & \ln(1-a-b) \end{pmatrix}\begin{pmatrix} \frac{1}{a+b} & \frac{1}{a+b} \\ \frac{b}{a+b} & -\frac{a}{a+b} \end{pmatrix},$$

или

$$\ln A=\begin{pmatrix} \frac{b\ln(1-a-b)}{a+b} & -\frac{a\ln(1-a-b)}{a+b} \\ -\frac{b\ln(1-a-b)}{a+b} & \frac{a\ln(1-a-b)}{a+b} \end{pmatrix}=$$

$$=\frac{\ln(1-a-b)}{a+b}\begin{pmatrix} b & -a \\ -b & a \end{pmatrix}.$$

Согласно определению функции от матрицы, если две различных функции $f(t)$ и $g(t)$ совпадают на спектре матрицы A, то $f(A)=g(A)$. Очевидно, это равенство остается справедливым, если

$$f(t)\equiv g(t).$$

Отсюда следует, что на функции от матрицы распространяются все тождества, связывающие функции скалярного аргумента (если, конечно, эти функции определены на спектре матрицы). Таким образом, справедливы, например, матричные тождества

$$\sin^2 A+\cos^2 A=E,\ e^A e^{-A}=E,\ e^{iA}=\cos A+i\sin A.$$

§ 4. СПЕЦИАЛЬНЫЕ МНОГОЧЛЕНЫ ОТ МАТРИЦЫ

Выведем два важных тождества, основанных на совпадении функций на спектре матрицы A.

1. Тождество Гамильтона — Кэли. Покажем, что всякая квадратная матрица удовлетворяет своему характеристическому уравнению, т. е., если

$$\Delta(\lambda)=|A-\lambda E|,\ \text{то}\ \Delta(A)=0.$$

Действительно, на спектре матрицы $\Delta(\lambda)=0$, так как, если λ_0 — корень кратности k, то $\Delta(\lambda_0)=\Delta'(\lambda_0)=\ldots=\Delta^{(k-1)}(\lambda_0)=0$, причем k не меньше порядка любой жордановой клетки, отвечающей корню $\lambda=\lambda_0$. Из равенства $\Delta(\lambda)=0$ на спектре матрицы следует $\Delta(A)=0$.

2. Формула Лагранжа — Сильвестра. Другим важным следствием определения функции от матрицы является формула Лагранжа — Сильвестра. Пусть $f(t)$ — некоторая функция, определенная на спектре матрицы A, тогда, если мы подберем многочлен $P(t)$, значения которого на спектре матрицы A совпадают с соответствующими значениями $f(t)$, то

$$f(A)=P(A),$$

т. е. произвольная функция от матрицы может быть представлена в виде многочлена от той же матрицы. Найдем явное выражение многочлена $P(t)$. Для простоты ограничимся случаем, когда матрица A имеем простую структуру. В этом случае $P(\lambda_i)$ должно равняться $f(\lambda_i)$

$i=1,\ldots\mu$. Легко проверить, что в данном случае

$$P(t)=\sum_{i=1}^{\mu}\frac{(t-\lambda_1)\ldots(t-\lambda_{i-1})(t-\lambda_{i+1})\ldots(t-\lambda_\mu)}{(\lambda_i-\lambda_1)\ldots(\lambda_i-\lambda_{i-1})(\lambda_i-\lambda_{i+1})\ldots(\lambda_i-\lambda_\mu)}f(\lambda_i),$$

где $\lambda_1, \lambda_2, \ldots, \lambda_\mu$ — все различные характеристические числа матрицы A. Действительно из последней формулы следует $P(\lambda_i)=f(\lambda_i)$ $i=1,2,\ldots,\mu$, т. е. значения функции $f(t)$ совпадают со значениями $P(t)$ на спектре матрицы A, а это означает, что

$$f(A)=P(A)=$$

$$=\sum_{i=1}^{\mu}\frac{(A-\lambda_1E)\ldots(A-\lambda_{i-1}E)(A-\lambda_{i+1}E)\ldots(A-\lambda_\mu E)}{(\lambda_i-\lambda_1)\ldots(\lambda_i-\lambda_{i-1})(\lambda_i-\lambda_{i+1})\ldots(\lambda_i-\lambda_\mu)}f(\lambda_i).$$

Эта формула позволяет находить $f(A)$, не находя матрицы T, преобразующей матрицу A к жордановой форме I. Многочлен $P(A)$ называется *интерполяционным многочленом Лагранжа — Сильвестра.*

Например:

$$A=\begin{pmatrix}1-b & a\\ b & 1-a\end{pmatrix};\ f(t)=e^t;$$

$$e^A=\frac{A-\lambda_2E}{\lambda_1-\lambda_2}e^{\lambda_1}+\frac{A-\lambda_1E}{\lambda_2-\lambda_1}e^{\lambda_2},$$

где

$$\lambda_1=1;$$

$$\lambda_2=1-a-b,$$

т. е.

$$e^A=\begin{pmatrix}a & a\\ b & b\end{pmatrix}\frac{e}{a+b}-\begin{pmatrix}-b & a\\ b & -a\end{pmatrix}\frac{e^{1-a-b}}{a+b},$$

или

$$e^A=\begin{pmatrix}\dfrac{ae+be^{1-a-b}}{a+b} & \dfrac{ae-ae^{1-a-b}}{a+b}\\ \dfrac{be-be^{1-a-b}}{a+b} & \dfrac{be+ae^{1-a-b}}{a+b}\end{pmatrix}.$$

Отметим, что в общем случае (когда матрица A не имеет простой структуры) многочлен $P(t)$ нужно выбирать так, чтобы, кроме равенств $f(\lambda_i)=P(\lambda_i)$, выполнялись соотношения

$$f'(\lambda_i)=P'(\lambda_i)\ldots f^{(n_i-1)}(\lambda_i)=P^{(n_i-1)}(\lambda_i).$$

В этом случае формула, выражающая $P(t)$, будет сложнее, однако, если известны значения функции $f(t)$ на спектре A, в каждом примере можно определить $P(t)$, например, используя метод неопределенных коэффициентов или формулу Тейлора.

Пример.

Имеем

$$A=\begin{pmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{pmatrix}. \text{ Здесь } \lambda_1=\lambda_2=\lambda_3=2.$$

Ранг матрицы $A-2E$, $r=1$, следовательно, матрица A имеет два линейно независимых собственных вектора и одну жорданову цепочку.

Для нахождения $f(A)$ достаточно подобрать многочлен $P(t)$, удовлетворяющий условиям

$$P(2)=f(2), \quad P'(2)=f'(2),$$

т. е.

$$P(t)=f(2)+f'(2)(t-2),$$
$$f(A)=f(2)E+f'(2)(A-2E),$$

в частности, если $f(t)=e^{\alpha t}$, то

$$e^{\alpha A}=e^{2\alpha}E+\alpha e^{2\alpha}(A-2E)=\begin{pmatrix} (1-2\alpha)e^{2\alpha} & \alpha e^{2\alpha} & 0 \\ -4\alpha e^{2\alpha} & (2\alpha+1)e^{2\alpha} & 0 \\ -2\alpha e^{2\alpha} & \alpha e^{2\alpha} & e^{2\alpha} \end{pmatrix}$$

Если $f(t)=\frac{1}{t}$, то $f(A)=A^{-1}$ — обратная матрица.

В данном примере

$$A^{-1}=\frac{1}{2}E-\frac{1}{4}(A-2E)=E-\frac{1}{4}A=\begin{pmatrix} 1 & -\frac{1}{4} & 0 \\ 1 & 0 & 0 \\ \frac{1}{2} & -\frac{1}{4} & \frac{1}{2} \end{pmatrix},$$

в чем легко убедиться непосредственно.

В заключение отметим, что представление функции от матрицы многочленом не однозначно. Степень и коэффициенты этого многочлена зависят не только от вида функции $f(t)$, но и от спектра матрицы A.

Так, из приведенных выше примеров следует, что функция $f(A)=e^A$ для матрицы $A=\begin{pmatrix} 1-b & a \\ b & 1-a \end{pmatrix}$ имеет вид

$$e^A=P_1(A)=\frac{e}{a+b}\left[(1-e^{-(a+b)})A-(1-a-b-e^{-(a+b)})E\right],$$

а для матрицы

$$A=\begin{pmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{pmatrix}$$

$$e^A = P_2(A) = e^2(A-E).$$

Причем многочлены $P_1(t)$ и $P_2(t)$ различны при всех значениях параметров a и b.

§ 5. РАЗЛОЖЕНИЕ ФУНКЦИИ ОТ МАТРИЦЫ В РЯД

Пусть функция $f(t)$ и последовательность функций $f_1(t), f_2(t), \ldots, f_s(t), \ldots$ определены на спектре матрицы A.

Функцию $f(t)$ будем называть *пределом последовательности* $\{f_s(t)\}$ *при* $s \to \infty$ *на спектре матрицы* A, если

$$\lim_{s\to\infty} f_s(\lambda_i) = f(\lambda_i),$$

$$\lim_{s\to\infty} f'_s(\lambda_i) = f'(\lambda_i), \ldots, \lim_{s\to\infty} f_s^{(n_i-1)}(\lambda_i) = f^{(n_i-1)}(\lambda_i)$$

для всех характеристических чисел λ_i *матрицы* A (напомним, что n_i — порядок жордановой клетки).

Очевидно, что приведенное определение эквивалентно условию

$$\lim_{s\to\infty} f_s(I) = f(I),$$

где I — жорданова матрица, подобная матрице A.

Тогда из соотношения (см. § 3)

$$f_s(A) = Tf_s(I)\,T^{-1}$$

(где T — постоянная матрица) следует, *что для того, чтобы последовательность матриц* $\{f_s(A)\}$ *стремилась при* $s \to \infty$ *к пределу* $f(A)$, *необходимо и достаточно, чтобы последовательность функций* $\{f_s(t)\}$ *стремилась к функции* $f(t)$ *на спектре матрицы* A.

Введем теперь ряд, составленный из матриц $u_s(A)$, каждая из которых является некоторой функцией матрицы A. Обозначим этот ряд, как обычно через

$$\sum_{s=1}^{\infty} u_s(A).$$

Тогда, если существует предел последовательности частичных сумм

$$\sigma_s(A) = \sum_{i=1}^{s} u_i(A)$$

этого ряда при $s \to \infty$, то ряд называется *сходящимся*, а его суммой является матрица

$$\sigma(A) = \lim_{s \to \infty} \sigma_s(A).$$

В силу приведенной выше теоремы для сходимости ряда

$$\sum_{s=1}^{\infty} u_s(A)$$

к матрице $\sigma(A)$ необходимо и достаточно, чтобы функциональный ряд

$$\sum_{s=1}^{\infty} u_s(t)$$

сходился на спектре матрицы A к функции $\sigma(t)$. Отсюда следует, что если функцию $f(t)$ разложить в ряд Тейлора

$$f(t) = \sum_{s=0}^{\infty} a_s (t-a)^s,$$

который сходится в круге $|t-a| < R$, то

$$f(A) = \sum_{s=0}^{\infty} a_s (A - aE)^s \tag{9.1}$$

при условии, что все характеристические числа λ_i матрицы A лежат внутри круга сходимости, т. е.

$$|\lambda_i - a| < R \qquad i = 1, 2, \ldots \mu. \tag{9.2}$$

Примеры.

1. Так как

$$e^t = \sum_{s=0}^{\infty} \frac{t^s}{s!}, \quad \cos t = \sum_{s=0}^{\infty} \frac{(-1)^s t^{2s}}{(2s)!}, \quad \sin t = \sum_{s=0}^{\infty} \frac{(-1)^s t^{2s+1}}{(2s+1)!}$$

при всех t $(R = \infty)$,

то

$$e^A = \sum_{s=0}^{\infty} \frac{A^s}{s!},$$

$$\cos A = \sum_{s=0}^{\infty} \frac{(-1)^s}{(2s)!} A^{2s},$$

$$\sin A = \sum_{s=0}^{\infty} \frac{(-1)^s}{(2s+1)!} A^{2s+1}$$

для любой квадратной матрицы A.

2. Так как

$$\ln(1+t) = \sum_{s=1}^{\infty} \frac{(-1)^{s-1}}{s} t^s$$

при $|t| < 1$,
то

$$\ln(E+A) = \sum_{s=1}^{\infty} \frac{(-1)^{s-1}}{s} A^s$$

для таких матриц A, все характеристические числа которых удовлетворяют условию

$$|\lambda_i| < 1 \qquad i = 1, 2, \ldots \mu.$$

Представление функции от матрицы $f(A)$ степенным рядом (9.1) отличается от ее представления многочленом (см. § 4) тем, что коэффициенты a_s ряда при выполнении условия (9.2) не зависят от характера матрицы A и полностью определяются видом функции $f(t)$. В частности, если $f(t)$ раскладывается в степенной ряд (по степеням t или $t-a$), сходящийся при всех t, то функция $f(A)$ может быть представлена степенным рядом (по степеням A или $A-aE$) одним и тем же для всех матриц A (что имеет место, например, для функций e^A, $\cos A$, $\sin A$ и др.).

Наряду с функциями от матриц в приложениях часто используются так называемые функциональные матрицы. Введем эти матрицы и рассмотрим некоторые операции над ними.

§ 6. ФУНКЦИОНАЛЬНЫЕ МАТРИЦЫ

Матрица $A(t)$ размеров $(m \times n)$ называется *функциональной*, если ее элементы являются функциями переменной t (или нескольких переменных), т. е.

$$A(t) = (a_{ij}(t)).$$

Пределом матрицы $A(t)=(a_{ij}(t))$ при $t\to t_0$ называется *матрица* $B=(b_{ij})$ *тех же размеров, каждый элемент которой*

$$b_{ij}=\lim_{t\to t_0} a_{ij}(t),$$

что обозначается так

$$B=\lim_{t\to t_0} A(t).$$

Из этого определения следует, что если $a_{ij}(t)$ непрерывны при $t=t_0$, то

$$\lim_{t\to t_0} A(t)=A(t_0).$$

Производная, неопределенный и определенный интегралы функциональной матрицы $A(t)$ по переменной t определяются соотношениями

$$A'(t)=\frac{dA}{dt}=(a'_{ij}(t))=\left(\frac{da_{ij}}{dt}\right),$$

$$\int A(t)\,dt=\left(\int a_{ij}(t)\,dt\right),$$

$$\int_\alpha^\beta A(t)\,dt=\left(\int_\alpha^\beta a_{ij}(t)\,dt\right).$$

Например:

$$A(t)=\begin{pmatrix} t^2 & \sin t \\ 1 & 2t \end{pmatrix},$$

$$A'(t)=\begin{pmatrix} 2t & \cos t \\ 0 & 2 \end{pmatrix}.$$

Заметим, что при дифференцировании по переменной t функции от матрицы, содержащей t, часто удобно пользоваться разложением ее в ряд. Например, если A — постоянная матрица (не зависящая от t), то

$$[e^{At}]'_t=\left[\sum_{s=0}^{\infty}\frac{A^s t^s}{s!}\right]'_t=\sum_{s=0}^{\infty}\frac{A^s s t^{s-1}}{s!}=A\sum_{s=1}^{\infty}\frac{A^{s-1}t^{s-1}}{(s-1)!}=Ae^{At},$$

т. е.

$$[e^{At}]'_t=Ae^{At}.$$

Точно также

$$[\cos At]'_t=\left[\sum_{s=0}^{\infty}\frac{(-1)^s A^{2s}t^{2s}}{(2s)!}\right]'_t=$$

$$=-A\sum_{s=1}^{\infty}\frac{(-1)^{s-1}A^{2s-1}t^{2s-1}}{(2s-1)!}=-A\sin At.$$

Упражнения

1. Используя тождество

$$\begin{pmatrix} 17 & -6 \\ 35 & -12 \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 5 & 7 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} -7 & 3 \\ 5 & -2 \end{pmatrix},$$

найти

$$\begin{pmatrix} 17 & -6 \\ 35 & -12 \end{pmatrix}^5.$$

2. Найти

$$\begin{pmatrix} 43 & -3 \\ 23 & -2 \\ 44 & -3 \end{pmatrix}^6.$$

3. Найти $f(A)$, если $A = \begin{pmatrix} 5 & 2 & -3 \\ 1 & 3 & -1 \\ 2 & 2 & -1 \end{pmatrix}$ и

$$f(x) = x^3 - 7x^2 + 15x - 5.$$

4. Проверить, что матрица $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ удовлетворяет уравнению

$$x^2 - (a + d)x + ad - bc = 0.$$

5. Дано:

$$A = \begin{pmatrix} 4 & 3 & -3 \\ 2 & 3 & -2 \\ 4 & 4 & -3 \end{pmatrix}.$$

Найти
а) e^A;
б) $\ln A$;
в) $\sin Ax$.

6. Доказать тождество $\sin 2A = 2 \sin A \cos A$, где A — любая квадратная матрица.

7. Доказать, что если матрица A — клеточно диагональная

$$A = \begin{pmatrix} A_1 & & 0 \\ & A_2 & \\ 0 & & \ddots A_n \end{pmatrix}$$

и функция $f(\lambda)$ определена на спектре матрицы A, то

$$f(A) = \begin{pmatrix} f(A_1) & & 0 \\ & f(A_2) & \\ 0 & & \ddots f(A_n) \end{pmatrix},$$

Глава 10

СИСТЕМЫ ЛИНЕЙНЫХ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ

В теории линейных дифференциальных уравнений применение матричного исчисления приводит к существенному сокращению и упрощению многих доказательств, связанных с исследованием систем линейных дифференциальных уравнений. Особенно эффективным оказывается использование функций от матриц при исследовании систем с постоянными коэффициентами. В настоящей главе рассматриваются некоторые основные вопросы, связанные с применением теории матриц к исследованию и решению систем линейных дифференциальных уравнений.

§ 1. СИСТЕМА ЛИНЕЙНЫХ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ

1. Матричная запись системы линейных дифференциальных уравнений. Как известно из теории дифференциальных уравнений, любую систему (обыкновенных) линейных дифференциальных уравнений (в том числе и одно уравнение n-го порядка) можно привести к системе вида

$$\left.\begin{aligned}
\frac{dy_1}{dx} &= p_{11}(x)\,y_1 + \ldots + p_{1n}(x)\,y_n + f_1(x),\\
\frac{dy_2}{dx} &= p_{21}(x)\,y_1 + \ldots + p_{2n}(x)\,y_n + f_2(x),\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
\frac{dy_n}{dx} &= p_{n1}(x)\,y_1 + \ldots + p_{nn}(x)\,y_n + f_n(x),
\end{aligned}\right\} \qquad (10.1)$$

где $p_{ij}(x)$ и $f_i(x)$ — заданные функции независимой переменной x. Функции $p_{ij}(x)$ называются *коэффициентами*, а $f_i(x)$ *свободными членами системы* (10.1).

Если все свободные члены

$$f_i(x) = 0,$$

то система называется *однородной*.

Решением системы (10.1) называется всякая совокупность из n функций $y_1(x), y_2(x), \ldots, y_n(x)$, подстановка которых в уравнения (10.1) обращает их в тождества.

Будем говорить, что решение системы удовлетворяет начальным условиям

$$x_0,\ y_{10},\ y_{20},\ \ldots\ y_{n0}$$

(которые называются *условиями Коши*), если

$$y_1(x_0)=y_{10},\ y_2(x_0)=y_{20},\ \ldots,\ y_n(x_0)=y_{n0}. \qquad (10.2)$$

Задача об отыскании решения системы (10.1), удовлетворяющего заданным начальным условиям Коши (10.2), называется *задачей* Коши.

Можно доказать*, что если коэффициенты $p_{ij}(x)$ и свободные члены $f_i(x)$ непрерывны на некотором интервале (a, b), то на нем существует и притом единственное решение задачи Коши, удовлетворяющее любым начальным условиям $y_{10}, y_{20}, \ldots y_{n0}$ при $a < x_0 < b$.

Если ввести матрицу коэффициентов

$$P(x)=\begin{pmatrix} p_{11}(x) & p_{12}(x) \ldots & p_{1n}(x) \\ p_{21}(x) & p_{22}(x) \ldots & p_{2n}(x) \\ \cdots & \cdots & \cdots \\ p_{n1}(x) & p_{n2}(x) \ldots & p_{nn}(x) \end{pmatrix}$$

и матрицы-столбцы

$$Y=\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \quad \text{и} \quad F(x)=\begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_n(x) \end{pmatrix},$$

то систему линейных дифференциальных уравнений можно записать в матричной форме

$$\frac{dY}{dx}=P(x)\,Y+F(x). \qquad (10.3)$$

Начальные условия (10.2) при этом примут вид

$$Y\,|_{x=x_0}=Y(x_0)=Y_0, \qquad (10.4)$$

где

$$Y_0=\begin{pmatrix} y_{10} \\ y_{20} \\ \vdots \\ y_{n0} \end{pmatrix}.$$

---

* Доказательство этой теоремы, называемой теоремой существования и единственности решения задачи Коши, читатель найдет в общих курсах теории дифференциальных уравнений.

2. Пространство решений однородной системы. Рассмотрим однородную систему линейных дифференциальных уравнений. Так как в этом случае $F(x)=0$, то согласно (10.3) она имеет вид

$$\frac{dY}{dx}=P(x)Y. \tag{10.5}$$

Покажем, что множество всех решений системы (10.5) является линейным пространством. В самом деле: пусть $Y_1(x)$ и $Y_2(x)$ — какие-либо решения системы (10.5), т. е.

$$\frac{dY_1}{dx}=P(x)Y_1, \quad \frac{dY_2}{dx}=P(x)Y_2.$$

Назовем суммой решений Y_1 и Y_2 матрицу

$$Z=Y_1+Y_2.$$

Так как

$$\frac{dZ}{dx}=\frac{dY_1}{dx}+\frac{dY_2}{dx}=P(x)Y_1+P(x)Y_2=P(x)Z,$$

то сумма решений есть решение системы (10.5).

Произведением решения Y системы (10.5) на число λ назовем матрицу

$$Z=\lambda Y$$

и так как

$$\frac{dZ}{dx}=\lambda\frac{dY}{dx}=\lambda P(x)Y=P(x)\lambda Y=P(x)Z,$$

то и произведение решения Y на число λ также есть решение системы (10.5).

Учитывая, что для суммы матриц и произведения матрицы на число выполняются все аксиомы линейного пространства, приходим к выводу, что множество всех решений однородной системы линейных уравнений образует линейное пространство. Для определения размерности этого пространства необходимо исследовать некоторые дополнительные свойства решений системы (10.5).

3. Определитель Вронского. Пусть

$$Y_1=\begin{pmatrix} y_{11}(x) \\ y_{21}(x) \\ \vdots \\ y_{n1}(x) \end{pmatrix}, \ldots, Y_n=\begin{pmatrix} y_{1n}(x) \\ y_{2n}(x) \\ \vdots \\ y_{nn}(x) \end{pmatrix}$$

произвольные n решений системы (10.5). Матрица

$$W=(Y_1Y_2, \ldots Y_n)=(y_{ij}(x))$$

n-го порядка называется *интегральной матрицей*. Определитель $|W|$ интегральной матрицы называется *определителем Вронского* системы решений.

Так как (см. введение п. 6)

$$\frac{d|W|}{dx}=\begin{vmatrix} y'_{11}y'_{12}\ldots y'_{1n} \\ y_{21}y_{22}\ldots y_{2n} \\ \cdots\cdots \\ \cdots\cdots \\ y_{n1}y_{n2}\ldots y_{nn} \end{vmatrix}+\ldots+\begin{vmatrix} y_{11}y_{12}\ldots y_{1n} \\ y_{21}y_{22}\ldots y_{2n} \\ \cdots\cdots \\ \cdots\cdots \\ y'_{n1}y'_{n2}\ldots y'_{nn} \end{vmatrix},$$

то, подставляя сюда y'_{ij} из системы уравнений (10.1) (при $f_{ij}=0$),

$$y'_{ij}=\sum_{k=1}^{n} p_{ik}(x)\, y_{kj},$$

и учитывая, что при прибавлении к элементам строки определителя соответствующих элементов другой строки, умноженных на один и тот же множитель, определитель не изменяется, после очевидных преобразований получаем

$$\frac{d|W|}{dx}=(p_{11}(x)+p_{22}(x)+\ldots+p_{nn}(x))\,|W|,$$

откуда следует, что

$$|W|=Ce^{\int_{x_0}^{x}[p_{11}(x)+p_{22}(x)+\ldots+p_{nn}(x)]\,dx}.$$

Последнее соотношение называется *тождеством Якоби*.

