САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ
УНИВЕРСИТЕТ
Факультет прикладной математики — процессов управления

А.Ю.УТЕШЕВ

ВЫСШАЯ АЛГЕБРА

Раздел II

Санкт-Петербург
2007

ББК 22.14
Д32
УДК 512.8

Р е ц е н з е н т ы:
докт. физ.-мат. наук, проф. *Е.И.Веремей*
(С.-Петерб. гос. ун-т.),
докт. физ.-мат. наук, проф. *В.Ф.Зайцев*
(Рос. гос. пед. ун-т им. А.И.Герцена)

Печатается по заказу профсоюзного бюро студентов
и по постановлению редакционно-издательского совета
факультета прикладной математики — процессов управления
Санкт-Петербургского государственного университета

Утешев А.Ю.

Д32 Высшая алгебра. Раздел II. Учеб. пособие. Издание второе, исправленное. — СПб.: Изд-во "Золотое Сечение", 2007. — 162 с.

Настоящее пособие составлено на основе лекций, читаемых автором студентам факультета прикладной математики — процессов управления (ПМ-ПУ) СПбГУ по специальности "Прикладная математика и информатика".

ББК 22.14

Раздел II
ЛИНЕЙНЫЕ ПРОСТРАНСТВА И ОТОБРАЖЕНИЯ

Оглавление

Глава 3
ЛИНЕЙНЫЕ ОТОБРАЖЕНИЯ

Глава 4
ПРИМЕНЕНИЯ ЖОРДАНОВОЙ НОРМАЛЬНОЙ ФОРМЫ

Глава 5
ЭЛЕМЕНТЫ ЧИСЛЕННЫХ МЕТОДОВ ЛИНЕЙНОЙ АЛГЕБРЫ

Раздел II
ЛИНЕЙНЫЕ ПРОСТРАНСТВА И ОТОБРАЖЕНИЯ

Основная задача раздела

В первом разделе курса высшей алгебы основной задачей этой науки была продекларирована задача поиска решений уравнений и систем уравнений.

К началу XX века в алгебре начало формироваться и другое направление исследований: классификация (установление свойств) многомерных отображений различной природы. Именно, выявление **свойств линейных отображений** составляет основную задачу настоящего раздела.

Общим у двух разделов алгебры является математический аппарат: теория матриц (ими линейные отображения описываются) и теория алгебраических уравнений от одной переменной (с их помощью эти отображения анализируются).

Кроме того, в настоящем разделе на первый план выдвигаются и прикладные аспекты алгебры: разработанные в ней алгоритмы используются в теории дифференциальных уравнений (в том числе, в механике) и в теории вероятностей.

Благодарность

Автор благодарит Е.А.Калинину за оформление всех рисунков настоящего пособия.

Глава 1. Линейные пространства и многообразия

1 Основные определения

Определение. Пусть дано множество $\mathbb{V} = \{X, Y, Z, U, \dots\}$ элементов произвольной природы. Пусть для элементов этого множества определены две операции: сложения $X+Y$ и умножения на любое вещественное число $\alpha\cdot X$ и множество $\mathbb{V}$ замкнуто относительно этих операций: $X+Y \in \mathbb{V},\ \alpha\cdot X \in \mathbb{V}$. Пусть эти операции подчиняются аксиомам:

1. $X + Y = Y + X$ для $\{X, Y\} \subset \mathbb{V}$.
2. $(X + Y) + Z = X + (Y + Z)$ для $\{X, Y, Z\} \subset \mathbb{V}$.
3. В $\mathbb{V}$ существует нулевой вектор $\mathbb{O}$ со свойством $X + \mathbb{O} = X$ для $\forall X \in \mathbb{V}$.
4. Для каждого $X \in \mathbb{V}$ существует обратный вектор $X' \in \mathbb{V}$ со свойством $X + X' = \mathbb{O}$.
5. $1 \cdot X = X$ для $\forall X \in \mathbb{V}$.
6. $\lambda\,(\mu X) = (\lambda\mu)\,X$ для $\forall X \in \mathbb{V}$, $\{\lambda, \mu\} \in \mathbb{R}$.
7. $(\lambda + \mu)X = \lambda X + \mu X$ для $\forall X \in \mathbb{V}$, $\{\lambda, \mu\} \in \mathbb{R}$.
8. $\lambda(X + Y) = \lambda X + \lambda Y$ для $\{X, Y\} \subset \mathbb{V}, \lambda \in \mathbb{R}$.

Тогда такое множество $\mathbb{V}$ называется **линейным (векторным) пространством**, его элементы называются **векторами**, и — чтобы подчеркнуть их отличие от чисел из $\mathbb{R}$ — последние называются **скалярами**. Пространство, состоящее из одного только нулевого вектора, называется **тривиальным**.

Замечание. Если в аксиомах 6–8 допустить умножение и на комплексные скаляры, то такое линейное пространство называется **комплексным**.

Элементарно доказывается единственность нулевого вектора, и единственность вектора, обратного вектору $X \in \mathbb{V}$: $X' = -1 \cdot X \stackrel{\text{def}}{=} -X$.

Определение. Подмножество лин.пространства $\mathbb{V}$, само являющееся лин. пространством, называется **линейным подпространством** пространства $\mathbb{V}$. **Тривиальными подпространствами** лин.пространства $\mathbb{V}$ называются само $\mathbb{V}$ и пространство, состоящее из одного нулевого вектора $\{\mathbb{O}\}$.

Приведем некоторые примеры лин.пространств и их подпространств.

Пространство $\mathbb{R}^3$ упорядоченных троек вещественных чисел $[a_1, a_2, a_3]$ с операциями, определяемыми равенствами:

$$[a_1, a_2, a_3]+[b_1, b_2, b_3] \stackrel{\text{def}}{=} [a_1+b_1, a_2+b_2, a_3+b_3],\ \lambda[a_1, a_2, a_3] \stackrel{\text{def}}{=} [\lambda a_1, \lambda a_2, \lambda a_3]\ .$$

Геометрическая интерпретация очевидна: вектор в пространстве, “привязанный” к началу координат, может быть задан в координатах своего конца $[a_1, a_2, a_3]$.

Естественным обобщением $\mathbb{R}^3$ служит пространство $\mathbb{R}^n$: векторное пространство строк $[a_1, \dots, a_n]$ или столбцов $[a_1, \dots, a_n]^\top$. Один из способов задания подпространства в $\mathbb{R}^n$ — задание набора ограничений. Так, рассмотрим систему линейных однородных уравнений:

$$\left\{ \begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n & = & 0, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n & = & 0, \\ \ldots & & \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n & = & 0 \end{array} \right. \iff AX = \mathbb{O} \ . \qquad (1.1)$$

Теорема 1.1. *Множество решений системы* (1.1) *образует линейное подпространство пространства* $\mathbb{R}^n$*.*

Доказательство следует из того, что если столбец $X_1 \in \mathbb{R}^n$ удовлетворяет системе (1.1), то и λX_1 удовлетворяет этой системе; если, вдобавок, и X_2 — решение системы, то

$$AX_1 = \mathbb{O},\ AX_2 = \mathbb{O} \ \Rightarrow \ A(X_1 + X_2) = AX_1 + AX_2 = \mathbb{O} \ .$$

□

Упражнение 1.1. *Почему множество решений системы неоднородных уравнений не образует линейного подпространства?*

Обобщая дальше, можем рассмотреть пространство “бесконечно-длинных” строк $[a_1, \dots, a_n, \dots]$, обычно являющееся объектом математического анализа — при рассмотрении последовательностей и рядов.

Множество полиномов степени не выше n

$$\mathbb{P}_n = \left\{ p(x) \in \mathbb{R}[x] \middle| \deg p(x) \le n \right\}$$

с обычными операциями сложения полиномов и умножения на число образует лин. пространство. Его очевидными подпространствами являются $\mathbb{P}_0, \mathbb{P}_1, \dots, \mathbb{P}_{n-1}$. Кроме того, подпространствами будут множество четных и множество нечетных полиномов степени не выше n. Множество всевозможных полиномов (без ограничения на степени) тоже образует линейное пространство.

Множество $m \times n$-матриц с элементами из поля $\mathbb{R}$ с операциями сложения матриц и умножения на вещественные числа образует лин. пространство.

В пространстве квадратных матриц порядка n можно выделить два подпространства: подпространство симметричных матриц и подпространство кососимметричных матриц.

Пусть имеются два линейных пространства: $\mathbb{V}$ с операцией $+$ и $\mathbb{W}$ с операцией $\oplus$. Можно ли их как-нибудь сравнивать? В теории групп две разные группы сравнивались через понятие изоморфизма. Используем то же слово и для лин.пространств, только на этот раз потребуем большего.

Определение. Пространства $\mathbb{V}$ и $\mathbb{W}$ **изоморфны** если между множествами их элементов можно установить такое взаимно-однозначное соответствие, что если $X \leftrightarrow X'$ и $Y \leftrightarrow Y'$ то $X + Y \leftrightarrow X' \oplus Y'$ и $\lambda X \leftrightarrow \lambda X'$.

Теорема 1.2. *При изоморфизме пространств $\mathbb{V}$ и $\mathbb{W}$ нулевому вектору одного пространства будет соответствовать нулевой вектор другого пространства.*

Доказательство . Действительно, если $X \leftrightarrow X'$ и $\mathbb{O} \leftrightarrow Y'$, то, по свойству изоморфизма, должно выполняться $X = X + \mathbb{O} \leftrightarrow X' \oplus Y'$. Таким образом, должно выполняться $X' = X' \oplus Y'$ для любого X'. Но это возможно тогда и только тогда, когда $Y' = \mathbb{O}'$ — на основании аксиомы 3. □

Пример 1.1. *Пространство $\mathbb{R}^n$ изоморфно пространству $\mathbb{P}_{n-1}$. В самом деле, изоморфизм устанавливается соответствием*

$$[a_1, \dots, a_n] \leftrightarrow a_1 + a_2 x + \dots + a_n x^{n-1} \quad .$$

Пример 1.2. *Пространство $m \times n$-матриц изоморфно пространству $\mathbb{R}^{mn}$. Изоморфизм устанавливается соответствием, которое мы проиллюстрируем для $m = 2, n = 3$:*

$$\left(\begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{array} \right) \leftrightarrow [a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23}]$$

(матрица “вытягивается” в одну строчку).

2 Линейная зависимость, базис

Определение. Множество всевозможных линейных комбинаций системы $\{X_1, \dots, X_m\}$

$$\left\{ \sum_{j=1}^{m} \alpha_j X_j \middle| [\alpha_1, \dots, \alpha_m] \in \mathbb{R}^m \right\}$$

образует линейное подпространство пространства $\mathbb{V}$. Оно называется **линейной оболочкой** векторов $X_1, \dots, X_m$ и обозначается $\mathcal{L}(X_1, \dots, X_m)$.

Все остальные определения и результаты этого пункта фактически повторяют те, что приведены в §11.1 главы 4 раздела I. Поэтому мы опускаем доказательства.

Определение. Система (совокупность, набор) m векторов

$$\{X_1, \dots, X_m\} \tag{2.1}$$

называется **линейно зависимой** (**л.з.**) если существуют числа $\alpha_1, \dots, \alpha_m$, хотя бы одно из которых отлично от нуля, такие что

$$\alpha_1 X_1 + \dots + \alpha_m X_m = \mathbb{O} \tag{2.2}$$

Если же равенство (2.2) возможно только при $\alpha_1 = 0, \dots, \alpha_m = 0$, то система (2.2) называется линейно независимой (**л.н.з.**).

Теорема 2.1. **а)** *Если система* (2.1) *содержит хотя бы один нулевой вектор, то она* **л.з.**

б) *Если система* (2.1) **л.н.з.**, *то и любой ее подсистема* **л.н.з.**

в) *При* $m > 1$ *система* (2.1) **л.з.** *тогда и только тогда, когда по меньшей мере один ее вектор линейно выражается через остальные, т.е. существуют* $j \in \mathbb{N}$ *и константы* $\gamma_1, \dots, \gamma_{j-1}, \gamma_{j+1}, \dots, \gamma_n$ *такие, что*

$$X_j = \gamma_1 X_1 + \dots + \gamma_{j-1} X_{j-1} + \gamma_{j+1} X_{j+1} + \dots + \gamma_m X_m \ .$$

Теорема 2.2. *Если каждый из векторов системы* (2.1) *линейно выражается через векторы* $B_1, \dots, B_k$ *и при этом* $k < m$, *то система* (2.1) *будет* **л.з.**

Определение. Две системы векторов называются **эквивалентными** если каждый вектор одной системы линейно выражается через векторы другой и обратно.

Теорема 2.3. *Системы векторов* $\{X_1, \dots, X_m\}$ *и* $\{Y_1, \dots, Y_k\}$ *будут эквивалентными тогда и только тогда когда совпадают линейные оболочки этих систем:*

$$\mathcal{L}(X_1, \dots, X_m) = \mathcal{L}(Y_1, \dots, Y_k) \ .$$

Теорема 2.4. *Если каждая из эквивалентных систем* $\{X_1, \dots, X_m\}$ *и* $\{Y_1, \dots, Y_k\}$ *является* **л.н.з**, *то эти системы состоят из одинакового числа векторов:* $m = k$.

Определение. Линейно независимая система векторов $\{X_j\} \subset \mathbb{V}$ называется **базисом** этого пространства если каждый $X \in \mathbb{V}$ можно представить в виде линейной комбинации указанных векторов:

$$X = \sum_j \alpha_j X_j \ . \tag{2.3}$$

При этом не подразумевается конечность системы, т.е. суммирование может распространяться на *бесконечное* число слагаемых. Так, например, пространство бесконечных строк $[a_1, a_2, \dots]$ имеет бесконечный базис, состоящий из векторов

$$[\underbrace{0, \dots, 0, 1}_{j}, 0, \dots] \quad \text{при } j \in \mathbb{N} \ .$$

Определение. В случае, когда базис пространства $\mathbb{V}$ конечен, то пространство $\mathbb{V}$ называется **конечномерным**, а число векторов базиса тогда называется **размерностью** пространства $\mathbb{V}$ и обозначается[1]: $\dim \mathbb{V}$. Также полагают $\dim\{\mathbb{O}\} \stackrel{\text{def}}{=} 0$.

Замечание. В дальнейшем мы будем рассматривать только конечномерные пространства.

Единственность базиса в пространстве не предполагается.

[1] **dimensio** (*лат.*) — обмер, измерение, протяжение; хлебная мера, солдатский паек.

Теорема 2.5. *Если* $\dim \mathbb{V} = d > 0$*, то любая система из* d *линейно независимых векторов пространства образует базис этого пространства.*

Доказательство . Пусть $\{Y_1, \dots, Y_d\}$ — **л.н.з.** система. Рассмотрим произвольный $X \in \mathbb{V}$. Если система $\{X, Y_1, \dots, Y_d\}$ **л.н.з.**, то $\dim \mathbb{V} \geq d + 1$, что противоречит условию теоремы. Следовательно, система линейно зависима: $\alpha_0 X + \alpha_1 Y_1 + \dots + \alpha_d Y_d = \mathbb{O}$ при каком-то из чисел α_j не равном нулю. Если $\alpha_0 = 0$, то $\alpha_1 Y_1 + \dots + \alpha_d Y_d = \mathbb{O}$ при каком-то ненулевом коэффициенте. Это означает, что система $\{Y_1, \dots, Y_d\}$ линейно зависима, что противоречит предположению. Следовательно $\alpha_0 \neq 0$, но тогда вектор X может быть представлен в виде линейной комбинации векторов $Y_1, \dots, Y_d$:

$$X = -{}^{\alpha_1}\!/_{\alpha_0} Y_1 - \dots - {}^{\alpha_d}\!/_{\alpha_0} Y_d \ .$$

По определению, система $\{Y_1, \dots, Y_d\}$ является базисом $\mathbb{V}$. $\square$

Найдем теперь размерности линейных пространств, упомянутых в §1.

Очевидно, $\dim \mathbb{R}^n = n$: строки

$$[1, 0, 0, \dots, 0], \ [0, 1, 0, \dots, 0], \ [0, 0, 1, \dots, 0], \ \dots, [0, 0, 0, \dots, 1]$$

образуют базис этого пространства.

В $\mathbb{R}^n$ имеется два способа задания линейных подпространств. Пусть

$$\mathbb{V}_1 \stackrel{\text{def}}{=} \mathcal{L}(A_1, \dots, A_k) \quad \text{при } \{A_1, \dots, A_k\} \subset \mathbb{R}^n \ .$$

В §11 главы 4 раздела I было установлено, что

$$\dim \mathbb{V}_1 = \operatorname{rank} \{A_1, \dots, A_k\} = \operatorname{rank} A \ ,$$

где A — матрица, составленная из строк (столбцов) $A_1, \dots, A_k$.

Пример 2.1. *Найти базис подпространства*

$$\mathcal{L}\left([1, 2, 1, 1], \, [-1, 0, -1, 0], \, [-1, 2, -1, 1], \, [0, 1, 0, 1]\right) \ .$$

Решение. Ищем

$$\operatorname{rank} \left(\begin{array}{rrrr} 1 & 2 & 1 & 1 \\ -1 & 0 & -1 & 0 \\ -1 & 2 & -1 & 1 \\ 0 & 1 & 0 & 1 \end{array} \right)$$

по методу окаймляющих миноров. Существует минор третьего порядка

$$\left| \begin{array}{rrr} 1 & 2 & 1 \\ -1 & 0 & 0 \\ 0 & 1 & 1 \end{array} \right|$$

отличный от нуля, а определитель самой матрицы равен нулю. Замечаем, что найденный отличный от нуля минор расположен в первой, второй и четвертой строках матрицы. Именно эти строки и образуют базис.

Ответ. Базис составляют, например, первая, вторая и четвертая строки.

Другим способом задания линейного подпространства в $\mathbb{R}^n$ может служить задание набора ограничений, которым должны удовлетворять векторы подпространства. Таким набором ограничений может являться, например, система уравнений (1.1). Какова размерность подпространства решений этой системы? На этот вопрос мы ответим сразу же, если вспомним определение фундаментальной системы решений (**ф.с.р.**). Именно, в терминах настоящей главы, **ф.с.р.** — набор линейно независимых решений, через которые линейно выражается любое решение системы (1.1). Иначе говоря, **ф.с.р.** является базисом подпространства решений. Теорема 12.4 главы 4, раздела I утверждала, что размерность этого подпространства равна $n - \mathfrak{r}$, где $\mathfrak{r} \stackrel{\mathrm{def}}{=} \operatorname{rank} A$.

В качестве следующего примера рассмотрим пространство полиномов $\mathbb{P}_n$. Его размерность равна $n+1$ поскольку любой полином степени не выше n может быть представлен в виде линейной комбинации степеней переменной $1, x, x^2, \dots, x^n$, а тождество $a_0x^n + a_1x^{n-1} + \dots + a_n \equiv 0$ возможно только при нулевом наборе коэффициентов.

Обращаясь теперь ко множеству $m \times n$ матриц, заметим, что каждая из них может быть представлена в виде линейной комбинации следующих матриц

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

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

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

Набор этих матриц линейно независим. Поэтому размерность линейного пространства $m \times n$-матриц равна mn.

Упражнение 2.1. *Найти размерности подпространств симметричных и кососимметричных матриц порядка n.*

Определение. Если $\dim \mathbb{V} = d$, то сумма (2.3) становится конечной:

$$X = \alpha_1 X_1 + \dots + \alpha_d X_d \ .$$

В этом случае говорят, что **вектор X разложен по базису** $\{X_1, \dots, X_d\}$, а числа $\alpha_1, \dots, \alpha_d$ называются **координатами** вектора X в этом базисе.

Теорема 2.6. *Любой вектор $X \in \mathbb{V}$ может быть разложен по фиксированному базису пространства единственным образом.*

Доказательство . Пусть для какого-то вектора X существует два разложения по базису $\{X_1, \dots, X_d\}$:

$$X = x_1 X_1 + \dots + x_d X_d = \widetilde{x}_1 X_1 + \dots + \widetilde{x}_d X_d \ ,$$

тогда $(x_1 - \widetilde{x}_1)X_1 + \dots + (x_d - \widetilde{x}_d)X_d = \mathbb{O}$. Поскольку векторы $\{X_1, \dots, X_d\}$ **л.н.з**, отсюда следует $x_1 - \widetilde{x}_1 = 0, \dots, x_d - \widetilde{x}_d = 0$. □

Иными словами, при фиксировании базиса $\mathbb{V}$, вектор $X \in \mathbb{V}$ будет однозначно описываться упорядоченным набором своих координат. Это обстоятельство позволяет свести исследование свойств *произвольного* векторного пространства к одному-единственному частному случаю — векторному пространству строк.

Теорема 2.7. *Любое векторное пространство $\mathbb{V}$ размерности d изоморфно $\mathbb{R}^d$.*

Доказательство . Изоморфизм можно установить следующим соответствием. Если $\{X_1, \dots, X_d\}$ — какой-то базис $\mathbb{V}$, то вектору $X \in \mathbb{V}$ поставим в соответствие набор его координат в этом базисе:

$$X = x_1 X_1 + \dots + x_d X_d \ \Rightarrow \ X \mapsto [x_1, \dots, x_d] \in \mathbb{R}^d \ .$$

На основании теоремы 2.6, такое соответствие будет взаимно-однозначным, а проверка двух свойств изоморфизма тривиальна. □

Теорема 2.8. *Пусть $\mathbb{V}_1$ — подпространство пространства $\mathbb{V}$, отличное от тривиального. Тогда произвольный базис подпространства $\mathbb{V}_1$ можно дополнить до базиса пространства $\mathbb{V}$.*

Доказательство . Пусть $\dim \mathbb{V}_1 = d_1$ и $\{X_1, \dots, X_{d_1}\}$ — какой-то базис этого подпространства. В пространстве $\mathbb{V}$ найдется такой вектор X_{d_1+1}, что набор $\{X_1, \dots, X_{d_1}, X_{d_1+1}\}$ будет **л.н.з.** (В противном случае, $\dim \mathbb{V} = d_1$, что противоречит условию.) Если $d_1 + 1 = d \overset{\text{def}}{=} \dim \mathbb{V}$, то, на основании теоремы 2.5, требуемый базис построен. Если же $d_1 + 1 < d$, то в пространстве $\mathbb{V}$ найдется такой вектор X_{d_1+2}, что набор $\{X_1, \dots, X_{d_1}, X_{d_1+1}, X_{d_1+2}\}$ будет **л.н.з.** И т.д. Процесс закончится за конечное число шагов. □

3 Сумма и пересечение линейных подпространств. Линейное многообразие

3.1 Сумма и пересечение

Определение. Пусть $\mathbb{V}_1$ и $\mathbb{V}_2$ — подпространства лин. пространства $\mathbb{V}$. Множество

$$\mathbb{V}_1 + \mathbb{V}_2 \overset{\text{def}}{=} \left\{ X_1 + X_2 \middle| X_1 \in \mathbb{V}_1, \ X_2 \in \mathbb{V}_2 \right\}$$

называется **суммой**, а множество

$$\mathbb{V}_1 \cap \mathbb{V}_2 \stackrel{\text{def}}{=} \left\{ X \middle| X \in \mathbb{V}_1,\ X \in \mathbb{V}_2 \right\}$$

—**пересечением** подпространств $\mathbb{V}_1$ и $\mathbb{V}_2$. Аналогично определяется сумма и пересечение $N > 2$ подпространств.

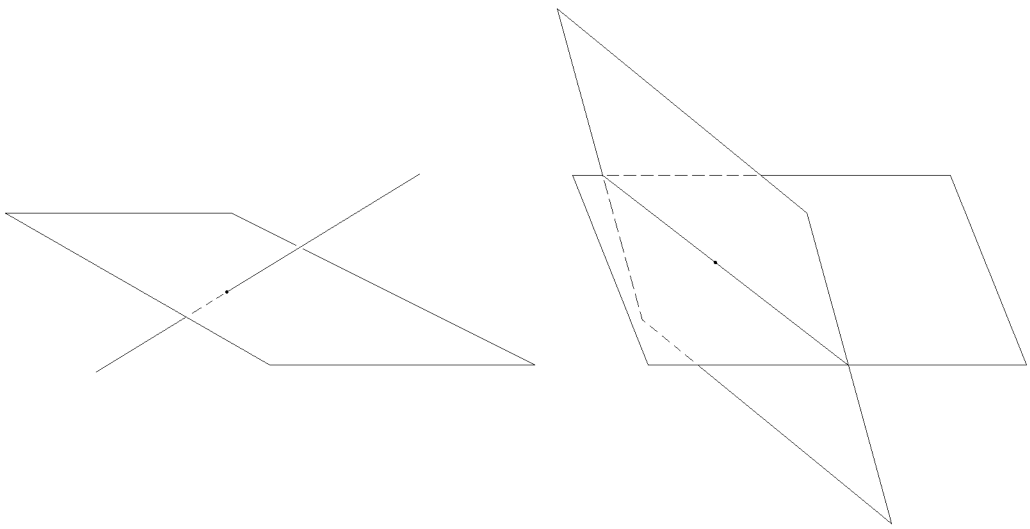


Рис. 1 а) Рис. 1 б)

Фактически понятие пересечения лин.подпространств совпадает с понятием пересечения множеств.

$$\boxed{\textbf{Как правило}, \ \mathbb{V}_1 + \mathbb{V}_2 \neq \mathbb{V}_1 \cup \mathbb{V}_2.}$$

Геометрическая интерпретация суммы подпространств в $\mathbb{R}^3$ дается рисунком 1. В обоих случаях $\mathbb{V}_1 + \mathbb{V}_2 = \mathbb{R}^3$: любой вектор $\mathbb{R}^3$ можно получить суммированием векторов, взятых из каждого подпространства.

Теорема 3.1. *$\mathbb{V}_1 + \mathbb{V}_2$ и $\mathbb{V}_1 \cap \mathbb{V}_2$ являются подпространствами лин. пространства $\mathbb{V}$.*

Теорема 3.2. *Имеет место формула:*

$$\dim \mathbb{V}_1 + \dim \mathbb{V}_2 = \dim (\mathbb{V}_1 \cap \mathbb{V}_2) + \dim (\mathbb{V}_1 + \mathbb{V}_2) \tag{3.1}$$

Доказательство . Пусть $d_1 \stackrel{\text{def}}{=} \dim \mathbb{V}_1$, $d_2 \stackrel{\text{def}}{=} \dim \mathbb{V}_2$, $p \stackrel{\text{def}}{=} \dim (\mathbb{V}_1 \cap \mathbb{V}_2)$ и $\{X_1, \dots, X_p\}$ — базис $\mathbb{V}_1 \cap \mathbb{V}_2$. На основании теоремы 2.8, этот базис можно дополнить до базисов каждого из подпространств $\mathbb{V}_j$: пусть

$$\begin{array}{rl} \{X_1, \dots, X_p, X_{p+1}, \dots, X_{d_1}\} & \text{— базис } \mathbb{V}_1 \quad \text{, а} \\ \{X_1, \dots, X_p, Y_{p+1}, \dots, Y_{d_2}\} & \text{— базис } \mathbb{V}_2 \end{array}$$

Докажем, что система

$$\{X_1, \dots, X_p, X_{p+1}, \dots, X_{d_1}, Y_{p+1}, \dots, Y_{d_2}\} \tag{3.2}$$

является базисом $\mathbb{V}_1 + \mathbb{V}_2$. Действительно, произвольный вектор $Z \in \mathbb{V}_1 + \mathbb{V}_2$ можно представить в виде $Z = Z_1 + Z_2$, где $Z_j \in \mathbb{V}_j$. А каждый из слагаемых векторов, в свою очередь, можно разложить в линейную комбинацию базисных векторов соответствующего подпространства.

Покажем, что система (3.2) **л.н.з.** Пусть

$$\underbrace{\alpha_1 X_1 + \dots + \alpha_p X_p + \alpha_{p+1} X_{p+1} + \dots + \alpha_{d_1} X_{d_1}}_{\stackrel{\text{def}}{=} U \in \mathbb{V}_1} + \beta_{p+1} Y_{p+1} + \dots + \beta_{d_2} Y_{d_2} = \mathbb{O} \ . \tag{3.3}$$

Из этого соотношения вектор U может быть выражен в виде:

$$U = -\beta_{p+1} Y_{p+1} - \dots - \beta_{d_2} Y_{d_2} \ \in \mathbb{V}_2 \tag{3.4}$$

и, таким образом, $U \in \mathbb{V}_1 \cap \mathbb{V}_2$. Тогда U выражается через базисные векторы пересечения:

$$U = -\beta_{p+1} Y_{p+1} - \dots - \beta_{d_2} Y_{d_2} = \tilde{\alpha}_1 X_1 + \dots + \tilde{\alpha}_p X_p \ .$$

Отсюда

$$\tilde{\alpha}_1 X_1 + \dots + \tilde{\alpha}_p X_p + \beta_{p+1} Y_{p+1} + \dots + \beta_{d_2} Y_{d_2} = \mathbb{O}$$

и, следовательно, $\tilde{\alpha}_1 = 0, \dots, \tilde{\alpha}_p = 0, \beta_{p+1} = 0, \dots, \beta_{d_2} = 0$, поскольку комбинируемые векторы являются базисными для $\mathbb{V}_2$. Итак, $U = \mathbb{O}$, но тогда из определения этого вектора по (3.3) вытекает, что $\alpha_1 = 0, \dots, \alpha_{d_1} = 0$, т.к. комбинируемые векторы составляют базис $\mathbb{V}_1$. Итак, соотношение (3.3) возможно только при нулевом наборе скаляров, что и означает линейную независимость системы (3.2).

Мы доказали, что система (3.2) образует базис пространства $\mathbb{V}_1 + \mathbb{V}_2$, тогда

$$\dim(\mathbb{V}_1 + \mathbb{V}_2) = d_1 + d_2 - p \ .$$

□

Замечание. Только что доказанная теорема может быть трактована как проявление того общего результата формальной логики, что был сформулирован в теореме 4.2 главы 1 раздела I. Эта аналогия позволяет распространить результат теоремы на случай пересечения трех и более подпространств.

Упражнение 3.1. *Показать справедливость формулы*

$$\dim \mathbb{V}_1 + \dim \mathbb{V}_2 + \dim \mathbb{V}_3 - \{\dim(\mathbb{V}_1 \cap \mathbb{V}_2) + \dim(\mathbb{V}_1 \cap \mathbb{V}_3) + \dim(\mathbb{V}_2 \cap \mathbb{V}_3)\} +$$
$$+ \dim(\mathbb{V}_1 \cap \mathbb{V}_2 \cap \mathbb{V}_3) = \dim(\mathbb{V}_1 + \mathbb{V}_2 + \mathbb{V}_3) \ .$$

Теорема 3.3.

$$\mathcal{L}(X_1, \dots, X_m) + \mathcal{L}(Y_1, \dots, Y_\ell) = \mathcal{L}(X_1, \dots, X_m, Y_1, \dots, Y_\ell) \ .$$

Пример 3.1. *Найти базис суммы и размерность пересечения*

$$\mathbb{V}_1=\mathcal{L}\left(\begin{bmatrix}0\\1\\1\\1\end{bmatrix},\begin{bmatrix}1\\1\\1\\2\end{bmatrix},\begin{bmatrix}-2\\0\\1\\1\end{bmatrix}\right)\quad \textit{и}\quad \mathbb{V}_2=\mathcal{L}\left(\begin{bmatrix}-1\\3\\2\\-1\end{bmatrix},\begin{bmatrix}1\\1\\0\\-1\end{bmatrix}\right)$$

Решение. Действуя согласно теореме 3.3, составляем матрицу из всех векторов

$$\begin{pmatrix}0&1&-2&-1&1\\1&1&0&3&1\\1&1&1&2&0\\1&2&1&-1&-1\end{pmatrix}$$

и ищем ее ранг методом окаймляющих миноров. Имеем: rank $=3$ и обеспечивается первыми тремя столбцами.

Ответ. Базис $\mathbb{V}_1+\mathbb{V}_2$ составляют векторы X_1,X_2,X_3; $\dim(\mathbb{V}_1\cap\mathbb{V}_2)=3+2-3=2$.

Алгоритм нахождения базиса $\mathcal{L}(X_1,\dots,X_m)\cap\mathcal{L}(Y_1,\dots,Y_\ell)$ проиллюстрируем на примере.

Пример 3.2. *Найти базис* $\mathbb{V}_1\cap\mathbb{V}_2$ *при*

$$\mathbb{V}_1=\mathcal{L}\left(\underset{X_1}{\begin{bmatrix}1\\-1\\1\\-1\\1\end{bmatrix}},\underset{X_2}{\begin{bmatrix}1\\2\\1\\2\\1\end{bmatrix}},\underset{X_3}{\begin{bmatrix}0\\1\\0\\1\\0\end{bmatrix}}\right),\quad \mathbb{V}_2=\mathcal{L}\left(\underset{Y_1}{\begin{bmatrix}1\\0\\0\\0\\1\end{bmatrix}},\underset{Y_2}{\begin{bmatrix}1\\1\\0\\1\\1\end{bmatrix}},\underset{Y_3}{\begin{bmatrix}0\\1\\1\\1\\0\end{bmatrix}}\right).$$

Решение. 1. Сначала найдем базисы каждого из подпространств:

$$\dim\mathbb{V}_1=2,\ \mathbb{V}_1=\mathcal{L}(X_1,X_2)\ ;\ \dim\mathbb{V}_2=3,\ \mathbb{V}_2=\mathcal{L}(Y_1,Y_2,Y_3)\ .$$

2. Произвольный вектор $Z\in\mathbb{R}^5$, принадлежащий $\mathbb{V}_1\cap\mathbb{V}_2$, должен раскладываться по базису каждого из подпространств:

$$Z=\alpha_1X_1+\alpha_2X_2=\beta_1Y_1+\beta_2Y_2+\beta_3Y_3\ .$$

Для определения неизвестных значений координат составляем систему уравнений

$$\begin{array}{c}\begin{array}{ccccc}X_1&X_2&&&\end{array}\\ \begin{pmatrix}1&1&-1&-1&0\\-1&2&0&-1&-1\\1&1&0&0&-1\\-1&2&0&-1&-1\\1&1&-1&-1&0\end{pmatrix}\\ \begin{array}{ccccc}&&-Y_1&-Y_2&-Y_3\end{array}\end{array}\begin{pmatrix}\alpha_1\\\alpha_2\\\beta_1\\\beta_2\\\beta_3\end{pmatrix}=\mathbb{O}_{5\times1}$$

и решаем ее по методу Гаусса.

$$\left(\begin{array}{rrrrr} 1 & 1 & -1 & -1 & 0 \\ 0 & 3 & -1 & -2 & -1 \\ 0 & 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)\left(\begin{array}{c} \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{array}\right) = \mathbb{O} \ \Rightarrow \ \begin{array}{rrr|rr} \multicolumn{3}{c}{\overbrace{\qquad\qquad\qquad}^{\text{зависимые}}} & \multicolumn{2}{c}{\overbrace{\qquad\quad}^{\text{основные}}} \\ \alpha_1 & \alpha_2 & \beta_1 & \beta_2 & \beta_3 \\ \hline -1/3 & 1/3 & -1 & 1 & 0 \\ 1/3 & 2/3 & 1 & 0 & 1 \end{array}$$

3. Получившиеся значения координат позволяют выразить базис пересечения — либо через базис подпространства $\mathbb{V}_1$ (если использовать α_1, α_2), либо через базис подпространства $\mathbb{V}_2$ (если использовать $\beta_1, \beta_2, \beta_3$). Например,

$$Z_1 = -1/3 X_1 + 1/3 X_2 = [0,1,0,1,0]^\top, \ Z_2 = 1/3 X_1 + 2/3 X_2 = [1,1,1,1,1]^\top .$$

Ответ. Например, $\left\{[0,1,0,1,0]^\top, [1,1,1,1,1]^\top\right\}$.

Пример 3.3. *Найти базисы суммы и пересечения подпространств*

$$\mathbb{V}_1 = \left\{ X \in \mathbb{R}^4 \left| \begin{array}{lllll} 2x_1 & +x_2 & +4x_3 & +x_4 & = 0, \\ 2x_1 & +x_2 & +3x_3 & & = 0 \end{array} \right. \right\}$$

и

$$\mathbb{V}_2 = \left\{ X \in \mathbb{R}^4 \left| \begin{array}{lllll} 3x_1 & +2x_2 & -x_3 & -6x_4 & = 0, \\ 2x_1 & & +8x_3 & +7x_4 & = 0 \end{array} \right. \right\} .$$

Решение. Проще определить базис пересечения. Можно было бы свести задачу к случаю, рассмотренному в примере 3.2 — посредством определения **ф.с.р.** для каждой из систем уравнений. Однако, по здравому размышлению о способе задания каждого из подпространств, можно предложить более простой алгоритм. Множество векторов $X \in \mathbb{R}^4$, принадлежащих $\mathbb{V}_j$, выделяется из пространства $\mathbb{R}^4$ наложением на эти векторы ограничений в виде равенств, которым они должны удовлетворять. Теперь очевидно, что $X \in \mathbb{V}_1 \cap \mathbb{V}_2$ тогда и только тогда, когда X одновременно удовлетворяет ограничениям как одного подпространства так и другого. Иными словами, он должен быть решением объединенной системы линейных уравнений

$$\left\{ \begin{array}{lllll} 2x_1 & +x_2 & +4x_3 & +x_4 & = 0, \\ 2x_1 & +x_2 & +3x_3 & & = 0, \\ 3x_1 & +2x_2 & -x_3 & -6x_4 & = 0, \\ 2x_1 & & +8x_3 & +7x_4 & = 0. \end{array} \right.$$

Решаем ее по методу Гаусса:

$$\left\{ \begin{array}{lllll} 2x_1 & +x_2 & +4x_3 & +x_4 & = 0, \\ & 1/2 x_2 & -7x_3 & -15/2 x_4 & = 0, \\ & & x_3 & +x_4 & = 0 \end{array} \right. \Rightarrow \begin{array}{rrr|r} x_1 & x_2 & x_3 & x_4 \\ \hline 1 & 1 & -1 & 1 \end{array} .$$

Итак, $\dim(\mathbb{V}_1 \cap \mathbb{V}_2) = 1$ и базисным вектором может быть взят $[1,1,-1,1]^\top$.

Для нахождения базиса $\mathbb{V}_1+\mathbb{V}_2$ нам все же придется определить базисы каждого из подпространств, т.е. решить каждую из систем уравнений:

$$\begin{array}{rcl} \mathbb{V}_1 & = & \mathcal{L}\left([-{}^1\!/_2,\, 1,\, 0,\, 0]^\top,\, [{}^3\!/_2,\, 0,\, -1,\, 1]^\top\right) \ , \\ \mathbb{V}_2 & = & \mathcal{L}\left([-8,\, {}^{25}\!/_2,\, 1,\, 0]^\top,\, [-7,\, {}^{27}\!/_2,\, 0,\, 1]^\top\right) \ . \end{array}$$

Объединяем базисные векторы в одну матрицу:

$$\left(\begin{array}{rrrr} -{}^1\!/_2 & {}^3\!/_2 & -8 & -7 \\ 1 & 0 & {}^{25}\!/_2 & {}^{27}\!/_2 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{array}\right) \ .$$

Теорема 3.1 позволяет предсказать величину ее ранга: $2+2-1=3$. Осталось только выяснить какими столбцами этот ранг обеспечивается:

$$\left|\begin{array}{rrr} 1 & 0 & {}^{25}\!/_2 \\ 0 & -1 & 1 \\ 0 & 1 & 0 \end{array}\right| \ne 0 \ \Rightarrow$$

$$\Rightarrow \ \mathbb{V}_1+\mathbb{V}_2=\mathcal{L}\left([-{}^1\!/_2,\, 1,\, 0,\, 0]^\top,\, [{}^3\!/_2,\, 0,\, -1,\, 1]^\top, [-8,\, {}^{25}\!/_2,\, 1,\, 0]^\top\right) \ .$$

3.2 Линейное многообразие

Определение. Пусть $\mathbb{V}_1$ — линейное подпространство пространства $\mathbb{V}$, а X_0 — произвольный вектор из $\mathbb{V}$. Множество

$$\mathbb{M} \stackrel{\mathrm{def}}{=} X_0+\mathbb{V}_1=\left\{X_0+Y \middle| Y\in\mathbb{V}_1\right\}$$

называется **линейным многообразием**. **Размерностью** этого многообразия называется $\dim \mathbb{V}_1$: $\dim \mathbb{M} \stackrel{\mathrm{def}}{=} \dim \mathbb{V}_1$. Пересечение многообразий определяется традиционным способом, а сумма многообразий не определяется.

Наиболее часто встречающимся примером многообразия является множество столбцов пространства $\mathbb{R}^n$, удовлетворяющих системе уравнений

$$\left\{\begin{array}{rcl} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n & = & b_1, \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n & = & b_2, \\ \ldots & & \ldots \\ a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n & = & b_m \end{array}\right. \iff AX=\mathcal{B} \ . \qquad (3.5)$$

При $\mathcal{B}\ne\mathbb{O}_{m\times 1}$ множество решений линейного пространства не образует. Структуру этого множества описывала теорема 12.3 главы 4, раздела I: если система (3.5) совместна, то ее общее решение можно представить как сумму какого-то одного ее решения и общего решения соответствующей однородной системы (1.1). Иными словами, множество решений неоднородной системы (3.5) образует линейное многообразие:

$$\mathbb{M}=X_0+\mathcal{L}(X_1,\ldots,X_{n-\mathfrak{r}}) \ .$$

Здесь X_0 означает частное решение системы (3.5): $AX_0 = \mathcal{B}$, а $\{X_1, \dots, X_{n-\mathfrak{r}}\}$ — **ф.с.р.** для системы (1.1). О таком многобразии говорят иногда как о $(n-\mathfrak{r})$-**мерной плоскости** (или **гиперплоскости**), а при $n-\mathfrak{r}=1$ — как о **прямой**:

$$\mathbb{M} = X_0 + tX_1 \quad \text{при } t \in \mathbb{R} \ ;$$

в последнем случае вектор X_1 называют **направляющим вектором** этой прямой.

Геометрический смысл этих определений поясним на примере $\mathbb{R}^3$. Уравнение

$$a_1x_1 + a_2x_2 + a_3x_3 = b \tag{3.6}$$

— хотя бы при одном из чисел a_1, a_2 или a_3 отличном от нуля — определяет плоскость. Если $b = 0$, то эта плоскость проходит через начало координат. Множество векторов, имеющих начало в начале координат, а концы — в точках плоскости, образуют линейное подпространство $\mathbb{V} \subset \mathbb{R}^3$. Этот факт означает замкнутость множества $\mathbb{V}$ относительно сложения векторов и умножения их на скаляр (“растяжения”): линейная комбинация векторов из $\mathbb{V}$ снова будет лежать в $\mathbb{V}$. **Ф.с.р.** однородного уравнения $a_1x_1 + a_2x_2 + a_3x_3 = 0$ определяет базис этой плоскости.

Если же в уравнении (3.6) имеем $b \neq 0$, то множество точек, координаты которых удовлетворяют этому уравнению, снова образуют плоскость $\mathbb{M}$ — но эта плоскость уже не проходит через начало координат. Множество векторов, имеющих начало в начале координат, а концы — в точках этой плоскости, теперь *не является* замкнутым относительно операции сложения векторов: сумма двух таких векторов даст вектор, конец которого не принадлежит $\mathbb{M}$. Очевидно, что плоскость $\mathbb{M}$ параллельна плоскости $\mathbb{V}$, т.е. может быть получена из последней сдвигом на некоторый вектор X_0, конец которого лежит на плоскости $\mathbb{M}$.

Любой вектор $X \in \mathbb{M}$ может быть представлен в виде $X = X_0 + \alpha_1X_1 + \alpha_2X_2$, где векторы X_1 и X_2 — базисные на плоскости $\mathbb{V}$.

Если рассмотреть теперь систему уравнений

$$\left\{ \begin{array}{lcl} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 & = & b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 & = & b_2, \end{array} \right. \tag{3.7}$$

то она также задает многобразие в $\mathbb{R}^3$. Его можно интерпретировать как пересечение двух многобразий

$$\mathbb{M}_1 = \left\{ X \in \mathbb{R}^3 \middle| a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \right\}$$

и

$$\mathbb{M}_2 = \left\{ X \in \mathbb{R}^3 \middle| a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \right\}$$

— т.е. плоскостей в пространстве. Это пересечение может оказаться пустым ($\dim(\mathbb{M}_1 \cap \mathbb{M}_2) = 0$), а может оказаться прямой ($\dim(\mathbb{M}_1 \cap \mathbb{M}_2) = 1$) или плоскостью ($\dim(\mathbb{M}_1 \cap \mathbb{M}_2) = 2$) — в зависимости от величин

$$\operatorname{rank} \left(\begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{array} \right) \quad \text{и} \quad \operatorname{rank} \left(\begin{array}{cccc} a_{11} & a_{12} & a_{13} & b_1 \\ a_{21} & a_{22} & a_{23} & b_2 \end{array} \right) \ .$$

Если оба этих числа равны 2, то получаем прямую — множество векторов $\{X_0 + \alpha X_1 \mid \alpha \in \mathbb{R}\}$, где X_0 — произвольное решение системы (3.7) (точка на прямой), а X_1 — произвольное нетривиальное решение однородной системы

$$\left\{ \begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 & = & 0, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 & = & 0, \end{array} \right.$$

— задает направляющий вектор этой прямой.

Замечание. Исходя из этой геометрической интерпретации, часто говорят о векторе X произвольного линейного пространства $\mathbb{V}$ как о “точке пространства $\mathbb{V}$”. Иногда эту точку называют “концом вектора X”. Кроме удобства ассоциативного восприятия, этим словам не придается никакого формального смысла: понятие “конец вектора” отсутствует в аксиоматике линейного пространства.

Пример 3.4. *Найти прямую, проходящую через точку* $X_0 = [6, 5, 1, -1]^\top$ *и пересекающую плоскости*

$$\mathbb{M}_1 = \left\{ X \in \mathbb{R}^4 \left| \begin{array}{rrrrl} -x_1 & +2\,x_2 & +x_3 & & = 1, \\ x_1 & & & +x_4 & = 1 \end{array} \right. \right\}$$

и

$$\mathbb{M}_2 = \left\{ X = [4+t, 4+2\,t, 5+3\,t, 4+4\,t]^\top \middle| t \in \mathbb{R} \right\} \quad .$$

Решение. Ищем прямую в виде

$$\mathbb{M} = \left\{ X = X_0 + \alpha Z \middle| \alpha \in \mathbb{R} \right\} \quad ,$$

где $Z = [z_1, z_2, z_3, z_4]^\top$ — направляющий вектор искомой прямой, а α — параметр. Условие пересечения $\mathbb{M}$ и $\mathbb{M}_1$ состоит в том, что при некотором значении $\alpha = \alpha_1$ совпадают координаты на плоскости $\mathbb{M}_1$ и вдоль прямой $\mathbb{M}$.

$$\left\{ \begin{array}{lllll} -(6+\alpha_1 z_1) & +2\,(5+\alpha_1 z_2) & +(1+\alpha_1 z_3) & & = 1, \\ \;\;(6+\alpha_1 z_1) & & & +(-1+\alpha_1 z_4) & = 1 \end{array} \right. \iff$$

$$\iff \left\{ \begin{array}{ll} \alpha_1(-z_1 + 2\,z_2 + z_3) & = -4, \\ \alpha_1(z_1 + z_4) & = -4. \end{array} \right.$$

Рассмотрев эти уравнения как систему относительно α_1, выпишем **Н.** и **Д.** условие ее совместности:

$$\operatorname{rank} \left(\begin{array}{c} -z_1 + 2\,z_2 + z_3 \\ z_1 + z_4 \end{array} \right) = \operatorname{rank} \left(\begin{array}{cc} -z_1 + 2\,z_2 + z_3, & -4 \\ z_1 + z_4, & -4 \end{array} \right) \quad .$$

Последнее удовлетворяется тогда и только тогда, когда

$$\left| \begin{array}{cc} -z_1 + 2\,z_2 + z_3, & -4 \\ z_1 + z_4, & -4 \end{array} \right| = 0 \iff 2\,z_1 - 2\,z_2 - z_3 + z_4 = 0 \quad . \qquad (3.8)$$

Условие пересечения прямых $\mathbb{M}$ и $\mathbb{M}_2$ выписывается аналогично: система

$$\left\{\begin{array}{rl} 6+\alpha_2 z_1 & = 4+t, \\ 5+\alpha_2 z_2 & = 4+2\,t, \\ 1+\alpha_2 z_3 & = 5+3\,t, \\ -1+\alpha_2 z_4 & = 4+4\,t, \end{array}\right. \iff \left\{\begin{array}{rr} \alpha_2 z_1 - \quad t = & -2, \\ \alpha_2 z_2 - 2\,t = & -1, \\ \alpha_2 z_3 - 3\,t = & 4, \\ \alpha_2 z_4 - 4\,t = & 5 \end{array}\right.$$

должна быть совместна относительно α_2 и t. **Н.** и **Д.** условие:

$$\operatorname{rank}\begin{pmatrix} z_1 & -1 \\ z_2 & -2 \\ z_3 & -3 \\ z_4 & -4 \end{pmatrix} = \operatorname{rank}\begin{pmatrix} z_1 & -1 & -2 \\ z_2 & -2 & -1 \\ z_3 & -3 & 4 \\ z_4 & -4 & 5 \end{pmatrix}.$$

Поскольку число слева не превосходит 2, а справа — больше или равно 2, то имеет смысл выяснить при каких условиях на z_1, z_2, z_3, z_4 расширенная матрица имеет ранг в точности равный 2.

$$\begin{pmatrix} z_1 & z_2 & z_3 & z_4 \\ -1 & -2 & -3 & -4 \\ -2 & -1 & 4 & 5 \end{pmatrix} \to \begin{pmatrix} -1 & -2 & -3 & -4 \\ -2 & -1 & 4 & 5 \\ z_1 & z_2 & z_3 & z_4 \end{pmatrix} \to$$

$$\to \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 3 & 10 & 13 \\ 0 & z_2-2\,z_1 & z_3-3\,z_1 & z_4-4\,z_1 \end{pmatrix} \to$$

$$\to \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 3 & 10 & 13 \\ 0 & 0 & {}^{11}\!/_3 z_1 - {}^{10}\!/_3 z_2 + z_3 & {}^{14}\!/_3 z_1 - {}^{13}\!/_3 z_2 + z_4 \end{pmatrix}.$$

Ранг этой матрицы будет равен 2 тогда и только тогда, когда

$$\left\{\begin{array}{lllll} {}^{11}\!/_3 z_1 & -{}^{10}\!/_3 z_2 & +z_3 & & =0, \\ {}^{14}\!/_3 z_1 & -{}^{13}\!/_3 z_2 & & +z_4 & =0. \end{array}\right.$$

Объединив полученные условия с (3.8), составим систему для определения компонент направляющего вектора прямой:

$$\left\{\begin{array}{rrrrl} 2\,z_1 & -2\,z_2 & -z_3 & +z_4 & =0, \\ 11\,z_1 & -10\,z_2 & +3\,z_3 & & =0, \\ 14\,z_1 & -13\,z_2 & & +3\,z_4 & =0 \end{array}\right. \iff \left\{\begin{array}{rrrrl} 2\,z_1 & -2\,z_2 & -z_3 & +z_4 & =0, \\ z_1 & & +8\,z_3 & -5\,z_4 & =0, \\ & z_2 & +7\,z_3 & -4\,z_4 & =0 \end{array}\right.$$

$$\iff \left\{\begin{array}{rrrrl} z_1 & & +8\,z_3 & -5\,z_4 & =0, \\ & z_2 & +7\,z_3 & -4\,z_4 & =0, \\ & & -3\,z_3 & +3\,z_4 & =0 \end{array}\right. \Rightarrow \begin{array}{ccc|c} z_1 & z_2 & z_3 & z_4 \\ \hline -3 & -3 & 1 & 1 \end{array}$$

Ответ. $\mathbb{M} = \left\{ [6-3\,\alpha,\ 5-3\,\alpha,\ 1+\alpha,\ -1+\alpha]^\top \;\middle|\; \alpha \in \mathbb{R} \right\}.$

4 Прямая сумма линейных подпространств

Определение. Пусть $\mathbb{V}_1$ и $\mathbb{V}_2$ — подпространства лин.пространства $\mathbb{V}$. Говорят, что $\mathbb{V}$ **раскладывается в прямую сумму** подпространств $\mathbb{V}_1$ и $\mathbb{V}_2$ если любой $X \in \mathbb{V}$ может быть представлен в виде $X = X_1 + X_2$, где $X_1 \in \mathbb{V}_1, X_2 \in \mathbb{V}_2$ и такое представление единственно. Пишут $\mathbb{V} = \mathbb{V}_1 \oplus \mathbb{V}_2$.

Для геометрической интерпретации понятия прямой суммы обратимся к рисунку 1. В обоих случаях $\mathbb{V}_1 + \mathbb{V}_2 = \mathbb{R}^3$. В случае рис. 1 б) эта сумма не является прямой суммой, поскольку каждый вектор X (с началом — в начале координат) пространства может быть представлен не единственным способом в виде суммы X_1+X_2, где $X_1 \in \mathbb{V}_1$, а $X_2 \in \mathbb{V}_2$. Несколько сложнее показать, что в случае рис. 1 а) сумма будет прямой, но это можно сделать с помощью приведенной ниже теоремы 4.1.

Пример 4.1. *Линейное пространство квадратных матриц порядка n раскладывается в прямую сумму подпространств: подпространства симметричных матриц и подпространства кососимметричных матриц. В самом деле, для матрицы $A_{n\times n}$ справедливо разложение*

$$A = {}^1\!/_2 \left(A + A^{\top}\right) + {}^1\!/_2 \left(A - A^{\top}\right)$$

и в правой части первая скобка дает симметричную матрицу, а вторая — кососимметричную. Покажите, что не существует иного разложения матрицы A в сумму симметричной и кососимметричной.

Теорема 4.1. *Сумма $\mathbb{V} = \mathbb{V}_1 + \mathbb{V}_2$ будет прямой тогда и только тогда, когда*

$$\mathbb{V}_1 \cap \mathbb{V}_2 = \{\mathbb{O}\} \quad \textit{и} \quad \dim \mathbb{V}_1 + \dim \mathbb{V}_2 = \dim \mathbb{V} \ .$$

Доказательство . **Н.** Пусть $\mathbb{V} = \mathbb{V}_1 \oplus \mathbb{V}_2$ и пусть $X \in \mathbb{V}_1 \cap \mathbb{V}_2$, т.е. $X \in \mathbb{V}_1$ и $X \in \mathbb{V}_2$. Тогда $(-X) \in \mathbb{V}_2$, и $\mathbb{O} = X + (-X)$ может быть представлен в виде суммы векторов, взятых по одному из каждого подпространства. По предположению, должно быть $X = \mathbb{O}$. Следовательно, $\dim(\mathbb{V}_1 \cap \mathbb{V}_2) = 0$ и на основании формулы (3.1) имеем $\dim(\mathbb{V}_1 + \mathbb{V}_2) = \dim \mathbb{V}_1 + \dim \mathbb{V}_2$.

Д. Если $\dim(\mathbb{V}_1 \cap \mathbb{V}_2) = 0$, то $\mathbb{V}_1 \cap \mathbb{V}_2 = \{\mathbb{O}\}$. Для произвольного вектора $X \in \mathbb{V}$ имеет место представление в виде суммы $X = X_1 + X_2$, где $X_j \in \mathbb{V}_j$. Надо доказать единственность такого представления. Пусть имеется еще одно: $X = Y_1 + Y_2$, где $Y_j \in \mathbb{V}_j$. Тогда

$$\underbrace{X_1 - Y_1}_{\in \mathbb{V}_1} = \underbrace{Y_2 - X_2}_{\in \mathbb{V}_2} \quad \Longrightarrow \quad X_1 - Y_1 \in \mathbb{V}_1 \cap \mathbb{V}_2$$

Однако, по предположению, $\mathbb{V}_1 \cap \mathbb{V}_2 = \{\mathbb{O}\}$. Следовательно, $X_1 - Y_1 = \mathbb{O}$, но тогда и $X_2 - Y_2 = \mathbb{O}$. $\square$

Следствие 1. *Сумма $\mathbb{V} = \mathbb{V}_1 + \mathbb{V}_2$ будет прямой тогда и только тогда, когда базис $\mathbb{V}$ может быть получен объединением базисов $\mathbb{V}_j$.*

Замечание. Сформулированное таким образом утверждение содержится во всех учебниках по линейной алгебре. Тем не менее, с формальной точки зрения, оно неверно. В самом деле, пусть $\mathbb{V}_1 = \mathcal{L}(X_1, X_2)$, $\mathbb{V}_2 = \mathcal{L}(X_2, X_3)$ при **л.н.з.** $\{X_1, X_2, X_3\}$. Очевидно базис $\mathbb{V}_1 + \mathbb{V}_2 = \mathcal{L}(X_1, X_2, X_3)$ получается объединением базисов $\mathbb{V}_1$ и $\mathbb{V}_2$. В то же время $\mathbb{V}_1 \cap \mathbb{V}_2 \neq \{\mathbb{O}\}$. Причина возникновения этой ошибки кроется в содержании термина "объединение базисов". С точки зрения терминологии теории *множеств*, во множестве не может содержаться одинаковых элементов (во множестве они неразличимы). Однако мы с самого начала изложения допустили, что в *систему* векторов могут входить одинаковые, которые различаются порядком своего расположения (хотя это особо и не подчеркивалось, векторы в системе всегда пронумерованы). Исходя из этих соображений, объединение базисов $\mathbb{V}_1$ и $\mathbb{V}_2$ будет пониматься[2] в настоящем параграфе (и кое-где далее) как система векторов, в которую входят последовательно векторы базисов $\mathbb{V}_1$ и $\mathbb{V}_2$ — с допуском дублей. В рамках такой договоренности, для приведенного примера получим: объединение базисов линейных подпространств $\mathcal{L}(X_1, X_2)$ и $\mathcal{L}(X_2, X_3)$ представляет систему $\{X_1, X_2, X_2, X_3\}$, которая, очевидно, не является базисом. Таким образом сумма $\mathcal{L}(X_1, X_2) + \mathcal{L}(X_2, X_3)$ не является прямой и результат следствия остается справедливым.

Доказательство . Если $\mathbb{V} = \mathbb{V}_1 \oplus \mathbb{V}_2$, то $\forall X \in \mathbb{V}$ может быть представлен в виде линейной комбинации объединения базисов

$$\{\underbrace{X_1, \dots, X_{d_1}}_{\mathbb{V}_1}, \underbrace{Y_1, \dots, Y_{d_2}}_{\mathbb{V}_2}\} \tag{4.1}$$

Эта система векторов **л.н.з** поскольку $\mathbb{V}_1 \cap \mathbb{V}_2 = \{\mathbb{O}\}$.

Обратно, если система (4.1) составляет базис $\mathbb{V}_1 + \mathbb{V}_2$, то $\dim \mathbb{V} = \dim \mathbb{V}_1 + \dim \mathbb{V}_2$ и, на основании формулы (3.1), $\mathbb{V}_1 \cap \mathbb{V}_2 = \{\mathbb{O}\}$. □

Пример 4.2. *Доказать, что сумма подпространств*

$$\mathbb{V}_1 = \mathcal{L}\left(\begin{bmatrix} 2 \\ 3 \\ 11 \\ 5 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 5 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right) \quad \text{и} \quad \mathbb{V}_2 = \mathcal{L}\left(\begin{bmatrix} 2 \\ 1 \\ 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 5 \\ 2 \\ 6 \\ 2 \end{bmatrix} \right)$$

будет прямой и найти разложение вектора $Z = [2, 0, 0, 3]^\top$ по этим подпространствам.

Решение. Базисы $\mathbb{V}_1$ и $\mathbb{V}_2$ составляют соответственно системы $\{X_2, X_3\}$ и $\{Y_1, Y_2\}$, т.е. $\dim \mathbb{V}_1 = \dim \mathbb{V}_2 = 2$. На основании следствия 1, нам достаточно установить, что объединенная система $\{X_2, X_3, Y_1, Y_2\}$ **л.н.з.** Для

[2] Для математической строгости здесь требуется введение отдельного определения — аналога встреченной нами в §2 главы 4, раздела I *конкатенации*, ... но из соображений экономии лучше слегка ослабим однозначность терминологии.

этого достаточно проверить, что определитель матрицы

$$A = \begin{pmatrix} 1 & 0 & 2 & 1 \\ 1 & 1 & 1 & 1 \\ 5 & 1 & 3 & 3 \\ 2 & 1 & 2 & 4 \end{pmatrix}$$

отличен от нуля. Поскольку это условие выполнено, то сумма $\mathbb{V}_1 + \mathbb{V}_2$ — прямая и базис этой суммы состоит из взятых векторов. Для нахождения разложения вектора X по этому базису решаем систему уравнений

$$A \begin{bmatrix} \alpha_2 \\ \alpha_3 \\ \beta_1 \\ \beta_2 \end{bmatrix} = Z$$

и получаем единственное решение: $\alpha_2 = -1, \alpha_3 = -1, \beta_1 = 1\,, \beta_2 = 1$. Разложение $Z = Z_1 + Z_2$ составляют векторы $Z_1 = \alpha_2 X_2 + \alpha_3 X_3$ и $Z_2 = \beta_1 Y_1 + \beta_2 Y_2$.

Ответ. $Z = [-1, -2, -6, -3]^\top + [3, 2, 6, 6]^\top$.

Определение. Прямая сумма подпространств $\mathbb{V} = \mathbb{V}_1 \oplus \mathbb{V}_2 \oplus \cdots \oplus \mathbb{V}_N$ для $N \geq 3$ определяется по аналогии со случаем $N = 2$:

$$\forall X \in \mathbb{V} \quad \text{может быть представлен в виде} \quad X = \underbrace{X_1}_{\in \mathbb{V}_1} + \underbrace{X_2}_{\in \mathbb{V}_2} + \cdots + \underbrace{X_N}_{\in \mathbb{V}_N}$$

и такое представление единственно.

Теорема 4.2. *Сумма $\mathbb{V} = \mathbb{V}_1 + \mathbb{V}_2 + \cdots + \mathbb{V}_N$ будет прямой тогда и только тогда, когда выполнено хотя бы одно из следующих условий.*

А. *Представление $\mathbb{O}$ в виде суммы векторов из $\mathbb{V}_j$ единственно.*

Б. *Представление хотя бы одного $X \in \mathbb{V}$ в виде суммы векторов из $\mathbb{V}_j$ единственно.*

В. $V_j \cap \sum_{k \neq j} V_k = \{\mathbb{O}\}$ *для* $\forall j \in \{1, 2, \ldots, N\}$.

Г. *Объединение базисов $\mathbb{V}_j$ дает базис $\mathbb{V}$.*

Доказательство . Если сумма $\mathbb{V} = \mathbb{V}_1 + \mathbb{V}_2 + \cdots + \mathbb{V}_N$ прямая, то каждый вектор, в том числе и $\mathbb{O}$ представи́м в виде суммы векторов из $\mathbb{V}_j$ единственным образом. Обратно, пусть представление $\mathbb{O}$ в виде суммы векторов из $\mathbb{V}_j$ единственно: $\mathbb{O} = \mathbb{O} + \mathbb{O} + \cdots + \mathbb{O}$, но для какого-то $X \in \mathbb{V}$ имеют место два представления:

$$X = X_1 + X_2 + \cdots + X_N = Y_1 + Y_2 + \cdots + Y_N \quad \text{при} \quad \{X_j, Y_j\} \subset \mathbb{V}_j \ .$$

Но тогда

$$\mathbb{O} = \underbrace{(X_1 - Y_1)}_{\in \mathbb{V}_1} + \underbrace{(X_2 - Y_2)}_{\in \mathbb{V}_2} + \cdots + \underbrace{(X_N - Y_N)}_{\in \mathbb{V}_N} \ .$$

По предположению $X_1 = Y_1, X_2 = Y_2, \ldots, X_N = Y_N$.

