Г.С.Шевцов

ЛИНЕЙНАЯ АЛГЕБРА

Теория и прикладные аспекты

$\boldsymbol{x}\cdot\boldsymbol{y}=\sum_{i=1}^{n}x_i y_i$

$\|\boldsymbol{x}\|_1=\sum_{i=1}^{n}|x_i|$

$\boldsymbol{x}\cdot\boldsymbol{y}=\sum_{i=1}^{n}x_i y_i$

$\|A\|=\sup_{\|x\|\neq 0}\frac{\|Ax\|}{\|x\|}$

$\|\boldsymbol{x}\|_1=\sum_{i=1}^{n}|x_i|$

$\|\boldsymbol{x}\|_1=\sum_{i=1}^{n}|x_i|$

$\|\boldsymbol{x}\|_1=\sum_{i=1}^{n}|x_i|$

$\|\boldsymbol{x}\|_2=\sum_{i=1}^{n}|x_i|^2$

$\|\boldsymbol{x}\|_2=\sum_{i=1}^{n}|x_i|^2$

$\|\boldsymbol{x}\|_2=\sum_{i=1}^{n}|x_i|^2$

УДК 512.64(075.8)
ББК 22.143я 73
Ш37

РЕЦЕНЗЕНТЫ:

кафедра высшей математики
Пермского государственного технического университета
(заведующий кафедрой — А.Р. Абдуллаев,
доктор физико-математических наук, профессор);

В.М. Суслонов,
доктор технических наук, профессор

Шевцов Г.С.

Ш37 Линейная алгебра: теория и прикладные аспекты: Учеб. пособие. – М.: Финансы и статистика, 2003. – 576 с.

ISBN 5-279-02557-7

Пособие охватывает основные разделы линейной алгебры, а также некоторые нетрадиционные: специальные разложения матриц, функции от матриц, псевдообратные матрицы, итерационные методы решения систем линейных уравнений, устойчивость решений систем линейных уравнений. Понятия и утверждения подробно разъясняются, иллюстрируются многочисленными примерами, указываются пути практического применения изучаемых фактов.

Для студентов, обучающихся по специальностям "Математика", "Прикладная математика", "Физика", "Экономика", "Экономическая кибернетика", "Инженерная технология", "Информатика" и др. Для специалистов, применяющих методы линейной алгебры в своей практической деятельности. Может быть использовано в качестве справочника.

$Ш \frac{1602020000 - 017}{010(01) - 2003}$ 242 – 2002

УДК 512.64(075.8)
ББК 22.143я 73

ISBN 5-279-02557-7

ПРЕДИСЛОВИЕ

Методы линейной алгебры широко применяются в различных областях науки, техники, экономики. Данный раздел математики занимает важное место в вузовском образовании математиков, механиков, физиков, инженеров, экономистов и других специалистов. Многолетний опыт преподавания автором этой абстрактной дисциплины показывает, что у студентов возникает немало трудностей при освоении как теоретической, так и практической части изучаемого предмета.

Линейной алгебре посвящено много прекрасных учебников и монографий. Однако практическая сторона этой дисциплины освещена в них недостаточно, по-видимому, из-за ограниченности объема имеющихся изданий. Попытку восполнить этот пробел автор предпринял несколько лет назад [35]. Настоящее пособие охватывает весь обязательный программный материал по курсу линейной алгебры; теоретические положения подкреплены подробными доказательствами; приведены методические разъяснения изучаемых понятий и фактов; даны рекомендации по решению практических задач.

Материал, выходящий за рамки обязательного курса, излагается менее строго, однако практическим вопросам, таким, как специальные разложения матриц, псевдообращение матриц, итерационные методы решения систем линейных уравнений, метод наименьших квадратов в задачах линейной алгебры, вопросы устойчивости решений систем линейных уравнений, уделяется большое внимание. Для того чтобы глубже раскрыть предмет, автор отдает предпочтение доказательствам, которые дают конкретные методы решения задач линейной алгебры. Значительное внимание уделено теории и практике применения симметрических (эрмитовых) и ортогональных (унитарных) операторов, квадратичным формам, решению полной и частичной проблем собственных значений и собственных векторов, элементам n-мерной аналитической геометрии. Представленный материал снабжен многочисленными примерами с подробными решениями. При этом каждый раз указывается на возможность практического применения рассматриваемых вопросов. Так, при обсуждении канонического разложения матрицы показана целесообразность его применения при вычислении степеней и корней из матрицы, что широко используется, например, в теории вероятностей и математической статистике, даются рекомендации по решению систем линейных уравнений с матрицей, представленной каноническим разложением. При рассмотрении сингулярного разложения матрицы предлагается возможность его применения для построения полярного разложения матрицы и псевдообратной матрицы, для отыскания псевдорешений и устойчивого решения системы линейных уравнений, для проведения сингулярного анализа при построении математических моделей.

Данное пособие состоит из 12 глав, в которых приведен традиционный (гл. 1–5, 8–12) и нетрадиционный материал (гл. 6, 7).

Глава 1 содержит общие сведения. Главы 2–5 знакомят читателя с системами линейных уравнений, матрицами и действиями над ними, линейными пространствами, линейными операторами в линейных пространствах. Главы 6, 7 посвящены канонической жордановой форме матриц и функциям от матриц. В главах 8–12 рассматриваются евклидовы и унитарные пространства и линейные операторы в них, решение систем линейных уравнений методом наименьших квадратов и итерационными методами, вопросы устойчивости решений систем линейных уравнений, квадратичные формы, приближенные методы вычисления собственных значений и собственных векторов, элементы n-мерной аналитической геометрии.

В пособии знаком $\triangleright$ обозначается начало доказательства теоремы или леммы, а знаком $\blacktriangleright$ — окончание доказательства.