Из тождества Якоби следует, что определитель Вронского $|W|$ либо тождественно равен нулю ($C=0$), либо отличен от нуля ($C\neq 0$) при всех значениях x из интервала, в котором $p_{ij}(x)$ — непрерывны.

Если $|W|\neq 0$, то матрица W называется *фундаментальной*, а матрицы

$$Y_1,\ Y_2, \ldots Y_n$$

называются фундаментальной *системой решений*.

4. Базис пространства решений. Покажем, что если $Y_1, Y_2, \ldots Y_n$ — фундаментальная система решений, то они

линейно независимы. Для этого предположим обратное, т. е., что существуют постоянные $c_1, c_2, \ldots c_n$, не все равные нулю, такие, что

$$c_1Y_1 + c_2Y_2 + \ldots + c_nY_n = 0.$$

Вводя матрицу

$$C = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix},$$

запишем последнее равенство в матричной форме

$$WC = 0$$

и, так как по условию $\det W \neq 0$, то $C = 0$, что противоречит предположению.

Для того, чтобы показать, что фундаментальная система образует базис пространства решений, достаточно показать, что любое решение $Z(x)$ системы (10.5) может быть представлено в виде линейной комбинации решений фундаментальной системы (разложено по базису).

Рассмотрим систему уравнений (относительно $c_1, c_2 \ldots c_n$)

$$c_1Y_1(x_0) + c_2Y_2(x_0) + \ldots + c_nY_n(x_0) = Z(x_0), \quad (10.6)$$

или в матричной форме

$$W(x_0)\,C = Z(x_0),$$

где x_0 — произвольная точка интервала (a, b).
Так как $\det W(x_0) \neq 0$, то система (10.6) имеет единственное решение

$$C^* = \begin{pmatrix} c_1^* \\ c_2^* \\ \vdots \\ c_n^* \end{pmatrix}.$$

Но тогда решение $Y(x) = c_1^*Y_1(x) + c_2^*Y_2(x) + \ldots + c_n^*Y_n(x)$ совпадает с решением $Z(x)$ в силу единственности решения задачи Коши, так как в силу (10.6) $Z(x_0) = Y(x_0)$ и значит решения $Z(x)$ и $Y(x)$ удовлетворяют одним и тем же начальным условиям.

Таким образом, произвольное решение $Z(x)$ можно представить в виде $Z(x) = c_1Y_1(x) + c_2Y_2(x) + \ldots + c_nY_n(x)$. Следовательно, $Y_1(x)$, $Y_2(x), \ldots, Y_n(x)$ — базис пространства решений. Отсюда следует утверждение.

Теорема. *Любая фундаментальная система решений $Y_1, Y_2, \ldots Y_n$ образует базис пространства решений.*

Ясно, что размерность этого пространства равна n. Из доказанной теоремы непосредственно следует, что общее решение системы (5) имеет вид

$$Y = c_1Y_1 + c_2Y_2 + \ldots + c_nY_n = W(x)\,C,$$

где W — фундаментальная матрица. Таким образом, для решения системы (10.5) достаточно найти ее любые n линейно независимых частных решений.

5. Неоднородная система. Перейдем к исследованию неоднородной системы линейных дифференциальных уравнений

$$\frac{dY}{dx} = P(x)\,Y + F(x). \qquad (10.7)$$

Для ее решения применим метод *вариации произвольных постоянных*, который состоит в следующем. Сначала решается однородная система

$$\frac{dZ}{dx} = P(x)\,Z$$

[соответствующая системе (10.7)], общее решение которой имеет вид

$$Z = W(x)\,C,$$

где $W(x)$ — фундаментальная матрица; C — постоянная матрица.

После этого решение неоднородной системы (10.7) ищем в виде

$$Y = W(x)\,C(x),$$

где

$$C(x) = \begin{pmatrix} c_1(x) \\ c_2(x) \\ \vdots \\ c_n(x) \end{pmatrix}$$ — неизвестная функциональная матрица.

Тогда

$$\frac{dY}{dx} = \frac{dW}{dx}\,C(x) + W\,\frac{dC(x)}{dx}$$

и, следовательно,

$$\frac{dW}{dx}C+W\frac{dC}{dx}=PWC+F.$$

Но так как

$$\frac{dW}{dx}=PW,$$

то первые члены левой и правой частей уничтожаются и мы приходим к равенству

$$W\frac{dC}{dx}=F.$$

Так как $|W|\neq 0$, то матрица W имеет обратную и, значит,

$$\frac{dC}{dx}=W^{-1}F.$$

Откуда следует, что

$$C(x)=\int_{x_0}^{x} W^{-1}(x)F(x)\,dx+C_0,$$

где

$$C_0=\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$ — произвольная постоянная матрица.

Таким образом, решение неоднородной системы (10.7) имеет вид

$$Y=W(x)C_0+W(x)\int_{x_0}^{x} W^{-1}(x)F(x)\,dx. \qquad (10.8)$$

Так как при $C_0=0$

$$Y=Y^*=W(x)\int_{x_0}^{x} W^{-1}(x)F(x)\,dx,$$

то это значит, что Y^* является частным решением неоднородной системы и, следовательно, решение (10.8) состоит из суммы некоторого частного решения Y^* неоднородной системы и общего решения $Z=WC_0$ соответствующей однородной системы. Решение (10.8) является общим, так как если $Z(x)$ — произвольное решение системы (10.7), то $Z(x)-Y^*(x)$, очевидно, удовлетворяет однородной си-

стеме (10.5) и, следовательно, имеет вид

$$Z(x) - Y^*(x) = W(x)\,C^*.$$

Значит, $Z(x)$ получится из (10.8) при $C_0 = C^*$.

6. Линейное уравнение n-го порядка. *Линейным дифференциальным уравнением* n-го порядка называется уравнение

$$y^{(n)} + p_1(x)\,y^{(n-1)} + \ldots + p_n(x)\,y = f(x), \qquad (10.9)$$

линейное относительно неизвестной функции и всех ее производных до n-го порядка включительно. Функция $f(x)$ — называется *правой частью уравнения*, а функции $p_i(x)$ $(i = 1, 2, \ldots n)$ — его *коэффициентами*. Если $f(x) = 0$, то уравнение называется *однородным*, если $f(x) \neq 0$, то — *неоднородным*.

Начальными условиями типа Коши уравнения (10.9) называются условия

$$\text{при } x = x_0, \quad y = y_0 \quad y' = y_0', \ldots y^{(n-1)} = y_0^{(n-1)}.$$

Всякая функция $y(x)$, подстановка которой в уравнение (10.9) обращает его в тождество, называется *решением уравнения*. Задача об отыскании решения, удовлетворяющего заданным начальным условиям типа Коши, называется *задачей Коши*.

Если коэффициенты $p_i(x)$ и правая часть $f(x)$ непрерывны на интервале (a, b), то задача Коши имеет и при том единственное решение, удовлетворяющее любым начальным условиям (типа Коши) при x_0 из интервала (a, b). Как уже упоминалось в п. 1, линейное уравнение n-го порядка сводится к системе линейных уравнений типа (10.1). Действительно, обозначим

$$y = y_1 \quad y' = y_2 \quad y'' = y_3, \ldots y^{(n-1)} = y_n.$$

В этом случае

$$\begin{aligned} y_1' &= y_2, \\ y_2' &= y_3, \\ &\cdots\cdots \\ &\cdots\cdots \\ y_{n-1}' &= y_n, \\ y_n' &= -p_n(x)\,y_1 - p_{n-1}(x)\,y_2 - \ldots - p_1(x)\,y_n + f(x). \end{aligned}$$

Мы получаем систему линейных дифференциальных уравнений, которая в матричной форме имеет вид

$$\frac{dY}{dx} = P(x)Y + F(x), \tag{10.10}$$

где

$$P(x) = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & 0 & 1 \\ -p_n & -p_{n-1} & \dots & & & -p_1 \end{pmatrix}; \quad F(x) = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ f(x) \end{pmatrix}.$$

Если $y(x)$ — решение уравнения (10.9), то

$$Y(x) = \begin{pmatrix} y(x) \\ y'(x) \\ \vdots \\ y^{(n-1)}(x) \end{pmatrix} \text{ — решение системы (10.10)}$$

и, наоборот, если $Y(x) = \begin{pmatrix} y_1(x) \\ y_2(x) \\ \vdots \\ y_n(x) \end{pmatrix}$ — решение системы (10.10), то $y = y_1(x)$ — решение уравнения (10.9). Это обстоятельство позволяет применить результаты теории системы линейных дифференциальных уравнений к исследованию уравнения (10.9). Вначале рассмотрим однородное уравнение $f(x) = 0$

$$y^{(n)} + p_1(x)\, y^{(n-1)} + \dots + p_n(x)\, y = 0. \tag{10.11}$$

Его общее решение будет первым элементом матрицы общего решения системы (10.10)

$$Y = W \cdot C = c_1 Y_1 + c_2 Y_2 + \dots + c_n Y_n$$

и поэтому имеет вид

$$y = c_1 y_1(x) + c_2 y_2(x) + \dots + c_n y_n(x),$$

причем функции $y_1(x)$, $y_2(x) \dots y_n(x)$, очевидно, линейно независимы, так как в противном случае все их производные также были бы линейно зависимыми с теми же коэффициентами. Последнее означает, что решения си-

стемы (10.10)

$$Y_1=\begin{pmatrix} y_1 \\ y_1' \\ \vdots \\ y_1^{(n-1)} \end{pmatrix}, \ldots, Y_n=\begin{pmatrix} y_n \\ y_n' \\ \vdots \\ y_n^{(n-1)} \end{pmatrix}$$

линейно зависимы, что противоречит фундаментальности системы $Y_1, Y_2, \ldots, Y_n$.

Таким образом, общее решение уравнения (10.9) представляет собой линейную комбинацию из его n линейно независимых частных решений. Заметим, что тождество Якоби для определителя Вронского решений уравнения (10.10) в данном случае имеет вид

$$|W|=ce^{-\int\limits_{x_0}^{x} p_1(x)\,dx}, \quad \text{где} \quad W=\begin{pmatrix} y_1 & y_2 & \ldots & y_n \\ y_1' & y_2' & \ldots & y_n' \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ y_1^{(n-1)} & y_2^{(n-1)} & \ldots & y_n^{(n-1)} \end{pmatrix}.$$

Если $f(x)\neq 0$, то очевидно (см. 10.8), что общее решение неоднородного уравнения является первым элементом матрицы:

$$Y=W(x)\,C+W(x)\int\limits_{x_0}^{x} W^{-1}(x)\,F(x)\,dx$$

и может быть записано в виде

$$y=c_1z_1(x)+\ldots+c_nz_n(x)+z_0(x).$$

Оно представляет собой сумму частного решения

$$y=z_0(x)$$

уравнения (10.9) и общего решения соответствующего ему однородного уравнения (10.11).

§ 2. ЛИНЕЙНЫЕ СИСТЕМЫ С ПОСТОЯННЫМИ КОЭФФИЦИЕНТАМИ

1. Фундаментальная матрица. В предыдущих параграфах мы выяснили, какова структура общего решения линейной системы дифференциальных уравнений. Главную трудность при ее решении составляет нахождение

фундаментальной системы решений. В общем виде задача нахождения фундаментальной системы решений решена лишь для некоторых частных случаев линейных систем. Мы рассмотрим один из них, а именно: случай, когда матрица системы $P(x)$ постоянна. В этом случае система (10.3) имеет вид

$$\frac{dY}{dx} = AY + F(x). \tag{10.12}$$

Вначале решим однородную систему

$$\frac{dY}{dx} = AY. \tag{10.13}$$

Покажем, что в качестве фундаментальной матрицы этой системы можно выбрать матрицу

$$W(x) = e^{Ax}.$$

Действительно, в силу результатов (§ 6, гл. 9)

$$\frac{dW}{dx} = (e^{Ax})'_x = Ae^{Ax} = AW$$

и, следовательно, $W(x)$ — интегральная матрица. Кроме того, $W(x)$ — невырожденная матрица, так как, в силу тождества

$$e^{Ax}e^{-Ax} = E, \quad \det e^{Ax} \neq 0$$

и, следовательно, $W(x)$ — фундаментальная матрица. Таким образом, общее решение системы (10.13) имеет вид

$$Y = WC = e^{Ax}C.$$

Запишем это решение подробнее.

Согласно определению

$$e^{Ax} = Te^{Jx}T^{-1},$$

причем

$$e^{Jx} = \begin{pmatrix} e^{J_{n_1}x} & 0 & 0 \\ 0 & e^{J_{n_2}x} & 0 \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ 0 & 0 & e^{J_{n_m}x} \end{pmatrix},$$

а

$$e^{I_{n_i}x}=\begin{pmatrix} e^{\lambda_i x} & xe^{\lambda_i x} & \dots & \frac{x^{n_i-1}}{(n-1)!}e^{\lambda_i x} \\ 0 & e^{\lambda_i x} & \dots & \frac{x^{n_i-2}}{(n_i-2)!}e^{\lambda_i x} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & e^{\lambda_i x} \end{pmatrix}.$$

Таким образом, общее решение системы (10.13) представляет собой линейную комбинацию функций вида

$$x^s e^{\lambda x},$$

где λ — характеристическое число матрицы A, а s — неотрицательное целое число, причем жордановой клетке порядка k, соответствующей характеристическому числу λ в общем решении отвечает выражение вида

$$(c_1x^{k-1}+c_2x^{k-2}+\dots+c_k)\,e^{\lambda x}.$$

При этом следует иметь в виду, что если среди характеристических чисел λ найдутся комплексные

$$\lambda=\alpha+\beta i,$$

то решение будет (в силу формулы Эйлера $e^{it}=\cos t+$ $+i\sin t$) содержать функции вида

$$x^s e^{\alpha x}\cos\beta x \quad \text{и} \quad x^s e^{\alpha x}\sin\beta x.$$

Примеры. Найти общее решение системы

$$\frac{dy_1}{dx}=y_2$$
$$\frac{dy_2}{dx}=-4y_1+4y_2$$
$$\frac{dy_3}{dx}=-2y_1+y_2+2y_3.$$

В данном случае матрица системы $A=\begin{pmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{pmatrix}$. В силу результатов примера (п. 2 § 4 гл. IX)

$$e^{Ax}=\begin{pmatrix} (1-2x)e^{2x} & xe^{2x} & 0 \\ -4xe^{2x} & (2x+1)e^{2x} & 0 \\ -2xe^{2x} & xe^{2x} & e^{2x} \end{pmatrix} \qquad (10.14)$$

и, следовательно, общее решение имеет вид:

$$Y = e^{Ax} \cdot C = \begin{pmatrix} (1-2x)\,e^{2x} & xe^{2x} & 0 \\ -4xe^{2x} & (2x+1)\,e^{2x} & 0 \\ -2xe^{2x} & xe^{2x} & e^{2x} \end{pmatrix} \cdot \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix},$$

т. е.

$$y_1 = c_1 (1-2x)\,e^{2x} + c_2 x e^{2x}$$
$$y_2 = -4c_1 x e^{2x} + c_2 (2x+1)\,e^{2x}$$
$$y_3 = -2c_1 x e^{2x} + c_2 x e^{2x} + c_3 e^{2x}.$$

Заметим, что если надо найти частное решение, удовлетворяющее начальным условиям

$$Y(x_0) = Y_0,$$

то из соотношения

$$Y(x_0) = e^{Ax_0} C = Y_0$$

находим

$$C = e^{-Ax_0} Y_0$$

и, следовательно, искомое решение будет

$$Y = e^{A(x-x_0)}\, Y_0. \tag{10.15}$$

Так, если еще надо найти решение, удовлетворяющее начальным условиям

$$x_0 = 0,\ \ y_{10} = 1;\ \ y_{20} = 0;\ \ y_{30} = -1,$$

т. е. $Y(0) = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$, то, используя выражения (10.14) и (10.15), получаем

$$Y = \begin{pmatrix} (1-2x)\,e^{2x} \\ -4xe^{2x} \\ -(2x+1)\,e^{2x} \end{pmatrix},$$

или

$$y_1 = (1-2x)\,e^{2x},$$
$$y_2 = -4xe^{2x},$$
$$y_3 = -(2x+1)\,e^{2x}.$$

Для решения неоднородной системы (10.12) с постоянными коэффициентами воспользуемся формулой (10.8), в которую вместо $W(x)$ подставим e^{Ax}. Тогда получим общее решение системы (10.12):

$$Y = e^{Ax} C + e^{Ax} \int_{x_0}^{x} e^{-A\tau} F(\tau)\, d\tau = e^{Ax} C + \int_{x_0}^{x} e^{A(x-\tau)} F(\tau)\, d\tau.$$

Пример. Найти общее решение системы

$$\frac{dy_1}{dx} = y_2$$
$$\frac{dy_2}{dx} = -4y_1 + 4y_2$$
$$\frac{dy_3}{dx} = -2y_1 + y_2 + 2y_3 + e^{3x}.$$

Матрица этой системы $A=\begin{pmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{pmatrix}$ совпадает с матрицей A предыдущего примера. Так как $F(x)=\begin{pmatrix} 0 \\ 0 \\ e^{3x} \end{pmatrix}$, то

$$e^{A(x-\tau)}F(\tau)=\begin{pmatrix} 0 \\ 0 \\ e^{2x+\tau} \end{pmatrix}.$$

Тогда, полагая $x_0=0$, получим

$$\int_0^x e^{A(x-\tau)}F(\tau)\,d\tau=\begin{pmatrix} 0 \\ 0 \\ \int_0^x e^{2x+\tau}\,d\tau \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ e^{3x}-e^{2x} \end{pmatrix}$$

и, значит, общее решение имеет вид

$$\begin{aligned} y_1 &= c_1(1-2x)e^{2x}+c_2xe^{2x} \\ y_2 &= -4c_1xe^{2x}+c_2(2x+1)e^{2x} \\ y_3 &= -2c_1xe^{2x}+c_2xe^{2x}+c_3e^{2x}+e^{3x}-e^{2x}. \end{aligned}$$

Заметим, что частное решение неоднородной системы (10.12), удовлетворяющее начальным условиям $Y(x_0)=Y_0$, очевидно, имеет вид

$$Y=e^{A(x-x_0)}Y_0+\int_{x_0}^x e^{A(x-\tau)}F(\tau)\,d\tau.$$

5. Линейное уравнение n-го порядка с постоянными коэффициентами. Применим полученные выше результаты к решению линейного однородного уравнения n-го порядка

$$y^{(n)}+a_1y^{(n-1)}+\ldots+a_ny=0,$$

которое согласно п. 6 § 1 приводится к системе

$$\frac{dY}{dx}=AY,$$

где

$$A=\begin{pmatrix} 0 & 1 & 0 \ldots & 0 \\ 0 & 0 & 1 \ldots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 \ldots & 1 \\ -a_n & -a_{n-1} & \ldots & -a_1 \end{pmatrix}.$$

Заметим, что ранг характеристической матрицы

$$A-\lambda E=\begin{vmatrix} -\lambda & 1 & 0 & \dots & 0 \\ 0 & -\lambda & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & & \dots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \dots & -a_1-\lambda \end{vmatrix}$$

при любом значении λ не меньше, чем $n-1$. Это вытекает из того, что минор элемента $-a_n$ равен

$$\begin{vmatrix} 1 & 0 & 0 & \dots & 0 \\ -\lambda & 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{vmatrix}=1$$

(т. е. отличен от нуля).

Но это значит, что ранг матрицы

$$A-\lambda_i E$$

(где λ_i — характеристическое число матрицы A), который всегда меньше, чем n, равен $n-1$. Отсюда следует, что число линейно независимых собственных векторов матрицы A равно числу различных корней характеристического многочлена $\Delta(\lambda)$, который в данном случае имеет вид*

$$\Delta(\lambda)=|A-\lambda E|=\lambda^n+a_1\lambda^{n-1}+\dots+a_n=0. \quad (10.16)$$

Таким образом, *число клеток жордановой матрицы I, подобной матрице A, равно числу различных корней характеристического многочлена*, а это означает, что *порядок каждой клетки равен кратности соответствующего корня.* Но в п. 1, § 2 было показано, что жордановой клетке порядка k, соответствующей характеристическому числу λ в общем решении отвечает выражение вида

$$(c_1x^{k-1}+c_2x^{k-2}+\dots+c_k)\,e^{\lambda x}. \quad (10.17)$$

---

* В этом легко убедиться, если каждый столбец с номером j умножить на λ^{j-1} и прибавить к нему все предыдущие столбцы, после чего раскрыть определитель по элементам последней строки $\left(\text{разделив его, конечно, на } \lambda^{1+2+\dots+(n-1)}=\lambda^{\frac{n(n-1)}{2}}\right)$.

Таким образом, общее решение линейного однородного уравнения с постоянными коэффициентами представляет собой сумму выражений вида (10.17), соответствующих различным корням характеристического многочлена (10.16).

Пример.

Найти общее решение уравнения

$$y''' + 2y'' + y' = 0.$$

Составим характеристический многочлен (10.16)

$$\Delta(\lambda) = \lambda^3 + 2\lambda^2 + \lambda = 0.$$

Его корни

$$\lambda_1 = \lambda_2 = -1 \text{ и } \lambda_3 = 0.$$

Следовательно,

$$y = (c_1 x + c_2) e^x + c_3.$$

Что касается неоднородного линейного уравнения с постоянными коэффициентами, то после того, как найдено общее решение соответствующего ему однородного уравнения, его решение в общем случае находится так же, как и решение неоднородного уравнения с произвольными коэффициентами (см. п. 6, § 1).

Упражнения

1. Найти общее решение системы уравнений

$$\frac{dx_1}{dt} = x_2 - 7x_1$$

$$\frac{dx_2}{dt} = -2x_1 - 5x_2.$$

2. Найти частное решение системы уравнений

$$\frac{dy_1}{dx} = y_2 + y_3,$$

$$\frac{dy_2}{dx} = y_3 + y_1,$$

$$\frac{dy_3}{dz} = y_1 + y_2$$

при начальных условиях

$$y_1(0) = -1, \quad y_2(0) = 1, \quad y_3 0 = 0.$$

3. Решить систему уравнений

$$y_1'' + 2y_2 = 0,$$

$$y_2'' - 2y_1 = 0, \quad \text{при начальных условиях:}$$

$$y_1(0) = 0, \quad y_1'(0) = 1, \quad y_2(0) = 0, \quad y_2'(0) = 0.$$

4. Найти общее решение уравнения

$$y^{IV} - y = 0.$$

5. Найти частное решение системы

$$\left.\begin{aligned} x_1' + 5x_1 - 2x_2 &= e^t \\ x_2^1 - x_1 + 6x_2 &= e^{-2t} \end{aligned}\right| \text{ при начальных условиях } x_1(0) = 0,\ x_2(0) = 0.$$

6. Найти частное решение уравнения $y'' + 4y' + 4y = f(x)$, удовлетворяющее нулевым начальным условиям.

Глава 11

ПРИЛОЖЕНИЕ МАТРИЦ К НЕКОТОРЫМ ВОПРОСАМ АНАЛИЗА

§ 1. НЕЛИНЕЙНЫЙ ОПЕРАТОР В ЕВКЛИДОВОМ ПРОСТРАНСТВЕ

1. Область. Рассмотрим n-мерное евклидово пространство $\mathscr{E}_n$, в котором выберем ортонормированный базис

$$\boldsymbol{e}_1,\ \boldsymbol{e}_2,\ \ldots,\ \boldsymbol{e}_n.$$

Напомним, что векторы $\boldsymbol{x}$ линейного пространства, а значит, и пространства $\mathscr{E}_n$ называют также точками и обозначают в этом случае x, или более подробно

$$x(x_1,\ x_2,\ \ldots,\ x_n),$$

где x_i — координаты вектора $\boldsymbol{x}$ в выбранном базисе. Множество точек x, расстояние которых до некоторой фиксированной точки x_0 удовлетворяет условию

$$\rho(x,\ x_0) = \rho(x,\ x_0) < r,$$

назовем *сферой радиуса* r с центром в точке x_0 или r — *окрестностью точки* x_0.

Множество точек пространства $\mathscr{E}_n$, которое вместе с любой точкой x содержит и некоторую (сколь угодно малую) ее окрестность, называется *областью* (открытой).