Докажем $\mathsf{Б} \Rightarrow \mathsf{А}$. Пусть это не так: $\mathbb{O} = Y_1 + \cdots + Y_N$ при $Y_j \neq \mathbb{O}$. Тогда $X = X - \mathbb{O} = (X_1 - Y_1) + \cdots + (X_N - Y_N)$ и мы получили еще одно разложение для X, что противоречит предположению.

Докажем $\mathsf{А} \Rightarrow \mathsf{В}$.

$$\text{Если} \quad X \in \mathbb{V}_j \cap \sum_{k\neq j} \mathbb{V}_k \quad \text{то} \quad X = \underbrace{X_j}_{\in \mathbb{V}_j} = \sum_{k\neq j} \underbrace{X_k}_{\in \mathbb{V}_k}$$

Тогда $\mathbb{O} = X - X = X_j - \sum_{k\neq j} X_k$, т.е. нулевой вектор представи́м в виде линейной комбинации ненулевых векторов. Это противоречит предположению.

Докажем $\mathsf{В} \Rightarrow \mathsf{А}$. Если $\mathbb{O} = X_1 + \cdots + X_N$ при $X_j \neq \mathbb{O}$, то $X_j = -\sum_{k\neq j} X_k \in \mathbb{V}_j \cap \sum_{k\neq j} \mathbb{V}_k$, т.е. нарушается предположение пункта $\mathsf{В}$.

Докажем $\mathsf{А} \iff \mathsf{Г}$. Пусть

$$\left\{ X_1^{(j)}, \ldots, X_{d_j}^{(j)} \right\} \tag{4.2}$$

— базис $\mathbb{V}_j$. Составим линейную комбинацию

$$\sum_{j=1}^{N} \underbrace{\left(\alpha_{j1} X_1^{(j)} + \cdots + \alpha_{jd_j} X_{d_j}^{(j)} \right)}_{\stackrel{\rm def}{=} X_j \in \mathbb{V}_j} \tag{4.3}$$

и приравняем $\mathbb{O}$. Если имеет место $\mathsf{А}$, то все X_j являются нулевыми векторами. Но тогда $\alpha_{j1} = 0, \ldots, \alpha_{jd_j} = 0$, поскольку векторы (4.2) — базисные. Обратно, если имеет место $\mathsf{Г}$, то комбинация (4.3) может давать $\mathbb{O}$ только при нулевом наборе скаляров, что и влечет за собой $\mathsf{А}$. $\square$

Следствие 1. *Если* $\mathbb{V} = \mathbb{V}_1 \oplus \mathbb{V}_2 \oplus \cdots \oplus \mathbb{V}_N$ *то*

$$\dim \mathbb{V} = \dim \mathbb{V}_1 + \dim \mathbb{V}_2 + \cdots + \dim \mathbb{V}_N \ .$$

5 Относительная линейная независимость. Факторпространство

Пусть $\mathbb{V}$ — линейное пространство ($\dim \mathbb{V} = n$), а $\mathbb{V}_1$ — его подпространство ($\dim \mathbb{V}_1 = d_1$).

Определение. Говорят, что векторы $X_1, \ldots, X_k$ **линейно независимы относительно** $\mathbb{V}_1$ если

$$\text{из условия} \quad \alpha_1 X_1 + \cdots + \alpha_k X_k \in \mathbb{V}_1 \quad \text{следует} \quad \alpha_1 = \cdots = \alpha_k = 0 \ .$$

Теорема 5.1. *Обозначим* $Y_1, \ldots, Y_{d_1}$ *— произвольный базис* $\mathbb{V}_1$. *Векторы* $X_1, \ldots, X_k$ **л.н.з.** *относительно* $\mathbb{V}_1$ *тогда и только тогда, когда векторы* $X_1, \ldots, X_k, Y_1, \ldots, Y_{d_1}$ *линейно независимы.*

Доказательство . **Н.** Пусть $X_1, \dots, X_k$ **л.н.з.** относительно $\mathbb{V}_1$ и

$$\alpha_1 X_1 + \dots + \alpha_k X_k + \beta_1 Y_1 + \dots + \beta_{d_1} Y_{d_1} = \mathbb{O} \Rightarrow$$

$$\Rightarrow \alpha_1 X_1 + \dots + \alpha_k X_k = -\beta_1 Y_1 - \dots - \beta_{d_1} Y_{d_1} \in \mathbb{V}_1 \ .$$

По определению относительной линейной независимости $\alpha_1 = 0, \dots, \alpha_k = 0$, но тогда и $\beta_1 = 0, \dots, \beta_{d_1} = 0$, т.к. векторы $Y_1, \dots, Y_{d_1}$ — базисные.

Д. Пусть $X_1, \dots, X_k, Y_1, \dots, Y_{d_1}$ **л.н.з.** и $X = \alpha_1 X_1 + \dots + \alpha_k X_k \in \mathbb{V}_1$, тогда X можно разложить по базису $\mathbb{V}_1$: $X = \gamma_1 Y_1 + \dots + \gamma_{d_1} Y_{d_1}$. Имеем

$$\alpha_1 X_1 + \dots + \alpha_k X_k - \gamma_1 Y_1 - \dots - \gamma_{d_1} Y_{d_1} = \mathbb{O},$$

но тогда $\alpha_1 = 0, \dots, \alpha_k = 0, \gamma_1 = 0, \dots, \gamma_{d_1} = 0$ из-за независимости комбинируемых векторов. □

Пример 5.1. *Найти все значения параметра α, при которых векторы*

$$X_1 = [1,\ 2,\ \alpha,\ 1]^\top \ \text{и}\ X_2 = [1,\ \alpha,\ 2,\ 1]^\top$$

л.н.з. *относительно подпространства*

$$\mathbb{V}_1 = \left\{ X \in \mathbb{R}^4 \left| \begin{array}{ll} x_1 + 2\,x_2 - 3\,x_3 + 4\,x_4 & = 0, \\ x_1 + x_2 - x_3 - x_4 & = 0 \end{array} \right. \right\} \ .$$

Решение. Базисом подпространства $\mathbb{V}_1$ является произвольная **ф.с.р.** заданной системы однородных уравнений, например $Y_1 = [-1, 2, 1, 0]^\top$, $Y_2 = [6, -5, 0, 1]^\top$. Теорема 5.1 утверждает, что векторы X_1 и X_2 **л.н.з.** относительно $\mathbb{V}_1$ тогда и только тогда, когда векторы X_1, X_2, Y_1, Y_2 **л.н.з.** в обычном понимании. Последнее равносильно тому, что матрица, составленная из этих векторов, должна иметь ранг равный 4.

$$\operatorname{rank} \begin{pmatrix} 1 & 1 & -1 & 6 \\ 2 & \alpha & 2 & -5 \\ \alpha & 2 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix} = 4 \iff \begin{vmatrix} 1 & 1 & -1 & 6 \\ 2 & \alpha & 2 & -5 \\ \alpha & 2 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{vmatrix} = \alpha^2 - 10\,\alpha + 16 \ne 0 \ .$$

Ответ. $\alpha \notin \{2,\ 8\}$.

Определение. Говорят, что векторы $X_1, \dots, X_k$ **образуют базис** $\mathbb{V}$ **относительно** $\mathbb{V}_1$ если они **л.н.з.** относительно $\mathbb{V}_1$ и любой вектор $X \in \mathbb{V}$ можно представить в виде

$$X = c_1 X_1 + \dots + c_k X_k + Y, \quad \text{где} \quad Y \in \mathbb{V}_1 \ .$$

Теорема 5.2. *Обозначим $Y_1, \dots, Y_{d_1}$ — произвольный базис $\mathbb{V}_1$. Векторы $X_1, \dots, X_k$ образуют базис $\mathbb{V}$ относительно $\mathbb{V}_1$ тогда и только тогда, когда вектора $X_1, \dots, X_k, Y_1, \dots, Y_{d_1}$ образуют базис $\mathbb{V}$.*

Доказательство . Действительно, любой вектор $X \in \mathbb{V}$ выражается через векторы $X_1, \dots, X_k, Y_1, \dots, Y_{d_1}$. По теореме 5.1 для линейной независимости этих векторов **Н.** и **Д.** относительной линейной независимости $X_1, \dots, X_k$. □

Следствие 1. *Базис* $\mathbb{V}$ *строится дополнением базиса* $\mathbb{V}_1$ *векторами* $X_1, \ldots, X_k$ *линейно независимыми относительно* $\mathbb{V}_1$*. Поэтому*

$$\text{число векторов относительного базиса } = \dim \mathbb{V} - \dim \mathbb{V}_1 \ .$$

Определение. Говорят, что векторы $X_1, X_2 \in \mathbb{V}$ **сравнимы по подпространству** $\mathbb{V}_1$, если $X_1 - X_2 \in \mathbb{V}_1$; этот факт записывают:

$$X_1 \equiv_{\mathbb{V}_1} X_2 \ .$$

Все пространство $\mathbb{V}$ раскладывается на объединение подмножеств, или классов векторов, сравнимых по подпространству $\mathbb{V}_1$. Если $X_1 \equiv_{\mathbb{V}_1} X_2$, $Y_1 \equiv_{\mathbb{V}_1} Y_2$, то $\alpha X_1 + \beta Y_1 \equiv_{\mathbb{V}_1} \alpha X_2 + \beta Y_2$. Два разных класса не пересекаются и полностью определяются заданием любого своего представителя. Поэтому их обозначают $\overline{X_1}, \overline{X_2}, \ldots$

Определение. Множество классов, сравнимых по подпространству $\mathbb{V}_1$ называется **факторпространством** $\mathbb{V}$ **над** $\mathbb{V}_1$: $\mathbb{V}/\mathbb{V}_1$.

Теорема 5.3. *Факторпространство* $\mathbb{V}/\mathbb{V}_1$ *является линейным пространством, базис которого состоит из классов, порожденных векторами, образующими базис* $\mathbb{V}$ *относительно* $\mathbb{V}_1$*. Обратно, если из каждого базисного класса факторпространства взять по одному вектору, то получим базис* $\mathbb{V}$ *относительно* $\mathbb{V}_1$*.*

Доказательство . Положим:

$$\alpha \overline{X} + \beta \overline{Y} \stackrel{\text{def}}{=} \overline{\alpha X + \beta Y} \ . \tag{5.1}$$

Введенная таким образом определение корректно, т.е. не зависит от выбора представителей класса:

$$\text{если } X_1 \equiv_{\mathbb{V}_1} X, \ Y_1 \equiv_{\mathbb{V}_1} Y, \text{ то } \alpha X_1 + \beta Y_1 \equiv_{\mathbb{V}_1} \alpha X + \beta Y \ .$$

Легко проверяются свойства линейного пространства.

Далее,

$$\alpha_1 X_1 + \cdots + \alpha_k X_k \in \mathbb{V}_1 \quad \Longleftrightarrow \quad \alpha_1 X_1 + \cdots + \alpha_k X_k \equiv_{\mathbb{V}_1} \mathbb{O}$$

и, на основании (5.1):

$$\Longleftrightarrow \ \alpha_1 \overline{X_1} + \cdots + \alpha_k \overline{X_k} = \overline{\mathbb{O}} \ .$$

Линейная независимость $X_1, \ldots, X_k$ относительно $\mathbb{V}_1$ эквивалентна линейной независимости классов $\overline{X_1}, \ldots, \overline{X_k}$ факторпространства. □

Следствие 1. $\dim \mathbb{V}/\mathbb{V}_1 = \dim \mathbb{V} - \dim \mathbb{V}_1$.

6 Преобразование координат при замене базиса

Пусть $\mathbb{V}$ — линейное пространство, $\dim \mathbb{V} = n$. Пусть

$$\{X_1, \dots, X_n\} \quad \text{и} \quad \{\mathfrak{X}_1, \dots, \mathfrak{X}_n\}$$

— два произвольных базиса $\mathbb{V}$ (“старый” и “новый”).

Задача. Вывести соотношения, связывающие координаты произвольного вектора $X \in \mathbb{V}$ в старом и новом базисах:

$$X = x_1 X_1 + \dots + x_n X_n = \mathfrak{x}_1 \mathfrak{X}_1 + \dots + \mathfrak{x}_n \mathfrak{X}_n \ .$$

Предположим, что нам известны координаты векторов нового базиса в старом:

$$\left\{ \begin{array}{lcl} \mathfrak{X}_1 & = & c_{11} X_1 + c_{21} X_2 + \dots + c_{n1} X_n, \\ \mathfrak{X}_2 & = & c_{12} X_1 + c_{22} X_2 + \dots + c_{n2} X_n, \\ \dots & & \qquad\qquad \dots \\ \mathfrak{X}_n & = & c_{1n} X_1 + c_{2n} X_2 + \dots + c_{nn} X_n. \end{array} \right. \tag{6.1}$$

Определение. Матрица

$$C = \left(\begin{array}{cccc} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \dots & & & \dots \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{array} \right),$$

по столбцам которой стоят координаты новых базисных векторов в старом базисе называется **матрицей перехода от старого базиса к новому**.

Теорема 6.1. *Матрица C неособенная.*

Доказательство . Сначала покажем справедливость утверждения в частном случае $\mathbb{V} = \mathbb{R}^n$. Вектора нового и старого базисов являются столбцами из n вещественных чисел, и равенства (6.1) можно переписать в матричном виде:

$$[\mathfrak{X}_1, \dots, \mathfrak{X}_n] = [X_1, \dots, X_n] \cdot C \ . \tag{6.2}$$

Поскольку системы $\{X_1, \dots, X_n\}$ и $\{\mathfrak{X}_1, \dots, \mathfrak{X}_n\}$ — базисные, то

$$\det [X_1, \dots, X_n] \ne 0, \quad \det [\mathfrak{X}_1, \dots, \mathfrak{X}_n] \ne 0 \ .$$

Из (6.2) тогда следует, что $\det C \ne 0$.

Теперь докажем теорему для случая произвольного пространства. Если $\det C = 0$, то столбцы матрицы C линейно зависимы, т.е. существует линейная комбинация

$$\alpha_1 c_{j1} + \dots + \alpha_n c_{jn} = 0 \quad \forall j \in \{1, \dots, n\}$$

и при некотором $\alpha_k \ne 0$. Но тогда из формул (6.1) следует, что

$$\alpha_1 \mathfrak{X}_1 + \dots + \alpha_n \mathfrak{X}_n = \mathbb{O} \ ,$$

что противоречит линейной независимости системы $\{\mathfrak{X}_1, \dots, \mathfrak{X}_n\}$. $\square$

Пример 6.1. *Найти матрицу перехода*

$$
\begin{array}{cc}
\textit{от базиса} & \textit{к базису} \\
[1,1,0,0,0]\,, & [1,1,1,1,1]\,, \\
[1,0,1,0,0]\,, & [1,1,1,1,0]\,, \\
[1,0,0,1,0]\,, & [1,1,1,0,0]\,, \\
[1,0,0,0,1]\,, & [1,1,0,0,0]\,, \\
[1,1,1,1,1] & [1,0,0,0,0]\,.
\end{array}
$$

Решение. Можно попытаться найти элементы матрицы C напрямую — устанавливая формулы (6.1). В нашем конкретном примере это не очень трудно сделать — первый и четвертый столбцы вообще очевидны ($\mathfrak{X}_1 = X_5$, $\mathfrak{X}_4 = X_1$). Но мы пойдем по формальному пути и воспользуемся определяющим матричным соотношением (6.2). Поставим координаты базисных векторов по столбцам соответствующих матриц:

$$
C = \left(\begin{array}{ccccc} 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{array} \right)^{-1} \left(\begin{array}{ccccc} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{array} \right) =
$$

$$
= \frac{1}{3} \left(\begin{array}{rrrrr} 1 & 2 & -1 & -1 & -1 \\ 1 & -1 & 2 & -1 & -1 \\ 1 & -1 & -1 & 2 & -1 \\ 1 & -1 & -1 & -1 & 2 \\ -1 & 1 & 1 & 1 & 1 \end{array} \right) \left(\begin{array}{ccccc} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{array} \right) =
$$

$$
= \left(\begin{array}{ccccc} 0 & 1/3 & 2/3 & 1 & 1/3 \\ 0 & 1/3 & 2/3 & 0 & 1/3 \\ 0 & 1/3 & -1/3 & 0 & 1/3 \\ 0 & -2/3 & -1/3 & 0 & 1/3 \\ 1 & 2/3 & 1/3 & 0 & -1/3 \end{array} \right) .
$$

$\triangle$

Теперь будем решать задачу, поставленную в начале параграфа. С помощью формул (6.1) получаем:

$$
\begin{array}{lcl}
X = x_1X_1 + \cdots + x_nX_n & = & \mathfrak{x}_1\mathfrak{X}_1 + \cdots + \mathfrak{x}_n\mathfrak{X}_n = \\
 & = & \mathfrak{x}_1(c_{11}X_1 + c_{21}X_2 + \cdots + c_{n1}X_n) + \\
 & + & \mathfrak{x}_2(c_{12}X_1 + c_{22}X_2 + \cdots + c_{n2}X_n) + \\
 & + & \cdots + \\
 & + & \mathfrak{x}_n(c_{1n}X_1 + c_{2n}X_2 + \cdots + c_{nn}X_n) = \\
= (c_{11}\mathfrak{x}_1 + c_{12}\mathfrak{x}_2 + \cdots + c_{1n}\mathfrak{x}_n)X_1 & + & \cdots + (c_{n1}\mathfrak{x}_1 + c_{n2}\mathfrak{x}_2 + \cdots + c_{nn}\mathfrak{x}_n)X_n
\end{array}
$$

Поскольку координаты вектора определяются однозначно (теорема 2.6), получаем равенства

$$\left\{\begin{array}{lcl} x_1 & = & c_{11}\mathfrak{x}_1 + c_{12}\mathfrak{x}_2 + \dots + c_{1n}\mathfrak{x}_n, \\ x_2 & = & c_{21}\mathfrak{x}_1 + c_{22}\mathfrak{x}_2 + \dots + c_{2n}\mathfrak{x}_n, \\ \dots & & \dots \\ x_n & = & c_{n1}\mathfrak{x}_1 + c_{n2}\mathfrak{x}_2 + \dots + c_{nn}\mathfrak{x}_n \end{array}\right. \quad \text{или} \quad \left(\begin{array}{c} x_1 \\ x_2 \\ \dots \\ x_n \end{array}\right) = C \left(\begin{array}{c} \mathfrak{x}_1 \\ \mathfrak{x}_2 \\ \dots \\ \mathfrak{x}_n \end{array}\right) \tag{6.3}$$

Теорема 6.2. *Координаты вектора в старом и новом базисах связаны посредством матрицы перехода C соотношением* (6.3).

Практическое значение последнего результата невелико, т.к. нас больше интересуют новые координаты.

Следствие 1. *Новые координаты выражаются через старые по формуле*

$$\left(\begin{array}{c} \mathfrak{x}_1 \\ \mathfrak{x}_2 \\ \dots \\ \mathfrak{x}_n \end{array}\right) = C^{-1} \left(\begin{array}{c} x_1 \\ x_2 \\ \dots \\ x_n \end{array}\right),$$

при этом матрицу C^{-1} можно интерпретировать как матрицу перехода от нового базиса к старому.

Упражнение 6.1. *Пусть в некотором "новейшем" базисе $\mathcal{X}_1, \dots, \mathcal{X}_n$ пространства $\mathbb{V}$ вектор X имеет координаты $(\varkappa_1, \dots, \varkappa_n)$. Как они связаны с координатами $(x_1, \dots, x_n)$, если известна матрица D перехода от базиса $\mathfrak{X}_1, \dots, \mathfrak{X}_n$ к базису $\mathcal{X}_1, \dots, \mathcal{X}_n$?*

Упражнение 6.2. *Проверить, что каждая из систем векторов*

$$\left\{ \left[\begin{array}{c} 1 \\ 2 \\ 1 \end{array}\right], \left[\begin{array}{c} 2 \\ 3 \\ 3 \end{array}\right], \left[\begin{array}{c} 3 \\ 7 \\ 1 \end{array}\right] \right\} \quad \text{и} \quad \left\{ \left[\begin{array}{c} 3 \\ 1 \\ 4 \end{array}\right], \left[\begin{array}{c} 5 \\ 2 \\ 1 \end{array}\right], \left[\begin{array}{c} 1 \\ 1 \\ -6 \end{array}\right] \right\}$$

является базисной в $\mathbb{R}^3$, и найти координаты вектора X

а) *во втором базисе, если $x_1 = 1, x_2 = 0, x_3 = 1$ — его координаты в первом;*

б) *в первом базисе, если $\mathfrak{x}_1 = 1, \mathfrak{x}_2 = -1, \mathfrak{x}_3 = 1$ — его координаты во втором.*

Упражнение 6.3. *Доказать, что для пространства $\mathbb{P}_3$ каждая из систем*

$$\begin{array}{rcl} \mathfrak{N} & = & \{1,\, x - x_1,\, (x-x_1)(x-x_2),\, (x-x_1)(x-x_2)(x-x_3)\}\,; \\ \mathfrak{L} & = & \big\{(x-x_2)(x-x_3)(x-x_4),\, (x-x_1)(x-x_3)(x-x_4), \\ & & \qquad (x-x_1)(x-x_2)(x-x_4),\, (x-x_1)(x-x_2)(x-x_3)\big\} \end{array}$$

является базисной при x_1, x_2, x_3, x_4 различных. Найти координаты произвольного полинома $p(x)$ в этих базисах. (Проиллюстрировать на примере $x_1 = 0, x_2 = 1, x_3 = 2, x_4 = 3$ и $p(x) = x^3 + x^2 - 3x + 1$, $p(x) = x^3 - x^2 - x + 1$). Найти матрицу перехода от первого базиса ко второму.

Глава 2. Евклидовы пространства

Одной из самых важных задач геометрии является задача измерения расстояния между двумя объектами. В произвольном линейном пространстве мы пока не можем определить насколько "близки" между собой объекты. В настоящей главе понятие расстояния между двумя векторами — элементами линейного пространства — будет вводиться посредством скалярного произведения векторов. Насколько обоснован такой порядок введения понятий: *скалярное произведение* $\to$ *длина*? Ведь в аналитической геометрии последовательность кажется более "естественной": скалярное произведение двух векторов X и Y определялось как произведение длин этих векторов на косинус угла между ними: $(X,Y) \stackrel{\text{def}}{=} |X|\cdot|Y|\cdot\cos(\widehat{X,Y})$. Тем не менее, не встретится формального противоречия и на обратном пути: если допустить, что скалярное произведение *любых* двух векторов может быть как-то вычислено (например, в $\mathbb{R}^3$ по формуле $(X,Y) \stackrel{\text{def}}{=} x_1y_1+x_2y_2+x_3y_3$ при заданных прямоугольных координатах (x_1,x_2,x_3) и (y_1,y_2,y_3) векторов X и Y), то и длину векторов и угол между ними можно выразить через подходящие скалярные произведения ($|X| = \sqrt{(X,X)}$, $\widehat{X,Y} = \arccos(X,Y)/\sqrt{(X,X)(Y,Y)}$).

1 Определения

1.1 Определения и примеры

Определение. Линейное пространство $\mathbb{E}$ называется **евклидовым**, если в этом пространстве определена операция, ставящая в соответствие паре векторов $X \in \mathbb{E}$ и $Y \in \mathbb{E}$ вещественное число, называемое **скалярным произведением векторов** X и Y, и обозначаемое (X,Y); при этом операция подчиняется аксиомам:

1. $(X,Y) = (Y,X)$ для $\{X,\,Y\} \subset \mathbb{E}$;
2. $(X_1+X_2,Y) = (X_1,Y)+(X_2,Y)$ для $\{X_1,\,X_2,\,Y\} \subset \mathbb{E}$;
3. $(\lambda\,X,Y) = \lambda\,(X,Y)$ для $\{X,Y\} \subset \mathbb{E},\ \lambda \in \mathbb{R}$;
4. $(X,X) > 0$ для $\forall X \neq \mathbb{O}$, $(\mathbb{O},\mathbb{O}) = 0$.

Из аксиом 1 и 2 вытекает свойство линейности скалярного произведения и по второму вектору:

2.′ $(X,Y_1+Y_2) = (X,Y_1)+(X,Y_2)$ для $\{X,Y_1,\,Y_2\} \subset \mathbb{E}$.

Теперь рассмотрим примеры введения скалярного произведения в пространствах из предыдущей главы.

Пространство $\mathbb{R}^n$.

$$(X,Y) \stackrel{\text{def}}{=} \sum_{j=1}^{n} x_jy_j \quad \text{при} \quad X = [x_1,\dots,x_n]^\top,\ Y = [y_1,\dots,y_n]^\top \ . \qquad (1.1)$$

Легко проверить выполнимость аксиом $1 - -4$. Однако определение скалярного произведения формулой (1.1) вовсе не является единственно допустимым, его можно ввести и другим способом. В самом деле, рассмотрим квадратную матрицу A порядка n и положим

$$\begin{array}{rcl} (X,Y) \stackrel{\text{def}}{=} X^{\top}AY & = & a_{11}x_1y_1 + a_{12}x_1y_2 + \dots + a_{1n}x_1y_n + \\ & + & a_{21}x_2y_1 + a_{22}x_2y_2 + \dots + a_{2n}x_2y_n + \\ & + & \dots \qquad + \\ & + & a_{n1}x_ny_1 + a_{n2}x_ny_2 + \dots + a_{nn}x_ny_n \quad . \end{array} \tag{1.2}$$

(Здесь векторы X и Y рассматриваются как столбцы.)

Выясним, при каких условиях на матрицу A операция, заданная формулой (1.2), определяет именно скалярное произведение. Аксиома 1 выполняется тогда и только тогда, когда матрица A симметрична: $A = A^{\top}$. В самом деле, если это условие выполнено, то $(X,Y) = X^{\top}AY = \left(X^{\top}AY\right)^{\top} = Y^{\top}A^{\top}X = Y^{\top}AX = (Y,X)$. Обратно, если условие $(X,Y) = (Y,X)$ выполняется для *любых* векторов X и Y, то оно выполняется и для

$$X_i \stackrel{\text{def}}{=} [\underbrace{0,\dots,0,1}_{i},0,\dots,0]^{\top} \quad \text{и} \quad Y_j \stackrel{\text{def}}{=} [\underbrace{0,\dots,0,1}_{j},0,\dots,0]^{\top} \ .$$

Тогда $(X_i,Y_j) = a_{ij}$, а $(Y_j,X_i) = a_{ji}$. Из условия равенства этих чисел и следует $A = A^{\top}$.

Теперь добьемся выполнения аксиомы 4: $X^{\top}AX > 0$ при $X \neq \mathbb{O}$. Это условие равносильно положительной определенности квадратичной формы $X^{\top}AX$. Критерий Сильвестра[3] гарантирует положительную определенность если положительны все главные миноры матрицы A. Аксиомы 2 и 3 выполняются автоматически. Следовательно формула (1.2) действительно определяет скалярное произведение в $\mathbb{R}^n$ если матрица A симметрична и все ее главные миноры положительны. Случай (1.1) получается из (1.2) при $A = E$.

<u>Пространство $\mathbb{P}_n$</u> (полиномов степени не выше n). Скалярное произведение полиномов $p(x) = a_0x^n + a_1x^{n-1} + \dots + a_n$ и $q(x) = b_0x^n + b_1x^{n-1} + \dots + b_n$ введем формулой

$$(p(x),q(x)) \stackrel{\text{def}}{=} \sum_{j=0}^{n} a_jb_j \ . \tag{1.3}$$

Легко проверить справедливость аксиом 1 – 4.

В том же пространстве укажем еще один способ задания скалярного произведения

$$(p(x),q(x)) \stackrel{\text{def}}{=} \int_a^b p(t)q(t)dt \ . \tag{1.4}$$

при некоторых фиксированных вещественных константах a и b, $a < b$. Аксиомы 1 – 3 следуют из свойств определенного интеграла. Для установления аксиомы IV прежде всего заметим, что для $\forall p(x) \in \mathbb{P}_n$ будет

[3] Раздел I, глава 7, теорема 4.2.

выполнено $\int_a^b p^2(t)d\,t \geq 0$. Предположим, что существует $p(x) \not\equiv 0$ такой, что $\int_a^b p^2(t)d\,t = 0$. По основной теореме высшей алгебры, число корней полинома $p(x)$, расположенных на $[a, b]$ конечно, и, следовательно, можно подобрать такой интервал $[a_1, b_1] \subset [a, b]$, на котором корней вовсе нет. На основании свойства аддитивности интеграла:

$$0 = \int_a^b p^2(t)d\,t = \int_a^{a_1} p^2(t)d\,t + \int_{a_1}^{b_1} p^2(t)d\,t + \int_{b_1}^{b} p^2(t)d\,t \ .$$

Из того, что каждое слагаемое неотрицательно следует, что все они должны обращаться в нуль. С другой стороны, по теореме о среднем имеем

$$0 = \int_{a_1}^{b_1} p^2(t)d\,t = p^2(\xi)(b_1 - a_1) \ , \text{ где } \xi \in [a_1, b_1] \ .$$

Следовательно, $p(\xi) = 0$, т.е. полином $p(x)$ имеет корень на $[a_1, b_1]$, что противоречит предположению.

Упражнение 1.1. *В пространстве $\mathbb{P}_n$ выделяется подпространство полиномов таких, что $p(a) = p(b)$ при $a < b$. Можно ли в этом подпространстве определить скалярное произведение формулой*

$$(p(x), q(x)) \stackrel{\text{def}}{=} \int_a^b p(t)d\,q(t) \ ?$$

Линейное пространство квадратных матриц порядка n.

Определение. Для матрицы A ее **следом**[4] называется сумма ее диагональных элементов:

$$\operatorname{Sp} A \stackrel{\text{def}}{=} a_{11} + \dots + a_{nn} \ .$$

Легко проверяются свойства:

$$\operatorname{Sp}(\alpha A) = \alpha \operatorname{Sp}(A), \quad \operatorname{Sp}(A + B) = \operatorname{Sp} A + \operatorname{Sp} B, \quad \operatorname{Sp}(AB) = \operatorname{Sp}(BA) \ .$$

Скалярное произведение матриц введем формулой

$$(A, B) \stackrel{\text{def}}{=} \operatorname{Sp}\left(A \cdot B^\top\right) = \sum_{j,k=1}^{n} a_{jk} b_{jk} \ . \tag{1.5}$$

Если вспомнить пример 1.2 из главы 1, то легко видеть, что это определение соответствует определению (1.1) скалярного произведения в $\mathbb{R}^n$. Свойства 1 – 3 скалярного произведения следуют из свойств Sp, а равенство

$$\operatorname{Sp}\left(A \cdot A^\top\right) = \sum_{j,k=1}^{n} a_{jk}^2$$

доказывает свойство 4.

[4]**Spur**(*нем.*) — след; в англоязычной литературе обозначается $\mathbf{tr}(A)$ — от английского trace.

1.2 Свойства

Теорема 1.1. *Имеет место* **неравенство Коши–Буняковского***:*

$$(X,Y)^2 \le (X,X)(Y,Y) \quad \textit{для } \forall\{X,Y\} \subset \mathbb{E} \ . \tag{1.6}$$

Доказательство для случая $\mathbb{R}^n$ следует из неравенства Коши[5]. Для доказательства общего случая используем одну вспомогательную конструкцию. Из аксиомы 4 следует, что для $\forall \lambda \in \mathbb{R}$ будет выполнено $(\lambda X - Y,\, \lambda X - Y) \ge 0$. Имеем:

$$0 \le (\lambda X - Y,\, \lambda X - Y) \le \lambda^2 (X,X) - 2\lambda (X,Y) + (Y,Y) \ .$$

Квадратное относительно λ неравенство будет выполнено при всех вещественных значениях этого параметра тогда и только тогда, когда дискриминант квадратного трехчлена будет отрицателен:

$$\mathcal{D} = (X,Y)^2 - (X,X)(Y,Y) \le 0 \ .$$

□

Определение. **Длиною** вектора X в евклидовом пространстве $\mathbb{E}$ называется число

$$|X| \stackrel{\text{def}}{=} \sqrt{(X,X)} \ .$$

Здесь квадратный корень понимается как корень арифметический: $|X| \ge 0$.

Замечание. Это формальное определение сразу же входит в противоречие со здравым смыслом: длина вектора в линейном пространстве оказывается **не фиксированной**, раз и навсегда, **величиной** — как мы привыкли в геометрии, — но зависящей от способа задания скалярного произведения в этом пространстве. На самом деле это противоречие не мешает нам даже в повседневной жизни. Так, никому и в голову не приходит, что расстояние между двумя городами фактически вычисляется "нечестно": не по прямой, их соединяющей (и, следовательно, проходящей сквозь Землю), а по дуге, проходящей по поверхности земного эллипсоида ... Вопрос о том, какое расстояние является *более истинным* подменяется другим: какое из расстояний позволяет решить ту или иную практическую задачу.

Пример 1.1. *В* $\mathbb{R}^n$ *при скалярном произведении, заданном формулой* (1.1), *длина вектора* $X = [x_1, \dots, x_n]$ *определяется привычным для геометрии способом:* $|X| = \sqrt{x_1^2 + \dots + x_n^2}$.

Замечание. В пространстве $\mathbb{P}_n$ со скалярным произведением (1.4) для избежания путаницы не используют словосочетание "длина полинома $p(x)$"; вместо этого величину $\sqrt{\int_a^b p^2(t)dt}$ называют **нормой** этого **полинома** и обозначают $\|p(x)\|$. Это же название распространяется и в линейное

[5] Раздел I, глава 4, теорема 8.2.

пространство квадратных матриц со скалярным произведением заданным формулой (1.5): величина

$$\|A\| \stackrel{\text{def}}{=} \sqrt{\mathrm{Sp}\,(A \cdot A^{\top})} = \sqrt{\sum_{j,k=1}^{n} a_{jk}^2}$$

называется **нормой матрицы** $A = [a_{jk}]_{j,k=1}^{n}$.

С помощью введенного определения неравенство (1.6) можно переписать в виде

$$|(X,Y)| \le |X| \cdot |Y| \quad \text{для } \forall \{X,Y\} \subset \mathbb{E} \ , \tag{1.7}$$

где $|\ \ |$ в левой части означает модуль, а в правой части — длину.

Теорема 1.2. *Имеет место неравенство*

$$|X+Y| \le |X| + |Y| \quad \textit{для } \forall \{X,Y\} \subset \mathbb{E} \ . \tag{1.8}$$

Доказательство .

$$\begin{array}{lll} 0 & \le & (X+Y,\, X+Y) = (X,X) + 2(X,Y) + (Y,Y) \stackrel{(1.7)}{\le} \\ & \stackrel{(1.7)}{\le} & |X|^2 + 2\,|X| \cdot |Y| + |Y|^2 = (|X| + |Y|)^2 \ . \end{array} \tag{1.9}$$

□

Определение. **Углом между векторами** X **и** Y называется угол

$$\varphi \stackrel{\text{def}}{=} \arccos \frac{(X,Y)}{|X| \cdot |Y|} \ .$$

Ввиду неравенства (1.7) это определение законно: дробь под знаком арккосинуса не превосходит 1 по абсолютной величине.

Определение. Векторы X и Y называются **ортогональными**: $X \bot Y$ если угол между ними равен $\pi/2$, или, что то же, $(X,Y) = 0$.

Теорема 1.3. *Если* $X \bot Y$, *то* $|X+Y|^2 = |X|^2 + |Y|^2$.

Доказательство следует из формулы (1.9). □

Замечание. Неравенство (1.8) иногда называется “неравенством треугольника”, а теорема 1.3 — “теоремой Пифагора”: по аналогии с известным в $\mathbb{R}^2$ результатом.

Следствие 1. *Если векторы* $X_1, \dots, X_k$ *попарно взаимно ортогональны, то*

$$|X_1 + \dots + X_k|^2 = |X_1|^2 + \dots + |X_k|^2 \ .$$

Определение. Расстоянием между векторами X и Y называется число $|X - Y|$.

Пример 1.2. *Найти расстояние между полиномами*

$$p(x) = x^{100} - {}^1\!/_2\, x^{85} - {}^1\!/_2\, x^{64} + 5\,x^{34} - 5\,x^{32} + 5\,x^2 + 1 \quad \text{и} \quad q(x) = 5\,x^2 + 1$$

если скалярное произведение задается формулой **а)** (1.3)*;* **б)** (1.4) *при* $a = -1$, $b = 1$.

Решение. Для случая **а)** нам достаточно просто вычислить сумму квадратов коэффициентов разности $p(x) - q(x)$: расстояние равно $\sqrt{103/2}$.

Для случая **б)** нам придется иметь дело с интегралом

$$\int_{-1}^{1} (p(t) - q(t))^2\, d\,t = \int_{-1}^{1} \left(t^{100} - {}^1\!/_2\, t^{85} - {}^1\!/_2\, t^{64} + 5\,t^{34} - 5\,t^{32}\right)^2 d\,t \ ,$$

который, несмотря на свою громоздкий вид, может быть вычислен элементарными приемами математического анализа. В этом случае расстояние будет равно $\sqrt{95965413818\,/16503052280715}$.

Ответ. **а)** ≈ 7.176 ; **б)** ≈ 0.076.

Теперь прокомментируем последний пример. Мы уже использовали в главах 3 и 8 один результат, который основывался на оценке близости двух полиномов: теорему о непрерывной зависимости корней полинома от его коэффициентов. Смысл этого результата заключался в следующем: если коэффициенты полиномов

$$f(x) = x^n + a_1 x^{n-1} + \dots + a_n \quad \text{и} \quad \widetilde{f}(x) = x^n + \widetilde{a}_1 x^{n-1} + \dots + \widetilde{a}_n$$

из $\mathbb{C}[x]$ близки, то и корни этих полиномов (при соответствующей нумерации) будут близки на комплексной плоскости. Мера близости полиномов вычислялась по формуле

$$\sqrt[n]{\sum_{k=1}^{n} |a_k - \widetilde{a}_k| \gamma^{n-k}} \ , \quad \text{где} \quad \gamma \stackrel{\text{def}}{=} \max_{j \in \{1,\dots,n\}} \left\{ \sqrt[j]{|a_j|}\ , \sqrt[j]{|\widetilde{a}_j|} \right\} \ ,$$

которая, хоть и не совпадает с формулой

$$\sqrt{\sum_{k=1}^{n} (a_k - \widetilde{a}_k)^2} \ ,$$

задающей расстояние в пространстве $\mathbb{P}_n$ со скалярным произведением (1.3), но идейно ей близка. Вычисленное в примере 1.2 **а)** расстояние между полиномами $p(x)$ и $q(x)$ оказывается достаточно большим в том смысле, что если для полинома $p(x)$ искать полином, имеющий почти такое же расположение корней на $\mathbb{C}$, то $q(x)$ оказывается неподходящим выбором. (Да и графики этих функций сильно различаются, если рассматривать их на достаточно большом интервале.)

Другое дело, если ставится задача приближения полинома $p(x)$ только на отрезке $[-1, 1]$ — тогда полином $q(x)$ вполне может оказаться полезным.

Выясним сначала природу интеграла, возникшего при решении. Пусть сначала $p(x)$ и $q(x)$ — произвольные, но (для простоты рассуждений) неотрицательные на отрезке $[a,b]$ полиномы. Геометрический смысл интеграла $\int_a^b p(t)d\,t$ — площадь криволинейной трапеции на плоскости (x,y), ограниченной прямыми $x=a$, $x=b$, $y=0$ и графиком $y=p(x)$. Следовательно, геометрический смысл интеграла

$$\int_a^b |p(t)-q(t)|\,d\,t \tag{1.10}$$

— площадь фигуры, ограниченной прямыми $x=a$, $x=b$ и графиками $y=p(x), y=q(x)$ (заштрихована на рис. 1).

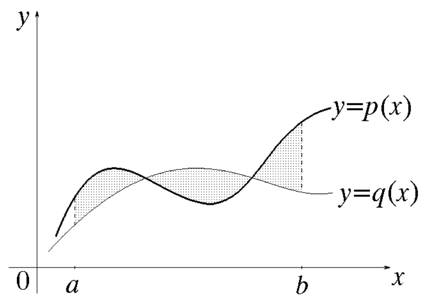


Рис. 1.

Чем меньше эта площадь, тем "теснее" друг к другу на отрезке $[a,b]$ расположены графики $y=p(x)$ и $y=q(x)$. Величина

$$\sqrt{\int_a^b (p(t)-q(t))^2\,d\,t}\ ,$$

вообще говоря, не совпадает с (1.10), но смысл ее тот же: она позволяет оценивать близость графиков на всем отрезке $[a,b]$. Ответ в примере 1.2 **б)** позволяет сделать заключение, что на отрезке $[-1,1]$ полином $p(x)$ неплохо приближается своими младшими членами, т.е. на указанном отрезке график $y=p(x)$ не должен слишком сильно отличаться от параболы $y=5\,x^2+1$.

Подводя итог приведенным рассуждениям, можно только повторить слова замечания со с. 34: метод, выбираемый для оценки близости между объектами, может зависеть от поставленной задачи. Микроскоп не пригоден для наблюдения за большими объектами, а телескоп — за малыми.

Историческая справка

Пифагор Самосский ($\Pi\upsilon\vartheta\alpha\gamma\acute{o}\rho\alpha\varsigma$, ок. 570 – ок. 497 до н.э.)
Его имя означает по гречески "*убеждающий речью*". Родился на острове Самос (по другим источникам – из рода тирренцев, т.е. этрусков).

“Юный, но жаждущий знания, он покинул отечество для посвящения во все таинства, как эллинские, так и варварские: он появился в Египте,... выучил египетский язык, он явился и к халдеям и к магам[6]. Потом на Крите он ... спустился в пещеру Иды, как и в Египте в тамошние святилища, и узнал о богах самое сокровенное. А вернувшись на Самос и застав отечество под тиранией Поликрата, он удалился в италийский Кротон; там он написал законы для италийцев и достиг у них великого почета вместе со своими учениками, числом до трехсот, которые вели государственные дела так отменно, что поистине это была аристократия, что значит «владычество лучших».

О себе он говорил,... что некогда был Эфалидом и почитался сыном Гермеса; и что Гермес предложил ему на выбор любой дар, кроме бессмертия, а он попросил оставить ему и живому и мертвому память о том, что с ним было. Поэтому и при жизни он помнил обо всем, и в смерти сохранил ту же память ...

Пять лет [его ученики] проводили в молчании, только внимая речам Пифагора, но не видя его, пока не проходили испытания; и лишь затем они допускались в его жилище и к его лицезрению ...

Это он довел до совершенства геометрию ... Больше всего внимания он уделял числовой стороне этой науки... А когда он нашел, что в прямоугольном треугольнике квадрат гипотенузы равен квадрату катетов, то принес богам гекатомбу[7]...

Говорят, он первый заявил, что душа совершает круг неизбежности, чередою облекаясь то в одну, то в другую жизнь...

Александр в «Преемствах философов» говорит, что в пифагорейских записках содержится также вот что. Начало всего — единица; единице, как причине подлежит как вещество неопределенная двоица; из единицы и неопределенной двоицы исходят числа; из чисел — точки; из точек — линии; из них — плоские фигуры; из плоских — объемные фигуры; из них — чувственновоспринимаемые тела, в которых четыре основы — огонь, вода, земля и воздух; перемещаясь и превращаясь целиком, они порождают мир — одушевленный, разумный, шаровидный, в середине которого — земля; и земля также шаровидна и населена со всех сторон. Существуют даже антиподы, и наш низ — для них верх.”

(**Диоген Лаэрций** “О жизни, учениях и изречениях знаменитых философов”.)

“Философия, которую он исповедовал, целью своей имела вызволить и высвободить врожденный наш разум от его оков и цепей; а без ума человек не познает ничего здравого, ничего истинного и даже неспособен ничего уловить какими бы то ни было чувствами, — только ум сам по себе все видит и все слышит, прочее же и слепо и глухо.

А для тех, кто уже совершил очищение, есть некоторые полезные приемы. Приемы он придумал такие: медленно и постепенно, всегда одним и тем же образом, начиная от все более мелкого, переводить себя к созерцанию вечного и сродного ему бестелесного... Вот почему для предварительной подготовки душевных очей к переходу от всего телесного ... к истинно сущему он обращался к математическим и иным предметам рассмотрения, лежащим на грани телесного и бестелесного (эти предметы трехмерны, как и все телесное, но плотности не имеют, как все бестелесное), — это как бы искусственно приводило душу к потребности в [настоящей ее] пище. Подводя с помощью такого приема к созерцанию истинно сущего, он дарил людям блаженство, — для этого и нужны были ему математические упражнения.

... Пифагор со всеми друзьями немалое время жил в Италии, пользуясь таким почтением, что целые государства вверяли себя его ученикам. Но в конце концов против них скопилась зависть и сложился заговор, а случилось это вот каким образом. Был в Кротоне человек по имени Килон, первый между гражданами и богатством, и знатностью, и славою своих предков, но сам обладавший нравом тяжелым и властным, а силою друзей своих и обилием богатств пользовавшийся не для добрых дел; вот и он-то, полагая себя достойным всего самого

[6] *Магами* греки называли жрецов персидской зороастрийской религии, *халдеями* — жрецов более древней вавилонской религии, *гимнософистами* (голыми мудрецами) — индийских брахманов, *друидами* — жрецов национальной кельтской религии.

[7] Жертва ста быков; впрочем эта легенда противоречит традиционному пифагорейскому вегетарианству и, по другим источникам, в жертву был принесен бык, сделанный из меда и муки.

лучшего, почел за нужнейшее причастится к Пифагоровой философии. Он пришел к Пифагору, похваляясь и притязая стать его другом. Но Пифагор сразу прочитал весь нрав этого человека по лицу его ... и поняв, что это за человек, велел ему идти прочь и не в свои дела не мешаться. Килон почел себя этим обиженным и оскорбился; а нрава он был дурного и в гневе безудержен. И вот, созвав своих друзей, он стал обличать пред ними Пифагора и готовить с ними заговор против философа и его учеников. И когда после этого друзья Пифагора сошлись на собрании в доме атлета Милона (а самого Пифагора, по этому рассказу, между нии не было ...), то дом был подожжен со всех сторон, и все собравшиеся погибли; только двое спаслись от пожара, Архипп и Лисид, и Лисид бежал в Элладу и стал там другом и учителем Эпаминонда. А по рассказу Дикеарха ... при этом покушении был и сам Пифагор; ... сорок друзей его были застигнуты в доме на собрании, остальные перебиты порознь в городе, а Пифагор, лишась друзей, пустился искать спасения сперва в гавань Кавлония, а затем в Локры. Локрийцы, узнав об этом, выслали к рубежу своей земли избранных своих старейшин с такими словами к Пифагору: «Мы знаем, Пифагор, что ты мудрец и человек предивный, но законы в нашем городе безупречные, и мы хотим при них жить, как жили, а ты возьми у нас, коли что надобно, и ступай отсюда прочь, куда знаешь». Повернув таким образом прочь от Локров, Пифагор поплыл в Тарент, а когда и в Таренте случилось такое же, как и в Кротоне, то перебрался в Метапонт[8]. Ибо повсюду вспыхивали великие мятежи, которые и посейчас у историков тех мест называются пифагорейскими...

Здесь, в Метапонте, Пифагор, говорят и погиб: он бежал от мятежа в святилище Муз и оставался там без пищи целых сорок дней. А другие говорят, что когда подожгли дом, где они собирались, то друзья его, бросившись в огонь, проложили в нем дорогу учителю, чтобы он по их телам вышел из огня как по мосту; но, спасшись из пожара и оставшись без товарищей, Пифагор так затосковал, что сам лишил себя жизни.

Бедствие это, обрушившись на людей, задело вместе с этим и науку их, потому что до этих пор они ее хранили неизреченно в сердцах своих, а вслух высказывали лишь темными намеками. И от Пифагора сочинений не осталось, а спасшиеся Архипп, Лисид и остальные, кто был тогда на чужбине, сберегли лишь немногие искры его философии, смутные и рассеянные. В одиночестве, угнетенные случившимся, скитались они где попало, чуждаясь людского общества. И тогда, чтобы не погибла в людях вовсе память о философии и чтобы за это не прогневались на них боги, стали они составлять сжатые записки, собирать сочинения старших и все, что сами помнили, и каждый оставлял это там, где случалось ему умереть, а сыновьям, дочерям и жене завещал никому это из дома не выносить; и это завещание они долго соблюдали, передавая от потомка к потомку..."

(**Порфирий** "Жизнь Пифагора".)

2 Ортогональный базис

На основании аксиом скалярного произведения, его вычисление для произвольных векторов X и Y может быть сведено к вычислению скалярных произведений векторов произвольного базиса. В самом деле, если система $\{X_1, \dots, X_n\}$ составляет базис $\mathbb{E}$, то, разложив оба вектора по этому базису

$$X = x_1X_1 + \cdots + x_nX_n \quad \text{и} \quad Y = y_1X_1 + \cdots + y_nX_n \ ,$$

получаем: $(X, Y) = (x_1X_1 + \cdots + x_nX_n, y_1X_1 + \cdots + y_nX_n) =$

$$\begin{aligned}
&= x_1y_1(X_1, X_1) + x_1y_2(X_1, X_2) + \cdots + x_1y_n(X_1, X_n) + \\
&+ x_2y_1(X_2, X_1) + x_2y_2(X_2, X_2) + \cdots + x_2y_n(X_2, X_n) + \\
&+ \cdots + \\
&+ x_ny_1(X_n, X_1) + x_ny_2(X_n, X_2) + \cdots + x_ny_n(X_n, X_n) =
\end{aligned}$$

[8]Город существует до наших дней.

$$
= (x_1, x_2, \dots, x_n) \begin{pmatrix} (X_1, X_1) & (X_1, X_2) & \dots & (X_1, X_n) \\ (X_2, X_1) & (X_2, X_2) & \dots & (X_2, X_n) \\ \dots & & & \dots \\ (X_n, X_1) & (X_n, X_2) & \dots & (X_n, X_n) \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} . \tag{2.1}
$$

Итак, при изменении векторов X и Y в последней формуле изменятся лишь строка и столбец координат, а промежуточная матрица останется неизменной. Задание этой матрицы, следовательно, полностью определит скалярное произведение в $\mathbb{R}^n$. Фактически задание скалярного произведения по формуле (1.2) и можно рассматривать как частный случай (2.1) при подходящем подборе базисных векторов. Согласно рассуждениям из §1, матрица из (2.1) должна обладать некоторыми дополнительными свойствами. Мы их обсудим позднее — в §4. А пока обратим внимание на тот факт, что формула (2.1) очень упростится, если матрица, в ней участвующая, станет диагональной.

Определение. Пусть $\dim \mathbb{E} = n$ и векторы $\{X_1, \dots, X_n\}$ составляют базис $\mathbb{E}$. Этот базис называется **ортогональным** если векторы попарно ортогональны: $X_i \perp X_j$; базис называется **нормированным** если каждый его вектор имеет единичную длину: $|X_i| = 1$; базис называется **ортонормированным** если он ортогонален и нормирован, т.е.

$$
(X_i, X_j) = \delta_{ij}, \quad i, j \in \{1, \dots, n\} \ .
$$

Ортогональный базис будем обозначать $\mathfrak{E}_1, \dots, \mathfrak{E}_n$.

В пространстве $\mathbb{R}^n$ стандартным ортогональным базисом является базис, состоящий из векторов

$$
\mathfrak{e}_i \stackrel{\text{def}}{=} \big[\underbrace{0, \dots, 0, 1}_{i}, 0, \dots, 0\big]^\top, \quad i \in \{1, \dots, n\} \ .
$$

Существование же ортогонального базиса в *произвольном* евклидовом пространстве еще требует доказательства. Предварительно установим следующий результат.

Теорема 2.1. *Если ненулевые векторы $X_1, \dots, X_n$ попарно ортогональны, то они* **л.н.з.**

Доказательство . В самом деле, если

$$
\lambda_1 X_1 + \dots + \lambda_n X_n = \mathbb{O} \ ,
$$

то, домножив это равенство скалярно на X_1, получим

$$
\lambda_1 (X_1, X_1) + \dots + \lambda_n (X_1, X_n) = 0 \ .
$$

Поскольку $(X_1, X_j) = 0$ для $j \in \{2, \dots, n\}$, то $\lambda_1 (X_1, X_1) = 0$, откуда $\lambda_1 = 0$. Аналогично показывается, что и все остальные λ_j равны 0. □

Теорема 2.2. *В любом евклидовом пространстве существует ортогональный базис.*

Доказательство теоремы основано на процедуре, имеющей самостоятельное значение. Пусть имеется произвольная система $\{X_1,\dots,X_k\}$ линейно независимых векторов. Требуется построить систему ортогональных векторов $\{\mathfrak{E}_1,\dots,\mathfrak{E}_k\}$ так, чтобы

$$\mathcal{L}\left(X_1,\dots,X_m\right)=\mathcal{L}\left(\mathfrak{E}_1,\dots,\mathfrak{E}_m\right) \quad \text{для } m\in\{1,\dots,k\} \ . \tag{2.2}$$

Для $m=1$ условие (2.2) будет очевидно выполнено если взять $\mathfrak{E}_1=X_1$: поскольку вектор X_1 базисный, то $\mathfrak{E}_1\neq\mathbb{O}$. Далее, будем искать $\mathfrak{E}_2$ в виде

$$\mathfrak{E}_2=X_2+\alpha_{21}\mathfrak{E}_1$$

при пока неопределенном коэффициенте α_{21}. Очевидно, что при таком выборе $\mathfrak{E}_2$ условие (2.2) будет выполняться. Подберем α_{21} так, чтобы выполнялось $\mathfrak{E}_2\perp\mathfrak{E}_1$.

$$0=(\mathfrak{E}_1,\mathfrak{E}_2)=(\mathfrak{E}_1,X_2)+\alpha_{21}(\mathfrak{E}_1,\mathfrak{E}_1) \ \Rightarrow \ \alpha_{21}=-(\mathfrak{E}_1,X_2)\big/(\mathfrak{E}_1,\mathfrak{E}_1) \ .$$

Таким образом, коэффициент α_{21}, а вместе с ним и вектор $\mathfrak{E}_2$ определяются единственным образом. При этом $\mathfrak{E}_2\neq\mathbb{O}$, ибо, в противном случае, векторы X_2 и $\mathfrak{E}_1=X_1$ были бы **л.з.**, что противоречит предположению о **л.н.з** системы $\{X_1,\dots,X_k\}$. Продолжаем процесс далее: вектор $\mathfrak{E}_3$ ищем в виде

$$\mathfrak{E}_3=X_3+\alpha_{31}\mathfrak{E}_1+\alpha_{32}\mathfrak{E}_2$$

при пока неопределенных коэффициентах α_{31} и α_{32}. Условие (2.2) выполняется поскольку

$$\alpha_{31}\mathfrak{E}_1+\alpha_{32}\mathfrak{E}_2\in\mathcal{L}(X_1,X_2)\subset\mathcal{L}(X_1,X_2,X_3) \ .$$

Подберем скаляры α_{31} и α_{32} так, чтобы выполнялось $\mathfrak{E}_3\perp\mathfrak{E}_1$ и $\mathfrak{E}_3\perp\mathfrak{E}_2$. Два этих условия задают систему линейных уравнений

$$\left\{\begin{array}{ll}(X_3,\mathfrak{E}_1)+\alpha_{31}(\mathfrak{E}_1,\mathfrak{E}_1)+\alpha_{32}(\mathfrak{E}_2,\mathfrak{E}_1) & =0,\\ (X_3,\mathfrak{E}_2)+\alpha_{31}(\mathfrak{E}_1,\mathfrak{E}_2)+\alpha_{32}(\mathfrak{E}_2,\mathfrak{E}_2) & =0,\end{array}\right. \iff \begin{array}{l}\alpha_{31}=-(X_3,\mathfrak{E}_1)\big/|\mathfrak{E}_1|^2\\ \alpha_{32}=-(X_3,\mathfrak{E}_2)\big/|\mathfrak{E}_2|^2\end{array}$$

Процесс продолжается далее аналогично. Допустим, что векторы $\mathfrak{E}_1,\dots,\mathfrak{E}_{k-1}$ уже построены, они ненулевые, попарно ортогональные и

$$\mathcal{L}\left(X_1,\dots,X_{k-1}\right)=\mathcal{L}\left(\mathfrak{E}_1,\dots,\mathfrak{E}_{k-1}\right) \ .$$

Вектор $\mathfrak{E}_k$ ищем в виде:

$$\mathfrak{E}_k=X_k+\alpha_{k1}\mathfrak{E}_1+\alpha_{k2}\mathfrak{E}_2+\dots+\alpha_{k,k-1}\mathfrak{E}_{k-1}$$

при пока неопределенных коэффициентах $\alpha_{k1},\dots,\alpha_{k,k-1}$. Условие (2.2) выполнено и, кроме того, $\mathfrak{E}_k\neq\mathbb{O}$ (в противном случае $X_k\in\mathcal{L}\left(\mathfrak{E}_1,\dots,\mathfrak{E}_{k-1}\right)=\mathcal{L}\left(X_1,\dots,X_{k-1}\right)$, т.е. система $\{X_1,\dots,X_{k-1},X_k\}$ **л.з.**). Коэффициенты α_{k1}, $\dots,\alpha_{k,k-1}$ подбираются из условий $\mathfrak{E}_k\perp\mathfrak{E}_1,\dots,\mathfrak{E}_k\perp\mathfrak{E}_{k-1}$. Получающаяся система линейных уравнений имеет единственное решение

$$\alpha_{k1}=-(X_k,\mathfrak{E}_1)\big/|\mathfrak{E}_1|^2,\dots,\alpha_{k,k-1}=-(X_k,\mathfrak{E}_{k-1})\big/|\mathfrak{E}_{k-1}|^2 \ ,$$

это решение определяет единственный вектор $\mathfrak{E}_k$.

Приведенная только что процедура называется **процессом ортогонализации Грама – Шмидта**[9] системы векторов $\{X_1, \dots, X_k\}$.

Доказательство же самой теоремы следует из возможности ортогонализации произвольного базиса пространства $\mathbb{E}$. $\square$

Пример 2.1. *Ортогонализовать систему векторов*

$$X_1 = [1, 0, 0, 0, 1]\,,\ X_2 = [1, 1, 0, 1, 1]\,,\ X_3 = [1, 1, 1, 1, 1]\ \ .$$

Решение. Имеем:

$$\begin{array}{lcll}
\mathfrak{E}_1 & = & X_1, & \Rightarrow \mathfrak{E}_1 = [1,0,0,0,1]\,, \\
\mathfrak{E}_2 & = & X_2 + \alpha_{21}\mathfrak{E}_1, & \\
 & & \qquad \alpha_{21} = -(X_2, \mathfrak{E}_1)/|\mathfrak{E}_1|^2 = -1 & \Rightarrow \mathfrak{E}_2 = [0,1,0,1,0]\,, \\
\mathfrak{E}_3 & = & X_3 + \alpha_{31}\mathfrak{E}_1 + \alpha_{32}\mathfrak{E}_2, & \\
 & & \qquad \alpha_{31} = -(X_3, \mathfrak{E}_1)/|\mathfrak{E}_1|^2 = -1, & \\
 & & \qquad \alpha_{32} = -(X_3, \mathfrak{E}_2)/|\mathfrak{E}_2|^2 = -1 & \Rightarrow \mathfrak{E}_3 = [0,0,1,0,0]\ \ .
\end{array}$$

Ответ. $\mathfrak{E}_1 = [1, 0, 0, 0, 1]\,,\ \mathfrak{E}_2 = [0, 1, 0, 1, 0]\,,\ \mathfrak{E}_3 = [0, 0, 1, 0, 0]$.

Пример 2.2. *Пусть в пространстве* $\mathbb{P}_n = \{p(x) \in \mathbb{R}[x] \mid \deg p(x) \le n\}$ *скалярное произведение задается формулой*

$$(p(x), q(x)) \stackrel{\text{def}}{=} \int_{-1}^{1} p(t)q(t)d\,t\ \ .$$

Построить ортогональный базис этого пространства.

Решение. Требуемый базис строится ортогонализацией канонического базиса $1, x, x^2, \dots, x^n$. В результате получаем **полиномы Лежандра**:

$$L_0(x) = 1,\ L_1(x) = x,\ L_2(x) = (3\,x^2 - 1)/2,\ L_3(x) = (5\,x^3 - 3\,x)/2,$$

$$L_4(x) = (35\,x^4 - 30\,x^2 + 3)/8, \dots, L_n(x) = \frac{1}{2^n}\sum_{k=0}^{\lfloor n/2 \rfloor} \frac{(-1)^k(2n-2k)!}{k!(n-k)!(n-2k)!}x^{n-2k}\ \ .$$

Рекуррентное соотношение

$$kL_k(x) - (2k-1)\,xL_{k-1}(x) + (k-1)\,L_{k-2}(x) \equiv 0, \quad k \ge 2\ \ ;$$

позволяет найти полином $L_k(x)$ если уже вычислены $L_{k-1}(x)$ и $L_{k-2}(x)$.
$\triangle$

В ортонормированном базисе пространства $\mathbb{E}$ матрица из формулы (2.1) становится единичной и в таком базисе формула скалярного произведения принимает привычный из геометрии вид:

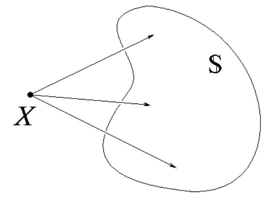
$$(X, Y) = \sum_{j=1}^{n} x_j y_j\ \ .$$

[9]Хотя, по-видимому, процесс использовался уже Лапласом и Коши.

3 Расстояние от точки до многообразия

Задача. Найти расстояние от заданного вектора X до заданного множества $\mathbb{S} \subset \mathbb{E}$.

Такая постановка требует немедленного уточнения: что такое расстояние вектора до множества? Обратясь за помощью к геометрии, мы можем ввести это понятие, основываясь на понятии расстояния между точками: например, расстояние от точки $X \in \mathbb{R}^2$ до множества $\mathbb{S} \subset \mathbb{R}^2$ определить как минимальное из возможных расстояний между точками X и Y, где $Y \in \mathbb{S}$. Следующий пример показывает, что наше определение неполно.



Пример 3.1. *Множество $\mathbb{S} = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\}$ не содержит своих граничных точек; требуемый минимум расстояний от точки $X = (2, 0)$ до точек круга не достигается: при любом выборе точки $Y_1 \in \mathbb{S}$ найдется более близкая к точке X точка $Y_2 \in \mathbb{S}$.*

Определение. Расстоянием от вектора $X \in \mathbb{E}$ до множества $\mathbb{S} \subset \mathbb{E}$ назовем число

$$\inf_{Y \in \mathbb{S}} |X - Y| \ .$$

Замечание. Повторим содержание замечания страницы 20. По традиции, идущей из геометрии, в алгебре также говорят о "расстоянии от конца вектора X ..." или о "расстоянии от точки X ...". Кроме удобства ассоциативного восприятия, этим словам не придается никакого формального смысла: понятие "конец вектора" отсутствует в аксиоматике евклидова пространства.

В случае произвольного множества $\mathbb{S}$ поставленная задача весьма сложна. В настоящей главе мы будем решать ее только для частного случая расстояния "от точки до плоскости", т.е. от вектора $X \in \mathbb{E}$ до линейного многообразия $\mathbb{M} \subset \mathbb{E}$.

3.1 Ортогональное дополнение

Определение. Говорят, что вектор X **ортогонален множеству** $\mathbb{S} \subset \mathbb{E}$, если X ортогонален любому вектору Y из $\mathbb{S}$.