Считаю приятным долгом выразить глубокую благодарность кандидату физико-математических наук, доценту МГТУ имени Н.Э. Баумана А.Н. Канатникову за неоценимую помощь при подготовке пособия к изданию. Замечания и рекомендации, сделанные им, во многом способствовали улучшению пособия при его подготовке к изданию.

Глава 1 ПЕРВОНАЧАЛЬНЫЕ СВЕДЕНИЯ

1.1. Множества, алгебраические операции, группы, кольца, поля

В любой области деятельности приходится рассматривать различные совокупности объектов, объединенные некоторым общим признаком. Такие совокупности объектов в математике принято называть ***множествами***, а сами объекты — ***элементами множеств***. Множества будем обозначать прописными латинскими буквами A, B, C, $\ldots$, X, Y, $\ldots$, а их элементы — строчными латинскими буквами a, b, c, $\ldots$, x, y, $\ldots$ При этом будем писать $x \in A$, если элемент x принадлежит множеству A, и $x \notin A$, если элемент x не принадлежит множеству A. Отметим так называемое ***пустое множество***, не содержащее ни одного элемента. Это множество будем обозначать символом $\varnothing$.

Пусть дано некоторое множество A, содержащее хотя бы один элемент. Будем говорить, что в множестве A определена ***алгебраическая операция***, если указан закон, по которому любой паре элементов a и b, взятых из этого множества в определенном порядке, ставится во взаимно однозначное соответствие некоторый элемент c, также принадлежащий этому множеству. Если эту операцию называют ***сложением***, то элемент c называют ***суммой*** элементов a и b и обозначают символом $a + b$; если операцию называют ***умножением***, то элемент c называют ***произведением*** элементов a и b и обозначают символом ab. В остальных случаях алгебраическую операцию будем обозначать символом $*$.

Алгебраическая операция $*$ называется ***коммутативной***, если результат ее применения не зависит от порядка элементов, т.е. для любых элементов a и b из рассматриваемого множества выполняется равенство $a * b = b * a$. Например, операции сложения и умножения чисел являются коммутативными операциями, а вычитание и деление — некоммутативными.

Алгебраическая операция $*$ определена для пар элементов. Исходя из трех элементов a, b, c, двукратным применением операции $*$ можно получить либо элемент $(a * b) * c$, либо элемент $a * (b * c)$. В общем случае эти элементы могут оказаться различными.

Алгебраическую операцию $*$ называют ***ассоциативной***, если для любых трех элементов a, b, c рассматриваемого множества выполняется равенство $(a * b) * c = a * (b * c)$. В этом случае результат операции $*$, примененный к элементам a, b, c, записывают в виде $a * b * c$, опуская в выражении скобки, указывающие порядок выполнения двух операций. Ассоциативная операция позволяет рассматривать выражения $a_1 * a_2 * \ldots * a_k$, содержащие произвольное конечное число элементов $a_1, a_2, \ldots, a_k$. Для ассоциативной операции результат вычисления такого выражения не зависит от расстановки скобок. Если операция ассоциативна и коммутативна, то результат не зависит и от порядка расположения элементов в этом выражении.

Для алгебраической операции $*$ часто приходится рассматривать наличие обратной операции, что равносильно решению уравнений

$$a * x = b, \qquad y * a = b$$

относительно элементов x и y из множества A. Решение этих уравнений приводит к ***правой*** и ***левой обратным операциям***. В случае их существования будем говорить, что операция $*$ имеет ***обратную операцию***. Наличие обратной операции равносильно существованию для любого элемента рассматриваемого множества ***правого*** и ***левого обратных элементов***.

Если для элемента a правый и левый обратные элементы совпадают, то этот единственный элемент называют ***обратным элементом*** к элементу a. В случае, когда алгебраическая операция названа сложением, обратный элемент к элементу a называют ***противоположным элементом*** для элемента a и обозначают символом $-a$; в случае, когда алгебраическая операция названа умножением, обратный элемент к элементу a обозначают символом a^{-1}. Это позволяет операцию, обратную к умножению, записать в виде $a/b = ab^{-1}$.

Группой называют множество с одной ассоциативной и обратимой операцией. Если алгебраическая операция в группе названа сложением, то группу называют ***аддитивной***; если алгебраическая операция в группе названа умножением, то группу называют ***мультипликативной***. Группу с коммутативной операцией называют ***абелевой***; группу, состоящую из конечного числа элементов, называют ***конечной группой***, а число элементов в группе — ***порядком группы***.

Примерами групп являются:

1. Множество всех целых чисел с операцией сложения чисел.

2. Множество всех четных чисел с операцией сложения чисел.

3. Множество всех чисел, кратных данному числу n, с операцией сложения чисел.

4. Множества всех рациональных, действительных и комплексных чисел с операцией сложения чисел.

5. Множества всех рациональных, действительных и комплексных чисел, отличных от нуля, с операцией умножения чисел.

6. Множество всех комплексных корней n-й степени из единицы с операцией умножения комплексных чисел.

7. Множества всех векторов на прямой, на плоскости и в пространстве, изучаемых в аналитической геометрии, с операцией сложения векторов.

Если дана какая-либо группа G и подмножество H, содержащееся в G, образует группу относительно алгебраической операции, заданной в G, то группу H называют ***подгруппой*** группы G. Например, аддитивная группа всех четных чисел является подгруппой аддитивной группы всех целых чисел, которая сама является подгруппой аддитивной группы всех действительных чисел.

Пусть в множестве K введены две операции — операция сложения и операция умножения. Говорят, что эти операции связаны ***законом дистрибутивности***, если для любых элементов a, b, c из K выполняются соотношения

$$(a+b)c = ac + bc, \qquad a(b+c) = ab + ac.$$

Множество K называют ***кольцом***, если в нем определены ассоциативные операции сложения и умножения, связанные законом дистрибутивности, причем операция сложения коммутативная и обладает обратной операцией — вычитанием.

Кольцо называют ***коммутативным***, если в нем операция умножения коммутативная, и ***некоммутативным*** — в противном случае. Заметим, что любое кольцо является абелевой группой по сложению.