Напомним, что расстояние между двумя точками x и y евклидова пространства было введено в главе 6 и определялось как норма разности векторов $\boldsymbol{x}$ и $\boldsymbol{y}$, т. е.

$$\rho(x,\ y) = |\boldsymbol{x} - \boldsymbol{y}| = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}.$$

2. Нелинейный оператор. Согласно общему определению оператора $\tilde{A}$ в линейном пространстве R, введенному в гл. 4, назовем *оператором*, отображающим евклидово пространство $\mathscr{E}_n$ в евклидово пространство $\mathscr{E}_m$, любое правило, которое каждому вектору $\boldsymbol{x} \in E_n$ ставит в соответствие единственный вектор $\boldsymbol{y} \in \mathscr{E}_m$. Если оператор $\tilde{A}$ нелинейный, то в отличие от принятой ранее записи $\boldsymbol{y} = \tilde{A}\boldsymbol{x}$ будем использовать обозначение

$$\boldsymbol{y} = f(\boldsymbol{x}), \text{ или } Y = f(X), \tag{11.1}$$

где X и Y — матрицы-столбцы

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}; \qquad Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}.$$

Указанное отображение устанавливает связь между векторами $\boldsymbol{x}$ и $\boldsymbol{y}$, а значит, и между их координатами. Можно задать это отображение и в координатной форме:

$$\begin{gathered} y_1 = f_1(x_1, x_2, \ldots x_n) \\ \cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots \\ y_m = f_m(x_1, x_2, \ldots x_n). \end{gathered} \tag{11.2}$$

Таким образом, задание оператора (11.1) эквивалентно заданию системы функций n переменных (11.2) (при условии, что выбраны соответствующие базисы в пространствах $\mathscr{E}_n$ и $\mathscr{E}_m$). Отметим, что если $f_i(x_1, x_2, \ldots x_n)$ $(i = 1, 2, \ldots m)$ — линейные однородные функции, то оператор будет линейным.

Примеры:

1) функция одной переменной $y = f(x)$, заданная для всех значений x, определяет оператор, отображающий $\mathscr{E}_1$ в $\mathscr{E}_1$;

2) функция двух переменных $z = f(x, y)$, заданная для всех значений x и y, определяет оператор, отображающий $\mathscr{E}_2$ в $\mathscr{E}_1$;

3) совокупность двух функций одной переменной $y = f_1(x)$ и $z = f_2(x)$, заданных для всех значений x, определяет оператор, отображающий $\mathscr{E}_1$ в $\mathscr{E}_2$.

Будем говорить, что оператор отображает область G пространства $\mathscr{E}_n$ в область D пространства $\mathscr{E}_m$, если каждому вектору $\bar{\boldsymbol{x}} \in G$ он ставит вектор $\boldsymbol{y} \in D$.

Отметим, что в частном случае область G или D может совпадать со всем пространством.

Например, функция $y=\frac{1}{\sqrt{1-x^2}}$, заданная в области $-1<x<1$, определяет оператор, отображающий эту область пространства E_1 в область $0\leqslant x<\infty$ — того же самого пространства, оператор

$$y=\sin x$$

отображает пространство $\mathscr{E}_1$ в область $-1\leqslant y\leqslant 1$, а оператор

$$y=x^3+1$$

отображает пространство $\mathscr{E}_1$ в $\mathscr{E}_1$.

Если все функции $f_1(x_1, x_2, \ldots, x_n)$ $(i=1, 2, \ldots m)$ непрерывны в области G, то оператор f называется *непрерывным оператором*.

3. **Дифференцируемый оператор.** Пусть $\boldsymbol{x}\in E_n$, $\boldsymbol{y}\in E_m$ и $\boldsymbol{y}=f(\boldsymbol{x})$, дадим вектору $\boldsymbol{x}$ приращение $\Delta\boldsymbol{x}$, тогда вектор $\boldsymbol{y}$ получит приращение

$$\Delta\boldsymbol{y}=f(\boldsymbol{x}+\Delta\boldsymbol{x})-f(\boldsymbol{x})$$

или в координатной форме

$$\Delta y_i=f_i(x_1+\Delta x_1, x_2+\Delta x_2, \ldots, x_n+\Delta x_n)-f_i(x_1, x_2, \ldots, x_n)$$
$$i=1, 2, \ldots m.$$

Оператор f называется *дифференцируемым в точке* $x(x_1, x_2, \ldots, x_n)$, *если все функции* $f_i(x_1, x_2, \ldots, x_n)$ *дифференцируемы в этой точке.*

Но в этом случае, как известно,

$$\Delta y_i=\frac{\partial f_i}{\partial x_1}\Delta x_1+\frac{\partial f_i}{\partial x_2}\Delta x_2+\ldots+\frac{\partial f_i}{\partial x_n}\Delta x_n+\rho\varepsilon_i;\ i=1, 2, \ldots m,$$

где

$$\rho=|\Delta\boldsymbol{x}|, \text{ а } \lim_{\rho\to 0}\varepsilon_i=0.$$

Таким образом, если оператор f дифференцируем, то выполняется матричное равенство

$$\Delta Y=A\cdot\Delta X+\rho\varepsilon, \tag{11.3}$$

где

$$\Delta X=\begin{pmatrix}\Delta x_1\\ \Delta x_2\\ \vdots\\ \Delta x_n\end{pmatrix};\qquad \Delta Y=\begin{pmatrix}\Delta y_1\\ \Delta y_2\\ \vdots\\ \Delta y_m\end{pmatrix};\qquad \varepsilon=\begin{pmatrix}\varepsilon_1\\ \varepsilon_2\\ \vdots\\ \varepsilon_n\end{pmatrix};$$

$$A=\begin{vmatrix}\frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial f_2} & \cdots & \frac{\partial f_1}{\partial x_n}\\ \cdot & \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot & \cdot\\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n}\end{vmatrix}=\left(\frac{\partial f_i}{\partial x_k}\right)_{m,n}.$$

Матрица A, элементами которой являются частные производные $\frac{\partial f_i}{\partial x_k}$ $(i=1, 2, \ldots, m, k=1, 2, \ldots, n)$, называется *якобиевой матрицей*.

Если ввести линейный оператор $\tilde{A}$, матрица которого (в соответствующем базисе) совпадает с якобиевой матрицей A, то равенство (11.3) можно записать в векторной форме

$$\Delta \boldsymbol{y}=\tilde{A}\Delta \boldsymbol{x}+\rho\boldsymbol{\varepsilon}. \qquad (11.4)$$

Если якобиева матрица A — квадратная $(m=n)$, то ее определитель называется *якобианом* и обозначается

$$I=\det A=\frac{D(f_1, f_2, \ldots, f_n)}{D(x_1, x_2, \ldots, x_n)}=\frac{D(y_1, y_2, \ldots, y_n)}{D(x_1, x_2, \ldots, x_n)}.$$

4. Сложный оператор. Пусть оператор f преобразует векторы $\boldsymbol{x}$ пространства $\mathscr{E}_n$ в векторы $\boldsymbol{y}$ пространства $\mathscr{E}_m$, а оператор φ преобразует векторы $\boldsymbol{y}$ в векторы $\boldsymbol{z}$ пространства $\mathscr{E}_p$. Тем самым задается оператор, преобразующий векторы $\boldsymbol{x}$ в векторы $\boldsymbol{z}$, который будем называть *сложным* или *составным оператором*.

Так как

$$\boldsymbol{y}=f(\boldsymbol{x}),\ \text{а}\ \boldsymbol{z}=\varphi(\boldsymbol{y}),\ \text{то}\ \boldsymbol{z}=\varphi[f(\boldsymbol{x})].$$

Предполагая операторы f и φ дифференцируемыми, найдем якобиеву матрицу C сложного оператора.

Учитывая, что

$$z_k=\varphi_k(y_1, y_2, \ldots, y_m)\qquad k=1, 2, \ldots p$$
$$y_j=f_j(x_1, x_2, \ldots, x_n)\qquad j=1, 2, \ldots m$$

и, используя правило дифференцирования сложной функции нескольких переменных, имеем

$$\frac{\partial z_k}{\partial x_i}=\sum_{j=1}^{m}\frac{\partial z_k}{\partial y_j}\frac{\partial y_j}{\partial x_i},$$

откуда непосредственно следует, что $C=BA$,

где $A=\left(\frac{\partial y_j}{\partial x_i}\right)$ — якобиева матрица оператора f;

$B=\left(\frac{\partial z_k}{\partial y_j}\right)$ — якобиева матрица оператора φ.

Таким образом, якобиева матрица сложного оператора равна произведению якобиевых матриц операторов φ и f. Если $p=m=n$, то все три якобиевы матрицы квадратны и на основании свойства определителя произведения матриц имеем

$$\frac{D(z_1, z_2, \ldots, z_n)}{D(x_1, x_2, \ldots, x_n)}=\frac{D(z_1, z_2, \ldots, z_n)}{D(y_1, y_2, \ldots, y_n)}\cdot\frac{D(y_1, y_2, \ldots, y_n)}{D(x_1, x_2, \ldots, x_n)}. \quad (11.5)$$

5. Обратный оператор. Рассмотрим случай, когда вектор z, полученный в результате сложного преобразования, совпадает с исходным вектором x, т. е.

$$x=z=\varphi[f(x)]$$

для любого x из области задания оператора f. В этом случае оператор φ называется *обратным* по отношению к оператору f и обозначается $\varphi=f^{-1}$.

Не всякий оператор f имеет обратный. Чтобы f^{-1} существовал, надо на оператор f наложить некоторые дополнительные условия.

Приведем лишь теорему, доказательство которой читатель может найти, например, в курсе анализа Г. М. Фихтенгольца.

Теорема. *Если якобиан оператора непрерывен в некоторой точке и ее окрестности и отличен от нуля в этой точке, то найдется такая (достаточно малая) окрестность этой точки, в которой существует дифференцируемый оператор, обратный данному.*

Не останавливаясь на доказательстве этой теоремы, поясним только суть одного из сформулированных условий.

Если оператор f дифференцируем в окрестности точки x, то с точностью до бесконечно малых более высокого порядка, чем $\rho=|\Delta x|$, преобразование с помощью опе-

ратора f (в силу равенства (11.4)) может быть заменено в окрестности точки x линейным оператором $\tilde{A}$. Но, как мы знаем, линейный оператор имеет обратный оператор $\tilde{A}^{-1}$, если его матрица A невырождена, т. е. если якобиан $I = \det A$ отличен от нуля.

Если обратный оператор существует, то из равенства (11.5) следует

$$1 = \frac{D(x_1, x_2, \ldots, x_n)}{D(y_1, y_2, \ldots, y_n)} \cdot \frac{D(y_1, y_2, \ldots, y_n)}{D(x_1, x_2, \ldots, x_n)}.$$

Очевидно, что в этом случае якобиевы матрицы операторов f и f^{-1} взаимно обратны.

Также очевидно *необходимым и достаточным условием существования обратного оператора f^{-1} является наличие взаимнооднозначного соответствия между векторами $\boldsymbol{x}$ и $\boldsymbol{y}$.*

§ 2. ЭКСТРЕМУМ ФУНКЦИИ НЕСКОЛЬКИХ ПЕРЕМЕННЫХ

1. Определение экстремума. Пусть функция нескольких переменных

$$y = f(x_1, x_2, \ldots, x_n) = f(\boldsymbol{x})$$

определена в области G евклидова пространства $\mathscr{E}_n$. Точка $x^0(x_1^0, x_2^0, \ldots, x_n^0)$ называется *точкой экстремума*, если в некоторой ее окрестности приращение функции

$$\Delta y = \Delta f = f(\boldsymbol{x}) - f(\boldsymbol{x}^0)$$

не меняет знака.

Если $\Delta f \leqslant 0$, то x^0 называется *точкой максимума*, а если $\Delta f \geqslant 0$, то — *точкой минимума*.

2. Необходимые условия экстремума. Теорема. *Если функция $f(\boldsymbol{x})$ дифференцируема и достигает в точке x^0 экстремального значения, то в этой точке равны нулю все ее частные производные первого порядка,* т. е.

$$\frac{\partial f(\boldsymbol{x}^0)}{\partial x_i} = 0; \quad i = 1, 2, \ldots n.$$

Действительно, по условию $y = f(\boldsymbol{x})$ — дифференцируемая функция. Поэтому

$$\Delta y = f(\boldsymbol{x}) - f(\boldsymbol{x}^0) = \frac{\partial f(\boldsymbol{x}^0)}{\partial x_1} \Delta x_1 + \ldots + \frac{\partial f(\boldsymbol{x}^0)}{\partial x_n} \Delta x_n + \rho\varepsilon.$$

Предположим, что хоть одна из частных производных в точке x^0 отлична от нуля.

Пусть

$$\frac{\partial f(x^0)}{\partial x_i} \neq 0.$$

Тогда при малом ρ величина $\rho\varepsilon$ не влияет на знак Δy. Положим

$$\Delta x_1 = \ldots = \Delta x_{i-1} = \Delta x_{i+1} = \ldots = \Delta x_n = 0.$$

В этом случае очевидно, что при изменении знака Δx_i приращение Δy изменит знак, что противоречит определению экстремума. В дальнейшем точки, в которых $\frac{\partial f}{\partial x_i} = 0$ $(i = 1, 2, \ldots n)$, будем называть *стационарными точками*.

3. Достаточные условия экстремума. *Если функция $y = f(x)$ дважды непрерывно дифференцируема в стационарной точке x^0 и ее дифференциал второго порядка в окрестности этой точки меньше (больше) нуля, то точка x^0 является точкой максимума (минимума).*

Так как функция $f(x)$ дважды непрерывно дифференцируема, то

$$\Delta y = df(x^0) + \frac{1}{2} d^2 f(x^0) + \varepsilon\rho^2.$$

По условию x^0 — стационарная точка, поэтому $df(x^0) = 0$. Так как

$$d^2 f(x^0) \neq 0,$$

то при достаточно малом ρ знак Δy совпадает со знаком $d^2 f(x^0)$.

Значит, при $d^2 f(x^0) < 0$ x^0 — точка максимума, а при $d^2 f(x^0) > 0$ x^0 — точка минимума. Найдем теперь условия, при которых $d^2 f(x^0)$ сохраняет знак в окрестности точки x^0.

Как известно,

$$d^2 f(x^0) = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f(x^0)}{\partial x_i \partial x_j} \Delta x_i \Delta x_j.$$

Правая часть этого равенства представляет собой квадратичную форму переменных $\Delta x_1, \Delta x_2, \ldots \Delta x_n$, матрица ко-

торой

$$H=\begin{pmatrix} \frac{\partial^2 f(x^0)}{\partial x_1^2} & \frac{\partial^2 f(x^0)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(x^0)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x^0)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x^0)}{\partial x_2^2} & \cdots & \frac{\partial^2 f(x^0)}{\partial x_2 \partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2 f(x^0)}{\partial x_n \partial x_1} & \frac{\partial^2 f(x^0)}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(x^0)}{\partial x_n^2} \end{pmatrix}.$$

Используя критерий Сильвестра (см. п. 2 § 4 гл. 7), получаем:

1) если главные диагональные миноры матрицы H положительны, то $d^2f(x^0)>0$ и x^0 — точка минимума;

2) если главные диагональные миноры матрицы $-H$ положительны, то $d^2f(x^0)<0$ и x^0 — точка максимума.

Замечание. Очевидно, что если $d^2f(x)$ не сохраняет знака в окрестности x^0, то экстремума в точке x^0 нет.

Если в окрестности точки x^0 $d^2f(x)\geqslant 0$ или $d^2f(x)\leqslant 0$, то вопрос о существовании экстремума в этой точке остается открытым.

4. **Экстремум функции двух переменных.** Пусть функция двух переменных

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

дважды непрерывно дифференцируема.

Для нахождения экстремальных точек надо предварительно найти стационарные точки, координаты которых удовлетворяют системе уравнений

$$f'_x(x,\ y)=0,$$
$$f'_y(x,\ y)=0.$$

Пусть x_0, y_0 — одно из вещественных решений этой системы. Тогда точка $P_0(x_0,\ y_0)$ — стационарная.

Вычислим частные производные второго порядка в этой точке. Обозначим

$$A=\frac{\partial^2 f(P_0)}{\partial x^2},\quad B=\frac{\partial^2 f(P_0)}{\partial x\,\partial y},\quad C=\frac{\partial^2 f(P_0)}{\partial y^2}.$$

Составим матрицу

$$H=\begin{pmatrix} A & B \\ B & C \end{pmatrix}.$$

Главными диагональными минорами матрицы H будут

$$A \quad \text{и} \quad \begin{vmatrix} A & B \\ B & C \end{vmatrix} = AC - B^2,$$

а матрицы $-H$

$$-A \quad \text{и} \quad \begin{vmatrix} -A & -B \\ -B & -C \end{vmatrix} = AC - B^2.$$

Следовательно, если

1) $AC - B^2 > 0$ и $A > 0$, то P_0 — точка минимума;

2) $AC - B^2 > 0$ и $A < 0$, то P_0 — точка максимума.

Легко показать, что если $AC - B^2 < 0$, то экстремума в точке P_0 нет, а если $AC - B^2 = 0$, то имеет место сомнительный случай.

§ 3. ЗАМЕНА ПЕРЕМЕННЫХ В ДИФФЕРЕНЦИАЛЬНЫХ ВЫРАЖЕНИЯХ

1. Преобразование переменных. Имеем две системы величин $x_1, x_2, \ldots, x_n$ и $y_1, y_2, \ldots, y_n$, связанные между собой соотношениями

$$y_i = f_i(x_1, x_2, \ldots, x_n) \qquad i = 1, 2, \ldots n. \tag{11.6}$$

Эти соотношения можно рассматривать как преобразование координат в пространстве $\mathscr{E}_n$ (или как замену переменных). Например, соотношения $x = r\cos\varphi$, $y = r\sin\varphi$ можно рассматривать как формулы перехода от полярных координат к декартовым. Соотношения (11.6) определяют некоторый оператор $\boldsymbol{y} = f(\boldsymbol{x})$, который отображает пространство $\mathscr{E}_n$ в $\mathscr{E}_n$.

Будем в дальнейшем считать, что оператор непрерывный, дифференцируемый и имеет обратный оператор в $\mathscr{E}_n$.

2. Преобразование производных. Пусть задана функция $z = z(x_1, x_2, \ldots, x_n)$, дифференцируемая необходимое число раз.

Сделаем замену переменных

$$x_i = f_i(t_1, t_2, \ldots, t_n) \qquad i = 1, 2, \ldots n.$$

Требуется найти формулы, связывающие производные от функции z по старым переменным x_i с производными от z по новым переменным t_i. По формуле производной слож-

ной функции имеем

$$\frac{\partial z}{\partial x_i}=\sum_{j=1}^{n}\frac{\partial z}{\partial t_j}\frac{\partial t_j}{\partial x_i}\qquad i=1,\ 2,\ \ldots\ n.$$

Запишем эти соотношения в матричном виде

$$\begin{pmatrix}\frac{\partial z}{\partial x_1}\\ \frac{\partial z}{\partial x_2}\\ \vdots\\ \frac{\partial z}{\partial x_n}\end{pmatrix}=A\begin{pmatrix}\frac{\partial z}{\partial t_1}\\ \frac{\partial z}{\partial t_2}\\ \vdots\\ \frac{\partial z}{\partial t_n}\end{pmatrix},\tag{11.7}$$

где

$$A=\begin{pmatrix}\frac{\partial t_1}{\partial x_1} & \frac{\partial t_2}{\partial x_1} & \cdots & \frac{\partial t_n}{\partial x_1}\\ \frac{\partial t_1}{\partial x_2} & \frac{\partial t_2}{\partial x_2} & \cdots & \frac{\partial t_n}{d x_t}\\ \cdot & \cdot & \cdot & \cdot\\ \frac{\partial t_1}{\partial x_n} & \frac{\partial t_2}{\partial x_n} & \cdots & \frac{\partial t_n}{\partial x_n}\end{pmatrix}=\left(\frac{\partial t_i}{\partial x_k}\right).$$

Матрица A — это транспонированная якобиева матрица оператора f^{-1}. Поэтому она является обратной матрицей для транспонированной якобиевой матрицы оператора f, т. е.

$$A=\left(\frac{\partial x_k}{\partial t_i}\right)^{-1}=\frac{1}{\frac{D(x_1,\ x_2,\ \ldots,\ x_n)}{D(t_1,\ t_2,\ \ldots,\ t_n)}}(A_{ik}),$$

где A_{ik} — алгебраическое дополнение элемента $\frac{\partial x_k}{\partial t_i}$ матрицы $\left(\frac{\partial x_k}{\partial t_i}\right)$. Подставляя найденное значение A в соотношение (11.7), получаем

$$\frac{\partial z}{\partial x_i}=\frac{\sum\limits_{k=1}^{n}A_{ik}\frac{\partial z}{\partial t_k}}{\frac{D(x_1,\ x_2,\ \ldots,\ x_n)}{D(t_1,\ t_2,\ \ldots,\ t_n)}}=\frac{\frac{D(x_1,\ x_2,\ \ldots,\ x_{i-1},\ z,\ x_{i+1},\ \ldots,\ x_n)}{D(t_1,\ t_2,\ \ldots,\ t_n)}}{\frac{D(x_1,\ x_2,\ \ldots,\ x_n)}{D(t_1,\ t_2,\ \ldots,\ t_n)}}.$$

Эти формулы выражают производные z по старым переменным через производные z по новым переменным.

Аналогично получаем

$$\frac{\partial^2 z}{\partial x_i\,\partial x_j}=\frac{\dfrac{D\left(x_1,\ldots,x_{j-1},\dfrac{\partial z}{x_i},x_{j+1},\ldots,x_n\right)}{D(t_1,t_2,\ldots,t_n)}}{\dfrac{D(x_1,x_2,\ldots,x_n)}{D(t_1,t_2,\ldots,t_n)}},$$

где вместо $\frac{\partial z}{\partial x_i}$ надо подставить его выражение из предыдущей формулы. Таким же образом можем найти производные более высоких порядков.

Полученные формулы используются при замене независимых переменных в дифференциальных выражениях, т. е. выражениях вида

$$F\left(x_1,\ x_2,\ldots,x_n,\ z,\ \frac{\partial z}{\partial x_1},\frac{\partial z}{\partial x_2},\ldots,\frac{\partial z}{\partial x_n},\frac{\partial^2 z}{\partial x_1^2},\ldots\right).$$

Пример.

В волновом уравнении $\frac{\partial^2 z}{\partial x^2}-\frac{\partial^2 z}{\partial y^2}=0$ произвести замену переменных

$$x=u+v,$$
$$y=u-v.$$

Имеем

$$\frac{D(x,y)}{D(u,v)}=\begin{vmatrix}1 & 1\\ 1 & -1\end{vmatrix}=-2,$$

$$\frac{\partial z}{\partial x}=\frac{\dfrac{D(z,y)}{D(u,v)}}{-2}=-\frac{1}{2}\begin{vmatrix}\dfrac{\partial z}{\partial u} & 1\\ \dfrac{\partial z}{\partial v} & -1\end{vmatrix}=\frac{1}{2}\left(\frac{\partial z}{\partial u}+\frac{\partial z}{\partial v}\right).$$

Аналогично $\frac{\partial z}{\partial y}=\frac{1}{2}\left(\frac{\partial z}{\partial u}-\frac{\partial z}{\partial v}\right)$,

$$\frac{\partial^2 z}{\partial x^2}=\frac{\dfrac{D\left(\dfrac{\partial z}{\partial x},y\right)}{\partial(u,v)}}{-2}=-\frac{1}{2}\begin{vmatrix}\dfrac{\partial}{\partial u}\left[\dfrac{1}{2}\left(\dfrac{\partial z}{\partial u}+\dfrac{\partial z}{\partial v}\right)\right] & 1\\ \dfrac{\partial}{\partial v}\left[\dfrac{1}{2}\left(\dfrac{\partial z}{\partial u}+\dfrac{\partial z}{\partial v}\right)\right] & -1\end{vmatrix}=$$

$$=\frac{1}{4}\left(\frac{\partial^2 z}{\partial u^2}+2\frac{\partial^2 z}{\partial u\,\partial v}+\frac{\partial^2 z}{\partial v^2}\right).$$

Аналогично $\frac{\partial^2 z}{\partial y^2}=\frac{1}{4}\left(\frac{\partial^2 z}{\partial u^2}-2\frac{\partial^2 z}{\partial u\,\partial v}+\frac{\partial^2 z}{\partial v^2}\right)$.