Теорема 3.1. *Для того, чтобы вектор X был ортогонален линейному подпространству $\mathbb{E}_1$* **Н.** *и* **Д.**, *чтобы X был ортогонален произвольному базису этого подпространства.*

Теорема 3.2. *Множество векторов X ортогональных подпространству $\mathbb{E}_1 \subset \mathbb{E}$ образует подпространство. Оно называется* **ортогональным дополнением** *подпространства $\mathbb{E}_1$ в $\mathbb{E}$ и обозначается $\mathbb{E}_1^{\perp}$. Справедливо равенство*

$$\mathbb{E} = \mathbb{E}_1 \oplus \mathbb{E}_1^{\perp} \ . \qquad (3.1)$$

Доказательство . Если $X \perp \mathbb{E}_1$, то $\lambda X \perp \mathbb{E}_1$ для $\forall \lambda \in \mathbb{R}$. Если, вдобавок, $Y \perp \mathbb{E}_1$, то и $(X+Y) \perp \mathbb{E}_1$. Следовательно, $\mathbb{E}_1^{\perp}$ является подпространством пространства $\mathbb{E}$.

Обозначим $\{X_1, \dots, X_k\}$ — произвольный базис $\mathbb{E}_1$. Согласно теореме 2.2, этот базис можно ортогонализовать:

$$\{X_1, \dots, X_k\} \ \to \ \{\mathfrak{E}_1, \dots, \mathfrak{E}_k\} \quad \text{где } (\mathfrak{E}_j, \mathfrak{E}_\ell) = 0 \text{ при } j \neq \ell$$

и

$$\mathbb{E}_1 = \mathcal{L}(X_1, \dots, X_k) = \mathcal{L}(\mathfrak{E}_1, \dots, \mathfrak{E}_k) \ .$$

Дополним ортогональный базис подпространства до ортогонального базиса всего пространства:

$$\mathbb{E} = \mathcal{L}(\mathfrak{E}_1, \dots, \mathfrak{E}_k, \mathfrak{E}_{k+1}, \dots, \mathfrak{E}_n) \ .$$

Иными словами[10], векторы $\{\mathfrak{E}_{k+1}, \dots, \mathfrak{E}_n\}$ образуют базис $\mathbb{E}$ относительно $\mathbb{E}_1$. Если мы докажем, что

$$\mathbb{E}_1^{\perp} = \mathcal{L}(\mathfrak{E}_{k+1}, \dots, \mathfrak{E}_n) \ , \tag{3.2}$$

то на основании следствия 1 к теореме 4.1 главы 1 будет справедливо равенство (3.1).

Для доказательства (3.2) заметим, что если, с одной стороны, $X \in \mathcal{L}(\mathfrak{E}_{k+1}, \dots, \mathfrak{E}_n)$, то

$$X = x_{k+1}\mathfrak{E}_{k+1} + \dots + x_n \mathfrak{E}_n \ \Rightarrow \ X \perp \mathfrak{E}_1, \dots, X \perp \mathfrak{E}_k \ \Rightarrow \ X \perp \mathbb{E}_1 \ .$$

Таким образом, $\mathcal{L}(\mathfrak{E}_{k+1}, \dots, \mathfrak{E}_n) \subset \mathbb{E}_1^{\perp}$. С другой стороны, если вектор

$$Y = y_1 \mathfrak{E}_1 + \dots + y_k \mathfrak{E}_k + y_{k+1}\mathfrak{E}_{k+1} + \dots + y_n \mathfrak{E}_n$$

принадлежит $\mathbb{E}_1^{\perp}$, то $Y \perp \mathfrak{E}_1, \dots, Y \perp \mathfrak{E}_k$, т.е.

$$0 = (Y, \mathfrak{E}_1) = y_1, \dots, 0 = (Y, \mathfrak{E}_k) = y_k \ \Rightarrow$$

$$\Rightarrow \ Y = y_{k+1}\mathfrak{E}_{k+1} + \dots + y_n \mathfrak{E}_n \ \Rightarrow \ Y \in \mathcal{L}(\mathfrak{E}_{k+1}, \dots, \mathfrak{E}_n) \ .$$

Следовательно, $\mathcal{L}(\mathfrak{E}_{k+1}, \dots, \mathfrak{E}_n) \supset \mathbb{E}_1^{\perp}$. Взаимное включение множеств $\mathcal{L}(\mathfrak{E}_{k+1}, \dots, \mathfrak{E}_n)$ и $\mathbb{E}_1^{\perp}$ друг в друга доказывает равенство (3.2). $\square$

Следствие 1. $\dim \mathbb{E}_1^{\perp} = \dim \mathbb{E} - \dim \mathbb{E}_1$.

[10] Глава 1, §5.

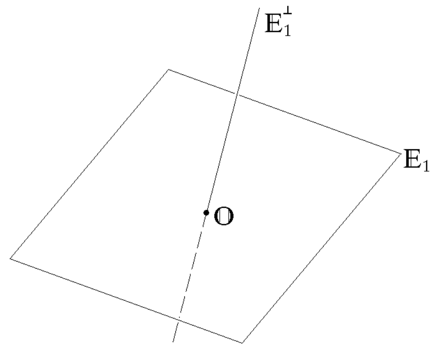


Рис. 3.

Пример 3.2. *Пусть в* $\mathbb{R}^4$ *скалярное произведение задано формулой* (1.1). *Построить* $\mathbb{E}_1^{\perp}$ *для*

$$\textbf{а)}\ \mathbb{E}_1 = \mathcal{L}\left(\begin{bmatrix} 1\\0\\1\\0\end{bmatrix}, \begin{bmatrix} 0\\1\\1\\-2\end{bmatrix}, \begin{bmatrix} 2\\-1\\1\\2\end{bmatrix}\right) \ ;$$

$$\textbf{б)}\ \mathbb{E}_1 = \left\{ X\in\mathbb{R}^4 \left| \begin{array}{rrrrl} 3\,x_1 & +2\,x_2 & -x_3 & -x_4 & =0,\\ x_1 & +x_2 & -3\,x_3 & -2\,x_4 & =0,\\ 2\,x_1 & +x_2 & +2\,x_3 & +x_4 & =0 \end{array}\right.\right\} \ .$$

Решение возможно было бы вести по алгоритму доказательства теоремы 3.2. Но мы упростим себе жизнь, поставив целью нахождение *произвольного* — т.е. не обязательно ортогонального — базиса $\mathbb{E}_1^{\perp}$.

Для случая **а)** вектор $X=[x_1,x_2,x_3,x_4]^{\top}$, принадлежащий $\mathbb{E}_1^{\perp}$, должен быть ортогонален всем векторам из $\mathbb{E}_1$, и, в частности, тем, на которые натянуто это подпространство. Из полученных равенств составляем систему

$$\left\{\begin{array}{rrrrl} x_1 & & +x_3 & & =0,\\ & x_2 & +x_3 & -2\,x_4 & =0,\\ 2\,x_1 & -x_2 & +x_3 & +2\,x_4 & =0, \end{array}\right.$$

не заботясь о том, что некоторые уравнения могут оказаться зависимыми от других. Множество решений этой системы как раз и будет совпадать с $\mathbb{E}_1^{\perp}$. Нам остается лишь найти базис, т.е. **ф.с.р.**:

$$\Longleftrightarrow \left\{\begin{array}{rrrl} x_1 & & +x_3 & & =0,\\ & x_2 & +x_3 & -2\,x_4 & =0 \end{array}\right. \Rightarrow \begin{array}{cc|cc} \multicolumn{2}{c}{\overbrace{}^{\text{зависимые}}} & \multicolumn{2}{c}{\overbrace{}^{\text{основные}}}\\ x_1 & x_2 & x_3 & x_4\\ \hline -1 & -1 & 1 & 0\\ 0 & 2 & 0 & 1 \end{array}$$

$$\mathbb{E}_1^{\perp} = \mathcal{L}\left([-1,-1,1,0]^{\top},[0,2,0,1]^{\top}\right) \ .$$

Решение примера **б)** можно было бы свести к предыдущему случаю. Однако базисные векторы $\mathbb{E}_1^{\perp}$ легко определяются из следующих соображений. Заметим, что каждое из уравнений системы, задающей $\mathbb{E}_1$, можно "перевести на язык" скалярного произведения

$$3\,x_1 + 2\,x_2 - x_3 - x_4 = 0 \iff \left([3, 2, -1, -1]^{\top}, [x_1, x_2, x_3, x_4]^{\top} \right) = 0 \ .$$

Формально это соотношение означает, что произвольный вектор X из $\mathbb{E}_1$ должен быть ортогонален вектору $[3, 2, -1, -1]^{\top}$. Но последний факт как раз и означает, что $[3, 2, -1, -1]^{\top} \in \mathbb{E}_1^{\perp}$. Аналогичные рассуждения справедливы и для других уравнений системы. Имеем:

$$\mathbb{E}_1^{\perp} = \mathcal{L} \left(\left[\begin{array}{r} 3 \\ 2 \\ -1 \\ -1 \end{array} \right], \left[\begin{array}{r} 1 \\ 1 \\ -3 \\ -2 \end{array} \right], \left[\begin{array}{r} 2 \\ 1 \\ 2 \\ 1 \end{array} \right] \right) = \mathcal{L} \left(\left[\begin{array}{r} 3 \\ 2 \\ -1 \\ -1 \end{array} \right], \left[\begin{array}{r} 1 \\ 1 \\ -3 \\ -2 \end{array} \right] \right)$$

$\triangle$

Упражнение 3.1. *Доказать следующие свойства операции* $\perp$*:*

а) $\left(\mathbb{E}_1^{\perp}\right)^{\perp} = \mathbb{E}_1$; **б)** $(\mathbb{E}_1 + \mathbb{E}_2)^{\perp} = \mathbb{E}_1^{\perp} \cap \mathbb{E}_2^{\perp}$; **в)** $(\mathbb{E}_1 \cap \mathbb{E}_2)^{\perp} = \mathbb{E}_1^{\perp} + \mathbb{E}_2^{\perp}$.

Упражнение 3.2. *Доказать, что в пространстве квадратных матриц со скалярным произведением, заданным формулой* (1.5)*, подпространство кососимметричных матриц является ортогональным дополнением подпространства симметричных матриц.*

3.2 Вычисление расстояния до многообразия

Теорема 3.2 позволяет сформулировать результат, на котором и будет основано решение задачи поиска расстояния.

Следствие 2. *Для любого вектора* $X \in \mathbb{E}$ *существует единственное представление его в виде*

$$X = X^{\parallel} + X^{\perp} \quad \textit{где } X^{\parallel} \in \mathbb{E}_1, X^{\perp} \in \mathbb{E}_1^{\perp} \ . \tag{3.3}$$

Определение. В разложении (3.3) вектор $X^{\parallel}$ называется **ортогональной проекцией вектора** X **на** $\mathbb{E}_1$, а вектор $X^{\perp}$ — **ортогональной составляющей вектора** X **относительно** $\mathbb{E}_1$ или же **перпендикуляром, опущенным из точки** X на **подпространство** $\mathbb{E}_1$.

Теорема 3.3. *Число* $\left|X^{\perp}\right|$ *будет расстоянием от точки* X *до* $\mathbb{E}_1$*, т.е.*

$$\left|X^{\perp}\right| = \min_{Y \in \mathbb{E}_1} |X - Y| \ .$$

Доказательство .

$$X^{\perp} = \left(X - X^{\parallel}\right) \perp \mathbb{E}_1 \ \Rightarrow \ X^{\perp} \perp \left(-Y + X^{\parallel}\right) \quad \text{для } \forall Y \ \in \mathbb{E}_1 \ .$$

По теореме 1.3 (Пифагора):

$$\left|X^{\perp}\right|^2 + \left|X^{\parallel} - Y\right|^2 = \left|X^{\perp} + X^{\parallel} - Y\right|^2 = |X - Y|^2 \ \Rightarrow$$

$$\Rightarrow \ \left|X^{\perp}\right|^2 \le |X - Y|^2 \ \Rightarrow \ \left|X^{\perp}\right| \le \min_{Y \in \mathbb{E}_1} |X - Y| \ .$$

С другой стороны, указанный минимум достигается при $Y = X^{\parallel}$:

$$\left|X^{\perp}\right| = \left|X - X^{\parallel}\right| \ .$$

□

Итак, задача, поставленная в начале параграфа, решается вычислением $\left|X^{\perp}\right|$. Для нахождения последнего числа сначала найдем базис $\{X_1, \dots, X_k\}$ подпространства $\mathbb{E}_1$. Далее, ищем $X^{\parallel}$, принадлежащий $\mathbb{E}_1$, в виде линейной комбинации базисных векторов:

$$X^{\parallel} = \alpha_1 X_1 + \dots + \alpha_k X_k \ .$$

Для нахождения скаляров $\alpha_1, \dots, \alpha_k$ используем тот факт, что вектор $X^{\perp} = X - X^{\parallel}$ должен быть ортогонален $\mathbb{E}_1$, а значит, ввиду теоремы 3.1, ортогонален каждому X_j:

$$(X - X^{\parallel}, X_j) = 0 \iff (X^{\parallel}, X_j) = (X, X_j) \ .$$

Получаем систему линейных уравнений:

$$\left\{ \begin{array}{lllll} \alpha_1(X_1, X_1) & +\alpha_2(X_1, X_2) & +\dots & +\alpha_k(X_1, X_k) & = (X, X_1), \\ \alpha_1(X_2, X_1) & +\alpha_2(X_2, X_2) & +\dots & +\alpha_k(X_2, X_k) & = (X, X_2), \\ \quad\dots & & & & \quad\quad\dots \\ \alpha_1(X_k, X_1) & +\alpha_2(X_k, X_2) & +\dots & +\alpha_k(X_k, X_k) & = (X, X_k). \end{array} \right. \tag{3.4}$$

Для однозначной разрешимости относительно $\alpha_1, \dots, \alpha_k$ **Н.** и **Д.**, чтобы матрица этой системы была невырожденной. Матрица такого вида уже встречалась нам в формуле (2.1).

Определение. **Матрицей Грама**[11] системы векторов $\{Y_1, \dots, Y_k\}$ называется матрица

$$G(Y_1, \dots, Y_k) \stackrel{\text{def}}{=} \left(\begin{array}{cccc} (Y_1, Y_1) & (Y_1, Y_2) & \dots & (Y_1, Y_k) \\ (Y_2, Y_1) & (Y_2, Y_2) & \dots & (Y_2, Y_k) \\ \dots & & & \dots \\ (Y_k, Y_1) & (Y_k, Y_2) & \dots & (Y_k, Y_k) \end{array} \right)$$

[11]**Грам Йорген Педерсен** (Gram Jørgen Pedersen, 1850 – 1916) — датский математик.

Ее определитель

$$\mathfrak{G}(Y_1,\dots,Y_k) \stackrel{\text{def}}{=} \det G(Y_1,\dots,Y_k) = \det\left[(Y_i,Y_j)\right]_{i,j=1}^k$$

называется **определителем Грама** или **грамианом** системы векторов $\{Y_1,\dots,Y_k\}$.

Упражнение 3.3. *Пусть $X_1,\dots,X_n$ и $\mathfrak{X}_1,\dots,\mathfrak{X}_n$ — два базиса пространства $\mathbb{E}$, а C — матрица перехода от одного базиса к другому. Доказать, что*

$$G(\mathfrak{X}_1,\dots,\mathfrak{X}_n) = C^{\top} G(X_1,\dots,X_n) C \ .$$

Матрица Грама, очевидно, симметрична и обращается в единичную если векторы $Y_1,\dots,Y_k$ входят в состав ортонормированного базиса пространства $\mathbb{E}$. Следовательно, по крайней мере в этом частном случае, система (3.4) будет иметь единственное решение. В следующем параграфе будет доказан и более общий факт:

$$\mathfrak{G}(Y_1,\dots,Y_k)=0 \iff \text{система } \{Y_1,\dots,Y_k\} \text{ линейно зависима} \ .$$

Этот факт позволяет нам заключить, что, поскольку векторы $\{X_1,\dots,X_k\}$ — базисные, то система уравнений (3.4) имеет единственное решение относительно $\alpha_1,\dots,\alpha_k$:

$$\alpha_1=\alpha_1^*,\dots,\alpha_k=\alpha_k^* \ .$$

Теперь может быть найдена проекция вектора X на $\mathbb{E}_1$:

$$X^{\parallel} = \alpha_1^* X_1 + \dots + \alpha_k^* X_k \ ,$$

а затем и составляющая: $X^{\perp} = X - X^{\parallel}$.

Пример 3.3. *Найти расстояние от точки $X=[1,1,2,2,2]$ до подпространства*

$$\mathbb{E}_1 = \left\{ X\in\mathbb{R}^5 \,\middle|\, \begin{array}{rrrrrl} x_1 & -x_2 & -x_3 & +x_4 & & =0, \\ 2\,x_1 & -x_2 & -x_3 & +2\,x_4 & -x_5 & =0 \end{array} \right\} \ ,$$

если скалярное произведение определяется формулой (1.1).

Решение. Базис $\mathbb{E}_1$ составляет **ф.с.р.** системы линейных уравнений, например:

$$X_1=[0,-1,1,0,0],\ X_2=[-1,0,0,1,0],\ X_3=[1,1,0,0,1] \ .$$

Составляем матрицу Грама этой системы векторов и выписываем систему (3.4):

$$\begin{pmatrix} 2 & 0 & -1 \\ 0 & 2 & -1 \\ -1 & -1 & 3 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 4 \end{pmatrix} \ \Rightarrow\ \alpha_1 = {}^7\!/_4,\ \alpha_2 = {}^7\!/_4,\ \alpha_3 = {}^5\!/_2 \ .$$

Ортогональная проекция вектора X на $\mathbb{E}_1$:

$$X^{\parallel} = {}^7\!/_4 X_1 + {}^7\!/_4 X_2 + {}^5\!/_2 X_3 = [{}^3\!/_4,\ {}^3\!/_4,\ {}^7\!/_4,\ {}^7\!/_4,\ {}^5\!/_2]\ ,$$

а

$$X^{\perp} = X - X^{\parallel} = [{}^1\!/_4,\ {}^1\!/_4,\ {}^1\!/_4,\ {}^1\!/_4,\ -{}^1\!/_2] \quad \Rightarrow \left| X^{\perp} \right| = \sqrt{{}^1\!/_2}\ .$$

Ответ. ${}^1\!/_{\sqrt{2}}$.

После того, как решена задача о нахождении расстояния от точки до линейного подпространства, можно решить и более общую задачу.

Теорема 3.4. *Расстояние от точки X до линейного многообразия $\mathbb{M} = X_0 + \mathbb{E}_1$ равно длине ортогональной составляющей вектора $X - X_0$ относительно подпространства $\mathbb{E}_1$.*

Доказательство . Имеем:

$$\min_{Y \in \mathbb{M}} |X - Y| = \min_{Z \in \mathbb{E}_1} |X - (X_0 + Z)| = \min_{Z \in \mathbb{E}_1} |(X - X_0) - Z)|\ .$$

Последняя величина — это расстояние от точки $X - X_0$ до $\mathbb{E}_1$; согласно теореме 3.3 оно равно длине ортогональной составляющей вектора $X - X_0$ относительно $\mathbb{E}_1$. □

На рис. 4 приведена геометрическая интерпретация этого результата в $\mathbb{R}^3$.

В следующем параграфе будет выведена универсальная формула для расстояния от точки до линейного многообразия.

3.3 Интерпретация одной интерполяционной задачи

В §2.2 главы 5 раздела I мы рассматривали задачу поиска псевдорешений системы линейных уравнений

$$\left\{ \begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n & = & b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n & = & b_2, \\ \ldots & & \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n & = & b_m \end{array} \right. \iff AX = \mathcal{B} \qquad (3.5)$$

при $m \geq n$. Подобные системы могут возникать при проведении экспериментов (измерений) с неизбежными для них погрешностями. Поэтому решение системы (3.5) ищется не в традиционном смысле — удовлетворить каждому из уравнений, а в смысле минимизации отклонений левых частей системы от правых. Именно, псевдорешением системы (3.5) назывался набор значений для $x_1, \ldots, x_n$, обеспечивающий минимум величине

$$\sum_{i=1}^{m} [b_i - (a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n)]^2\ . \qquad (3.6)$$

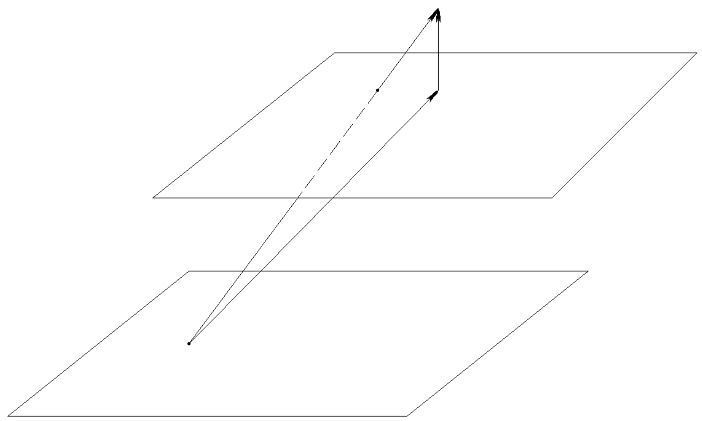


Рис. 4 а)

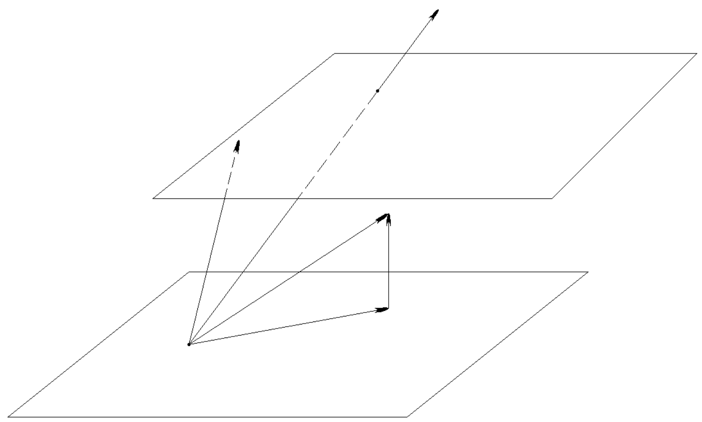


Рис. 4 б)

Оказывается, что задачу поиска псевдорешения можно интерпретировать как задачу поиска проекции вектора на линейное подпространство в $\mathbb{R}^m$. В самом деле, рассмотрим в этом пространстве векторы — столбцы $A_{[1]}, \dots, A_{[n]}$ матрицы A. Левая часть системы (3.5) является, таким образом, линейной комбинацией $x_1 A_{[1]} + \dots + x_n A_{[n]}$, а выражение (3.6) — квадратом расстояния от столбца правых частей $\mathcal{B}$ до этой комбинации. Требование, чтобы величина (3.6) была минимальной эквивалентно следующему: подобрать числа $x_1, \dots, x_n$ так, чтобы расстояние между векторами $\mathcal{B}$ и $x_1 A_{[1]} + \dots + x_n A_{[n]}$ было наименьшим. Обращаясь к предыдущему пункту, видим, что в такой постановке задача эквивалентна поиску проекции вектора $\mathcal{B}$ на подпространство $\mathcal{L}(A_{[1]}, \dots, A_{[n]}) \subset \mathbb{R}^m$. Для нахождения этой проекции мы должны составить систему вида (3.4), т.е., в нашем случае

$$\left\{ \begin{array}{llll} x_1(A_{[1]}, A_{[1]}) & + \dots & + x_n(A_{[1]}, A_{[n]}) & = (\mathcal{B}, A_{[1]}), \\ \quad \dots & & & \quad \dots \\ x_1(A_{[n]}, A_{[1]}) & + \dots & + x_n(A_{[n]}, A_{[n]}) & = (\mathcal{B}, A_{[n]}). \end{array} \right. \tag{3.7}$$

Если мы распишем выражения для участвующих в этих формулах скалярных произведений и вспомним правило умножения матриц, то получим ту же самую **нормальную систему**

$$\left[A^{\top} A \right] X = A^{\top} \mathcal{B} \ , \tag{3.8}$$

которую мы получили в разделе I. В том же разделе было доказано, что эта система имеет единственное решение если $\operatorname{rank} A = n$, т.е. если система столбцов $\{A_{[1]}, \dots, A_{[n]}\}$ **л.н.з.**; этот факт будет подтвержден в следующем параграфе как следствие более общего результата.

4 Свойства определителя Грама

Пусть

$$\mathfrak{G}(X_1, \dots, X_m) \stackrel{\text{def}}{=} \left| \begin{array}{cccc} (X_1, X_1) & (X_1, X_2) & \dots & (X_1, X_m) \\ (X_2, X_1) & (X_2, X_2) & \dots & (X_2, X_m) \\ \dots & & & \dots \\ (X_m, X_1) & (X_m, X_2) & \dots & (X_m, X_m) \end{array} \right| \tag{4.1}$$

Теорема 4.1. $\mathfrak{G}(X_1, \dots, X_m) = 0$ *тогда и только тогда, когда система векторов* $\{X_1, \dots, X_m\}$ *линейно зависима.*

Доказательство . **Н.** Если система $\{X_1, \dots, X_m\}$ **л.з.**, то какой-то из векторов должен линейно выражаться через остальные. Пусть для определенности

$$X_m = \gamma_1 X_1 + \dots + \gamma_{m-1} X_{m-1} \ .$$

Вычтем из последнего столбца (4.1) первый столбец, домноженный на γ_1, второй, домноженный на γ_2, и т.д., $(m-1)$-й, домноженный на γ_{m-1}:

$$\mathfrak{G} = \left| \begin{array}{cccc} * & * & \dots & (X_1, X_m - \gamma_1 X_1 - \dots - \gamma_{m-1} X_{m-1}) \\ * & * & \dots & (X_2, X_m - \gamma_1 X_1 - \dots - \gamma_{m-1} X_{m-1}) \\ \dots & & & \dots \\ * & * & \dots & (X_m, X_m - \gamma_1 X_1 - \dots - \gamma_{m-1} X_{m-1}) \end{array} \right|$$

Все элементы последнего столбца обращаются в нуль, поэтому $\mathfrak{G} = 0$.

Д. Пусть $\mathfrak{G}(X_1, \dots, X_m) = 0$. Рассмотрим уравнение

$$\alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_m X_m = \mathbb{O}$$

относительно скаляров $\alpha_1, \alpha_2, \dots, \alpha_m$. Если мы домножим его (скалярно) на векторы $X_1, X_2, \dots, X_m$, то получим систему линейных уравнений

$$\left(\begin{array}{cccc} (X_1, X_1) & (X_1, X_2) & \dots & (X_1, X_m) \\ (X_2, X_1) & (X_2, X_2) & \dots & (X_2, X_m) \\ \dots & & & \dots \\ (X_m, X_1) & (X_m, X_2) & \dots & (X_m, X_m) \end{array} \right) \left(\begin{array}{c} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_m \end{array} \right) = \mathbb{O}_{m \times 1}$$

с определителем равным нулю по предположению. На основании следствия 1 к теореме 12.2, главы 4 такая система имеет нетривиальное решение: $\alpha_1 = \alpha_1^*, \alpha_2 = \alpha_2^*, \dots, \alpha_m = \alpha_m^*$, $\exists \alpha_j^* \neq 0$. Покажем, что линейная комбинация

$$X^* \stackrel{\text{def}}{=} \alpha_1^* X_1 + \alpha_2^* X_2 + \dots + \alpha_m^* X_m$$

дает нулевой вектор. В самом деле,

$$\begin{aligned} (X^*, X^*) &= (\alpha_1^*, \alpha_2^*, \dots, \alpha_m^*) \left(\begin{array}{cccc} (X_1, X_1) & (X_1, X_2) & \dots & (X_1, X_m) \\ (X_2, X_1) & (X_2, X_2) & \dots & (X_2, X_m) \\ \dots & & & \dots \\ (X_m, X_1) & (X_m, X_2) & \dots & (X_m, X_m) \end{array} \right) \left(\begin{array}{c} \alpha_1^* \\ \alpha_2^* \\ \dots \\ \alpha_m^* \end{array} \right) = \\ &= (\alpha_1^*, \alpha_2^*, \dots, \alpha_m^*) \cdot \mathbb{O}_{m \times 1} = 0, \end{aligned}$$

откуда и следует, что $X^* = \mathbb{O}$, т.е. система $\{X_1, \dots, X_m\}$ **л.з.** □

Теорема 4.2. $\mathfrak{G}(X_1, \dots, X_m) \geq 0$ *для любой системы векторов* $\{X_1, \dots, X_m\}$.

Доказательство . Если система $\{X_1, \dots, X_m\}$ **л.з.**, то $\mathfrak{G}(X_1, \dots, X_m) = 0$ по теореме 4.1. Пусть $\{X_1, \dots, X_m\}$ **л.н.з.** Это означает, что при любом ненулевом наборе скаляров α_j вектор $Y = \alpha_1 X_1 + \dots + \alpha_m X_m$ будет ненулевым: $Y \neq \mathbb{O}$. Следовательно

$$0 < (Y, Y) = (\alpha_1, \dots, \alpha_m) \left(\begin{array}{ccc} (X_1, X_1) & \dots & (X_1, X_m) \\ \dots & & \dots \\ (X_m, X_1) & \dots & (X_m, X_m) \end{array} \right) \left(\begin{array}{c} \alpha_1 \\ \vdots \\ \alpha_m \end{array} \right) \qquad (4.2)$$

при любых $(\alpha_1, \dots, \alpha_m) \neq \mathbb{O}$. Это означает положительную определенность квадратичной формы (4.2). По критерию Сильвестра[12] должны быть выполнены неравенства

$$\mathfrak{G}(X_1) > 0, \mathfrak{G}(X_1, X_2) > 0, \dots, \mathfrak{G}(X_1, X_2, \dots, X_m) > 0 \ .$$

□

[12] Раздел I, глава 7, теорема 4.2.

Замечание. Неравенство (1.6) (Коши–Буняковского) является частным случаем доказанной теоремы.

Теорема 4.3. *Пусть* $X_m^{\perp}$ *означает ортогональную составляющую вектора* X_m *относительно* $\mathcal{L}(X_1, \dots, X_{m-1})$. *Тогда*

$$\mathfrak{G}(X_1, \dots, X_{m-1}, X_m) = \mathfrak{G}(X_1, \dots, X_{m-1}) \left| X_m^{\perp} \right|^2 . \tag{4.3}$$

Доказательство . Пусть $X_m^{\parallel}$ означает ортогональную проекцию вектора X_m на $\mathcal{L}(X_1, \dots, X_{m-1})$: $X_m = X_m^{\parallel} + X_m^{\perp}$. Поскольку $X_m^{\perp} \perp \mathcal{L}(X_1, \dots, X_{m-1})$, то имеем:

$$(X_j, X_m) = \left(X_j, X_m^{\parallel} + X_m^{\perp}\right) = \left(X_j, X_m^{\parallel}\right) + \left(X_j, X_m^{\perp}\right) = \left(X_j, X_m^{\parallel}\right)$$

при $j \in \{1, \dots, m-1\}$,

$$\begin{aligned} (X_m, X_m) &= \left(X_m^{\parallel} + X_m^{\perp}, X_m^{\parallel} + X_m^{\perp}\right) = \left(X_m^{\parallel}, X_m^{\parallel}\right) + 2\left(X_m^{\parallel}, X_m^{\perp}\right) + \\ &+ \left(X_m^{\perp}, X_m^{\perp}\right) = \left(X_m^{\parallel}, X_m^{\parallel}\right) + \left(X_m^{\perp}, X_m^{\perp}\right). \end{aligned}$$

С помощью этих равенств преобразуем грамиан $\mathfrak{G}(X_1, \dots, X_{m-1}, X_m) =$

$$\begin{aligned} &= \begin{vmatrix} (X_1, X_1) & \dots & (X_1, X_{m-1}) & (X_1, X_m) \\ \dots & & & \dots \\ (X_{m-1}, X_1) & \dots & (X_{m-1}, X_{m-1}) & (X_{m-1}, X_m) \\ (X_m, X_1) & \dots & (X_m, X_{m-1}) & (X_m, X_m) \end{vmatrix} = \\ &= \begin{vmatrix} (X_1, X_1) & \dots & (X_1, X_{m-1}) & (X_1, X_m^{\parallel}) \\ \dots & & & \dots \\ (X_{m-1}, X_1) & \dots & (X_{m-1}, X_{m-1}) & (X_{m-1}, X_m^{\parallel}) \\ (X_m^{\parallel}, X_1) & \dots & (X_m^{\parallel}, X_{m-1}) & (X_m^{\parallel}, X_m^{\parallel}) + (X_m^{\perp}, X_m^{\perp}) \end{vmatrix} = \\ &= \underbrace{\mathfrak{G}(X_1, \dots, X_{m-1}, X_m^{\parallel})}_{=0 \text{ по теореме 4.1}} + \begin{vmatrix} (X_1, X_1) & \dots & (X_1, X_{m-1}) & 0 \\ \dots & & & \dots \\ (X_{m-1}, X_1) & \dots & (X_{m-1}, X_{m-1}) & 0 \\ (X_m, X_1) & \dots & (X_m, X_{m-1}) & (X_m^{\perp}, X_m^{\perp}) \end{vmatrix} = \\ &= \mathfrak{G}(X_1, \dots, X_{m-1}) \left| X_m^{\perp} \right|^2 . \end{aligned}$$

□

Теперь мы можем обеспечить теоремы 3.3 и 3.4 красивыми формулами.

Следствие 1. *При* **л.н.з.** *системе* $\{X_1, \dots, X_k\}$ *расстояние от точки* X *до* $\mathcal{L}(X_1, \dots, X_k)$ *вычисляется по формуле*

$$\sqrt{\mathfrak{G}(X_1, \dots, X_k, X) \Big/ \mathfrak{G}(X_1, \dots, X_k)} \ ,$$

а до линейного многообразия $X_0 + \mathcal{L}(X_1, \dots, X_k)$ — по формуле

$$\sqrt{\mathfrak{G}(X_1, \dots, X_k, X - X_0) \Big/ \mathfrak{G}(X_1, \dots, X_k)} \ .$$

Следствие 2.

$$\mathfrak{G}(X_1, \dots, X_{m-1}, X_m) \le |X_1|^2 \times \dots \times |X_{m-1}|^2 \, |X_m|^2 \tag{4.4}$$

Доказательство следует из формулы (4.3) и из того, что

$$|X_m|^2 = \left| X_m^{\parallel} + X_m^{\perp} \right|^2 = \left| X_m^{\parallel} \right|^2 + \left| X_m^{\perp} \right|^2 \ge \left| X_m^{\perp} \right|^2 \ \Rightarrow$$

$$\mathfrak{G}(X_1, \dots, X_{m-1}, X_m) \le \mathfrak{G}(X_1, \dots, X_{m-1}) \, |X_m|^2 \ .$$

При этом равенство возможно тогда и только тогда, когда $X_m \perp X_1, \dots, X_m \perp X_{m-1}$. Продолжая далее процесс вынесения векторов из-под знака грамиана, окончательно придем к неравенству (4.4); в нем равенство возможно тогда и только тогда, когда либо все векторы попарно ортогональны, либо один из них — нулевой. □

Следствие 3. *Для произвольной квадратной матрицы*

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

справедливо следующее **неравенство Адамара**[13]*:*

$$|\det A| \le \sqrt{\sum_{j=1}^n a_{1j}^2} \sqrt{\sum_{j=1}^n a_{2j}^2} \times \dots \times \sqrt{\sum_{j=1}^n a_{nj}^2} \ . \tag{4.5}$$

Иными словами: модуль определителя матрицы не превосходит произведения "длин" его строк. Аналогичное утверждение справедливо и относительно столбцов матрицы.

Доказательство . Обозначим j−ю строку матрицы A через $A^{[j]}$. Тогда

$$(\det A)^2 = \det \left(A \cdot A^{\top} \right) = \det \begin{bmatrix} \left(A^{[1]}, A^{[1]}\right) \left(A^{[1]}, A^{[2]}\right) \dots \left(A^{[1]}, A^{[n]}\right) \\ \left(A^{[2]}, A^{[1]}\right) \left(A^{[2]}, A^{[2]}\right) \dots \left(A^{[2]}, A^{[n]}\right) \\ \dots \qquad\qquad\qquad \dots \\ \left(A^{[n]}, A^{[1]}\right) \left(A^{[n]}, A^{[2]}\right) \dots \left(A^{[n]}, A^{[n]}\right) \end{bmatrix} =$$

$$= \mathfrak{G}\left(A^{[1]}, A^{[2]}, \dots, A^{[n]} \right) \overset{(4.4)}{\le} \left| A^{[1]} \right|^2 \left| A^{[2]} \right|^2 \times \dots \times \left| A^{[n]} \right|^2 \ .$$

Равенство возможно тогда и только тогда, когда либо все строки попарно ортогональны, либо хотя бы одна строка — нулевая. □

[13] **Адамар Жак Саломон** (Hadamard Jacques Salomon, 1865–1963) — французский математик.

Определение. **Объемом параллелепипеда**, натянутого на систему векторов $\{X_1, \dots, X_{m-1}, X_m\}$, называется величина, определяемая рекурсивно по m:

$$\begin{aligned} \mathbf{V}_1(X_1) &\stackrel{\text{def}}{=} |X_1| \ , \\ \mathbf{V}_m(X_1, \dots, X_{m-1}, X_m) &\stackrel{\text{def}}{=} \mathbf{V}_{m-1}(X_1, \dots, X_{m-1}) \left| X_m^{\perp} \right| \ . \end{aligned} \qquad (4.6)$$

Здесь $X_m^{\perp}$ означает ортогональную составляющую вектора X_m относительно $\mathcal{L}(X_1, \dots, X_{m-1})$.

Теорема 4.4. *Имеет место равенство:*

$$\mathbf{V}_m(X_1, \dots, X_m) = \sqrt{\mathfrak{G}(X_1, \dots, X_m)} \ .$$

Доказательство проводится индукцией по m. Для $m = 1$:

$$\mathbf{V}_1(X_1) = |X_1| = \sqrt{(X_1, X_1)} = \sqrt{\mathfrak{G}(X_1)} \ .$$

Пусть утверждение справедливо для $m - 1$:

$$\mathbf{V}_{m-1}(X_1, \dots, X_{m-1}) = \sqrt{\mathfrak{G}(X_1, \dots, X_{m-1})} \ .$$

Тогда

$$\mathbf{V}_m(X_1, \dots, X_{m-1}, X_m) \stackrel{(4.6)}{=} \mathbf{V}_{m-1}(X_1, \dots, X_{m-1}) \left| X_m^{\perp} \right| =$$

$$= \sqrt{\mathfrak{G}(X_1, \dots, X_{m-1}) \left| X_m^{\perp} \right|^2} \stackrel{(4.3)}{=} \sqrt{\mathfrak{G}(X_1, \dots, X_{m-1}, X_m)} \ .$$

□

Следствие 1. *В пространстве $\mathbb{R}^n$ модуль определителя матрицы A равен объему параллелепипеда, натянутого на его строки:*

$$\left| \det \left(\begin{array}{cccc} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & & & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{array} \right) \right| = \mathbf{V}_n \left(A^{[1]}, A^{[2]}, \dots, A^{[n]} \right) \ .$$

Глава 3. Линейные отображения

1 Пространство линейных отображений

1.1 Линейное отображение

Определение. **Линейным отображением** векторного пространства $\mathbb{V}$ в векторное пространство $\mathbb{W}$ называется функция (соответствие) $\mathscr{A} : \mathbb{V} \longmapsto \mathbb{W}$ (т.е. определенная на $\mathbb{V}$, имеющая значения в $\mathbb{W}$), обладающая свойством линейности:

$$\mathscr{A}(\alpha_1 X_1 + \alpha_2 X_2) = \alpha_1 \mathscr{A}(X_1) + \alpha_2 \mathscr{A}(X_2) \text{ для } \left\{ \begin{array}{lll} \forall \{X_1, X_2\} & \subset & \mathbb{V} \\ \forall \{\alpha_1, \alpha_2\} & \subset & \mathbb{R} \text{ или } \mathbb{C} \end{array} \right. \tag{1.1}$$

(здесь α_1, α_2 — константы из $\mathbb{R}$ если оба пространства вещественны, и из $\mathbb{C}$, если хотя бы одно из пространств комплексное). Если $Y = \mathscr{A}(X)$, то говорят, что Y — образ вектора X, а X — прообраз вектора Y при отображении $\mathscr{A}$. Пространство $\mathbb{V}$ называется областью определения отображения $\mathscr{A}$.

Пример 1.1. *Линейная форма:*

$$\mathscr{A}(x_1, \dots, x_n) = a_1 x_1 + \dots + a_n x_n, \quad \{a_j\}_{j=1}^n \subset \mathbb{R}$$

является примером линейного отображения $\mathbb{R}^n$ в $\mathbb{R}$. Этот пример показывает, что прообразов у одного и того же элемента из $\mathbb{W}$ может быть несколько:

$$\mathscr{A}(x_1, x_2) = 2x_1 - x_2 \text{ отображает вектора } X_1 = [0, 0] \text{ и } X_2 = [1, 2] \text{ в } 0 \ .$$

Пример 1.2. *Обобщением предыдущего примера является отображение $\mathscr{A} : \mathbb{R}^n \longmapsto \mathbb{R}^m$, задаваемое*

$$\mathscr{A} \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & & & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}$$

при произвольной вещественной матрице.

Пример 1.3. *Предыдущим примерам можно дать и геометрическую интерпретацию. Так, линейное отображение $\mathbb{R}^3 \longmapsto \mathbb{R}^3$:*

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \longmapsto \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$$

задает ортогональную проекцию вектора $X = (x, y, z)$ на плоскость (x, y). (Можно рассматривать его и как отображение $\mathbb{R}^3 \longmapsto \mathbb{R}^2$.) Проектиро-

вание же на произвольное подпространство может быть задано с помощью матрицы. Так, например, отображение

$$\left(\begin{array}{c} x \\ y \\ z \end{array} \right) \longmapsto {}^1\!/_3 \left(\begin{array}{rrr} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{array} \right) \left(\begin{array}{c} x \\ y \\ z \end{array} \right)$$

задает ортогональную проекцию вектора X на многообразие $x+y+z=0$.

Пример 1.4. *Рассмотрим линейное пространство полиномов степени не выше n: $\mathbb{P}_n = \{p(x) \in \mathbb{R}[x] \mid \deg p(x) \le n\}$. Операция дифференцирования*

$$\frac{d}{d\,x}: \ p(x) \longmapsto p'(x)$$

является отображением $\mathbb{P}_n$ в $\mathbb{P}_{n-1}$ линейным поскольку

$$\frac{d}{d\,x}(\alpha_1 p_1(x) + \alpha_2 p_2(x)) = \alpha_1 \frac{d}{d\,x} p_1(x) + \alpha_2 \frac{d}{d\,x} p_2(x) \ .$$

Прообраз любого элемента $\mathbb{P}_{n-1}$ неединствен: $\frac{d}{d\,x}({}^1\!/_2 x^2 + \ const) = x$.

Пример 1.5. *Операцию нахождения первообразной:*

$$\begin{array}{ccc} \displaystyle\int_0^x : p(x) & \longmapsto & \displaystyle\int_0^x p(t) d\,t \\ a_0 x^n + a_1 x^{n-1} + \cdots + a_n & \longmapsto & \dfrac{a_0}{n+1} x^{n+1} + \dfrac{a_1}{n} x^n + \cdots + a_n x \end{array}$$

тоже можно рассматривать как линейное отображение $\mathbb{P}_n \longmapsto \mathbb{P}_{n+1}$. При этом прообраз каждого полинома из $\mathbb{P}_{n+1}$ (если существует) будет единствен.

Упражнение 1.1. *Доказать, что линейное отображение $\mathscr{A}$ отображает произвольную прямую пространства $\mathbb{V}$ в прямую или точку пространства $\mathbb{W}$.*

Упражнение 1.2. *Является ли линейным отображение*

$$A \longmapsto \mathrm{Sp}(A) \ ,$$

определенное в пространстве квадратных матриц порядка n?

1.2 Свойства линейных отображений

Определение. Два линейных отображения $\mathscr{A}$ и $\mathscr{B}$ из $\mathbb{V}$ в $\mathbb{W}$ называются равными если $\mathscr{A}(X) = \mathscr{B}(X)$ для любого $X \in \mathbb{V}$. **Нулевое** отображение определяется условием

$$\mathcal{O}(X) = \mathbb{O}' \ \forall \ X \in \mathbb{V} \ .$$

Теорема 1.1. *Для любого линейного отображения $\mathscr{A}(X)$:*

а) $\mathscr{A}(\mathbb{O}) = \mathbb{O}'$;

б) *если $X_1, \dots, X_k$ — линейно зависимы, то и $\mathscr{A}(X_1), \dots, \mathscr{A}(X_k)$ — линейно зависимы;*

в) *если $\mathscr{A}(X_1), \dots, \mathscr{A}(X_k)$ — линейно независимы, то и $X_1, \dots, X_k$ — линейно независимы.*

Доказательство . Имеем на основании равенства (1.1):

$$\mathscr{A}(\mathbb{O}) = \mathscr{A}(2 \cdot \mathbb{O}) = 2\mathscr{A}(\mathbb{O}) \ ,$$

что возможно только при $\mathscr{A}(\mathbb{O}) = \mathbb{O}'$.

Тогда

$$\alpha_1 X_1 + \dots + \alpha_k X_k = \mathbb{O} \quad \Longrightarrow \quad \alpha_1 \mathscr{A}(X_1) + \dots + \alpha_k \mathscr{A}(X_k) = \mathbb{O}',$$

из чего и следуют утверждения пунктов **б)** и **в)**. □

Теорема 1.2. *Пусть $\{X_1, \dots, X_n\}$ — произвольный базис $\mathbb{V}$, а $Y_1, \dots, Y_n$ — произвольные векторы из $\mathbb{W}$. Существует единственное линейное отображение $\mathscr{A} : \mathbb{V} \longmapsto \mathbb{W}$ такое, что*

$$\mathscr{A}(X_1) = Y_1, \dots, \mathscr{A}(X_n) = Y_n \ .$$

Доказательство . Поскольку векторы $X_1, \dots, X_n$ — базисные, то существует и единственно разложение любого $X \in \mathbb{V}$: $X = x_1 X_1 + \dots + x_n X_n$. Зададим отображение $\mathscr{A} : \mathbb{V} \longmapsto \mathbb{W}$ формулой

$$\mathscr{A}(X) \stackrel{\text{def}}{=} x_1 Y_1 + \dots + x_n Y_n \ .$$

Легко проверить свойство его линейности (1.1). Кроме того:

$$\mathscr{A}(X_j) = \mathscr{A}(0 \cdot X_1 + \dots + 1 \cdot X_j + \dots + 0 \cdot X_n) = 0 \cdot Y_1 + \dots + 1 \cdot Y_j + \dots + 0 \cdot Y_n = Y_j,$$

т.е. оно удовлетворяет условиям теоремы.

Предположим теперь, что существует еще одно отображение $\mathscr{B}(X)$, удовлетворяющее этим условиям: $\mathscr{B}(X_j) = Y_j$. Тогда

$$\mathscr{A}(X) = x_1 Y_1 + \dots + x_n Y_n = x_1 \mathscr{B}(X_1) + \dots + x_n \mathscr{B}(X_n) = \mathscr{B}(X),$$

и на основании определения $\mathscr{A}(X) = \mathscr{B}(X)$. □

Определение. Отображение $\mathscr{S} : \mathbb{V} \longmapsto \mathbb{W}$ называется **суммой** линейных отображений $\mathscr{A}$ и $\mathscr{B}$ если $\mathscr{S}(X) = \mathscr{A}(X) + \mathscr{B}(X)$ для любого $X \in \mathbb{V}$. Отображение $\mathscr{F} : \mathbb{V} \longmapsto \mathbb{W}$ называется **произведением** линейного отображения $\mathscr{A}$ на число λ если $\mathscr{F}(X) = \lambda \cdot \mathscr{A}(X)$ для любого $X \in \mathbb{V}$.

Теорема 1.3. *Отображения $\mathscr{S}$ и $\mathscr{F}$ — линейные.*

Пример 1.6. *В пространстве полиномов* $\mathbb{P}_n$ *операцию нахождения второй производной*

$$\frac{d^2}{d\,x^2} : p(x) \longmapsto p''(x)$$

тоже можно рассматривать как линейное отображение $\mathbb{P}_n \longmapsto \mathbb{P}_{n-1}$. *Линейным также будет и отображение*

$$\frac{d^2}{d\,x^2} \times \Box + 2\frac{d}{d\,x} \times \Box : \; p(x) \;\longmapsto\; p''(x) + 2p'(x).$$

.

Теорема 1.4. *Множество* $\mathcal{H}om(\mathbb{V}, \mathbb{W})$ *всех линейных отображений из* $\mathbb{V}$ *в* $\mathbb{W}$ *образует линейное пространство и*

$$\dim \mathcal{H}om(\mathbb{V}, \mathbb{W}) = \dim \mathbb{V} \cdot \dim \mathbb{W} \;.$$

Доказательство . Пусть $\{X_1, \ldots, X_n\}$ — базис $\mathbb{V}$, а $\{Y_1, \ldots, Y_m\}$ — базис $\mathbb{W}$. Определим отображение $\mathscr{A}_{jk}$ условиями

$$\begin{aligned} \mathscr{A}_{jk}(X_1) = \mathbb{O}', \ldots, \quad & \mathscr{A}_{jk}(X_{k-1}) = \mathbb{O}', \mathscr{A}_{jk}(X_k) = Y_j, \\ & \mathscr{A}_{jk}(X_{k+1}) = \mathbb{O}', \ldots, \mathscr{A}_{jk}(X_n) = \mathbb{O}'. \end{aligned} \tag{1.2}$$

По теореме 1.2 существует единственное линейное отображение с таким свойством. Любое другое линейное отображение $\mathscr{A}$ можно представить в виде линейной комбинации отображений

$$\{\mathscr{A}_{jk}\}_{j=1,\ldots,m;\; k=1,\ldots,n} \;. \tag{1.3}$$

Действительно, пусть имеют место равенства

$$\left\{ \begin{array}{lcr} \mathscr{A}(X_1) & = & \alpha_{11}Y_1 + \alpha_{21}Y_2 + \cdots + \alpha_{m1}Y_m, \\ \ldots & & \ldots, \\ \mathscr{A}(X_n) & = & \alpha_{1n}Y_1 + \alpha_{2n}Y_2 + \cdots + \alpha_{mn}Y_m. \end{array} \right. \tag{1.4}$$

Определим тогда новое линейное отображение $\mathscr{A}_*$ формулой:

$$\mathscr{A}_* \stackrel{\mathrm{def}}{=} \sum_{\substack{1 \le j \le m \\ 1 \le k \le n}} \alpha_{jk} \mathscr{A}_{jk} \;,$$

и найдем образ базисного вектора X_ℓ. Согласно формулам (1.2): $\mathscr{A}_{jk}(X_\ell) = \mathbb{O}'$ при $k \ne \ell$ и $\mathscr{A}_{j\ell}(X_\ell) = Y_j$. Следовательно,

$$\mathscr{A}_*(X_\ell) = \sum_{\substack{1 \le j \le m \\ 1 \le k \le n}} \alpha_{jk} \mathscr{A}_{jk}(X_\ell) = \sum_{1 \le j \le m} \alpha_{j\ell} \mathscr{A}_{j\ell}(X_\ell) = \alpha_{1\ell}Y_1 + \cdots + \alpha_{m\ell}Y_m.$$

Правая часть оказалась равной $\mathscr{A}(X_\ell)$, ввиду формул (1.4). Итак, образы базисных векторов одинаковы при отображениях $\mathscr{A}$ и $\mathscr{A}_*$. По теореме 1.2: $\mathscr{A} = \mathscr{A}_*$.

Осталось доказать линейную независимость множества (1.3). Предположим, что

$$\sum_{\substack{1\le j\le m\\ 1\le k\le n}} \beta_{jk}\mathscr{A}_{jk} = \mathcal{O}.$$

и найдем образ вектора X_ℓ ($\ell \in \{1,\dots,n\}$) под действием левого и правого отображения. На основании равенств (1.2):

$$\beta_{1\ell}\mathscr{A}_{1\ell}(X_\ell) + \dots + \beta_{m\ell}\mathscr{A}_{m\ell}(X_\ell) = \beta_{1\ell}Y_1 + \dots + \beta_{m\ell}Y_m = \mathbb{O}' \ ,$$

что возможно лишь при $\beta_{1\ell} = 0, \dots, \beta_{m\ell} = 0$.

Итак, отображения (1.3) — базисные для $\mathcal{H}om(\mathbb{V},\mathbb{W})$. Их число равно $n\cdot m = \dim\mathbb{V}\cdot\dim\mathbb{W}$. □

2 Ядро и образ линейного отображения

Определение. Для линейного отображения $\mathscr{A}$ его **ядром** называется множество векторов из $\mathbb{V}$, отображающихся в $\mathbb{O}' \in \mathbb{W}$:

$$\mathcal{K}er(\mathscr{A}) \stackrel{\text{def}}{=} \left\{X \in \mathbb{V} \middle| \mathscr{A}(X) = \mathbb{O}'\right\} \ ;$$

а его **образом** называется множество всех векторов из $\mathbb{W}$, для каждого из которых существует прообраз из $\mathbb{V}$:

$$\mathcal{I}m(\mathscr{A}) \stackrel{\text{def}}{=} \{Y \in \mathbb{W} \mid \exists X \in \mathbb{V},\ \mathscr{A}(X) = Y\} \ .$$

Теорема 2.1. *$\mathcal{K}er(\mathscr{A})$ и $\mathcal{I}m(\mathscr{A})$ являются линейными подпространствами соответствующих пространств.*

Теорема 2.2. *Если $X_1,\dots,X_n$ — произвольный базис $\mathbb{V}$, то $\mathcal{I}m(\mathscr{A})$ совпадает с линейной оболочкой образов этих векторов*

$$\mathcal{I}m(\mathscr{A}) = \mathcal{L}\left(\mathscr{A}(X_1),\dots,\mathscr{A}(X_n)\right) \ .$$

Доказательство . Действительно, любой вектор Y из $\mathcal{I}m(\mathscr{A})$ является образом какого-то вектора $X = x_1X_1 + \dots + x_nX_n$, тогда на основании линейности отображения:

$$Y = \mathscr{A}(X) = x_1\mathscr{A}(X_1) + \dots + x_n\mathscr{A}(X_n) \in \mathcal{L}\left(\mathscr{A}(X_1),\dots,\mathscr{A}(X_n)\right) \ .$$

Таким образом

$$\mathcal{I}m(\mathscr{A}) \subset \mathcal{L}\left(\mathscr{A}(X_1),\dots,\mathscr{A}(X_n)\right) \ .$$

Обратно, поскольку векторы $\mathscr{A}(X_1),\dots,\mathscr{A}(X_n)$ принадлежат $\mathcal{I}m(\mathscr{A})$, то по теореме 2.1 и любая линейная комбинация этих векторов должна принадлежать $\mathcal{I}m(\mathscr{A})$:

$$\mathcal{L}\left(\mathscr{A}(X_1),\dots,\mathscr{A}(X_n)\right) \subset \mathcal{I}m(\mathscr{A}) \ .$$

Из двух взаимных включений множеств следует их равенство. □

Пример 2.1. *Найти ядро и образ отображения* $\mathbb{R}^3 \longmapsto \mathbb{R}^4$

$$\mathscr{A}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_3 \\ 0 \\ x_1+x_2+x_3 \\ x_1+x_2-x_3 \end{pmatrix} .$$

Решение. Для определения $\mathcal{K}er(\mathscr{A})$ найдем фундаментальную систему решений системы уравнений

$$\left\{ \begin{array}{rcl} x_3 & = & 0 \\ 0 & = & 0 \\ x_1+x_2+x_3 & = & 0 \\ x_1+x_2-x_3 & = & 0 \end{array} \right. \quad \Longrightarrow X_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$$

Имеем $\dim \mathcal{K}er(\mathscr{A}) = 1$ и $\{X_1\}$ — базис.

Теперь для нахождения $\mathcal{I}m(\mathscr{A})$ воспользуемся теоремой 2.2: базис следует искать среди векторов

$$Y_1 = \mathscr{A}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix},\ Y_2 = \mathscr{A}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix},\ Y_3 = \mathscr{A}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ -1 \end{pmatrix} .$$

Имеем: $\dim \mathcal{I}m(\mathscr{A}) = 2$ и $\{Y_1, Y_3\}$ — базис. △

Упражнение 2.1. *Найти ядра и образы линейных отображений примеров* 1.3 – 1.6.

Определение. Для линейного отображения $\mathscr{A}$ его

дефектом	называется размерность	ядра	$\operatorname{dfc}(\mathscr{A})$	$\stackrel{\text{def}}{=}$	$\dim(\mathcal{K}er(\mathscr{A}))$
рангом		образа	$\operatorname{rank}(\mathscr{A})$	$\stackrel{\text{def}}{=}$	$\dim(\mathcal{I}m(\mathscr{A}))$.

Отображение называется **невырожденным** если $\operatorname{dfc}(\mathscr{A}) = 0$.

Теорема 2.3.

$$\dim \mathbb{V} = \operatorname{dfc}(\mathscr{A}) + \operatorname{rank}(\mathscr{A}) \ . \tag{2.1}$$

Доказательство . Пусть $\operatorname{rank}(\mathscr{A}) = \mathfrak{r}$ и система $\{Y_1, \dots, Y_\mathfrak{r}\}$ составляет базис $\mathcal{I}m(\mathscr{A})$. Тогда, согласно пункту **в)** теоремы 1.1, их прообразы $X_1, \dots, X_\mathfrak{r}$ $(Y_j = \mathscr{A}(X_j))$ линейно независимы. Пусть $\operatorname{dfc}(\mathscr{A}) = d$ и система $\{X_1^*, \dots, X_d^*\}$ составляет базис $\mathcal{K}er(\mathscr{A})$. Докажем, что система

$$\{X_1, \dots, X_\mathfrak{r}, X_1^*, \dots, X_d^*\} \tag{2.2}$$

является базисом $\mathbb{V}$.

Образ любого вектора $X \in \mathbb{V}$ представи́м в виде линейной комбинации $Y_1, \dots, Y_\mathfrak{r}$:

$$\mathscr{A}(X) = \beta_1 Y_1 + \dots + \beta_\mathfrak{r} Y_\mathfrak{r} \ .$$

Тогда вектор

$$\widetilde{X} \stackrel{\text{def}}{=} X - (\beta_1 X_1 + \dots + \beta_{\mathfrak{r}} X_{\mathfrak{r}}) \tag{2.3}$$

должен принадлежать $\mathcal{Ker}(\mathscr{A})$:

$$\mathscr{A}(\widetilde{X}) = \mathscr{A}(X) - \beta_1 \mathscr{A}(X_1) - \dots - \beta_{\mathfrak{r}} \mathscr{A}(X_{\mathfrak{r}}) = \mathscr{A}(X) - \beta_1 Y_1 - \dots - \beta_{\mathfrak{r}} Y_{\mathfrak{r}} = \mathbb{O}'.$$

Следовательно $\widetilde{X}$ представи́м в виде линейной комбинации $X_1^*, \dots, X_d^*$:

$$\widetilde{X} = \alpha_1 X_1^* + \dots + \alpha_d X_d^* \ . \tag{2.4}$$

Из двух равенств (2.3) и (2.4) вытекает, что вектор X представи́м в виде линейной комбинации векторов (2.2).

Осталось показать, что вектора (2.2) линейно независимы. Пусть

$$\gamma_1 X_1 + \dots + \gamma_{\mathfrak{r}} X_{\mathfrak{r}} + \delta_1 X_1^* + \dots + \delta_d X_d^* = \mathbb{O} \tag{2.5}$$

при некотором наборе констант. Тогда действие $\mathscr{A}$ на обе части равенства приводит к

$$\gamma_1 Y_1 + \dots + \gamma_{\mathfrak{r}} Y_{\mathfrak{r}} + \mathbb{O}' + \dots + \mathbb{O}' = \mathbb{O}' \ ,$$

что возможно только при $\gamma_1 = 0, \dots, \gamma_{\mathfrak{r}} = 0$ поскольку $\{Y_1, \dots, Y_{\mathfrak{r}}\}$ — базис $\mathcal{Im}(\mathscr{A})$. Таким образом равенство (2.5) вырождается в $\delta_1 X_1^* + \dots + \delta_d X_d^* = \mathbb{O}$, которое также возможно только при $\delta_1 = 0, \dots, \delta_d = 0$, поскольку $\{X_1^*, \dots, X_d^*\}$ — базис $\mathcal{Ker}(\mathscr{A})$. Равенство (2.5) оказывается возможным только при нулевом наборе коэффициентов.

Мы доказали, что вектора (2.2) — базисные для $\mathbb{V}$, но тогда $\mathfrak{r} + d = \dim \mathbb{V}$. □

Замечание. Утверждение $\mathbb{V} = \mathcal{Ker}(\mathscr{A}) \oplus \mathcal{Im}(\mathscr{A})$ вообще говоря неверно!

Следствие 1. *Линейное отображение $\mathscr{A}$ невырождено тогда и только тогда, когда у каждого образа существует единственный прообраз.*

Доказательство . Н. Если $\mathscr{A}$ невырождено, то $\mathcal{Ker}(\mathscr{A}) = \{\mathbb{O}\}$, т.е. единственным вектором из $\mathbb{V}$, отображающимся в $\mathbb{O}'$ должен быть $\mathbb{O}$. Если предположить неединственность прообраза для какого-то $Y \in \mathbb{W}$: $Y = \mathscr{A}(X_1) = \mathscr{A}(X_2)$ при $X_1 \neq X_2$, то

$$\mathbb{O}' = \mathscr{A}(X_1) - \mathscr{A}(X_2) = \mathscr{A}(X_1 - X_2)$$

и получаем противоречие с единственностью прообраза у $\mathbb{O}'$.

Д. Пусть $\mathscr{A}(X_1) \neq \mathscr{A}(X_2)$ для любых $X_1 \neq X_2$. Если бы $\mathcal{Ker}(\mathscr{A})$ имело ненулевую размерность, то существовал бы $X \neq \mathbb{O}$ такой, что $\mathscr{A}(X) = \mathbb{O}'$, что противоречило бы предыдущей фразе: $\mathscr{A}(X) = \mathscr{A}(\mathbb{O})$. □

Упражнение 2.2. *Какие из отображений примеров* 1.3 – 1.6 *невырождены?*

Теорема 2.4. *Пусть $\mathbb{V}_1$ — линейное подпространство $\mathbb{V}$, а $\mathbb{W}_1$ — линейное подпространство $\mathbb{W}$, причем*

$$\dim \mathbb{V}_1 + \dim \mathbb{W}_1 = \dim \mathbb{V} \ . \tag{2.6}$$

Тогда существует линейное отображение $\mathscr{A} : \mathbb{V} \longmapsto \mathbb{W}$ такое, что

$$\mathcal{K}er(\mathscr{A}) = \mathbb{V}_1, \quad \mathcal{I}m(\mathscr{A}) = \mathbb{W}_1. \tag{2.7}$$

Доказательство . Пусть $\dim \mathbb{V} = n, \dim \mathbb{V}_1 = n_1, \dim \mathbb{W}_1 = m_1$, имеем по (2.6): $n = n_1 + m_1$. Подберем базис $X_1, \ldots, X_n$ пространства $\mathbb{V}$ так, чтобы векторы $X_1, \ldots, X_{n_1}$ образовывали базис $\mathbb{V}_1$ (см. главу 1, §5). Пусть $\{Y_1, \ldots, Y_{m_1}\}$ — базис $\mathbb{W}_1$. По теореме 1.2 существует единственное линейное отображение $\mathscr{A} : \mathbb{V} \longmapsto \mathbb{W}$ такое, что

$$\mathscr{A}(X_1) = \mathbb{O}', \ldots, \mathscr{A}(X_{n_1}) = \mathbb{O}', \mathscr{A}(X_{n_1+1}) = Y_1, \ldots, \mathscr{A}(X_n) = Y_{m_1} \ .$$

Это отображение — именно то, что нам требуется. Действительно, на основании теоремы 2.2 имеем

$$\mathcal{I}m(\mathscr{A}) = \mathcal{L}\left(\mathscr{A}(X_1), \ldots, \mathscr{A}(X_{n_1}), \mathscr{A}(X_{n_1+1}), \ldots, \mathscr{A}(X_n)\right) = \mathcal{L}\left(Y_1, \ldots, Y_{m_1}\right) = \mathbb{W}_1 \ .$$

А тогда $\mathcal{K}er(\mathscr{A}) = \mathbb{V}_1$ по способу построения. $\square$

3 Матрица линейного отображения

3.1 Определение

Рассмотрим линейное отображение $\mathscr{A} : \mathbb{V} \longmapsto \mathbb{W}$, и пусть $\{X_1, \ldots, X_n\}$ — базис $\mathbb{V}$, $\{Y_1, \ldots, Y_m\}$ — базис $\mathbb{W}$. Найдем координаты векторов $\mathscr{A}(X_1), \ldots, \mathscr{A}(X_n)$ в базисе $\{Y_1, \ldots, Y_m\}$: составим формулы (1.4).