Коммутативное кольцо P, в котором есть единичный элемент и каждый ненулевой элемент имеет обратный элемент относительно умножения, называют ***полем***.

Примерами полей являются:

1. Множества всех рациональных, действительных, комплексных чисел с операциями сложения и умножения чисел.

2. Множество всех чисел вида $a + b\sqrt{2}$, где a и b — рациональные числа, с операциями сложения и умножения чисел.

3. Множество, состоящее из двух элементов 0 и 1, с операциями сложения и умножения, заданными равенствами

$$0+0=0, \qquad 0+1=1+0=1, \qquad 1+1=0$$

и

$$0 \cdot 0 = 0, \qquad 0 \cdot 1 = 1 \cdot 0 = 0, \qquad 1 \cdot 1 = 1.$$

В любом поле множество всех элементов является *абелевой группой с операцией сложения*, а множество всех ненулевых элементов — *абелевой группой с операцией умножения*. Отсюда следует, что операции сложения и умножения имеют обратные операции — вычитание и деление. Для этих операций используют те же обозначения, что и для одноименных числовых операций. Кроме того, для четырех операций (сложение, умножение, вычитание, деление) сохраняются обычные правила преобразования выражений, а именно:

$$\frac{a}{b} \pm \frac{c}{d} = \frac{ad \pm bc}{bd}, \qquad \frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}, \qquad \frac{-a}{b} = -\frac{a}{b}.$$

Отметим также, что равенство $\frac{a}{b} = \frac{c}{d}$ верно тогда и только тогда, когда $ad = bc$. Таким образом, с точки зрения правил обращения с выражениями, включающими четыре операции, поля аналогичны полям чисел (полю рациональных, действительных или комплексных чисел). Поэтому элементы любого поля станем называть числами, если такое название не будет приводить к путанице или какой-либо двусмысленности. При этом нулевой элемент поля будем обозначать символом 0, а единичный элемент — символом 1.

Перечислим нужные нам в дальнейшем общие факты об элементах произвольного поля.

А. В поле P определена операция, называемая сложением, которая каждой паре элементов a и b поля P ставит в соответствие элемент $a + b$ из P, называемый суммой элементов a и b. При этом:

1) сложение коммутативно, т.е. $a + b = b + a$ для любых элементов a и b из P;
2) сложение ассоциативно, т.е. $a + (b + c) = (a + b) + c$ для любых элементов a, b, c из P;
3) в поле P существует единственный элемент 0, называемый нулевым, такой, что $a + 0 = a$ для любого элемента a из P;

4) для любого элемента a из P существует единственный элемент $-a$, называемый противоположным, такой, что $a+(-a)=0$ (это свойство обеспечивает существование операции, обратной к сложению, — вычитания).

B. В поле P определена операция, называемая умножением, которая каждой паре элементов a и b из P ставит в соответствие элемент ab из P, называемый произведением элементов a и b. При этом:

1) умножение коммутативно, т.е. $ab = ba$ для любых элементов a и b из P;
2) умножение ассоциативно, т.е. $a(bc) = (ab)c$ для любых элементов a, b, c из P;
3) в поле P существует единственный элемент 1, называемый единичным, такой, что $a \cdot 1 = 1 \cdot a = a$ для любого элемента a из P;
4) для каждого ненулевого элемента a из P существует единственный элемент a^{-1}, называемый обратным, такой, что $a \cdot a^{-1} = a^{-1} \cdot a = 1$ (это свойство обеспечивает существование в поле P операции, обратной к умножению, — *деления*).

C. В поле P сложение и умножение связаны законом дистрибутивности, т.е. для любых элементов a, b, c из P выполняется соотношение $(a+b)c = ac+bc$.

1.2. Простые и двойные суммы

В математике часто приходится рассматривать суммы большого количества чисел или элементов некоторого множества с операцией сложения в нем, когда слагаемые имеют одинаковый вид и различаются лишь индексами, например суммы вида

$$a_k + a_{k+1} + \ldots + a_p, \qquad \alpha_1\beta_1 + \alpha_2\beta_2 + \ldots + \alpha_n\beta_n.$$

Кратко такие суммы записывают следующим образом:

$$a_k + a_{k+1} + \ldots + a_p = \sum_{i=k}^{p} a_i, \qquad \alpha_1\beta_1 + \alpha_2\beta_2 + \ldots + \alpha_n\beta_n = \sum_{i=1}^{n} \alpha_i\beta_i,$$

где символ $\sum$ — ***символ суммы***; i — ***индекс суммирования***.

Индекс суммирования можно обозначать любой буквой, т.е.

$$\sum_{i=k}^{p} a_i = \sum_{j=k}^{p} a_j.$$

Множитель, не зависящий от индекса суммирования, можно выносить за знак суммы, т.е.

$$\sum_{k=p}^{n} \alpha x^k = \alpha \sum_{k=p}^{n} x^k.$$

Действительно,

$$\sum_{k=p}^{n} \alpha x^k = \alpha x^p + \alpha x^{p+1} + \ldots + \alpha x^n =$$
$$= \alpha\left(x^p + x^{p+1} + \ldots + x^n\right) = \alpha \sum_{k=p}^{n} x^k.$$

Часто также приходится суммировать слагаемые по двум индексам, каждый из которых независимо пробегает определенные значения. Это приводит к двойным суммам типа

$$\sum_{i=k}^{p} \sum_{j=m}^{n} a_{ij}.$$

При действиях с двойными суммами можно изменять порядок суммирования, т.е.