Таким образом,

$$\frac{\partial^2 z}{\partial x^2}-\frac{\partial^2 z}{\partial y^2}=\frac{\partial^2 z}{\partial u\,\partial v}.$$

Получаем

$$\frac{\partial^2 z}{\partial u\, \partial v} = 0.$$

3. **Криволинейные координаты.** В § 1 настоящей главы система m функций n переменных, заданных в пространстве $\mathscr{E}_n$, трактовалась как некоторый оператор, отображающий пространство $\mathscr{E}_n$ в пространство $\mathscr{E}_m$ (при условии, что выбраны соответствующие базисы в пространстве $\mathscr{E}_n$ и $\mathscr{E}_m$). Однако при $m = n$ и некоторых дополнительных условиях этой системе функций можно дать иную, чисто геометрическую трактовку, связанную с переходом от одной системы координат в $\mathscr{E}_n$ к другой системе координат в этом же пространстве. Так как практически нас будет интересовать случай перехода от одной системы координат к другой только в плоскости и обычном пространстве, то мы ограничимся лишь значениями $n = 2$ и $n = 3$, разобрав каждый из этих случаев достаточно подробно.

Рассмотрим замену переменных вида

$$\begin{aligned} u &= u(x, y), \\ v &= v(x, y). \end{aligned} \tag{11.8}$$

Как и раньше будем считать, что функции u и v однозначные, непрерывно дифференцируемые и имеют обратные функции

$$\begin{aligned} x &= x(u, v), \\ y &= y(u, v). \end{aligned} \tag{11.9}$$

Мы знаем, что соотношения (11.8) определяют оператор, отображающий (взаимно однозначно) $\mathscr{E}_2$ в $\mathscr{E}_2$. Иначе, соотношения (11.8) каждой точке $P(x, y)$ пространства xOy ставят в соответствие единственную точку $Q(u, v)$ пространства $uO'v$ (и наоборот).

Будем считать, что x и y — декартовы прямоугольные координаты точки P плоскости xOy, а соответствующие им u и v — декартовы прямоугольные координаты точки Q плоскости $uO'v$. Координаты u и v точки Q в дальнейшем будем называть *криволинейными координатами точки* P. Очевидно, что если дана точка P, то по формулам (11.8) можно найти ее криволинейные координаты, и, наоборот, если известны криволинейные координаты u и v, то по

формулам (11.9) однозначно определяются её декартовы координаты.

Зафиксируем теперь криволинейную координату u (положим $u = u_0 = \text{const}$). В этом случае в плоскости xOy получим линию, уравнение которой $u(x, y) = u_0$ и будем называть ее *координатной линией*. Меняя параметр u_0, получим семейство координатных линий.

Аналогично получаем семейство координатных линий

$$v(x, y) = v_0.$$

Очевидно, этим координатным линиям соответствуют прямые, параллельные осям координат ($u = u_0$, $v = v_0$) в плоскости $uO'v$ (см. рис. 6).

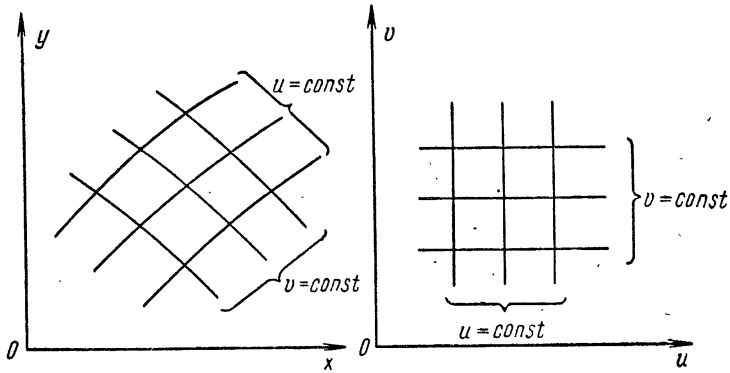


Рис. 6

Так как по условию соответствие (11.8) взаимно однозначное, то две координатные линии разных семейств пересекаются только в одной точке, а координатные линии одного семейства не пересекаются. Любая точка P является точкой пересечения двух координатных линий различных семейств, соответствующие им значения параметров ($u = u_1$, $v = v_1$) и служат криволинейными координатами точки P (отсюда термин криволинейные координаты).

Таким образом, соотношения (11.8) можно трактовать двояко:

во-первых, как соответствие между точками (P и Q) разных плоскостей;

во-вторых, как формулы, преобразующие систему координат в одной и той же плоскости.

Аналогично вводятся криволинейные координаты в пространстве соотношениями

$$u = u(x, y, z),$$
$$v = v(x, y, z),$$
$$w = w(x, y, z),$$

где функции u, v, и w — однозначные, непрерывно дифференцируемые и имеют обратные функции.

Координатными поверхностями называются поверхности

$$u(x, y, z) = u_0, \qquad v = (x, y, z) = v_0 \qquad w = (x, y, z) = w_0.$$

Линия пересечения двух координатных поверхностей различных семейств называется *координатной линией*.

Например, поверхности $u(x, y, z) = u_0$ и $v(x, y, z) = v_0$ пересекаются по координатной линии, которую назовем *w линией*. Каждая точка есть точка пересечения трех координатных поверхностей. Введем базис в криволинейной системе координат:

$$\begin{aligned} \boldsymbol{e}_u &= \frac{1}{L_u}\frac{\partial \boldsymbol{r}}{\partial u} = \frac{1}{L_u}\left(\frac{\partial x}{\partial u}\boldsymbol{i} + \frac{\partial y}{\partial u}\boldsymbol{j} + \frac{\partial z}{\partial u}\boldsymbol{k}\right), \\ \boldsymbol{e}_v &= \frac{1}{L_v}\frac{\partial \boldsymbol{r}}{\partial v} = \frac{1}{L_v}\left(\frac{\partial x}{\partial v}\boldsymbol{i} + \frac{\partial y}{\partial v}\boldsymbol{j} + \frac{\partial z}{\partial v}\boldsymbol{k}\right), \\ \boldsymbol{e}_w &= \frac{1}{L_w}\frac{\partial \boldsymbol{r}}{\partial w} = \frac{1}{L_w}\left(\frac{\partial x}{\partial w}\boldsymbol{i} + \frac{\partial y}{\partial w}\boldsymbol{j} + \frac{\partial z}{\partial w}\boldsymbol{k}\right), \end{aligned} \tag{11.10}$$

где

$$L_u = \left|\frac{\partial \boldsymbol{r}}{\partial u}\right|; \quad L_v = \left|\frac{\partial \boldsymbol{r}}{\partial v}\right|; \quad L_w = \left|\frac{\partial \boldsymbol{r}}{\partial w}\right|;$$
$$\boldsymbol{r} = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}.$$

Эти единичные векторы (как известно из курса анализа) направлены по касательным к координатным линиям, пересекающимся в точке P.

Если векторы $\boldsymbol{e}_u$, $\boldsymbol{e}_v$, $\boldsymbol{e}_w$ попарно ортогональны в каждой точке, то криволинейные координаты u, v, w называются *ортогональными*.

Рассмотрим вектор $\boldsymbol{F} = F_x(x, y, z)\boldsymbol{i} + F_y(x, y, z)\boldsymbol{j} + F_z(x, y, z)\boldsymbol{k}$. Найдем его координаты в базисе $\boldsymbol{e}_u$, $\boldsymbol{e}_v$, $\boldsymbol{e}_w$. Учитывая формулы (11.10) и соотношение (4.10),

имеем

$$\begin{pmatrix} F_x \\ F_y \\ F_z \end{pmatrix} = T \begin{pmatrix} F_u \\ F_v \\ F_w \end{pmatrix},$$

где

$$T = \begin{pmatrix} \frac{1}{L_u}\frac{\partial x}{\partial u} & \frac{1}{L_v}\frac{\partial x}{\partial v} & \frac{1}{L_w}\frac{\partial x}{\partial w} \\ \frac{1}{L_u}\frac{\partial y}{\partial u} & \frac{1}{L_v}\frac{\partial y}{\partial v} & \frac{1}{L_w}\frac{\partial y}{\partial w} \\ \frac{1}{L_u}\frac{\partial z}{\partial u} & \frac{1}{L_v}\frac{\partial z}{\partial v} & \frac{1}{L_w}\frac{\partial z}{\partial w} \end{pmatrix}$$ — матрица перехода от базиса $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ к базису $\boldsymbol{e}_u, \boldsymbol{e}_v, \boldsymbol{e}_w$.

Отсюда

$$\begin{pmatrix} F_u \\ F_v \\ F_w \end{pmatrix} = T^{-1} \begin{pmatrix} F_x \\ F_y \\ F_z \end{pmatrix}. \tag{11.11}$$

Если криволинейные координаты ортогональны, то матрица T — ортогональна и $T^{-1} = T'$.

Пример.

Найти $\boldsymbol{f} = \operatorname{grad} \varphi(M) = \frac{\partial \varphi}{\partial x}\boldsymbol{i} + \frac{\partial \varphi}{\partial y}\boldsymbol{j} + \frac{\partial \varphi}{\partial z}\boldsymbol{k}$ в криволинейных координатах.

По формуле (11.7)

$$\begin{pmatrix} \frac{\partial \varphi}{\partial x} \\ \frac{\partial \varphi}{\partial y} \\ \frac{\partial \varphi}{\partial z} \end{pmatrix} = A \begin{pmatrix} \frac{\partial \varphi}{\partial u} \\ \frac{\partial \varphi}{\partial v} \\ \frac{\partial \varphi}{\partial w} \end{pmatrix}.$$

Тогда по формуле (11.11)

$$\begin{pmatrix} F_u \\ F_v \\ F_w \end{pmatrix} = T^{-1} A \begin{pmatrix} \frac{\partial \varphi}{\partial u} \\ \frac{\partial \varphi}{\partial v} \\ \frac{\partial \varphi}{\partial w} \end{pmatrix}.$$

Если система координат — ортогональная, то, как легко видеть

$$T^{-1}A=\begin{pmatrix}\frac{1}{L_u} & 0 & 0\\ 0 & \frac{1}{L_v} & 0\\ 0 & 0 & \frac{1}{L_w}\end{pmatrix}.$$

В этом случае

$$f(u, v, w)=\operatorname{grad}\varphi(M)=\frac{1}{L_u}\frac{\partial\varphi}{\partial u}e_u+\frac{1}{L_v}\frac{\partial\varphi}{\partial v}e_v+\frac{1}{L_w}\frac{\partial\varphi}{\partial w}e_w.$$

§ 4. ЗАМЕНА ПЕРЕМЕННЫХ В КРАТНЫХ ИНТЕГРАЛАХ

Рассмотренные в предыдущем параграфе криволинейные координаты находят широкое применение в самых различных разделах анализа, к которым прежде всего относятся кратные интегралы. Во многих случаях замена декартовых координат специально выбранными криволинейными координатами удобна при вычислении кратных интегралов.

Указанный переход от декартовых к криволинейным координатам называется *заменой переменных в кратных интегралах*.

Так же, как и при рассмотрении криволинейных координат, мы ограничимся пространствами $\mathscr{E}_2$ и $\mathscr{E}_3$ и рассмотрим только двойные и тройные интегралы.

1. Коэффициент искажения. Введем криволинейные координаты на плоскости соотношениями

$$u=u(x, y),\qquad v=v(x, y). \tag{11.12}$$

При этом точка $P_0(x_0, y_0)$ плоскости xOy отображается в точку $Q_0(u_0, v_0)$ плоскости $uO'v$.

Рассмотрим прямоугольник с вершиной в точке P_0 и сторонами, параллельными осям x и y. Его площадь обозначим s_P.

В плоскости $uO'v$ этому прямоугольнику будет соответствовать некоторый «криволинейный» параллелограмм с вершиной в точке Q_0. Его площадь обозначим s_Q.

Коэффициентом искажения $K(x_0, y_0)$ в точке $P_0(x_0, y_0)$ назовем *предел отношения площади s_Q к площади s_P* при условии, что $s_P\to 0$ и прямоугольник стягивается к точке P_0. Найдем $K(x_0, y_0)$, предварительно рассмотрев

частный случай

$$u = a_1 x + b_1 y + c_1,$$
$$v = a_2 x + b_2 y + c_2,$$

или в матричной форме

$$\begin{pmatrix} u \\ v \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} c_1 \\ c_2 \end{pmatrix},$$

где

$$A = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} \quad \text{и} \quad \Delta = \det A = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \neq 0.$$

Такое преобразование называется аффинным.

Произведем параллельный перенос системы координат по формулам

$$u = u_1 + c_1,$$
$$v = v_1 + c_2.$$

Очевидно, это преобразование не изменяет площади (следовательно, и коэффициент искажения остается неизменным).

Тогда

$$u_1 = a_1 x + b_1 y,$$
$$v_1 = a_2 x + b_2 y.$$

И так как это есть линейное преобразование, то (см. п. 2, § 5, гл. IV)

$$s_Q = |\Delta| s_P.$$

Отсюда

$$K(x_0, y_0) = |\Delta|.$$

Таким образом, коэффициент искажения при аффинном преобразовании постоянный и равен модулю определителя матрицы A. Переходя к общему случаю, заметим, что преобразование (11.12) в силу формулы (11.3) в окрестности точки P_0 можно записать в виде

$$\begin{pmatrix} u - u_0 \\ v - v_0 \end{pmatrix} = \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{pmatrix} \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} + \rho \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}.$$

Поэтому с точностью до бесконечно малых более высокого порядка, чем ρ, можно преобразование (11.12) в окрестности точки считать аффинным.

Следовательно:

$$K(x_0, y_0) = \lim_{\substack{s_P \to 0 \\ \rho \to 0}} \frac{s_Q}{s_P} = |I|,$$

где $I = \frac{D(u, v)}{D(x, y)}$ — якобиан преобразования (11.12) в точке P_0.

В отличие от аффинного преобразования в общем случае коэффициент искажения есть функция точки.

В пространственном случае

$$u = u(x, y, z),$$
$$v = v(x, y, z),$$
$$w = w(x, y, z).$$

Прямоугольный параллелепипед с вершиной в точке $P_0(x_0, y_0, z_0)$ и гранями, параллельными координатным плоскостям xOy, yOz и zOx (объем которого v_P) отображается в «криволинейный» параллелепипед с вершиной в точке $Q_0(u_0, v_0, w_0)$ (объем которого v_Q).

В этом случае коэффициентом искажения $K(x_0, y_0, z_0)$ в точке P_0 назовем *предел отношения объемов*

$$K(x_0, y_0, z_0) = \lim_{v_P \to 0} \frac{v_Q}{v_P}.$$

Полностью повторяя предшествующие рассуждения, получаем

$$K(x_0, y_0, z_0) = \left| \frac{D(u, v, w)}{D(x, y, z)} \right|.$$

2. Замена переменных в кратных интегралах. Имеем двойной интеграл

$$\iint_G f(x, y)\, dx\, dy.$$

По определению

$$\iint_G f(x, y)\, dx\, dy = \lim_{\lambda \to 0} \sum_{i=1}^{n} f(P_i)\, \Delta s_i,$$

где Δs_i — площади, а λ максимальный диаметр частичных областей, на которые разбита область G.

Как известно, этот предел не зависит от способа разбиения области G на части и выбора точек P_i в каждой части.

Произведем замену переменных

$$x = x(u, v),$$
$$y = y(u, v).$$

Предполагаем, что эти функции непрерывно дифференцируемы и осуществляют взаимно однозначное соответствие между точками области G и некоторой области D плоскости $uO'v$.

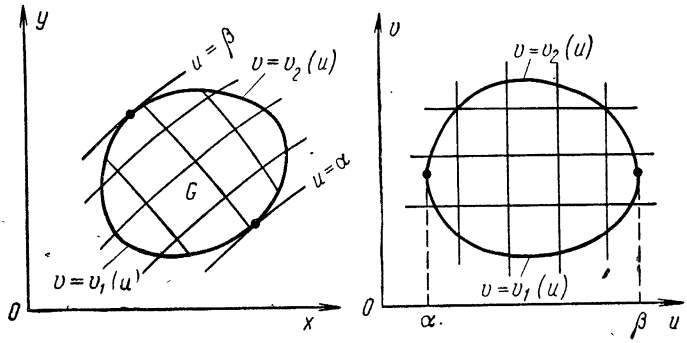


Рис. 7

Разобьем область D на частичные прямыми, параллельными осям координат (рис. 7). Площадь каждой частичной области обозначим $\Delta\sigma_i$.

Площади соответствующих областей плоскости xOy обозначим Δs_i.

Составим интегральную сумму

$$\sum_{i=1}^{n} f(P_i)\,\Delta s_i.$$

На основании геометрического смысла якобиана имеем

$$\Delta s_i = (|I(Q_i)| + \varepsilon_i)\,\Delta\sigma_i.$$

Поэтому

$$\sum_{i=1}^{n} f(P_i)\,\Delta s_i = \sum_{i=1}^{n} f[x(Q_i)\,y(Q_i)]\,(|I(Q_i)| + \varepsilon_i)\,\Delta\sigma_i.$$

Значит,

$$\iint\limits_G f(x, y)\,dx\,dy = \lim_{\lambda \to 0} \sum_{i=1}^{n} f(P_i)\,\Delta s_i =$$
$$= \iint\limits_D f[x(u, v),\ y(u, v)]\,|I|\,du\,dv,$$

где

$$I = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \frac{D(x, y)}{D(u, v)}.$$

Пример.

Вычислить $\iint\limits_G \sqrt{x^2+y^2}\,dx\,dy$, где G — область, заключенная между линиями

$$y = \sqrt{1-x^2},\ y = x,\ y = 0;\ x \geqslant 0.$$

Делаем замену

$$x = r\cos\varphi,$$
$$y = r\sin\varphi,$$
$$I = \begin{vmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & +r\cos\varphi \end{vmatrix} = r,$$
$$\iint\limits_G \sqrt{x^2+y^2}\,dx\,dy = \iint\limits_D \sqrt{r^2\cos^2\varphi + r^2\sin^2\varphi}\,r\,dr\,d\varphi =$$
$$= \int_0^{\pi/4} d\varphi \int_0^1 r^2\,dr = \frac{\pi}{12}.$$

Аналогично получаем

$$\iiint\limits_G f(x, y, z)\,dx\,dy\,dz = \iiint\limits_D f[x(u, v, w),\ y(u, v, w),\ z(u, v, w)] \times$$
$$\times |I|\,du\,dv\,dw,$$

где

$$I = \frac{D(x, y, z)}{D(u, v, w)}.$$

Замечание. Как отмечалось выше, соотношения

$$u = u(x, y),$$
$$v = v(x, y)$$

можно интерпретировать либо как формулы, преобразующие область, что использовалось при выводе формулы замены переменных в двойном интеграле, либо как формулы, преобразующие систему координат.

Поэтому можно в двойном интеграле произвести замену переменных, не переходя к новой области D.

Действительно, как мы видели, прямым, параллельным координатным осям в плоскости $vO'w$, соответствуют две системы координатных линий в плоскости xOy. Следовательно, при выбранном разбиении области D на части область G разобьется на элементарные области координатными линиями.

Исходя из этих соображений, получаем (см. рис. 7)

$$\iint\limits_G f(x, y)\,dx\,dy = \iint\limits_G f[x(u, v), y(u, v)]\,|I|\,du\,dv =$$

$$= \int\limits_\alpha^\beta du \int\limits_{v_1(u)}^{v_2(u)} f[x(u, v), y(u, v)]\,|I|\,dv,$$ где $v_1(u)$ и $v_2(u)$ уравнения границы области G в криволинейных координатах.

Аналогичный результат справедлив и для тройных интегралов:

$$\iiint\limits_G f(x, y, z)\,dx\,dy\,dz =$$

$$= \int\limits_\alpha^\beta du \int\limits_{v_1(u)}^{v_2(u)} dv \int\limits_{w_1(u, v)}^{w_2(u, v)} f[x(u, v, w), y(u, v, w), z(u, v, w)]\,|I|\,dw.$$

Глава 12

ЦЕПИ МАРКОВА

§ 1. ЦЕПИ МАРКОВА С ДИСКРЕТНЫМ ВРЕМЕНЕМ

Рассмотрим некоторую систему, которая может находиться в одном и только одном из k несовместных состояний $E_1, E_2, \ldots, E_k$. Предположим, что эти состояния могут сменяться одно другим (в частности, могут и сохраняться) только в заранее фиксированные моменты времени $t_0, t_1, t_2, \ldots, t_n, \ldots$

Тогда, если условная вероятность $p_{ij}(t_n)$ того, что система на интервале времени $[t_n, t_{n+1}]$ (или что то же самое в момент времени t_n) будет находиться в состоянии E_j при условии, что в момент времени t_{n-1} она находилась в состоянии E_i, определяется только этим последним состоянием и не изменяется от дополнительной информации о тех состояниях, в которых система находилась в более ранние моменты времени, предшествующие

моменту t_{n-1}, то случайный процесс смены состояний такой системы называется *простой* (или *односвязной*) *цепью Маркова*.

Короче, последовательность состояний некоторой системы образует простую цепь Маркова, если состояние системы в момент t_n стохастически полностью определяется ее состоянием в предшествующий момент времени.

Вероятности $p_{ij}(t_n)$ называются *вероятностями перехода* (из состояния i в состояние j в момент t_n). Они определяют квадратную матрицу

$$P(t_n) = \begin{pmatrix} p_{11}(t_n) & p_{12}(t_n) & p_{1k}(t_n) \\ p_{21}(t_n) & p_{22}(t_n) & p_{2k}(t_n) \\ \cdots & \cdots & \cdots \\ p_{k1}(t_n) & p_{2k}(t_n) & p_{kk}(t_n) \end{pmatrix},$$

которая называется *матрицей перехода*. Очевидно, что для матрицы перехода выполняются условия

$$0 \leqslant p_{ij}(t_n) \leqslant 1,$$

$$\sum_{j=1}^{k} p_{ij}(t_n) = 1.$$

Матрицы, элементы которых удовлетворят этим условиям, называются *стохастическими матрицами*. В общем случае, когда матрица перехода $P(t_n)$ зависит от t_n, цепь Маркова называется неоднородной. Если матрица перехода не зависит от t_n

$$P(t_n) = P = (p_{ij}) \quad i, j = 1, 2, \ldots k,$$

то цепь Маркова называется *однородной*.

Для полного описания простой цепи Маркова необходимо, кроме матрицы перехода, задать еще вероятности состояний системы в начальный момент времени t_0, $p_i(t_0)$; $i = 1, 2 \ldots k$ (или p_{0i}), которые называются *начальными вероятностями*.

Легко видеть, что простая цепь Маркова представляет собой естественное обобщение схемы независимых испытаний, состоящее в том, что вероятность исхода E_j в n-ом испытании зависит только от исхода в предшествующем ему $n-1$ испытании и не изменяется от дополнительных сведений относительно исходов в более ранних испытаниях.

Простейшим примером простой однородной цепи Маркова может служить следующая урновая схема: в двух урнах — белой и черной содержится соответственно m_1 белых, $n_1 - m_1$ черных и m_2 белых, $n_2 - m_2$ черных шаров. Выбирая наугад урну, извлекаем из нее шар (с возвращением) и фиксируем его цвет. Затем каждый раз выбираем урну, цвет которой совпадает с цветом вынутого шара и т. д. Назовем состоянием системы в момент t_n цвет вынутого шара (белый — *первое состояние*, черный — *второе состояние*) при n-ом испытании. Тогда очевидно начальные вероятности

$$p_{01} = p_{02} = 0{,}5,$$

а матрица перехода —

$$P = \begin{pmatrix} \frac{m_1}{n_1} & \frac{n_1 - m_1}{n_1} \\ \frac{m_2}{n_2} & \frac{n_2 - m_2}{n_2} \end{pmatrix}.$$

Один из основных вопросов при изучении простых цепей Маркова состоит в следующем: зная начальные вероятности и матрицу перехода, определить вероятности $p_i(t_n)$ состояний; E_i, $i = 1, 2 \ldots k$ в любой момент t_n и исследовать их поведение при $n \to \infty$.

Так как в дальнейшем мы будем рассматривать только простые, однородные цепи Маркова, то будем говорить просто цепь Маркова, кроме того, учитывая, что сами значения моментов времени t_n никакой роли не играют, будем вместо t_n писать n и называть эти моменты *шагами*. Таким образом, $p_i(t_n) = p_i(n)$ означает вероятность системы на n-ом шаге находиться в состоянии E_i.