Определение. Матрица

$$\mathbf{A} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2n} \\ \ldots & & & \ldots \\ \alpha_{m1} & \alpha_{m2} & \ldots & \alpha_{mn} \end{pmatrix}_{m \times n} ,$$

по столбцам которой стоят координаты образов базисных векторов, называется **матрицей линейного отображения** $\mathscr{A}$ в выбранных базисах.

Теорема 3.1. *Координаты произвольного вектора $X = x_1 X_1 + \cdots + x_n X_n$ и его образа $\mathscr{A}(X) = y_1 Y_1 + \cdots + y_m Y_m$ связаны формулой:*

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \mathbf{A} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \ . \tag{3.1}$$

Доказательство . С помощью формул (1.4) получаем:

$$\begin{aligned}
\mathscr{A}(X) &= \mathscr{A}(x_1X_1+\cdots+x_nX_n) = x_1\mathscr{A}(X_1)+\cdots+x_n\mathscr{A}(X_n) = \\
&= x_1(\alpha_{11}Y_1+\cdots+\alpha_{m1}Y_m)+\cdots+x_n(\alpha_{1n}Y_1+\cdots+\alpha_{mn}Y_m) = \\
&= \underbrace{(x_1\alpha_{11}+\cdots+x_n\alpha_{1n})}_{y_1}Y_1+\cdots+\underbrace{(x_1\alpha_{m1}+\cdots+x_n\alpha_{mn})}_{y_m}Y_m,
\end{aligned}$$

откуда и следует утверждение теоремы. □

Теорема 3.2. *При фиксированных базисах матрица линейного отображения определяется единственным образом и*

$$\operatorname{rank}(\mathscr{A}) = \operatorname{rank} A,\ \operatorname{dfc}(\mathscr{A}) = n - \operatorname{rank} A\ .$$

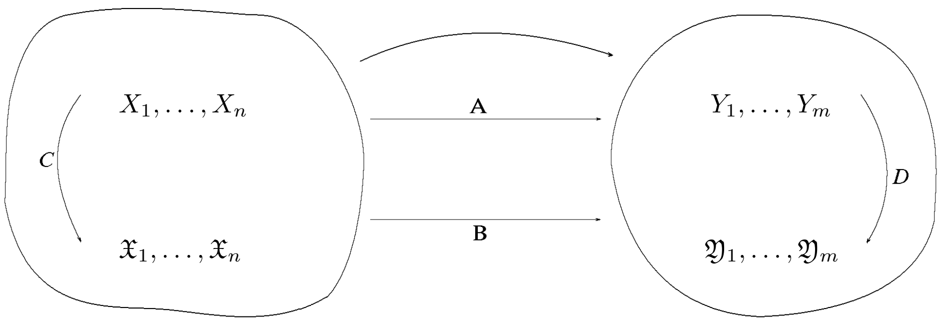
Теорема 3.3. *Существует изоморфизм между линейными пространствами $\mathcal{H}om(\mathbb{V},\mathbb{W})$ (линейных отображений из $\mathbb{V}$ в $\mathbb{W}$) и $\mathbb{M}_{m\times n}$ ($m\times n$ матриц над соответствующим полем).*

Доказательство . Зафиксируем некоторые базисы пространств $\mathbb{V}$ и $\mathbb{W}$. В этих базисах каждому отображению $\mathscr{A}\in\mathcal{H}om(\mathbb{V},\mathbb{W})$ соответствует его матрица $\mathbf{A}\in\mathbb{M}_{m\times n}$, и она единственна. Такое соответствие взаимно-однозначно: если бы матрица $\mathbf{A}$ соответствовала еще одному отображению $\mathscr{A}_1$, то на основании формулы (3.1), было бы выполнено $\mathscr{A}(X)=\mathscr{A}_1(X), \forall X$. Но тогда $\mathscr{A}=\mathscr{A}_1$ по определению.

Линейной комбинации отображений будет соответствовать та же линейная комбинация их матриц. □

Как изменяется матрица линейного отображения $\mathscr{A}$ при изменении базисов?

Пусть $\{\mathfrak{X}_1,\ldots,\mathfrak{X}_n\}$ — новый базис $\mathbb{V}$, $\{\mathfrak{Y}_1,\ldots,\mathfrak{Y}_m\}$— новый базис $\mathbb{W}$, и в этих базисах линейное отображение $\mathscr{A}$ имеет матрицу $\mathbf{B}$.



Теорема 3.4. *Если C — матрица перехода от старого базиса к новому в пространстве $\mathbb{V}$, а D — матрица перехода от старого базиса к новому в пространстве $\mathbb{W}$, то*

$$\mathbf{B} = D^{-1}\cdot\mathbf{A}\cdot C\ . \tag{3.2}$$

Доказательство . Действительно, координаты произвольного вектора

$$X = x_1X_1 + \cdots + x_nX_n = \mathfrak{x}_1\mathfrak{X}_1 + \cdots + \mathfrak{x}_n\mathfrak{X}_n \ ,$$

и его образа

$$Y = \mathscr{A}(X) = y_1Y_1 + \cdots + y_mY_m = \mathfrak{y}_1\mathfrak{Y}_1 + \cdots + \mathfrak{y}_m\mathfrak{Y}_m$$

связаны следующими соотношениями: с одной стороны, на основании теоремы 3.1,

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \mathbf{A} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \qquad \begin{pmatrix} \mathfrak{y}_1 \\ \vdots \\ \mathfrak{y}_m \end{pmatrix} = \mathbf{B} \begin{pmatrix} \mathfrak{x}_1 \\ \mathfrak{x}_2 \\ \vdots \\ \mathfrak{x}_n \end{pmatrix} \ ;$$

с другой стороны, на основании результатов §6 главы 1,

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = C \begin{pmatrix} \mathfrak{x}_1 \\ \mathfrak{x}_2 \\ \vdots \\ \mathfrak{x}_n \end{pmatrix}, \qquad \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = D \begin{pmatrix} \mathfrak{y}_1 \\ \vdots \\ \mathfrak{y}_m \end{pmatrix}.$$

Получаем цепочку равенств:

$$\mathbf{B} \begin{pmatrix} \mathfrak{x}_1 \\ \mathfrak{x}_2 \\ \vdots \\ \mathfrak{x}_n \end{pmatrix} = \begin{pmatrix} \mathfrak{y}_1 \\ \vdots \\ \mathfrak{y}_m \end{pmatrix} = D^{-1} \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = D^{-1}\mathbf{A} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = D^{-1}\mathbf{A}C \begin{pmatrix} \mathfrak{x}_1 \\ \mathfrak{x}_2 \\ \vdots \\ \mathfrak{x}_n \end{pmatrix}.$$

Поскольку равенство справедливо для любого столбца координат, то на основании теоремы 3.2 имеем справедливость формулы (3.2). □

3.2 Канонический вид матрицы линейного отображения

Задача. Подобрать базисы пространств $\mathbb{V}$ и $\mathbb{W}$ так, чтобы матрица заданного линейного отображения $\mathscr{A}$ имела наиболее простой вид.

Воспользуемся понятием относительной линейной независимости из §5 главы 1. Найдем относительный базис $\mathbb{V}$ над $\mathcal{K}er(\mathscr{A})$, т.е. базис $\mathcal{K}er(\mathscr{A})$ дополним до базиса $\mathbb{V}$:

$$\begin{array}{lll} X_1, \dots, X_{\mathfrak{r}}, & & \leftarrow \text{ относит. базис } \mathbb{V} \text{ над } \mathcal{K}er(\mathscr{A}) \\ & X_{\mathfrak{r}+1}, \dots, X_n & \leftarrow \text{ базис } \mathcal{K}er(\mathscr{A}) \end{array} \tag{3.3}$$

Покажем, что система $\{\mathscr{A}(X_1), \dots, \mathscr{A}(X_{\mathfrak{r}})\} \subset \mathbb{W}$ является базисом $\mathcal{I}m(\mathscr{A})$. Рассмотрим произвольный $Y \in \mathcal{I}m(\mathscr{A})$, пусть $X = x_1X_1 + \cdots + x_nX_n$ его прообраз, тогда

$$Y = \mathscr{A}(X) = \sum_{j=1}^{\mathfrak{r}} x_j\mathscr{A}(X_j) + \sum_{j=\mathfrak{r}+1}^{n} x_j\mathscr{A}(X_j) = \sum_{j=1}^{\mathfrak{r}} x_j\mathscr{A}(X_j).$$

Следовательно, $\mathcal{Im}(\mathscr{A}) = \mathcal{L}(\mathscr{A}(X_1), \dots, \mathscr{A}(X_\mathfrak{r}))$. Покажем, что векторы **л.н.з.** Если

$$\alpha_1\mathscr{A}(X_1) + \dots + \alpha_\mathfrak{r}\mathscr{A}(X_\mathfrak{r}) = \mathbb{O}' \text{ то } \mathscr{A}(\alpha_1 X_1 + \dots + \alpha_\mathfrak{r} X_\mathfrak{r}) = \mathbb{O}'$$

и, значит, $\alpha_1 X_1 + \dots + \alpha_\mathfrak{r} X_\mathfrak{r} \in \mathcal{Ker}(\mathscr{A})$. Тогда, по определению относительного базиса, должно быть $\alpha_1 = 0, \dots, \alpha_\mathfrak{r} = 0$, и, значит, векторы $\mathscr{A}(X_1), \dots, \mathscr{A}(X_\mathfrak{r})$ **л.н.з.**

Составим базис $\mathbb{W}$ их дополнением:

$$\begin{array}{rl} \mathscr{A}(X_1), \dots, \mathscr{A}(X_\mathfrak{r}), \qquad\qquad & \leftarrow \text{ базис } \mathcal{Im}(\mathscr{A}) \\ Y_{\mathfrak{r}+1}, \dots, Y_m & \leftarrow \text{ относит. базис } \mathbb{W} \text{ над } \mathcal{Im}(\mathscr{A}) \end{array} \tag{3.4}$$

и найдем матрицу отображения $\mathscr{A}$ в базисах (3.3) и (3.4). Образы базисных векторов (3.3):

$$\begin{array}{rcl} \mathscr{A}(X_1) & = & 1 \cdot \mathscr{A}(X_1) + 0 \cdot \mathscr{A}(X_2) + \dots + 0 \cdot \mathscr{A}(X_\mathfrak{r}) + 0 \cdot Y_{\mathfrak{r}+1} + \dots + 0 \cdot Y_m, \\ \mathscr{A}(X_2) & = & 0 \cdot \mathscr{A}(X_1) + 1 \cdot \mathscr{A}(X_2) + \dots + 0 \cdot \mathscr{A}(X_\mathfrak{r}) + 0 \cdot Y_{\mathfrak{r}+1} + \dots + 0 \cdot Y_m, \\ \dots & & \dots \\ \mathscr{A}(X_\mathfrak{r}) & = & 0 \cdot \mathscr{A}(X_1) + 0 \cdot \mathscr{A}(X_2) + \dots + 1 \cdot \mathscr{A}(X_\mathfrak{r}) + 0 \cdot Y_{\mathfrak{r}+1} + \dots + 0 \cdot Y_m, \end{array}$$

а $\mathscr{A}(X_{\mathfrak{r}+1}) = \dots = \mathscr{A}(X_m) = \mathbb{O}'$ по определению $\mathcal{Ker}(\mathscr{A})$.

Теорема 3.5. *В базисах* (3.3) *и* (3.4) *матрица линейного отображения* $\mathscr{A}$ *имеет следующий* **канонический вид***:*

$$\mathbf{B} = \begin{array}{c} \mathfrak{r}\left\{ \vphantom{\begin{matrix} 1 \\ 1 \\ \ddots \\ 1 \end{matrix}} \right. \\ \vphantom{\mathbb{O}} \end{array} \!\! \begin{pmatrix} 1 & & & & \\ & 1 & & & \mathbb{O} \\ & & \ddots & & \\ & & & 1 & \\ & & & & \\ & \mathbb{O} & & & \mathbb{O} \end{pmatrix} = \begin{pmatrix} E_{\mathfrak{r}\times\mathfrak{r}} & \mathbb{O}_{\mathfrak{r}\times(n-\mathfrak{r})} \\ \mathbb{O}_{(m-\mathfrak{r})\times\mathfrak{r}} & \mathbb{O}_{(m-\mathfrak{r})\times(n-\mathfrak{r})} \end{pmatrix}$$

Здесь $\mathfrak{r} \stackrel{\text{def}}{=} \operatorname{rank}(\mathscr{A})$.

4 Линейный оператор

4.1 Основные свойства

Определение. Линейное отображение векторного пространства $\mathbb{V}$ в себя $\mathscr{A} : \mathbb{V} \longmapsto \mathbb{V}$ называется **линейным преобразованием** $\mathbb{V}$ или **оператором** на $\mathbb{V}$.

Характерное свойство оператора, выделяющее его из всего множества линейных отображений, это то, что $\mathcal{Im}(\mathscr{A}) \subset \mathbb{V}$.

Упражнение 4.1. *Доказать, что для оператора в* $\mathbb{R}^4$

$$\mathscr{A}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_3 \\ x_4 \\ 0 \\ 0 \end{pmatrix}$$

имеет место равенство $\mathcal{K}er(\mathscr{A}) = \mathcal{I}m(\mathscr{A})$.

Определение. Отображение $\mathscr{P} : \mathbb{V} \longmapsto \mathbb{V}$ называется **произведением** оператора $\mathscr{A}$ на оператор $\mathscr{B}$: $\mathscr{P} = \mathscr{A}\mathscr{B}$, если $\mathscr{P}(X) = \mathscr{A}(\mathscr{B}(X))$ для любого $X \in \mathbb{V}$.

Теорема 4.1. *Произведение операторов является оператором на* $\mathbb{V}$. *Операция произведения ассоциативна.*

Доказательство . Имеем на основании свойства линейности

$$\mathscr{P}(\alpha_1 X_1 + \alpha_2 X_2) = \mathscr{A}(\mathscr{B}(\alpha_1 X_1 + \alpha_2 X_2)) = \mathscr{A}(\alpha_1 \mathscr{B}(X_1) + \alpha_2 \mathscr{B}(X_2)) =$$
$$= \alpha_1 \mathscr{A}(\mathscr{B}(X_1)) + \alpha_2 \mathscr{A}(\mathscr{B}(X_2)) = \alpha_1 \mathscr{P}(X_1) + \alpha_2 \mathscr{P}(X_2).$$

Далее, для любого вектора X:

$$\mathscr{A}_1(\mathscr{A}_2\mathscr{A}_3(X)) = \mathscr{A}_1(\mathscr{A}_2(\mathscr{A}_3(X))) = \mathscr{A}_1\mathscr{A}_2(\mathscr{A}_3(X)) \ ,$$

откуда и следует ассоциативность. □

Определение. Говорят, что операторы $\mathscr{A}$ и $\mathscr{B}$ **коммутируют** если $\mathscr{A}\mathscr{B} = \mathscr{B}\mathscr{A}$.

Пример 4.1. *В пространстве полиномов* $\mathbb{P}_n$ *рассмотрим дифференциальный оператор*

$$\mathscr{A} \stackrel{\mathrm{def}}{=} x\frac{d}{d\,x} \times \Box - 1 \times \Box \ : \ p(x) \longmapsto xp'(x) - p(x) \ .$$

Этот оператор не коммутирует с обычным оператором дифференцирования $\mathscr{B} \stackrel{\mathrm{def}}{=} {}^{d}/_{d\,x}$:

$$\mathscr{A}(x^2) = x^2, \quad \mathscr{B}(\mathscr{A}(x^2)) = 2x, \quad \mathscr{B}(x^2) = 2x, \quad \mathscr{A}(\mathscr{B}(x^2)) = 0 \ .$$

Определение. Оператор $\mathscr{E} : X \longmapsto X$ (отображающий произвольный вектор $X \in \mathbb{V}$ в себя) называется **тождественным** на $\mathbb{V}$. Оператор $\mathscr{B}$ называется **обратным** оператору $\mathscr{A}$, если $\mathscr{B}\mathscr{A} = \mathscr{E}$. В этом случае оператор $\mathscr{A}$ называют обратимым и записывают: $\mathscr{B} = \mathscr{A}^{-1}$.

Не всякий оператор обратим.

Пример 4.2. *Для оператора проектирования на многообразие (пример* 1.3*) обратного не существует, т.к.* $\mathscr{A}(0,0,1) = (0,0,0)$ *и ни при каком выборе оператора* $\mathscr{B}$ *нельзя добиться выполнения равенства* $\mathscr{B}(0,0,0) = (0,0,1)$.

Упражнение 4.2. *Показать, что обратным для оператора*

$$\frac{1}{x}\int_0^x \ : p(x) \longmapsto \frac{1}{x}\int_0^x p(t)d\,t \ ,$$

на $\mathbb{P}_n$ *является оператор*

$$\frac{d}{d\,x}\,(x \times \Box) \ : p(x) \longmapsto (xp(x))' \ .$$

Теорема 4.2. *Оператор* $\mathscr{A}$ *обратим тогда и только тогда, когда когда он невырожден:* $\operatorname{dfc}\mathscr{A} = 0$. *В этом случае* $\mathscr{A}^{-1}$ *единствен и коммутирует с* $\mathscr{A}$.

Доказательство . **Н.** Пусть $\operatorname{dfc}\mathscr{A} \neq 0$. Это значит, что существует $X_1 \neq \mathbb{O}$ такой, что $\mathscr{A}(X_1) = \mathbb{O}$. Но тогда, на основании пункта а) теоремы 1.1, при любом линейном операторе $\mathscr{B}$: $\mathscr{B}(\mathscr{A}(X_1)) = \mathbb{O} \neq X_1$, т.е. оператор $\mathscr{A}$ необратим.

Д. Пусть $\operatorname{dfc}\mathscr{A} = 0$, тогда, на основании теоремы 2.3, $\dim \mathbb{V} = \dim(\mathcal{I}m\,\mathscr{A})$, т.е. $\mathcal{I}m(\mathscr{A}) = \mathbb{V}$. Следовательно, для любого $Y \in \mathbb{V}$ существует прообраз и этот прообраз будет единственным (почему?). Для оператора $\mathscr{A} : X \longmapsto Y$ существует обратное отображение — ставящее в соответствие каждому образу его прообраз: $\mathscr{B} : Y \longmapsto X$. Докажем линейность этого отображения. Действительно, для любых Y_1, Y_2 из $\mathbb{V}$ существуют прообразы и, на основании линейности $\mathscr{A}$:

$$\text{если } Y_1 = \mathscr{A}(X_1), Y_2 = \mathscr{A}(X_2), \text{ то } \alpha_1 Y_1 + \alpha_2 Y_2 = \mathscr{A}(\alpha_1 X_1 + \alpha_2 X_2) \ .$$

Тогда обратное отображение должно действовать так:

$$\mathscr{B}(Y_1) = X_1, \mathscr{B}(Y_2) = X_2, \ \mathscr{B}(\alpha_1 Y_1 + \alpha_2 Y_2) = \alpha_1 X_1 + \alpha_2 X_2 \ ,$$

откуда и вытекает его линейность: $\mathscr{B}(\alpha_1 Y_1 + \alpha_2 Y_2) = \alpha_1 \mathscr{B}(Y_1) + \alpha_2 \mathscr{B}(Y_2)$.

Для доказательства оставшейся части теоремы, предположим, что, наряду с только что найденным $\mathscr{B}$ существует еще один оператор $\mathscr{B}_1$ со свойством $\mathscr{A}\mathscr{B}_1 = \mathscr{E}$. Тогда домножив слева на $\mathscr{B}$ мы должны получить

$$\mathscr{B}(\mathscr{A}\mathscr{B}_1) = \mathscr{B}\mathscr{E} \quad \Longleftrightarrow \quad \mathscr{E}\mathscr{B}_1 = \mathscr{B} \quad \Longleftrightarrow \quad \mathscr{B}_1 = \mathscr{B} \ ,$$

т.е. "левый обратный" оператор должен совпадать с "правым": $\mathscr{A}^{-1}\mathscr{A} = \mathscr{A}\mathscr{A}^{-1}$. Если теперь предположить, что существует $\mathscr{B}_2$ такой, что $\mathscr{B}_2\mathscr{A} = \mathscr{E}$, то домножив это равенство на $\mathscr{A}^{-1}$ справа получим: $\mathscr{B}_2 = \mathscr{A}^{-1}$, т.е. единственность обратного оператора. □

Определение. K-я степень оператора $\mathscr{A}$ определяется рекурсивной формулой

$$\mathscr{A}^K = \mathscr{A}(\mathscr{A}^{K-1}), \ (K \in \mathbb{N}, \ \text{а если, вдобавок, } \exists \mathscr{A}^{-1}, \ \text{то } K \in \mathbb{Z}) \ .$$

Полагают также $\mathscr{A}^0 \overset{\text{def}}{=} \mathscr{E}$ для любого $\mathscr{A} \neq \mathcal{O}$.

Теорема 4.3. *Степени оператора* $\mathscr{A}$ *коммутируют:*

$$\mathscr{A}^K \mathscr{A}^L = \mathscr{A}^L \mathscr{A}^K = \mathscr{A}^{K+L} \ .$$

Доказательство . Следствие теоремы 4.1: $\mathscr{A}^2\mathscr{A} = \mathscr{A}\mathscr{A}(\mathscr{A}) = \mathscr{A}(\mathscr{A}\mathscr{A}) = \mathscr{A}\mathscr{A}^2$. Для общего случая доказывается по индукции. □

Пример 4.3. *В пространстве полиномов $\mathbb{P}_n$ K-й степенью оператора дифференцирования (пример 1.4) будет оператор нахождения $K-$й производной:*

$$\left(\frac{d}{d\,x}\right)^K \stackrel{\text{def}}{=} \frac{d^K}{d\,x^K}\ .$$

Заметим, что при $K > n$ этот оператор будет нулевым.

Определение. Пусть задан произвольный полином $g(x) = b_0x^m + b_1x^{m-1} + \dots + b_m$ из $\mathbb{R}[x]$ или $\mathbb{C}[x]$. Выражение

$$g(\mathscr{A}) \stackrel{\text{def}}{=} b_0\mathscr{A}^m + b_1\mathscr{A}^{m-1} + \dots + b_m\mathscr{E}$$

будем называть **операторным полиномом**.

Упражнение 4.3. *Доказать, что операторные полиномы коммутируют: $g_1(\mathscr{A})g_2(\mathscr{A}) = g_2(\mathscr{A})g_1(\mathscr{A})$.*

Упражнение 4.4. *Доказать, что для любого $\mathscr{A} \in \mathcal{H}om(\mathbb{V},\mathbb{V})$ всегда найдется полином $g(x)$, $\deg g \le n^2 + 1$ такой, что $g(\mathscr{A}) = \mathcal{O}$.*

4.2 Матрица оператора

Рассмотрим оператор $\mathscr{A}$ на $\mathbb{V}$ и пусть $\{X_1,\dots,X_n\}$ — базис $\mathbb{V}$. Являясь частным случаем линейного отображения, оператор должен обладать и соответствующей матрицей. Существенной особенностью, отличающей наш случай от рассмотренного в §3, является невозможность *произвола* при выборе базиса для $\mathcal{I}m\,\mathscr{A}$. Поскольку $\mathcal{I}m\,\mathscr{A}$ является подпространством $\mathbb{V}$, то было бы слишком большой роскошью иметь два разных базиса для одного и того же пространства.

Найдем координаты векторов $\mathscr{A}(X_1),\dots,\mathscr{A}(X_n)$ в базисе $\{X_1,\dots,X_n\}$:

$$\left\{\begin{array}{lcl} \mathscr{A}(X_1) & = & \alpha_{11}X_1 + \alpha_{21}X_2 + \dots + \alpha_{n1}X_n, \\ \mathscr{A}(X_2) & = & \alpha_{12}X_1 + \alpha_{22}X_2 + \dots + \alpha_{n2}X_n, \\ \dots & & \dots, \\ \mathscr{A}(X_n) & = & \alpha_{1n}X_1 + \alpha_{2n}X_2 + \dots + \alpha_{nn}X_n. \end{array}\right. \tag{4.1}$$

Определение. Матрица

$$\mathbf{A} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \dots & & & \dots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{pmatrix}_{n\times n},$$

в которой по столбцам стоят координаты образов базисных векторов, называется **матрицей оператора** $\mathscr{A}$ в выбранном базисе.

Пример 4.4. *Известны образы базисных векторов $\mathbb{R}^3$ под действием оператора $\mathscr{A}$:*

$$\mathscr{A}\begin{pmatrix} 5 \\ 3 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix},\ \mathscr{A}\begin{pmatrix} 1 \\ -3 \\ -2 \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \\ 0 \end{pmatrix},\ \mathscr{A}\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ -3 \\ 0 \end{pmatrix}\ .$$

Найти матрицу этого оператора в исходном базисе.

Решение. Компоненты матрицы $\mathbf{A}$ ищутся по формулам (4.1), которые можно записать в матричном виде:

$$[X_1, \dots, X_n]\,\mathbf{A} = [\mathscr{A}(X_1), \dots, \mathscr{A}(X_n)] \ .$$

Откуда

$$\mathbf{A} = [X_1, \dots, X_n]^{-1}\,[\mathscr{A}(X_1), \dots, \mathscr{A}(X_n)] \ ,$$

и для нашего примера эта формула дает

$$\begin{aligned}\mathbf{A} &= \begin{pmatrix} 5 & 1 & 1 \\ 3 & -3 & 2 \\ 1 & -2 & 1 \end{pmatrix}^{-1} \begin{pmatrix} -2 & -1 & -2 \\ 1 & 3 & -3 \\ 0 & 0 & 0 \end{pmatrix} = \\ &= \begin{pmatrix} 1 & -3 & 5 \\ -1 & 4 & -7 \\ -3 & 11 & -18 \end{pmatrix} \begin{pmatrix} -2 & -1 & -2 \\ 1 & 3 & -3 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} -5 & -10 & 7 \\ 6 & 13 & -10 \\ 17 & 36 & -27 \end{pmatrix}.\end{aligned}$$

Чему равны $\operatorname{rank}\mathscr{A}$ и $\operatorname{dfc}\mathscr{A}$? $\triangle$

Как следствие теоремы 3.1 имеем: координаты произвольного вектора $X = x_1X_1 + \dots + x_nX_n$ и его образа $Y = \mathscr{A}(X) = y_1X_1 + \dots + y_nX_n$ связаны формулой

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \mathbf{A} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \ . \tag{4.2}$$

Для матриц операторов справедливы аналоги теорем 3.2 и 3.3, но, помимо того, и следующий результат, касающийся произведений.

Теорема 4.4. *В любом базисе пространства*

а) *матрица тождественного оператора $\mathscr{E}$ совпадает с единичной $\mathbf{E}$;*

б) *матрица оператора $\mathscr{A}^{-1}$ совпадает с $\mathbf{A}^{-1}$;*

в) *матрица произведения операторов $\mathscr{A}_1\mathscr{A}_2$ совпадает с $\mathbf{A}_1\mathbf{A}_2$;*

г) *матрица операторного полинома $g(\mathscr{A})$ совпадает с матричным полиномом $g(\mathbf{A})$;*

д) *коммутирующим операторам соответствуют коммутирующие матрицы.*

Таким образом, свойства операторов абсолютно сходны свойствам квадратных матриц, что и позволяет свести исследование первых к исследованию последних — если фиксирован базис пространства.

Как изменяется матрица оператора при переходе к новому базису?

Теорема 4.5. *Если C — матрица перехода от старого базиса к новому, то матрицы $\mathbf{A}$ и $\mathbf{B}$ оператора в старом и новом базисах связаны формулой:*

$$\mathbf{B} = C^{-1} \cdot \mathbf{A} \cdot C \ . \tag{4.3}$$

Доказательство . Пусть $\{\mathfrak{X}_1, \ldots, \mathfrak{X}_n\}$ — новый базис и

$$X = x_1X_1 + \cdots + x_nX_n = \mathfrak{x}_1\mathfrak{X}_1 + \cdots + \mathfrak{x}_n\mathfrak{X}_n,$$

$$Y = \mathscr{A}(X) = y_1X_1 + \cdots + y_nX_n = \mathfrak{y}_1\mathfrak{X}_1 + \cdots + \mathfrak{y}_n\mathfrak{X}_n.$$

На основании результатов §6 главы 1,

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = C \begin{pmatrix} \mathfrak{x}_1 \\ \vdots \\ \mathfrak{x}_n \end{pmatrix}, \qquad \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = C \begin{pmatrix} \mathfrak{y}_1 \\ \vdots \\ \mathfrak{y}_n \end{pmatrix}.$$

Получаем цепочку равенств:

$$\mathbf{B} \begin{pmatrix} \mathfrak{x}_1 \\ \vdots \\ \mathfrak{x}_n \end{pmatrix} = \begin{pmatrix} \mathfrak{y}_1 \\ \vdots \\ \mathfrak{y}_n \end{pmatrix} = C^{-1} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = C^{-1}\mathbf{A} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = C^{-1}\mathbf{A}C \begin{pmatrix} \mathfrak{x}_1 \\ \vdots \\ \mathfrak{x}_n \end{pmatrix}.$$

Равенство имеет место для любых столбцов $[\mathfrak{x}_1, \ldots, \mathfrak{x}_n]^\top$, следовательно и для столбцов

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Объединяя полученные n равенств в одно матричное, получим $\mathbf{B} \cdot E = C^{-1} \cdot \mathbf{A} \cdot C \cdot E$, откуда и следует (4.3). □

Пример 4.5. *Оператор* $\mathscr{A}$ *в базисе*

$$\begin{pmatrix} 8 \\ -6 \\ 7 \end{pmatrix}, \begin{pmatrix} -16 \\ 7 \\ -13 \end{pmatrix}, \begin{pmatrix} 9 \\ -3 \\ 7 \end{pmatrix} \quad \begin{matrix} \textit{имеет} \\ \textit{матрицу} \end{matrix} \quad \begin{pmatrix} 1 & -18 & 15 \\ -1 & -22 & 20 \\ 1 & -25 & 22 \end{pmatrix}.$$

Найти его матрицу в базисе

$$\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}.$$

Решение. Матрица C перехода от старого базиса к новому находится по формуле (6.2) главы 1:

$$\begin{aligned} C &= \begin{pmatrix} 8 & -16 & 9 \\ -6 & 7 & -3 \\ 7 & -13 & 7 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 3 & 2 \\ -2 & -1 & 1 \\ 1 & 2 & 2 \end{pmatrix} = \\ &= \begin{pmatrix} 2 & -1 & -3 \\ 21/5 & -7/5 & -6 \\ 29/5 & -8/5 & -8 \end{pmatrix} \begin{pmatrix} 1 & 3 & 2 \\ -2 & -1 & 1 \\ 1 & 2 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & -3 \\ 1 & 2 & -5 \\ 1 & 3 & -6 \end{pmatrix} \end{aligned}$$

$$
\begin{aligned}
\mathbf{B} = C^{-1}\mathbf{A}C \quad &= \begin{pmatrix} 3 & -3 & 1 \\ 1 & -3 & 2 \\ 1 & -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & -18 & 15 \\ -1 & -22 & 20 \\ 1 & -25 & 22 \end{pmatrix} \begin{pmatrix} 1 & 1 & -3 \\ 1 & 2 & -5 \\ 1 & 3 & -6 \end{pmatrix} = \\
&= \begin{pmatrix} 7 & -13 & 7 \\ 6 & -2 & -1 \\ 4 & 1 & -3 \end{pmatrix} \begin{pmatrix} 1 & 1 & -3 \\ 1 & 2 & -5 \\ 1 & 3 & -6 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 2 \\ 3 & -1 & -2 \\ 2 & -3 & 1 \end{pmatrix}.
\end{aligned}
$$

$\triangle$

Определение. Матрицы $\mathbf{A}$ и $\mathbf{B}$, связанные соотношением (4.3) (при какой-то неособенной матрице C) называются **подобными**: $\mathbf{A} \doteq \mathbf{B}$.

Упражнение 4.5. *Доказать, что отношение подобия есть отношение эквивалентности, и если* $\mathbf{A} \doteq \mathbf{B}$ *то* $g(\mathbf{A}) \doteq g(\mathbf{B})$ *при любом полиноме* $g(x)$.

5 Инвариантные подпространства оператора

Задача. Для данного оператора $\mathscr{A}$ выбрать такой базис пространства $\mathbb{V}$, чтобы матрица оператора имела наиболее простой вид. Хотя эта задача похожа на решенную в §3.2, но теперь мы ограничены в маневре: если для линейного *отображения* закон изменения его матрицы $\mathbf{B} = D^{-1} \cdot \mathbf{A} \cdot C$ допускал произвол в выборе **двух** матриц D и C, то для *оператора* $\mathbf{B} = C^{-1} \cdot \mathbf{A} \cdot C$, и мы имеем возможность выбора лишь **одной** матрицы.

5.1 Инвариантное подпространство

Исследуем действие оператора $\mathscr{A}$ на произвольное подпространство $\mathbb{V}_1 \subset \mathbb{V}$:

$$\mathscr{A}(\mathbb{V}_1) = \{ Y \in \mathbb{V} \mid Y = \mathscr{A}(X),\ X \in \mathbb{V}_1 \} \quad .$$

Вообще говоря, множества $\mathbb{V}_1$ и $\mathscr{A}(\mathbb{V}_1)$ будут различными, и $\exists X_1 \in \mathbb{V}_1$ такой, что $\mathscr{A}(X_1) \notin \mathbb{V}_1$.

Определение. Подпространство $\mathbb{V}_1$ называется **инвариантным** подпространством оператора $\mathscr{A}$, если оно отображается в себя: $\mathscr{A}(\mathbb{V}_1) \subset \mathbb{V}_1$.

Определение. $\mathbb{V}_1 = \{\mathbb{O}\}$ и $\mathbb{V}_1 = \mathbb{V}$ — **тривиальные** инвариантные подпространства.

Нас будут интересовать нетривиальные инвариантные подпространства.

Пример 5.1. *Оператор*

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \longmapsto \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

задает в пространстве поворот вокруг оси $\mathbb{O}z$ *на угол* $+\pi/4$. *Инвариантными подпространствами будут*

а) $\mathbb{V}_1 = \{(0,0,z)^{\top} \mid z \in \mathbb{R}\}$, $\dim \mathbb{V}_1 = 1$ *(ось вращения) и*

б) $\mathbb{V}_2 = \{(x,y,0)^{\top} \mid x \in \mathbb{R}, y \in \mathbb{R}\}$, $\dim \mathbb{V}_2 = 2$ *(плоскость, перпендикулярная оси вращения).*

Пример 5.2. *Оператор*

$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} \lambda_1 x \\ \lambda_2 y \end{pmatrix}$$

задает на плоскости "растяжение": x-компонента увеличивается в λ_1 раз, а y-компонента — в λ_2 раз. При любой комбинации коэффициентов растяжения координатные оси будут инвариантными подпространствами. Но в частном случае $\lambda_1 = \lambda_2$ инвариантной будет также любая прямая, проходящая через начало координат.

Пример 5.3. *Оператор в $\mathbb{R}^n$ задан блочной матрицей*

$$X \longmapsto \begin{pmatrix} \mathbf{A}_1 & * \\ \mathbb{O} & \mathbf{A}_2 \end{pmatrix} X$$

где $\mathbf{A}_1$ — $n_1 \times n_1$-матрица, $\mathbf{A}_2$ — $(n-n_1)\times(n-n_1)$-матрица. Множество столбцов

$$\mathbb{V}_1 = \left\{ X = [x_1, \dots, x_{n_1}, 0, \dots, 0]^\top \middle| x_1 \in \mathbb{R}, \dots, x_{n_1} \in \mathbb{R} \right\}$$

образует инвариантное подпространство, $\dim \mathbb{V}_1 = n_1$. *Если же, вдобавок, матрица, обозначенная $*$ — нулевая, то вторым инвариантным подпространством будет*

$$\mathbb{V}_2 = \left\{ X = [0, \dots, 0, x_{n_1+1}, \dots, x_n]^\top \middle| x_{n_1+1} \in \mathbb{R}, \dots, x_n \in \mathbb{R} \right\} \ .$$

Упражнение 5.1. *В пространстве полиномов $\mathbb{P}_n$ найти инвариантное подпространство размерности $k \in \{1, \dots, n\}$ для оператора дифференцирования.*

Теорема 5.1. *$\mathcal{K}er\, \mathscr{A}$ и $\mathcal{I}m\, \mathscr{A}$ — инвариантные подпространства оператора $\mathscr{A}$.*

Доказательство . На основании определения $\mathcal{I}m\, \mathscr{A}$, имеем

$$\mathcal{I}m\, \mathscr{A} = \mathscr{A}(\mathbb{V}),\ \mathcal{I}m\, \mathscr{A} \subset \mathbb{V} \Longrightarrow \mathscr{A}(\mathcal{I}m\, \mathscr{A}) \subset \mathscr{A}(\mathbb{V}) = \mathcal{I}m\, \mathscr{A}.$$

Далее: $\mathscr{A}(\mathcal{K}er\, \mathscr{A}) = \{\mathbb{O}\} \subset \mathcal{K}er\, \mathscr{A}$. □

Теорема 5.2. *Если пространство $\mathbb{V}$ раскладывается в прямую сумму подпространств, инвариантных относительно оператора $\mathscr{A}$, то существует базис пространства, в котором матрица оператора будет блочно-диагональной.*

Доказательство . Пусть $\mathbb{V} = \mathbb{V}_1 \oplus \mathbb{V}_2$. Искомый базис $\mathbb{V}$ составим объединением базисов подпространств. Если $\{X_1, \dots, X_{d_1}\}$ — базис $\mathbb{V}_1$, а

$\{X_{d_1+1}, \ldots, X_n\}$ — базис $\mathbb{V}_2$, то, ввиду инвариантности этих подпространств, имеем $\mathscr{A}(X_1) \in \mathbb{V}_1, \ldots, \mathscr{A}(X_{d_1}) \in \mathbb{V}_1$:

$$\begin{array}{lcl} \mathscr{A}(X_1) &=& \alpha_{11}X_1 + \ldots + \alpha_{1d_1}X_{d_1} + 0 \cdot X_{d_1+1} + \ldots + 0 \cdot X_n, \\ & & \ldots \qquad\qquad\qquad \ldots \\ \mathscr{A}(X_{d_1}) &=& \alpha_{d_1 1}X_1 + \ldots + \alpha_{d_1 d_1}X_{d_1} + 0 \cdot X_{d_1+1} + \ldots + 0 \cdot X_n, \end{array}$$

а также $\mathscr{A}(X_{d_1+1}) \in \mathbb{V}_2, \ldots, \mathscr{A}(X_n) \in \mathbb{V}_2$:

$$\begin{array}{lcl} \mathscr{A}(X_{d_1+1}) &=& 0 \cdot X_1 + \ldots + 0 \cdot X_{d_1} + \alpha_{d_1+1,d_1+1}X_{d_1+1} + \cdots + \alpha_{d_1+1,n}X_n, \\ & & \ldots \qquad\qquad \ldots \\ \mathscr{A}(X_n) &=& 0 \cdot X_1 + \ldots + 0 \cdot X_{d_1} + \alpha_{n,d_1+1}X_{d_1+1} + \cdots + \alpha_{nn}X_n. \end{array}$$

Следовательно, в указанном базисе матрица оператора имеет структуру

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_1 & \mathbb{O} \\ \mathbb{O} & \mathbf{A}_2 \end{pmatrix}$$

и размеры блоков равны размерностям инвариантных подпространств. □

Теорема обобщается очевидным образом на произвольное число слагаемых подпространств: $\mathbb{V} = \mathbb{V}_1 \oplus \mathbb{V}_2 \oplus \cdots \oplus \mathbb{V}_k$. Если при этом $\dim \mathbb{V}_1 = \cdots = \dim \mathbb{V}_k = 1$, то матрица $\mathbf{A}$ становится диагональной!

5.2 Собственные числа и собственные векторы

Рассмотрим оператор над комплексным пространством $\mathbb{V}$.

Определение. Вектор $X \in \mathbb{V}$ называется **собственным вектором** оператора $\mathscr{A}$, если

$$\textbf{а)}\, X \neq \mathbb{O}, \quad \text{и} \quad \textbf{б)}\, \exists\, \lambda \in \mathbb{C} \text{ такое, что } \mathscr{A}(X) = \lambda X \ .$$

В этом случае число λ называется **собственным** (или **характеристическим**) **числом** оператора, соответствующим данному собственному вектору; обратно, говорят, что вектор X **принадлежит собственному числу** λ.

Теорема 5.3. *Любой собственный вектор оператора порождает его одномерное инвариантное подпространство, и обратно: любой ненулевой вектор одномерного инвариантного подпространства оператора является собственным вектором.*

Доказательство . Пусть $\mathscr{A}(X) = \lambda X$ при некоторых $\lambda \in \mathbb{C}$ и $X \neq \mathbb{O}$. Тогда $\mathbb{V}_1 = \mathcal{L}(X)$ — инвариантное подпространство для $\mathscr{A}$:

$$\mathbb{V}_1 = \mathcal{L}(X) \stackrel{\text{def}}{=} \{\alpha X \mid \alpha \in \mathbb{C}\} \Longrightarrow \mathscr{A}(\mathbb{V}_1) = \{\mathscr{A}(\alpha X) \mid \alpha \in \mathbb{C}\} =$$

$$= \{\alpha \mathscr{A}(X) \mid \alpha \in \mathbb{C}\} = \{\alpha\lambda X \mid \alpha \in \mathbb{C}\} = \{\alpha_1 X \mid \alpha_1 \in \mathbb{C}\} = \mathcal{L}(X) \ .$$

Обратно, если $\dim \mathbb{V}_1 = 1$, то любой ненулевой вектор $X \in \mathbb{V}_1$ можно взять в качестве базисного вектора подпространства. Если $\mathscr{A}(\mathbb{V}_1) = \mathbb{V}_1$, то X отображается в то же подпространство:

$$\mathscr{A}(X) \in \mathbb{V}_1 \Longrightarrow \exists \lambda \in \mathbb{C} \ : \mathscr{A}(X) = \lambda X \ .$$

Значит X — собственный вектор. □

Геометрический смысл вещественных собственных чисел и векторов проясняет пример 5.2. Собственный вектор задает направление, на котором действие оператора сводится к растяжению, тогда коэффициент растяжения и будет собственным числом.

Теорема 5.4. *В комплексном пространстве любой оператор имеет по крайней мере один собственный вектор.*

Доказательство . Пусть $\{X_1, \ldots, X_n\}$ — произвольный базис $\mathbb{V}$ и $\mathbf{A}$ — матрица оператора $\mathscr{A}$ в этом базисе. Тогда для того чтобы вектор $X = x_1X_1 + \cdots + x_nX_n \neq \mathbb{O}$ был собственным, принадлежащим собственному числу λ, **Н.** и **Д.** чтобы выполнялось равенство

$$\mathbf{A}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \Longleftrightarrow \quad \begin{pmatrix} \alpha_{11}-\lambda & \alpha_{12} & \ldots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22}-\lambda & \ldots & \alpha_{2n} \\ \ldots & & & \ldots \\ \alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nn}-\lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \mathbb{O}_{n\times 1} \tag{5.1}$$

Покажем, что существуют комплексные числа λ и не все нулевые $x_1, \ldots, x_n$, удовлетворяющие системе (5.1). Необходимым условием существования нетривиального решения у однородной системы (5.1) является равенство нулю ее определителя:

$$\det(\mathbf{A} - \lambda E) = \begin{vmatrix} \alpha_{11}-\lambda & \alpha_{12} & \ldots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22}-\lambda & \ldots & \alpha_{2n} \\ \ldots & & & \ldots \\ \alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nn}-\lambda \end{vmatrix} = 0 \ . \tag{5.2}$$

Этот определитель является полиномом степени n по λ. По основной теореме высшей алгебры этот полином имеет по крайней мере один комплексный корень $\lambda = \lambda_1$. Подставив его в (5.1), получаем однородную систему уравнений с нулевым определителем. У такой системы всегда существует нетривиальное решение $[x_1^*, \ldots, x_n^*]^\top$, но тогда вектор $\mathfrak{X}_1 \stackrel{\text{def}}{=} x_1^*X_1 + \cdots + x_n^*X_n$ будет собственным вектором оператора $\mathscr{A}$, принадлежащим λ_1. $\square$

Определение. Уравнение (5.2) называется **характеристическим** или **вековым** уравнением, а полином в левой его части — **характеристическим** полиномом матрицы $\mathbf{A}$. Набор всех корней хар.полинома (с учетом кратностей) называется **спектром** матрицы.

Следствие 1. *Любой корень хар.полинома является собственным числом оператора* $\mathscr{A}$ *и обратно: любое собственное число оператора* $\mathscr{A}$ *является корнем хар.полинома.*

Пример 5.4. *Найти спектр и собственные векторы матрицы*

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 2 & 3 \\ -1 & 0 & 4 & 7 \\ -2 & -4 & 0 & 2 \\ -3 & -7 & -2 & 0 \end{pmatrix} .$$

Решение. Хар.полином $\det(\mathbf{A}-\lambda E)=\lambda^4+83\lambda^2$. Спектр матрицы:

$$\{0,0,\boldsymbol{i}\sqrt{83},-\boldsymbol{i}\sqrt{83}\}.$$

Для нахождения собственных векторов подставим каждое из этих значений в систему (5.1):

$$(\mathbf{A}-0\cdot E)X=\mathbb{O} \quad \Longrightarrow \textbf{ф.с.р.}=\left\{\mathfrak{X}_1=\begin{pmatrix}4\\-2\\1\\0\end{pmatrix},\ \mathfrak{X}_2=\begin{pmatrix}7\\-3\\0\\1\end{pmatrix}\right\}.$$

Любой вектор вида $\alpha_1\mathfrak{X}_1+\alpha_2\mathfrak{X}_2$ будет собственным, принадлежащим $\lambda=0$.

$$(\mathbf{A}-\boldsymbol{i}\sqrt{83}E)X=\mathbb{O} \qquad\qquad (\mathbf{A}+\boldsymbol{i}\sqrt{83}E)X=\mathbb{O}$$

$$\Downarrow \qquad\qquad\qquad\qquad \Downarrow$$

$$\mathfrak{X}_3=\begin{pmatrix}1-\boldsymbol{i}\sqrt{83}\\8-2\boldsymbol{i}\sqrt{83}\\12\\17+\boldsymbol{i}\sqrt{83}\end{pmatrix} \qquad \mathfrak{X}_4=\begin{pmatrix}1+\boldsymbol{i}\sqrt{83}\\8+2\boldsymbol{i}\sqrt{83}\\12\\17-\boldsymbol{i}\sqrt{83}\end{pmatrix}.$$

$\triangle$

Теорема 5.5. *Хар.полиномы подобных матриц одинаковы.*

Доказательство . $\mathbf{A}\doteq\mathbf{B} \iff \exists$ неособенная матрица C, такая что $\mathbf{B}=C^{-1}\mathbf{A}C$. Имеем:

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

$$=\det(C^{-1}\mathbf{A}C-\lambda C^{-1}EC)=\det C^{-1}(\mathbf{A}-\lambda E)C=\det(\mathbf{A}-\lambda E)\ .$$

$\square$

Иначе говоря, для данного оператора $\mathscr{A}$ хар.полином его матрицы не зависит от выбора базиса пространства. Поэтому можно говорить о хар.полиноме *оператора* $\mathscr{A}$.

Упражнение 5.2. *Доказать, что хар.полином матрицы не изменится при ее транспонировании.*

5.3 Диагонализуемость матрицы оператора

Теорема 5.6. *Собственные векторы оператора, принадлежащие различным собственным числам, линейно независимы.*

Доказательство . Пусть $\lambda_1, \dots, \lambda_k$ — различные собственные числа оператора $\mathscr{A}$, а $\mathfrak{X}_1, \dots, \mathfrak{X}_k$ — принадлежащие им собственные векторы: $\mathscr{A}(\mathfrak{X}_j) = \lambda_j \mathfrak{X}_j$. Докажем теорему индукцией по k. Для $k = 1$ утверждение очевидно. Пусть оно верно для $(k-1)$-го векторов, но неверно для k векторов:

$$\alpha_1 \mathfrak{X}_1 + \dots + \alpha_{k-1} \mathfrak{X}_{k-1} + \alpha_k \mathfrak{X}_k = \mathbb{O} \tag{5.3}$$

при каком-то из коэффициентов отличном от нуля. Пусть $\alpha_1 \neq 0$.

К обеим частям равенства (5.3) применим оператор $\mathscr{A}$. Получим

$$\mathscr{A}(\alpha_1 \mathfrak{X}_1 + \dots + \alpha_{k-1} \mathfrak{X}_{k-1} + \alpha_k \mathfrak{X}_k) = \mathbb{O} \implies$$

$$\implies \alpha_1 \lambda_1 \mathfrak{X}_1 + \dots + \alpha_{k-1} \lambda_{k-1} \mathfrak{X}_{k-1} + \alpha_k \lambda_k \mathfrak{X}_k = \mathbb{O} \ .$$

Домножим равенство (5.3) на λ_k и вычтем из последнего:

$$\alpha_1(\lambda_1 - \lambda_k)\mathfrak{X}_1 + \dots + \alpha_{k-1}(\lambda_{k-1} - \lambda_k)\mathfrak{X}_{k-1} = \mathbb{O}.$$

Здесь $\alpha_1(\lambda_1 - \lambda_k) \neq 0$ т.к. $\lambda_1 \neq \lambda_k$. Векторы $\mathfrak{X}_1, \dots, \mathfrak{X}_{k-1}$ получились линейно зависимыми, что противоречит индукционному предположению. $\square$

Теорема 5.7. *Если оператор имеет $n = \dim \mathbb{V}$ линейно независимых собственных векторов, то в базисе ими образуемом матрица оператора диагональна. Обратно: если матрица оператора в некотором базисе диагональна, то каждый вектор этого базиса является собственным для оператора.*

Доказательство . Если

$$\mathscr{A}(\mathfrak{X}_1) = \lambda_1 \mathfrak{X}_1, \dots, \mathscr{A}(\mathfrak{X}_n) = \lambda_n \mathfrak{X}_n \tag{5.4}$$

и система $\{\mathfrak{X}_1, \dots, \mathfrak{X}_n\}$ **л.н.з**, то взяв ее в качестве базиса пространства $\mathbb{V}$ получим соответствующую матрицу оператора $\mathscr{A}$ в виде (см. §4.2):

$$\mathbf{A}_{diag} = \begin{pmatrix} \lambda_1 & & \mathbb{O} \\ & \ddots & \\ \mathbb{O} & & \lambda_n \end{pmatrix}. \tag{5.5}$$

Обратно, если матрица оператора в некотором базисе $\{\mathfrak{X}_1, \dots, \mathfrak{X}_n\}$ имеет вид (5.5), то это означает, например, что $\mathscr{A}(\mathfrak{X}_1) = \lambda_1 \mathfrak{X}_1$, т.е. $\mathfrak{X}_1$ — собственный вектор, принадлежащий λ_1. Аналогично доказывается и для оставшихся векторов. $\square$

Определение. Базис линейного пространства, состоящий из собственных векторов оператора $\mathscr{A}$, называется **каноническим**.

Следствие 1 (матричный аналог теоремы). *Пусть $\mathbf{A}$ — матрица оператора $\mathscr{A}$ в заданном базисе. Неособенная матрица C, удовлетворяющая равенству*

$$C^{-1}\mathbf{A}C = \mathbf{A}_{diag}$$

существует тогда и только тогда, когда существует базис пространства $\mathbb{C}^n$, состоящий из собственных векторов. Тогда матрица C является матрицей перехода от заданного базиса к каноническому базису, а на диагонали $\mathbf{A}_{diag}$ стоят собственные числа матрицы $\mathbf{A}$.

Определение. При выполнении условия предыдущего следствия говорят, что матрица **A диагонализуема** (или **приводится к диагональной форме**).

Теорема 5.6 позволяет сформулировать достаточное условие диагонализуемости.

Теорема 5.8. *Если хар.полином оператора не имеет кратных корней, то матрица оператора диагонализуема.*

Условие теоремы проверяется чисто алгебраически: вычислением дискриминанта[14] хар.полинома: $\mathcal{D}(\det(\mathbf{A} - \lambda E)) \neq 0$. Это условие не является необходимым, как показывает пример 5.4. Матрица

$$C = \left(\begin{array}{cccc} 4 & 7 & (1 - \boldsymbol{i}\sqrt{83}) & (1 + \boldsymbol{i}\sqrt{83}) \\ -2 & -3 & (8 - 2\boldsymbol{i}\sqrt{83}) & (8 + 2\boldsymbol{i}\sqrt{83}) \\ 1 & 0 & 12 & 12 \\ 0 & 1 & 17 + \boldsymbol{i}\sqrt{83} & 17 - \boldsymbol{i}\sqrt{83} \end{array} \right) \quad \begin{array}{c} \text{приводит} \\ \text{матрицу } \mathbf{A} \\ \text{к виду} \end{array} \quad \left(\begin{array}{cccc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \boldsymbol{i}\sqrt{83} & 0 \\ 0 & 0 & 0 & -\boldsymbol{i}\sqrt{83} \end{array} \right)$$

Случай существования кратных корней у хар.полинома является “критическим”: имеются примеры как диагонализуемых (см. выше), так и недиагонализуемых матриц. Так, для матриц

$$\mathbf{A} = \left(\begin{array}{rr} 0 & 1 \\ -1 & 2 \end{array} \right) \quad \text{и} \quad \mathbf{A} = \left(\begin{array}{cc} 1 & 0 \\ 1 & 1 \end{array} \right)$$

при попытке подобрать матрицу C, удовлетворяющую равенству

$$\mathbf{A}C = C \left(\begin{array}{cc} \alpha_1 & 0 \\ 0 & \alpha_2 \end{array} \right) \qquad \text{для } \forall \{\alpha_1, \alpha_2\} \subset \mathbb{C}$$

получим: $\det C = 0$.

Каждый кратный корень исследуется отдельно на количество линейно-независимых собственных векторов, ему принадлежащих. Позднее будет показано, что число таких векторов не превосходит кратности собственного числа в хар.полиноме. Таким образом, матрица оператора диагонализуема тогда и только тогда, когда для каждого собственного числа выполнено:

$$\operatorname{dfc}(\mathbf{A} - \lambda_j E) = \text{кратность } \lambda_j. \tag{5.6}$$

Пример 5.5. *Найти все вещественные значения параметра α, при которых матрицы*

$$\textbf{а)} \left(\begin{array}{rrr} 1 & 2\alpha & \alpha - 2 \\ -1 & 2 & 1 \\ 2 & 0 & -3 \end{array} \right); \quad \textbf{б)} \left(\begin{array}{rrrr} 1 & -3 & 0 & 3 \\ -2 & -2\alpha & 0 & 13 \\ 0 & -3 & 1 & 3 \\ -1 & -\alpha - 1 & 0 & 8 \end{array} \right)$$

диагонализуемы.

[14] Раздел I, глава 6, §2.

Решение. **а)** Хар.полином $f(\lambda) = -\lambda^3 + 3\,\lambda - 2\,(3\,\alpha - 1)$ имеет кратные корни только тогда когда его дискриминант $\mathcal{D}(f) = -324\,\alpha(3\,\alpha - 2)$ обращается в нуль. При $\alpha = 0$ корень $\lambda = -1$ имеет кратность 2. Найдем дефект матрицы $\mathbf{A} + E$:

$$\begin{pmatrix} 2 & 0 & -2 \\ -1 & 3 & 1 \\ 2 & 0 & -2 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & -1 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{pmatrix} \Longrightarrow \operatorname{rank} = 2, \operatorname{dfc} = 1$$

Условие (5.6) не выполнено. Оно не будет выполнено и при $\alpha = {}^2\!/_3$ (здесь корень $\lambda = 1$ имеет кратность 2).

б) Хар.полином $(\lambda-1)^2(\lambda^2-2\,(4-\alpha)\lambda+(10-3\,\alpha))$ всегда имеет кратный корень $\lambda = 1$, который при $\alpha \ne 3$ имеет вторую кратность, а при $\alpha = 3$ — третью. Ищем дефект матрицы $\mathbf{A} - E$:

$$\begin{pmatrix} 0 & -3 & 0 & 3 \\ -2 & (-2\alpha - 1) & 0 & 13 \\ 0 & -3 & 0 & 3 \\ -1 & (-\alpha - 1) & 0 & 7 \end{pmatrix} \longrightarrow \begin{pmatrix} 0 & -1 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ -1 & (-\alpha - 1) & 0 & 7 \end{pmatrix} \longrightarrow \begin{pmatrix} -1 & (-\alpha - 1) & 0 & 7 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Видим, что при любом значении параметра $\operatorname{dfc}(\mathbf{A} - E) = 2$.

Хар.полином имеет кратный корень еще и при $\alpha = 2$: это $\lambda = 2$ кратности 2. Для него $\operatorname{dfc}(\mathbf{A} - 2\,E) = 1$.

Ответ. Матрица диагонализуема при всех значениях параметра, за исключением **а)** $\alpha = 0$ и $\alpha = {}^2\!/_3$; **б)** $\alpha = 2$ и $\alpha = 3$. □

Упражнение 5.3. *Найти все вещественные значения параметра α, при которых матрица*

$$\begin{pmatrix} 3-\alpha & \alpha-5 & \alpha \\ -\alpha & \alpha-2 & \alpha \\ 5 & -5 & -2 \end{pmatrix}$$

диагонализуема.

6 Структура и свойства хар.полинома

В настоящем параграфе мы установим некоторые свойства хар. полинома произвольной матрицы A (на время отвлекаясь от ее интерпретации как матрицы некоторого оператора). Сначала попытаемся найти каноническое представление для хар.полинома, т.е. аналитические выражения для его коэффициентов при степенях λ через элементы матрицы A.

6.1 Каноническое представление хар.полинома

Пример 6.1. а) *Для $n = 2$:*

$$\det(A - \lambda E) = \begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} = \lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21}) \ .$$

б) *Для $n = 3$:*

$$\det(A - \lambda E) = \begin{vmatrix} a_{11} - \lambda & a_{12} & a_{13} \\ a_{21} & a_{22} - \lambda & a_{23} \\ a_{31} & a_{32} & a_{33} - \lambda \end{vmatrix} =$$

$$= -\lambda^3 + (a_{11} + a_{22} + a_{33})\lambda^2 - \left\{ \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} \right\} \lambda + \det A \ .$$

Обобщая эмпирические результаты, приходим к следующему каноническому представлению хар.полинома.

Теорема 6.1. $\det(A - \lambda E) =$

$$= (-1)^n \Big(\lambda^n - \lambda^{n-1} \sum_{1 \le j \le n} a_{jj} + \lambda^{n-2} \sum_{1 \le j_1 < j_2 \le n} \begin{vmatrix} a_{j_1 j_1} & a_{j_1 j_2} \\ a_{j_2 j_1} & a_{j_2 j_2} \end{vmatrix} - \dots +$$

$$+ (-1)^k \lambda^{n-k} \sum_{1 \le j_1 < j_2 < \dots < j_k \le n} \begin{vmatrix} a_{j_1 j_1} & a_{j_1 j_2} & \dots & a_{j_1 j_k} \\ a_{j_2 j_1} & a_{j_2 j_2} & \dots & a_{j_2 j_k} \\ & \dots & & \dots \\ a_{j_k j_1} & a_{j_k j_2} & \dots & a_{j_k j_k} \end{vmatrix} + \dots + (-1)^n \det A \Big) \ .$$

Образно говоря, коэффициент при $(-1)^k \lambda^{n-k}$ получается суммированием всех миноров k-го порядка матрицы A, построенных на элементах ее главной диагонали.

Следствие 1. *Если $\lambda_1, \dots, \lambda_n$ — собственные числа матрицы A (с учетом их кратностей), то*

$$\sum_{1 \le j \le n} \lambda_j = \sum_{1 \le j \le n} a_{jj} = \operatorname{Sp} A \ , \qquad \prod_{1 \le j \le n} \lambda_j = \det A \ . \tag{6.1}$$

Заметим, что каноническое представление для хар.полинома, даваемое теоремой 6.1 имеет скорее теоретическое, чем прикладное значение, и практические вычисления производятся иными способами[15]. Важным теоретическим следствием теоремы является тот факт, что коэффициенты хар.полинома являются непрерывными функциями элементов матрицы. На основании теоремы о непрерывной зависимости корней полинома от его коэффициентов[16] имеем следующий результат.

Теорема 6.2. *Собственные числа матрицы являются непрерывными функциями ее элементов. Иначе: пусть*

$$A = [a_{jk}]_{j,k=1}^n \quad , \quad B = [b_{jk}]_{j,k=1}^n \ .$$

Обозначим

$$M \stackrel{\text{def}}{=} \max_{j,k \in \{1,\dots,n\}} \{|a_{jk}|, |b_{jk}|\} \quad , \quad \delta \stackrel{\text{def}}{=} \frac{1}{nM} \sum_{j,k=1}^n |a_{jk} - b_{jk}| \ .$$

[15]См. далее главу 5.

[16]Раздел I, глава 3, теорема 7.8.

Тогда любому собственному числу λ_1 матрицы A можно поставить в соответствие такое собственное число μ_1 матрицы B, что

$$|\lambda_1-\mu_1| \le (n+2)M\sqrt[n]{\delta} \ .$$

Для некоторых матриц хар.полином вычисляется достаточно просто.

Пример 6.2. *Если $A_1, A_2, \dots, A_k$ — произвольные квадратные матрицы, то*

$$\det\left(\left[\begin{array}{llll} A_1 & * & \dots & * \\ \mathbb{O} & A_2 & \dots & * \\ \dots & & \ddots & \\ \mathbb{O} & & \mathbb{O} & A_k \end{array}\right]-\lambda E\right)=\prod_{1\le j\le k}\det(A_j-\lambda E) \ ;$$

в частности,

$$\det\left(\left[\begin{array}{cccc} a_{11} & a_{12} & \dots & a_{1n} \\ & a_{22} & \dots & a_{2n} \\ & \mathbb{O} & \ddots & \vdots \\ & & & a_{nn} \end{array}\right]-\lambda E\right)=\prod_{1\le j\le n}(a_{jj}-\lambda) \ .$$

Пример 6.3. *Для матрицы* **Фробениуса**

$$\mathfrak{F} \stackrel{\mathrm{def}}{=} \left(\begin{array}{lllllll} 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 & 0 \\ \dots & & & & & & \dots \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 \\ a_n & a_{n-1} & a_{n-2} & & \dots & a_2 & a_1 \end{array}\right)_{n\times n}$$

хар. полином равен $(-1)^n(\lambda^n-a_1\lambda^{n-1}-\dots-a_n)$.

Теперь приведем один результат, касающийся оценки расположения собственных чисел матрицы A.