$$\sum_{i=k}^{p} \sum_{j=m}^{n} a_{ij} = \sum_{j=m}^{n} \sum_{i=k}^{p} a_{ij}.$$

Действительно,

$$\sum_{i=k}^{p}\sum_{j=m}^{n} a_{ij} = \sum_{i=k}^{p}\left(a_{im} + a_{i,m+1} + \ldots + a_{in}\right) =$$

$$\begin{array}{llll} = a_{km} & + a_{k,m+1} & + \ldots + a_{kn} & + \\ + a_{k+1,m} & + a_{k+1,m+1} & + \ldots + a_{k+1,n} & + \\ \ldots & \ldots & \ldots & \\ + a_{pm} & + a_{p,m+1} & + \ldots + a_{pn} & = \end{array}$$

$$= \sum_{i=k}^{p} a_{im} + \sum_{i=k}^{p} a_{i,m+1} + \ldots + \sum_{i=k}^{p} a_{i,n} = \sum_{j=m}^{n}\left(\sum_{i=k}^{p} a_{ij}\right) = \sum_{j=m}^{n}\sum_{i=k}^{p} a_{ij}.$$

Отмеченное правило распространяется на случай суммирования по любому конечному числу индексов.

1.3. Перестановки и подстановки

Пусть дано конечное множество M, состоящее из n элементов. Элементы этого множества можно перенумеровать натуральными числами $1, 2, \ldots, n$. Поскольку нас не будут интересовать свойства элементов множества, то можно принять, что элементами множества M являются сами эти числа. Всякое расположение чисел $1, 2, \ldots, n$ в определенном порядке называют ***перестановкой*** из n чисел (или из n символов). В общем случае перестановку из n символов записывают в виде $i_1\ i_2\ \ldots\ i_n$, где каждое число i_s, $s = 1, 2, \ldots, n$, есть одно из чисел $1, 2, \ldots, n$ и ни одно из этих чисел не повторяется.

Теорема 1.1. *Общее количество различных перестановок из n символов равно $n! = 1 \cdot 2 \cdot 3 \cdots n$.*

$\triangleright$ Действительно, в перестановке $i_1\ i_2\ \ldots\ i_n$ за i_1 можно принять любое из чисел $1, 2, \ldots, n$. Таких возможностей всего n. При выбранном значении i_1 за i_2 можно принять одно из $n-1$ не совпадающих с i_1 чисел. Значит, различных возможностей выбора пары символов i_1 и i_2 существует $n(n-1)$. Продолжая подсчет вариантов, через n шагов получим общее количество перестановок, равное $n \cdot (n-1) \cdots 3 \cdot 2 \cdot 1 = n!$ ▶

Говорят, что числа i и j, стоящие в перестановке, составляют ***инверсию***, или ***беспорядок***, если $i > j$, в то время как i в этой перестановке стоит раньше (левее) числа j. Так, в перестановке 2 1 5 4 3 числа 2 и 1, 5 и 4, 4 и 3, 5 и 3 составляют инверсии.

Перестановку называют ***четной***, если число всех инверсий в ней четно, и ***нечетной***, если число всех инверсий в ней нечетно.

Подсчет числа инверсий в перестановке удобно проводить следующим образом. Сначала подсчитывают, сколько чисел в перестановке, бóльших единицы, стоит левее единицы. Затем подсчитывают, сколько чисел, бóльших двух, стоит левее двойки и т.д. Пусть левее единицы стоит k_1 чисел, бóльших единицы, левее двух — k_2 чисел, бóльших двух и т.д. Наконец, пусть левее числа $n-1$ стоит k_{n-1} чисел, бóльших $n-1$. Тогда общее число инверсий в перестановке будет равно

$$k_1 + k_2 + \ldots + k_{n-1}.$$

Так, при подсчете числа инверсий в перестановке 3 2 1 5 4 получим: $k_1 = 2$, $k_2 = 1$, $k_3 = 10$, $k_4 = 1$. Поэтому общее число инверсий в этой перестановке равно $2 + 1 + 0 + 1 = 4$, а перестановка является четной.

Перемену местами каких-либо двух символов в перестановке (не обязательно соседних) называют ***транспозицией*** этих символов. Транспозицию символов i и j в перестановке обозначают через (i, j). Транспозиция в перестановке из n символов приводит к другой перестановке из тех же символов. При помощи последовательности транспозиций можно перейти от одной перестановки из n символов к любой другой перестановке из тех же символов.

Теорема 1.2. *Всякая транспозиция меняет четность перестановки.*

▷ Рассмотрим случай транспозиции рядом стоящих символов i и j, т.е. случай транспозиции этих символов в перестановке вида $A\ i\ j\ B$, где A — группа символов, стоящих в перестановке слева от символа i, а B — группа символов, стоящих в перестановке справа от символа j. Транспозиция (i, j) переводит данную перестановку в перестановку $A\ j\ i\ B$. В обеих перестановках символ i составляет одни и те же инверсии с символами из групп A и B. То же самое справедливо и для символа j. Если в данной перестановке символы i и j не составляли инверсии, то в новой перестановке появится одна новая инверсия, т.е. число инверсий увеличится на единицу. Если же символы i и j в данной перестановке составляли инверсию, то в новой перестановке она пропадет, т.е. число инверсий уменьшится на единицу. В обоих случаях четность перестановки меняется.