Для решения поставленной выше задачи введем сначала вероятности $p_{ij}(m, n)$ перехода из состояния i на m-ом шаге в состояние j на n-ом шаге. Очевидно, $p_{ij}(n-1, n) = p_{ij}$ и не зависит от n (в силу однородности).

Введем произвольный промежуточный шаг s $(m < s < n)$, на котором система примет одно из возможных состояний, например E_l. Вероятность этого события равна $p_{il}(m, s)$. Тогда вероятность перехода из состояния E_l на s-ом шаге в состояние E_j на n-ом шаге будет равна $p_{lj}(s, n)$ и согласно формуле полной вероятности

$$p_{ij}(m, n) = \sum_{l=1}^{k} p_{il}(m, s)\, p_{lj}(s, n). \qquad (12.1)$$

Совокупность соотношений (12.1) (для всех i, j, фиксированных m и n и произвольном s), называемых *уравнениями Маркова*, удобно представить в матричной форме

$$P(m, n) = P(m, s)\, P(s, n), \tag{12.2}$$

где

$$P(m, n) = \begin{pmatrix} p_{11}(m, n) & p_{12}(m, n) & \dots & p_{1k}(m, n) \\ p_{21}(m, n) & p_{22}(m, n) & \dots & p_{2k}(m, n) \\ \dots & \dots & \dots & \dots \\ p_{k1}(m, n) & p_{k2}(m, n) & \dots & p_{kk}(m, n) \end{pmatrix}.$$

Матричное уравнение (12.2) легко решается. Для этого воспользовавшись тем, что s произвольно, положим $s = m + 1$ и так как

$$p_{il}(m, m+1) = p_{il},$$

то $P(m, m+1) = P$ — матрица перехода и, значит, $P(m, n) = P(m, m+1)\, P(m+1, n) = P \cdot P(m+1, n)$. Но отсюда следует

$$P(m, n) = P^2 P(m+2, n) = P^3 P(m+3, n) = \dots = P^{n-m}.$$

Таким образом, окончательно находим

$$P(m, n) = P^{n-m}. \tag{12.3}$$

Как видно из полученного решения в силу однородности цепи Маркова, переходные вероятности из состояния E_i на m-ом шаге в состояние E_j на n-ом шаге зависят только от разности $n - m$ (т. е. от числа шагов), поэтому их называют *вероятностями перехода* за $s = n - m$ шагов и обозначают

$$p_{ij}(m, n) = p_{ij}(n - m) = p_{ij}(s).$$

Соответствующая этим вероятностям матрица

$$P(s) = \begin{pmatrix} p_{11}(s) & p_{12}(s) & \dots & p_{1k}(s) \\ p_{21}(s) & p_{22}(s) & \dots & p_{2k}(s) \\ \dots & \dots & \dots & \dots \\ p_{k1}(s) & p_{k2}(s) & \dots & p_{kk}(s) \end{pmatrix}$$

равна s-ой степени матрицы перехода P. После нахождения вероятностей перехода за s шагов $p_{ij}(s)$ вероятности $p_i(n)$ вычисляются совершенно элементарно. В самом деле, чтобы система на n-ом шаге была в состоянии E_i, надо, чтобы из начального состояния (которое может

быть любым, например E_l, вероятность чего равна p_{ol}) она за n шагов перешла в состояние E_i, вероятность чего равна $p_{li}(n)$. Тогда по формуле полной вероятности

$$p_i(n) = \sum_{l=1}^{k} p_{ol} p_{li}(n)$$

или в матричной форме

$$P_n = P'(n) P_0 = (P^n)' P_0,$$

где

$$P_0 = \begin{pmatrix} p_{01} \\ p_{02} \\ \vdots \\ p_{0k} \end{pmatrix}; \quad P_n = \begin{pmatrix} p_1(n) \\ p_2(n) \\ \vdots \\ p_k(n) \end{pmatrix}.$$

Таким образом, матрица-столбец P_n выражена через матрицу перехода P и матрицу начальных вероятностей P_0.

Рассмотрим простейший пример цепи Маркова с двумя состояниями, обозначая вероятности перехода $p_{11} = 1 - \alpha$, $p_{12} = \alpha$; $p_{21} = \beta$; $p_{22} = 1 - \beta$, а начальные вероятности $p_{01} = p$ и $p_{02} = q = 1 - p$.

Тогда матрица $P(n)$ перехода за n шагов

$$P(n) = \begin{pmatrix} p_{11}(n) & p_{12}(n) \\ p_{21}(n) & p_{22}(n) \end{pmatrix} = P^n = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}^n.$$

Используя результаты возведения матрицы второго порядка совпадающей с матрицей P', приведенные в § 1 главы 9, имеем

$$P_n = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}^n = \frac{1}{\alpha+\beta} \begin{pmatrix} \beta + \alpha\mu^n & \alpha(1-\mu^n) \\ \beta(1-\mu^n) & \alpha + \beta\mu^n \end{pmatrix},$$

где

$$\mu = 1 - \alpha - \beta.$$

Соответственно матрица

$$P_n = \begin{pmatrix} p_1(n) \\ p_2(n) \end{pmatrix} = P'(n) \cdot P_0 = \frac{1}{\alpha+\beta} \begin{pmatrix} \beta + \alpha\mu^n, & \beta(1-\mu^n) \\ \alpha(1-\mu^n), & \alpha + \beta\mu^n \end{pmatrix} \cdot \begin{pmatrix} p \\ q \end{pmatrix}$$

или

$$P_n = \frac{1}{\alpha+\beta} \begin{pmatrix} p(\beta + \alpha\mu^n) + q\beta(1-\mu^n) \\ p\alpha(1-\mu^n) + q(\alpha + \beta\mu^n) \end{pmatrix}.$$

Таким образом,

$$p_1(n)=\frac{\beta}{\alpha+\beta}+\frac{p\alpha-(1-p)\beta}{\alpha+\beta}(1-\alpha-\beta)^n,$$

$$p_2(n)=1-p_1(n)=\frac{\alpha}{\alpha+\beta}-\frac{p\alpha-(1-p)\beta}{\alpha+\beta}(1-\alpha-\beta)^n.$$

На этом примере рассмотрим поведение матрицы $P(n)$ и $P(n)$ при больших значениях n, которое, очевидно, существенно зависит от величины $\mu=1-\alpha-\beta$. Так как $0\leqslant\alpha\leqslant1$, $0\leqslant\beta\leqslant1$, то $-1\leqslant\mu\leqslant1$

Разберем три случая:

1. $|\mu|<1$, тогда при $n\to\infty$, $\mu^n\to0$,

$$\lim_{n\to\infty}P(n)=\begin{pmatrix}\frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta}\\ \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta}\end{pmatrix},\quad \lim_{n\to\infty}P_n=\begin{pmatrix}\frac{\beta}{\alpha+\beta}\\ \frac{\alpha}{\alpha+\beta}\end{pmatrix},$$

т. е. в этом случае при $n\to\infty$ вероятность $p_{i,1}(n)$ имеет предел, который не зависит от i, а вероятность $p_1(n)$ имеет предел, который не зависит от начальных вероятностей и эти пределы равны. То же относится и к вероятностям $p_{i2}(n)$ и $p_2(n)$.

Пределы $\lim\limits_{n\to\infty}p_1(n)=\frac{\beta}{\alpha+\beta}$ и $\lim\limits_{n\to\infty}p_2(n)=\frac{\alpha}{\alpha+\beta}$ называются *финальными вероятностями* состояний E_1 и E_2 соответственно. Таким образом, если $|\mu|<1$, то при больших n вероятности $p_i(n)$ $(i=1,\ 2)$ практически не зависят от начального состояния системы и близки к финальным вероятностям.

Заметим, что можно так подобрать начальные вероятности p и $q=1-p$, чтобы вероятности $p_i(n)$ при всех n совпадали с финальными. Для этого должно выполняться очевидное равенство

$$p\alpha-(1-p)\beta=0$$

или

$$p=\frac{\beta}{\alpha+\beta},\ q=1-p=\frac{\alpha}{\alpha+\beta},$$

т. е. начальные вероятности состояний E_1 и E_2 должны совпадать с финальными. При выполнении этого условия цепь Маркова часто называют *стационарной*.

2. $\mu=1$. Это возможно только, если

$$\alpha=\beta=0$$

и матрица перехода является единичной матрицей

$$P=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}=E.$$

Тогда $\quad P(n)=P^n=E$

$$P_n=E\cdot P_0=\begin{pmatrix}p\\ 1-p\end{pmatrix}.$$

Таким образом, в этом случае пределы

$$\lim_{n\to\infty} p_2(n)=1-p \text{ и } \lim_{n\to\infty} p_1(n)=p$$

существуют, но они не совпадают, причем предел вероятностей $p_i(n)$ $(i=1,\ 2)$ существенно зависит от начальных вероятностей (совпадает с начальной вероятностью соответствующего состояния). Этот результат ясен и непосредственно, так как если матрица перехода является единичной, то это значит, что система с вероятностью, равной единице, сохраняет свое состояние.

3. $\mu=-1$. Это возможно только при условии $\alpha=\beta=1$.

В этом случае матрица перехода имеет вид

$$P=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}.$$

Это означает, что система на каждом шаге с вероятностью, равной единице, изменяет свое состояние.

Тогда

$$p_1(n)=\frac{1}{2}+\frac{2p-1}{2}(-1)^n,$$

$$p_2(n)=\frac{1}{2}-\frac{2p-1}{2}(-1)^n$$

и если $p\neq\frac{1}{2}$, то пределы $p_i(n)$ при $n\to\infty$ не существуют. Эти пределы существуют только, если $p=\frac{1}{2}$. Тогда

$$\lim_{n\to\infty} p_1(n)=\lim_{n\to\infty} p_2(n)=\frac{1}{2}.$$

Таким образом, при $k=2$ $\lim\limits_{n\to\infty} p_i(n)$ $(i=1,\ 2)$ существует и не зависит от начальных вероятностей тогда и только тогда, когда

$$|\mu|=|1-\alpha-\beta|<1.$$

Если финальные вероятности состояний существуют и не зависят от начальных вероятностей, то говорят, что

цепь Маркова обладает эргодическим свойством. Следовательно, при $k=2$ цепь Маркова эргодична только, если $|\mu|<1$.

Исследованию условий, при которых цепь Маркова при $k>2$ является эргодичной, посвящен следующий параграф.

§ 2. ФИНАЛЬНЫЕ ВЕРОЯТНОСТИ ДЛЯ ЦЕПЕЙ МАРКОВА С КОНЕЧНЫМ ЧИСЛОМ СОСТОЯНИЙ

1. Общие замечания. Рассмотрим, при каких условиях существуют финальные вероятности, т. е. при каких условиях существует предел при $n\to\infty$ матрицы

$$P_n = P'(n)\,P_0,$$

не зависящий от матрицы начальных вероятностей.

При решении этого вопроса полезным является следующее (довольно очевидное) утверждение.

Теорема. *Для существования финальных вероятностей необходимо и достаточно, чтобы существовал предел при $n\to\infty$ матрицы $P(n)$ перехода за n шагов при условии, что элементы предельной матрицы не зависят от первого индекса (т. е. все строки предельной матрицы одинаковы).*

Для доказательства достаточности допустим, что

$$\lim_{n\to\infty} p_{ij}(n) = p_j \text{ при всех } i \text{ и } j.$$

Это значит, что

$$P'(n)\to\begin{pmatrix} p_1 & p_1 & \dots & p_1 \\ p_2 & p_2 & \dots & p_2 \\ \vdots & \vdots & & \vdots \\ p_k & p_k & \dots & p_k \end{pmatrix},$$

но тогда существует предел

$$P_n\to\begin{pmatrix} p_1 & p_1 & \dots & p_1 \\ p_2 & p_2 & \dots & p_2 \\ \cdot & \cdot & \cdot & \cdot \\ p_k & p_k & \dots & p_k \end{pmatrix}\begin{pmatrix} p_{01} \\ p_{02} \\ \vdots \\ p_{0k} \end{pmatrix} =$$

$$=\begin{pmatrix} p_1(p_{01}+p_{02}+\dots+p_{0k}) \\ p_2(p_{01}+p_{02}+\dots+p_{0k}) \\ \cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot \\ p_k(p_{01}+p_{02}+\dots+p_{0k}) \end{pmatrix} = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_k \end{pmatrix},$$

(так как $p_{01}+p_{02}+\ldots+p_{0k}=1$).

Для доказательства необходимости предположим, что существуют финальные вероятности, т. е., что при $n\to\infty$

$$P_n=P'(n)\,P_0\to\begin{pmatrix}p_1\\p_2\\\vdots\\p_k\end{pmatrix}.$$

Так как $p_1, p_2, \ldots, p_k$ не должны зависеть от n при любых матрицах P_0, то элементы матрицы $P(n)$ при $n\to\infty$ должны быть независимыми от n. Если при этом $p_{ij}(n)\to p_{ij}$, т. е. пределы элементов этой матрицы зависят от первого индекса, то

$$\begin{pmatrix}p_1\\p_2\\\vdots\\p_k\end{pmatrix}=\begin{pmatrix}p_{01}p_{11}+p_{02}p_{21}+\ldots+p_{0k}p_{k1}\\p_{01}p_{12}+p_{02}p_{22}+\ldots+p_{0k}p_{k2}\\\cdots\cdots\cdots\cdots\cdots\\p_{02}p_{1k}+p_{02}p_{2k}+\ldots+p_{0k}p_{kk}\end{pmatrix}.$$

Пусть хотя бы для одного j найдутся такие i и s, что

$$p_{ij}\neq p_{sj},$$

тогда, выбирая начальные вероятности соответственно

$$p_{0i}=1;\ p_{0\nu}=0 \quad \text{при } \nu\neq i$$

и

$$p_{0s}=1;\ p_{0\nu}=0 \quad \text{при } \nu\neq s$$

получаем в первом случае, что

$$\lim_{n\to\infty} p_j(n)=p_{ij},$$

а во втором —

$$\lim_{n\to\infty} p_j(n)=p_{sj},$$

что противоречит условию о независимости

$$\lim_{n\to\infty} p_j(n)$$

от начальных вероятностей.

Для исследования условий, при которых существует предел матрицы

$$P(n)=P^n,$$

при $n \to \infty$ необходимо выяснить некоторые свойства характеристических чисел матрицы перехода P.

Прежде всего заметим, что так как элементами матрицы $P(n)$ являются вероятности $p_{ij}(n)$ перехода за n шагов, то выполняются условия

$$0 \leqslant p_{ij}(n) \leqslant 1,$$

$$\sum_{j=1}^{k} p_{ij}(n) = 1$$

и, значит, любая степень матрицы перехода P является стохастической матрицей.

Явное выражение для матрицы P^n можно получить, зная характеристические числа матрицы P. Как было показано в § I гл. 9, если λ_i $(i = 1, 2, \ldots s \leqslant k)$ — различные характеристические числа матрицы P, то матрица

$$P(n) = P^n$$

может быть представлена в виде

$$P^n = T^{-1} J^n T, \tag{12.4}$$

где T — матрица перехода к жорданову базису (не зависящая от n);

J — жорданова матрица, эквивалентная матрице P.

Напомним, что J^n — клеточная матрица, состоящая из клеток вида

$$J_i^n = \begin{pmatrix} \lambda_i^n, & n\lambda_i^{n-1}, & \frac{n(n-1)}{2}\lambda_i^{n-2}, \ldots 1\,0 \ldots 0 \\ 0, & \lambda_i^n, & n\lambda_i^{n-1}, \ldots\ldots 1 \ldots 0 \\ \ldots & \ldots & \ldots\ldots\ldots\ldots \\ 0, & 0, & 0, \ldots\ldots\ldots\ldots \lambda_i^n \end{pmatrix}. \tag{12.5}$$

2. Спектральные свойства стохастической матрицы. Необходимые для дальнейшего некоторые свойства характеристических чисел матрицы перехода P определяются следующими теоремами.

Теорема 1. *Характеристические числа матрицы перехода P не могут быть по модулю больше единицы.*

Действительно, если $|\lambda_i| > 1$, то из выражения (12.4) и (12.5) и определения предела матрицы (см. § 6 гл. 9) следует, что при неограниченном увеличении n элементы матрицы $P(n) = P^n$ будут неограниченно увеличиваться,

что невозможно, так как эта матрица — стохастическая и все ее элементы не превосходят единицы.

Теорема 2. *Если* $|\lambda_i|=1$, *то ранг матрицы* $P-\lambda_i E$ *равен* $k-k_i$, *где* k_i — *кратность корня* λ_i.

Действительно, если ранг матрицы $P-\lambda_i E$ отличен от $k-k_i$, то ему будет соответствовать хотя бы одна жорданова клетка порядка больше единицы и тогда согласно выражению (12.5) при неограниченном увеличении n элементы $n\lambda_i^{n-1}$, $\frac{n(n-1)}{2}\lambda_i^{n-2}$, ... будут неограниченно увеличиваться, что невозможно. Таким образом, все жордановы клетки, соответствующие корню λ_i, модуль которого равен единице, могут быть только первого порядка, так что в клеточной матрице J^n они образуют клетку вида

$$\begin{pmatrix} \lambda_i^n, & 0, & \dots & 0 \\ 0, & \lambda_i^n, & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0, & 0, & \dots & \lambda_i^n \end{pmatrix},$$

порядок которой равен кратности k_i этого корня.

Теорема 3. *Матрица перехода* P *всегда имеет характеристическое число, равное единице.*

Действительно, так как P — стохастическая матрица, то сумма элементов любой ее строки равна единице. Отсюда следует, что

$$PZ=Z,$$

где

$$Z=\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}.$$

Но это значит, что оператор, соответствующий матрице P, преобразует вектор

$$z(1, 1, 1, \dots, 1)\neq 0$$

в коллинеарный ему вектор с коэффициентом пропорциональности $\lambda=1$. Следовательно, матрица P имеет характеристическое число, равное единице, а вектор $z(1, 1, \dots, 1)$ является собственным вектором матрицы P, соответствующим этому характеристическому числу.

3. Правильные и регулярные цепи Маркова. Если матрица перехода P, кроме характеристического числа $\lambda = 1$ (которое всегда существует), не имеет других характеристических чисел, модуль которых равен единице, то соответствующую этой матрице цепь Маркова будем называть *правильной*. (Заметим, что характеристическое число $\lambda = 1$ может быть кратным.) Если, кроме того, $\lambda = 1$ является простым корнем характеристического уравнения матрицы P, то цепь Маркова будем называть регулярной. Для правильных цепей Маркова имеет место следующее утверждение.

Теорема. *Матрица*

$$P(n) = P^n$$

имеет предел при $n \to \infty$ *тогда и только тогда, когда цепь Маркова правильна.*

Действительно,

$$\lim_{n \to \infty} P^n = P^\infty$$

существует тогда и только тогда, когда существует предел функции

$$f(\lambda) = \lambda^n$$

на спектре матрицы P (см. гл. 9, § 5).

Для характеристических чисел λ_j матрицы P, модуль которых

$$|\lambda_j| < 1,$$

$$\lim_{n \to \infty} f(\lambda_j) = 0; \quad \lim_{n \to \infty} f'(\lambda_j) = 0, \quad \ldots \quad \lim_{n \to \infty} f^{(r)}(\lambda_j) = 0.$$

Что касается характеристических чисел λ_i, для которых

$$|\lambda_j| = 1,$$

то для них указанные пределы существуют только, если $\lambda_j = 1$. Действительно, из теоремы 2 жорданова клетка матрицы P^n, соответствующая корню λ_j ($|\lambda_j| = 1$), будет (λ_j^n).

Так как $\lambda_j = \cos\varphi + i\sin\varphi$, где $\varphi = \arg\lambda_j$, то $\lambda_j^n = \cos n\varphi + i\sin n\varphi$.
Поэтому $\lim\limits_{n \to \infty} \lambda_j^n$ существует только при $\lambda_j = 1$ (т. е. при $\varphi = 0$).

Следует, однако, иметь в виду, что предельная матрица

$$P^{\infty}$$

может иметь различные строки, а это значит, что хотя предел матрицы

$$P_n = P'(n) P_0$$

и существует, но он зависит от матрицы начальных вероятностей, т. е. что цепь Маркова не обладает эргодическим свойством.

Вопрос об эргодичности цепи Маркова решается следующим утверждением:

Теорема. *Для того чтобы цепь Маркова обладала эргодическим свойством, необходимо и достаточно, чтобы она была регулярной.*

Действительно, чтобы все строки матрицы P^{∞} были одинаковы, т. е. чтобы финальные вероятности не зависели от начальных вероятностей, необходимо и достаточно, чтобы ранг предельной матрицы P^{∞} (которая является стохастической матрицей) был равен единице. Но это имеет место тогда и только тогда, когда корень $\lambda = 1$ характеристического уравнения матрицы P является простым, т. е. когда его кратность равна единице. Эта теорема обычно называется эргодической теоремой для цепей Маркова. Рекомендуем читателю проверить две последние теоремы на рассмотренном выше примере цепи Маркова с двумя состояниями.

Рассмотрим условия, которым удовлетворяют предельные вероятности однородной правильной цепи Маркова.

В формуле $P_n = (P^n)' P_0$ перейдем к пределу при $n \to \infty$. Так как по условию предельные вероятности существуют, то получаем

$$P_{\infty} = (P^{\infty})' P_0,$$

где

$$P_{\infty} = \lim_{n \to \infty} P_n, \text{ а } P^{\infty} = \lim_{n \to \infty} P^n.$$

Далее перейдем к пределу при $n \to \infty$ в формуле

$$P' P_n = (P^{n+1})' P_0,$$

имеем

$$P' P_{\infty} = (P^{\infty})' P_0.$$

Из полученных соотношений следует, что

$$P'P_\infty = P_\infty, \text{ или } (P' - E)\, P_\infty = 0.$$

Значит, вектор P_∞ является собственным вектором матрицы P', соответствующим характеристическому числу $\lambda = 1$. Кроме того, элементы p_i матрицы P_∞ должны удовлетворять очевидному соотношению

$$\sum_{i=1}^{k} p_i = 1,$$

которое назовем *условием нормировки*.

Следствие. *Если цепь Маркова регулярна, то предельные (финальные) вероятности являются координатами собственного вектора матрицы P', соответствующего характеристическому числу $\lambda = 1$, и удовлетворяющего условию нормировки.* Эти финальные вероятности определяются из системы уравнений

$$\begin{aligned} &P'P_\infty = P_\infty, \\ &\sum_{i=1}^{k} p_i = 1. \end{aligned} \tag{12.6}$$

Причем в данном случае система (12.6) имеет единственное решение. Действительно ранг матрицы $P' - E$ равен $k - 1$ (так как $\lambda = 1$ теперь простой корень). Поэтому все решения системы (12.6) отличаются лишь множителем, который находится однозначно из условия нормировки.

Пример. На телефонную станцию в моменты времени t_0, $t_0 + \Delta t$, $t_0 + 2\Delta t$, ... поступают вызовы. Поступивший вызов занимает станцию на время Δt, $2\Delta t$, ... в зависимости от продолжительности разговора, которая является случайной величиной. Если в предыдущий отрезок времени вызов не был обслужен, то станция будет занята в последующий отрезок времени Δt с вероятностью $1 - p$ и будет свободна с вероятностью p. Если же в предыдущий отрезок времени Δt станция была свободна, то в последующий отрезок времени она окажется занята с вероятностью λ и останется свободной с вероятностью $1 - \lambda$. Определить финальную вероятность p_1 того, что станция будет занята. Или, как обычно говорят, найти вероятность p_1 занятости станции в стационарном состоянии.