Теорема 6.3 (Гершгорин). *Обозначим $\mathbb{D}_j$ круг на комплексной плоскости $\mathbb{C}$ с центром в точке a_{jj} и радиуса $r_j=\sum_{\ell\ne j}|a_{j\ell}|$. Тогда все собственные числа матрицы A лежат внутри объединения этих кругов:*

$$\{\lambda_1,\dots,\lambda_n\}\subset\bigcup_{j=1}^{n}\mathbb{D}_j \ .$$

Доказательство . Пусть $\mathfrak{X}=[\mathfrak{x}_1,\dots,\mathfrak{x}_n]^{\top}$ — собственный вектор матрицы A, соответствующий собственному числу λ: $A\mathfrak{X}=\lambda\mathfrak{X}$. Выделим у этого вектора максимальную по абсолютной величине координату: пусть

$$|\mathfrak{x}_j|=\max_{1\le k\le n}|\mathfrak{x}_k| \ .$$

Рассмотрим тогда j-ю компоненту равенства $A\mathfrak{X} = \lambda\mathfrak{X}$:

$$\sum_{k=1}^{n} a_{jk}\mathfrak{x}_k = \lambda\mathfrak{x}_j \ \Rightarrow \ \sum_{k=1}^{n} a_{jk}\mathfrak{x}_k - a_{jj}\mathfrak{x}_j = \lambda\mathfrak{x}_j - a_{jj}\mathfrak{x}_j \ .$$

Переходя к модулям, получаем:

$$|\lambda - a_{jj}| \cdot |\mathfrak{x}_j| = \left| \sum_{k\ne j} a_{jk}\mathfrak{x}_k \right| \le \sum_{k\ne j} |a_{jk}| \cdot |\mathfrak{x}_k| \le |\mathfrak{x}_j| \sum_{k\ne j} |a_{jk}| \ .$$

Поделив обе части неравенства на $|\mathfrak{x}_j|$, придем к $|\lambda - a_{jj}| \le r_j$; последнее неравенство говорит о том, что $\lambda \in \mathbb{D}_j \ \Rightarrow \ \lambda \in \bigcup_{j=1}^{n} \mathbb{D}_j$. $\square$

Пример 6.4. *Построить круги Гершгорина для матрицы*

$$A = \left(\begin{array}{ccc} -2 & 4+7\boldsymbol{i} & -3-7\boldsymbol{i} \\ -1 & 6+2\boldsymbol{i} & -5-2\boldsymbol{i} \\ -3 & 4+7\boldsymbol{i} & -2-7\boldsymbol{i} \end{array} \right) \ .$$

Решение.

$$|\lambda+2| \le |4+7\boldsymbol{i}\,|+|-3-7\boldsymbol{i}\,| = \sqrt{65}+\sqrt{58},\ |\lambda-6-2\boldsymbol{i}\,| \le 1+\sqrt{29},\ |\lambda+2+7\boldsymbol{i}\,| \le 3+\sqrt{65} \ .$$

Проверка. Спектр матрицы A: $\{1,\ -2\boldsymbol{i}\,,\ 1-3\boldsymbol{i}\,\}$.

6.2 Теорема Гамильтона–Кэли

Теорема 6.4 (Гамильтон, Кэли). *Если* $f(\lambda) = \det(A - \lambda E)$*, то* $f(A) = \mathbb{O}_{n\times n}$*. Иначе говорят: матрица A является корнем своего хар. полинома.*

Пример 6.5. *Для $n = 2$ справедливость утверждения проверяется непосредственно:*

$$\left(\begin{array}{cc} a_{11} & a_{12} \\ a_{21} & a_{22} \end{array} \right)^2 - (a_{11}+a_{22}) \left(\begin{array}{cc} a_{11} & a_{12} \\ a_{21} & a_{22} \end{array} \right) + (a_{11}a_{22} - a_{12}a_{21}) \left(\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right) = \left(\begin{array}{cc} 0 & 0 \\ 0 & 0 \end{array} \right) \ .$$

Доказательство . Рассмотрим хар. матрицу $A - \lambda E$ и вычислим ей союзную[17]. Для $n = 3$:

$$B(\lambda) = \left(\begin{array}{ccc} \left| \begin{array}{cc} a_{22}-\lambda & a_{23} \\ a_{32} & a_{33}-\lambda \end{array} \right| & -\left| \begin{array}{cc} a_{12} & a_{13} \\ a_{32} & a_{33}-\lambda \end{array} \right| & \left| \begin{array}{cc} a_{12} & a_{13} \\ a_{22}-\lambda & a_{23} \end{array} \right| \\ -\left| \begin{array}{cc} a_{21} & a_{23} \\ a_{31} & a_{33}-\lambda \end{array} \right| & \left| \begin{array}{cc} a_{11}-\lambda & a_{13} \\ a_{31} & a_{33}-\lambda \end{array} \right| & -\left| \begin{array}{cc} a_{11}-\lambda & a_{13} \\ a_{21} & a_{23} \end{array} \right| \\ \left| \begin{array}{cc} a_{21} & a_{22}-\lambda \\ a_{31} & a_{32} \end{array} \right| & -\left| \begin{array}{cc} a_{11}-\lambda & a_{12} \\ a_{31} & a_{32} \end{array} \right| & \left| \begin{array}{cc} a_{11}-\lambda & a_{12} \\ a_{21} & a_{22}-\lambda \end{array} \right| \end{array} \right) \ .$$

[17] Матрицей, союзной матрице $\mathfrak{A} = \left[\mathfrak{a}_{jk}\right]_{j,k=1}^{n}$ называется матрица $\left(\left[\mathfrak{A}_{jk}\right]_{j,k=1}^{n}\right)^{\top}$, где $\mathfrak{A}_{jk}$ — алгебраическое дополнение к элементу $\mathfrak{a}_{jk}$. Раздел I, глава 4, §10.

На основании известных свойств союзной матрицы имеем:

$$B(\lambda)(A-\lambda E) \equiv E\det(A-\lambda E) = Ef(\lambda)\ , \quad (A-\lambda E)B(\lambda) \equiv Ef(\lambda)\ . \tag{6.2}$$

Поскольку элементами матрицы $B(\lambda)$ являются полиномы степеней $\le n-1$, то $B(\lambda)$ можно представить в виде матричного полинома

$$B(\lambda) = B_0 + \lambda B_1 + \cdots + \lambda^{n-1}B_{n-1}.$$

Так, например, для случая $n=3$ имеем:

$$\begin{aligned} B(\lambda) \quad &= \quad \begin{pmatrix} a_{22}a_{33}-a_{23}a_{32} & a_{32}a_{13}-a_{12}a_{33} & a_{12}a_{23}-a_{13}a_{22} \\ a_{31}a_{23}-a_{21}a_{33} & a_{11}a_{33}-a_{13}a_{31} & a_{13}a_{21}-a_{11}a_{23} \\ a_{21}a_{32}-a_{22}a_{31} & a_{31}a_{12}-a_{11}a_{32} & a_{11}a_{22}-a_{12}a_{21} \end{pmatrix} + \\ &+ \quad \lambda \begin{pmatrix} -a_{22}-a_{33} & a_{12} & a_{13} \\ a_{21} & -a_{11}-a_{33} & a_{23} \\ a_{31} & a_{32} & -a_{11}-a_{22} \end{pmatrix} + \lambda^2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad . \end{aligned}$$

Из первого из тождеств (6.2) получаем уравнения для определения неопределенных матриц $B_0, B_1, \dots, B_{n-1}$:

$$B_0A + \lambda(B_1A - B_0) + \cdots + \lambda^{n-1}(B_{n-1}A - B_{n-2}) - \lambda^n B_{n-1} =$$

$$= E(a_n + a_{n-1}\lambda + \cdots + a_1\lambda^{n-1} + a_0\lambda^n).$$

Коэффициент при				домножим на
1	B_0A	$=$	Ea_n	E
λ	$B_1A - B_0$	$=$	Ea_{n-1}	A
λ^2	$B_2A - B_1$	$=$	Ea_{n-2}	A^2
$\dots$	$\dots$	$=$	$\dots$	$\dots$
λ^{n-1}	$B_{n-1}A - B_{n-2}$	$=$	Ea_1	A^{n-1}
λ^n	$-B_{n-1}$	$=$	Ea_0	A^n
				и просуммируем:

$$\mathbb{O}_{n\times n} = Ea_n + Aa_{n-1} + A^2a_{n-2} + \cdots + A^na_0\ \ .$$

□

Следствие 1. *Пусть каноническое разложение хар. полинома над* $\mathbb{C}$ *имеет вид:*

$$f(\lambda) \equiv (-1)^n(\lambda-\lambda_1)(\lambda-\lambda_2)\times\cdots\times(\lambda-\lambda_n)\ \ .$$

Тогда

$$(A-\lambda_1E)(A-\lambda_2E)\times\cdots\times(A-\lambda_nE) = \mathbb{O}_{n\times n}\ \ , \tag{6.3}$$

причем порядок следования сомножителей значения не имеет.

Следствие 2. *Матрицы* $E, A, A^2, \dots, A^n$ *всегда линейно зависимы, но существуют матрицы* A*, для которых линейно зависимы* $E, A, A^2, \dots, A^k$ *при* $k<n$.

Следствие 3. *Любая матрица A^{n+k} при $k \in \mathbb{N}$ линейно выражается через*

$$E, A, A^2, \dots, A^{n-1} \tag{6.4}$$

Если же $a_n \neq 0$, то и A^{-k} при $k \in \mathbb{N}$ линейно выражается через те же матрицы.

Доказательство . Домножим равенство $f(A) = \mathbb{O}$ на A^k:

$$(-1)^n A^{n+k} + a_1 A^{n+k-1} + \dots + a_n A^k = \mathbb{O} \quad \Rightarrow$$

$$\Rightarrow A^{n+k} = (-1)^{n+1} \left(a_1 A^{n+k-1} + \dots + a_n A^k \right) \ .$$

Получили рекуррентное соотношение, связывающее степень A^{n+k} с n предшествующими степенями. Продолжая процесс выражения далее, в конце концов опустимся до набора (6.4).

Поскольку $a_n = \det A$, то условие $a_n \neq 0$ влечет существование A^{-1}. Домножим равенство $f(A) = \mathbb{O}$ на A^{-1}:

$$A^{-1} = -\frac{1}{a_n} \left(A^{n-1} + a_1 A^{n-2} + \dots + a_{n-1} E \right) \ ,$$

т.е. обратная матрица линейно выражается через матрицы (6.4). Вычисление A^{-k} сводится к тем же матрицам по рекурсивному закону:

$$A^{-k} = -\frac{1}{a_n} \left(A^{n-k} + a_1 A^{n-k-1} + \dots + a_{n-1} A^{-k+1} \right) \ .$$

□

После того, как спектр[18] матрицы установлен (приближенно или точно), можно найти и соответствующие собственные векторы — через нахождение **ф.с.р.** систем линейных уравнений, как это делалось в пункте 5.2. Покажем еще один способ нахождения собственного вектора, основанный на теореме 6.4.

Теорема 6.5. *Обозначим частное от деления хар. полинома на его k-й линейный множитель через $f_k(\lambda)$:*

$$f_k(\lambda) \stackrel{\text{def}}{=} f(\lambda)/(\lambda - \lambda_k) \ .$$

Тогда любой ненулевой столбец матрицы $f_k(A)$ является собственным вектором, принадлежащим λ_k.

Доказательство следует из равенства (6.3)

$$(A - \lambda_k E) f_k(A) = \mathbb{O}_{n \times n} \ .$$

На основании определения любой ненулевой столбец $f_k(A)$ должен быть собственным вектором. □

[18] Определение — на с. 75.

Пример 6.6. *Найти собственные векторы матрицы*

$$A = \begin{pmatrix} 9 & 22 & -6 \\ -1 & -4 & 1 \\ 8 & 16 & -5 \end{pmatrix} .$$

Решение. $\det(A - \lambda E) = -(\lambda - 3)(\lambda + 2)(\lambda + 1)$. Пренебрегая знаком $-$, имеем:

$$f_1(\lambda) = \lambda^2 + 3\lambda + 2 \quad \text{и} \quad f_1(A) = \begin{pmatrix} 40 & 80 & -20 \\ 0 & 0 & 0 \\ 40 & 80 & -20 \end{pmatrix} ,$$

$$f_2(\lambda) = \lambda^2 - 2\lambda - 3 \quad \text{и} \quad f_2(A) = \begin{pmatrix} -10 & -30 & 10 \\ 5 & 15 & -5 \\ 0 & 0 & 0 \end{pmatrix} ,$$

$$f_3(\lambda) = \lambda^2 - \lambda - 6 \quad \text{и} \quad f_3(A) = \begin{pmatrix} -4 & -8 & 4 \\ 4 & 8 & -4 \\ 8 & 16 & -8 \end{pmatrix} .$$

Ответ. $\mathfrak{X}_1 = [1, 0, 1]^\top$ соответствует $\lambda_1 = 3$, $\mathfrak{X}_2 = [-2, 1, 0]^\top$ соответствует $\lambda_2 = -2$, $\mathfrak{X}_3 = [-1, 1, 2]^\top$ соответствует $\lambda_3 = -1$.

Следствие 1. *Если λ_k является простым корнем хар. полинома, то ненулевые столбцы $f_k(A)$ будут пропорциональными.*

Рассмотренный только что пример показывает, однако, нечто большее — ненулевые *строки* матрицы $f_k(A)$ тоже пропорциональны!

Упражнение 6.1. *Доказать, что любая ненулевая строка матрицы $f_k(A)$ является собственным вектором матрицы $A^\top$, принадлежащим λ_k.*

Пусть теперь $g(x) = b_0 x^m + b_1 x^{m-1} + \dots + b_m \in \mathbb{C}$ — произвольный полином. Он порождает матричный полином

$$g(A) = b_0 A^m + b_1 A^{m-1} + \dots + b_m E .$$

Как связаны между собой спектры матриц A и $g(A)$?

Теорема 6.6. *Если $\mathfrak{X}_1 \in \mathbb{C}^n$ — собственный вектор матрицы A, соответствующий собственному числу λ_1, то он же будет собственным и для матрицы $g(A)$, соответствующим собственному числу $g(\lambda_1)$. Если $\{\lambda_1, \dots, \lambda_n\}$— спектр A, то $\{g(\lambda_1), \dots, g(\lambda_n)\}$— спектр $g(A)$.*

Доказательство . Домножим равенство $A\mathfrak{X}_1 = \lambda_1 \mathfrak{X}_1$ слева на матрицу A:

$$A^2 \mathfrak{X}_1 = \lambda_1 A \mathfrak{X}_1 = \lambda_1^2 \mathfrak{X}_1 .$$

По индукции доказывается и общее равенство: $A^k \mathfrak{X}_1 = \lambda_1^k \mathfrak{X}_1$. Домножим его на b_{m-k} и просуммируем по k от 0 до m: $g(A)\mathfrak{X}_1 = g(\lambda_1)\mathfrak{X}_1$, что и доказывает первую часть утверждения теоремы.

Вторая часть теоремы доказывается тривиально при дополнительном предположении, что все числа $g(\lambda_1), \dots, g(\lambda_n)$ различны. Действительно, поскольку $\deg(\det[g(A) - \lambda E]) = n$, то иных корней у хар. полинома матрицы $g(A)$ просто быть не может.

Для доказательства в общем случае приходится использовать вспомогательный результат.

Лемма 1. *Для произвольного полинома $G(x) \in \mathbb{C}[x]$ имеем: $\det G(A)$ с точностью до знака совпадает с результантом*[19] *полиномов f и G:*

$$\det G(A) = (-1)^{mn} \mathcal{R}(f, G) \ .$$

Доказательство леммы. Разложим $G(x)$ на линейные множители:

$$G(x) \equiv b_0(x - \mu_1) \times \cdots \times (x - \mu_m)$$

тогда $G(A) = b_0(A - \mu_1 E) \times \cdots \times (A - \mu_m E)$ и $\det G(A) =$

$$= b_0^n \det(A - \mu_1 E) \times \cdots \times \det(A - \mu_m E) = b_0^n f(\mu_1) \times \cdots \times f(\mu_m) = (-1)^{mn} \mathcal{R}(f, G) \ .$$

□

Для завершения доказательства теоремы достаточно применить лемму, взяв $G(x) \stackrel{\text{def}}{=} g(x) - \lambda$:

$$\det(g(A) - \lambda E) = (-1)^{mn} \mathcal{R}(f(x), g(x) - \lambda) =$$

и воспользоваться представлением результанта через корни полинома $f(x)$:

$$= (g(\lambda_1) - \lambda) \times \cdots \times (g(\lambda_n) - \lambda) \ .$$

□

Следствие 1. *Справедливо равенство:*

$$\operatorname{Sp} A^k = \sum_{1 \le j \le n} \lambda_j^k \ . \tag{6.5}$$

6.3 Диагонализуемость матрицы оператора над $\mathbb{R}$

Во всех предшествующих параграфах мы рассматривали операторы над полем $\mathbb{C}$. Как следствие, элементы матриц были, в общем случае, комплексными. Однако на практике чаще встречаются операторы над полем $\mathbb{R}$ и интерес представляет возможность их диагонализации над этим же полем. Следствие 1 к теореме 5.7 дает необходимое и достаточное условие диагонализуемости матрицы $\mathbf{A}$ оператора $\mathscr{A}$ над $\mathbb{R}$: все собственные числа матрицы должны быть вещественными.

Теорема 5.8 позволяет сформулировать и достаточный критерий диагонализуемости матрицы оператора $\mathscr{A}$ над $\mathbb{R}$.

[19] Раздел I, глава 6, §1.

Теорема 6.7. *Если хар.полином оператора имеет только простые вещественные корни, то матрица оператора диагонализуема над* $\mathbb{R}$.

Доказательство . Действительно, линейная система для нахождения собственного вектора будет иметь вид: $(\mathbf{A}-\lambda_j E)X=\mathbb{O}$. Поскольку $\lambda_j\in\mathbb{R}$, то и решение $\mathfrak{X}_j$ этой системы может быть выбрано из $\mathbb{R}^n$. На основании теоремы 5.6, *вещественная* неособенная матрица $C=[\mathfrak{X}_1,\dots,\mathfrak{X}_n]$ приводит $\mathbf{A}$ к диагональному виду. □

Этот критерий допускает возможность алгебраической проверки. Напомним, что условие различности и вещественности корней хар. полинома

$$f(\lambda)=(-1)^n(\lambda^n+a_1\lambda^{n-1}+\dots+a_n)$$

можно установить по коэффициентам этого полинома. Воспользуемся, например, теоремой Якоби[20]. Напомним, что по коэффициентам $a_1,\dots,a_n$ можно определить сумму Ньютона полинома $f(\lambda)$, т.е. величину

$$s_k=\sum_{1\le j\le n}\lambda_j^k\ . \tag{6.6}$$

Далее, после нахождения этих сумм при $k\in\{0,\dots,2n-2\}$, из них составляется матрица

$$S=[s_{j+k}]_{j,k=0}^{n-1} \tag{6.7}$$

и вычисляются ее главные миноры $S_1,\dots,S_n$. Для вещественности всех корней $f(\lambda)$ необходимо и достаточно, чтобы

$$S_1\ge 0,\dots,S_n\ge 0\ ; \tag{6.8}$$

все эти корни будут различными тогда и только тогда, когда $S_n>0$.

Пример 6.7. *Установить условия диагонализуемости над* $\mathbb{R}$ *матрицы из примера* 5.5 **а)**.

Решение. На основании теоремы 6.7 нам нужно установить условия вещественности корней хар.полинома $f(\lambda)=-\lambda^3+3\,\lambda-2\,(3\,\alpha-1)$. Вычисляем суммы Ньютона: $s_0=3,\ s_1=0,\ s_2=6,\ s_3=18\,\alpha-6,\ s_4=18$, а затем главные миноры матрицы S:

$$S_1=3,\ S_2=18,\ S_3=-324\,\alpha\,(3\,\alpha-2)=\mathcal{D}(f)\ .$$

При $\alpha\ne 0$ и $\alpha\ne {}^2\!/_3$ все собственные числа различны, условие (6.8) выполняется при $0<\alpha<{}^2\!/_3$. Граничные точки последнего интервала следовало бы исследовать отдельно: хотя этим значениям параметра и соответствует случай кратных вещественных корней хар. полинома, но матрица $\mathbf{A}$ может оказаться диагонализуемой на основании следствия 1 к теореме 5.7, т.е. при выполнении условия (5.6). Но на с. 79 мы уже установили, что это условие не выполнено.

Ответ. Матрица диагонализуема над $\mathbb{R}$ при $0<\alpha<{}^2\!/_3$.

[20] Раздел I, глава 8, §2.

Пример 6.8. *Выяснить, диагонализуема ли над $\mathbb{R}$ матрица*

$$A = \begin{pmatrix} -13 & 29 & -15 & -12 \\ -14 & 30 & -15 & -14 \\ -13 & 27 & -13 & -14 \\ -7 & 14 & -6 & -8 \end{pmatrix} \tag{6.9}$$

Решение проведем, несколько изменив алгоритм решения предыдущего примера. Заметим, что знание коэффициентов хар. полинома вовсе не необходимо для построения матрицы S. Действительно, на основании формулы (6.5), суммы Ньютона (6.6) можно найти посредством вычисления степеней матрицы A: $s_1 = \operatorname{Sp} A = -4$,

$$A^2 = \begin{pmatrix} 42 & -80 & 27 & 56 \\ 55 & -107 & 39 & 70 \\ 58 & -114 & 43 & 72 \\ 29 & -57 & 21 & 36 \end{pmatrix} \quad \Rightarrow \quad s_2 = \operatorname{Sp} A^2 = 14.$$

Продолжая вычисления степеней далее, получим

$$A^3 = \begin{pmatrix} -169 & 331 & -117 & -210 \\ -214 & 418 & -147 & -268 \\ -221 & 431 & -151 & -278 \\ -104 & 202 & -69 & -132 \end{pmatrix} \quad \Rightarrow \quad s_3 = -34,$$

$s_4 = 90,\ s_5 = -234,\ s_6 = 614$. Теперь составим матрицу (6.7) и найдем ее главные миноры:

$$S_1 = 4,\ S_2 = 40,\ S_3 = 40,\ S_4 = 96\ .$$

Все $S_j > 0$, следовательно все собственные числа вещественны и различны. Матрица диагонализуема над $\mathbb{R}$.

Развивая этот подход, можно даже локализовать собственные числа — воспользовавшись теоремой Йоахимшталя. Для этого понадобится произвести еще одно вычисление:

$$s_7 = \operatorname{Sp} A^7 = \operatorname{Sp} \begin{pmatrix} -16265 & 31219 & -9657 & -21142 \\ -21266 & 40826 & -12651 & -27632 \\ -22265 & 42747 & -13255 & -28926 \\ -9934 & 19062 & -5883 & -12920 \end{pmatrix} = -1614\ .$$

Далее, составляем матрицу

$$H(t) = [s_{j+k}t - s_{j+k+1}]_{j,k=0}^{3}\ .$$

Подставляя сюда вместо t конкретные значения $t = a$ и $t = b$ и вычисляя соответствующие числовые миноры получающихся матриц, можем найти количество собственных чисел матрицы на $]a, b[$:

$$\operatorname{nrr}\{\det(A - \lambda E) = 0 \mid a < \lambda < b\} =$$

$$= \mathcal{V}(1, H_1(a), \ldots, H_4(a)) - \mathcal{V}(1, H_1(b), \ldots, H_4(b)), \tag{6.10}$$

Например, для $t = 0$: $\mathcal{V}(1, 4, -60, -64, 192) = 2$, т.е. два собственных числа отрицательны и два положительны. Далее находим грубую нижнюю их оценку, и локализуем подсчетом знаков миноров $H_j(t)$: по два корня на $\big] -3, -2\big[$ и на $]0, 1[$.

Проверка. Собственные числа этой матрицы: $-1 + \sqrt{3}$, $-1 - \sqrt{3}$, $-1 + \sqrt{2}$, $-1 - \sqrt{2}$.

Упражнение 6.2. *Локализовать комплексные собственные числа матрицы*

$$A = \begin{pmatrix} -160 & 0 & 161 \\ 89 & 1 & -88 \\ -149 & -1 & 149 \end{pmatrix}$$

с помощью теоремы 6.3 *(Гершгорина), а ее вещественные собственные числа — с помощью теоремы Йоахимшталя.*

Историческая справка

Кэли Артур (Cayley Arthur, 1821–1895) родился в Ричмонде, графство Сюррей. До 1829 г. жил в Санкт-Петербурге, где работал купцом его отец. В 1842 окончил Trinity Colledge первым в выпуске и в течение 4 лет получал стипендию для исследований в Кэмбридже. Нужно было зарабатывать деньги, он выбрал адвокатуру и в 1849 году был принят в коллегию адвокатов. В течение 14 лет работал адвокатом, считался крупным специалистом по составлению нотариальных актов о передаче имущества. Кэли приписывают следующую фразу

The object of law was to say a thing in the greatest number of words, and of mathematics to say in the fewest
(Задача юриспруденции заключается в описании факта наибольшим количеством слов, в то время как математики — наименьшим.)

от авторства которой он, впрочем, неоднократно отрекался. За эти годы опубликовал около 250 математических работ.

В 1863 году получил приглашение стать профессором Pure Mathematics в Кэмбридже и немедленно его принял, хотя это приводило к значительному понижению доходов — его заработок уменьшался в несколько раз.

Общее количество публикаций — около 900. Разработал теорию определителей (первым ввел привычное обозначение для определителя), и распространил ее на $n-$мерные матрицы. Работы по алгебре, геометрии, дифференциальным уравнениям, астрономии.

7 Диагонализуемость симметричной матрицы над $\mathbb{R}$

Для одного достаточно широкого класса вещественных матриц условие диагонализуемости над $\mathbb{R}$ всегда гарантировано. Это — симметричные матрицы: $\mathbf{A} = \mathbf{A}^\top$.

7.1 Свойства собств. чисел и собств. векторов

Теорема 7.1. *Все собств.числа симметричной матрицы* $\mathbf{A}$ *вещественны.*

Доказательство I. Если $\lambda \in \mathbb{C}$ — собств.число, соответствующее собств.вектору $\mathfrak{X} \in \mathbb{C}^n$ ($\mathfrak{X} \ne \mathbb{O}$), то $\mathbf{A}\mathfrak{X} = \lambda\mathfrak{X}$. Обозначим $\overline{\mathfrak{X}}$ вектор комплексно-сопряженный $\mathfrak{X}$. Вычислим число $a = (\overline{\mathfrak{X}})^\top \mathbf{A}\mathfrak{X} \in \mathbb{C}$

$$a = (\overline{\mathfrak{X}})^\top A\mathfrak{X} = (\overline{\mathfrak{X}})^\top \lambda\mathfrak{X} = \lambda(\overline{\mathfrak{X}})^\top \mathfrak{X} = \lambda \underbrace{(|x_1|^2 + \dots + |x_n|^2)}_{\in\mathbb{R}}.$$

Если мы теперь докажем, что $a \in \mathbb{R}$, то тогда и $\lambda \in \mathbb{R}$. Имеем:

$$\overline{a} = (\overline{a})^\top = \left(\overline{(\overline{\mathfrak{X}})^\top \mathbf{A}\mathfrak{X}}\right)^\top = \left(\mathfrak{X}^\top \overline{\mathbf{A}\mathfrak{X}}\right)^\top = (\overline{\mathfrak{X}})^\top \mathbf{A}^\top \mathfrak{X} = (\overline{\mathfrak{X}})^\top \mathbf{A}\mathfrak{X} = a.$$

Но это и означает, что $a \in \mathbb{R}$. □

Итак, нам удалось установить вещественность всех корней хар. полинома симметричной матрицы — и установить это косвенным путем. Приведем еще одно доказательство этого результата, на этот раз непосредственно связанное с общим результатом о **Н.** и **Д.** условиях вещественности всех корней полинома. Это — уже использованная в предыдущем параграфе теорема Якоби.

Доказательство II. В евклидовом пространстве квадратных матриц порядка n со скалярным произведением, введенным формулой (1.5) главы 2:

$$(A, B) \stackrel{\text{def}}{=} \operatorname{Sp}\left(A \cdot B^\top\right) \quad ,$$

рассмотрим систему матриц $\left\{E, \mathbf{A}, \mathbf{A}^2, \dots, \mathbf{A}^{n-1}\right\}$. Составим матрицу Грама[21] указанной системы:

$$G(E, \mathbf{A}, \mathbf{A}^2, \dots, \mathbf{A}^{n-1}) = \left(\begin{array}{cccc} (E, E) & (E, \mathbf{A}) & \dots & (E,\, \mathbf{A}^{n-1}) \\ (\mathbf{A}, E) & (\mathbf{A}, \mathbf{A}) & \dots & (\mathbf{A},\, \mathbf{A}^{n-1}) \\ \dots & & & \dots \\ (\mathbf{A}^{n-1}, E) & (\mathbf{A}^{n-1}, \mathbf{A}) & \dots & (\mathbf{A}^{n-1}, \mathbf{A}^{n-1}) \end{array} \right) . \tag{7.1}$$

При симметричной матрице $\mathbf{A}$ любая ее степень также является симметричной матрицей. Тогда $(\mathbf{A}^j, \mathbf{A}^k) = \operatorname{Sp}(\mathbf{A}^{j+k}) \stackrel{(6.5)}{=} s_{j+k}$, т.е. элементы матрицы (7.1) являются суммами Ньютона хар. полинома матрицы $\mathbf{A}$. Следовательно, сама матрица (7.1) совпадает с матрицей

$$S = \left(\begin{array}{cccc} s_0 & s_1 & \dots & s_{n-1} \\ s_1 & s_2 & \dots & s_n \\ \dots & & & \dots \\ s_{n-1} & s_n & \dots & s_{2n-2} \end{array} \right)$$

[21] Глава 2, §3.2.

из теоремы Якоби.

По теореме 4.2 главы 2, все главные миноры S_j этой матрицы неотрицательны. Следовательно, по теореме Якоби, все корни хар. полинома матрицы $\mathbf{A}$ вещественны. $\square$

Следствие 1. *Для каждого собств.числа симметричной матрицы найдется соответствующий ему вещественный собств.вектор.*

Доказательство . Действительно, линейная система для нахождения этого вектора будет иметь вид: $(\mathbf{A} - \lambda_j E)X = \mathbb{O}$. Поскольку, по доказанному, $\lambda_j \in \mathbb{R}$, то и решение $X = \mathfrak{X}_j$ этой системы может быть выбрано из $\mathbb{R}^n$. Более того, можно нормировать $\mathfrak{X}_j$, т.е. сделать $|\mathfrak{X}_j| = 1$. $\square$

Теорема 7.2. *Собств. векторы, принадлежащие различным собств.числам симметричной матрицы* $\mathbf{A}$ *попарно ортогональны.*

Доказательство . Если $\mathbf{A}\mathfrak{X}_1 = \lambda_1\mathfrak{X}_1, \mathbf{A}\mathfrak{X}_2 = \lambda_2\mathfrak{X}_2$ и $\lambda_1 \neq \lambda_2$, то

$$a \overset{\text{def}}{=} \mathfrak{X}_2^\top \mathbf{A}\mathfrak{X}_1 = \begin{cases} \mathfrak{X}_2^\top \lambda_1 \mathfrak{X}_1 & = \lambda_1 \mathfrak{X}_2^\top \mathfrak{X}_1; \\ \mathfrak{X}_2^\top \mathbf{A}^\top \mathfrak{X}_1 = (\mathbf{A}\mathfrak{X}_2)^\top \mathfrak{X}_1 & = \lambda_2 \mathfrak{X}_2^\top \mathfrak{X}_1. \end{cases}$$

Тогда $0 = a - a = (\lambda_1 - \lambda_2)\mathfrak{X}_2^\top \mathfrak{X}_1 \ \Rightarrow\ \mathfrak{X}_2^\top \mathfrak{X}_1 = 0$, т.е. $\mathfrak{X}_1 \perp \mathfrak{X}_2$. $\square$

Определение. Вещественная матрица $P = [P_1, \dots, P_n]$ называется **ортогональной**, если ее столбцы образуют ортонормированный базис $\mathbb{R}^n$, т.е. удовлетворяют условию

$$(P_j, P_k) = P_j^\top P_k = \delta_{jk} = \begin{cases} 1 & \text{если } j = k; \\ 0 & \text{если } j \neq k. \end{cases}$$

Иначе:

$$P \cdot P^\top = E \quad \Longleftrightarrow \quad P^{-1} = P^\top$$

(иногда последнее равенство берут в качестве определения ортогональной матрицы).

Теорема 7.3. *Множество ортогональных матриц порядка* n *образует подгруппу группы* $\mathbf{GL}(n, \mathbb{R})$.

Доказательство . Если $PP^\top = E,\ QQ^\top = E$, то и

$$(PQ)^\top PQ = Q^\top P^\top PQ = Q^\top EQ = Q^\top Q = E \quad \text{и} \quad (Q^{-1})^{-1} = Q = (Q^{-1})^\top \ .$$

Таким образом, произведение ортогональных матриц будет ортогональной и обратная к ортогональной — тоже ортогональная. $\square$

Упражнение 7.1. *Будет ли эта подгруппа абелевой?*

Упражнение 7.2. *Доказать, что*

а) *определитель любой ортогональной матрицы равен* $+1$ *или* -1, *а любое ее собственное число по модулю равно* 1;

б) *в определении ортогональной матрицы можно свойство ортогональности строк заменить на ортогональность столбцов;*

в) *произведение симметричных матриц — не обязательно симметричная матрица.*

7.2 Диагонализуемость

Теорема 7.4. *Существует ортогональная матрица P, приводящая симметричную матрицу $\mathbf{A}$ к диагональному виду:*

$$P^{-1}\mathbf{A}P = P^{\top}\mathbf{A}P = \begin{pmatrix} \lambda_1 & & & \mathbb{O} \\ & \lambda_2 & & \\ & & \ddots & \\ \mathbb{O} & & & \lambda_n \end{pmatrix}. \tag{7.2}$$

Доказательство особенно просто в случае когда все собств.числа различны. На основании теорем 6.7 и 7.1 матрица $\mathbf{A}$ диагонализуема над $\mathbb{R}$ и на основании теоремы 7.2 матрица P, приводящая к диагональному виду может быть выбрана ортогональной.

Для общего случая докажем индукцией по порядку n матрицы $\mathbf{A}$. При $n = 1$ утверждение тривиально. Пусть оно справедливо для любой симметричной матрицы порядка $n-1$. По теореме 7.1 у матрицы $\mathbf{A}$ существует хотя бы одно собственное число $\lambda_1 \in \mathbb{R}$, и ему соответствует хотя бы один собственный вектор-столбец $\mathfrak{X}_1 \in \mathbb{R}^n$. Будем считать $|\mathfrak{X}_1| = 1$. Рассмотрим произвольную ортогональную матрицу

$$Q = [\mathfrak{X}_1, Q_2, \dots, Q_n] \ , \quad (\mathfrak{X}_1, Q_2) = 0, \dots, (\mathfrak{X}_1, Q_n) = 0, \quad (Q_j, Q_k) = \delta_{jk} \ ,$$

дополнив $\mathfrak{X}_1$ до ортонормированного базиса $\mathbb{R}^n$. Имеем:

$$\begin{aligned} Q^{\top}\mathbf{A}Q &= \begin{bmatrix} \mathfrak{X}_1^{\top} \\ Q_2^{\top} \\ \vdots \\ Q_n^{\top} \end{bmatrix} \mathbf{A}\,[\mathfrak{X}_1, Q_2, \dots, Q_n] = \begin{bmatrix} \mathfrak{X}_1^{\top} \\ Q_2^{\top} \\ \vdots \\ Q_n^{\top} \end{bmatrix} [\mathbf{A}\mathfrak{X}_1, \mathbf{A}Q_2, \dots, \mathbf{A}Q_n] = \\ &= \begin{bmatrix} \mathfrak{X}_1^{\top} \\ Q_2^{\top} \\ \vdots \\ Q_n^{\top} \end{bmatrix} [\lambda_1\mathfrak{X}_1, \mathbf{A}Q_2, \dots, \mathbf{A}Q_n] = \begin{bmatrix} \lambda_1\mathfrak{X}_1^{\top}\mathfrak{X}_1 & \mathfrak{X}_1^{\top}\mathbf{A}Q_2 & \dots & \mathfrak{X}_1^{\top}\mathbf{A}Q_n \\ \lambda_1 Q_2^{\top}\mathfrak{X}_1 & Q_2^{\top}\mathbf{A}Q_2 & \dots & Q_2^{\top}\mathbf{A}Q_n \\ \dots & & & \dots \\ \lambda_1 Q_n^{\top}\mathfrak{X}_1 & Q_n^{\top}\mathbf{A}Q_2 & \dots & Q_n^{\top}\mathbf{A}Q_n \end{bmatrix} = \\ &= \begin{pmatrix} \lambda_1 & 0 \ \dots \ 0 \\ 0 & \\ \vdots & \underbrace{\begin{bmatrix} \ \widetilde{\mathbf{A}} \ \end{bmatrix}}_{(n-1)\times(n-1)} \\ 0 & \end{pmatrix} \qquad \text{здесь } \widetilde{\mathbf{A}} = \widetilde{\mathbf{A}}^{\top} \ . \end{aligned}$$

Очевидно, что $\det(\mathbf{A} - \lambda E) = (\lambda_1 - \lambda)\det(\widetilde{\mathbf{A}} - \lambda E)$ и спектр матрицы $\widetilde{\mathbf{A}}$ получается из спектра $\mathbf{A}$ отбрасыванием собственного числа λ_1. По индукционному предположению для матрицы $\widetilde{\mathbf{A}}$ существует ортогональная матрица $\widetilde{Q}_{(n-1)\times(n-1)}$, приводящая ее к диагональному виду:

$$\widetilde{Q}^{\top}\widetilde{\mathbf{A}}\widetilde{Q}=\left(\begin{array}{ccc}\lambda_2 & & \mathbb{O}\\ & \ddots & \\ \mathbb{O} & & \lambda_n\end{array}\right)_{(n-1)\times(n-1)}.$$

Но тогда матрица

$$Q_1 \stackrel{\text{def}}{=} \left(\begin{array}{cccc} 1 & 0 & \dots & 0\\ 0 & & & \\ \vdots & & \widetilde{Q} & \\ 0 & & & \end{array}\right)$$

тоже ортогональная и

$$Q_1^{\top}\left(\begin{array}{cccc} \lambda_1 & 0 & \dots & 0\\ 0 & & & \\ \vdots & & \widetilde{\mathbf{A}} & \\ 0 & & & \end{array}\right)Q_1=\left(\begin{array}{cccc}\lambda_1 & & & \mathbb{O}\\ & \lambda_2 & & \\ & & \ddots & \\ \mathbb{O} & & & \lambda_n\end{array}\right).$$

Матрица $P \stackrel{\text{def}}{=} QQ_1$ является искомой матрицей из формулы (7.2). Она будет ортогональной на основании теоремы 7.3. □

Теорема утверждает что и при наличии кратных корней у хар.полинома

$$f(\lambda)=(-1)^n(\lambda-\lambda_1)^{\mathfrak{m}_1}\times\dots\times(\lambda-\lambda_{\mathfrak{r}})^{\mathfrak{m}_{\mathfrak{r}}} \quad (\mathfrak{m}_1+\dots+\mathfrak{m}_{\mathfrak{r}}=n,\ \lambda_k\ne\lambda_\ell \text{ при } k\ne\ell)$$

симметричной матрицы $\mathbf{A}$ условие диагонализуемости (5.6):

$$\operatorname{dfc}(\mathbf{A}-\lambda_j E)=\mathfrak{m}_j$$

будет выполнено. Иначе говоря, размерность подпространства

$$\left\{X\in\mathbb{R}^n \,\middle|\, (\mathbf{A}-\lambda_j E)X=\mathbb{O}_{n\times 1}\right\}$$

равна $\mathfrak{m}_j$. При нахождении **ф.с.р.** указанной системы уравнений мы получим $\mathfrak{m}_j$ **л.н.з.** собственных векторов $\mathfrak{X}_{j1},\dots,\mathfrak{X}_{j\mathfrak{m}_j}$, принадлежащих λ_j. Однако при традиционном способе построения **ф.с.р.** вовсе не гарантирована ортогональность этих векторов. Как построить **ф.с.р.**, так, чтобы она удовлетворяла условию теоремы 7.4, т.е. была ортонормированной? Воспользуемся для этого процессом ортогонализации из §2 главы 2, применив его к системе $\{\mathfrak{X}_{j1},\dots,\mathfrak{X}_{j\mathfrak{m}_j}\}$. Результатом процесса будет новая система векторов $\{\mathfrak{Y}_{j1},\dots,\mathfrak{Y}_{j\mathfrak{m}_j}\}$ таких что

$$\mathcal{L}\left(\mathfrak{Y}_{j1},\dots,\mathfrak{Y}_{j\mathfrak{m}_j}\right)=\mathcal{L}\left(\mathfrak{X}_{j1},\dots,\mathfrak{X}_{j\mathfrak{m}_j}\right) \quad \text{и} \quad (\mathfrak{Y}_{jk},\mathfrak{Y}_{j\ell})=0 \text{ при } k\ne\ell\ .$$

Нормировав новые векторы, мы получим требуемую теоремой систему из $\mathfrak{m}_j$ ортогональных столбцов матрицы P, соответствующих кратному собств. числу λ_j.

Пример 7.1. *Найти ортогональную матрицу, приводящую матрицу*

$$\mathbf{A} = \begin{pmatrix} 2 & 2 & -1 \\ 2 & -1 & 2 \\ -1 & 2 & 2 \end{pmatrix}$$

к диагональному виду.

Решение. Хар.полином $f(\lambda) = -(\lambda-3)^2(\lambda+3)$. Простому собств.числу $\lambda = -3$ соответствует собств.вектор $\mathfrak{X}_1 = [1,-2,1]^\top$, собств.числу $\lambda = 3$ второй кратности соответствуют два собств.вектора $\mathfrak{X}_2 = [2,1,0]^\top$ и $\mathfrak{X}_3 = [-1,0,1]^\top$. Очевидно $\mathfrak{X}_1 \perp \mathfrak{X}_2, \mathfrak{X}_1 \perp \mathfrak{X}_3$, но $(\mathfrak{X}_2, \mathfrak{X}_3) \neq 0$. Ортогонализуем систему векторов $\{\mathfrak{X}_2, \mathfrak{X}_3\}$:

$$\mathfrak{Y}_2 = \mathfrak{X}_2, \mathfrak{Y}_3 = \mathfrak{X}_3 + \alpha \mathfrak{Y}_2 \quad \text{и} \ (\mathfrak{Y}_2, \mathfrak{Y}_3) = 0 \implies \alpha = -\frac{(\mathfrak{X}_2, \mathfrak{X}_3)}{(\mathfrak{X}_2, \mathfrak{X}_2)} = \frac{2}{5}$$

и $\mathfrak{Y}_3 = [-1/5, 2/5, 1]^\top$. После нормирования, получаем ортогональную матрицу

$$P = \begin{pmatrix} 1/\sqrt{6} & 2/\sqrt{5} & -1/\sqrt{30} \\ -2/\sqrt{6} & 1/\sqrt{5} & 2/\sqrt{30} \\ 1/\sqrt{6} & 0 & 5/\sqrt{30} \end{pmatrix} \quad \begin{matrix} \text{приводящую} \\ \mathbf{A} \\ \text{к виду} \end{matrix} \quad \begin{pmatrix} -3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

$\triangle$

Замечание. Поскольку **ф.с.р.** для кратного собств.числа можно строить разными способами, то у последней задачи имеется бесконечное множество ответов. Так, например, в качестве еще одной ортогональной матрицы можно взять

$$P = \begin{pmatrix} 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \\ -2/\sqrt{6} & 0 & 1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \end{pmatrix}.$$

Замечание. В §3 главы 6 раздела I вводилось понятие конгруэнтности симметричных матриц:

$$\mathbf{A} \cong \mathbf{B} \quad \iff \quad \exists\, C \in \mathbf{GL}(n, \mathbb{R}) \ : C^\top \mathbf{A} C = \mathbf{B} \ .$$

На с.72 было введено понятие подобия произвольных матриц:

$$\mathbf{A} \doteq \mathbf{B} \quad \iff \quad \exists\, C \in \mathbf{GL}(n, \mathbb{R}) \ : C^{-1} \mathbf{A} C = \mathbf{B} \ .$$

Два этих, вообще говоря различных, понятия для симметричных матриц идейно близки: при ортогональности матрицы C конгруэнтность эквивалентна подобию. Для прояснения этого переформулируем теорему 7.4 в терминах преобразований кв. форм.

Следствие 1. *Существует ортогональная замена переменных $X = PY$, приводящая кв.форму $\mathfrak{f}(X) = X^\top \mathbf{A} X$ к каноническому виду:*

$$X^\top \mathbf{A} X = Y^\top P^\top \mathbf{A} P Y = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 \ .$$

Итак, наряду с методом Лагранжа приведения кв. формы к каноническому виду[22], существует еще один метод решения той же задачи, даваемый теоремой 7.4. Если в методе Лагранжа матрица C была, как правило, верхнетреугольной, то в последнем методе она является ортогональной. Метод Лагранжа имеет значительно большее практическое значение, т.к. не требует явного определения собств. чисел; более того, с его помощью можно получить информацию о знаках этих чисел. Действительно, на основании закона инерции кв. форм, число ненулевых λ_j должно равняться $\mathfrak{r} \stackrel{\text{def}}{=} \operatorname{rank}(\mathfrak{f}) = \operatorname{rank}\mathbf{A}$, а число положительных (отрицательных) — положительному (отрицательному) индексу инерции $n_+(\mathfrak{f})$ ($n_-(\mathfrak{f})$). Для определения этих характеристик кв. формы можно воспользоваться формулой Якоби.

Теорема 7.5. *Если главные миноры $\mathbf{A}_1,\dots,\mathbf{A}_{\mathfrak{r}}$ симметричной матрицы $\mathbf{A}$ все ненулевые, то число ее положительных (отрицательных) собств. чисел равно*

$$\mathcal{P}(1,\mathbf{A}_1,\mathbf{A}_2,\dots,\mathbf{A}_{\mathfrak{r}}) \quad (\text{соответственно } \mathcal{V}(1,\mathbf{A}_1,\mathbf{A}_2,\dots,\mathbf{A}_{\mathfrak{r}})) \ .$$

Здесь $\mathcal{P}$ — число знакопостоянств, а $\mathcal{V}$ — число знакоперемен в указанной последовательности.

Еще один способ определения числа положительных и отрицательных собств. чисел симметричной матрицы $\mathbf{A}$ дает правило знаков Декарта[23].

Упражнение 7.3. *Доказать, что если $(-1)^n\left(\lambda^n + a_1\lambda^{n-1} + \dots + a_n\right)$ хар. полином (симметричной) матрицы $\mathbf{A}$ кв. формы $\mathfrak{f}(X) = X^{\top}\mathbf{A}X$, то*

$$\textbf{а)}\ \operatorname{rank}\mathfrak{f} = \mathfrak{r} \iff a_n = a_{n-1} = \dots = a_{\mathfrak{r}+1} = 0, a_{\mathfrak{r}} \neq 0 \ .$$

В этом случае будет также выполнено

$$\textbf{б)}\ n_+(\mathfrak{f}) = \mathcal{V}(1,a_1,\dots,a_{\mathfrak{r}}), \quad n_-(\mathfrak{f}) = \mathcal{P}(1,a_1,\dots,a_{\mathfrak{r}}) \ .$$

7.3 Локализация собств.чисел

В пункте 6.3 был указан способ решения задачи локализации вещественных собств. чисел произвольной вещественной матрицы. Для симметричной матрицы решение этой задачи упрощается.

Теорема 7.6 (Коши). *Для вещественной симметричной матрицы $\mathbf{A}$ число ее собственных чисел, лежащих на интервале $]a,b[$ равно*

$$\mathcal{P}(1,\mathbf{H}_1(a),\mathbf{H}_2(a),\dots,\mathbf{H}_n(a)) - \mathcal{P}(1,\mathbf{H}_1(b),\mathbf{H}_2(b),\dots,\mathbf{H}_n(b)) \ .$$

Здесь $\mathbf{H}_1(\lambda),\mathbf{H}_2(\lambda),\dots,\mathbf{H}_n(\lambda)$ — главные миноры матрицы $\mathbf{A}-\lambda E$.

[22] Раздел I, глава 7, §1.
[23] Раздел I, глава 3, §7.

Доказательство . Пусть по-прежнему, $f(\lambda) \stackrel{\text{def}}{=} \det(\mathbf{A} - \lambda\, E)$. Имеем:

$$\operatorname{nrr}\left\{f(\lambda)=0 \mid \lambda>t\right\} \stackrel{\lambda=\mu+t}{=} \operatorname{nrr}\left\{\widetilde{f}(\mu) \stackrel{\text{def}}{=} f(\mu+t)=0 \mid \mu>0\right\} \ .$$

С другой стороны $\widetilde{f}(\mu)$ — это хар. полином матрицы $\mathbf{A} - t\, E$, т.к.

$$\det\left((\mathbf{A}-t\,E)-\mu\,E\right)=\det\left(\mathbf{A}-(t+\mu)\,E\right)=f(\mu+t)=\widetilde{f}(\mu) \ .$$

На основании теоремы 7.5:

$$\operatorname{nrr}\left\{\widetilde{f}(\mu)=0 \mid \mu>0\right\}=n_{+}(\mathbf{A}-t\,E)=\mathcal{P}\left(1,\mathbf{H}_1(t),\mathbf{H}_2(t),\dots,\mathbf{H}_n(t)\right)$$

и тогда

$$\begin{array}{rcl}\operatorname{nrr}\left\{f(\lambda)=0 \mid \lambda>a\right\} &=& \mathcal{P}\left(1,\mathbf{H}_1(a),\mathbf{H}_2(a),\dots,\mathbf{H}_n(a)\right) \ , \\ \operatorname{nrr}\left\{f(\lambda)=0 \mid \lambda>b\right\} &=& \mathcal{P}\left(1,\mathbf{H}_1(b),\mathbf{H}_2(b),\dots,\mathbf{H}_n(b)\right) \ .\end{array}$$

Разность двух последних чисел даст требуемое $\operatorname{nrr}\left\{f(\lambda)=0 \mid a<\lambda<b\right\}$.
□

Итак, главные миноры матрицы $\mathbf{A} - \lambda\, E$ играют роль системы полиномов Штурма для хар. полинома симметричной матрицы $\mathbf{A}$.

Пример 7.2. *Локализовать собственные числа матриц*

$$\textbf{а)}\left(\begin{array}{rrr} 11 & 2 & -8 \\ 2 & 2 & 10 \\ -8 & 10 & 5 \end{array}\right) \quad ; \quad \textbf{б)}\left(\begin{array}{rrr} 1 & -2 & 2 \\ -2 & -2 & 4 \\ 2 & 4 & -2 \end{array}\right) \ .$$

Решение. **а)**

$\mathbf{H}_1(\lambda)=11-\lambda,\ \mathbf{H}_2(\lambda)=\lambda^2-13\,\lambda+18,\ f(\lambda)=\mathbf{H}_3(\lambda)=-\lambda^3+18\,\lambda^2+81\,\lambda-1458$.

λ	1	$\mathbf{H}_1(\lambda)$	$\mathbf{H}_2(\lambda)$	$\mathbf{H}_3(\lambda)$	$\mathcal{P}$	Комментарии
0	+	+	+	−	2	число положит. =2
−10	+	+	+	+	3	собств. число
−5	+	+	+	−	2	лежит на $]-10,-5[$
5	+	+	−	−	2	собств. число
10	+	+	−	+	1	лежит на $]5,10[$
15	+	−	−	+	1	собств. число
20	+	−	+	−	0	лежит на $]15,20[$

б)

$\mathbf{H}_1(\lambda)=1-\lambda,\ \mathbf{H}_2(\lambda)=\lambda^2+\lambda-6,\ f(\lambda)=\mathbf{H}_3(\lambda)=-\lambda^3-3\,\lambda^2+24\,\lambda-28$.

λ	1	$\mathbf{H}_1(\lambda)$	$\mathbf{H}_2(\lambda)$	$\mathbf{H}_3(\lambda)$	$\mathcal{P}$	Комментарии
0	+	+	−	−	2	число положит. =2
−8	+	+	+	+	3	собств. число
−6	+	+	+	−	2	лежит на $]-8,-6[$
1.5	+	−	−	−	2	два собств. числа
3	+	−	+	−	0	лежат на $]1.5,3[$

Никаким дроблением интервала $]1.5\,,\,3[$ не удается отделить два вещественных собств. числа. Вывод: имеется кратное собств. число. $\triangle$

Проверка. Спектры матриц **а)** $\{-9, 9, 18\}$; **б)** $\{-7, 2, 2\}$.

7.4 Экстремальное свойство собств.чисел

Задача. Найти условные экстремумы квадратичной формы $\mathfrak{f}(X) = X^{\top}\mathbf{A}X$ на единичной сфере $\mathbb{S} : x_1^2 + \dots + x_n^2 = 1$:

$$\max_{\mathbb{S}} X^{\top}\mathbf{A}X \quad \text{и} \quad \min_{\mathbb{S}} X^{\top}\mathbf{A}X \ .$$

В курсе математического анализа показывается, что, во-первых, указанные экстремумы существуют[24], и, во-вторых, могут быть найдены применением метода множителей Лагранжа.

Теорема 7.7. *Имеем:*

$$\max_{\mathbb{S}} X^{\top}\mathbf{A}X = \lambda_{\max}, \qquad \min_{\mathbb{S}} X^{\top}\mathbf{A}X = \lambda_{\min} \ .$$

Здесь $\lambda_{\max}$ — максимальное, а $\lambda_{\min}$ — минимальное собственные числа матрицы $\mathbf{A}$. Указанные экстремумы $\mathfrak{f}(X)$ достигает на соответствующих собственных векторах матрицы $\mathbf{A}$ единичной длины.

Доказательство . Применяем метод множителей Лагранжа, т.е. составляем функцию

$$F(X, \lambda) \stackrel{\text{def}}{=} \mathfrak{f}(X) - \lambda(x_1^2 + \dots + x_n^2 - 1)$$

и ищем ее <u>абсолютные</u> экстремумы (как функции $(n+1)$-го аргумента). На основании теоремы 2.0 главы 5, раздела I эти экстремумы должны достигаться на вещественных решениях системы уравнений

$$\left\{ \begin{array}{lll} \partial F/\partial x_1 = & 2\,(a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n) - 2\lambda x_1 & = 0, \\ \dots & & \dots \\ \partial F/\partial x_n = & 2\,(a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n) - 2\lambda x_n & = 0, \\ \partial F/\partial \lambda = & x_1^2 + \dots + x_n^2 - 1 & = 0 \ . \end{array} \right. \tag{7.3}$$

Решаем эту систему. Первые n уравнений перепишем в матричном виде

$$\mathbf{A}X - \lambda X = \mathbb{O} \iff (\mathbf{A} - \lambda E)X = \mathbb{O} \ .$$

Из последнего уравнения системы (7.3) следует, что $X \neq \mathbb{O}$. Следовательно, решениями системы (7.3) будут исключительно только собственные векторы $\mathfrak{X}_j$ матрицы $\mathbf{A}$, при λ равном соответствующему собственному числу λ_j этой матрицы. При $X = \mathfrak{X}_j$ и $\lambda = \lambda_j$ получаем экстремальные значения функции $\mathfrak{f}(X)$:

$$\mathfrak{f}(\mathfrak{X}_j) = \mathfrak{X}_j^{\top}\mathbf{A}\mathfrak{X}_j = \lambda_j \mathfrak{X}_j^{\top}\mathfrak{X}_j = \lambda_j \ .$$

Откуда и следует утверждение теоремы. $\square$

[24] Теорема Вейерштрасса.

Упражнение 7.4. *Пусть уравнение $X^{\top}\mathbf{A}X = 1$ задает n-мерный эллипсоид в $\mathbb{R}^n$. Найти расстояние до него от начала координат $\mathbb{O}$.*

8 Жорданова нормальная форма оператора в $\mathbb{C}$

В настоящем параграфе пространство $\mathbb{V}$ предполагается комплексным.

8.1 Общая схема

Задача. Найти базис пространства $\mathbb{V}$, в котором матрица оператора $\mathscr{A}$ имеет наиболее простой вид.

В §5 мы установили, что если возможно найти базис $\mathbb{V}$, состоящий из собственных векторов оператора, то соответствующая матрица будет диагональной. В частности, существование всегда гарантировано в случае, когда хар. полином оператора имеет только простые корни. Случай наличия кратных корней

$$f(\lambda) = (-1)^n(\lambda-\lambda_1)^{\mathfrak{m}_1}\times\cdots\times(\lambda-\lambda_{\mathfrak{r}})^{\mathfrak{m}_{\mathfrak{r}}} \quad (\mathfrak{m}_1+\cdots+\mathfrak{m}_{\mathfrak{r}} = n,\ \lambda_k \ne \lambda_\ell \text{ при } k \ne \ell) \tag{8.1}$$

при хотя бы одном $\mathfrak{m}_j > 1$ оказывается пограничным: оператор может оказаться и недиагонализуемым (см. примеры на с. 78).

Общая схема действий: пространство $\mathbb{V}$ удается разбить в прямую сумму подпространств

$$\mathbb{V} = \mathbb{V}_1 \oplus \cdots \oplus \mathbb{V}_{\mathfrak{r}}\ , \quad \dim \mathbb{V}_j = \mathfrak{m}_j \tag{8.2}$$

инвариантных относительно $\mathscr{A}$: $\mathscr{A}(\mathbb{V}_j) \subset \mathbb{V}_j$. При этом $\mathbb{V}_j$ обязательно будет включать собственные векторы, принадлежащие λ_j, но, помимо них — в случае когда

$$\mathfrak{m}_j > \ell_j \stackrel{\text{def}}{=} \operatorname{dfc}(\mathscr{A} - \lambda_j\mathscr{E})$$

— и другие: так называемые, **корневые**. На основании теоремы 5.2 в базисе $\mathbb{V}$, составленном объединением базисов $\mathbb{V}_j$ матрица оператора будет иметь блочно–диагональный вид

$$\left(\begin{array}{cccc} \mathsf{A}_1 & \mathbb{O} & \dots & \mathbb{O} \\ \mathbb{O} & \mathsf{A}_2 & \dots & \mathbb{O} \\ & & \ddots & \\ \mathbb{O} & \mathbb{O} & \dots & \mathsf{A}_{\mathfrak{r}} \end{array}\right) \quad , \quad \text{здесь } \mathsf{A}_j\text{—матрица } \mathfrak{m}_j \times \mathfrak{m}_j\ . \tag{8.3}$$

Каждый из базисов составляющих $\mathbb{V}$ подпространств $\mathbb{V}_j$ удается подобрать так, чтобы матрица A_j имела снова блочно-диагональный вид

$$\mathsf{A}_j = \left(\begin{array}{cccc} \mathbf{A}_{j1} & \mathbb{O} & \dots & \mathbb{O} \\ \mathbb{O} & \mathbf{A}_{j2} & \dots & \mathbb{O} \\ & & \ddots & \\ \mathbb{O} & \mathbb{O} & \dots & \mathbf{A}_{j\ell_j} \end{array}\right)$$

где на диагонали стоят матрицы вида

$$\mathfrak{J}_k(\lambda_j) \stackrel{\text{def}}{=} \begin{pmatrix} \lambda_j & 0 & 0 & \dots & 0 \\ 1 & \lambda_j & 0 & \dots & 0 \\ 0 & 1 & \lambda_j & \dots & 0 \\ & & \ddots & \ddots & \\ 0 & 0 & \dots & 1 & \lambda_j \end{pmatrix}_{k\times k}$$

называемые (нижними) **клетками Жордана**.

ОПРЕДЕЛЕНИЕ. Указанный вид матрицы оператора $\mathscr{A}$ называется канонической **формой Жордана** или **жордановой нормальной формой**, а соответствующий базис пространства — **каноническим базисом**.

Частным видом **ж.н.ф.** является диагональный: когда все клетки Жордана — первого порядка.

ЗАМЕЧАНИЕ. Даже при формальном совпадении хар.полиномов двух операторов $\mathscr{A}_1$ и $\mathscr{A}_2$ их **ж.н.ф.** могут быть различными. Однако для каждого оператора **ж.н.ф.** определяется единственным образом — с точностью до перестановки клеток на диагонали.

8.2 Аннулирующий полином

Пусть $g(x), g_1(x), g_2(x)$ — произвольные полиномы над $\mathbb{C}$.

Теорема 8.1. *Если* НОД $(g_1, g_2) = 1$*, то существуют полиномы* $p_1(x), p_2(x) \in \mathbb{C}[x]$ *такие, что*

$$p_1(\mathscr{A})g_1(\mathscr{A}) + p_2(\mathscr{A})g_2(\mathscr{A}) = \mathscr{E} \ . \tag{8.4}$$

Доказательство . В §4.2 главы 3 раздела I доказывалось, что в предположении НОД $(g_1, g_2) = 1$ существуют $\{p_1(x), p_2(x)\} \subset \mathbb{C}[x]$, удовлетворяющие тождеству Безу

$$p_1(x)g_1(x) + p_2(x)g_2(x) \equiv 1 \ .$$

Легко видеть, что это тождество должно выполняться и при замене x на $\mathscr{A}$, поскольку операторные полиномы коммутируют. □

ОПРЕДЕЛЕНИЕ. Говорят, что операторный полином $g(\mathscr{A})$ — **аннулирующий** для вектора $X \in \mathbb{V}$ если $g(\mathscr{A})(X) = \mathbb{O}$.

Очевидно, что множество векторов $X \in \mathbb{V}$ аннулируемых $g(\mathscr{A})$ образует линейное подпространство: $\mathcal{K}er\, g(\mathscr{A})$.

Теорема 8.2. *Если* НОД $(g_1, g_2) = 1$*, то подпространства векторов, аннулируемых* $g_1(\mathscr{A})$ *и* $g_2(\mathscr{A})$ *имеют тривиальное пересечение.*

Доказательство . Если существует $X \in \mathbb{V}$ такой, что $g_1(\mathscr{A})(X) = \mathbb{O}$ и $g_2(\mathscr{A})(X) = \mathbb{O}$, то из равенства (8.4) следует, что

$$p_1(\mathscr{A})g_1(\mathscr{A})(X) + p_2(\mathscr{A})g_2(\mathscr{A})(X) = \mathscr{E}(X) \quad \Longrightarrow \mathbb{O} = X \ .$$

□

Теорема 8.3. *Если вектор $X \neq \mathbb{O}$ аннулируется произведением $g_1(\mathscr{A})g_2(\mathscr{A})$ где* НОД $(g_1, g_2) = 1$, *то его можно представить в виде суммы $X = X_1 + X_2$, где X_j аннулируется $g_j(\mathscr{A})$.*

Доказательство . Применим операторное равенство (8.4) к вектору X:

$$\underbrace{p_1(\mathscr{A})g_1(\mathscr{A})(X)}_{\stackrel{\text{def}}{=} X_2} + \underbrace{p_2(\mathscr{A})g_2(\mathscr{A})(X)}_{\stackrel{\text{def}}{=} X_1} = \mathscr{E}(X) \ .$$

Тогда $g_2(\mathscr{A})(X_2) = p_1(\mathscr{A})g_1(\mathscr{A})g_2(\mathscr{A})(X) = p_1(\mathscr{A})(\mathbb{O}) = \mathbb{O}$, т.е. X_2 аннулируется $g_2(\mathscr{A})$. Аналогично доказывается, что $g_1(\mathscr{A})(X_1) = \mathbb{O}$. $\square$

Следствие 1. *Если вектор X аннулируется произведением $g(\mathscr{A}) = g_1(\mathscr{A})\times \dots \times g_{\mathfrak{r}}(\mathscr{A})$ где полиномы $g_1(x), \dots, g_{\mathfrak{r}}(x)$ взаимно просты, то его можно представить в виде суммы $X = X_1 + \dots + X_{\mathfrak{r}}$, где X_j аннулируется $g_j(\mathscr{A})$.*

Рассмотрим теперь хар. полином $f(\lambda)$ оператора $\mathscr{A}$. На основании теоремы 6.4 (Гамильтона–Кэли) он обладает тем свойством, что $f(\mathscr{A}) = \mathcal{O}$, т.е. операторный полином $f(\mathscr{A})$ аннулирует любой вектор $X \in \mathbb{V}$. Если воспользоваться разложением (8.1) то последнее следствие утверждает, что этот вектор X может быть представлен в виде суммы

$$X = X_1 + \dots + X_{\mathfrak{r}} \ , \qquad \text{где } X_j \quad \text{аннулируется } (\mathscr{A} - \lambda_j\mathscr{E})^{\mathfrak{m}_j} \ .$$

На основании теоремы 8.1 и теоремы 4.2 главы 1 можно утверждать, что такое представление единственно, т.е. $\mathbb{V}$ разлагается в прямую сумму (8.2) где $\mathbb{V}_j$ аннулируется $(\mathscr{A} - \lambda_j\mathscr{E})^{\mathfrak{m}_j}$.