Теперь рассмотрим общий случай, когда символы i и j в перестановке разделены группой из s символов ($s > 0$). Здесь перестановка имеет вид:

$$A\ i\ k_1\ k_2\ \ldots\ k_s\ j\ B\,.$$

После транспозиции $(i,\ j)$ эта перестановка примет вид:

$$A\ j\ k_1\ k_2\ \ldots\ k_s\ i\ B\,.$$

Такое преобразование можно рассматривать как последовательность транспозиций, сперва символа i с символами k_1, $k_2, \ldots, k_s, j$, а затем символа j с символами $k_s, k_{s-1}, \ldots, k_1$, поэтому рассматриваемая транспозиция эквивалентна последовательности из $2s+1$ транспозиций стоящих рядом символов. При этом четность перестановки будет меняться нечетное число $2s+1$ раз, т.е. данная и полученная перестановки имеют разную четность. ▶

Теорема 1.3. *Число четных перестановок из n символов при $n \geqslant 2$ равно числу нечетных перестановок из n символов и составляет $n!/2$.*

▷ Пусть a — общее число всех четных, а b — общее число всех нечетных перестановок из n символов. В каждой четной перестановке проведем транспозицию первого и второго символов. Все полученные после этого перестановки станут нечетными и среди них не будет равных. Общее число полученных перестановок равно количеству четных перестановок a. Поэтому $a \leqslant b$. Аналогично, выполняя транспозицию первого и второго символов в нечетных перестановках, получим, что $b \leqslant a$. Из этих неравенств и из того, что общее число всех перестановок равно $n!$, следует утверждение теоремы. ▶

Всякое взаимно однозначное отображение множества чисел 1, 2, …, n на себя называют ***подстановкой*** n-й степени. Если при данной подстановке символ i_k отображается в символ α_k, то эту подстановку целесообразно записывать в виде

$$\begin{pmatrix} i_1 & i_2 & \dots & i_n \\ \alpha_{i_1} & \alpha_{i_2} & \dots & \alpha_{i_n} \end{pmatrix}. \tag{1.1}$$

Если в подстановке (1.1) совершать транспозиции (перестановки) столбцов, то будет изменяться лишь форма записи этой подстановки. Само же отображение остается тем же. Поэтому любую подстановку A можно записать в виде

$$A = \begin{pmatrix} 1 & 2 & \dots & n \\ \alpha_1 & \alpha_2 & \dots & \alpha_n \end{pmatrix}. \tag{1.2}$$

Представленные в таком виде подстановки n-й степени отличаются только второй строкой, которая фактически есть перестановка из n символов. Отсюда следует, что общее число различных подстановок n-й степени равно $n!$ Подстановку называют ***четной***, если ее верхняя и нижняя строки определяют перестановки одной четности (в этом случае сумма всех инверсий двух перестановок четная). Если подстановка записана в виде (1.1), то ее четность есть четность перестановки в нижней строке. Отсюда вытекает, что число четных подстановок n-й степени равно $n!/2$.

Глава 2 СИСТЕМЫ ЛИНЕЙНЫХ УРАВНЕНИЙ. ОПРЕДЕЛИТЕЛИ

2.1. Метод последовательного исключения неизвестных (метод Гаусса)

Пусть дана система линейных уравнений

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m. \end{cases} \tag{2.1}$$

Коэффициенты при неизвестных составляют прямоугольную таблицу

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

называемую ***матрицей системы***. Первый индекс у коэффициента a_{ij} означает номер уравнения, второй — номер неизвестного, при котором стоит этот коэффициент. Коэффициенты b_1, b_2, …, b_m называются ***свободными членами уравнений системы***. Если свободные члены равны нулю, система называется ***однородной***, в противном случае —

неоднородной. Матрицу

$$\dot{B} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix}$$

называют ***расширенной матрицей системы*** (2.1).

Решение системы (2.1) — это любой упорядоченный набор $(x_1, x_2, \dots, x_n)$ из n чисел, при подстановке которых в уравнения системы вместо соответствующих неизвестных каждое уравнение системы превращается в тождество. Система, не имеющая ни одного решения, называется ***несовместной***, или ***противоречивой***. Система, имеющая хотя бы одно решение, называется ***совместной***.

Совместные системы подразделяют на ***определенные***, обладающие единственным решением, и ***неопределенные***, обладающие множеством решений. Однородная система всегда совместна, так как имеет по крайней мере нулевое решение $x_1 = x_2 = \ldots = x_n = 0$.

Выражения (формулы), содержащие неизвестные x_1, x_2, $\dots$, x_n и некоторый набор произвольных постоянных, из которых при соответствующем выборе значений произвольных постоянных можно получить любое конкретное решение системы, называют ***общим решением системы***, а любое конкретное решение системы — ее ***частным решением***. Две системы с одними и теми же неизвестными ***эквивалентны*** (***равносильны***), если каждое решение одной из них является решением другой или обе системы несовместны.

Над уравнениями системы обычно приходится проводить следующие ***элементарные преобразования***:

1) умножение обеих частей какого-либо уравнения на число, отличное от нуля;

2) прибавление (вычитание) к одному уравнению другого, умноженного на некоторое число;

3) перестановку уравнений;

4) вычеркивание уравнений вида $0 \cdot x_1 + 0 \cdot x_2 + \ldots + 0 \cdot x_n = 0$, т.е. тождеств $0 = 0$;

5) перестановку неизвестных в системе уравнений.

В результате элементарных преобразований система преобразуется в эквивалентную. Общий способ отыскания решений обычно основывается на последовательном переходе с помощью элементарных преобразований от данной системы к такой эквивалентной системе, для которой решение находится просто. Одним из таких способов является ***метод последовательного исключения неизвестных*** (***метод Гаусса***). Алгоритм этого метода состоит в следующем.

Предположим, что коэффициент a_{11} системы (2.1) отличен от нуля. Этого всегда можно добиться, переставляя в случае необходимости уравнения системы или неизвестные в ней и меняя нумерацию неизвестных. Умножим первое уравнение на a_{21}/a_{11} и вычтем из второго уравнения, затем на a_{31}/a_{11} и вычтем из третьего уравнения и т.д. Наконец, умножим первое уравнение на a_{m1}/a_{11} и вычтем из последнего уравнения. В результате неизвестное x_1 будет исключено из всех уравнений, кроме первого, и система примет вид:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ \qquad\quad a'_{22}x_2 + \ldots + a'_{2n}x_n = b'_2, \\ \qquad\quad \ldots\ldots\ldots\ldots\ldots \\ \qquad\quad a'_{m2}x_2 + \ldots + a'_{mn}x_n = b'_m. \end{cases} \tag{2.2}$$

В системе (2.2) следует вычеркнуть уравнения вида $0 \cdot x_1 + 0 \cdot x_2 + \ldots + 0 \cdot x_n = 0$, если такие появились. На этом первый шаг метода Гаусса заканчивается. Элемент a_{11} называют ***ведущим элементом*** этого шага.