Последовательность состояний станции (E_1 — станция занята, — E_2 — станция свободна) образует простую однородную цепь Мар-

кова с матрицей перехода

$$P=\begin{pmatrix}1-p & p\\ \lambda & 1-\lambda\end{pmatrix}.$$

Предположим дополнительно, что $p\neq 0$ и $\lambda\neq 0$. Так как характеристические числа матрицы P равны

$$\lambda_1=1 \text{ и } \lambda_2=1-p-\lambda,$$

то при $p\neq 0$ и $\lambda\neq 0$,

$$\lambda_2\neq 1$$

и, следовательно, цепь является регулярной. Но это значит, что в данном случае финальные вероятности существуют и определяются из системы уравнений

$$P'P_\infty=P_\infty,$$

$$p_1+p_2=1,$$

где

$$P=\begin{pmatrix}p_1\\ p_2\end{pmatrix}.$$

Тогда имеем

$$(1-p)p_1+\lambda p_2=p_1,$$

$$pp_1+(1-\lambda)p_2=p_2.$$

Отсюда

$$pp_1=\lambda p_2=\lambda(1-p_1)$$

или окончательно

$$p_1=\frac{\lambda}{p+\lambda},$$

$$p_2=\frac{p}{p+\lambda}.$$

§ 3. ЦЕПИ МАРКОВА С НЕПРЕРЫВНЫМ ВРЕМЕНЕМ

1. Определение цепи Маркова с непрерывным временем. При определении цепи Маркова с дискретным временем предполагалось, что смена состояний системы может происходить только в заранее фиксированные (заданные) моменты времени t_0, t_1, ... t_n ...

Если отказаться от этого допущения и считать возможным смену состояний в любой заранее неизвестный момент времени $t>t_0$, то мы приходим к цепи Маркова с непрерывным временем. Точнее, если вероятность $p_{ij}(s, t)$ того, что система, находясь в момент времени s в состоянии E_i перейдет в момент времени t в состояние E_j (для любых $t_0<s<t$), полностью определяется

ее состоянием в момент времени s и не изменяется от дополнительной информации относительно состояний системы в любые моменты времени $\tau < s$ (предшествующие времени s), то процесс смены состояний системы называется *простой цепью Маркова с непрерывным временем.* Очевидно, что если мы заранее произвольно выберем и зафиксируем моменты времени $t_0, t_1, \ldots, t_n, \ldots$, то мы получим рассмотренную ранее цепь Маркова с дискретным временем. Отсюда следует, что все соотношения, приведенные в предыдущих параграфах для цепей Маркова с дискретным временем справедливы и для цепей Маркова с непрерывным временем.

В частности, имеют место следующие основные равенства:

$$P(s, t) = P(s, \tau) P(\tau, t), \tag{12.7}$$

$$P(t) = P'(t_0, t) P(t_0), \tag{12.8}$$

где $P(s, t)$ — стохастическая матрица, элементами которой являются вероятности $p_{ij}(s, t)$, а $P(t)$ и $P(t_0)$ — матрицы-столбцы, составленные из вероятностей и $p_i(t)$ и $p_i(t_0)$;

$p_i(t)$ — вероятность того, что система в момент времени t находится в состоянии E_i;

$p_i(t_0)$ — начальные вероятности, т. е. вероятности того, что в начальный момент t_0 система находится в состоянии E_i.

Из соотношения (12.8) видно, что если известна матрица $P(s, t)$ для любых моментов времени $t_0 \leqslant s \leqslant t$ и матрица начальных вероятностей $P(t_0)$, то вероятности различных состояний в произвольный момент времени t находятся совершенно элементарно. Однако на практике обычно матрица $P(s, t)$ может быть задана лишь для достаточно близких моментов времени t и $t + \Delta t$. Другими словами, заранее известен характер матрицы $P(t, t + \Delta t)$ при достаточно малых Δt.

В ряде случаев этого оказывается достаточно для определения матрицы $P(s, t)$ для любых моментов времени. Сформулируем основные предположения о характере матрицы $P(t, t + \Delta t)$, которые в некотором смысле можно рассматривать как требования «непрерывности во времени» процесса смены состояний системы.

2. Основные допущения. Будем считать, что элементы матрицы $P(t, t + \Delta t)$ при $\Delta t \to 0$ и любом $t \geqslant t_0$ удовле-

творяют следующим условиям:

$$p_{i,i}(t,\ t+\Delta t)=1-c_i(t)\,\Delta t+o(\Delta t);$$
$$p_{i,j}(t,\ t+\Delta t)=c_i(t)\,q_{ij}(t)\,\Delta t+o(\Delta t), \qquad (12.9)$$

где $c_i(t)\geqslant 0$; $1\geqslant q_{ij}(t)\geqslant 0$ — непрерывные функции времени t, причем $\sum\limits_{i\neq j}^{k} q_{ij}(t)=1$;

$o(\Delta t)$ — бесконечно малые величины более высокого порядка, нежели Δt.

Если величины $c_i(t)$ и $q_{ij}(t)$ — постоянны (не зависят от t), то соответствующая цепь Маркова с непрерывным временем называется однородной. Эти допущения означают, что при достаточно малых Δt вероятность изменения состояния системы пропорциональна Δt (т. е. является величиной того же порядка малости, что и Δt).

Если учесть, что вероятность изменения состояния E_i в интервале времени $(t,\ t+\Delta t)$ при малом Δt приближенно равна $c_i(t)\,\Delta t$, то величину $q_{ij}(t)$ можно трактовать, как условную вероятность перехода системы из состояния E_i в состояние E_j (в указанном интервале) при условии, что имел место переход из состояния E_i в некоторое другое состояние.

Принятые упрощения о характере матрицы $P(t,\ t+\Delta t)$ при $\Delta t\to 0$ позволяют составить систему дифференциальных уравнений относительно элементов матрицы $P(s,\ t)$. Вывод этих уравнений принадлежит А. Н. Колмогорову, а сами уравнения называются *уравнениями Колмогорова.*

3. **Уравнения Колмогорова.** Согласно соотношениям (12.9) матрица $P(t,\ t+\Delta t)$ может быть представлена в виде

$$P(t,\ t+\Delta t)=E+R(t)\,\Delta t+A\Delta t, \qquad (12.10)$$

где E — единичная матрица,

$$R(t)=\begin{pmatrix} -c_1(t) & c_1(t)\,q_{12}(t) & \dots & c_1(t)\,q_{1k}(t) \\ c_2(t)\,q_{21}(t) & -c_2(t) & \dots & c_2(t)\,q_{2k}(t) \\ \dots & \dots & \dots & \dots \\ c_k(t)\,q_{k1}(t) & c_k(t)\,q_{k2}(t) & \dots & -c_k(t) \end{pmatrix}$$

$$A=(\alpha_{ij});\quad \alpha_{ij}=\alpha_{ij}(t,\ \Delta t)\to 0 \quad \text{при } \Delta t\to 0.$$

Для однородной цепи Маркова с непрерывным временем матрица $R(t)=\text{const}$ (не зависит от t).

Из основного уравнения Маркова (12.7) следует, что

$$P(t_0, t+\Delta t) = P(t_0, t)\, P(t, t+\Delta t).$$

Тогда, используя равенство (12.10), получаем

$$P(t_0, t+\Delta t) = P(t_0, t)\,[E + R(t)\,\Delta t + A\Delta t].$$

Откуда

$$\frac{P(t_0, t+\Delta t) - P(t_0, t)}{\Delta t} = P(t_0, t)\, R(t) + P(t_0, t)\, A.$$

Переходя к пределу и учитывая, что $\lim\limits_{\Delta t \to 0} A = 0$, убеждаемся в существовании предела правой части. Следовательно, существует также предел и левой части, который, очевидно, представляет собой

$$\frac{\partial P(t, t_0)}{\partial t} = P(t_0, t)\, R(t).$$

Так как величина t_0 является фиксированной постоянной, то

$$\frac{\partial P(t_0, t)}{\partial t} = \frac{dP(t, t_0)}{dt},$$

и мы приходим к системе k^2 однородных линейных дифференциальных уравнений относительно $p_{ij}(t_0, t)$ $(i, j = 1, 2, \ldots k)$, которая в матричной форме имеет вид

$$\frac{dP(t_0, t)}{dt} = P(t_0, t)\, R(t). \tag{12.11}$$

При $t = t_0$ решение системы (12.11) должно удовлетворять очевидному начальному условию (вытекающему из соотношений (12.10))

$$P(t_0, t_0) = E.$$

Решив систему дифференциальных уравнений (12.11), находим матрицу

$$P(t_0, t),$$

зная которую, можно найти матрицу $P(s, t)$ для любых $s < t$, используя равенство

$$P(t_0, t) = P(t_0, s)\, P(s, t),$$

откуда

$$P(s, t) = P^{-1}(t_0, s)\, P(t_0, t),$$

и матрицу

$$P(t) = P'(t_0, t)\, P(t_0).$$

Однако практически матрицу $P(t)$ удобно находить иначе. Транспонируем матричное равенство (12.11)

$$\frac{dP'(t_0, t)}{dt} = R'(t) P'(t_0, t)$$

и умножим обе его части справа на $P(t_0)$. Так как $P(t_0)$ не зависит от t, то получаем

$$\frac{d[P'(t_0, t) P(t_0)]}{dt} = R'(t) P'(t_0, t) P(t_0),$$

откуда

$$\frac{dP(t)}{dt} = R'(t) P(t). \qquad (12.12)$$

Соотношение (12.12) представляет собой матричную запись системы k линейных однородных дифференциальных уравнений относительно элементов матрицы $P(t)$. Решение этой системы должно удовлетворять очевидным начальным условиям

$$P(t)|_{t=t_0} = P(t_0).$$

Если цепь Маркова однородна ($R = \text{const}$), то решение системы (12.12) согласно результатам главы 11 имеет вид

$$P(t) = e^{R'(t-t_0)} P(t_0).$$

4. Финальная вероятность. При рассмотрении финальных вероятностей для однородных цепей Маркова с дискретным временем было показано, что вероятность $P_i(n)$ того, что система на n-м шаге находится в состоянии E_i, имеет предел при $n \to \infty$ тогда и только тогда, когда цепь является правильной. При этом в общем случае эти пределы могут зависеть от начальных вероятностей.

Как будет показано ниже, для однородных цепей Маркова с непрерывным временем аналогичные вероятности $p_i(t)$ имеют предел при $t \to \infty$, если выполнены основные допущения (12.9). Таким образом, выполнение этих допущений заменяет в непрерывном случае требование правильности цепи Маркова.

Полагая $R = \text{const}$ (цепь по условию однородна), докажем предварительно три вспомогательные теоремы.

Теорема 1. *Матрица R всегда имеет характеристическое число $\lambda = 0$.*

Действительно, так как

$$R=\begin{pmatrix} -c_1 & c_1q_{12} & \cdots & cq_{1k} \\ c_2q_{21} & -c_2 & \cdots & c_2q_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ c_kq_{k1} & c_kq_{k2} & \cdots & -c_k \end{pmatrix},$$

то

$$R\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}=0,$$

откуда следует, что $\lambda=0$ является характеристическим числом матрицы R.

Теорема 2. *Вещественные части характеристических чисел матрицы R либо отрицательны, либо равны нулю.* В последнем случае характеристическое число равно нулю.

Пусть

$$u=\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_k \end{pmatrix}$$

собственный вектор матрицы R, соответствующий характеристическому числу λ, а u_m — наибольшее по модулю из чисел $u_1, u_2, \ldots u_k$. Так как u — собственный вектор матрицы R, то

$$(R-\lambda E)\,u=0$$

и m-ый элемент матрицы-столбца, стоящей в левой части этого равенства, имеет вид

$$u_1c_mq_m+\ldots+u_{m-1}c_mq_{m-1,\,m}-$$
$$-u_m(c_m+\lambda+u_{m+1}c_mq_{m+1,\,m}+\ldots+u_kc_mq_{km})=0.$$

Отсюда

$$u_m(c_m+\lambda)=\sum_{i\neq m}^{k}u_ic_mq_{im},$$

т. е.

$$|u_m||c_m+\lambda|\leqslant c_m\sum_{i\neq m}^{k}|u_i|\,q_{im}\leqslant c_m|u_m|\sum_{i\neq m}^{k}q_{im}\leqslant c_m|u_m|,$$

так как

$$|t_m| \geqslant |t_i| \text{ и } \sum_{i \neq m}^{k} q_{im} = 1.$$

Таким образом,

$$|c_m + \lambda| \leqslant c_m \text{ и, полагая } \lambda = \alpha + i\beta,$$

имеем

$$\sqrt{(c_m + \alpha)^2 + \beta^2} \leqslant c_m.$$

Отсюда $\alpha \leqslant 0$, причем, если $\alpha = 0$, то $\beta = 0$.

Теорема 3. *Если характеристическое число $\lambda = 0$ имеет кратность k_0, то ранг матрицы R равен $k - k_0$.*

Доказательство этой теоремы аналогично доказательству теоремы 2, приведенной на стр. 221.

Следствие. Жорданова клетка, соответствующая характеристическому числу $\lambda = 0$ матрицы R', имеет вид

$$J_j = (0).$$

Поэтому

$$e^{J_j(t-t_0)} = E.$$

Докажем существование предельных вероятностей. Так как

$$P(t) = e^{R'(t-t_0)} P(t_0)$$

и

$$e^{R'(t-t_0)} = T^{-1} e^{J(t-t_0)} T,$$

где J — жорданова нормальная форма матрицы R', то

$$P(t) = (T^{-1} e^{J(t-t_0)} T)' P(t_0).$$

Все клетки матрицы $e^{J(t-t_0)}$, соответствующие характеристическим числам с отрицательными вещественными частями при $t \to \infty$, стремятся к нулевой клетке. Кроме того,

$$e^{J_j(t-t_0)} = E$$

при всех t, поэтому предел матрицы $e^{R'(t-t_0)}$ при $t \to \infty$ существует. Следовательно, существует и

$$\lim_{t \to \infty} P(t) = P_\infty.$$

Найдем условия, которым удовлетворяют элементы матрицы P_∞.

Переходя к пределу при $t \to \infty$ в системе дифференциальных уравнений (12.12) и учитывая, что $\frac{dP_\infty}{dt} = 0$,

получаем

$$0 = R'P_{\infty}.$$

Значит матрице P_{∞} соответствует собственный вектор матрицы R с характеристическим числом $\lambda = 0$.

Кроме того, очевидно, что

$$\sum_{i=1}^{k} p_i(\infty) = \sum_{i=1}^{k} p_i = 1.$$

Решим вопрос, когда предельные вероятности не зависят от начальных.

Обозначая

$$\lim_{t \to \infty} e^{R'(t-t_0)} = P^{\infty}, \quad P(t_0) = P_0,$$

получаем

$$P_{\infty} = P^{\infty} P_0.$$

Но, как было показано выше (п. 1 § 2), справедлив вывод, что *предельные вероятности не зависят от начальных тогда и только тогда, когда все строки матрицы P^{∞} одинаковы.* Последнее же возможно лишь, если корень $\lambda = 0$ имеет кратность 1. В этом случае цепь Маркова является эргодической и соответствующие финальные вероятности определяются из линейной системы алгебраических уравнений

$$R'P_{\infty} = 0,$$

$$\sum_{i=1}^{k} p_i = 1.$$

5. **Уравнения «рождения» и «гибели».** Рассмотрим важный для приложений частный случай системы (12,12), который получается, если мы в основных допущениях (12.10) положим

1) $c_i = \lambda_i + \mu_i$,

2) $q_{i,\,i+1} = \dfrac{\lambda_i}{\lambda_i + \mu_i}$; $q_{i,\,i-1} = \dfrac{\mu_i}{\lambda_i + \mu_i}$; $\lambda_k = \mu_1 = 0$,

3) $q_{i,\,j} = 0$ при $|i - j| > 1$.

Это значит, что в рассматриваемом случае за время Δt (с точностью до бесконечно малых величин более высокого порядка, чем Δt) невозможны переходы от состояния E_i к состоянию E_j, если $|i - j| > 1$. Тогда система

дифференциальных уравнений (12.12) принимает вид

$$\begin{cases} \dfrac{dp_1(t)}{dt} = -\lambda_1 p_1(t) + \mu_2 p_2(t), \\ \dfrac{dp_2(t)}{dt} = \lambda_1 p_1(t) - (\lambda_2 + \mu_2) p_2(t) + \mu_3 p_3(t), \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ \dfrac{dp_i(t)}{dt} = \lambda_{i-1} p_{i-1}(t) - (\lambda_i + \mu_i) p_i(t) + \mu_{i+1} p_{i+1}(t), \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ \dfrac{dp_{k-1}(t)}{dt} = \lambda_{k-2} p_{k-2}(t) - (\lambda_{k-1} + \mu_{k-1}) p_{k-1}(t) + \mu_k p_k(t), \\ \dfrac{dp_k(t)}{dt} = \lambda_{k-1} p_{k-1}(t) - \mu_k p_k(t). \end{cases}$$

Полученные уравнения называются уравнениями «рождения» и «гибели», что связано со следующей вероятностной трактовкой рассматриваемого процесса. Будем считать, что система находится в состоянии E_ν, если в некотором объеме находится ν частиц. Тогда приведенные выше условия означают, что за время Δt (с точностью до бесконечно малых более высокого порядка, чем Δt) в этот объем может добавиться одна частица с вероятностью $\lambda_\nu \Delta t$ («рождение») или исчезнуть одна частица с вероятностью $\mu_\nu \Delta t$ («гибель»).

Найдем теперь условия, при которых существует

$$\lim_{t \to \infty} p_i(t),$$

не зависящий от начальных вероятностей. Как было показано в п. 4 настоящего параграфа, для этого ранг матрицы

$$R = \begin{pmatrix} -\lambda_1 & \lambda_1 & 0 & \dots & 0 & 0 \\ \mu_2 & -\lambda_2 - \mu_2 & \lambda_2 & \dots & 0 & 0 \\ 0 & \mu_3 & -\lambda_3 - \mu_3 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -\lambda_{k-1}\mu_k & \lambda_{k-1} \\ 0 & 0 & 0 & \dots & \mu_k & -\mu_k \end{pmatrix}$$

должен быть равен $k-1$.

Производя элементарные преобразования, которые состоят в суммировании первых $i-1$ столбцов, и добав-

ления этой суммы к i столбцу, получаем

$$R \sim \begin{pmatrix} -\lambda_1, & 0, & 0 \ldots & 0 & 0 \\ \mu_2, & -\lambda_2, & 0 \ldots & 0 & 0 \\ \cdots & & & & \\ \cdots & & & & \\ 0 & 0 & 0 \ldots & -\lambda_{k-1} & 0 \\ 0 & 0 & 0 \ldots & \mu_k & 0 \end{pmatrix}.$$

Отсюда видно, что ранг матрицы R равен $k-1$, если, например, все $\lambda_i \neq 0$ $(i=1,2,\ldots k-1)$ или все $\mu_i \neq 0$ $(i=2,3,\ldots k)$.

Если процесс обладает эргодическим свойством, то финальные вероятности определяются однозначно из линейной системы алгебраических уравнений

$$R' P_\infty = 0$$

$$\sum_{i-1}^{k} p_i = 1.$$

Запишем эту систему подробнее

$$\begin{cases} -\lambda_1 p_1 + \mu_2 p_2 = 0, \\ \lambda_1 p - (\lambda_2 + \mu_2)\, p_2 + \mu_3 p_3 = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \lambda_{k-2} p_{k-2}, \; -(\lambda_{k-1} + \mu_{k-1})\, p_{k-1}, \; + \mu_k p_k, \; = 0, \\ \lambda_{k-1} p_{k-1}, \; - \mu_k p_k = 0, \\ p_1 + p_2 + \ldots + p_k = 1. \end{cases}$$

Будем считать для определенности, что все $\mu_k \neq 0$, тогда, применяя метод Гаусса, получаем

$$p_2 = \frac{\lambda_1}{\mu_2}\, p_1,$$

$$p_3 = \frac{\lambda_2}{\mu_3}\, p_2,$$

$$\ldots\ldots\ldots$$

$$p_k = \frac{\lambda_{k-1}}{\mu_k}\, p_{k-1}.$$

Отсюда непосредственно находим

$$p_i = \frac{\lambda_1 \lambda_2 \ldots \lambda_{i-1}}{\mu_2 \mu_3 \ldots \mu_i}\, p_1 \quad i = 2,3 \ldots k.$$

Подставляя эти соотношения в условие нормировки

$$\sum_{i=1}^{k} p_i = 1,$$

имеем

$$p_i + \sum_{i=2}^{k} \frac{\lambda_1\lambda_2, \ldots \lambda_{i-1}}{\mu_2\mu_3, \ldots \mu_i} p_1 = 1.$$

Следовательно,

$$p_1 = \frac{1}{1 + \sum\limits_{i=2}^{k} \frac{\lambda_1\lambda_2 \ldots \lambda_{i-1}}{\mu_2\mu_3 \ldots \mu_i}}.$$

И, значит,

$$p_i = \frac{\frac{\lambda_1\lambda_2 \ldots \lambda_{i-1}}{\mu_2\mu_3 \ldots \mu_i}}{1 + \sum\limits_{j=2}^{k} \frac{\lambda_1\lambda_2 \ldots \lambda_{j-1}}{\mu_2\mu_3 \ldots \mu_j}}.$$

В частности, если

$$\lambda_i = \lambda \qquad i = 1, 2, \ldots k-1$$

$$\mu_i = (i-1)\,\mu \quad i = 2, 3, \ldots k,$$

последние соотношения принимают вид

$$p_i = \frac{\frac{1}{(i-1)!}\left(\frac{\lambda}{\mu}\right)^{i-1}}{\sum\limits_{j=1}^{k} \frac{1}{(j-1)!}\left(\frac{\lambda}{\mu}\right)^{j-1}}.$$

Эти соотношения называются *формулами Эрланга*. Они используются при расчете необходимого числа каналов телефонной станции.

Поясним кратко содержание задачи Эрланга, в результате решения которой получаются формулы Эрланга.

Пусть на телефонную станцию поступают независимые вызовы, образующие пуассоновский поток, т. е. пусть вероятность поступления m вызовов за время t равна

$$\frac{(\lambda t)^m}{m!} e^{-\lambda t}.$$

Телефонная станция имеет $k-1$ канал. Поступивший вызов обслуживается свободным каналом, который он

занимает случайное время τ с показательным законом распределения

$$P(\tau < x) = 1 - e^{-\mu x}.$$

Если в момент поступления некоторого вызова все каналы заняты, то он теряется (т. е. этот вызов не обслуживается).

Задача Эрланга состоит в определении финальных вероятностей p_i того, что ровно $i - 1$ — линия окажется занятой *. Приведенные выше формулы Эрланга и решают эту задачу.

---

* Вывод уравнений «рождения» и «гибели», соответствующих задаче Эрланга, читатель может найти в учебниках по теории вероятностей, например, в книге Е. С. Вентцель «Теория вероятностей». М., Физматгиз, 1962.

ОТВЕТЫ

Глава 1

1. а) $\begin{pmatrix} 11 & -22 & 29 \\ 9 & -27 & 32 \\ 13 & -17 & 26 \end{pmatrix}$; б) $\begin{pmatrix} 2 & -1 & 5 & -1 \\ 1 & -3 & -10 & 7 \end{pmatrix}$; в) $\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$. **2.** $A^3 = \begin{pmatrix} 13 & -14 \\ 21 & -22 \end{pmatrix}$ **3.** а) $\begin{pmatrix} \cos 2\alpha & -\sin 2\alpha \\ \sin 2\alpha & \cos 2\alpha \end{pmatrix}$; б) $\begin{pmatrix} \cos 3\alpha & -\sin 3\alpha \\ \sin 3\alpha & \cos 3\alpha \end{pmatrix}$; в) $\begin{pmatrix} \cos n\alpha & -\sin n\alpha \\ \sin n\alpha & \cos n\alpha \end{pmatrix}$. **4.** $A^5 = \begin{pmatrix} 304 & -61 \\ 305 & -62 \end{pmatrix}$. **5.** $A^n = \begin{pmatrix} \lambda^n & n\lambda^{n-1} & \frac{n(n-1)}{2}\lambda^{n-2} \\ 0 & \lambda^n & n\lambda^{n-1} \\ 0 & 0 & \lambda^n \end{pmatrix}$. **7.** $\begin{pmatrix} 3197 & -1266 \\ 7385 & -2922 \end{pmatrix}$.

Указание. Использовать результаты задачи **6.**

8. а) $\begin{pmatrix} a & b \\ c & a \end{pmatrix}$, где a, b, c — любые числа, удовлетворяющие соотношению $a^2 + bc = 0$; б) $\pm E$ или $\begin{pmatrix} a & b \\ c & -a \end{pmatrix}$, где $a^2 + bc = 1$.

9. а) $\begin{pmatrix} 7 & -4 \\ -5 & 3 \end{pmatrix}$; б) $\begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}$; в) $\frac{1}{9}\begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{pmatrix}$.