Теорема 8.4. *Линейное подпространство векторов аннулируемых $(\mathscr{A} - \lambda_j\mathscr{E})^{\mathfrak{m}_j}$ инвариантно относительно $\mathscr{A}$.*

Доказательство . Надо доказать, что если X принадлежит подпространству, то и $\mathscr{A}(X)$ принадлежит тому же подпространству. В самом деле:

$$(\mathscr{A} - \lambda_j\mathscr{E})^{\mathfrak{m}_j}(X) = \mathbb{O} \Longrightarrow (\mathscr{A} - \lambda_j\mathscr{E})^{\mathfrak{m}_j}(\mathscr{A}(X)) = \mathscr{A}\left((\mathscr{A} - \lambda_j\mathscr{E})^{\mathfrak{m}_j}(X)\right) = \mathbb{O} \ ,$$

т.е. $\mathscr{A}(X)$ также аннулируется $(\mathscr{A} - \lambda_j\mathscr{E})^{\mathfrak{m}_j}$. $\square$

Итак, на основании теоремы 5.2 и теоремы 5.2 главы I можно утверждать, что если базис $\mathbb{V}$ составить объединением базисов $\mathbb{V}_j$, то в нем матрица оператора $\mathscr{A}$ будет иметь блочно-диагональный вид (8.3).

8.3 Корневые векторы

Теперь наша задача — построить такой базис $\mathbb{V}_j$, в котором составляющий блок $\mathbf{A}_j$ будет состоять из клеток Жордана.

Определение. Вектор $X \in \mathbb{V}$ называется **корневым** для оператора $\mathscr{A}$, принадлежащим собственному числу λ_j если он аннулируется операторным полиномом $(\mathscr{A} - \lambda_j\mathscr{E})^k$ при некотором $k \in \mathbb{N}$: $(\mathscr{A} - \lambda_j\mathscr{E})^k(X) = \mathbb{O}$.

Наименьший из показателей k с таким свойством называется **высотой** корневого вектора X:

$$\text{Высота } (X) = h \iff (\mathscr{A} - \lambda_j \mathscr{E})^h(X) = \mathbb{O},\ (\mathscr{A} - \lambda_j \mathscr{E})^{h-1}(X) \neq \mathbb{O} \ .$$

Для определенности будем считать высоту $\mathbb{O}$ равной 0.

Пример 8.1. *Любой собственный вектор оператора $\mathscr{A}$ будет его корневым высоты* 1.

Теорема 8.5. *Высота корневого вектора, принадлежащего λ_j не превосходит кратности этого числа в хар. полиноме.*

Доказательство . Пусть $\exists X \in \mathbb{V}$ такой, что

$$(\mathscr{A} - \lambda_j \mathscr{E})^h(X) = \mathbb{O}, \quad (\mathscr{A} - \lambda_j \mathscr{E})^{\mathfrak{m}_j}(X) \neq \mathbb{O}$$

при $h > \mathfrak{m}_j$. Обозначим

$$\widetilde{X} \stackrel{\text{def}}{=} (\mathscr{A} - \lambda_j \mathscr{E})^{\mathfrak{m}_j}(X), \quad g(\lambda) \stackrel{\text{def}}{=} f(\lambda)/(\lambda - \lambda_j)^{\mathfrak{m}_j} \ .$$

По определению, вектор $\widetilde{X}$ — корневой, принадлежащий λ_j, высоты $h - \mathfrak{m}_j$. Поскольку $g(\lambda)$ не имеет корнем λ_j, то $\mathsf{НОД}\,(g(\lambda), (\lambda - \lambda_j)^{h-\mathfrak{m}_j}) = 1$. По теореме 8.2: $g(\mathscr{A})(\widetilde{X}) \neq \mathbb{O}$. Но тогда

$$g(\mathscr{A})(\mathscr{A} - \lambda_j \mathscr{E})^{\mathfrak{m}_j}(X) \neq \mathbb{O} \Longrightarrow f(\mathscr{A})(X) \neq \mathbb{O} \ ,$$

что противоречит тому, что $f(\mathscr{A}) = \mathcal{O}$. $\square$

Теорема 8.6. *Корневые векторы, принадлежащие λ_j образуют линейное подпространство $\mathbb{V}_j$. При этом $\mathbb{V}_j \bigcap \mathbb{V}_k = \{\mathbb{O}\}$ если $j \neq k$.*

Следствие 1. *Пространство $\mathbb{V}$ разлагается в прямую сумму корневых подпространств: $\mathbb{V} = \mathbb{V}_1 \oplus \cdots \oplus \mathbb{V}_{\mathfrak{r}}$.*

Для построения базиса корневого подпространства $\mathbb{V}_j$ выделим в нем подпространства корневых векторов высот $\le s$:

$$\mathbb{Q}_s \stackrel{\text{def}}{=} \mathcal{K}er(\mathscr{A} - \lambda_j\, \mathscr{E})^s \ , \ \mathbb{Q}_0 \stackrel{\text{def}}{=} \{\mathbb{O}\} \ .$$

Понятно, что имеет место вложенность

$$\mathbb{Q}_0 \subset \mathbb{Q}_1 \subset \cdots \subset \mathbb{Q}_p = \mathbb{V}_j \quad (p \le \mathfrak{m}_j) \ .$$

Теорема 8.7. *Если векторы $X_1, \dots, X_k$ принадлежат $\mathbb{Q}_s$ и* **л.н.з.** *относительно $\mathbb{Q}_{s-1}$, то $(\mathscr{A} - \lambda_j \mathscr{E})(X_1), \dots, (\mathscr{A} - \lambda_j \mathscr{E})(X_k)$ принадлежат $\mathbb{Q}_{s-1}$ и* **л.н.з.** *относительно $\mathbb{Q}_{s-2}$.*

Доказательство . Если $X \in \mathbb{Q}_s$ то $(\mathscr{A} - \lambda_j\mathscr{E})^s(X) = \mathbb{O}$, т.е.

$$(\mathscr{A} - \lambda_j\mathscr{E})^{s-1}\left((\mathscr{A} - \lambda_j\mathscr{E})(X)\right) = \mathbb{O}\ ,$$

но это и означает, что $(\mathscr{A} - \lambda_j\mathscr{E})(X) \in \mathbb{Q}_{s-1}$.

Предположим теперь, что существуют скаляры $c_1, \dots, c_k$ такие, что

$$\begin{array}{lll} & c_1(\mathscr{A} - \lambda_j\mathscr{E})(X_1) + \dots + c_k(\mathscr{A} - \lambda_j\mathscr{E})(X_k) \in \mathbb{Q}_{s-2} & \Longleftrightarrow \\ \Longleftrightarrow (\mathscr{A} - \lambda_j\mathscr{E})^{s-2} & (c_1(\mathscr{A} - \lambda_j\mathscr{E})(X_1) + \dots + c_k(\mathscr{A} - \lambda_j\mathscr{E})(X_k)) = \mathbb{O} & \Longleftrightarrow \\ (\mathscr{A} - \lambda_j\mathscr{E})^{s-1} & (c_1X_1 + \dots + c_kX_k) = \mathbb{O}\ . & \end{array}$$

По условию теоремы последнее соотношение возможно только при $c_1 = 0, \dots, c_k = 0$. $\square$

Искомый базис $\mathbb{V}_j$ строим следующим алгоритмом. Обозначим

$$\mathscr{B} \stackrel{\text{def}}{=} \mathscr{A} - \lambda_j\mathscr{E}$$

1. Выберем $X_{11}, \dots, X_{1k_1}$ — относительный базис $\mathbb{Q}_p = \mathbb{V}_j$ над $\mathbb{Q}_{p-1}$.
2. По теореме 8.7 векторы $\mathscr{B}(X_{11}), \dots, \mathscr{B}(X_{1k_1})$ принадлежат $\mathbb{Q}_{p-1}$ и **л.н.з.** относительно $\mathbb{Q}_{p-2}$. Дополним эти векторы до относительного базиса $\mathbb{Q}_{p-1}$ над $\mathbb{Q}_{p-2}$: пусть

   $$\mathscr{B}(X_{11}), \dots, \mathscr{B}(X_{1k_1}), X_{21}, \dots, X_{2k_2}$$

   этот базис.
3. По теореме 8.7 векторы

   $$\mathscr{B}^2(X_{11}), \dots, \mathscr{B}^2(X_{1k_1}), \mathscr{B}(X_{21}), \dots, \mathscr{B}(X_{2k_2})$$

   принадлежат $\mathbb{Q}_{p-2}$ и **л.н.з.** относительно $\mathbb{Q}_{p-3}$. Дополним эти векторы до относительного базиса $\mathbb{Q}_{p-2}$ над $\mathbb{Q}_{p-3}$ и т.д.
4. … p-1.

p. На p-м шаге получим векторы

   $$\mathscr{B}^{p-1}(X_{11}), \dots, \mathscr{B}^{p-1}(X_{1k_1}), \mathscr{B}^{p-2}(X_{21}), \dots, \mathscr{B}^{p-2}(X_{2k_2}), \dots,$$
   $$\mathscr{B}(X_{p-1,1}), \dots, \mathscr{B}(X_{p-1,k_{p-1}}),$$

   принадлежащие $\mathbb{Q}_1$ и **л.н.з.** относительно $\mathbb{Q}_0$, т.е. **л.н.з.** в обычном понимании. Дополним эти векторы до базиса $\mathbb{Q}_1$: пусть

   $$\mathscr{B}^{p-1}(X_{11}), \dots, \mathscr{B}^{p-1}(X_{1k_1}), \mathscr{B}^{p-2}(X_{21}), \dots, \mathscr{B}^{p-2}(X_{2k_2}), \dots,$$
   $$\mathscr{B}(X_{p-1,1}), \dots, \mathscr{B}(X_{p-1,k_{p-1}}), X_{p1}, \dots, X_{pk_p}$$

   этот базис.

Теорема 8.8. *Справедливо равенство:*

$$k_1 + k_2 + \cdots + k_p = \mathrm{dfc}\,(\mathscr{A} - \lambda_j \mathscr{E}) \quad ,$$

т.е. равно числу **л.н.з.** *собственных векторов, принадлежащих* λ_j.

Доказательство . Действительно, $\mathbb{Q}_1$, по определению, состоит из собственных векторов оператора $\mathscr{A}$, принадлежащих λ_j. □

Теорема 8.9. *Базис* $\mathbb{V}_j$ *получается объединением всех векторов, встречаемых в алгоритме.*

Доказательство . Действительно

$$\begin{array}{rcl} \text{базис } \mathbb{Q}_2 & = & \{\text{базис } \mathbb{Q}_1\} \bigcup \{\text{относит. базис } \mathbb{Q}_2 \text{ над } \mathbb{Q}_1\} \quad , \\ \text{базис } \mathbb{Q}_3 & = & \{\text{базис } \mathbb{Q}_2\} \bigcup \{\text{относит. базис } \mathbb{Q}_3 \text{ над } \mathbb{Q}_2\} \quad , \\ \dots & & \dots \\ \text{базис } \underbrace{\mathbb{Q}_p}_{=\mathbb{V}_j} & = & \big\{\text{базис } \mathbb{Q}_{p-1}\big\} \bigcup \big\{\text{относит. базис } \mathbb{Q}_p \text{ над } \mathbb{Q}_{p-1}\big\} \quad . \end{array}$$

Теорема 8.10. *Размерность корневого подпространства равна кратности собственного числа в хар. полиноме:* $\dim \mathbb{V}_j = \mathfrak{m}_j$.

Нам осталось выяснить каким образом преобразуются векторы построенного базиса под действием оператора $\mathscr{A}$ — тогда мы сможем построить матрицу.

8.4 Циклическое подпространство

Пусть в пространстве $\mathbb{V}$ действует оператор $\mathscr{A}_1$. Для любого $X \in \mathbb{V}$ построим наименьшее инвариантное подпространство, содержащее X. Рассмотрим последовательность $X, \mathscr{A}_1(X), \mathscr{A}_1(\mathscr{A}_1(X)) = \mathscr{A}_1^2(X), \dots ,$ и продолжим ее до тех пор, пока не возникнет линейная зависимость: пусть $X, \mathscr{A}_1(X), \dots, \mathscr{A}_1^{k-1}(X)$ еще линейно независимы, а $X, \mathscr{A}_1(X), \dots, \mathscr{A}_1^{k-1}(X)$, $\mathscr{A}_1^{k}(X)$ уже линейно зависимы. Следовательно,

$$\mathscr{A}_1^k(X) = -\alpha_1 X - \alpha_2 \mathscr{A}_1(X) - \ldots - \alpha_k \mathscr{A}_1^{k-1}(X). \tag{8.5}$$

Теорема 8.11. *Линейная оболочка векторов* $X, \mathscr{A}_1(X), \dots, \mathscr{A}_1^{k-1}(X)$

$$\widetilde{\mathbb{V}} \stackrel{\mathrm{def}}{=} \mathcal{L}(X, \mathscr{A}_1(X), \dots, \mathscr{A}_1^{k-1}(X)) \tag{8.6}$$

является инвариантным подпространством. При этом $\widetilde{\mathbb{V}}$ *будет минимальным инвариантным подпространством, содержащим* X, *т.е. если* $\widetilde{\widetilde{\mathbb{V}}}$ *— произвольное инвариантное подпространство, содержащее* X, *то* $\widetilde{\widetilde{\mathbb{V}}} \supset \widetilde{\mathbb{V}}$.

Доказательство . Рассмотрим произвольный вектор из $\widetilde{\mathbb{V}}$:

$$
\begin{array}{rcl}
Y & = & c_1X + c_2\mathscr{A}_1(X) + \ldots + c_k\mathscr{A}_1^{k-1}(X), \\
\mathscr{A}_1(Y) & = & c_1\mathscr{A}_1(X) + c_2\mathscr{A}_1^2(X) + \ldots + c_k\mathscr{A}_1^k(X) = \\
& \overset{(8.5)}{=} & -\alpha_1c_kX + (c_1 - \alpha_2c_k)\mathscr{A}_1(X) + \cdots + (c_{k-1} - \alpha_kc_k)\mathscr{A}_1^{k-1}(X) \ \in \widetilde{\mathbb{V}}
\end{array}
$$

т.к. все слагаемые принадлежат $\widetilde{\mathbb{V}}$.

Если $\widetilde{\widetilde{\mathbb{V}}}$ — еще какое-то инвариантное подпространство, содержащее X, то оно должно содержать и $\mathscr{A}(X)$, но тогда — и $\mathscr{A}(\mathscr{A}(X)) = \mathscr{A}^2(X)$ и т.д., а, значит, и $\widetilde{\mathbb{V}}$. □

Определение. Подпространство (8.6) называется **циклическим** подпространством, порожденным вектором X.

Рассмотрим теперь оператор $\mathscr{A}_1 \overset{\text{def}}{=} \mathscr{B} = \mathscr{A} - \lambda_j\mathscr{E}$ и его циклическое подпространство, порожденное произвольным корневым вектором X высоты s оператора $\mathscr{A}$.

Теорема 8.12. *В базисе пространства* $\mathbb{V}$*, составленном дополнением векторов*

$$
Y_1 = X,\, Y_2 = \mathscr{B}(Y_1) = \mathscr{B}(X),\, Y_3 = \mathscr{B}(Y_2) = \mathscr{B}^2(X), \ldots,
$$

$$
Y_s = \mathscr{B}(Y_{s-1}) = \mathscr{B}^{\,s-1}(X) \ ,
$$

матрица оператора $\mathscr{A}$ *имеет вид:*

$$
\begin{pmatrix}
\lambda_j & 0 & 0 & \ldots & 0 & * & * & * \\
1 & \lambda_j & 0 & \ldots & 0 & * & * & * \\
0 & 1 & \lambda_j & & 0 & * & * & * \\
\vdots & & \ddots & \ddots & & & & \vdots \\
0 & 0 & \ldots & 1 & \lambda_j & * & * & * \\
 & & & & & * & * & * \\
 & & \mathbb{O} & & & & \ldots & \\
 & & & & & * & * & *
\end{pmatrix}
$$

Доказательство . Действительно, $\mathscr{A} = \mathscr{B} + \lambda_j\mathscr{E}$ и тогда

$$
\mathscr{A}(Y_1) = \mathscr{B}(Y_1) + \lambda_j\mathscr{E}(Y_1) = \lambda_jY_1 + Y_2,\ \mathscr{A}(Y_2) = \mathscr{B}(Y_2) + \lambda_j\mathscr{E}(Y_2) = \lambda_jY_2 + Y_3, \ldots,
$$

$$
\mathscr{A}(Y_s) = \mathscr{B}(Y_s) + \lambda_j\mathscr{E}(Y_s) = \lambda_jY_s,
$$

($\mathscr{B}(Y_s) = \mathscr{B}^s(X) = \mathbb{O}$ поскольку, по условию, X — корневой высоты s). □

Эта теорема устанавливает структуру матрицы оператора в базисе, построенном в предыдущем пункте.

Пример 8.2. *Для матрицы*

$$\mathbf{A} = \begin{pmatrix} -1 & 0 & -9 & 0 & 0 & 0 \\ 0 & 3 & 0 & 1 & 0 & -1 \\ 1 & 0 & 5 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 2 & 0 & 0 & 0 & 2 \end{pmatrix}$$

построить **ж.н.ф.** *и матрицу C, к ней приводящую.*

Решение. **1.** Вычисляем хар. полином $\det(\mathbf{A}-\lambda\, E) = (\lambda-2)^6$. Он имеет единственный корень $\lambda_1 = 2$ кратности $\mathfrak{m}_1 = 6$.

2. Ищем $\mathbb{Q}_1$, т.е. подпространство корневых векторов высоты 1, принадлежащих λ_1. Для этого составляем матрицу

$$\mathbf{B} = \mathbf{A} - 2\,E = \begin{pmatrix} -3 & 0 & -9 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & -1 \\ 1 & 0 & 3 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \end{pmatrix}$$

и ищем **ф.с.р.** для системы $\mathbf{B}X = \mathbb{O}$. В результате гауссовского исключения приходим к системе

$$\left\{ \begin{array}{lllll} x_1 & & +3x_3 & & = 0 \\ & x_2 & & & = 0 \\ & & x_4 & -x_6 & = 0 \end{array} \right. \quad \Rightarrow \quad \begin{array}{ccc|ccc} \multicolumn{3}{c}{\overbrace{\hphantom{x_1 \; x_2 \; x_4}}^{\text{зависимые}}} & \multicolumn{3}{c}{\overbrace{\hphantom{x_3 \; x_5 \; x_6}}^{\text{основные}}} \\ x_1 & x_2 & x_4 & x_3 & x_5 & x_6 \\ \hline -3 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ \multicolumn{3}{c}{} & \multicolumn{3}{c}{\underbrace{\hphantom{0 \quad 0 \quad 1}}_{E}} \end{array}$$

и $\mathbb{Q}_1 = \mathcal{L}(\mathfrak{e}_5, \mathfrak{e}_4 + \mathfrak{e}_6, -3\mathfrak{e}_1 + \mathfrak{e}_3)$.

Вывод. Собственному числу $\lambda_1 = 2$ соответствуют $\ell_1 = 3$ клетки Жордана в **ж.н.ф.** Матрица $\mathbf{A}$ недиагонализуема.

3. Ищем $\mathbb{Q}_2$, т.е. подпространство корневых векторов высоты ≤ 2, принадлежащих λ_1. Для этого вычисляем матрицу

$$\mathbf{B}^2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 2 & 0 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 2 & 0 & -2 \end{pmatrix}$$

и ищем **ф.с.р.** для системы $\mathbf{B}^2X = \mathbb{O}$. Эта система вырождается в единственное уравнение

$$x_2 + x_4 - x_6 = 0 \ ,$$

для которого **ф.с.р.** можно строить произвольным образом. Мы, однако же, построим ее дополнением **ф.с.р.**, полученной на шаге **2**:

$$\begin{array}{r|rrrrrl}
x_4 & x_3 & x_5 & x_6 & x_1 & x_2 & \\ \hline
0 & 1 & 0 & 0 & -3 & 0 & \\
0 & 0 & 1 & 0 & 0 & 0 & \Bigg\}\text{базис } \mathbb{Q}_1 \\
1 & 0 & 0 & 1 & 0 & 0 & \\
0 & 0 & 0 & 0 & 1 & 0 & \\
-1 & 0 & 0 & 0 & \multicolumn{2}{c}{\underbrace{0 \quad\ 1}_{E}} &
\end{array}$$

(переменные x_1 и x_2, которые были зависимыми на шаге **2**, переведены в разряд основных).

Таким образом $\mathbb{Q}_2 = \mathcal{L}(\mathfrak{e}_5, \mathfrak{e}_4 + \mathfrak{e}_6, -3\mathfrak{e}_1 + \mathfrak{e}_3, \underbrace{\mathfrak{e}_2 - \mathfrak{e}_4, \mathfrak{e}_1}_{\mathbb{Q}_2 \dot{\ } \mathbb{Q}_1})$.

4. Ищем $\mathbb{Q}_3$, т.е. подпространство корневых векторов высоты ≤ 3, принадлежащих λ_1. Матрица $\mathbf{B}^3$ оказывается нулевой, следовательно $\mathbb{Q}_3 = \mathbb{R}^6$ и базис $\mathbb{Q}_3$ построим дополнением базиса $\mathbb{Q}_2$:

$$\mathbb{Q}_3 = \mathcal{L}(\mathfrak{e}_5, \mathfrak{e}_4 + \mathfrak{e}_6, -3\mathfrak{e}_1 + \mathfrak{e}_3, \mathfrak{e}_2 - \mathfrak{e}_4, \mathfrak{e}_1, \underbrace{\ \mathfrak{e}_4\ }_{\mathbb{Q}_3 \dot{\ } \mathbb{Q}_2}) \ .$$

5. Поскольку число векторов в базисе $\mathbb{Q}_3$ совпало с кратностью $\mathfrak{m}_1 = 6$ собственного числа $\lambda_1 = 2$, то на этом процесс вычисления корневых векторов останавливается. Применение первой части теоремы 8.13 дает информацию о структуре клеток Жордана, соответствующих λ_1: $p = 3, k_1 = 1$, следовательно имеется одна клетка 3×3; в относительном базисе $\mathbb{Q}_2$ над $\mathbb{Q}_1$ содержатся 2 вектора, т.е. $k_1 + k_2 = 2$, следовательно $k_2 = 1$, т.е. имеется одна клетка 2×2; в базисе $\mathbb{Q}_1$ содержатся 3 вектора, т.е. $k_1 + k_2 + k_3 = 3$, следовательно $k_3 = 1$, т.е. имеется одна клетка 1×1.

Можно представить этот алгоритм в виде схемы 1, в ней каждый этаж показывает число корневых векторов, добавляемых на каждом шаге. Стоящие друг над другом квадраты образуют башни, высоты которых дают размерности клеток Жордана.

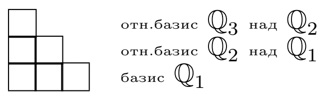
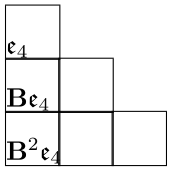


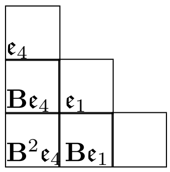
Схема 1.

$$
\mathbf{A}_{\mathfrak{J}} = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{pmatrix} \quad . \tag{8.7}
$$

6. Для построения канонического базиса обратимся к схеме 1, и будем заполнять ее квадраты, начиная с самого верхнего. Согласно теореме 8.13, для построения базиса циклического подпространства размерности 3 мы должны взять произвольный вектор из относительного базиса $\mathbb{Q}_3$ над $\mathbb{Q}_2$. Этот вектор единствен: $\mathfrak{e}_4$. Далее, домножаем его на матрицы $\mathbf{B}$ и $\mathbf{B}^2$. Три полученных вектора $\mathfrak{e}_4, \mathfrak{e}_2 - \mathfrak{e}_4, 2(\mathfrak{e}_4 + \mathfrak{e}_6)$ являются первыми векторами канонического базиса.



Больше циклических подпространств размерности 3 не имеется, и мы начинаем искать базис циклических подпространств размерности 2. Согласно теореме 8.13, мы должны взять произвольный вектор из относительного базиса $\mathbb{Q}_2$ над $\mathbb{Q}_1$, линейно независимый с тем, что получен ранее, т.е. с $(\mathfrak{e}_2 - \mathfrak{e}_4)$. Такой вектор единствен: $\mathfrak{e}_1$. Домножим его на матрицу $\mathbf{B}$. Два вектора: $\mathfrak{e}_1, -3\mathfrak{e}_1 + \mathfrak{e}_3$ являются следующими векторами канонического базиса и соответствуют клетке Жордана 2×2.



Осталось одномерное циклическое подпространство. Его базис выбирается из $\mathbb{Q}_1$. Из базисных векторов $\mathbb{Q}_1$ можно взять только $\mathfrak{e}_5$ (т.к. векторы $-3\mathfrak{e}_1+\mathfrak{e}_3$ и $\mathfrak{e}_4+\mathfrak{e}_6$ уже содержатся среди канонических). Итак, канонический базис пространства $\mathbb{V}$ задается

$$
\{\mathfrak{e}_4, \mathfrak{e}_2 - \mathfrak{e}_4, 2(\mathfrak{e}_4 + \mathfrak{e}_6), \mathfrak{e}_1, -3\mathfrak{e}_1 + \mathfrak{e}_3, \mathfrak{e}_5\}
$$

т.е. матрица

$$
C = \begin{pmatrix} 0 & 0 & 0 & 1 & -3 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & -1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 2 & 0 & 0 & 0 \end{pmatrix}
$$

приводит матрицу $\mathbf{A}$ к **ж.н.ф.** (8.7): $C^{-1}\mathbf{A}C = \mathbf{A}_{\mathfrak{J}}$.

Замечание. Канонический базис и, следовательно, матрица C определяются не единственным способом. Поэтому актуальна проверка правильности вычислений. Такая проверка может быть проведена посредством проверки более простого условия:

$$
\mathbf{A}C = C\mathbf{A}_{\mathfrak{J}} \quad . \tag{8.8}
$$

Следует, однако, иметь в виду, что последнее условие является необходимым, но не достаточным. Так, справедливо равенство

$$\underbrace{\begin{pmatrix} 0 & -1 & 1 \\ 2 & -5 & 3 \\ 6 & -13 & 7 \end{pmatrix}}_{\mathbf{A}} \underbrace{\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 1 & 4 \end{pmatrix}}_{C} = \underbrace{\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 1 & 4 \end{pmatrix}}_{C} \begin{pmatrix} 0 & & \\ & 0 & \\ & & 2 \end{pmatrix}$$

тем не менее истинная **ж.н.ф.** матрицы $\mathbf{A}$ недиагональна:

$$\mathbf{A}_{\mathfrak{J}} = \begin{pmatrix} 0 & & \\ 1 & 0 & \\ & & 2 \end{pmatrix} \quad \text{при} \quad C = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 2 \\ 2 & 1 & 4 \end{pmatrix} .$$

Объяснение этой кажущейся неоднозначности заключается в том, что первая матрица C является особенной: $\det C = 0$, и C^{-1} не существует.

9 Жорданова нормальная форма оператора в $\mathbb{R}$

Рассмотрим оператор $\mathscr{A}$ над вещественным пространством $\mathbb{V}$.

В произвольном базисе $\mathbb{V}$ матрица оператора $\mathbf{A}$ будет вещественной. В то же время, в произвольном каноническом базисе того же пространства, **ж.н.ф.** матрицы оператора не обязательно сохранит вещественный же вид. Действительно, основная теорема высшей алгебры гарантирует существование *комплексных* собственных чисел; иными словами дает возможность представить хар.полином в виде

$$\det(\mathbf{A}-\lambda E) = (-1)^n (x-\lambda_1)^{\mathfrak{m}_1} \times \dots \times (x-\lambda_{\mathfrak{r}})^{\mathfrak{m}_{\mathfrak{r}}} (x^2+p_1x+q_1)^{\mathfrak{M}_1} \times \dots \times (x^2+p_\ell x+q_\ell)^{\mathfrak{M}_\ell}$$

где $\{\lambda_1, \dots, \lambda_{\mathfrak{r}}\}$ — различные вещественные числа, а квадратные трехчлены $\{x^2+p_1x+q_1, \dots, x^2+p_\ell x+q_\ell\} \subset \mathbb{R}[x]$ — различные с отрицательными дискриминантами $\mathcal{D}_j = p_j^2 - 4q_j < 0$. Мы будем искать такие базисы пространства, в которых матрица оператора имеет наиболее простой вид, оставаясь при этом вещественной. Для того, чтобы увидеть конечную цель, обратимся к примеру 5.4 на с. 75. Собственные векторы $\mathfrak{X}_3$ и $\mathfrak{X}_4$, принадлежащие соответственно собственным числам $\boldsymbol{i}\sqrt{83}$ и $(-\boldsymbol{i}\sqrt{83})$, оказались комплексно–сопряженными:

$$\begin{aligned} \mathfrak{X}_3 &= \left[1-\boldsymbol{i}\sqrt{83}\,,\, 8-2\boldsymbol{i}\sqrt{83}\,,\, 12\,,\, 17+\boldsymbol{i}\sqrt{83}\right]^\top , \\ \mathfrak{X}_4 &= \left[1+\boldsymbol{i}\sqrt{83}\,,\, 8+2\boldsymbol{i}\sqrt{83}\,,\, 12\,,\, 17-\boldsymbol{i}\sqrt{83}\right]^\top . \end{aligned}$$

Выделим у этих векторов вещественную и мнимую части:

$$\mathfrak{Y}_3 \stackrel{\text{def}}{=} {}^1\!/_2\,(\mathfrak{X}_3+\mathfrak{X}_4) = \Re e\,\mathfrak{X}_3\ , \quad \mathfrak{Y}_4 \stackrel{\text{def}}{=} 1/(2\boldsymbol{i})\,(\mathfrak{X}_3-\mathfrak{X}_4) = \Im m\,\mathfrak{X}_3\ ;$$

новые векторы дополним до базиса $\mathbb{R}^4$ собственными векторами $\mathfrak{X}_1$ и $\mathfrak{X}_2$, принадлежащими собственному числу $\lambda = 0$. Матрица

$$\widetilde{C} \stackrel{\text{def}}{=} \begin{pmatrix} 4 & 7 & 1 & -\sqrt{83} \\ -2 & -3 & 8 & -2\sqrt{83} \\ 1 & 0 & 12 & 0 \\ 0 & 1 & 17 & \sqrt{83} \end{pmatrix} \quad \begin{matrix} \text{приводит} \\ \text{матрицу } \mathbf{A} \\ \text{к виду} \end{matrix} \quad \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sqrt{83} \\ 0 & 0 & -\sqrt{83} & 0 \end{pmatrix}$$

Хотя последняя матрица и не является диагональной, тем не менее она имеет достаточно простой вид. Вещественные собственные числа продолжают оставаться на диагонали, а векторы канонического базиса им соответствующие, строятся по алгоритму предыдущего параграфа. Комплексно-сопряженной паре собственных чисел $\pm \boldsymbol{i}\sqrt{83}$ в новой матрице соответствует блок 2×2 при соответствующих векторах канонического базиса, равных вещественной и мнимой части собственного вектора, принадлежащего $\boldsymbol{i}\sqrt{83}$. Теперь осталось подтвердить это предположение для общего случая.

Рассмотрим сначала случай когда квадратный трехчлен $x^2 + p\,x + q$ $(p^2 - 4\,q < 0)$ входит в разложение хар.полинома в первой степени, т.е. его корни $\lambda = \alpha + \boldsymbol{i}\,\beta$ и $\overline{\lambda} = \alpha - \boldsymbol{i}\,\beta$ являются простыми. Если вектор $\mathfrak{X}$ является собственным, принадлежащим λ, то $\overline{\mathfrak{X}}$ будет собственным, принадлежащим $\overline{\lambda}$; эти два вектора будут **л.н.з.** поскольку принадлежат различным собственным числам. Возьмем

$$\mathfrak{A} \stackrel{\text{def}}{=} \Re e\,(\mathfrak{X}) \ , \quad \mathfrak{B} \stackrel{\text{def}}{=} \Im m\,(\mathfrak{X}) \ .$$

Тогда

$$\mathscr{A}(\mathfrak{X}) = \lambda \mathfrak{X} \iff \mathscr{A}(\mathfrak{A} + \boldsymbol{i}\,\mathfrak{B}) = (\alpha + \boldsymbol{i}\,\beta)(\mathfrak{A} + \boldsymbol{i}\,\mathfrak{B}) \iff$$

$$\iff \left\{ \begin{array}{lcl} \mathscr{A}(\mathfrak{A}) & = & \alpha\mathfrak{A} - \beta\mathfrak{B}\,, \\ \mathscr{A}(\mathfrak{B}) & = & \beta\mathfrak{A} + \alpha\mathfrak{B}\,. \end{array} \right.$$

Линейная оболочка векторов $\mathfrak{A}$ и $\mathfrak{B}$ является инвариантным подпространством оператора $\mathscr{A}$. Если дополнить эти векторы до произвольного базиса пространства $\mathbb{V}$, то, в соответствии с теоремой 5.2, в этом базисе матрица оператора будет иметь вид

$$\begin{pmatrix} \alpha & \beta & * & \dots & * \\ -\beta & \alpha & * & \dots & * \\ 0 & 0 & * & \dots & * \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \underbrace{* & \dots & *}_{n-2} \end{pmatrix} \ .$$

В случае, когда хар.полином не имеет кратных корней, то такой алгоритм, примененный к каждому собственному числу, позволит построить базис, в котором матрица оператора будет иметь вид

$$\begin{pmatrix} \alpha_1 & \beta_1 & & & & & & & \\ -\beta_1 & \alpha_1 & & & & \mathbb{O} & & & \\ & & \ddots & & & & & & \\ & & & \alpha_\ell & \beta_\ell & & & & \\ & \mathbb{O} & & -\beta_\ell & \alpha_\ell & & & & \\ & & & & & \lambda_1 & & \mathbb{O} & \\ & & & & & & \ddots & & \\ & & & & & \mathbb{O} & & \lambda_{\mathfrak{r}} \end{pmatrix},$$

который и принимают за **ж.н.ф.** в $\mathbb{R}$.

Упражнение 9.1. *Построить* **ж.н.ф.** *в* $\mathbb{R}$ *для матрицы оператора*

$$\begin{pmatrix} 3 & -1 & 0 \\ 6 & -3 & 2 \\ 8 & -6 & 5 \end{pmatrix} .$$

Рассмотрим теперь случай, когда хар.полином имеет кратный квадратичный множитель, т.е. когда собственные числа $\lambda = \alpha + \boldsymbol{i}\,\beta$ и $\overline{\lambda} = \alpha - \boldsymbol{i}\,\beta$ являются кратными кратности $\mathfrak{m}$. Действуя по алгоритму пунктов 8.3 и 8.4, разбиваем подпространство векторов, аннулируемых $\left(\mathscr{A}^2 + p\,\mathscr{A} + q\mathscr{E}\right)^{\mathfrak{m}}$, на циклические подпространства. Пусть, например X — корневой вектор высоты s, принадлежащий собственному числу λ, тогда $\overline{X}$ будет корневым вектором той же высоты, принадлежащим собственному числу $\overline{\lambda}$. Легко проверить, что порождаемые этими векторами “башни” состоят из комплексно-сопряженных векторов: если

$$X,\ X_1 = (\mathscr{A} - \lambda\mathscr{E})(X),\ X_2 = (\mathscr{A} - \lambda\mathscr{E})(X_1), \dots, X_{s-1} = (\mathscr{A} - \lambda\mathscr{E})(X_{s-2})$$

— одна из них, то

$$\overline{X},\ \overline{X}_1 = (\mathscr{A} - \overline{\lambda}\mathscr{E})(\overline{X}),\ \overline{X}_2 = (\mathscr{A} - \overline{\lambda}\mathscr{E})(\overline{X}_1), \dots, \overline{X}_{s-1} = (\mathscr{A} - \overline{\lambda}\mathscr{E})(\overline{X}_{s-2})$$

— вторая. В вещественный базис суммы порожденных этими векторами циклических подпространств возьмем вещественные и мнимые части векторов. Пусть

$$X = U + \boldsymbol{i}\,V,\ X_1 = U_1 + \boldsymbol{i}\,V_1,\ X_2 = U_2 + \boldsymbol{i}\,V_2, \dots,\ X_{s-1} = U_{s-1} + \boldsymbol{i}\,V_{s-1}\ ,$$

тогда система векторов

$$U,\, V,\, U_1,\, V_1,\, U_2,\, V_2, \dots,\, U_{s-1},\, V_{s-1} \tag{9.1}$$

будет **л.н.з.** как неособенная комбинация линейно-независимых векторов. Выясним как преобразуются эти векторы под действием оператора $\mathscr{A}$.

$$\begin{aligned} \mathscr{A}(U + \boldsymbol{i}\,V) = \mathscr{A}(X) &= \lambda\,X + X_1 = (\alpha + \boldsymbol{i}\,\beta)(U + \boldsymbol{i}\,V) + (U_1 + \boldsymbol{i}\,V_1) = \\ &= (\alpha\,U - \beta\,V + U_1) + \boldsymbol{i}\ (\beta U + \alpha V + V_1) \end{aligned}$$

$$\Longleftrightarrow \left\{ \begin{array}{rl} \mathscr{A}(U) = & \alpha\, U - \beta\, V + 1 \cdot U_1 + 0 \cdot V_1, \\ \mathscr{A}(V) = & \beta\, U + \alpha\, V + 0 \cdot U_1 + 1 \cdot V_1. \end{array} \right.$$

По аналогии получим

$$\begin{array}{rl} \mathscr{A}(U_k) = & \alpha\, U_k - \beta\, V_k + 1 \cdot U_{k+1} + 0 \cdot V_{k+1}, \\ \mathscr{A}(V_k) = & \beta\, U_k + \alpha\, V_k + 0 \cdot U_{k+1} + 1 \cdot V_{k+1} \end{array}$$

для $k \in \{2, \ldots, s-2\}$ и

$$\begin{array}{rl} \mathscr{A}(U_{s-1}) = & \alpha\, U_{s-1} - \beta\, V_{s-1}, \\ \mathscr{A}(V_{s-1}) = & \beta\, U_{s-1} + \alpha\, V_{s-1}. \end{array}$$

Если дополнить векторы (9.1) до произвольного базиса пространства $\mathbb{V}$, то, в соответствии с теоремой 5.2, в этом базисе матрица оператора будет иметь вид

$$\left(\begin{array}{cccccccccccc} \alpha & \beta & 0 & 0 & & & & & & * & \ldots & * \\ -\beta & \alpha & 0 & 0 & & & & \mathbb{O} & & * & \ldots & * \\ 1 & 0 & \alpha & \beta & 0 & 0 & & & & * & \ldots & * \\ 0 & 1 & -\beta & \alpha & 0 & 0 & & & & * & \ldots & * \\ 0 & 0 & 1 & 0 & \alpha & \beta & & & & * & \ldots & * \\ 0 & 0 & 0 & 1 & -\beta & \alpha & & & & * & \ldots & * \\ \vdots & & & & & \vdots & \ddots & & & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \ldots & \alpha & \beta & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & \ldots & -\beta & \alpha & * & \ldots & * \\ 0 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & * & \ldots & * \\ \vdots & & & & & \vdots & & \vdots & & \vdots & & \\ 0 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & * & \ldots & * \end{array} \right)$$

Такой алгоритм, примененный к каждому кратному собственному числу, позволит построить базис, в котором матрица оператора будет иметь четырехдиагональую форму, причем на главной диагонали будут находиться вещественные части собственных чисел. Эта форма матрицы и считается **ж.н.ф.** в $\mathbb{R}$.

Глава 4. Применения жордановой нормальной формы

Все применения **ж.н.ф.** основаны на формуле

$$C^{-1}\mathbf{A}C = \mathbf{A}_{\mathfrak{z}} \quad \Longleftrightarrow \quad \mathbf{A} = C\mathbf{A}_{\mathfrak{z}}C^{-1}$$

Эти формулы означают, что матрица $\mathbf{A}$ подобна своей форме Жордана:

$$\mathbf{A} \doteq \mathbf{A}_{\mathfrak{z}} \ . \tag{9.2}$$

Но тогда любая мыслимая операция над матрицей $\mathbf{A}$ может быть сведена к этой же операции над формой Жордана:

$$g(\mathbf{A}) = Cg(\mathbf{A}_{\mathfrak{z}})C^{-1} \ .$$

Мы будем рассматривать матрицы и их **ж.н.ф.** над $\mathbb{C}$.

1 Матричный полином

1.1 Структура степенной функции от матрицы

Задача. Вычислить $\mathbf{A}^N$ для $N \in \mathbb{N}$.

В §4 главы 3 утверждалось, что матричные полиномы от подобных матриц будут также подобны, в частности, из (9.2) следует:

$$\mathbf{A}^N = C\mathbf{A}_{\mathfrak{z}}^N C^{-1} \ . \tag{1.1}$$

Лемма 1. *Если матрица* D *— блочно-диагональная, то и* D^N *— блочно-диагональная:*

$$\begin{pmatrix} \mathsf{D}_1 & \mathbb{O} & \dots & \mathbb{O} \\ \mathbb{O} & \mathsf{D}_2 & \dots & \mathbb{O} \\ \vdots & & \ddots & \vdots \\ \mathbb{O} & \mathbb{O} & \dots & \mathsf{D}_r \end{pmatrix}^N = \begin{pmatrix} \mathsf{D}_1^N & \mathbb{O} & \dots & \mathbb{O} \\ \mathbb{O} & \mathsf{D}_2^N & \dots & \mathbb{O} \\ & & \ddots & \\ \mathbb{O} & \mathbb{O} & \dots & \mathsf{D}_r^N \end{pmatrix}$$

здесь D_j *— матрица* $\mathfrak{m}_j \times \mathfrak{m}_j$.

Лемма 2. *Если матрица* L *— левая (нижняя) треугольная, то и* L^N *— левая (нижняя) треугольная.*

Формула (1.1), а также леммы 1 и 2 сводят вычисление степени матрицы $\mathbf{A}$ к вычислению степени ее клеток Жордана $\mathfrak{J}_k(\lambda_j)$.

Теорема 1.1.

$$\begin{pmatrix} \lambda & & & & \\ 1 & \lambda & & \mathbb{O} & \\ 0 & 1 & \lambda & & \\ \vdots & & \ddots & \ddots & \\ 0 & 0 & \dots & 1 & \lambda \end{pmatrix}^N_{k\times k} =$$

$$
= \begin{pmatrix} \lambda^N & & & & \\ \mathsf{C}_N^1\lambda^{N-1} & \lambda^N & & \mathbb{O} & \\ \mathsf{C}_N^2\lambda^{N-2} & \mathsf{C}_N^1\lambda^{N-1} & \lambda^N & & \\ \dots & & \dots & \ddots & \\ \mathsf{C}_N^{k-1}\lambda^{N-k+1} & \dots & & \mathsf{C}_N^1\lambda^{N-1} & \lambda^N \end{pmatrix} \tag{1.2}
$$

(считаем здесь элементы с отрицательными показателями λ равными нулю). Более корректная запись матрицы в правой части возможна с помощью матрицы

$$
H_j \stackrel{\text{def}}{=} \begin{array}{r} {\scriptstyle 0\to} \\ {\scriptstyle 1\to} \\ \vdots \\ {\scriptstyle j\to} \\ \vdots \\ {\scriptstyle k-1\to} \end{array} \begin{pmatrix} 0 & & & & \\ 0 & & \mathbb{O} & & \\ \vdots & & & & \\ 1 & & & & \\ \vdots & \ddots & & & \\ 0 & 1 & \dots & 0 & 0 \end{pmatrix} = \begin{bmatrix} \mathbb{O}_{j\times(k-j)} & \mathbb{O}_{j\times j} \\ E_{(k-j)\times(k-j)} & \mathbb{O}_{(k-j)\times j} \end{bmatrix}_{k\times k}
$$

(единицами заполнена j-я диагональ, считая от нулевой — главной):

$$
[\mathfrak{J}_k(\lambda)]^N =
$$

$$
= \lambda^N E + \mathsf{C}_N^1\lambda^{N-1}H_1 + \mathsf{C}_N^2\lambda^{N-2}H_2 + \dots + \begin{cases} \mathsf{C}_N^N H_N & \text{при} \quad N < k, \\ \mathsf{C}_N^{k-1}\lambda^{N-k+1}H_{k-1} & \text{при} \quad N \ge k. \end{cases} \tag{1.3}
$$

Доказательство проводится индукцией по N. Для $N = 1$ утверждение очевидно. Заметим, что

$$
\mathfrak{J}_k(\lambda) = \lambda\, E + H_1 \quad \text{и} \quad H_1^j = H_j \ .
$$

Если предположить справедливость формулы (1.3) для N, то

$$
(\lambda\, E + H_1)^{N+1} = (\lambda\, E + H_1)\left(\lambda^N E + \mathsf{C}_N^1\lambda^{N-1}H_1 + \dots + \mathsf{C}_N^m\lambda^{N-m}H_m + \dots\right) =
$$

$$
\begin{aligned} = \lambda^{N+1}E \quad &+ \mathsf{C}_N^1\lambda^N H_1 \quad + \dots + \quad \mathsf{C}_N^m\lambda^{N-m+1}H_m + \dots + \\ &+ \lambda^N H_1 \qquad + \dots + \quad \mathsf{C}_N^{m-1}\lambda^{N-m+1}H_1H_{m-1} + \dots \end{aligned}
$$

С помощью формулы $\mathsf{C}_{N+1}^m = \mathsf{C}_N^m + \mathsf{C}_N^{m-1}$ получаем справедливость формулы (1.3) и для показателя $N + 1$. □

Пример 1.1. *Вычислить*

$$
\begin{pmatrix} 1 & 2 & 1 & 0 \\ 1 & 1 & 0 & -1 \\ -2 & 0 & 1 & 2 \\ 0 & 2 & 1 & 1 \end{pmatrix}^{100} .
$$

Решение. Сначала найдем **ж.н.ф.** и матрицу C.

$$\det(\mathbf{A} - \lambda E) = (\lambda - 1)^4 \ .$$

Ищем корневые векторы, принадлежащие $\lambda_1 = 1$:

$$\mathbf{B} = \mathbf{A} - E = \begin{pmatrix} 0 & 2 & 1 & 0 \\ 1 & 0 & 0 & -1 \\ -2 & 0 & 0 & 2 \\ 0 & 2 & 1 & 0 \end{pmatrix} \quad \to \quad \begin{pmatrix} 0 & 2 & 1 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Имеем: $\mathbb{Q}_1 = \mathcal{L}\left([0,-1,2,0]^\top, [1,0,0,1]^\top\right)$, следовательно матрица недиагонализуема.

$$\mathbf{B}^2 = \mathbb{O}_{4\times 4} \quad \Longrightarrow \quad \mathbb{Q}_2 = \mathcal{L}\left([0,-1,2,0]^\top, [1,0,0,1]^\top, [1,0,0,0]^\top, [0,1,0,0]^\top\right)$$

В **ж.н.ф.** имеются 2 клетки 2×2. Для построения канонического базиса берем векторы из относительного базиса $\mathbb{Q}_2$ над $\mathbb{Q}_1$ и домножаем их на $\mathbf{B}$ (заполняем башни сверху вниз). Получаем

$\mathfrak{e}_1$	$\mathfrak{e}_2$
$\mathbf{B}\mathfrak{e}_1$	$\mathbf{B}\mathfrak{e}_2$

$$C = \begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 1 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \quad , \mathbf{A}_{\mathfrak{J}} = \begin{pmatrix} 1 & & & \\ 1 & 1 & & \\ & & 1 & \\ & & 1 & 1 \end{pmatrix}$$

Теперь воспользуемся формулой (1.1):

$$\mathbf{A}^{100} = C \begin{pmatrix} 1 & & & \\ 1 & 1 & & \\ & & 1 & \\ & & 1 & 1 \end{pmatrix}^{100} C^{-1} =$$

$$= \begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 1 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & & & \\ 100 & 1 & & \\ & & 1 & \\ & & 100 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & -1/2 & 0 \\ 0 & 1 & 1/2 & 0 \\ 0 & 0 & 0 & 1/2 \end{pmatrix} =$$

$$= \begin{pmatrix} 1 & 200 & 100 & 0 \\ 100 & 1 & 0 & -100 \\ -200 & 0 & 1 & 200 \\ 0 & 200 & 100 & 1 \end{pmatrix}$$

1.2 Вычисление матричного полинома

Рассмотрим произвольный полином $g(x) = b_0 x^m + b_1 x^{m-1} + \dots + b_m \in \mathbb{C}[x]$. Вычисление $g(\mathbf{A})$ производится по тому же алгоритму, что и вычисление $\mathbf{A}^N$: $g(\mathbf{A}) = C g(\mathbf{A}_{\mathfrak{J}}) C^{-1}$ и леммы аналогичные леммам 1 и 2 сводят вычисление к вычислению $g(x)$ от клеток Жордана. Поскольку

$$g\left(\mathfrak{J}_k(\lambda)\right) = b_0 \mathfrak{J}_k(\lambda)^m + b_1 \mathfrak{J}_k(\lambda)^{m-1} + \dots + b_m E_{m\times m}$$

то на основании формулы (1.2) получаем:

$$g\left(\begin{bmatrix} \lambda & & & & \\ 1 & \lambda & & & \\ 0 & 1 & \lambda & & \\ \vdots & & \ddots & \ddots & \\ 0 & 0 & \dots & 1 & \lambda \end{bmatrix}_{k\times k}\right) = \begin{bmatrix} g(\lambda) & & & & \\ g'(\lambda) & g(\lambda) & & & \\ \frac{g''(\lambda)}{2!} & g'(\lambda) & g(\lambda) & & \\ \vdots & & \ddots & \ddots & \\ \frac{g^{(k-1)}(\lambda)}{(k-1)!} & \frac{g^{(k-2)}(\lambda)}{(k-2)!} & \dots & g'(\lambda) & g(\lambda) \end{bmatrix} = \tag{1.4}$$

$$= g(\lambda)E + g'(\lambda)H_1 + \frac{g''(\lambda)}{2!}H_2 + \dots + \frac{g^{(k-1)}(\lambda)}{(k-1)!}H_{k-1}$$

Одно обстоятельство упрощает вычисление $g(\mathbf{A})$ в случае когда $\deg g \ge n$ (порядок матрицы $\mathbf{A}$).

Теорема 1.2. *Обозначим $g_1(x)$ остаток от деления $g(x)$ на хар. полином $f(x) = \det(\mathbf{A} - xE)$. Тогда*

$$g(\mathbf{A}) = g_1(\mathbf{A}) \ .$$

Доказательство . Действительно, указанное равенство последует из теоремы Гамильтона–Кэли[25] при подстановке матрицы $\mathbf{A}$ в формулу

$$g(x) \equiv f(x)q(x) + g_1(x) \quad , \deg g_1 < n \tag{1.5}$$

(здесь $q(x)$ — частное от деления $g(x)$ на $f(x)$). □

Эта теорема позволяет свести вычисление $g(\mathbf{A})$ к случаю полинома $g(x)$ степени меньшей порядка матрицы. Следующее упрощение возможно в том случае, когда нам известны собственные числа матрицы.

Теорема 1.3. *Если все собственные числа $\lambda_1, \dots, \lambda_n$ матрицы $\mathbf{A}$ различны, то*

$$g(\mathbf{A}) = \sum_{k=1}^{n} g(\lambda_k)\frac{f_k(\mathbf{A})}{f_k(\lambda_k)} = \sum_{k=1}^{n} g(\lambda_k)\frac{f_k(\mathbf{A})}{f'(\lambda_k)} \quad \text{где} \quad f_k(x) \stackrel{\mathrm{def}}{=} \frac{f(x)}{x-\lambda_k} \ . \tag{1.6}$$

Доказательство . При подстановке λ_k в (1.5) получаем: $g_1(\lambda_k) = g(\lambda_k)$ при $k = 1, \dots, n$. Поскольку $\deg g_1 < n$ то этим набором значений полином $g_1(x)$ определяется однозначно. Его выражение можно найти по формуле Лагранжа[26]:

$$g_1(x) \equiv \sum_{k=1}^{n} g_1(\lambda_k)\frac{(x-\lambda_1)\dots(x-\lambda_{k-1})(x-\lambda_{k+1})\dots(x-\lambda_n)}{(\lambda_k-\lambda_1)\dots(\lambda_k-\lambda_{k-1})(\lambda_k-\lambda_{k+1})\dots(\lambda_k-\lambda_n)}$$

При подстановке сюда матрицы $\mathbf{A}$ фактически и получаем (1.6). □

[25] Глава 3,§6.2.

[26] Раздел I, глава 5, §1.

Каждое из суммируемых выражений в формуле (1.6) состоит из двух сомножителей: один из них зависит от полинома $g(x)$, а второй — не зависит. Таким образом, при изменении полинома происходит лишь пересчет значений $g(\lambda_1), \dots, g(\lambda_n)$, а их матричные сомножители не меняются. Структура этих последних уже была нами установлена в §6 главы 3: каждая матрица $f_k(\mathbf{A})$ имеет свои столбцы пропорциональными.

Пример 1.2. *Выписать формулу* (1.6) *для произвольного* $g(x) \in \mathbb{C}[x]$ *и*

$$\mathbf{A} = \left(\begin{array}{rrr} 9 & 22 & -6 \\ -1 & -4 & 1 \\ 8 & 16 & -5 \end{array} \right) .$$

Решение. Все необходимые вычисления были произведены при решении примера 6.6 главы 3:

$$g(\mathbf{A}) = g(3)\frac{\begin{pmatrix} 40 & 80 & -20 \\ 0 & 0 & 0 \\ 40 & 80 & -20 \end{pmatrix}}{20} + g(-2)\frac{\begin{pmatrix} -10 & -30 & 10 \\ 5 & 15 & -5 \\ 0 & 0 & 0 \end{pmatrix}}{5} + g(-1)\frac{\begin{pmatrix} -4 & -8 & 4 \\ 4 & 8 & -4 \\ 8 & 16 & -8 \end{pmatrix}}{-4}$$

2 Линейное разностное уравнение.

Определение. Уравнение

$$x_{n+K} = a_1 x_{n+K-1} + \dots + a_n x_K, \quad a_j, x_p \in \mathbb{C} \tag{2.1}$$

называется **линейным разностным** (или **возвратным**) уравнением n-го порядка. Пусть числа $x_0, \dots, x_{n-1}$ заданы. Тогда уравнение (2.1) определяет **линейную рекуррентную** (или **возвратную**) последовательность n-го порядка: каждый элемент этой последовательности определяется через n предшествующих.

Пример 2.1. *Уравнение второго порядка*

$$x_{K+2} = x_{K+1} + x_K$$

определяет при $x_0 = x_1 = 1$ *последовательность чисел* **Фибоначчи**

$$\{x_K\}_{K=1}^{\infty} = \{1, 1, 2, 3, 5, 8, 13, 21, 34, \dots\} .$$

2.1 Аналитика

Задача. Решить уравнение (2.1), т.е. найти выражение для x_{n+K} в виде явной функции от номера K и "начальных данных" $x_0, \dots, x_{n-1}$. Будем говорить об **общем решении**, если $x_0, \dots, x_{n-1}$ считаются произвольными.

Для решения рассмотрим векторную последовательность в $\mathbb{C}^n$:

$$X_0 = \begin{pmatrix} x_0 \\ \vdots \\ x_{n-1} \end{pmatrix}, X_1 = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \dots, X_K = \begin{pmatrix} x_K \\ \vdots \\ x_{n+K-1} \end{pmatrix}, X_{K+1} = \begin{pmatrix} x_{K+1} \\ \vdots \\ x_{n+K} \end{pmatrix} \dots$$

с компонентами из элементов последовательности. Найдем линейный оператор, отображающий X_K в X_{K+1}. Легко проверяется равенство

$$X_{K+1} = \mathfrak{F} X_K \quad \text{при} \quad \mathfrak{F} = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 & 0 \\ \dots & & & & & & \dots \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 \\ a_n & a_{n-1} & a_{n-2} & & \dots & a_2 & a_1 \end{pmatrix}_{n\times n}$$

Матрица $\mathfrak{F}$ называется матрицей **Фробениуса**[27]. Имеем:

$$X_{K+1} = \mathfrak{F} X_K = \mathfrak{F}(\mathfrak{F} X_{K-1}) = \dots = \mathfrak{F}^{k+1} X_0 \ .$$

Мы свели поставленную задачу к возведению в степень матрицы $\mathfrak{F}$. Искомое выражение для x_{n+K} получится умножением последней строки матрицы $\mathfrak{F}^{K+1}$ на столбец "начальных данных" X_0.

Для нахождения $\mathfrak{F}^{K+1}$ воспользуемся результатами §1. Если найдем жорданову нормальную форму $\mathfrak{F}_{\mathfrak{Z}}$ и соответствующую матрицу преобразования базиса C, то получим

$$\mathfrak{F}_{\mathfrak{Z}} = C^{-1}\mathfrak{F} C \Longrightarrow \mathfrak{F}^{K+1} = C \mathfrak{F}_{\mathfrak{Z}}^{K+1} C^{-1} \ . \tag{2.2}$$

Хар. полином матрицы Фробениуса найден в примере 6.3 главы 3:

$$\det(\mathfrak{F} - \lambda E) = (-1)^n (\lambda^n - a_1\lambda^{n-1} - \dots - a_n) \ .$$

Определение. Полином $f(\lambda) = \lambda^n - a_1\lambda^{n-1} - \dots - a_n$ называется **характеристическим** полиномом уравнения (последовательности) (2.1). Он получается из (2.1) формальной заменой

$$x_{n+K} \to \lambda^n, x_{n+K-1} \to \lambda^{n-1}, \dots, x_K \to 1 \ .$$

Теорема 2.1. *Если все корни $\lambda_1, \dots, \lambda_n$ хар. полинома $f(\lambda)$ различны, то общее решение уравнения (2.1) имеет вид:*

$$x_{n+K} = C_1\lambda_1^{n+K} + \dots + C_n\lambda_n^{n+K} \ , \tag{2.3}$$

где $C_1, \dots, C_n$ — величины, зависящие от $x_0, \dots, x_{n-1}$ (и не зависящие от K).

Доказательство . При условии теоремы формула (2.2) имеет место при

$$\mathfrak{F}_{\mathfrak{Z}} = \begin{pmatrix} \lambda_1 & & & \mathbb{O} \\ & \lambda_2 & & \\ & & \ddots & \\ \mathbb{O} & & & \lambda_n \end{pmatrix} , \qquad C = \begin{pmatrix} 1 & 1 & \dots & 1 \\ \lambda_1 & \lambda_2 & \dots & \lambda_n \\ \dots & & & \dots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \dots & \lambda_n^{n-1} \end{pmatrix} .$$

[27]Глава 3, пример 6.3.

Действительно, пользуясь тем, что $\lambda_j^n = a_1\lambda_j^{n-1} + \cdots + a_n$ получаем:

$$\mathfrak{F}\begin{pmatrix} 1 \\ \lambda_j \\ \lambda_j^2 \\ \vdots \\ \lambda_j^{n-1} \end{pmatrix} = \begin{pmatrix} \lambda_j \\ \lambda_j^2 \\ \vdots \\ \lambda_j^{n-1} \\ a_n + \cdots + a_1\lambda_j^{n-1} \end{pmatrix} = \begin{pmatrix} \lambda_j \\ \lambda_j^2 \\ \vdots \\ \lambda_j^{n-1} \\ \lambda_j^n \end{pmatrix} = \lambda_j \begin{pmatrix} 1 \\ \lambda_j \\ \lambda_j^2 \\ \vdots \\ \lambda_j^{n-1} \end{pmatrix},$$

т.е. столбцами матрицы Вандермонда C являются собственные векторы матрицы $\mathfrak{F}$.

По формуле (2.2):

$$\mathfrak{F}^{K+1} = C \begin{pmatrix} \lambda_1^{K+1} & & \mathbb{O} \\ & \lambda_2^{K+1} & & \\ & & \ddots & \\ \mathbb{O} & & & \lambda_n^{K+1} \end{pmatrix} C^{-1} = \begin{pmatrix} \lambda_1^{K+1} & \lambda_2^{K+1} & \ldots & \lambda_n^{K+1} \\ \lambda_1^{K+2} & \lambda_2^{K+2} & \ldots & \lambda_n^{K+2} \\ \ldots & & & \ldots \\ \lambda_1^{K+n} & \lambda_2^{K+n} & \ldots & \lambda_n^{K+n} \end{pmatrix} C^{-1}$$

Последний элемент столбца $X_{K+1} = \mathfrak{F}^{K+1}X_0$, т.е. искомый x_{n+K} получается по формуле:

$$x_{n+K} = \left[\lambda_1^{n+K}, \lambda_2^{n+K}, \ldots, \lambda_n^{n+K}\right] C^{-1}X_0 \ .$$

Если теперь обозначить

$$\begin{pmatrix} C_1 \\ C_2 \\ \vdots \\ C_n \end{pmatrix} = C^{-1} \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_{n-1} \end{pmatrix} \tag{2.4}$$

то и получим представление (2.3). □

Замечание. Константы $C_1, \ldots, C_n$ находятся по формуле (2.4), т.е. как решения системы линейных уравнений

$$\begin{pmatrix} 1 & 1 & \ldots & 1 \\ \lambda_1 & \lambda_2 & \ldots & \lambda_n \\ \ldots & & \ldots & \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \ldots & \lambda_n^{n-1} \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \\ \vdots \\ C_n \end{pmatrix} = \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_{n-1} \end{pmatrix}$$

Уравнения этой системы получаются формальными подстановками $K = -n, -n+1, \ldots, -1$ в общую формулу (2.3).