Следующие шаги прямого хода метода Гаусса осуществляются аналогично. Так, на втором шаге при $a'_{22} \neq 0$ последовательно умножаем второе уравнение на a'_{32}/a'_{22}, a'_{42}/a'_{22}, …, a'_{m2}/a'_{22} и вычитаем его из 3-го, 4-го, …, m-го уравнений. В результате неизвестное x_2 исключается из всех уравнений, кроме 1-го и 2-го. На третьем шаге неизвестное x_3 исключается из всех уравнений, кроме первых трех, и т.д.

Возможно, что на некотором шаге прямого хода метода Гаусса встретится уравнение вида

$$0\cdot x_1 + 0\cdot x_2 + \ldots + 0\cdot x_n = b_i, \quad b_i \neq 0. \tag{2.3}$$

Тогда рассматриваемая система несовместна, и дальнейшее ее решение прекращается. Если же при выполнении прямого хода метода Гаусса не встретятся уравнения вида (2.3), то рассматриваемая система не более чем через m шагов прямого хода преобразуется в эквивалентную систему вида

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1r}x_r + \ldots + a_{1n}x_n = b_1, \\ \qquad\quad a_{22}x_2 + \ldots + a_{2r}x_r + \ldots + a_{2n}x_n = b_2, \\ \qquad\qquad \ldots\ldots\ldots\ldots\ldots \\ \qquad\qquad\qquad a_{rr}x_r + \ldots + a_{rn}x_n = b_r. \end{cases} \tag{2.4}$$

Для упрощения записи в системе (2.4) штрихи над коэффициентами опущены. В ней не более m уравнений, т.е. $r \leqslant m$, так как некоторые уравнения, возможно, были приведены к виду $0 = 0$ и вычеркнуты.

При $r = n$ система (2.4) имеет треугольный вид:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ \qquad\quad a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \qquad\qquad \ldots\ldots \\ \qquad\qquad\qquad a_{nn}x_n = b_n, \end{cases} \tag{2.5}$$

и в ней легко совершить обратный ход метода Гаусса. Для этого из последнего уравнения этой системы найдем значение неизвестного x_n. Подставив его в предпоследнее уравнение, найдем значение x_{n-1}. Продолжая так далее, однозначно определим все неизвестные $x_1, x_2, \ldots, x_n$. Следовательно, если система (2.1) при прямом ходе метода Гаусса сводится к системе треугольного вида, то такая система определенная, т.е. имеет единственное решение.

При $r < n$ система (2.4) имеет вид трапеции. В ней неизвестные $x_1, x_2, \ldots, x_r$ принимают за главные, а неизвестные $x_{r+1}, x_{r+2}, \ldots, x_n$ — за свободные. Свободные неизвестные могут принимать любые фиксированные значения. Полагая $x_{r+1}=\gamma_{r+1}$, $x_{r+2}=\gamma_{r+2}$, ..., $x_n=\gamma_n$, где $\gamma_{r+1}, \ldots, \gamma_n$ — произвольные постоянные, и проведя в системе обратный ход метода Гаусса, получим формулы:

$$\left\{\begin{array}{ll} x_1 & =\beta_1+\alpha_{1,r+1}\gamma_{r+1}+\ldots+\alpha_{1n}\gamma_n, \\ x_2 & =\beta_2+\alpha_{2,r+1}\gamma_{r+1}+\ldots+\alpha_{2n}\gamma_n, \\ \multicolumn{2}{l}{\ldots\ldots\ldots\ldots\ldots\ldots\ldots} \\ x_r & =\beta_r+\alpha_{r,r+1}\gamma_{r+1}+\ldots+\alpha_{rn}\gamma_n, \\ x_{r+1} & =\gamma_{r+1}, \\ \multicolumn{2}{l}{\ldots\ldots} \\ x_n & =\gamma_n, \end{array}\right. \tag{2.6}$$

которые составляют общее решение системы (2.1). Из общего решения (2.6) при конкретных значениях γ_{r+1}, γ_{r+2}, ..., γ_n будут получаться частные решения системы (2.1). Так как каждое свободное неизвестное может принимать бесчисленное множество значений, система (2.1) при $r<n$, т.е. в случае, когда она приводится к трапецеидальному виду, обладает бесчисленным множеством решений. Это справедливо для совместных систем, имеющих меньше уравнений, чем неизвестных, и, в частности, для однородных, имеющих меньше уравнений, чем неизвестных.

На практике метод Гаусса обычно реализуют в матричной форме. Для этого выписывают расширенную матрицу системы, в которой для удобства отделяют вертикальной чертой столбец свободных членов, и преобразования проводят над этой матрицей, затем над полученной и т.д. При этом матрицы эквивалентных систем также считают эквивалентными.

Пример 2.1. Методом Гаусса решить систему уравнений

$$\left\{\begin{array}{l} x_1-2x_2+4x_3=3, \\ 3x_1-x_2+5x_3=2, \\ 2x_1+x_2+x_3=-1, \\ 2x_1-4x_2+3x_3=1. \end{array}\right.$$

Решение. Оставляя в расширенной матрице системы

$$\left(\begin{array}{ccc|c} 1 & -2 & 4 & 3 \\ 3 & -1 & 5 & 2 \\ 2 & 1 & 1 & -1 \\ 2 & -4 & 3 & 1 \end{array}\right)$$

первую строку без изменения и вычитая утроенную первую строку из второй, удвоенную первую строку из третьей и четвертой, придем к эквивалентной матрице

$$\left(\begin{array}{ccc|c} 1 & -2 & 4 & 3 \\ 0 & 5 & -7 & -7 \\ 0 & 5 & -7 & -7 \\ 0 & 0 & -5 & -5 \end{array}\right).$$