Глава 2

1. $\begin{pmatrix} 2 & 1 & -1 \\ 0 & -2 & 3 \\ 6 & 1 & 2 \end{pmatrix} \sim \begin{pmatrix} 2 & 1 & -1 \\ 0 & -2 & 3 \\ 0 & -2 & 5 \end{pmatrix} \sim \begin{pmatrix} 2 & 1 & -1 \\ 0 & -2 & 3 \\ 0 & 0 & 2 \end{pmatrix} \sim \begin{pmatrix} 2 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \sim \begin{pmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

5. а) $A = \begin{pmatrix} 2 & 1 & -1 \\ 0 & -2 & 3 \\ 6 & 1 & 2 \end{pmatrix}$. Эта матрица в примере 1 приведена к единичной.

Произведем те же преобразования над $E=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}$

$$\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}\sim\begin{pmatrix}1&0&0\\0&1&0\\-3&0&1\end{pmatrix}\sim\begin{pmatrix}1&0&0\\0&1&0\\-3&-1&1\end{pmatrix}\sim\begin{pmatrix}-\frac{1}{2}&-\frac{1}{2}&\frac{1}{2}\\\frac{9}{2}&\frac{5}{2}&-\frac{3}{2}\\-3&-1&1\end{pmatrix}\sim$$

$$\sim\begin{pmatrix}\frac{7}{4}&\frac{3}{4}&-\frac{1}{4}\\\frac{9}{2}&\frac{5}{2}&-\frac{3}{2}\\-3&-1&1\end{pmatrix}\sim\begin{pmatrix}\frac{7}{8}&\frac{3}{8}&-\frac{1}{8}\\-\frac{9}{4}&-\frac{5}{4}&\frac{3}{4}\\-\frac{3}{2}&-\frac{1}{2}&\frac{1}{2}\end{pmatrix}$$

$$A^{-1}=\begin{pmatrix}\frac{7}{8}&\frac{3}{8}&-\frac{1}{8}\\-\frac{9}{4}&-\frac{5}{4}&\frac{3}{4}\\-\frac{3}{2}&-\frac{1}{2}&\frac{1}{2}\end{pmatrix};$$

б) $A^{-1}=\frac{1}{15}\begin{pmatrix}-7&3&-13&41\\-13&-3&8&16\\18&3&-3&6\\3&3&-3&6\end{pmatrix}$. **6.** а) $r=3$; б) $r=2$. **7.** $x_1=1$, $x_2=2$, $x_3=3$. **8.** $x_1=-2$, $x_2=0$, $x_3=1$, $x_4=-1$, $x_5=3$.

9. а) $x_1=\frac{9-x_3}{4}$, $x_2=\frac{3+x_3}{2}$; б) система несовместна; в) $x_3=\frac{196-108x_1+68x_2}{10}$, $x_4=\frac{-38+19x_1-19x_2}{10}$, $x_5=\frac{-56+13x_1-13x_2}{10}$. **10.** $\lambda_1=1$, $\lambda_2=2$, $\lambda_3=3$.

11. При $\lambda=-2$ система несовместна.

При $\lambda=1$ система имеет бесчисленное множество решений ($x_1=1-x_2-x_3$).

При $\lambda\neq1$, $\lambda\neq-2$ система имеет единственное решение

$$x_1=x_2=x_3=\frac{1}{\lambda+2}.$$

12. Общее решение $\begin{aligned}x_1&=8x_3-7x_4\\x_2&=-6x_3+5x_4.\end{aligned}$

Фундаментальная система $\begin{pmatrix}8\\-6\\1\\0\end{pmatrix}$, $\begin{pmatrix}-7\\5\\0\\1\end{pmatrix}$.

13. Указание. Обозначим столбцы матрицы A, A_1, A_2 ... A_n, а столбцы матрицы E — с теми же номерами E_1, E_2, ... E_n, тогда,

если X — решение системы, то

$$A_1X_1 + A_2X_2 + \ldots + A_nX_n = 0,$$

такая же линейная комбинация столбцов матрицы E дает

$$E_1X_1 + E_2X_2 + \ldots + E_nX_n = X.$$

Легко показать, что сказанное справедливо и в случае, когда число неизвестных не равно числу уравнений. **14.** Например: $x_1 = 1$; $x_2 = -1$; $x_3 = 1$; $x_4 = 1$.

Указание:

$$\begin{pmatrix} 1 & 2 & 4 & -3 \\ 3 & 5 & 6 & -4 \\ 4 & 5 & -2 & 3 \\ 3 & 8 & 24 & -19 \end{pmatrix} \sim \begin{pmatrix} 0 & 2 & 4 & -3 \\ 0 & 5 & 6 & -4 \\ 0 & 5 & -2 & 3 \\ 0 & 8 & 24 & -19 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}.$$

К элементам 1-го столбца добавим соответствующие элементы 2-го столбца, умноженные на —1, и соответствующие элементы 3-го и 4-го столбцов.

Глава 3

1. а) нет; б) да; в) нет. **3.** Нет. **4.** Нет. **5.** а) все пространство; б) множество точек (векторов) плоскости, проходящей через начало координат; в) множество точек прямой, проходящей через начало координат; г) начало координат. **6.** Да. **7.** а) нет, так как ранг матрицы $\begin{pmatrix} 1 & 3 \\ 2 & 6 \\ 3 & 7 \end{pmatrix}$ равен двум; б) да, так как ранг матрицы $\begin{pmatrix} 5 & 3 & 8 \\ 4 & 3 & 1 \\ 3 & 2 & 3 \end{pmatrix}$ равен двум. **8.** а) Да, так как векторы (многочлены) $1 + t^2$, $1 - t^2$ и $1 + 2t$ линейно независимые; б) нет. **9.** В задаче 13 из гл. 2 было показано, что столбцы G, соответствующие нулевым столбцам E, будут решениями системы. Так как пространство решений имеет размерность $n - r$, то остается показать, что рассматриваемые столбцы линейно независимые. Последнее следует из того, что матрица G невырождена (она эквивалентна E), и потому ее столбцы линейно независимые. **10.** Решение. Вместо того, чтобы производить элементарные преобразования над столбцами матрицы A, а затем E, удобно производить эти преобразования над матрицей $\begin{pmatrix} E \\ A \end{pmatrix}$. Имеем

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \\ 3 & 1 & -1 & 3 \\ 0 & 1 & -1 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix},$$

отсюда видим, что ранг матрицы системы равен 2, а фундаментальная система решений имеет вид:

$$\begin{pmatrix}0\\1\\1\\0\end{pmatrix},\ \begin{pmatrix}-1\\0\\0\\1\end{pmatrix}.$$

11. Смотреть решение примера 5. а) Размерность 3. Базис образуют любые 3 некомпланарных вектора; б) размерность 2. Базис образуют любые 2 неколлинеарных вектора плоскости; в) размерность 1. Базис — любой ненулевой вектор этой прямой. **12.** n^2. **13.** $x=(e_1+e_2+e_3)$ иначе: $x=\{1, 1, 1\}$. **14.** Укажем одно из возможных решений задачи. Пусть например $a_{i1}\neq 0$. Тогда, применяя метод Гаусса с помощью 1-го столбца, образуем нули в ν-ой строке.

С помощью ненулевого элемента следующего столбца образуем нули в соответствующей ему строке и т. д. В результате получим матрицу $B\infty A$, у которой будет r ненулевых столбцов, а остальные столбцы — нулевые.

Столбцы A, соответствующие (т. е. имеющие те же номера) ненулевым столбцам B, будут искомыми. **15.** Из координат данных векторов составим матрицу. Задача сводится к нахождению ее линейно независимых столбцов, что сделано в **14**. **16.** Используя результаты **15**, имеем:

$$\begin{pmatrix}1&0&2&2&1\\1&0&2&-1&2\\1&1&3&0&1\end{pmatrix}\infty\begin{pmatrix}1&0&0&0&0\\1&1&0&-2&1\\1&1&1&-1&0\end{pmatrix}\infty\begin{pmatrix}1&0&0&0&0\\1&0&0&-2&1\\1&1&0&0&0\end{pmatrix}\infty$$
$$\begin{pmatrix}1&0&0&0&0\\1&0&0&-2&0\\1&1&0&0&0\end{pmatrix},$$

следовательно, $r=3$ и векторы x_1, x_2, x_4 линейно независимы.

Глава 4

1. $\tilde{A}(\lambda x)=(\lambda x_2+\lambda x_3,\ 2\lambda x_1+\lambda x_3,\ 3\lambda x_1-\lambda x_2+\lambda x_3)=\lambda(x_2+x_3,\ 2x_1+x_3,\ 3x_1-x_2+x_3)=\lambda\tilde{A}x$; $\tilde{A}(x+y)=(x_2+y_2+x_3+y_3,\ 2x_1+2y_1+x_3+y_3,\ 3x_1+3y_1-x_2-y_2+x_3+y_3)=(x_2+x_3,\ 2x_1+x_3,\ 3x_1-x_2+x_3)+(y_2+y_3,\ 2y_1+y_3,\ 3y_1-y_2+y_3)=\tilde{A}x+\tilde{A}y$.

Следовательно, $\tilde{A}$ — линейный оператор

$\tilde{A}e_1=2e_2+3e_3$,

$\tilde{A}e_2=e_1-e_3$,

$\tilde{A}e_3=e_1+e_2+e_3$.

Значит, $A=\begin{pmatrix}0&1&1\\2&0&1\\3&-1&1\end{pmatrix}$ **2.** $\begin{pmatrix}1&0&0\\0&0&0\\0&0&0\end{pmatrix}$. **3.** $\begin{pmatrix}a&b&0&0\\c&d&0&0\\0&0&a&b\\0&0&c&d\end{pmatrix}$. **4.** $\begin{pmatrix}2&-11&6\\1&-7&4\\2&-1&0\end{pmatrix}$.

5. $X=TX^*$, где $T=\begin{pmatrix}2&1\\3&1\end{pmatrix}$.

Следовательно: $x_1 = 2x_1^* + x_2^*$,
$x_2 = 3x_1^* + x_3^*$.

6. $T = \begin{pmatrix} 1 & -1 & \frac{1}{2} \\ 0 & 1 & -1 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}$,

а) $T = \begin{pmatrix} \cos\alpha & \cos\alpha_1 & \cos\alpha_2 \\ \cos\beta & \cos\beta_1 & \cos\beta_2 \\ \cos\gamma & \cos\gamma_1 & \cos\gamma_2 \end{pmatrix}$; б) Очевидно операцию поворота можно заменить тремя последовательными вращениями старой системы координат. 1. Поворот относительно оси z на угол φ. Соответствующая матрица

$T_1 = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}$. 2. Поворот относительно оси l на угол Θ. Соответствующая матрица

$T_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\Theta & -\sin\Theta \\ 0 & \sin\Theta & \cos\Theta \end{pmatrix}$. 3. Поворот относительно оси z' на угол ψ. Соответствующая матрица

$T_3 = \begin{pmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Поэтому матрица результирующего преобразования $T = T_1 T_2 T_3$ или

$$\begin{pmatrix} \cos\varphi\cos\psi - \sin\varphi\sin\psi\cos\Theta, & -\cos\varphi\sin\psi - \sin\varphi\cos\psi\cos\Theta, & \sin\varphi\cos\Theta \\ \sin\varphi\cos\psi + \cos\psi\sin\psi\cos\Theta, & -\sin\varphi\sin\psi + \cos\psi\cos\psi\cos\Theta, & -\cos\varphi\sin\Theta \\ \sin\psi\sin\Theta, & \cos\psi\cos\Theta, & \cos\Theta \end{pmatrix}.$$

8. а) $\begin{pmatrix} 1 & 0 & 2 & 1 \\ 2 & 3 & 5 & 1 \\ 3 & -1 & 0 & 2 \\ 1 & 1 & 2 & 3 \end{pmatrix}$ б) $\begin{pmatrix} -2 & 0 & 1 & 0 \\ 1 & -4 & -8 & -7 \\ 1 & 4 & 6 & 4 \\ 1 & 3 & 4 & 7 \end{pmatrix}$. **9.** $\begin{pmatrix} 1 & 2 & 2 \\ 3 & -1 & -2 \\ 2 & -3 & 1 \end{pmatrix}$

10. Пусть $f(t) = at^2 + bt + c$, $\tilde{A} = \frac{d}{dt}$, $\tilde{A}f(t) = \frac{d}{dt}(at^2 + bt + c) = 2at + b$. а) $\frac{d}{dt}1 = 0$, $\frac{d}{dt}t = 1$, $\frac{d}{dt}t^2 = 2t$, $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$; б) $B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$; в) $B = T^{-1}AT$, где $T = \begin{pmatrix} 1 & -1 & \frac{1}{2} \\ 0 & 1 & -1 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}$. **11.** Линейность

оператора следует из свойств векторного произведения

$$A=\begin{pmatrix} 1 & -z_3 & z_2 \\ z_3 & 0 & -z_1 \\ -z_2 & z_1 & 0 \end{pmatrix}.$$

Глава 5

1. а) векторы, коллинеарные вектору $\{1, 1, 1\}$ и векторы вида $\alpha_1\{1, 1, 0\}+\alpha_2\{1, 0, -3\}$; б) векторы, коллинеарные вектору $\{1, 1, -1\}$; в) векторы, коллинеарные вектору $\{3, 5, 6\}$, и векторы, линейно зависимые с векторами $\{2, 1, 0\}$ и $\{1, 0, -1\}$. **2.** Матрицы примера 1 а и в. **3.** а) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ в) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$. **4.** Простейший вид: $\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix}$, преобразующая матрица $T=\begin{pmatrix} 1 & 1 & 1 & -1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$,

5. $\begin{pmatrix} 1 & & & O \\ & 2 & & \\ & & 3 & \\ & & & \ddots \\ O & & & & n \end{pmatrix}$. **6.** а) В поле вещественных чисел подобна матрице $\begin{pmatrix} 2 & 0 & 0 \\ 0 & \sqrt{3} & 0 \\ 0 & 0 & -\sqrt{3} \end{pmatrix}$; б) В поле комплексных чисел подобна матрице $\begin{pmatrix} -1 & 0 & 0 \\ 0 & 2+3i & 0 \\ 0 & 0 & 2-3i \end{pmatrix}$; в) В поле рациональных чисел подобна матрице $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$; г) Не подобна диагональной матрице ни в каком поле (так как в поле комплексных чисел не имеет простой структуры).

Глава 6

1. $x_1x_2=1\cdot 1+1\cdot 2+1\cdot 3-2\cdot 3=0$, $x_3=(\alpha, \beta, \gamma, 0)$, $x_3x_1=0$, $x_3x_2=0$, $\left.\begin{matrix} \alpha+\beta+\gamma=0 \\ \alpha+2\beta-3\gamma=0 \end{matrix}\right|$, следовательно, $x_3=(1, -2, 1, 0)$, $x_4=(\alpha, \beta, \gamma, \delta)$, $x_4x_1=0$; $x_4x_2=0$; $x_4x_3=0$, $\left.\begin{matrix} \alpha+\beta+\gamma+2\delta=0 \\ \alpha+\beta+3\gamma-3\delta=0 \\ \alpha-2\beta+\gamma=0 \end{matrix}\right|$, отсюда $x_4=(-25, -4, 17, 6)$. **2.** $(2, 1, 3, -1)$, $(3, 2, -3, -1)$, $(1, 5, 1, 10)$.

3. а) $e=\left(\frac{2}{3}, \frac{2}{3}, \frac{1}{3}\right)$; $e_2=\left(\frac{2}{3}, -\frac{1}{3}, -\frac{2}{3}\right)$, $e_3=\left(\frac{1}{3}, -\frac{2}{3}, \frac{2}{3}\right)$;
б) $\begin{pmatrix} 9 & 0 & 0 \\ 0 & 18 & 0 \\ 0 & 0 & -9 \end{pmatrix}$. **5.** Диагональные элементы равны ± 1. **8.** Если диагональные элементы равны по модулю 1. **9.** Пусть x — собственный вектор с собственным числом λ. Из инвариантности скалярного произведения следует $(\tilde{A}x, \tilde{A}x)=(x, x)$, или $(\lambda x, \lambda x)=\lambda^2(x, x)$ и $\lambda^2=1$.

Глава 7

1. $4y_1^2+4y_2^2-2y_3^2$ **2.** $3y_1^2+6y_2^2+9y_3^2$, $T=\begin{pmatrix} \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \end{pmatrix}$.

4. а) $4x_1^2+8y_1^2-2z_1^2-5=0$; б) $x_1^2+2y_1^2-2=0$; в) $2x_1^2+5y_1^2-5\sqrt{2}\,z_1=0$. **5.** При $\lambda>2$.

Глава 8

1. а) $\begin{pmatrix} 3 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}$; б) $\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$; в) $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$. **2.** $I=\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$,
$T=\begin{pmatrix} 0 & -2 & 0 \\ 1 & -3 & 0 \\ -3 & -3 & -1 \end{pmatrix}$ **4.** Указание. Воспользоваться результатами задачи **3.** **5.** Подобны. **6.** Подобны. $T=\begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix}$. **7.** $\begin{pmatrix} \lambda^2 & 1 & 0 \\ 0 & \lambda^2 & 1 \\ 0 & 0 & \lambda^2 \end{pmatrix}$.

8. $I^k=\begin{pmatrix} \lambda^k & k\lambda^{k-1} & \frac{k(k-1)}{2!}\lambda^{k-2} & \dots C_k^n\lambda^{k-n} \\ 0 & \lambda^k & k\lambda^{k-1} & \dots C_k^{n-1}\lambda^{k-n+1} \\ \dots & & & \\ 0 & 0 & 0 & \dots \lambda^k \end{pmatrix}$. **9.** $T^{-1}AT=$

$=\begin{pmatrix} T_1^{-1}A_1T_1 & & & O \\ & T_2^{-1}A_2T_2 & & \\ & & \ddots & \\ O & & & T_n^{-1}A_nT_n \end{pmatrix}$, где преобразующая матрица

$T=\begin{pmatrix} T_1 & & O \\ & T_2 & \\ & & \ddots \\ O & & T_n \end{pmatrix}$, $I_i=T_i^{-1}A_iT_i$.

Глава 9

1. $\begin{pmatrix} 3197 & -1266 \\ 7385 & -922 \end{pmatrix}$. **2.** $\begin{pmatrix} 190 & 189 & -189 \\ 126 & 127 & -126 \\ 252 & 252 & -251 \end{pmatrix}$. **3.** $f(A) = 2A$.

5. а) $e^A = \begin{pmatrix} 3e-2, & 3e-3, & -3e+3 \\ 2e-2, & 2e-1, & -2e+2 \\ 4e-4, & 4e-4, & -4e+5 \end{pmatrix} e$; б) $\ln A = \begin{pmatrix} 3 & 3 & -3 \\ 2 & 2 & -2 \\ 4 & 4 & -4 \end{pmatrix} \ln 2$;

в) $\sin Ax = \begin{pmatrix} 3\sin 2x - 2\sin x & 3\sin 2x - 3\sin x & -3\sin 2x + 3\sin x \\ 2\sin 2x - 2\sin x & 2\sin 2x - \sin x & -2\sin 2x + 2\sin x \\ 4\sin 2x - 4\sin x & 4\sin 2x - 4\sin x & -4\sin 2x + 5\sin x \end{pmatrix}$.

7. Воспользоваться решением задачи **9** гл. 8.

Глава 10

1. $A = \begin{pmatrix} -7 & 1 \\ -2 & -5 \end{pmatrix}$, $\begin{vmatrix} -7-\lambda & 1 \\ -2 & -5-\lambda \end{vmatrix} = 0$, $\lambda_{1,2} = -6 \pm i$,

$$f(A) = \frac{A - \lambda_1 E}{\lambda_2 - \lambda_1} f(\lambda_2) + \frac{A - \lambda_2 E}{\lambda_1 - \lambda_2} f(\lambda_1),$$

$$e^{At} = \frac{\begin{pmatrix} -1-i & 1 \\ -2 & 1-i \end{pmatrix}}{2i} e^{(-6-i)t} + \frac{\begin{pmatrix} -1+i & 1 \\ -2 & 1+i \end{pmatrix}}{-2i} e^{(-6+i)t}.$$

Используя формулы Эйлера, получаем

$$e^{At} = e^{-6t} \begin{pmatrix} \sin t - \cos t & -\sin t \\ 2\sin t & -\sin t - \cos t \end{pmatrix}, \quad X = e^{At} C, \quad \text{где } C = \begin{pmatrix} C_1 \\ C_2 \end{pmatrix}.$$

Откуда $x_1 = e^{-6t}[(c_1 - c_2)\sin t - c_1 \cos t]$ $x_2 = e^{-6t}[(2c_1 - c_2)\sin t - c_2 \cos t]$. **2.** $y_1 = -e^{-x}$, $y_2 = e^{-x}$, $y_3 = 0$. **3.** $y_1 = \frac{1}{2}(\cos x \operatorname{ch} x + \sin x \operatorname{ch} x)$, $y_2 = \frac{1}{2}(\sin x \operatorname{ch} x - \cos x \operatorname{sh} x)$. Указание. Обозначить $y_1' = z_1$, $y_2' = z_2$ и решить систему четырех уравнений. **4.** $y = c_1 e^x + c_2 e^{-x} + c_3 \cos x + c_4 \sin x$. **5.** $x_1 = \frac{7}{40} e^t + \frac{1}{5} e^{-2t} - \frac{7}{15} e^{-4t} + \frac{11}{120} e^{-7t}$, $x_2 = \frac{1}{40} e^t + \frac{3}{10} e^{-2t} - \frac{7}{30} e^{-4t} - \frac{11}{120} e^{-7t}$. **6.** $y = \int_0^x t e^{-2t} f(x-t)\, dt = \int_0^x (x-t) e^{-2(x-t)} f(t)\, dt$.

ЗАДАЧИ И УПРАЖНЕНИЯ

1) Вычислить определитель

$$\Delta = \begin{vmatrix} 1 & -1 & 2 \\ 3 & -1 & 1 \\ 2 & 1 & 4 \end{vmatrix}.$$

Отв. $\Delta = 15$.

Решение: укажем несколько способов, основанных на использовании свойств определителей.

а) Согласно определению имеем

$$\Delta = 1\cdot(-1)\cdot 4 + (-1)\cdot 1\cdot 2 + 3\cdot 1\cdot 2 - 2(-1)2 - 1\cdot 1\cdot 1 - \\ - 3\cdot(-1)4 = 15.$$

б) Используя свойство 9, разложим определитель по элементам первой строки:

$$\Delta = 1\cdot\begin{vmatrix} -1 & 1 \\ 1 & 4 \end{vmatrix} - (-1)\begin{vmatrix} 3 & 1 \\ 2 & 4 \end{vmatrix} + 2\begin{vmatrix} 3 & -1 \\ 2 & 1 \end{vmatrix} = \\ = -5 + 10 + 10 = 15.$$

в) Используя свойство 7, прибавим к элементам первой и второй строк соответствующие элементы третьей строки. Получим

$$\Delta = \begin{vmatrix} 3 & 0 & 6 \\ 5 & 0 & 5 \\ 2 & 1 & 4 \end{vmatrix} = -1\begin{vmatrix} 3 & 6 \\ 5 & 5 \end{vmatrix} = 15.$$

2) Вычислить определители

а) $\begin{vmatrix} a & 1 & 0 & -1 \\ -1 & a & 0 & 1 \\ 1 & -1 & 0 & a \\ 1 & 2 & 3 & 4 \end{vmatrix}$; б) $\begin{vmatrix} 2 & 1 & -1 & 1 \\ 1 & 3 & 3 & 3 \\ 0 & 1 & -2 & 2 \\ -4 & -2 & 2 & -2 \end{vmatrix}$;

в) $\Delta_n = \begin{vmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{nn} \end{vmatrix}$.

Отв. а) $-3a^2(a+3)$, б) 0 (см. свойство 5), в) $= \Delta_n = = a_{11}a_{22}\dots a_{nn}$.

3) Вычислить определитель Вандермонда

$$\Delta_n = \begin{vmatrix} 1 & 1 & \dots & 1 \\ k_1 & k_2 & \dots & k_n \\ k_1^2 & k_2^2 & \dots & k_n^2 \\ \dots & \dots & \dots & \dots \\ k_1^{n-1} & k_2^{n-1} & \dots & k_n^{n-1} \end{vmatrix}.$$

Отв. $\Delta_n = (k_2 - k_1)(k_3 - k_1)\dots(k_n - k_1)(k_3 - k_2)\dots(k_n - k_3)\dots$ $\dots(k_n - k_{n-1} = \prod_{j<i}(k_i - k_j)$.