Для последовательности чисел Фибоначчи (пример 2.1) имеем:

$$f(\lambda) = \lambda^2 - \lambda - 1, \quad \lambda_{1,2} = \frac{1 \pm \sqrt{5}}{2} \ ,$$

и решение уравнения дается формулой Бине[28]

$$x_K = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^{K+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{K+1}\right] \ .$$

Пример 2.2. *Решить уравнение третьего порядка*

$$x_{K+3} = x_K \ .$$

Решение. При любом наборе x_0, x_1, x_2 уравнение определяет периодическую последовательность. Хар. полином $\lambda^3 - 1$ имеет корни $\lambda_1 = 1, \lambda_{2,3} = -{}^1\!/_2 \pm i\,{}^{\sqrt{3}}\!/_2$. Хотя матрицы $\mathfrak{F}_3$ и C получаются комплексными, тем не менее при вещественных начальных данных общее решение можно записать в вещественном виде. Так, для $x_0 = 1, x_1 = 3, x_2 = 2$ с помощью формулы Муавра получаем ответ в виде:

$$x_{K+3} = 2 + (-1)^K \frac{2}{\sqrt{3}} \sin\frac{(K+4)\pi}{3} \ .$$

Лемма 1. *Если корень λ_1 характеристического полинома $f(\lambda)$ имеет кратность $\mathfrak{m}_1$, то в $\mathfrak{F}_3$ ему соответствует единственная клетка Жордана порядка $\mathfrak{m}_1$.*

Доказательство . Поскольку матрица

$$\mathbf{B} = \mathfrak{F} - \lambda_1 E = \left(\begin{array}{cccccccc} -\lambda_1 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & -\lambda_1 & 1 & 0 & \dots & 0 & 0 \\ \dots & & & & & & \dots \\ 0 & 0 & 0 & 0 & \dots & -\lambda_1 & 1 \\ a_n & a_{n-1} & a_{n-2} & & \dots & a_2 & a_1 - \lambda_1 \end{array}\right)_{n\times n} \qquad (2.5)$$

имеет $\operatorname{rank}\mathbf{B} = n-1$ ($\det\mathbf{B} = 0$, а минор $(n-1)$-го порядка, стоящий в правом верхнем углу отличен от нуля), то согласно теореме 8.13 главы 3 имеется единственная клетка Жордана, соответствующая λ_1. Тогда ее порядок совпадает с кратностью $\mathfrak{m}_1$. □

Теорема 2.2. *Если хар. полином имеет следующее разложение на линейные множители:*

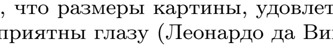
$$f(\lambda) = (\lambda - \lambda_1)^{\mathfrak{m}_1} \times \dots \times (\lambda - \lambda_{\mathfrak{r}})^{\mathfrak{m}_{\mathfrak{r}}}, \quad \lambda_k \neq \lambda_\ell \text{ при } k \neq \ell \ ,$$

то общее решение уравнения (2.1) *имеет вид:*

$$x_{n+K} = L_1(K)\lambda_1^{n+K-\mathfrak{m}_1} + \dots + L_{\mathfrak{r}}(K)\lambda_{\mathfrak{r}}^{n+K-\mathfrak{m}_r}, \qquad (2.6)$$

где $L_1(K), \dots, L_{\mathfrak{r}}(K)$ — полиномы от K: $L_p(K) \in \mathbb{C}[K]$, $\deg L_p < \mathfrak{m}_p$.

[28] Число $\lambda_1 = (1+\sqrt{5})/2 \approx 1.618034$ известно как число “золотого сечения”. Целое относится к большей своей части, как бо́льшая часть относится к меньшей:



$$\frac{|AB|}{|AC|} = \frac{|AC|}{|CB|} \implies \frac{|AB|}{|AC|} = \lambda_1$$

Было известно еще в древнем Египте и древней Греции (Фидий). В эпоху Возрождения считалось, что размеры картины, удовлетворяющие соотношению *ширина/высота*$=\lambda_1$ наиболее приятны глазу (Леонардо да Винчи).

Пример 2.3. *Решить уравнение четвертого порядка*

$$x_{K+4} = 4\,x_{K+3} - 6\,x_{K+2} + 4\,x_{K+1} - x_K \ .$$

Решение. Хар. полином $(\lambda - 1)^4$ имеет единственный корень $\lambda_1 = 1$. Общее решение ищется в виде $x_{K+4} = C_1 + C_2K + C_3K^2 + C_4K^3$, а при заданных начальных данных константы C_p определяются из системы линейных уравнений, получающейся из общего решения формальной подстановкой в него $K = -4, -3, -2, -1$. Так, при $x_0 = 1, x_1 = 8, x_2 = 27, x_3 = 64$ получаем следующую систему

$$\left(\begin{array}{llll} 1 & -4 & 16 & -64 \\ 1 & -3 & 9 & -27 \\ 1 & -2 & 4 & -8 \\ 1 & -1 & 1 & -1 \end{array}\right)\left(\begin{array}{c} C_1 \\ C_2 \\ C_3 \\ C_4 \end{array}\right) = \left(\begin{array}{c} 1 \\ 8 \\ 27 \\ 64 \end{array}\right),$$

решением которой являются $C_1 = 125, C_2 = 75, C_3 = 15, C_4 = 1$. Выражение для общего члена рекуррентной последовательности $x_{K+4} = (K+5)^3$.
$\square$

Замечание. На примере решения возвратного уравнения (2.1) мы показали обычную роль жордановой нормальной формы: роль лесов при строительстве здания. Необходимость в них пропадает как только здание построено. Именно это и произошло при решении нашей задачи: теоретический результат — т.е. вид формулы — был получен посредством **ж.н.ф.**, но вот конкретные значения для коэффициентов в этой формуле были получены применением другого алгоритма — метода неопределенных коэффициентов.

2.2 Асимптотика

Как ведет себя рекуррентная последовательность $\{x_K\}_{K=0}^{\infty}$ при возрастании K?

Определение. Если при любых $x_0, \dots, x_{n-1}$ решение x_K уравнения (2.1) ограничено, то будем называть это уравнение **устойчивым**. Устойчивое уравнение называется **асимптотически устойчивым** если

$$x_K \underset{K\to+\infty}{\to} 0 \ .$$

Уравнение называется **неустойчивым**, если существует хотя бы один набор $x_0, \dots, x_{n-1}$, для которого соответствующая последовательность x_K неограничена.

Теорема 2.3. *Уравнение* (2.1) *будет*

а) *устойчиво тогда и только тогда, когда* $|\lambda_j| \le 1$ *для любого* j, *и собственные числа с* $|\lambda_j| = 1$ *— простые для хар. полинома;*

б) *асимптотически устойчиво тогда и только тогда, когда* $|\lambda_j| < 1$ *для любого* j.

Конструктивная проверка условий теоремы возможна чисто алгебраическими методами: проверкой n ограничений критерия Шура–Кона[29] на коэффициенты хар.полинома.

Пример 2.4. *Найти все значения параметра α, при которых уравнение*

$$x_{K+5} = {}^{1}\!/_{5}(x_{K+4} - (\alpha - 4)x_{K+3} + (\alpha - 2)x_{K+2} + x_{K+1} + x_K)$$

будет устойчиво.

Решение. Хар.полином

$$\begin{array}{rcl} f(\lambda) & = & {}^{1}\!/_{5}\left(5\,\lambda^5 - \lambda^4 + (-4+\alpha)\,\lambda^3 + (-\alpha+2)\,\lambda^2 - \lambda - 1\right) = \\ & = & {}^{1}\!/_{5}(\lambda - 1)\underbrace{\left(5\,\lambda^4 + 4\,\lambda^3 + \alpha\,\lambda^2 + 2\,\lambda + 1\right)}_{\stackrel{\mathrm{def}}{=} f_1(\lambda)} \ . \end{array}$$

К полиному $f_1(\lambda)$ нужно применить алгоритм теоремы Шура–Кона. В примере 3.8 главы 8, раздела I мы уже сделали это: условие $0 < \alpha < {}^{27}\!/_{4}$ является необходимым и достаточным для того, чтобы все корни $f_1(\lambda)$ лежали внутри единичного круга. При $\alpha = 0$ полином $f_1(\lambda)$ имеет корень $\lambda = -1$; при $\alpha = {}^{27}\!/_{4}$ полином $f_1(\lambda)$ будет обладать комплексно-сопряженными корнями с модулями равными 1: $\lambda_{1,2} = -{}^{1}\!/_{4} \pm i\,{}^{\sqrt{15}}\!/_{4}$. В обоих этих “пограничных” случаях полином $f(\lambda)$ удовлетворяет условию **а)** теоремы 2.3.

Ответ. Уравнение устойчиво при $0 \le \alpha \le {}^{27}\!/_{4}$.

Случай существования простого собственного числа равного ± 1 или пары простых комплексно-сопряженных чисел равных по модулю 1, при прочих меньших 1 по модулю, оказывается “пограничным” при исследовании сходимости: решение уравнения оказывается ограниченным, но существует ли у него предел? В терминах матрицы Фробениуса $\mathfrak{F}$ вопрос этот равносилен существованию $\lim\limits_{K\to+\infty} \mathfrak{F}^K$. В следующем параграфе мы исследуем этот критический случай на примере другой матрицы — стохастической.

Упражнение 2.1. *Найти* $\lim\limits_{K\to+\infty} x_K$*, где*

$$x_{K+3} = {}^{1}\!/_{3}(x_{K+2} + x_{K+1} + x_K)\ , \text{и} \ \ x_0, x_1, x_2 \ \text{известны}\ .$$

Упражнение 2.2 (А.А.Марков). *Пусть хар.полином $f(\lambda)$ последовательности* (2.1) *имеет простой корень $\lambda_1 = 1$, а все остальные корни* $|\lambda_j| < 1$. *Доказать, что*

$$\lim_{K\to+\infty} x_K = -\frac{f(0)}{f'(0)}(x_{n-1} + b_1 x_{n-2} + \dots + b_{n-1}x_0)\ .$$

Здесь числа b_j — коэффициенты частного от деления $f(\lambda)$ на $(\lambda - 1)$: $f(\lambda) \equiv (\lambda - 1)(\lambda^{n-1} + b_1\lambda^{n-2} + \dots + b_{n-1})$.

[29] Раздел I, глава 8, §3.

3 Применения ж.н.ф. в теории вероятностей

3.0 Вспомогательные результаты

Определение. **Вероятностью события** называется отношение числа случаев благоприятных его появлению, к числу всех равновозможных случаев.

Из этого определения следует, что вероятность всегда выражается правильной дробью. Так, например, пусть имеется колода из 52 карт, хорошо смешанных, и из этой колоды вынимают одну карту, тогда число случаев благоприятных тому, что карта будет бубновой масти есть 13, общее число всех равновозможных случаев есть 52. Поэтому вероятность равна $^{13}/_{52} = {}^{1}/_{4}$.

Исчисление вероятностей основано на двух довольно наглядных результатах.

Теорема 3.1 (сложения вероятностей). *Если некоторое событие разбивается на несколько несовместимых видов, то его вероятность равна сумме вероятностей всех этих видов.*

Предположим, например, что в ящике содержатся n_1 белых шаров, на которых написан номер 1, и n_2 белых шаров, на которых написан номер 2, а также m_1 черных шаров, на которых написан номер 1, и m_2 черных шаров, на которых написан номер 2. Вероятность того, что случайным образом выбранный шар окажется белым равна $(n_1+n_2)/(n_1+m_1+n_2+m_2)$. Тот же самый результат получится если мы сложим вероятность появления белого шара с номером 1 с вероятностью появления белого шара с номером 2.

Теорема 3.2 (умножения вероятностей). *Вероятность совместного осуществления двух событий равна произведению вероятности одного из них на вероятность другого, вычисленную в предположении, что первое осуществилось.*

Доказательство . Пусть имеется некоторая физическая система, которая по своей природе ведет себя случайным образом и находится в одном (и только в одном) из n возможных, несовместимых и равновероятных состояний

$$S_1, \ldots, S_m, S_{m+1}, \ldots, S_{m_1}, S_{m_1+1}, \ldots S_n \ .$$

При этом нахождение системы в первых m_1 состояний

$$S_1, \ldots, S_m, S_{m+1}, \ldots, S_{m_1} \tag{3.1}$$

благоприятствует некоторому событию A, а остальные — не благоприятствуют. Пусть далее, из состояний (3.1) первые m, т.е. $S_1, \ldots, S_m$ благоприятствуют другому событию B, а остальные — не благоприятствуют.

При таких условиях вероятность события A выражается дробью $^{m_1}/_{n}$. Вероятность же события B, когда известно, что событие A уже произошло, выражается дробью $^{m}/_{m_1}$, т.к. при осуществлении события A состояния $S_{m_1+1}, \ldots, S_n$ невозможны, а состояния $S_1, \ldots, S_{m_1}$ остаются по-прежнему равновероятными. Наконец, вероятность того, что состояние, в котором оказалась система, одновременно обеспечивает события A и B выражаются дробью ${}^{m}\!/_{n}$, т.к. оба события появляются только в состояниях $S_1, \dots, S_m$. Очевидно:

$$ {}^{m}\!/_{n} = {}^{m_1}\!/_{n} \cdot {}^{m}\!/_{m_1} \ , $$

что и доказывает теорему. □

Пример 3.1. *В семье — двое детей. Какова вероятность того, что оба ребенка мальчики, если известно, что в семье уже есть мальчик?*

Решение. Большинство людей ожидают ответ $1/2$. Но это не так. Обозначим буквами М и Д соответственно мальчика и девочку, и на первом месте будем указывать старшего ребенка. Имеем четыре возможности: ММ, МД, ДМ, ДД. Этим событиям припишем вероятность $1/4$. Условие того, что один из детей — мальчик, сужает возможность вариаций: выбор возможен из трех равновероятных событий ММ, МД, ДМ. В соответствии с теоремой 3.2 получаем правильный

Ответ. $1/3$.

Определение. События называются **независимыми** друг от друга, если вероятность каждого из них не зависит от существования или несуществования остальных.

Следствие 1. *Вероятность совместного появления двух независимых событий равна произведению вероятностей появления каждого из них в отдельности.*

Следствие обобщается и на любое число K независимых событий, вероятность их появления будет равна произведению $p_1 \cdot p_2 \times \dots \times p_K$, где p_j — вероятность появления j-го события.

Упражнение 3.1. *Вероятность прихода в течение четверти часа автобуса 404 к остановке у станции метро "Автово" равна $1/6$, автобуса 424 — $1/2$, а автобуса 224 — $1/3$. Какова вероятность того, что студент, ожидающий любого из указанных видов транспорта уедет в университет в течение четверти часа?*

Приведем пример использования обеих теорем.

Пример 3.2 *Студент выходит из здания университета и едет на "Балтийскую" на одном из видов транспорта: автобусе или электричке. В среднем, в трех случаях из четырех он идет на платформу, а в одном случае — на автобусную остановку*[30]*. Ждет транспорта 20 минут. Вероятность прихода автобуса — $1/2$, электрички — $1/3$. Какова вероятность, что студент уедет в город в течение этих 20 минут?*

[30]Для простоты считаем, что расстояния до этих объектов одинаковы.

Решение. Итак, событие “отъезд в город” складывается из двух независимых: “отъезд в город на автобусе” и “отъезд в город на электричке”. Вероятность каждого из этих двух событий вычисляется по теореме 3.2: вероятность первого равна $1/4 \times 1/2$, а второго — $3/4 \times 1/3$. В соответствии с теоремой 3.1, вероятность отъезда получится сложением этих двух величин.

Ответ. $3/8$.

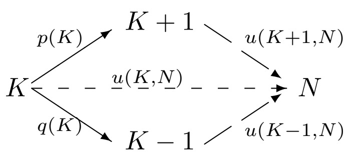
3.1 Задача о разорении игрока

Задача. Пусть игрок обладает конечным капиталом $K \in \mathbb{N}$ и участвует с ним в игре, где имеется вероятность $p(K)$ увеличения его капитала до $K+1$ и вероятность $q(K) = 1-p(K)$ сокращения его до $K-1$. Пусть, кроме того, игрок согласен делать ставки до тех пор, пока он либо накопит капитал N, либо потеряет все свои деньги: $K = 0$. Какова вероятность того, что при заданных $p(1), \ldots, p(N-1)$ игрок достигнет своей цели прежде, чем разорится?

Введем в рассмотрение функцию

$u(K,N)$ = вероятность того, что игрок, начиная с капиталом K, накопит капитал N прежде, чем разорится.

На каждом шаге игрок либо выигрывает с вероятностью $p(K)$ и тогда вероятность успешного окончания игры становится равной $u(K+1,N)$, либо проигрывает с вероятностью $q(K)$ и тогда вероятность успешного окончания игры становится равной $u(K-1,N)$.



В соответствии с теоремами 3.2 и 3.1 для вероятности $u(K,N)$ получаем следующую формулу:

$$u(K,N) = p(K)u(K+1,N) + q(K)u(K-1,N) \ . \tag{3.2}$$

При $K = 0$ и $K = N$ задаются ограничения “окончания игры”:

$$\begin{cases} u(0,N) &= 0 \quad \text{(игрок проиграл все деньги)} \\ u(N,N) &= 1 \quad \text{(игрок выиграл капитал } N\text{).} \end{cases} \tag{3.3}$$

Уравнение (3.2) — разностное второго порядка :

$$u(K+1,N) = \frac{1}{p(K)}u(K,N) - \frac{q(K)}{p(K)}u(K-1,N) \ ,$$

но в отличие от предыдущего параграфа, вместо начальных условий здесь задаются граничные (3.3). Мы будем решать задачу в ее “стационарном” варианте, т.е. считая вероятность $p(K)$ не зависящей от K:

$$p(K) = p, \ q(K) = q = 1 - p \ ;$$

тогда

$$u(K+1,N)=\frac{1}{p}u(K,N)-\frac{q}{p}u(K-1,N)\ . \tag{3.4}$$

Следуя общей схеме решения такого уравнения, найдем корни характеристического полинома

$$f(\lambda)=\lambda^2-\frac{1}{p}\lambda+\frac{q}{p}=(\lambda-1)\left(\lambda-\frac{q}{p}\right)\ .$$

Дальнейший ход решения задачи (3.4)–(3.3) зависит от того, является ли 1 простым или кратным корнем полинома.

1. Пусть $q\neq p$. Применяем теорему 2.1:

$$u(K,N)=C_1 1^K+C_2(q/p)^K=C_1+C_2(q/p)^K\ .$$

Константы C_j находим из условий (3.3):

$$\left\{\begin{array}{ccc} C_1+C_2 & = & 0\\ C_1+(q/p)^N C_2 & = & 1\end{array}\right. \Longrightarrow C_1=-C_2=\frac{1}{1-(q/p)^N}$$

$$\Longrightarrow u(K,N)=\frac{1-(q/p)^K}{1-(q/p)^N}\quad \text{при } K\in\{0,\dots,N\}\ .$$

2. Пусть теперь $q=p={}^1\!/_2$ (проигрыш и выигрыш равновероятны). Применяем теорему 2.2: $u(K,N)=(C_1+C_2K)\cdot 1^{K-2}$; константы находим из условий (3.3): $C_1=0, C_2=1/N$. Следовательно:

$$u(K,N)=K/N\quad \text{при } K\in\{0,\dots,N\}\ .$$

Теперь проанализируем полученные решения. Во втором случае выигрыш тем вероятнее, чем больше стартового капитала имеет игрок, и при $K>N/2$ игрок скорее выиграет. В первом случае анализ ситуации несколько сложнее, хотя зависимость вероятности выигрыша от размера стартового капитала сохраняется. Рассмотрим более интересный случай $p<q$: на каждом шаге проигрыш вероятнее выигрыша (“игра наперекор судьбе”). Пусть N достаточно велико. Существуют ли капиталы K при которых $u(K,N)$ становится больше ${}^1\!/_2$? Оказывается, не всегда. Так, при $N=10, p={}^1\!/_4, q={}^3\!/_4$ размер нужного капитала K должен быть равен, фактически, 10, но тогда можно и не играть! При $N=100, p={}^2\!/_5, q={}^3\!/_5$ и при капитале $K=99$ игроку имеет смысл вести игру: его выигрыш до 100 более вероятен, чем разорение “до нитки”. А вот при $K=98$ лучше не рисковать!

Упражнение 3.2. *При каком соотношении p и q ($p<q$) существуют капиталы $K\in\mathbb{N}$, с которыми можно вступать в игру? Конечный капитал $N\in\mathbb{N}$ считается достаточно большим.*

3.2 Цепи Маркова

Пусть имеется некоторая физическая система, которая по своей природе ведет себя случайным образом: в каждый момент времени находится в одном из n возможных и несовместимых состояний $S_1, \ldots, S_n$, которые сменяют друг друга через конечные промежутки времени по следующему закону: в некоторый начальный момент t_0 вероятности этих состояний равны соответственно $p_{01}, \ldots, p_{0n}$, а в каждый момент $t_\ell = t_1, t_2, \ldots$ вероятность системе, находившейся в состоянии S_j, перейти в состояние S_k в следующий момент $t_{\ell+1}$ равна числу p_{jk}, $0 \le p_{jk} \le 1$. Это число p_{jk} не зависит от состояний системы во все предшествующие моменты времени $t_{\ell-1}, t_{\ell-2}, \ldots, t_0$. Известно также, что в одно из состояний система перейдет наверняка:

$$p_{j1} + \cdots + p_{jn} = 1 \quad \text{при } j \in \{1, \ldots, n\} \ . \tag{3.5}$$

Составим из этих вероятностей матрицу.

Определение. Матрица из схемы

$$\begin{array}{c} \\ S_1 \\ S_2 \\ \vdots \\ S_n \end{array} \begin{array}{c} \begin{array}{cccc} S_1 & S_2 & \ldots & S_n \end{array} \\ \left(\begin{array}{cccc} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \cdots & & & \cdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{array} \right) \end{array}$$

называется матрицей **переходных вероятностей**, будем обозначать ее $\mathfrak{P}$; величины $p_{01}, \ldots, p_{0n}$ назовем **начальными вероятностями**. Определенная этими правилами смена состояний системы называется простой однородной **цепью Маркова** $\mathfrak{C}_n$ (с конечным числом состояний).

Пример 3.3. *В так называемом братско–сестринском скрещивании скрещиваются две особи, и среди их прямых потомков случайны образом выбираются две особи разного пола. Они вновь скрещиваются, и процесс этот продолжается бесконечно. Имея три генотипа* AA, Aa, aa *для каждого родителя, мы должны различать шесть комбинаций родителей, которые пометим следующим образом:*

$$S_1 = \mathsf{AA}\times\mathsf{AA},\ S_2 = \mathsf{AA}\times\mathsf{Aa},\ S_3 = \mathsf{Aa}\times\mathsf{Aa},\ S_4 = \mathsf{Aa}\times\mathsf{aa},\ S_5 = \mathsf{aa}\times\mathsf{aa},\ S_6 = \mathsf{AA}\times\mathsf{aa}\ .$$

Генотип потомка зависит от случайного процесса. При любых обстоятельствах каждый родительский ген может передаваться с вероятностью $1/2$, *и последовательные испытания независимы. Иначе говоря, генотипы потомков можно представить как результат независимых испытаний, каждое из которых эквивалентно бросанию пары монет. Например, при скрещивании особей* Aa × Aa *могут появиться генотипы* AA, Aa *и* aa *с вероятностями равными соответственно* $1/4$, $1/2$, $1/4$. *Скрещивание* AA × aa *может привести к появлению лишь особей* Aa *и т.п. Матрица* *переходных вероятностей будет иметь вид*

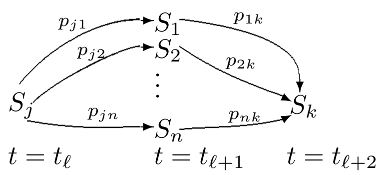
$$\mathfrak{P} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1/4 & 1/2 & 1/4 & 0 & 0 & 0 \\ 1/16 & 1/4 & 1/4 & 1/4 & 1/16 & 1/8 \\ 0 & 0 & 1/4 & 1/2 & 1/4 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \ .$$

Пример 3.4. *Предположим, что студенческая жизнь в каждый момент времени состоит из трех возможных состояний:* **лекция**, **сон**, **буфет**. *Пусть известны привычки студента, т.е. задана матрица переходных вероятностей* $\mathfrak{P}$*:*

$$\begin{array}{r c} & \begin{array}{ccc} \text{лекция} & \text{сон} & \text{буфет} \end{array} \\ \begin{array}{r} \text{лекция} \\ \text{сон} \\ \text{буфет} \end{array} & \begin{pmatrix} 1/2 & 1/4 & 1/4 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/6 & 1/2 \end{pmatrix} \end{array} \ .$$

Требуется определить "судьбу" студента, т.е. вероятность нахождения в одном из состояний по истечении достаточно большого промежутка времени.

Переходные вероятности p_{jk} можно назвать вероятностями перехода системы из S_j в S_k за один шаг. Найдем теперь вероятности перехода $p_{jk}^{(2)}$ системы из S_j в S_k за два шага.



На основании теоремы 3.2 вероятности каждого перехода:

переход	вероятность
$S_j \to S_1 \to S_k$	$p_{j1}p_{1k}$
$S_j \to S_2 \to S_k$	$p_{j2}p_{2k}$
$\dots$	$\dots$
$S_j \to S_n \to S_k$	$p_{jn}p_{nk}$

В соответствии с теоремой 3.1, чтобы получить вероятность перехода из S_j в S_k за два шага надо просуммировать полученные выражения:

$$p_{jk}^{(2)} = \sum_{i=1}^{n} p_{ji}p_{ik} \ . \tag{3.6}$$

Вывод: матрица $\mathfrak{P}^{(2)}$ переходных вероятностей за два шага состоит из элементов, которые формируются по закону перемножения матриц $\mathfrak{P} \cdot \mathfrak{P}$:

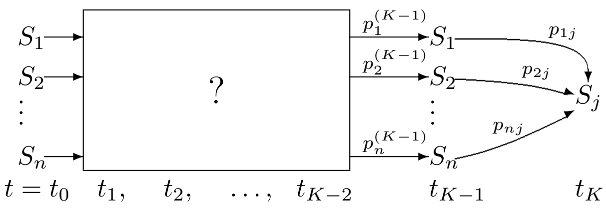
$$\mathfrak{P}^{(2)} = \mathfrak{P}^2 \ .$$

Этот результат обобщается на случай когда ищутся переходные вероятности за K шагов:

$$\mathfrak{P}^{(K)} = \mathfrak{P}^K \ , \quad K \in \mathbb{N} \ . \tag{3.7}$$

Определение. **Абсолютной вероятностью** $p_j^{(K)}$ состояния S_j системы в момент $t = t_K$ называется вероятность этого состояния независимо от того, в каких состояниях находилась система в моменты времени $t = t_{K-1}, \dots, t_1$.

Закон формирования $p_j^{(K)}$ очевиден:



то есть:

$$p_j^{(K)} = p_1^{(K-1)} p_{1j} + \dots + p_n^{(K-1)} p_{nj} = \left[p_1^{(K-1)}, \dots, p_n^{(K-1)} \right] \begin{pmatrix} p_{1j} \\ \dots \\ p_{nj} \end{pmatrix} \ .$$

Собираем абсолютные вероятности для всех состояний $S_1, \dots, S_n$ в вектор:

$$\left[p_1^{(K)}, \dots, p_n^{(K)} \right] = \left[p_1^{(K-1)}, \dots, p_n^{(K-1)} \right] \mathfrak{P} \ .$$

Рекурсивно по K применяем последнюю формулу:

$$\left[p_1^{(K)}, \dots, p_n^{(K)} \right] = \left[p_1^{(K-2)}, \dots, p_n^{(K-2)} \right] \mathfrak{P}^2 = \dots = [p_{01}, \dots, p_{0n}] \mathfrak{P}^K \ . \tag{3.8}$$

Задача. Найти $\lim\limits_{K \to +\infty} p_j^{(K)}$ для $j \in \{1, \dots, n\}$.

Определение. Если этот предел существует, то его величина p_j^{∞} называется **финальной** (предельной) абсолютной вероятностью.

Из формулы (3.8) следует, что существование p_j^{∞} зависит от существования $\lim\limits_{K \to +\infty} \mathfrak{P}^K$, т.е. мы пришли снова к задаче §2.2. Исследование начинаем с собственных чисел матрицы $\mathfrak{P}$.

Определение. Матрица A называется **неотрицательной** (**положительной**) если все ее элементы неотрицательны (положительны): $A \ge \mathbb{O}$ ($A > \mathbb{O}$). Квадратная неотрицательная матрица A называется **стохастической** если равна 1 сумма элементов каждой ее строки[31].

Матрица $\mathfrak{P}$ переходных вероятностей — стохастическая (и в дальнейшем всякую стохастическую матрицу будем обозначать $\mathfrak{P}$).

Теорема 3.3. *Если $\mathfrak{P}$ — стохастическая, то и $\mathfrak{P}^K$ — стохастическая.*

[31] Иногда добавляется еще требование: в каждом столбце существует хотя бы один ненулевой элемент.

Доказательство . Элементы j-й строки матрицы $\mathfrak{P}^2$ можно найти по формуле (3.6). Их неотрицательность очевидна. Просуммируем:

$$\sum_{\ell=1}^{n} p_{j\ell}^{(2)} = \sum_{\ell=1}^{n}\sum_{i=1}^{n} p_{ji}p_{i\ell} = \sum_{i=1}^{n}\sum_{\ell=1}^{n} p_{ji}p_{i\ell} = \sum_{i=1}^{n}\left(p_{ji}\sum_{\ell=1}^{n} p_{i\ell}\right) \overset{(3.5)}{=} \sum_{i=1}^{n} p_{ji} \overset{(3.5)}{=} 1 \ .$$

Следовательно, $\mathfrak{P}^2$ — стохастическая. Доказательство для общего случая проводится индукцией по K. □

Теорема 3.4. *Все собственные числа стохастической матрицы не превосходят по модулю* 1, *и по крайней мере одно из них равно* 1.

Доказательство . Действительно, согласно теореме 6.3 главы 3 (Гершгорина), собственное число $\lambda \in \mathbb{C}$ стохастической матрицы $\mathfrak{P}$ должно удовлетворять неравенству

$$|\lambda - p_{jj}| \le \sum_{k\ne j} |p_{jk}| = 1 - p_{jj}$$

хотя бы при одном j. Воспользовавшись следствием к неравенству треугольника имеем:

$$|\lambda| - |p_{jj}| \le |\lambda - p_{jj}| \le 1 - p_{jj} \ \Rightarrow \ |\lambda| \le 1 \ .$$

Далее,

$$\mathfrak{P}\begin{pmatrix}1\\1\\ \vdots\\1\end{pmatrix} = \begin{pmatrix}p_{11}+\dots+p_{1n}\\ p_{21}+\dots+p_{2n}\\ \dots\\ p_{n1}+\dots+p_{nn}\end{pmatrix} \overset{(3.5)}{=} \begin{pmatrix}1\\1\\ \vdots\\1\end{pmatrix},$$

т.е. $\lambda = 1$ — собственное число, соответствующее собственному вектору $X = [1,1,\dots,1]^{\top}$. □

Теорема 3.5. *Если собственное число* 1 *— простое для матрицы* $\mathfrak{P}$, *а модули всех остальных собственных чисел меньше* 1, *то существует конечная матрица*

$$\mathfrak{P}^{\infty} \overset{\mathrm{def}}{=} \lim_{K\to+\infty} \mathfrak{P}^K \ .$$

Доказательство . По теореме 3.4 имеем следующую структуру для **ж.н.ф.** $\mathfrak{P}_{\mathfrak{J}}$ и матрицы перехода к каноническому базису:

$$\mathfrak{P}_{\mathfrak{J}} = \begin{pmatrix}1 & 0 & \dots & 0\\ 0 & & \text{клетки} & \\ \vdots & & \text{для } \lambda_j & \\ 0 & & |\lambda_j|<1 & \end{pmatrix}, \quad C = \begin{pmatrix}1 & \text{корневые}\\ 1 & \text{векторы}\\ \vdots & \text{для } \lambda_j\\ 1 & |\lambda_j|<1\end{pmatrix}$$

На основании формулы (1.2):

$$\mathfrak{P}_{\mathfrak{J}}^{\infty} \overset{\mathrm{def}}{=} \lim_{K\to+\infty} \mathfrak{P}_{\mathfrak{J}}^K = \begin{pmatrix}1 & 0 & \dots & 0\\ 0 & & & \\ \vdots & & \mathbb{O} & \\ 0 & & & \end{pmatrix}.$$

Но тогда существует $\lim\limits_{K\to+\infty} \mathfrak{P}^K$ и он равен

$$C\mathfrak{P}_{\mathfrak{J}}^{\infty}C^{-1} = \begin{pmatrix} \tau_{11} & \tau_{12} & \dots & \tau_{1n} \\ \tau_{11} & \tau_{12} & \dots & \tau_{1n} \\ \dots & & & \dots \\ \tau_{11} & \tau_{12} & \dots & \tau_{1n} \end{pmatrix},$$

где $[\tau_{11}, \tau_{12}, \dots, \tau_{1n}]$ — первая строка матрицы C^{-1}. Получили, на первый взгляд, довольно неожиданный ответ: все строки матрицы $\mathfrak{P}^{\infty}$ должны быть одинаковыми! Кроме того, на основании теоремы 3.3, свойство стохастичности $\mathfrak{P}^{\infty}$ должно сохраняться: $\tau_{11} + \tau_{12} + \dots + \tau_{1n} = 1$. □

Определение. Стохастическая матрица $\mathfrak{P}$, удовлетворяющая условию теоремы 3.5 называется **регулярной** (и такая же — цепь $\mathfrak{C}_n$).

Следствие 1. *Для регулярной цепи $\mathfrak{C}_n$ финальные абсолютные вероятности находятся по формулам*

$$p_j^{\infty} = \tau_{1j} \quad \text{для } j \in \{1, \dots, n\} \ , \tag{3.9}$$

и они не зависят от начальных $p_{01}, \dots, p_{0n}$.

Иначе: для момента времени t_K достаточно удаленного от любого исходного момента поведение системы становится почти независимым от исходного состояния.

Пример 3.5. *Для стохастической матрицы*

$$\mathfrak{P} = \begin{pmatrix} 0.54580 & 0.36663 & 0.03519 & 0.05238 \\ 0.04406 & 0.00738 & 0.13600 & 0.81256 \\ 0.42015 & 0.27010 & 0.12121 & 0.18854 \\ 0.85136 & 0.00749 & 0.00123 & 0.13992 \end{pmatrix}$$

имеем:

$$\mathfrak{P}^{32} = \begin{pmatrix} 0.50875694_{79} & 0.203905925_{3} & 0.0522576617_{4} & 0.23507946_{60} \\ 0.50875694_{76} & 0.203905925_{0} & 0.0522576617_{9} & 0.23507946_{63} \\ 0.50875694_{78} & 0.203905925_{2} & 0.0522576617_{5} & 0.23507946_{61} \\ 0.50875694_{81} & 0.203905925_{1} & 0.0522576617_{2} & 0.23507946_{59} \end{pmatrix} .$$

Реальный расчет финальных вероятностей по формулам (3.9) затруднителен, более конструктивен следующий результат.

Теорема 3.6. *Пусть $\mathfrak{P}$ — регулярна. Тогда*

$$\mathfrak{P}^{\infty} = \frac{f_1(\mathfrak{P})}{f'(1)} \quad , \text{где} \quad f_1(\lambda) \stackrel{\text{def}}{=} \frac{\det(\mathfrak{P} - \lambda E)}{\lambda - 1} \ .$$

Доказательство проведем в дополнительном предположении простоты всех собственных чисел $f(\lambda)$. Воспользуемся формулой (1.6) в применении к полиному $g(x) \equiv x^K$. Положим $\lambda_1 = 1$, $f_j(\lambda) = \det(\mathfrak{P} - \lambda E)/(\lambda - \lambda_j)$:

$$\mathfrak{P}^K = 1^K \frac{f_1(\mathfrak{P})}{f'(1)} + \sum_{j=2}^{n} \lambda_j^K \frac{f_j(\mathfrak{P})}{f'(\lambda_j)} \ .$$

Поскольку $|\lambda_j| < 1$ при $j \in \{2, \dots, n\}$, то при $K \to \infty$ каждое слагаемое в сумме стремится к нулю. □

Следствие 1. *Если хар. полином матрицы* $\mathfrak{P}$ *равен*

$$f(\lambda) = a_0\lambda^n + a_1\lambda^{n-1} + \dots + a_n \ (a_0 = (-1)^n) \ ,$$

то $f_1(\lambda) = a_0\lambda^{n-1} + (a_0 + a_1)\lambda^{n-2} + \dots + (a_0 + a_1 + \dots + a_{n-1})$.

С помощью результата упражнения 6.1 главы 3 можем сформулировать еще один способ нахождения финальных абсолютных вероятностей.

Следствие 2. *Вектор* $[p_1^\infty, \dots, p_n^\infty]^\top$ *является собственным вектором матрицы* $\mathfrak{P}^\top$*, принадлежащим собственному числу* $\lambda = 1$.

Особо подчеркнем: из всего множества собственных векторов $\mathfrak{P}^\top$, принадлежащих $\lambda = 1$, нас интересует тот единственный, что имеет сумму компонент равной единице.

Пример 3.6. *Найти "судьбу" студента из примера* 3.4*, т.е. определить финальные абсолютные вероятности.*

Решение. **1-й способ.** $f(\lambda) = \det(\mathfrak{P} - \lambda E) = -\lambda^3 + {}^4\!/_3\,\lambda^2 - {}^{13}\!/_{36}\,\lambda + {}^1\!/_{36}$,

$$f'(1) = -{}^{25}\!/_{36} \ne 0 \quad \text{и} \quad f_1(\lambda) = -\lambda^2 + {}^1\!/_3\,\lambda - {}^1\!/_{36}$$

имеет все корни по модулю не превосходящими 1 (в данном примере это устанавливается непосредственным вычислением: $\lambda_{2,3} = {}^1\!/_6$; в общем же случае можно воспользоваться критерием Шура–Кона). Следовательно, матрица $\mathfrak{P}$ регулярна. Применение теоремы 3.6 дает:

$$\mathfrak{P}^\infty = \begin{pmatrix} 2/5 & 6/25 & 9/25 \\ 2/5 & 6/25 & 9/25 \\ 2/5 & 6/25 & 9/25 \end{pmatrix} .$$

2-й способ. Следствие 2 дает вектор финальных абсолютных вероятностей непосредственно из системы $(\mathfrak{P}^\top - E)X = \mathbb{O}$:

$$\begin{pmatrix} -1/2 & 1/3 & 1/3 \\ 1/4 & -2/3 & 1/6 \\ 1/4 & 1/3 & -1/2 \end{pmatrix} X = \mathbb{O} \qquad \Rightarrow \quad X = \alpha \begin{pmatrix} 10 \\ 6 \\ 9 \end{pmatrix} \ .$$

Из дополнительного условия $x_1 + x_2 + x_3 = 1$ получим $\alpha = {}^1\!/_{25}$.

Ответ. $p_1^\infty = {}^2\!/_5$, $p_2^\infty = {}^6\!/_{25}$, $p_3^\infty = {}^9\!/_{25}$. Студент попался трудолюбивый: наибольшая вероятность застать его на лекции.

Упражнение 3.3. *Определить финальные абсолютные вероятности для цепи Маркова, заданной матрицей*

$$\mathfrak{P} = {}^1\!/_{15} \begin{pmatrix} 1 & 2 & 4 & 6 & 2 \\ 4 & 1 & 2 & 2 & 6 \\ 2 & 4 & 6 & 1 & 2 \\ 2 & 1 & 4 & 2 & 6 \\ 6 & 2 & 2 & 4 & 1 \end{pmatrix} .$$

Для нерегулярной цепи финальные абсолютные вероятности будут зависеть от начальных.

Пример 3.7. *Определить финальные абсолютные вероятности для примера* 3.3*, если* $[{}^1\!/_4, {}^1\!/_3, {}^1\!/_6, {}^1\!/_6, 0, {}^1\!/_{12}]$ *— вектор начальных вероятностей распределения генотипов в популяции.*

Решение. $\det(\mathfrak{P} - \lambda E) = (\lambda - {}^1\!/_4)\,(\lambda - {}^1\!/_2)\,\left(\lambda^2 - {}^1\!/_2\,\lambda - {}^1\!/_4\right)(\lambda - 1)^2$.

$$\mathfrak{P}^\infty = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ {}^3\!/_4 & 0 & 0 & 0 & {}^1\!/_4 & 0 \\ {}^1\!/_2 & 0 & 0 & 0 & {}^1\!/_2 & 0 \\ {}^1\!/_4 & 0 & 0 & 0 & {}^3\!/_4 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ {}^1\!/_2 & 0 & 0 & 0 & {}^1\!/_2 & 0 \end{pmatrix} .$$

Ответ. $[{}^2\!/_3, 0, 0, 0, {}^1\!/_3, 0]$. И это при том, что исходно в популяции не предполагается скрещивания особей с генотипами, соответствующими S_5, т.е. $\mathsf{aa} \times \mathsf{aa}$.

Историческая справка

Марков Андрей Андреевич (1856–1922), ученик П.Л.Чебышева. Работы по теории чисел, интерполированию, теории вероятностей. Его брат Владимир (1871–1897), и сын Андрей (1903–1979) тоже стали математиками.

Человек открытый, прямой, смелый. Его гражданское мужество было очень стойким: он не считался ни с лицами, против которых выступал, ни с последствиями, которые его высказывания могли иметь для него самого. Когда ему однажды возразили, что его предложение идет вразрез с “высочайшим постановлением”, он во всеуслышание сказал:

“Я вам дело говорю, а вы мне — высочайшее постановление!”

Как-то петербургский городской голова приветствовал “от имени всех жителей Санкт-Петербурга” приехавшую в столицу государыню — императрицу Александру Федоровну. Текст приветствия был отпечатан в газетах. А.А.Марков послал в редакцию опровержение, в котором просил сообщить, что он, Андрей Андреевич Марков, вовсе не уполномочивал городского голову приветствовать императрицу от его имени.

Из рассказов академика В.И.Смирнова (1887-1974): “А знаете у кого я слушал курс теории вероятностей? У самого Андрея Андреевича Маркова (старшего). О, это был человек крутого нрава! Он начал свою первую лекцию так:

"Господа! Некто Гильберт недавно выдвинул перед математиками всего мира 23 задачи, объявив их почему-то наиболее важными[32] *. И среди них имелось предложение переделать теорию вероятностей на аксиоматической основе. К сожалению, наше математическое общество сочло возможным согласиться с господином Гильбертом и рекомендовало вниманию своих членов все без исключения придуманные им задачи. В знак протеста я немедленно вышел из этого общества!"*

Здесь Андрей Андреевич медленно обвел глазами аудиторию и, немного помолчав, сказал:

"А теперь приступим к делу." ...

Синод Русской Православной Церкви отлучил от церкви Л.Н.Толстого. 12 февраля 1912 г. Марков сам обратился в синод с просьбой отлучить его от церкви:

"... покорнейше прошу принять во внимание, что я не усматриваю существенной разницы между иконами и идолами,... и не сочувствую всем религиям, которые, подобно православию, поддерживаются огнем и мечом и сами служат им".

После революции семья бедствовала. 5 марта 1921 г. Марков сообщает Академии, что он не может посещать ее заседания из-за отсутствия обуви. Спустя пару недель Комиссия по улучшению быта ученых, заседавшая под председательством М.Горького, удовлетворила просьбу известного математика. Марков:

"Наконец я получил обувь; но она не только дурно сшита, но совершенно мне не подходит по своим размерам ... Я предлагаю поместить [ее] в Этнографическом музее как образец материальной культуры настоящего момента, ради чего я готов ею пожертвовать ... "

Но жизненные невзгоды не сломили его духа, не изменили его убеждения. В 1921 г. большевики попытались организовать "рабоче-крестьянские"наборы в университеты. Группа профессоров Петроградского университета, включавшая А.А.Маркова, В.А.Стеклова и Н.М.Гюнтера подписала заявление:

"Ввиду того что для успешности занятий в университете студенты должны иметь лишь соответствующую подготовку, прием слушателей в университет должен производиться согласно их знаниям, а не по каким либо классовым или политическим соображениям".

4 Матричный степенной ряд

4.1 Норма матрицы. Матричный ряд

Мы рассмотрим сначала произвольные (не обязательно квадратные) матрицы над $\mathbb{C}$.

ОПРЕДЕЛЕНИЕ. **Нормой** матрицы $\mathbf{A}$ называется неотрицательное число $\|\mathbf{A}\|$, удовлетворяющее условиям (аксиомам):

1. $\|\mathbf{A}\| = 0$ тогда и только тогда, когда $\mathbf{A} = \mathbb{O}_{n\times n}$;

2. для $\forall \alpha \in \mathbb{C}$ справедливо $\|\alpha\mathbf{A}\| = |\alpha| \cdot \|\mathbf{A}\|$;

[32]**Гильберт Давид** (Hilbert David, 1862–1943), крупнейший немецкий математик. В докладе "Математические проблемы", прочитанном 8 августа 1900 г. на II Международном Конгрессе математиков в Париже привел список из 23-х проблем, " ..., *исследование которых может значительно стимулировать дальнейшее развитие науки*". 6-я из этих проблем : "Математическое изложение аксиом физики" ставила целью создание аксиоматики теории вероятностей и механики.

3. для любых матриц $\mathbf{A}$ и $\mathbf{B}$ одинаковых порядков справедливо $\|\mathbf{A}+\mathbf{B}\| \le \|\mathbf{A}\| + \|\mathbf{B}\|$;
4. для любых матриц $\mathbf{A}$ и $\mathbf{B}$, допускающих умножение, справедливо $\|\mathbf{AB}\| \le \|\mathbf{A}\| \cdot \|\mathbf{B}\|$.

Норму можно вводить разными способами

I. Евклидова норма:

$$\|\mathbf{A}\| \overset{\text{def}}{=} \sqrt{\sum_{j,k} |a_{jk}|^2} \overset{\mathbb{R}}{=} \sqrt{\mathrm{Sp}\,(A \cdot A^{\top})} \ .$$

Выполнение для нее аксиом 1 и 2 очевидно, справедливость 3 следует из неравенства треугольника: $|a_{jk}+b_{jk}|^2 \le |a_{jk}|^2 + |b_{jk}|^2$, а 4 следует из неравенства Коши[33]:

$$|a_{j1}b_{1k} + \dots + a_{jn}b_{nk}|^2 \le \left(|a_{j1}|^2 + \dots + |a_{jn}|^2\right)\left(|a_{1k}|^2 + \dots + |a_{nk}|^2\right) \ .$$

Эту норму можно рассматривать как обобщение понятия длины вектора в $\mathbb{R}^n$: $\|X_{n\times 1}\| = \sqrt{x_1^2 + \dots + x_n^2} = |X|$. Вообще, любая формула, задающая скалярное произведение в евклидовом пространстве матриц, порождает и норму матрицы: $\|\mathbf{A}\| = \sqrt{(A,A)}$. Только что введенная норма соответствует формуле (1.5) главы 2.

II. $\|\mathbf{A}\| \overset{\text{def}}{=} \max\limits_{j} \sum\limits_{k} |a_{jk}|$. Тогда $\|X_{n\times 1}\| = \max\limits_{j=1,\dots,n} |x_j|$.

III. $\|\mathbf{A}\| \overset{\text{def}}{=} \max\limits_{k} \sum\limits_{j} |a_{jk}|$. Тогда $\|X_{n\times 1}\| = \sum\limits_{j=1}^{n} |x_j|$.

Введение понятия нормы позволяет упростить рассуждения о сходимости матричной последовательности $\{\mathbf{A}_N\}_{N=1}^{\infty}$. В предыдущих параграфах эта сходимость понималась в смысле существования пределов у последовательностей элементов $a_{jk}^{(N)}$ одновременно для всех индексов j и k. Норма позволяет объединить исследование этих последовательностей на сходимость в изучение одной.

Теорема 4.1. *Последовательность $\{\mathbf{A}_N\}_{N=1}^{\infty}$ сходится к матрице $\mathbf{A}$ тогда и только тогда, когда*

$$\lim_{N\to\infty} \|\mathbf{A}_N - \mathbf{A}\| = 0 \ .$$

Доказательство . Действительно, например для нормы I:

$$\|\mathbf{A}_N - \mathbf{A}\| = \sqrt{\sum_{jk} |a_{jk}^{(N)} - a_{jk}|^2} \underset{N\to\infty}{\to} 0$$

тогда и только тогда, когда $a_{jk}^{(N)} \underset{N\to\infty}{\to} a_{jk}$ для всех j и k. □

[33]Раздел I, глава 4, теорема 8.2.

Определение. Матричный ряд

$$\sum_{j=1}^{\infty} \mathbf{A}_j \stackrel{\text{def}}{=} \mathbf{A}_1 + \cdots + \mathbf{A}_N + \ldots \tag{4.1}$$

называется **сходящимся** если существует конечный предел последовательности

$$\left\{ \mathbf{S}_N \stackrel{\text{def}}{=} \sum_{j=1}^{N} \mathbf{A}_j = \mathbf{A}_1 + \cdots + \mathbf{A}_N \right\}_{N=1}^{\infty}$$

его частичных сумм. Тогда величина этого предела называется **суммой** матричного ряда:

$$\sum_{j=1}^{\infty} \mathbf{A}_j \stackrel{\text{def}}{=} \mathbf{S} = \lim_{N\to\infty} \mathbf{S}_N \ .$$

Матричный ряд (4.1) называется **расходящимся** если хотя бы для одной пары индексов j и k последовательность $s_{jk}^{(N)}$ расходится. Матричный ряд (4.1) называется **абсолютно сходящимся** если сходится числовой ряд его норм: $\sum\limits_{j=1}^{\infty} \|\mathbf{A}_j\|$.

Теорема 4.2. *Если ряд* (4.1) *сходится абсолютно, то он сходится и в обычном смысле.*

Теорема 4.3 (Признак сравнения). *Если ряд* $\sum\limits_{j=1}^{\infty} \mathbf{A}_j$ *абсолютно сходится и* $\|\mathbf{A}_j\| \ge \|\mathbf{B}_j\|$ *для* $\forall j \in \mathbb{N}$, *то и ряд* $\sum\limits_{j=1}^{\infty} \mathbf{B}_j$ *сходится абсолютно.*

4.2 Матричный степенной ряд

Рассмотрим теперь степенной ряд

$$\sum_{j=0}^{\infty} b_j z^j = b_0 + b_1 z + b_2 z^2 + \ldots \quad \text{при} \quad z = x + iy \in \mathbb{C} \tag{4.2}$$

и соответствующий матричный **степенной** ряд

$$\sum_{j=0}^{\infty} b_j \mathbf{Z}^j = b_0 E + b_1 \mathbf{Z} + b_2 \mathbf{Z}^2 + \ldots \quad \text{при} \quad \mathbf{Z} - n \times n \quad \text{матрице над } \mathbb{C} \tag{4.3}$$

Пусть R — радиус сходимости ряда (4.2), т.е. ряд сходится при $|z| < R$ и расходится при $|z| > R$. Известно, что

$$R = \overline{\lim_{n\to\infty}} \left| \frac{b_n}{b_{n+1}} \right| = \overline{\lim_{n\to\infty}} \frac{1}{\sqrt[n]{|b_n|}} \ .$$

Теорема 4.4. *Матричный степенной ряд* (4.3) *сходится абсолютно для любой матрицы* $\mathbf{Z}$ *такой, что* $\|\mathbf{Z}\| < R$.

Доказательство . Имеем

$$\sum_{j=0}^{\infty} \|b_j \mathbf{Z}^j\| \le \sum_{j=0}^{\infty} |b_j| \, \|\mathbf{Z}\|^j$$

Числовой ряд в правой части сходится при $\|\mathbf{Z}\| < R$. На основании теорем 4.2 и 4.3 сходится и ряд (4.3). □

Пример 4.1. *Ряд*

$$\sum_{j=0}^{\infty} z^j = 1 + z + z^2 + \dots + z^j + \dots$$

сходится при $|z| < 1$ *к функции* $1/(1-z)$. *На основании теоремы* 4.4 *можно утверждать, что матричный ряд*

$$\sum_{j=0}^{\infty} \mathbf{Z}^j = E + \mathbf{Z} + \mathbf{Z}^2 + \dots + \mathbf{Z}^j + \dots \tag{4.4}$$

сходится абсолютно для любой матрицы $\mathbf{Z}$ *такой, что* $\|\mathbf{Z}\| < 1$.

Следствие 1. *Если ряд* (4.2) *сходится при* $\forall z \in \mathbb{C}$, *то и ряд* (4.3) *сходится при любой матрице* $\mathbf{Z}$.

Определение. Сумма сходящегося степенного ряда (4.3) будет функцией переменной матрицы $\mathbf{Z}$. Такая функция называется **аналитической** функцией матрицы.

Если обозначить сумму ряда (4.4) через $F(\mathbf{Z})$, то легко проверяется свойство: $F(\mathbf{Z})(E-\mathbf{Z}) = E$. Таким образом, ряд (4.4) сходится к матрице $(E-\mathbf{Z})^{-1}$ при $\|\mathbf{Z}\| < 1$.

Теорема 4.5. *Если* $\mathbf{A} \doteq \mathbf{B}$ *и* F *— аналитическая функция, определенная для* $\mathbf{A}$, *то тогда она будет определена и для* $\mathbf{B}$ *и* $F(\mathbf{A}) \doteq F(\mathbf{B})$.

Доказательство . По условию $\mathbf{B} = C^{-1}\mathbf{A}C$ при некоторой неособенной матрице C. Рассмотрим N-ю частичную сумму ряда (4.3):

$$F_N(Z) \stackrel{\mathrm{def}}{=} \sum_{j=0}^{N} a_j \mathbf{Z}^j$$

Для матричных полиномов $F_N(\mathbf{A})$ и $F_N(\mathbf{B})$ будет выполнено

$$F_N(\mathbf{B}) = C^{-1} F_N(\mathbf{A}) C \ .$$

Переходя к пределу, получаем

$$\lim_{N\to\infty} F_N(\mathbf{B}) = C^{-1} \left(\lim_{N\to\infty} F_N(\mathbf{A}) \right) C = C^{-1} F(\mathbf{A}) C \ .$$

Таким образом, функция F определена и для $\mathbf{B}$ и $F(\mathbf{B}) = C^{-1}F(\mathbf{A})C$. □

Теорема 4.6. *Ряд (4.3) сходится для любой матрицы* $\mathbf{A}$ *чей спектр лежит внутри круга сходимости:* $|\lambda_j| < R$ *при* $j \in \{1, \dots, n\}$*; и расходится если хотя бы одно число оказывается за пределами этого круга:* $\exists j \colon |\lambda_j| > R$.

Доказательство . Для матрицы $\mathbf{A}$ найдем **ж.н.ф.**:

$$C^{-1}\mathbf{A}C = \mathbf{A}_{\mathfrak{J}} = \left(\begin{array}{cccc} \mathsf{A}_1 & \mathbb{O} & \dots & \mathbb{O} \\ \mathbb{O} & \mathsf{A}_2 & \dots & \mathbb{O} \\ & & \ddots & \\ \mathbb{O} & \mathbb{O} & \dots & \mathsf{A}_{\mathfrak{r}} \end{array} \right)$$

где каждая из составляющих матриц A_j включает в себя некоторое количество клеток Жордана вида

$$\mathfrak{J}_k(\lambda_j) = \left(\begin{array}{ccccc} \lambda_j & & & & \\ 1 & \lambda_j & & \mathbb{O} & \\ 0 & 1 & \lambda_j & & \\ \vdots & & \ddots & \ddots & \\ 0 & 0 & \dots & 1 & \lambda_j \end{array} \right)_{k\times k} .$$

Если мы докажем, что существует $F\left(\mathbf{A}_{\mathfrak{J}}\right)$ то утверждение теоремы будет следовать из теоремы 4.5. Чтобы избежать громоздкости будем опускать индекс у λ_j. Вновь рассмотрим N-ю частичную сумму ряда (4.3). Вычисление матричного полинома $F_N\left(\mathbf{A}_{\mathfrak{J}}\right)$ сводится к вычислению его значения на клетке Жордана. На основании формулы (1.4) получаем: $F_N\left(\mathfrak{J}_k(\lambda)\right) =$

$$= \left[\begin{array}{ccccc} F_N(\lambda) & & & & \\ F_N'(\lambda) & F_N(\lambda) & & \mathbb{O} & \\ \frac{F_N''(\lambda)}{2!} & F_N'(\lambda) & F_N(\lambda) & & \\ \vdots & & \ddots & \ddots & \\ \frac{F_N^{(k-1)}(\lambda)}{(k-1)!} & \frac{F_N^{(k-2)}(\lambda)}{(k-2)!} & \dots & F_N'(\lambda) & F_N(\lambda) \end{array} \right] \underset{N\to\infty}{\longrightarrow} \left[\begin{array}{ccccc} F(\lambda) & & & & \\ F'(\lambda) & F(\lambda) & & \mathbb{O} & \\ \frac{F''(\lambda)}{2!} & F'(\lambda) & F(\lambda) & & \\ \vdots & & \ddots & \ddots & \\ \frac{F^{(k-1)}(\lambda)}{(k-1)!} & \frac{F^{(k-2)}(\lambda)}{(k-2)!} & \dots & F'(\lambda) & F(\lambda) \end{array} \right] \tag{4.5}$$

при $|\lambda| < R$. Если же хоть одно собственное число лежит вне круга сходимости, то соответствующая последовательность $\{F_N\left(\mathfrak{J}_k(\lambda)\right)\}_{N=1}^{\infty}$ будет расходящейся.

Итак, при $|\lambda_1| < R, \dots, |\lambda_n| < R$ существует

$$F(\mathbf{A}) = \lim_{N\to\infty} F_N(\mathbf{A}) = C \left(\begin{array}{cccc} F(\mathsf{A}_1) & \mathbb{O} & \dots & \mathbb{O} \\ \mathbb{O} & F(\mathsf{A}_2) & \dots & \mathbb{O} \\ & & \ddots & \\ \mathbb{O} & \mathbb{O} & \dots & F(\mathsf{A}_{\mathfrak{r}}) \end{array} \right) C^{-1} \tag{4.6}$$

при блоках матриц $F(\mathsf{A}_j)$ определяемых с помощью (4.5). □

Следствие 1. *Если* $|\lambda_1| < R, \dots, |\lambda_n| < R$*, то собственными числами матрицы* $F(\mathbf{A})$ *являются* $F(\lambda_j)$.

Пример 4.2. *Вычислить сумму ряда*

$$\sum_{j=1}^{\infty}\frac{1}{j}\begin{pmatrix} -3/2 & 0 & -4 \\ -7 & 1/2 & -14 \\ 1 & 0 & 5/2 \end{pmatrix}^j .$$

Решение. Соответствующий степенной ряд сходится при $|z|<1$ и имеет суммой $(-\ln(1-z))$. Хар. полином $f(\lambda)=-(\lambda-1/2)^3$. По теореме 4.6 матричный ряд сходится. Для вычислений по формуле (4.6) нам потребуется **ж.н.ф.** Опуская промежуточные вычисления:

$$\mathbf{A}=\underbrace{\begin{pmatrix} -2 & -1 & -2 \\ 0 & 0 & -7 \\ 1 & 1 & 1 \end{pmatrix}}_{C}\underbrace{\begin{pmatrix} 1/2 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 1 & 1/2 \end{pmatrix}}_{\mathbf{A}_{\mathfrak{J}}}\underbrace{\begin{pmatrix} -1 & 1/7 & -1 \\ 1 & 0 & 2 \\ 0 & -1/7 & 0 \end{pmatrix}}_{C^{-1}}$$

$$-\ln(E-\mathbf{A})=\begin{pmatrix} -2 & -1 & -2 \\ 0 & 0 & -7 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} \ln 2 & 0 & 0 \\ 0 & \ln 2 & 0 \\ 0 & 2 & \ln 2 \end{pmatrix}\begin{pmatrix} -1 & 1/7 & -1 \\ 1 & 0 & 2 \\ 0 & -1/7 & 0 \end{pmatrix}$$

Ответ.

$$\begin{pmatrix} \ln 2-4 & 0 & -8 \\ -14 & \ln 2 & -28 \\ 2 & 0 & \ln 2+4 \end{pmatrix} .$$

Формула

$$F(\mathbf{Z})=CF(\mathbf{Z}_{\mathfrak{J}})C^{-1} \quad \text{при} \quad F\left(\mathfrak{J}_k(\lambda)\right)=\left[\begin{array}{ccccc} F(\lambda) & & & & \\ F'(\lambda) & F(\lambda) & & \mathbb{O} & \\ \frac{F''(\lambda)}{2!} & F'(\lambda) & F(\lambda) & & \\ \vdots & & \ddots & \ddots & \\ \frac{F^{(k-1)}(\lambda)}{(k-1)!} & \frac{F^{(k-2)}(\lambda)}{(k-2)!} & \dots & F'(\lambda) & F(\lambda) \end{array}\right] \tag{4.7}$$

может быть использована и для доопределения функции от матрицы для тех матриц, которые не удовлетворяют условиям теорем 4.4 и 4.5. Действительно, если ряд $\sum_{j=1}^{\infty}\frac{z^j}{j}$ представляет функцию $(-\ln(1-z))$ только при $|z|<1$, то сама функция существует и для $z\le -1$. Если договориться положить формулу (4.7) в качестве определения функции от матрицы $\mathbf{Z}$, то эта формула будет вычислять $F(\mathbf{Z})$ для всех тех матриц, на спектрах которых определены значения функции $F(z)$ и ее производных до порядков, соответствующих порядкам клеток Жордана.

С использованием этой договоренности возможно и дальнейшее упрощение вычислений $F(\mathbf{A})$, основанное на теореме Гамильтона-Кэли. В самом деле, поскольку степени матрицы $\mathbf{A}^n,\mathbf{A}^{n+1},\dots$ линейно выражаются через степени[34] $E=\mathbf{A}^0,\mathbf{A},\mathbf{A}^2,\dots,\mathbf{A}^{n-1}$, то бесконечный ряд (4.3) может

[34]Глава 3, §6.2, следствие 3.

быть свернут в матричный полином степени не выше $n-1$. Нахождение этого полинома производится обобщением теоремы 1.2.

Лемма 1. *Если $\{\lambda_1, \ldots, \lambda_n\} \subset \mathbb{C}$ — различные корни аналитической функции $\Phi(z)$, лежащие внутри ее круга сходимости, то существует единственное представление функции в виде:*

$$\Phi(z) \equiv (z-\lambda_1) \times \cdots \times (z-\lambda_n) Q(z) \ , \tag{4.8}$$

где $Q(z)$ — аналитическая функция с тем же радиусом сходимости.

Теорема 4.7. *Если все собственные числа $\lambda_1, \ldots, \lambda_n$ матрицы $\mathbf{A}$ различны, то*

$$F(\mathbf{A}) = g(\mathbf{A}) \ ,$$

где $g(x)$ — интерполяционный полином, построенный по таблице

$$\begin{array}{c|ccc} x & \lambda_1 & \ldots & \lambda_n \\ \hline y & F(\lambda_1) & \ldots & F(\lambda_n) \end{array} \ .$$

Доказательство . Рассмотрим функцию $\Phi(z) \stackrel{\text{def}}{=} F(z) - g(z)$. Поскольку $g(\lambda_j) = F(\lambda_j)$, то $\Phi(\lambda_j) = 0$ для всех $j \in \{1, \ldots, n\}$. Тогда для $\Phi(z)$ будет выполнено (4.8). Подставим матрицу $\mathbf{A}$ в это тождество:

$$\Phi(\mathbf{A}) \equiv \underbrace{(\mathbf{A}-\lambda_1 E) \times \cdots \times (\mathbf{A}-\lambda_n E)}_{=\mathbb{O} \text{ по т. Гамильтона-Кэли}} Q(\mathbf{A})$$

Отсюда и следует равенство $F(\mathbf{A}) = g(\mathbf{A})$. □

Теорема утверждает универсальность формулы (1.6): она остается справедливой для любой функции, для которой определены значения на спектре матрицы $\mathbf{A}$. Иногда для вычисления $F(\mathbf{A})$ не требуется знания спектра, а достаточно найти только хар.полином для матрицы $\mathbf{A}$.

4.3 Дифференцирование матрицы

Рассмотрим (не обязательно квадратную) матрицу элементами которой являются функции, зависящие от переменной $t \in]a, b[\subset \mathbb{R}$: $\mathbf{A}(t) = [a_{jk}(t)]_{jk}$.

Определение. **Производной** матрицы $\mathbf{A}(t)$ по t будем называть матрицу

$$\frac{d\,\mathbf{A}(t)}{d\,t} = \mathbf{A}'(t) \stackrel{\text{def}}{=} \left[\frac{d\,a_{jk}(t)}{d\,t}\right]_{jk}$$

в предположении, что все производные в правой части существуют.

Таким образом, дифференцирование матрицы сводится к поэлементному дифференцированию. Поэтому свойства дифференцирования матриц легко выводятся из соответствующих свойств производных скалярных функций.