Вычитая в этой матрице вторую строку из третьей и оставляя другие строки без изменения, получим матрицу

$$\left(\begin{array}{ccc|c} 1 & -2 & 4 & 3 \\ 0 & 5 & -7 & -7 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -5 & -5 \end{array}\right).$$

Вычеркивая здесь третью строку, придем к матрице

$$\left(\begin{array}{ccc|c} 1 & -2 & 4 & 3 \\ 0 & 5 & -7 & -7 \\ 0 & 0 & -5 & -5 \end{array}\right),$$

которая соответствует системе

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

Отсюда, совершая обратный ход метода Гаусса, найдем

$$x_1 = -1, \qquad x_2 = 0, \qquad x_3 = 1.$$

Пример 2.2. Методом Гаусса решить систему уравнений

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

Решение. Если в расширенной матрице системы

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 2 & 4 & 6 & 3 \\ 3 & 1 & -1 & 1 \end{array}\right)$$

первую строку оставить без изменения, удвоенную первую строку вычесть из второй, утроенную первую строку вычесть из третьей, то получим матрицу

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & -5 \\ 0 & -5 & -10 & -11 \end{array}\right).$$

Строка $(0\quad 0\quad 0\mid -5)$ соответствует уравнению $0\cdot x_1+$ $+0\cdot x_2+0\cdot x_3=-5$. Наличие такого уравнения указывает на несовместность рассматриваемой системы.

Пример 2.3. Методом Гаусса решить систему уравнений

$$\begin{cases} x_1+2x_2+3x_3=4, \\ 2x_1+\ \ x_2-\ \ x_3=3, \\ 3x_1+3x_2+2x_3=7. \end{cases}$$

Решение. Элементарные преобразования прямого хода метода Гаусса над строками расширенной матрицы системы дают следующую цепочку эквивалентных матриц:

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 2 & 1 & -1 & 3 \\ 3 & 3 & 2 & 7 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & -3 & -7 & -5 \\ 0 & -3 & -7 & -5 \end{array}\right) \sim$$

$$\sim \left(\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & -3 & -7 & -5 \\ 0 & 0 & 0 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & -3 & -7 & -5 \end{array}\right).$$

Последняя матрица этой цепочки соответствует системе

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

Полагая здесь $x_3 = \gamma_3$ (γ_3 — произвольная постоянная) и проводя обратный ход метода Гаусса, получим общее решение:

$$x_1 = \frac{1}{3}\big(2 + 5\gamma_3\big), \quad x_2 = \frac{1}{3}\big(5 - 7\gamma_3\big), \quad x_3 = \gamma_3.$$

Для повышения эффективности и устойчивости метода Гаусса его модифицируют различными способами. Например, часто применяют схему, в которой на каждом шаге прямого хода ведущий коэффициент выбирают наибольшим по модулю среди коэффициентов при неизвестных в выбранном уравнении или в подсистеме, с которой работают на данном этапе.

При решении систем „вручную" методом Гаусса, чтобы избежать сложных вычислений, иногда в промежутках между шагами прямого хода метода Гаусса или до его начала целесообразно проделывать дополнительные элементарные преобразования над некоторыми уравнениями системы. Например, при решении „вручную" системы

$$\begin{cases} 5x_1 + 9x_2 + 13x_3 = 1, \\ 6x_1 + 5x_2 - 10x_3 = 3, \\ 2x_1 + 4x_2 + 3x_3 = 2 \end{cases}$$

елесообразно сначала из первого уравнения системы вычесть двоенное третье, а остальные оставить без изменения. Тогда олучим систему

$$\begin{cases} x_1 + x_2 + 7x_3 = -3, \\ 6x_1 + 5x_2 - 10x_3 = 3, \\ 2x_1 + 4x_2 + 3x_3 = 2, \end{cases}$$

которой метод Гаусса проводится уже легко. Дополнитель-ые преобразования совершаются также над матрицами.

В заключение отметим, что метод Гаусса и его модификации находят самое широкое применение в вычислительной практике. Для его реализации на ЭВМ можно использовать стандартные программы, которые включены практически в любой пакет программ для решения математических задач.

2.2. Определители

Определителем (*детерминантом*) ***n-го порядка***, соответствующим квадратной матрице

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

называется алгебраическая сумма $n!$ членов вида

$$a_{1\alpha_1} a_{2\alpha_2} \dots a_{n\alpha_n}.$$

Эти члены представляют собой всевозможные произведения по n элементов матрицы A, взятых по одному из каждой строки и каждого столбца матрицы. Произведения берутся со знаком $(-1)^t$, где t — число инверсий в перестановке $\alpha_1\ \alpha_2\ \dots\ \alpha_n$, составленной из вторых индексов элементов матрицы, входящих в рассматриваемое произведение. Определитель обозначают одним из следующих символов:

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}, \qquad \det A, \qquad |A|, \qquad \Delta.$$

Кратко:

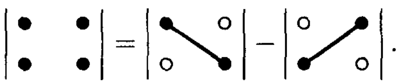
$$|A| = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \sum (-1)^t a_{1\alpha_1} a_{2\alpha_2} \dots a_{n\alpha_n},$$

где суммирование проводится по всевозможным перестановкам $\alpha_1\ \alpha_2\ \ldots\ \alpha_n$ из n символов $1, 2, \ldots, n$; t — число инверсий в перестановке $\alpha_1\ \alpha_2\ \ldots\ \alpha_n$. Поскольку число перестановок из n символов равно $n!$, то определитель n-го порядка состоит из $n!$ слагаемых, причем половина из них, т.е. $n!/2$, входит в определитель со знаком „плюс“, а половина — со знаком „минус“.