4) Не вычисляя определителя, решить уравнение

$$\begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 1-x & 1 & \dots & 1 \\ 1 & 1 & 2-x & \dots & 1 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & 1 & \dots & n-x \end{vmatrix}.$$

Отв. $x = 0, 1, 2, \dots, (n-1)$.

5) Вычислить

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^n.$$

Отв. $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ при $n = 4k$; $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ при $n = 4k+1$; $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ при $n = 4k+2$; $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ при $n = 4k+3$.

6) Доказать, что если матрицы A и B перестановочны, то

а) $A^2 - B^2 = (A-B)(A+B)$;

б) $A^{-1}B = BA^{-1}$.

7) Найти обратную матрицу для треугольной матрицы

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{pmatrix}.$$

Отв. Если $a_{11}a_{22}a_{33} \neq 0$, то

$$A^{-1} = \begin{pmatrix} \frac{1}{a_{11}} & -\frac{a_{12}}{a_{12}a_{12}} & \frac{a_{12}a_{23} - a_{13}a_{22}}{a_{11}a_{22}a_{33}} \\ 0 & \frac{1}{a_{22}} & -\frac{a_{23}}{a_{22}k_{33}} \\ 0 & 0 & \frac{1}{a_{33}} \end{pmatrix}.$$

Если $a_{11}a_{22}a_{33} = 0$, то обратная матрица не существует.

8) Найти обратную матрицу для матрицы

$$A=\begin{pmatrix}0 & 1 & 1 \dots 1\\ 1 & 0 & 1 \dots 1\\ \dots & \dots & \dots\\ 1 & 1 & 1 \dots 0\end{pmatrix}.$$

Отв.

$$A^{-1}=\frac{1}{n-1}\begin{pmatrix}-n+2 & 1 & 1 \dots & 1\\ 1 & -n+2 & 1 \dots & 1\\ \dots & \dots & \dots & \dots\\ 1 & 1 & 1 \dots & -n+2\end{pmatrix}.$$

9) Подобрать λ так, чтобы система уравнений имела решение

$$2x_1-x_2+x_3+x_4=1,\quad x_1+2x_2-x_3+4x_4=2,$$
$$x_1+7x_2-4x_3+11x_4=\lambda.$$

Отв. $\lambda=5$.

10) Известно, что система

$$ax_1+ax_3=b,\quad ax_2+bx_1=c,\quad bx_3+cx_2=a.$$

Имеет единственное решение. Доказать, что $a\cdot b\cdot c\neq 0$ и найти решение.

Отв. $x_1=\dfrac{b^2+c^2-a^2}{2bc}$; $x_2=\dfrac{a^2+c^2-b^2}{2ac}$; $x_3=\dfrac{a^2+b^2-c^2}{2ab}$.

11) Исследовать систему при различных значениях λ

$$x_1+\lambda x_2+x_3=1,\quad 2x_1+x_2+x_3=\lambda,\quad x_1+x_2+\lambda x_3=\lambda^2.$$

Отв. Если $\lambda\neq 1$ и $\lambda\neq -2$, то система имеет единственное решение. Если $\lambda=1$, то система неопределенная и имеет два свободных неизвестных. Если $\lambda=2$, то система несовместна.

12) При каком условии n точек плоскости $M_1(x_1, y_1)$, $M_2(x_2, y_2)$ ... $M_n(x_n, y_n)$ лежат на одной прямой.

Отв. Если ранг матрицы

$$\begin{pmatrix}x_1 & y_1 & 1\\ x_2 & y_2 & 1\\ \dots & \dots & \dots\\ x_n & y_n & 1\end{pmatrix}$$

меньше единицы.

13) Даны два линейных преобразования

$$x_1=2y_1+y_2,\quad x_2=y_1-2y_2;\qquad \text{и}\qquad z_1=2x_1+x_2,\quad z_2=x_2+3x_2.$$

Выразить переменные z_1 и z_2 через y_1 и y_2.

Отв. $z_1=5y_1$, $z_2=5y_1-5y_2$.

14) Дано линейное преобразование

$$y_1=x_2+x_3+x_4,\quad y_2=x_1+x_3+x_4,$$
$$y_3=x_1+x_2+x_4,\quad y_4=x_1+x_2+x_3.$$

Найти обратное преобразование.

Отв.

$$x_1 = -\frac{2}{3}y_1 + \frac{1}{3}y_2 + \frac{1}{3}y_3 + \frac{1}{3}y_4,\ x_2 = \frac{1}{3}y_1 - \frac{2}{3}y_2 + \frac{1}{3}y_3 + \frac{1}{3}y_4,$$

$$x_3 = \frac{1}{3}y_1 + \frac{1}{3}y_2 - \frac{2}{3}y_3 + \frac{1}{3}y_4,\ x_4 = \frac{1}{3}y_1 + \frac{1}{3}y_2 + \frac{1}{3}y_3 - \frac{2}{2}y_4.$$

15) Доказать, что множество функций вида

$$f(t) = a_1 e^t + a_2 e^{2t} + \ldots + a_n e^{nt},$$

где $a_1, a_2, \ldots, a_n$ — произвольные вещественные числа; образует линейное пространство. Найти его размерность и базис.

Отв. Пространство n-мерное, его базис образуют, например, функции $e^t, e^{2t}, \ldots, e^{nt}$.

16) Доказать, что линейное пространство $C(a, b)$ непрерывных на отрезке $[a, b]$ функций — бесконечномерное.

Указание. Воспользоваться результатами задачи 15.

17) Доказать, что если

$$B = PAQ,$$

где A, B, P, Q — квадратные матрицы одного и того же порядка и P и Q — невырожденные, то ранги матриц A и B одинаковы.

Указание. Воспользоваться теоремой о ранге матрицы линейного оператора.

18) Доказать, что каждая из двух систем векторов

$$e_1 = (1, 1, 1, 1);\ e_2 = (1, 2, 1, 1),\ e_3 = (1, 1, 2, 1),\ e_4 = (1, 3, 2, 3);$$

$$e_1^* = (1, 0, 3, 3);\ e_2^* = (-2, -3, -5, -4);\ e_3^* = (2, 2, 5, 4),$$

$$e_4^* = (-3, 3, -4, -4)$$

является базисом и найти матрицу перехода от одного базиса к другому.

Отв.

$$T = \begin{pmatrix} 2 & 0 & 1 & -1 \\ -3 & 1 & -2 & 1 \\ 1 & -2 & 2 & -1 \\ 1 & -1 & 1 & -1 \end{pmatrix}.$$

19) Найти координаты вектора

$$x = f(t) = a_0 + a_1 t + \ldots + a_n t^n$$

пространства многочленов степени не выше n в базисе

$$e_1 = 1;\ e_2 = t - \alpha;\ e_3 = (t - \alpha)^2; \ldots e_n (t - \alpha)^n.$$

Отв. $x = \left(f(\alpha),\ f'(\alpha),\ \frac{f''(\alpha)}{\varepsilon}, \ldots, \frac{f^{(n)}(\alpha)}{n!}\right).$

Указание. Воспользоваться формулой Тейлора.

20) Найти характеристические числа и собственные векторы операторов, матрицы которых в некотором базисе имеют вид

$$а)\begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix};\quad б)\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Отв. а) $\lambda_1 = ai$, $\lambda_2 = -ai$, $x_1 = \alpha(1, i)$ $x_2 = \beta(1, -i)$;
б) $\lambda_{1,2} = 1$ $\lambda_3 = -1$, $x_{1,2} = \alpha(1, 0, 1) + \beta(0, 1, 0)$,
$x = \gamma(1, 0, -1)$,

где α, β и γ — произвольные числа.

21) Следом квадратной матрицы A называется сумма его диагональных элементов:

$$\operatorname{Sp} A = a_{11} + a_{22} + \ldots + a_{nn}.$$

Доказать, что подобные матрицы имеют одинаковый след.

Указание. Воспользоваться инвариантностью характеристического многочлена $\Delta(\lambda)$ и найти коэффициент при λ^{n-1}.

22) Зная характеристические числа $\lambda_1, \lambda_2, \ldots, \lambda_n$ матрицы A, найти характеристические числа матриц: а) A^2; б) A^{-1}, если среди характеристических чисел матрицы нет нулевых.

Отв. а) $\lambda_1^2, \lambda_2^2, \ldots, \lambda_n^2$; б) $\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \ldots, \frac{1}{\lambda_n}$.

Указание. Воспользоваться тождествами

$$а)\ A^2 - \lambda^2 E = (A - \lambda E)(A + \lambda E);\ б)\ A^{-1} - \frac{1}{\lambda} E = -\frac{A^{-1}}{\lambda}(A - \lambda E).$$

23) Произвести ортогонализацию системы векторов

$$e_1 = (1, 0, 1);\ e_2 = (1, 1, 0);\ e_3 = (2, 3, 1).$$

Отв. $e_1^* = (1, 0, 1)$; $e_2^* = (-1, -2, 1)$, $e_3^* = (2, -2, -2)$.

24) Построить ортонормированный базис пространства решений системы уравнений

$$3x_1 - x_2 - x_3 - x_4 = 0,\ x_1 + 2x_2 - x_3 - x_4 = 0.$$

Отв. Например,

$$X_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix},\quad X_2 = \frac{1}{\sqrt{150}}\begin{pmatrix} 6 \\ 4 \\ 7 \\ 7 \end{pmatrix}.$$

25) Доказать, что множество векторов евклидова пространства ортогональных данному вектору образует подпространство.

26) Используя результат задачи 24, показать, что симметричный оператор в n-мерном пространстве имеет n попарно ортогональных собственных векторов.

Указание. Рассмотреть подпространство ортогональное произвольному собственному вектору, в этом подпространстве выбрать также произвольно другой собственный вектор. Затем рассмотреть подпространство, ортогональное двум выбранным векторам и т. д.

27) Найти ортогональные преобразования, приводящие следующие квадратичные формы к каноническому виду:

а) $11x_1^2+5x_2^2+2x_3^2+16x_1x_2+4x_1x_3-2x_2x_3$;

б) $x_1^2+x_2^2+x_3^2+4x_1x_2+4x_1x_3+4x_2x_3$;

в) $2x_1x_2-6x_1x_3-6x_2x_4-2x_3x_4$.

Отв. а) $9y_1^2+18y_2^2-9y_3^2$,

$$x_1=\frac{2}{3}y_1+\frac{2}{3}y_2-\frac{1}{3}y_3,$$

$$x_2=-\frac{1}{3}y_1+\frac{2}{3}y_2+\frac{2}{3}y_3,$$

$$x_3=\frac{2}{3}y_1-\frac{1}{3}y_2+\frac{2}{3}y_3;$$

б) $5y_1^2-y_2^2-y_3^2$,

$$x_1=\frac{1}{\sqrt{3}}y_1+\frac{1}{\sqrt{6}}y_2+\frac{1}{\sqrt{2}}y_3,$$

$$x_2=\frac{1}{\sqrt{3}}y_1+\frac{1}{\sqrt{6}}y_2-\frac{1}{\sqrt{2}}y_3,$$

$$x_3=\frac{1}{\sqrt{3}}y_1-\frac{2}{\sqrt{6}}y_2;$$

в) $2y_1^2+4y_2^2-2y_3^2-4y_4^2$,

$$x_1=\frac{1}{2}(y_1+y_2+y_3+y_4),$$

$$x_2=\frac{1}{2}(-y_1+y_2+y_3-y_4),$$

$$x_3=\frac{1}{2}(-y_1-y_2+y_3+y_4),$$

$$x_4=\frac{1}{2}(y_1-y_2+y_3-y_4).$$

28) Найти жордановы цепочки векторов оператора, матрица которого в некотором базисе имеет вид:

$$A=\begin{pmatrix} 3 & -1 & 1 & -7 \\ 9 & -3 & -7 & -1 \\ 0 & 0 & 4 & -8 \\ 0 & 0 & 2 & -4 \end{pmatrix}.$$

Отв. Имеется две цепочки. Например,

$$\boldsymbol{x}_1=(-1,\ -3,\ 0,\ 0);\ \boldsymbol{y}_1=(0,\ 1,\ 0,\ 0)$$

и

$$\boldsymbol{x}_2=(1,\ -7,\ 4,\ 2);\ \boldsymbol{y}_2=(0,\ 0,\ 1,\ 0).$$

29) Привести матрицу

$$A=\begin{pmatrix} \alpha & a_{12} & a_{13} & a_{14} \\ 0 & \alpha & a_{23} & a_{24} \\ 0 & 0 & \alpha & a_{34} \\ 0 & 0 & 0 & \alpha \end{pmatrix}$$

к простейшему виду при условии, что

$$a_{12}\cdot a_{23}\cdot a_{34}\neq 0.$$

Отв.

$$A=\begin{pmatrix} \alpha & 1 & 0 & 0 \\ 0 & \alpha & 1 & 0 \\ 0 & 0 & \alpha & 1 \\ 0 & 0 & 0 & \alpha \end{pmatrix}.$$

30) Пусть функция $f(t)$ определена на спектре матрицы A, характеристические числа которой $\lambda_1, \lambda_2, \ldots, \lambda_n$. Найти характеристические числа матрицы $f(A)$.

Отв.

Примечание. Сравнить с задачей 21.

31) Найти определитель матрицы

$$e^A,$$

где A — квадратная матрица n-го порядка.

Отв. $|e^A| = e^{\mathrm{Sp}\,A} = e^{a_{11}+a_{22}+\ldots+a_{nn}}$.

Указание. Использовать результаты задачи 21 и формулу Виетта для суммы корней многочлена.

32) Доказать, что матрица e^A существует и невырождена для любой квадратной матрицы A.

33) Найти общие решения следующих систем дифференциальных уравнений

а) $\dfrac{dy_1}{dx}=10y_1-3y_2-9y_3,$

$\dfrac{dy_2}{dx}=-18y_1+7y_2+18y_3,$

$\dfrac{dy_3}{dx}=18y_1-6y_2-17y_3;$

б) $\dfrac{dy_1}{dx}=5y_1-y_2-4y_3,$

$\dfrac{dy_2}{dx}=-12y_1+5y_2+12y_3,$

$\dfrac{dy_3}{dx}=10y_1-3y_2-9y_3.$

Отв. а) $y_1=C_1e^{-2x}+C_2e^x+C_3e^x,$
$y_2=-2C_1e^{-2x}+3C_2e^x,$
$y_3=2C_1e^{-2x}+C_3e^x;$

б) $y_1=C_1e^x+C_2(x+1)e^x+C_3e^{-x},$
$y_2=3C_2e^x-2C_3e^{-x},$
$y_3=C_1e^x+C_2xe^x+2C_3e^{-x}.$

34) Найти выражение оператора Лапласа

$$\Delta U=\frac{\partial^2 U}{\partial x^2}+\frac{\partial^2 U}{\partial y^2}+\frac{\partial^2 U}{\partial z^2}$$

а) в цилиндрических координатах

$$x = r\cos\varphi,\quad y = r\sin\varphi,\quad z = z;$$

б) в сферических координатах

$$x = \rho\sin\theta\cos\varphi,\quad y = \rho\sin\theta\sin\varphi,\quad z = \rho\cos\theta.$$

Отв. а) $\Delta U = \frac{\partial^2 U}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2 U}{\partial\varphi^2} + \frac{\partial^2 U}{\partial z^2} + \frac{1}{r}\frac{\partial U}{\partial r}$;

б) $\Delta U = \frac{\partial^2 U}{\partial\rho^2} + \frac{1}{\rho^2}\frac{\partial^2 U}{\partial\theta^2}\frac{1}{\rho^2\sin^2\theta}\frac{\partial^2 U}{\partial\varphi^2} + \frac{2}{\rho}\frac{\partial U}{\partial\rho} + \frac{1}{\rho^2\operatorname{tg}\theta}\frac{\partial U}{\partial\theta}$.

35) Приняв u и v за новые независимые переменные, преобразовать уравнение

$$\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} + m^2 z = 0,\qquad x = \frac{u^2 - v^2}{2};\qquad y = uv.$$

Отв. $\frac{\partial^2 z}{\partial u^2} + \frac{\partial^2 z}{\partial v^2} + m^2(u^2 + v^2)z = 0.$

36) Приняв u и v за новые независимые уравнения, а w за новую функцию, преобразовать уравнение

$$\frac{\partial z}{\partial y} + \frac{1}{2}y\frac{\partial^2 z}{\partial y^2} = \frac{1}{x},\qquad u = \frac{x}{y},\qquad v = x,\qquad w = xz - y.$$

Отв. $\frac{\partial^2 w}{\partial u^2} = 0.$

37) Найти экстремальные значения следующих функций

а) $z = x^2 + xy + y^2 - 3ax - 3by$; б) $u = xyz(4a - x - y - z)$.

Отв. а) Минимум $z = -3(a^2 + b^2 - ab)$ при $x = 2a - b$, $y = 2b - a$. б) Максимум $u = a^4$ при $x = y = z = a$ (в критической точке $x = y = z = 0$ нет экстремума).

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

Алгоритм Гаусса 28

Базис линейного пространства 56
— ортонормированный 112

Вектор 50
— собственный линейного оператора 86
— , координаты 57
— , норма 108
— ортогональный 110
Векторы линейно независимые и линейно зависимые 55
— , угол 109
Вероятность начальная 212
— переходная 214
— финальная 216
Вычитание матриц 13

Жорданов базис 141, 151
Жорданова матрица 145
— клетка матрицы 146

Зависимость линейная векторов 55

Интеграл от матрицы 173

Квадратичная форма 121
Координаты вектора 57
— криволинейные 202
Критерий Сильвестра 136

Линейный оператор 65
Линейное преобразование 45
— пространство 50
Линия координатная 203, 204
— второго порядка 129

Матрица 10
— вырожденная 12
— диагональная 11
— единичная 11
— жорданова 22, 145
— интегральная 178
Матрица квадратная 10
— клеточная 21
— , характеристический многочлен 86
— квазидиагональная 22
— кососимметрическая 121
— линейного оператора 66
— невырожденная 12
— обратная 17
— ортогональная 118
— особенная (вырожденная) 12
— преобразующая 80
— простой структуры 96
— прямоугольная 10
— симметрическая 16
— столбцевая 10
— стохастическая 212
— строчная 10
— транспонированная 16
— унитарная 121
— фундаментальная 178
— , интеграл 173
— , минор 35
— , многочлен от нее 163
— , производная 173
— , ранг 36
— , степень 20
— , функция от нее 165
Матрицы перестановочные 14
— подобные 81
— , вычитание 13
— , сложение 12
— , умножение 13
— , умножение на число 12
Метод Гаусса 28
Минор 6
Многочлен Лагранжа — Сильвестра 168
— неоднородный 125
— от матрицы 163
— нецентральный 125
— характеристический 86
— — , инвариантность 86
— центральный 125

Неравенство Коши — Буняковского 109
— треугольника 111
Норма вектора 108

Оператор дифференцируемый 193
— линейный 65
— —, матрица 66
— — невырожденный 69
— — нулевой 70
— — простой структуры 96
— — самосопряженный 116
— — симметричный 117
— — сопряженный 116
— —, собственные векторы 87
— —, собственные значения 86
— — тождественный 70
— обратный 76, 195
Операторы линейные, сложение 74
— —, умножение 75
— —, умножение на число 74
Определитель матрицы 11
Ортогонализация системы векторов 114

Поверхность координатная 204
— второго порядка 130
Подпространство пространства 53
Порядок матрицы 10
Предел последовательности матриц 170
Преобразование координат 77
— ортогональное 118
— элементарное 32
Приведение квадратичной формы к каноническому виду 123
Проектирования оператор 77
Пространство линейное 50
— евклидово 106
— —, ортонормированный базис 114
— унитарное 106

Ранг линейного оператора 68
— матрицы 36

Система линейных дифференциальных уравнений 175
— — — — с постоянными коэффициентами 184
Система линейных уравнений 26
Сложение линейных операторов 74
— — преобразований 46
— матриц 12

Теорема Гамильтона — Кэли 167
— Кронекера — Капелли 38

Угол между векторами 109
Умножение линейных операторов 75
— матриц 13
— — на число линейного оператора 74
— — — матрицы 12
Уравнение линейное дифференциальное 182
— Колмогорова 227
— характеристическое 86

Форма квадратичная 121
— — отрицательно определенная 135
— — положительно определенная 135
— нормальная жорданова 145, 156
Формулы Крамера 9, 18
— Эрланга 235
Функция от матрицы 165

Характеристическое число 86

Цепочка векторов жорданова 143
Цепь Маркова 211
— — однородная 212
— — правильная 222
— — простая 212
— — регулярная 222
— — стационарная 216

Число собственное (характеристическое) линейного оператора 86
— характеристическое матрицы 86
Числовое поле 49

Эквивалентные матрицы 32

Эргодическое свойство 218

Якобиева матрица 194

ОГЛАВЛЕНИЕ

Предисловие 3
Введение. Определители n-го порядка 5
Глава 1. Матрицы 10
§ 1. Основные определения 10
§ 2. Операции над матрицами 12
§ 3. Клеточные матрицы 21
Глава 2. Решение системы линейных уравнений . . . 26
§ 1. Формулы Крамера 26
§ 2. Метод Гаусса 28
§ 3. Эквивалентные матрицы 32
§ 4. Ранг матрицы 35
§ 5. Критерий совместности системы 38
§ 6. Однородная система линейных уравнений 39
Глава 3. Линейные преобразования и линейные пространства 45
§ 1. Линейные преобразования 45
§ 2. Линейные пространства 49
§ 3. Базис линейного пространства 54
§ 4. Действия над векторами в координатной форме . . 57
Глава 4. Линейный оператор 65
§ 1. Линейный оператор 65
§ 2. Действия с линейными операторами 74
§ 3. Преобразование координат 77
§ 4. Изменение матрицы линейного оператора при переходе к новому базису 80
§ 5. Геометрический смысл определителя матрицы оператора в пространствах V_3 и V_2 81
Глава 5. Собственные векторы и собственные числа линейного оператора 85
§ 1. Характеристический многочлен и характеристическое уравнение матрицы 85
§ 2. Собственный вектор и собственное число линейного оператора 86
§ 3. Оператор простой структуры 96
Глава 6. Пространства со скалярным произведением 105
§ 1. Евклидово пространство 105
§ 2. Метрические свойства векторов 108
§ 3. Ортонормированный базис 112
§ 4. Самосопряженный оператор 116
§ 5. Симметричный оператор 117
§ 6. Ортогональные преобразования 118

Глава 7. Квадратичные формы 121
§ 1. Квадратичные формы и приведение их к каноническому виду 121
§ 2. Неоднородный многочлен 2-ой степени 124
§ 3. Упрощение уравнений линий и поверхностей 2-го порядка 129
§ 4. Знакоопределенные квадратичные формы 135
Глава 8. Жорданова нормальная форма матриц 139
§ 1. Вспомогательный базис специального вида 139
§ 2. Жорданов базис в частном случае 141
§ 3. Жорданов базис в общем случае 151
Глава 9. Функции от матриц 160
§ 1. Возведение квадратной матрицы в целую положительную степень 160
§ 2. Многочлены от матриц 163
§ 3. Функции от матрицы 165
§ 4. Специальные многочлены от матрицы 167
§ 5. Разложение функции от матрицы в ряд 170
§ 6. Функциональные матрицы 172
Глава 10. Системы линейных дифференциальных уравнений 175
§ 1. Системы линейных дифференциальных уравнений . . 175
§ 2. Линейные системы с постоянными коэффициентами 184
Глава 11. Приложение матриц к некоторым вопросам анализа 191
§ 1. Нелинейный оператор в евклидовом пространстве . . 191
§ 2. Экстремум функции нескольких переменных 196
§ 3. Замена переменных в дифференциальных выражениях 199
§ 4. Замена переменных в кратных интегралах 206
Глава 12. Цепи Маркова 211
§ 1. Цепи Маркова с дискретным временем 211
§ 2. Финальные вероятности для цепей Маркова с конечным числом состояний 218
§ 3. Цепи Маркова с непрерывным временем 225
Ответы 237
Задачи и упражнения 245
Предметный указатель 253

ЗБК
Б705

Цена 30 коп.

ИЗДАТЕЛЬСТВО ВЫСШАЯ ШКОЛА
МОСКВА
1971 ГОД