Теорема 4.8. *Если соответствующие матричные действия имеют смысл, то справедливы следующие соотношения*

$$\frac{d\,\mathbf{C}}{d\,t} = \mathbb{O}$$

при постоянной матрице $\mathbf{C}$ *(элементы которой не зависят от* t*);*

$$\frac{d\,(\mathbf{A}(t)+\mathbf{B}(t))}{d\,t} = \frac{d\,\mathbf{A}(t)}{d\,t} + \frac{d\,\mathbf{B}(t)}{d\,t} \ ;$$

$$\frac{d\,\mathbf{C}\mathbf{A}(t)}{d\,t} = \mathbf{C}\frac{d\,\mathbf{A}(t)}{d\,t} \ , \quad \frac{d\,\mathbf{A}(t)\mathbf{C}}{d\,t} = \frac{d\,\mathbf{A}(t)}{d\,t}\mathbf{C} \ ;$$

$$\frac{d\,(\mathbf{A}(t)\mathbf{B}(t))}{d\,t} = \frac{d\,\mathbf{A}(t)}{d\,t}\mathbf{B}(t) + \mathbf{A}(t)\frac{d\,\mathbf{B}(t)}{d\,t} \ .$$

Для квадратной матрицы $\mathbf{A}(t)$ *и* $K \in \mathbb{N}$*:*

$$\frac{d\,\mathbf{A}(t)^K}{d\,t} = \mathbf{A}'(t)\mathbf{A}(t)^{K-1} + \mathbf{A}(t)\mathbf{A}'(t)\mathbf{A}(t)^{K-2} + \cdots + \mathbf{A}(t)^{K-1}\mathbf{A}'(t) \ ,$$

а если она коммутирует со своей производной $\mathbf{A}(t)\mathbf{A}'(t) = \mathbf{A}'(t)\mathbf{A}(t)$*, то*

$$\frac{d\,\mathbf{A}(t)^K}{d\,t} = K\mathbf{A}'(t)\mathbf{A}(t)^{K-1} \ .$$

Для матрицы $\mathbf{A}(t)$*, неособенной при любых рассматриваемых значениях* t*:*

$$\frac{d\,\mathbf{A}^{-1}(t)}{d\,t} = -\mathbf{A}^{-1}(t)\frac{d\,\mathbf{A}(t)}{d\,t}\mathbf{A}^{-1}(t) \ .$$

Теорема 4.9. *Если матричный ряд* $\sum\limits_{\ell=1}^{\infty}\mathbf{A}_\ell(t)$ *сходится при* $t \in]a,b[$*, а ряд производных* $\sum\limits_{\ell=1}^{\infty}\mathbf{A}_\ell(t)'$ *сходится равномерно на* $]a,b[$ *(т.е. равномерно сходятся все функциональные ряды* $\sum\limits_{\ell=1}^{\infty}\frac{d}{d\,t}a_{jk}^{(\ell)}(t)$*) то справедлива формула*

$$\frac{d}{d\,t}\sum_{\ell=1}^{\infty}\mathbf{A}_\ell(t) = \sum_{\ell=1}^{\infty}\frac{d}{d\,t}\mathbf{A}_\ell(t) \ .$$

4.4 Экспоненциал матрицы

Определение. **Экспоненциалом** матрицы $\mathbf{Z}$ называется матричный ряд

$$\exp(\mathbf{Z}) = e^{\mathbf{Z}} \stackrel{\text{def}}{=} \sum_{j=0}^{\infty}\frac{\mathbf{Z}^j}{j!} \ . \tag{4.9}$$

На основании следствия 1 к теореме 4.4 этот ряд сходится и притом абсолютно для любой матрицы $\mathbf{Z}$.

Теорема 4.10. *Если матрицы* $\mathbf{A}$ *и* $\mathbf{B}$ *коммутируют, то коммутируют и* $e^{\mathbf{A}}$ *и* $e^{\mathbf{B}}$*:*

$$e^{\mathbf{A}}e^{\mathbf{B}} = e^{\mathbf{B}}e^{\mathbf{A}} = e^{\mathbf{A}+\mathbf{B}} \quad . \tag{4.10}$$

Доказательство . Если $\mathbf{AB} = \mathbf{BA}$, то

$$\mathbf{B}e^{\mathbf{A}} = \mathbf{B}\left(\sum_{j=0}^{\infty}\frac{\mathbf{A}^j}{j!}\right) = \sum_{j=0}^{\infty}\frac{\mathbf{BA}^j}{j!} = \sum_{j=0}^{\infty}\frac{\mathbf{A}^j\mathbf{B}}{j!} = \left(\sum_{j=0}^{\infty}\frac{\mathbf{A}^j}{j!}\right)\mathbf{B} = e^{\mathbf{A}}\mathbf{B}$$

Аналогично доказывается, что $\mathbf{B}^N e^{\mathbf{A}} = e^{\mathbf{A}}\mathbf{B}^N$, и $p(\mathbf{B})e^{\mathbf{A}} = e^{\mathbf{A}}p(\mathbf{B})$ при любом полиноме $p(x) \in \mathbb{C}[x]$. Выбрав в последнем равенстве

$$p(x) = 1 + x + \frac{x^2}{2!} + \dots + \frac{x^N}{N!}$$

и переходя в нем к пределу при $N \to +\infty$, получаем справедливость первого из равенств (4.10).

Для доказательства второго равенства воспользуемся возможностью почленного перемножения двух абсолютно сходящихся рядов:

$$e^{\mathbf{A}}e^{\mathbf{B}} = \sum_{j=0}^{\infty}\frac{\mathbf{A}^j}{j!}\sum_{k=0}^{\infty}\frac{\mathbf{B}^k}{k!} = E + (\mathbf{A}+\mathbf{B}) + \left(\frac{\mathbf{A}^2}{2} + \mathbf{AB} + \frac{\mathbf{B}^2}{2}\right) +$$

$$+\left(\frac{\mathbf{A}^3}{3!} + \frac{\mathbf{A}^2\mathbf{B}}{2!} + \frac{\mathbf{AB}^2}{2!} + \frac{\mathbf{B}^3}{3!}\right) + \left(\frac{\mathbf{A}^4}{4!} + \frac{\mathbf{A}^3\mathbf{B}}{3!} + \frac{\mathbf{A}^2\mathbf{B}^2}{2!2!} + \frac{\mathbf{AB}^3}{3!} + \frac{\mathbf{B}^4}{4!}\right) + \dots +$$

$$+\left(\frac{\mathbf{A}^N}{N!} + \frac{\mathbf{A}^{N-1}\mathbf{B}}{(N-1)!} + \frac{\mathbf{A}^{N-2}\mathbf{B}^2}{(N-2)!2!} + \dots + \frac{\mathbf{A}^{N-K}\mathbf{B}^K}{(N-K)!K!} + \dots + \frac{\mathbf{AB}^{N-1}}{(N-1)!} + \frac{\mathbf{B}^N}{N!}\right) + \dots =$$

$$= E + (\mathbf{A}+\mathbf{B}) + \frac{1}{2}\left(\mathbf{A}^2 + 2\mathbf{AB} + \mathbf{B}^2\right) + \frac{1}{3!}\left(\mathbf{A}^3 + 3\mathbf{A}^2\mathbf{B} + 3\mathbf{AB}^2 + \mathbf{B}^3\right) +$$

$$+\frac{1}{4!}\left(\mathbf{A}^4 + 4\mathbf{A}^3\mathbf{B} + \frac{4!}{2!2!}\mathbf{A}^2\mathbf{B}^2 + 4\mathbf{AB}^3 + \mathbf{B}^4\right) + \dots +$$

$$+\frac{1}{N!}\left(\mathbf{A}^N + N\mathbf{A}^{N-1}\mathbf{B} + \frac{N!}{2!(N-2)!}\mathbf{A}^{N-2}\mathbf{B}^2 + \dots + \underbrace{\frac{N!}{K!(N-K)!}}_{=C_N^K}\mathbf{A}^{N-K}\mathbf{B}^K + \dots + \mathbf{B}^N\right) + \dots =$$

Ввиду коммутативности матриц $\mathbf{A}$ и $\mathbf{B}$, каждую из скобок можно свернуть по правилу бинома Ньютона:

$$= E + (\mathbf{A}+\mathbf{B}) + \frac{(\mathbf{A}+\mathbf{B})^2}{2} + \frac{(\mathbf{A}+\mathbf{B})^3}{3!} + \frac{(\mathbf{A}+\mathbf{B})^4}{4!} + \dots + \frac{(\mathbf{A}+\mathbf{B})^N}{N!} + \dots = e^{\mathbf{A}+\mathbf{B}}$$

□

Следствие 1. $\left(e^{\mathbf{A}}\right)^{-1} = e^{-\mathbf{A}}$.

Особое прикладное значение имеет матрица

$$e^{t\mathbf{A}} = \sum_{j=0}^{\infty} \frac{t^j \mathbf{A}^j}{j!} \ . \tag{4.11}$$

зависящая от параметра $t \in \mathbb{C}$. Если продифференцировать ряд (4.11) по этому параметру, то на основании теоремы 4.9 получим:

$$\frac{d\, e^{t\mathbf{A}}}{d\, t} = \sum_{j=1}^{\infty} \frac{j t^{j-1} \mathbf{A}^j}{j!} = \mathbf{A} \sum_{j=1}^{\infty} \frac{t^{j-1}\mathbf{A}^{j-1}}{(j-1)!} = \mathbf{A} e^{t\mathbf{A}}$$

Таким образом, $e^{t\mathbf{A}}$ является решением матричного дифференциального уравнения

$$\frac{d\, X}{d\, t} = \mathbf{A} X \ . \tag{4.12}$$

Обычно это уравнение рассматривают не относительно квадратной матрицы X, а относительно вектора-столбца $X_{n\times 1}$, т.е. формула (4.12) представляет матричную запись системы обыкновенных дифференциальных уравнений:

$$\left\{ \begin{array}{lclllll} d\, x_1/d\, t & = & a_{11}x_1 + & a_{12}x_2 + & \dots + & a_{1n}x_n & , \\ d\, x_2/d\, t & = & a_{21}x_1 + & a_{22}x_2 + & \dots + & a_{2n}x_n & , \\ \quad \dots & & & & \dots & & \\ d\, x_n/d\, t & = & a_{n1}x_1 + & a_{n2}x_2 + & \dots + & a_{nn}x_n & . \end{array} \right.$$

При задании **начальных условий**

$$x_1(t_0) = x_{10}, x_2(t_0) = x_{20}, \dots, x_n(t_0) = x_{n0} \quad \Longleftrightarrow \quad X(t_0) = X_0 \tag{4.13}$$

решение уравнения (4.12) представляет естественное обобщение одномерного (скалярного) случая:

$$X(t) = e^{(t-t_0)\mathbf{A}} X_0 \ . \tag{4.14}$$

В §4.2 были указаны два способа вычисления аналитической функции матрицы. Применим их для нахождения $e^{t\mathbf{A}}$.

$$e^{t\mathbf{A}} = C e^{t\mathbf{A}_{\mathfrak{J}}} C^{-1} \tag{4.15}$$

и вычисление $e^{t\,x}$ от клетки Жордана производится по формуле (4.7):

$$\exp\left(t \cdot \left[\begin{array}{cccccc} \lambda & & & & & \\ 1 & \lambda & & \mathbb{O} & & \\ 0 & 1 & \lambda & & & \\ \vdots & & \ddots & \ddots & & \\ 0 & 0 & \dots & & 1 & \lambda \end{array} \right]_{k\times k} \right) = \left[\begin{array}{cccc} e^{\lambda t} & & & \\ te^{\lambda t} & e^{\lambda t} & & \mathbb{O} \\ \frac{t^2}{2!}e^{\lambda t} & te^{\lambda t} & e^{\lambda t} & \\ \vdots & & \ddots & \ddots \\ \frac{t^{k-1}}{(k-1)!}e^{\lambda t} & \frac{t^{k-2}}{(k-2)!}e^{\lambda t} & \dots & te^{\lambda t} \ e^{\lambda t} \end{array} \right] =$$

$$= e^{\lambda t}\left(E + tH_1 + \frac{t^2}{2!}H_2 + \dots + \frac{t^{k-1}}{(k-1)!}H_{k-1} \right) \ . \tag{4.16}$$

Пример 4.3. *Для системы дифференциальных уравнений*

$$\begin{cases} d\,x_1/d\,t &= & & x_2 & +x_3, \\ d\,x_2/d\,t &= & x_1 & +x_2 & -x_3, \\ d\,x_3/d\,t &= & & x_2 & +x_3 \end{cases}$$

а) *найти решение при* $x_1(0) = 1, x_2(0) = 1, x_3(0) = 1$*;*

б) *найти все начальные условия* $x_1(0) = x_{10}, x_2(0) = x_{20}, x_3(0) = x_{30}$, *для которых решения будут ограниченными при* $t \to +\infty$.

Решение. Здесь $t_0 = 0$, $\det(\mathbf{A} - \lambda E) = -\lambda(\lambda - 1)^2$.

$$\mathbf{A} = \underbrace{\begin{pmatrix} 2 & -1 & 1 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix}}_{C} \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}}_{\mathbf{A}_{\mathfrak{J}}} \underbrace{\begin{pmatrix} 1 & 0 & -1 \\ 1 & 1 & -1 \\ 0 & 1 & 1 \end{pmatrix}}_{C^{-1}}$$

$$e^{t\mathbf{A}} = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^t & 0 \\ 0 & te^t & e^t \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 \\ 1 & 1 & -1 \\ 0 & 1 & 1 \end{pmatrix} =$$

$$\begin{pmatrix} 2 - e^t + te^t & t\,e^t & -2 + 2\,e^t - t\,e^t \\ -1 + e^t & e^t & 1 - e^t \\ 1 - e^t + t\,e^t & t\,e^t & -1 + 2\,e^t - t\,e^t \end{pmatrix} . \tag{4.17}$$

Домножая эту матрицу справа на столбец $[1, 1, 1]^{\top}$, получаем

$$x_1(t) = e^t(1 + t),\ x_2(t) = e^t,\ x_3(t) = e^t(1 + t) \ .$$

Для решения **б)** домножим матрицу (4.17) справа на столбец $[x_{10}, x_{20}, x_{30}]^{\top}$. Получившийся столбец можно представить в виде линейной комбинации трех:

$$\begin{pmatrix} 2x_{10} - 2x_{30} \\ -x_{10} + x_{30} \\ x_{10} - x_{30} \end{pmatrix} + e^t \begin{pmatrix} -x_{10} + 2\,x_{30} \\ x_{10} + x_{20} - x_{30} \\ -x_{10} + 2\,x_{30} \end{pmatrix} + t\,e^t \begin{pmatrix} x_{10} + x_{20} - x_{30} \\ 0 \\ x_{10} + x_{20} - x_{30} \end{pmatrix}$$

Решение может быть ограниченным лишь при условии, когда столбцы при e^t и $t\,e^t$ будут нулевыми. Решаем получившуюся систему линейных однородных уравнений относительно x_{10}, x_{20} и x_{30}, получаем, что при любом x_{30} решение, проходящее через точку $[2\,x_{30}, -x_{30}, x_{30}]^{\top}$ будет ограниченным; более того, оно будет **стационарным**, т.е. не покинет этой точки при изменении t.

Теорема 4.11. $\det e^{t\mathbf{A}} = e^{t(a_{11}+a_{22}+\cdots+a_{nn})} = e^{t\mathrm{Sp}\,\mathbf{A}}$.

Доказательство следует из (4.15) и (4.16):

$$\det e^{t\mathbf{A}} = \det C \det e^{t\mathbf{A}_{\mathfrak{J}}} \det C^{-1} = \det e^{t\mathbf{A}_{\mathfrak{J}}} = \prod_{k=1}^{n} e^{t\,\lambda_k} = e^{t\,(\lambda_1+\cdots+\lambda_n)} = e^{t\mathrm{Sp}\,\mathbf{A}}$$

Здесь последнее равенство следует из формулы (6.1) главы 3. □

Определение. Если при любых начальных условиях (4.13) решение уравнения (4.12) ограничено при $t \to +\infty$, то будем называть это уравнение **устойчивым по Ляпунову**. Устойчивое уравнение называется **асимптотически устойчивым** если все решения стремятся к $\mathbb{O}_{n\times 1}$ при $t \to +\infty$:

$$\|X(t)\| \underset{t\to+\infty}{\to} 0 \ .$$

Уравнение называется **неустойчивым**, если существует хотя бы один набор начальных условий, для которого соответствующее решение $X(t)$ неограничено при $t \to +\infty$.

Представление решения по формулам (4.15) и (4.16) позволяет свести исследование уравнения (4.12) на устойчивость к анализу знаков собственных чисел $\lambda_1, \dots, \lambda_n$ матрицы $\mathbf{A}$.

Теорема 4.12. *Уравнение* (4.12) *будет*

а) *устойчиво тогда и только тогда, когда* $\Re e(\lambda_j) \le 0$ *для любого* j, *и клетки Жордана в* **ж.н.ф.** $\mathbf{A}_{\mathfrak{Z}}$, *соответствующие собственным числам с* $\Re e(\lambda_j) = 0$ *имеют размерность* 1*;*

б) *асимптотически устойчиво тогда и только тогда, когда* $\Re e(\lambda_j) < 0$ *для любого* j.

Конструктивная проверка условия $\Re e(\lambda_1) < 0, \dots, \Re e(\lambda_n) < 0$ возможна чисто алгебраическими методами. Действительно, это условие означает устойчивость хар. полинома $f(\lambda) = \det(\mathbf{A} - \lambda E)$ и может быть проверено с помощью критериев Рауса или Льенара–Шипара[35] за конечное число операций над коэффициентами.

4.5 Другие специальные функции матрицы

Решение дифференциального уравнения второго порядка[36]

$$\frac{d^2x}{dt^2} = a\,x \ , \ \{x, a, t\} \subset \mathbb{R}, \ x(t_0) = x_0, \left. \frac{d\,x(t)}{d\,t} \right|_{t=t_0} = x_0'$$

представимо в виде

$$x(t) = \left\{ \begin{array}{lll} \dfrac{\sqrt{a}x_0 + x_0'}{2\sqrt{a}} e^{\sqrt{a}(t-t_0)} + \dfrac{x_0 - x_0'}{2\sqrt{a}} e^{-\sqrt{a}(t-t_0)} & \text{при} & a > 0, \\ x_0 \cos\left(\sqrt{|a|}(t-t_0)\right) + \dfrac{x_0'}{\sqrt{|a|}} \sin\left(\sqrt{|a|}(t-t_0)\right) & \text{при} & a < 0, \\ x_0 + x_0'(t-t_0) & \text{при} & a = 0 \ . \end{array} \right.$$

По аналогии поставим задачу поиска решения для системы уравнений

$$\frac{d^2X}{d\,t^2} = \mathbf{A}X \tag{4.18}$$

[35] Раздел I, глава 8, §3.

[36] описывающего, например, закон Гука

при вещественной $n\times n$ матрице $\mathbf{A}$. Формулы, позволяющие найти решение остаются практически прежними, если нам удастся найти матричные аналоги для используемых операций.

$$\begin{aligned}
\sin\mathbf{Z} &\stackrel{\text{def}}{=} \sum_{j=1}^{\infty}(-1)^{j-1}\frac{\mathbf{Z}^{2j-1}}{(2j-1)!}=\mathbf{Z}-\frac{\mathbf{Z}^3}{3!}+\frac{\mathbf{Z}^5}{5!}-\dots\\
\cos\mathbf{Z} &\stackrel{\text{def}}{=} \sum_{j=0}^{\infty}(-1)^{j}\frac{\mathbf{Z}^{2j}}{(2j)!}=E-\frac{\mathbf{Z}^2}{2!}+\frac{\mathbf{Z}^4}{4!}-\dots
\end{aligned}$$

На основании следствия 1 к теореме 4.4 оба ряда сходятся при любой матрице Z.

Сложности возникают при определении $\sqrt{\mathbf{Z}}$. Формально, корнем квадратным из матрицы $\mathbf{A}$ называется решение матричного уравнения

$$X^2=\mathbf{A}\ . \tag{4.19}$$

Уже для матриц второго порядка это уравнение не всегда разрешимо: например, не существует решения при

$$\mathbf{A}=\left(\begin{array}{cc}0&0\\1&0\end{array}\right)\ .$$

С другой стороны, при

$$\mathbf{A}=\left(\begin{array}{rr}-1&0\\0&-1\end{array}\right)$$

существует бесконечное множество решений

$$X=\left(\begin{array}{cc}\alpha&-(\alpha^2+1)/\beta\\ \beta&-\alpha\end{array}\right)\ ,$$

задаваемое двумя параметрами α и $\beta\neq 0$.

Теорема 4.13. *Решение уравнения* (4.19)

а) *всегда существует при* $\det\mathbf{A}\neq 0$*;*

б) *всегда существует среди вещественных симметричных матриц, если* $\mathbf{A}$ *— симметричная и положительно определенная.*

Доказательство проведем только для **б)**. Симметричная матрица $\mathbf{A}$ приводится к диагональному виду с помощью ортогональной матрицы[37]:

$$P^{\top}\mathbf{A}P=\left(\begin{array}{cccc}\lambda_1&&&\mathbb{O}\\&\lambda_2&&\\&&\ddots&\\ \mathbb{O}&&&\lambda_n\end{array}\right)\quad \text{здесь}\quad P^{\top}=P^{-1}\ .$$

[37] Глава 3, §7.2.

Легко видеть, что матрица

$$X \overset{\mathrm{def}}{=} P \begin{pmatrix} \sqrt{\lambda_1} & & & \mathbb{O} \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ \mathbb{O} & & & \sqrt{\lambda_n} \end{pmatrix} P^{\top} . \tag{4.20}$$

является решением уравнения (4.19). Здесь под $\sqrt{}$ понимается произвольное значение квадратного корня. Это значение будет вещественным при $\lambda_j \geq 0$. Если матрица $\mathbf{A}$ положительно определена, то все λ_j положительны[38]. Таким образом формула (4.20) определяет вещественное решение; матрица X будет симметричной и может быть выбрана положительно определенной. $\square$

Пример 4.4. *Вычислить*

$$\sqrt{\begin{pmatrix} 24 & 6 & -12 \\ 6 & 33 & 6 \\ -12 & 6 & 24 \end{pmatrix}} .$$

Решение. $\det(\mathbf{A} - \lambda E) = -(\lambda - 9)(\lambda - 36)^2$. Поскольку матрица $\mathbf{A}$ симметричная, то привести ее к диагональному виду можно с помощью ортогональной матрицы C. Мы, однако же, используем общий алгоритм диагонализации из §5 главы 3 чтобы получить возможно большее число ответов:

$$\mathbf{A} = \underbrace{\begin{pmatrix} 1 & 0 & 2 \\ 2 & 2 & -1 \\ 0 & 1 & 2 \end{pmatrix}}_{C} \underbrace{\begin{pmatrix} 36 & 0 & 0 \\ 0 & 36 & 0 \\ 0 & 0 & 9 \end{pmatrix}}_{\mathbf{A}_{\mathfrak{Z}}} \underbrace{\begin{pmatrix} 5/9 & 2/9 & -4/9 \\ -4/9 & 2/9 & 5/9 \\ 2/9 & -1/9 & 2/9 \end{pmatrix}}_{C^{-1}}$$

$$\sqrt{\mathbf{A}} = \underbrace{\begin{pmatrix} 1 & 0 & 2 \\ 2 & 2 & -1 \\ 0 & 1 & 2 \end{pmatrix}}_{C} \underbrace{\begin{pmatrix} \pm 6 & 0 & 0 \\ 0 & \pm 6 & 0 \\ 0 & 0 & \pm 3 \end{pmatrix}}_{\sqrt{\mathbf{A}_{\mathfrak{Z}}}} \underbrace{\begin{pmatrix} 5/9 & 2/9 & -4/9 \\ -4/9 & 2/9 & 5/9 \\ 2/9 & -1/9 & 2/9 \end{pmatrix}}_{C^{-1}}$$

Ответ. $\sqrt{\mathbf{A}} =$

$$\underbrace{\pm\frac{1}{3}\begin{pmatrix} 14 & 2 & -4 \\ 2 & 17 & 2 \\ -4 & 2 & 14 \end{pmatrix}}_{+++\ \sqrt{\mathbf{A}_{\mathfrak{Z}}}} ; \underbrace{\pm\frac{1}{3}\begin{pmatrix} 14 & 2 & -4 \\ 34 & 1 & -38 \\ 12 & -6 & -6 \end{pmatrix}}_{+-+\ \sqrt{\mathbf{A}_{\mathfrak{Z}}}} ; \underbrace{\pm\begin{pmatrix} -2 & -2 & 4 \\ -2 & -5 & -2 \\ 4 & -2 & -2 \end{pmatrix}}_{--+\ \sqrt{\mathbf{A}_{\mathfrak{Z}}}} ;$$

$$\underbrace{\pm\frac{1}{3}\begin{pmatrix} 6 & 6 & -12 \\ 38 & -1 & -34 \\ 4 & -2 & -14 \end{pmatrix}}_{+--\ \sqrt{\mathbf{A}_{\mathfrak{Z}}}} .$$

[38] Глава 3, теорема 7.5.

Глава 5. Элементы численных методов

Вычисление коэффициентов хар. полинома матрицы A непосредственным разложением определителя $\det(A - \lambda E)$ на $n!$ слагаемых — крайне неэффективно. Элементами этого разложения являются выражения, полиномиально зависящие от параметра λ. На каждом этапе вычислений мы получаем проблему **символьных** вычислений: хранения таких полиномов и действий над ними.

Источником этой трудности является неудобное расположение параметра λ — на диагонали матрицы. Поэтому крайне желательно предварительно преобразовать определитель $\det(A - \lambda E) = 0$ к виду, когда параметр оказывается "выметенным" с диагонали на крайний ряд (в столбец или в строку). Тогда разложение по этому ряду сводит задачу к нахождению **числового** определителя.

1 Метод Леверье

Метод основан на формуле (6.5) главы 3:

$$\mathrm{Sp}(A^k) = \lambda_1^k + \dots + \lambda_n^k = s_k \ ,$$

т.е. след k-й степени матрицы A равен k-й сумме Ньютона ее хар. полинома $f(\lambda) = \det(A - \lambda E)$. Вычисляем последовательные степени матрицы A:

$$s_1 = \mathrm{Sp}(A),\ s_2 = \mathrm{Sp}(A^2),\ \dots, s_n = \mathrm{Sp}(A^n) \ .$$

Неизвестные коэффициенты $f(\lambda) = (-1)^n(\lambda^n + a_1\lambda^{n-1} + \dots + a_n)$ находим по рекуррентным формулам Ньютона:

$$\begin{array}{l}
s_1 = -a_1 \qquad\qquad\quad \Longrightarrow a_1 = -s_1 \\
s_2 = -(a_1 s_1 + 2a_2) \Longrightarrow a_2 = -(s_2 + a_1 s_1)/2 \\
\dots \\
s_k = -(a_1 s_{k-1} + a_2 s_{k-2} + \dots + a_{k-1}s_1 + k a_k) \Longrightarrow \\
\qquad\qquad \Longrightarrow a_k = -(s_k + a_1 s_{k-1} + a_2 s_{k-2} + \dots + a_{k-1}s_1)/k \ (k \le n)
\end{array}$$

Очевидно, что при вычислении матрицы A^n достаточно обойтись лишь элементами на главной диагонали.

Пример 1.1. *Найти хар. полином матрицы*[39]

$$A = \left(\begin{array}{rrrr}
-5.509882 & 1.870086 & 0.422908 & 0.008814 \\
0.287865 & -11.811654 & 5.711900 & 0.058717 \\
0.049099 & 4.308033 & -12.970687 & 0.229326 \\
0.006235 & 0.269851 & 1.397369 & -17.596207
\end{array}\right) \qquad (1.1)$$

[39] Взята из работы Леверье

Решение.

$$A^2 = \begin{pmatrix} 30.91795128 & -30.56848188 & 2.878480155 & 0.0031325713 \\ -4.705449283 & 164.6764010 & -141.3504639 & -0.4143169528 \\ 0.3341843103 & -106.6094396 & 193.1869924 & -6.756396001 \\ 0.0022236138 & -1.904168948 & -41.16923134 & 309.9628536 \end{pmatrix},$$

$$A^3 = \begin{pmatrix} -179.0125092 & 431.2849919 & -198.8601505 & -0.9173897610 \\ 66.38829278 & -2562.954533 & 2771.458834 & -15.49709921 \\ -23.08728044 & 2090.291485 & -3124.010318 & 156.9329019 \\ -0.649145142 & -71.21907809 & 956.2502143 & -5463.723497 \end{pmatrix},$$

$$A^4 = \begin{pmatrix} 1100.720103 & * & * & * \\ * & 42332.23816 & * & * \\ * & * & 52669.62534 & * \\ * & * & * & 96355.91518 \end{pmatrix}.$$

Вычисляем Sp:

$$s_1 = -47.888430,\ s_2 = 698.7441983,\ s_3 = -11329.70086,\ s_4 = 192458.4988\ ,$$

и по формулам Ньютона получаем:

$$a_1 = 47.888430,\ a_2 = 797.278764_8,\ a_3 = 5349.45551_3,\ a_4 = 12296.550_{68}\ .$$

$\triangle$

После нахождения коэффициентов хар. полинома можно найти его корни каким-либо приближенным методом. Если λ_k — одно из собственных чисел, то для нахождения соответствующего собственного вектора воспользуемся теоремой 6.5 главы 3, дополнительно предположив, что все собственные числа различны. Обозначим

$$f_k(\lambda) \stackrel{\text{def}}{=} f(\lambda)/(\lambda - \lambda_k) = (-1)^n(\lambda^{n-1} + p_1\lambda^{n-2} + \cdots + p_{n-1}) \tag{1.2}$$

т.е. частное от деления $f(\lambda)$ на $\lambda - \lambda_k$. Тогда любой столбец матрицы $f_k(A)$ будет собственным вектором, принадлежащим λ_k. В частности, собственным вектором будет комбинация

$$\left[\begin{array}{l}\text{первый}\\ \text{столбец } A^{n-1}\end{array}\right] + p_1 \left[\begin{array}{l}\text{первый}\\ \text{столбец } A^{n-2}\end{array}\right] + \cdots + p_{n-1}\left[\begin{array}{l}\text{первый}\\ \text{столбец } E\end{array}\right].$$

Степени матрицы A уже посчитаны.

Для примера 1.1 находим собственные числа:

$$\lambda_1 = -17.86326,\ \lambda_2 = -17.15242,\ \lambda_3 = -7.57404,\ \lambda_4 = -5.29869\ . \tag{1.3}$$

Коэффициенты $f_1(\lambda)$ можно определить по схеме Хорнера:

	1	47.888430	797.2787648	5349.455513	12296.55068
-17.86326	1	$\underbrace{30.025170}_{p_1}$	$\underbrace{260.9313465}_{p_2}$	$\underbrace{688.371028}_{p_3}$	≈ 0

Собственным вектором, соответствующим λ_1, будет

$$[-0.0256_{67},\ 0.21938_0,\ -0.24187_1,\ 1.044526]^\top\ .$$

$\triangle$

Упражнение 1.1. *Получить следующее явное представление хар.полинома:*

$$f(\lambda)=\frac{1}{n!}\left|\begin{array}{llllll} s_1 & 1 & & & & \\ s_2 & s_1 & 2 & & \mathbb{O} & \\ s_3 & s_2 & s_1 & 3 & & \\ \dots & & & & \ddots & \\ s_n & s_{n-1} & \dots & & s_1 & n \\ \lambda^n & \lambda^{n-1} & \dots & & \lambda & 1 \end{array}\right|_{(n+1)\times(n+1)} .$$

Указание. Воспользоваться идеей доказательства теоремы 2.1 (см. следующий параграф).

Историческая справка

Леверье Урбен Жан Жозеф (Le Verrier Urbain Jean Joseph, 1811–1877)

В науке и технике бывали иногда открытия и изобретения, совершенные почти в одно время независимо друг от друга разными лицами. Примером этого может быть открытие на основании вычислений планеты Нептун Адамсом и Леверье.

В 1846 г. Леверье был еще молодым доцентом по математике и астрономии в Ecole Polytechnique. Адамс только что окончил курс Кембриджского университета и был оставлен при нем для приготовления к профессорскому званию. Оба они стали впоследствии знаменитыми астрономами. Леверье был характера твердого, энергичного; став директором Парижской обсерватории он беспрестанно увольнял сотрудников; они шли жаловаться военному министру маршалу Вальяну, как главному начальнику Геодезического управления, который говорил:

Обсерватория невозможна без Леверье, а Леверье еще более невозможен в обсерватории.

Виктор Гюго писал про него в своем романе “Наполеон малый” (памфлет на Наполеона III):

У него есть свой Лаплас, который отвечает на имя Леверье, но который не создал небесной механики.

Но здесь поэт жестоко ошибся. Леверье в течение больше чем 25 лет переработал целиком всю небесную механику Лапласа, включив в нее и планету Нептун, им открытую.

Адамс был характера скромного, можно сказать даже робкого. Пятьдесят лет он работал в Кембридже, преподавая математику и астрономию. Как вычислитель, он был также изумителен.

Адамс и Леверье искали сперва элементы орбиты неизвестной планеты, затем по этим элементам ее место, причем Адамс принял круговую орбиту, но важно было найти не элемент орбиты, а самую планету, и каждый ее нашел. У Адамса она оказалась на полтора градуса от места, им указанного. У Леверье — 52′. В Париже не было достаточно сильного инструмента, чтобы заметить диск планеты, поэтому Леверье написал в Берлин астроному Галле, который в первый же вечер и открыл планету, отметив ее ясно видимый диск.

Адамс, получив элементы орбиты, отнес свой расчет и теорию его королевскому астроному Эри. Эри был занят, и Адамса не принял. Через неделю он опять зашел к Эри, Эри опять был занят и его не принял. Тогда он оставил свой мемуар у Эри и больше к нему не приходил. Это было в сентябре 1845 г.

В Кембридже была университетская обсерватория, но со слабым инструментом. Директор этой обсерватории Чаллиз по просьбе Адамса обследовал указанную ему область неба, несколько раз наблюдал искомую планету, но счел ее за неподвижную звезду. После того как планета была открыта Галле, произошло невероятное смятение. Эри опубликовал мемуар Адамса, как приложение к "Nautical Almanach", но упущенного не воротишь, и слава открытия осталась за Леверье.

(А.Н.Крылов "Попов и Маркони" , в кн. "Воспоминания и очерки" М.АН СССР, 1956, с.505-506)

2 Метод Крылова

2.1 Идея

Идея А.Н.Крылова (1931 г.) заключалась в предварительном преобразовании характеристического уравнения

$$f(\lambda) = \det(A - \lambda E) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \dots & & & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{vmatrix} = 0 \qquad (2.1)$$

в эквивалентное ему

$$D(\lambda) = \begin{vmatrix} b_{11} - \lambda & b_{12} & \dots & b_{1n} \\ b_{21} - \lambda^2 & b_{22} & \dots & b_{2n} \\ \dots & & & \dots \\ b_{n1} - \lambda^n & b_{n2} & \dots & b_{nn} \end{vmatrix} = 0\ . \qquad (2.2)$$

Последнее можно разложить по первому столбцу, при этом миноры $(n-1)$-го порядка будут **числовыми**.

К этой идее А.Н.Крылов пришел, решая следующую задачу из теории линейных дифференциальных уравнений.

Пример 2.1. *Исключить переменные x_2 и x_3 из системы*

$$\left\{ \begin{array}{rcl} {}^{d\,x_1}\!/_{d\,t} & = & a_{11}x_1 + a_{12}x_2 + a_{13}x_3, \\ {}^{d\,x_2}\!/_{d\,t} & = & a_{21}x_1 + a_{22}x_2 + a_{23}x_3, \\ {}^{d\,x_3}\!/_{d\,t} & = & a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \end{array} \right. \qquad (2.3)$$

т.е. выписать дифференциальное уравнение относительно только переменной x_1.

Решение. Дифференцируем первое уравнение системы (2.3) по t:

$$\frac{d^2\,x_1}{d\,t^2} = a_{11}\frac{d\,x_1}{d\,t} + a_{12}\frac{d\,x_2}{d\,t} + a_{13}\frac{d\,x_3}{d\,t} =$$

и подставим в правую часть представления для производных из той же системы:

$$\begin{aligned} &= a_{11}\left(a_{11}x_1 + a_{12}x_2 + a_{13}x_3\right) + a_{12}\left(a_{21}x_1 + a_{22}x_2 + a_{23}x_3\right) + a_{13}\left(a_{31}x_1 + a_{32}x_2 + a_{33}x_3\right) \\ &= \underbrace{\left(a_{11}^2 + a_{12}a_{21} + a_{13}a_{31}\right)}_{b_{21}} x_1 + \underbrace{\left(a_{11}a_{12} + a_{12}a_{22} + a_{13}a_{32}\right)}_{b_{22}} x_2 + \underbrace{\left(a_{11}a_{13} + a_{12}a_{23} + a_{13}a_{33}\right)}_{b_{23}} x_3 \end{aligned}$$

Дифференцируем еще раз:

$$\frac{d^3\,x_1}{d\,t^3} = b_{21}\frac{d\,x_1}{d\,t} + b_{22}\frac{d\,x_2}{d\,t} + b_{23}\frac{d\,x_3}{d\,t} =$$

и снова подставим в правую часть представления для производных из системы (2.3):

$$= \underbrace{\left(a_{11}b_{21} + a_{21}b_{22} + a_{31}b_{23}\right)}_{b_{31}} x_1 + \underbrace{\left(a_{12}b_{21} + a_{22}b_{22} + a_{32}b_{23}\right)}_{b_{32}} x_2 + \underbrace{\left(a_{13}b_{21} + a_{23}b_{22} + a_{33}b_{23}\right)}_{b_{33}} x_3 \quad .$$

Объединим полученные уравнения в систему, дополнив ее тождеством $x_1 = x_1$:

$$\left\{ \begin{array}{lllll} -x_1 & +x_1 & & & = 0 \\ -x_1' & +a_{11}x_1 & +a_{12}x_2 & +a_{13}x_3 & = 0 \\ -x_1'' & +b_{21}x_1 & +b_{22}x_2 & +b_{23}x_3 & = 0 \\ -x_1''' & +b_{31}x_1 & +b_{32}x_2 & +b_{33}x_3 & = 0 \end{array} \right. \qquad (2.4)$$

Система (2.4) — однородная относительно столбца переменных $[1, x_1, x_2, x_3]^\top$. Ее определитель должен обращаться в нуль, поскольку она имеет нетривиальное решение (первая компонента этого столбца ненулевая):

$$\left| \begin{array}{cccc} -x_1 & 1 & 0 & 0 \\ -x_1' & a_{11} & a_{12} & a_{13} \\ -x_1'' & b_{21} & b_{22} & b_{23} \\ -x_1''' & b_{31} & b_{32} & b_{33} \end{array} \right| = 0 \quad . \qquad (2.5)$$

Раскладывая определитель по первому столбцу получаем требуемое дифференциальное уравнение относительно x_1:

$$\left| \begin{array}{cc} a_{12} & a_{13} \\ b_{22} & b_{23} \end{array} \right| \cdot x_1''' - \left| \begin{array}{cc} a_{12} & a_{13} \\ b_{32} & b_{33} \end{array} \right| \cdot x_1'' + \left| \begin{array}{cc} b_{22} & b_{23} \\ b_{32} & b_{33} \end{array} \right| \cdot x_1' - \left| \begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{array} \right| x_1 = 0 \quad .$$

□

Теперь вспомним, что система (2.3) всегда имеет частное решение $x_1 = Ce^{\lambda t}, x_2 \equiv 0, x_3 \equiv 0$, где λ — произвольное собственное число матрицы

A, а $C = const \neq 0$. Но тогда подстановка его в (2.5) должна сохранить равенство

$$\begin{vmatrix} -1 & 1 & 0 & 0 \\ -\lambda & a_{11} & a_{12} & a_{13} \\ -\lambda^2 & b_{21} & b_{22} & b_{23} \\ -\lambda^3 & b_{31} & b_{32} & b_{33} \end{vmatrix} = \begin{vmatrix} -1 & 0 & 0 & 0 \\ -\lambda & a_{11}-\lambda & a_{12} & a_{13} \\ -\lambda^2 & b_{21}-\lambda^2 & b_{22} & b_{23} \\ -\lambda^3 & b_{31}-\lambda^3 & b_{32} & b_{33} \end{vmatrix} = - \begin{vmatrix} a_{11}-\lambda & a_{12} & a_{13} \\ b_{21}-\lambda^2 & b_{22} & b_{23} \\ b_{31}-\lambda^3 & b_{32} & b_{33} \end{vmatrix} = 0 \tag{2.6}$$

И **любое** собственное число должно удовлетворять уравнению (2.6). Поскольку его степень — не выше порядка матрицы A, то можно утверждать, что по крайней мере при различности всех собственных чисел, мы получили то же самое характеристическое уравнение. Но форма, которую оно приобрело — именно та, что нам нужна: характеристическое (вековое) уравнение (2.1) преобразовано к виду (2.2).

Проясним теперь алгебраический смысл произведенных преобразований. Обратим внимание на закон формирования коэффициентов:

$$\begin{aligned} [b_{21}, b_{22}, b_{23}] &= [a_{11}, a_{12}, a_{13}] \cdot A \\ [b_{31}, b_{32}, b_{33}] &= [b_{21}, b_{22}, b_{23}] \cdot A = [a_{11}, a_{12}, a_{13}] \cdot A^2 \end{aligned}$$

и если ввести в рассмотрение строку $X_0 \stackrel{\text{def}}{=} [1, 0, 0]$, то увидим, что коэффициенты определителя в уравнении (2.5) получаются по формулам:

$$X_0 \cdot A, \quad X_0 \cdot A^2, \quad X_0 \cdot A^3 \ .$$

Возникает гипотеза, что выбор X_0 может быть произволен.

2.2 Реализация

Теорема 2.1. *Пусть* $X_0 \stackrel{\text{def}}{=} [x_{01}, x_{02}, \dots, x_{0n}] \in \mathbb{C}^n$ *и*

$$B^{[j]} \stackrel{\text{def}}{=} [b_{j1}, \dots, b_{jn}] \stackrel{\text{def}}{=} X_0 \cdot A^j \quad \text{для } j \in \{1, \dots, n\} \ .$$

Тогда определитель

$$\begin{vmatrix} 1 & x_{01} & \dots & x_{0n} \\ \lambda & b_{11} & \dots & b_{1n} \\ \dots & & & \dots \\ \lambda^n & b_{n1} & \dots & b_{nn} \end{vmatrix}_{(n+1)\times(n+1)}$$

отличается от характеристического полинома (2.1) *лишь числовым множителем.*

Доказательство . Если $f(\lambda) = \det(A - \lambda E) = (-1)^n(\lambda^n + a_1\lambda^{n-1} + \dots + a_n)$ (коэффициенты a_j мы и хотим определить), то по теореме Гамильтона–Кэли:

$$A^n + a_1 A^{n-1} + \dots + a_n E = \mathbb{O}_{n\times n} \Longrightarrow X_0 \cdot A^n + a_1 X_0 \cdot A^{n-1} + \dots + a_n X_0 \cdot E = \mathbb{O}_{1\times n}$$

$$\iff a_nX_0 + a_{n-1}B^{[1]} + \cdots + a_1B^{[n-1]} + B^{[n]} = \mathbb{O}_{1\times n}$$

Последнее равенство представляет линейную систему относительно неизвестных коэффициентов. Можно ее решать по формулам Крамера, но мы пойдем другим путем. Дополним эту систему тождеством

$$a_n + a_{n-1}\lambda + \cdots + a_1\lambda^{n-1} + (\lambda^n + (-1)^{n-1}f(\lambda)) \equiv 0$$

и рассмотрим получившуюся систему как линейную однородную относительно строки неизвестных $[a_n, a_{n-1}, \ldots, a_1, 1] \in \mathbb{C}^{n+1}$:

$$[a_n, a_{n-1}, \ldots, a_1, 1] \left[\begin{array}{ll} X_0 & 1 \\ B^{[1]} & \lambda \\ \vdots & \vdots \\ B^{[n-1]} & \lambda^{n-1} \\ B^{[n]} & \lambda^n + (-1)^{n-1}f(\lambda) \end{array}\right] = \mathbb{O}_{1\times(n+1)}$$

Поскольку она имеет нетривиальное решение (последняя компонента строки неизвестных равна 1), то ее определитель должен обращаться в нуль:

$$0 = \det \left[\begin{array}{ll} X_0 & 1 \\ B^{[1]} & \lambda \\ \vdots & \vdots \\ B^{[n-1]} & \lambda^{n-1} \\ B^{[n]} & \lambda^n + (-1)^{n-1}f(\lambda) \end{array}\right] =$$

$$= \det \left[\begin{array}{ll} X_0 & 0 \\ B^{[1]} & 0 \\ \vdots & \vdots \\ B^{[n-1]} & 0 \\ B^{[n]} & (-1)^{n-1}f(\lambda) \end{array}\right] + \det \left[\begin{array}{ll} X_0 & 1 \\ B^{[1]} & \lambda \\ \vdots & \vdots \\ B^{[n-1]} & \lambda^{n-1} \\ B^{[n]} & \lambda^n \end{array}\right] =$$

$$= (-1)^{n-1}f(\lambda) \det \left[\begin{array}{l} X_0 \\ B^{[1]} \\ \vdots \\ B^{[n-1]} \end{array}\right] + \det \left[\begin{array}{ll} X_0 & 1 \\ B^{[1]} & \lambda \\ \vdots & \vdots \\ B^{[n-1]} & \lambda^{n-1} \\ B^{[n]} & \lambda^n \end{array}\right]$$

Переставив во втором определителе столбцы, получим:

$$f(\lambda) \left| \begin{array}{ccc} x_{01} & \ldots & x_{0n} \\ b_{11} & \ldots & b_{1n} \\ \ldots & & \ldots \\ b_{n-1,1} & \ldots & b_{n-1,n} \end{array} \right|_{n\times n} = \left| \begin{array}{cccc} 1 & x_{01} & \ldots & x_{0n} \\ \lambda & b_{11} & \ldots & b_{1n} \\ \ldots & & & \ldots \\ \lambda^{n-1} & b_{n-1,1} & \ldots & b_{n-1,n} \\ \lambda^n & b_{n1} & \ldots & b_{nn} \end{array} \right|_{(n+1)\times(n+1)}, \qquad (2.7)$$

что и доказывает теорему. □

При выборе $X_0 \stackrel{\text{def}}{=} [1, 0, \ldots, 0]$ в (2.7) мы можем преобразовать правую часть к определителю n-го порядка вида (2.2).

Пример 2.2. *Найти хар.полином для матрицы из примера* (1.1).

Решение. Для $X_0 = [1, 0, 0, 0]$, получаем

$$\begin{array}{lllllr}
B^{[1]} = & X_0 A = & [-5.509882, & 1.870086, & 0.422908, & 0.008814] \\
B^{[2]} = & B^{[1]} A = & [30.91795128, & -30.56848188, & 2.878480155, & 0.0031325713] \\
B^{[3]} = & B^{[2]} A = & [-179.0125092, & 431.2849919, & -198.8601505, & -0.9173897610] \\
B^{[4]} = & B^{[3]} A = & [1100.720101, & -6285.901537, & 4965.821760, & -5.71527813]
\end{array}$$

По формуле (2.2):

$$D(\lambda) = 27.57521777\lambda^4 + 1320.533885\lambda^3 + 21985.13553\lambda^2 + 147512.4007\lambda + 339080.0599 \ ,$$

откуда

$$f(\lambda) = \lambda^4 + 47.8884_{2996}\,\lambda^3 + 797.27876_{35}\,\lambda^2 + 5349.45551_2\,\lambda + 12296.5505_7 \ .$$

$\triangle$

Замечание. Поскольку[40] характеристический полином матрицы не меняется при ее транспонировании, можем в рассуждениях теоремы 2.1 заменить строки на столбцы. Именно, взяв произвольный столбец $X_0 = [x_{10}, \ldots, x_{n0}]^{\top}$ и определив столбцы

$$B_{[j]} \stackrel{\text{def}}{=} (b_{1j}, \ldots, b_{nj})^{\top} \stackrel{\text{def}}{=} A^j \cdot X_0 \quad \text{для } j \in \{1, \ldots, n\} \ ,$$

получим новую версию равенства (2.7):

$$f(\lambda) \det \left[X_0, B_{[1]}, \ldots, B_{[n-1]}\right] = \det \left[\begin{array}{ccccc} X_0 & B_{[1]} & \ldots & B_{[n-1]} & B_{[n]} \\ 1 & \lambda & \ldots & \lambda^{n-1} & \lambda^n \end{array} \right] \ .$$

Этот способ следует считать более удобным если задачей ставится ещ и вычисление собственных векторов. Воспользуемся тем же приемом, что и в методе Леверье. Рассмотрим полином $f_k(x)$, определенный формулой (1.2). Поскольку любой столбец матрицы $f_k(A)$ является собственным вектором, принадлежащим λ_k, то и произвольная комбинация этих столбцов тоже будет собственным вектором. В частности, собственным будет и столбец

$$\begin{array}{rcl}
(-1)^n f_k(A) X_0 & = & A^{n-1} X_0 + p_1 A^{n-2} X_0 + \cdots + p_{n-1} X_0 = \\
 & = & p_{n-1} X_0 + \cdots + p_1 B_{[n-2]} + B_{[n-1]} \ .
\end{array}$$

Комбинируемые вектора уже посчитаны, а коэффициенты p_j можно определить по схеме Хорнера.

Теперь обсудим исключительные случаи. Во-первых, при неудачном выборе X_0 определитель из левой части (2.7) может обратиться в нуль. Тогда

[40] Глава 3, упражнение 5.2.

это равенство становится тождеством $0 = 0$, из которого невозможно получить информацию о хар. полиноме. Эта неприятность может произойти, например, если наш выбор пал на вектор X_0, такой что столбец $X_0^\top$ является собственным для матрицы $A^\top$.

Вероятность такого события ничтожно мала. Действительно, трудно ожидать, чтобы n строк $X_0, X_0 \cdot A, \dots, X_0 \cdot A^{n-1}$ оказались зависимыми, если только сама матрица A не обладает скрытыми дефектами.

Пример 2.3. *Найти уравнение* (2.2) *для матрицы*

$$A = \left(\begin{array}{ccc} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{array} \right) \ .$$

Решение. Проделав те же вычисления, что и в примере 2.1, получаем уравнение (2.6):

$$\left| \begin{array}{ccc} 2-\lambda & 1 & 1 \\ 6-\lambda^2 & 5 & 5 \\ 22-\lambda^3 & 21 & 21 \end{array} \right| = 0 \qquad \Longleftrightarrow \qquad 0 = 0 \ .$$

Тот же результат получится и при использовании формулы (2.7) при любом выборе $X_0 = (x_1, x_2, x_3)$. △

Объяснение этого феномена в том, что для матрицы A ее аннулирующий полином имеет степень меньшую ее порядка:

$$A^m + b_1 A^{m-1} + \dots + b_{m-1} E = \mathbb{O}_{n\times n} \quad \text{при} \quad m < n \ .$$

Понятно, что домножив это уравнение на любой вектор X_0 мы получим линейное соотношение, связывающее вектора $X_0, X_0A, \dots, X_0A^m$ и определитель из левой части (2.7) обратится в нуль. Такая ситуация возможна только тогда, когда хар. полином $f(\lambda)$ имеет кратные корни, и на практике встречается редко.

3 Частичная проблема собственных чисел

В двух предыдущих параграфах нас интересовал весь спектр матрицы A, т.е. полный набор ее собственных чисел. В отличие от этой задачи, в настоящем пункте нас будут интересовать только часть из них. Методы решения тоже меняются — они становятся итерационными.

Задача. Найти максимальное по модулю собственное число матрицы A.

Будем считать, что

$$|\lambda_1| = \max_{j\in\{1,\dots,n\}} |\lambda_j|$$

и такое число единственно: $|\lambda_1| > |\lambda_j|$ при $j \in \{2, \dots, n\}$. Для произвольного столбца $Y_0 = [y_{10}, \dots, y_{n0}]^\top$ составим итерационную векторную последовательность

$$Y_1 = A \cdot Y_0,\ Y_2 = A \cdot Y_1, \dots,\ Y_K = A \cdot Y_{K-1}, \dots \tag{3.1}$$

и выделим последовательность первых компонент этих векторов:

$$y_{11}, y_{12}, \dots, y_{1K}, \dots$$

Теорема 3.1. *Как правило*

$$\lim_{K\to+\infty} \frac{y_{1,K+1}}{y_{1K}} = \lambda_1 \ . \tag{3.2}$$

Доказательство основано на соображениях аналогичных используемым в §2 главы 4. Для простоты предположим также, что все собственные числа различны. Тогда имеет место представление

$$y_{1K} = C_1\lambda_1^K + C_2\lambda_2^K + \dots + C_n\lambda_n^K, \tag{3.3}$$

где $C_1, \dots, C_n$ — константы, не зависящие от K. Действительно, существует базис пространства, состоящий из собственных векторов A: $\mathfrak{X}_1, \dots, \mathfrak{X}_n$, $(A\mathfrak{X}_j = \lambda_j \mathfrak{X}_j)$. Вектор Y_0 можно разложить по этому базису:

$$Y_0 = \alpha_1\mathfrak{X}_1 + \dots + \alpha_n\mathfrak{X}_n \ .$$

Тогда получаем

$$\begin{array}{rcl} Y_1 = AY_0 & = & \alpha_1\lambda_1\mathfrak{X}_1 + \dots + \alpha_n\lambda_n\mathfrak{X}_n, \\ \dots & & \dots \\ Y_K = A^K Y_0 & = & \alpha_1\lambda_1^K\mathfrak{X}_1 + \dots + \alpha_n\lambda_n^K\mathfrak{X}_n \end{array}$$

откуда и следует равенство (3.3).

Поскольку $|\lambda_j/\lambda_1|^K \to 0$ при $K \to +\infty$ для $j \in \{2, \dots, n\}$, то

$$\lim_{K\to+\infty} \frac{y_{1,K+1}}{y_{1K}} = \lim_{K\to+\infty} \frac{\lambda_1^{K+1}\left[C_1 + C_2(\lambda_2/\lambda_1)^{K+1} + \dots + C_n(\lambda_n/\lambda_1)^{K+1}\right]}{\lambda_1^K\left[C_1 + C_2(\lambda_2/\lambda_1)^K + \dots + C_n(\lambda_n/\lambda_1)^K\right]} = \lambda_1$$

если только $C_1 \ne 0$ (именно это и скрывается под словами “как правило” в формулировке теоремы). □

Следующий пример показывает, насколько медленной может быть сходимость, если имеются несколько корней, близких по модулю $\max\limits_j |\lambda_j|$.

Пример 3.1. *Найти максимальное по модулю собственное число матрицы* (1.1).

Решение. Для столбца $Y_0 = [1, 0, 0, 0]^\top$ имеем

$$y_{1,101}/y_{1,100} = -17.8_{3113} \ ,$$

т.е. на 100-й итерации получаем лишь 3 верных знака. При этом компонентами векторов являются числа порядка 10^{123}. Если мы посмотрим на ответ (1.3), то увидим, что имеются два корня, почти равных по модулю. △

Замечание. При соответствующем предположении равенство (3.2) будет выполнено для любой компоненты векторов (3.1)

$$\lim_{K\to+\infty} \frac{y_{j,K+1}}{y_{jK}} = \lambda_1 \quad \text{для } j \in \{1,\ldots,n\} \ . \tag{3.4}$$

Следствие 1. *Если* $\lim\limits_{K\to+\infty} y_{1K} \neq 0$, *то вектор*

$$\left[1, \lim_{K\to+\infty} \frac{y_{2K}}{y_{1K}}, \ldots, \lim_{K\to+\infty} \frac{y_{nK}}{y_{1K}}\right]^\top$$

будет собственным, принадлежащим λ_1.

Доказательство . Действительно, разделим равенство $Y_{K+1} = AY_K$ на y_{1K}

$$\frac{1}{y_{1K}} Y_{K+1} = A\left(\frac{1}{y_{1K}} Y_K\right) \tag{3.5}$$

и перейдем к пределу при $K \to +\infty$. Обозначим пределы справа

$$\tau_{j1} \stackrel{\text{def}}{=} \lim_{K\to+\infty} y_{jK} \big/ y_{1K} \qquad \text{для } j \in \{2,\ldots,n\}$$

(существуют поскольку для y_{jK} имеют место представления вида (3.3)); положим $\tau_{11} \stackrel{\text{def}}{=} 1$. Вычислим теперь пределы в левой части (3.5):

$$\lim_{K\to+\infty} \frac{y_{j,K+1}}{y_{1K}} = \lim_{K\to+\infty} \frac{y_{j,K+1}}{y_{jK}} \frac{y_{jK}}{y_{1K}} = \lambda_1 \tau_{j1}$$

на основании (3.4). Итак, в равенстве (3.5) существуют пределы слева и справа при $K \to +\infty$. Поскольку при переходе к пределу равенство должно сохраняться, имеем справедливость следствия. □

Вычисление Google **PageRank**

Поиск в интернете страниц, соответствующих запросу, выдает ссылки на сотни соответствующих страниц. Их надлежащее упорядочивание представляет важную и нетривиальную задачу. Оригинальная идея Google, предложенная в 1998 г., заключалась в том, чтобы упорядочивать результаты запроса в соответствии с **PageRank** — рангом страницы, отражающим ее популярность. **PageRank** определяется следующим образом. Обозначим общее число страниц интернета через n, пронумеруем их как-нибудь, и определим $n \times n$-матрицу гиперссылок H: если i-я страница имеет ровно $k > 0$ исходящих гиперссылок на страницы $j_1, \ldots, j_k$, то $h_{ij_1} = \cdots = h_{ij_k} = {}^1\!/\!_k$ и $h_{ij} = 0$ при $j \notin \{j_1, \ldots, j_k\}$. (В случае, когда исходящих ссылок с i-й страницы нет, возможны разные подходы. Например, можно положить $h_{ii} = 1$ и $h_{ij} = 0$ при $j \neq i$.) Предполагается, что при случайном блуждании по интернету пользователь попадает с некоторой вероятностью на произвольную страницу.

При таких предположениях **PageRank** определяется как вектор-строка

$$\mathbf{PageRank} = [p_1^\infty, \dots, p_i^\infty, \dots, p_n^\infty] \ , \qquad (3.6)$$

финальных абсолютных вероятностей марковской цепи, пространство состояний которой совпадает со множеством всех страниц интернета, а матрица переходных вероятностей

$$\mathfrak{P} \stackrel{\mathrm{def}}{=} c\,H + \frac{1-c}{n} \left(\begin{array}{cccc} 1 & 1 & \dots & 1 \\ \dots & & & \dots \\ 1 & 1 & \dots & 1 \end{array} \right)_{n\times n} . \qquad (3.7)$$

Здесь $c \in]0,1[$ — вероятность "непрыгания" на случайную страницу — считается равной 0.85. В предположении, что пользователь с вероятностью c следует предлагаемым на очередной странице ссылкам, а с вероятностью $(1-c)$ уходит со страницы на совершенно случайную, можно интерпретировать i-ю компоненту вектора (3.6), т.е. p_i^∞, как вероятность того, что пользователь интернета находится на i-й странице.

Можно показать, что матрица (3.7) — регулярная. Тогда, в соответствии с результатами §3.2 главы 4, столбец $\mathbf{PageRank}^\top$ является собственным вектором матрицы $\mathfrak{P}^\top$, соответствующим собственному числу $\lambda = 1$, максимальному по модулю среди других чисел спектра $\mathfrak{P}$.

В связи с вычислением **PageRank** возникают две проблемы. С точки зрения создателя i-й страницы, имеет смысл грамотно организовать ссылки со своей страницы на другие — с целью увеличения p_i^∞. Для собственно же поисковой системы Google проблема заключается в организации регулярного (ежемесячного) эффективного вычисления **PageRank**.

Последняя задача весьма ресурсоемка поскольку число n громадно. Это обстоятельство полностью блокирует применение традиционных методов поиска собственного вектора, основанных на решении системы линейных уравнений или вычислении характеристического полинома. На помощь приходит следствие 1 к теореме 3.1: вектор $\mathfrak{P}^\top$ как раз является собственным вектором матрицы $\mathfrak{P}$, принадлежащим максимальному по модулю ее собственному числу. Поскольку структура матрицы $\mathfrak{P}$ довольно проста, то реально организовать и вычисление последовательности (3.1), распараллелив, вдобавок, вычисления по компонентам векторов ...

приведем цитату из книги А.Н.Крылова:

В 1869 г. возникла полемика о возрасте Земли между знаменитым физиком XIX столетия сэром Вильямом Томсоном (впоследствии лорд Кельвин) и не менее знаменитым геологом Гексли. Томсон исходил из физических расчетов о повышении температуры с глубиной... Гексли руководствовался законами геологии. Томсон по своим расчетам определил возраст Земли не более как в 50 млн. лет. Гексли не менее чем в 500 млн. Архиепископ иоркский Вильберфорс, усмотрев такую ересь, привел библейское исчисление в 7377 лет. Когда оба противника продолжали свою полемику в самых вежливых и галантных тонах, разъяренный епископ посоветовал им обоим прочесть в Библии из книги притчей Соломоновых:

"Яко пес возвращается на блевотину свою, так и глупец повторяет глупость свою" (2.26)

Не будем следовать ни за гневным архипастырем, ни за знаменитым физиком, но будем помнить слова Гексли:

"Математику можно сравнить с мельницей превосходного устройства, которая перемалывает что угодно до любой тонкости; тем не менее, то, что вы получаете, зависит от того, что вы засыпаете, и как великолепнейшая в мире мельница не доставит вам пшеничной крупчатки из лебеды, так и страницы формул не доставят вам определенного результата по сомнительным данным".

Подсказки и ответы к упражнениям

Глава 1

6.3 Интерполяционные полиномы, соответственно, Ньютона и Лагранжа. Матрица перехода:

$$C = \begin{pmatrix} (x_1 - x_2)(x_1 - x_3)(x_1 - x_4) & 0 & 0 & 0 \\ (x_1 - x_3)(x_1 - x_4) & (x_2 - x_3)(x_2 - x_4) & 0 & 0 \\ x_1 - x_4 & x_2 - x_4 & x_3 - x_4 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

Глава 3

5.2 $\alpha = 0$.
6.2 $\{s_k\}_{k=0}^5 = \{3,\ -10,\ 0,\ 200,\ -1000,\ 0\}$
7.1 Нет. **7.4** $1/\sqrt{\lambda_{\max}}$, где $\lambda_{\max}$ — максимальное собственное число матрицы A.
9.1

$$\mathbf{A}_{\mathfrak{J}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & 1 & 2 \end{pmatrix}$$

Глава 4

2.1 $^1\!/_6(x_0 + 2\,x_1 + 3\,x_2)$.
3.3 $^1\!/_{956}[187, 138, 242, 188, 201]$

Литература

ОСНОВНАЯ ЛИТЕРАТУРА

1. Гантмахер Ф.Р. *Теория матриц.* М. Наука, $\geq$ 1966
2. Гельфанд И.М. *Лекции по линейной алгебре.* М.Наука. 1971

ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА

Главы 3 и 5

Фаддеев Д.К., Фаддеева В.Н. *Вычислительные методы линейной алгебры.* М. ГИФМЛ, 1960

Глава 4

1. Феллер В. *Введение в теорию вероятностей и ее приложения.* Т.1. М.Мир. 1984
2. Романовский В.И. *Дискретные цепи Маркова.* Гостехиздат. 1948

ИСТОРИЧЕСКИЕ СПРАВКИ

составлены на основании следующих источников

1. Гродзенский С.Я. *Андрей Андреевич Марков.* М.Наука. 1987
2. *Академик А.Н.Крылов. Воспоминания и очерки.* М. АН СССР. 1956
3. Новожилов В.В. *Вопросы механики сплошной среды.* Л. Судостроение. 1989