Для определителя второго порядка непосредственно по определению получаем формулу

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

которую легко запомнить по следующей схеме:



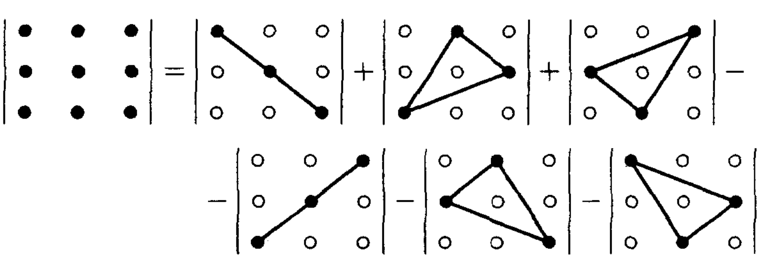
Так, например,

$$\begin{vmatrix} 3 & 5 \\ 2 & 7 \end{vmatrix} = 3\cdot 7 - 2\cdot 5 = 21 - 10 = 11.$$

Для определителя третьего порядка также непосредственно из определения получаем:

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

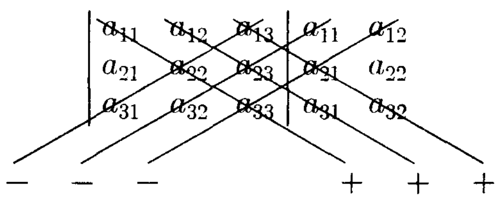
Для запоминания этой формулы удобно пользоваться правилом треугольников, которое схематично записывается так:



Например,

$$\begin{vmatrix} 2 & 3 & 4 \\ 5 & 7 & 8 \\ 3 & 2 & 1 \end{vmatrix} = 2\cdot7\cdot1+3\cdot8\cdot3+4\cdot5\cdot2-4\cdot7\cdot3-3\cdot5\cdot1-2\cdot8\cdot2=-5.$$

При выписывании формулы для вычисления определителя третьего порядка поступают и так. К определителю приписывают справа его первый и второй столбцы. Затем записывают нужную формулу, последовательно составляя произведения по три элемента, стоящих на одной диагонали, и приписывая знаки этим произведениям по следующей схеме:



2.3. Свойства определителей

Вычисление определителей выше третьего порядка непосредственно по определению весьма затруднительно. Поэтому необходимо сначала изучить свойства определителей, а затем выработать методы вычисления определителей.

Преобразование, при котором строки определителя заменяются на его соответствующие столбцы, называют ***транспонированием определителя***.

Свойство 2.1. Определитель не меняется при транспонировании.

Действительно, любой член определителя

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

имеет вид $a_{1\alpha_1}a_{2\alpha_2}\dots a_{n\alpha_n}$. Но это произведение является также членом определителя

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

и наоборот.

Знак рассматриваемого произведения в исходном определителе определяется числом инверсий в подстановке

$$\begin{pmatrix} 1 & 2 & \dots & n \\ \alpha_1 & \alpha_2 & \dots & \alpha_n \end{pmatrix},$$

а знак этого же произведения в транспонированном определителе определяется числом инверсий в подстановке

$$\begin{pmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_n \\ 1 & 2 & \dots & n \end{pmatrix},$$

но число инверсий в обеих подстановках одинаково. Поэтому рассматриваемое произведение имеет в обоих определителях один и тот же знак. Следовательно, исходный и транспонированный определители являются суммами одинаковых членов с одинаковыми знаками, т.е. равны друг другу.

Из свойства 2.1 вытекает, что в определителе строки и столбцы равноправны, т.е. всякое утверждение относительно строк определителя верно и для его столбцов, и наоборот. Поэтому в дальнейшем будем формулировать и доказывать свойства определителя только для его строк. Аналогичные свойства определителя для столбцов также будут выполняться.

Свойство 2.2. Если одна из строк определителя состоит из нулей, то определитель равен нулю.

Это следует из того, что каждый член определителя будет иметь множитель нуль из этой строки.

Свойство 2.3. При перестановке двух строк определитель меняет знак.

Действительно, пусть дан определитель

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \cdot & \cdot & \cdot & \cdot \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

Поменяв в нем местами строки с номерами i и j, получим новый определитель

$$\Delta_1 = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \cdot & \cdot & \cdot & \cdot \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

Рассмотрим произвольный член определителя Δ

$$a_{1\alpha_1} a_{2\alpha_2} \ldots a_{i\alpha_i} \ldots a_{j\alpha_j} \ldots a_{n\alpha_n}.$$

Все его множители и в определителе Δ_1 также находятся на разных строках и в разных столбцах. Следовательно, это произведение будет членом определителя Δ_1. Поэтому определители Δ и Δ_1 состоят из одних и тех же членов. В определителе Δ рассматриваемому члену соответствует подстановка

$$\begin{pmatrix} 1 & 2 & \ldots & i & \ldots & j & \ldots & n \\ \alpha_1 & \alpha_2 & \ldots & \alpha_i & \ldots & \alpha_j & \ldots & \alpha_n \end{pmatrix},$$

а в определителе Δ_1 — подстановка

$$\begin{pmatrix} 1 & 2 & \ldots & j & \ldots & i & \ldots & n \\ \alpha_1 & \alpha_2 & \ldots & \alpha_j & \ldots & \alpha_i & \ldots & \alpha_n \end{pmatrix}.$$

Количество инверсий в этих подстановках отличается на нечетное число. Поэтому члены определителей Δ и Δ_1 различаются лишь знаками.

Свойство 2.4. Определитель, содержащий две одинаковые строки, равен нулю.

Действительно, пусть в определителе Δ две строки одинаковые. Поменяем эти строки местами. Тогда получим определитель Δ_1, который, с одной стороны, равен определителю Δ, так как поменяли местами одинаковые строки, а, с другой стороны, по предыдущему свойству определитель Δ_1 отличается от определителя Δ знаком. Таким образом, $\Delta = -\Delta$ и $\Delta = 0$.

Свойство 2.5. Если все элементы некоторой строки умножить на число k, то определитель умножится на это число.

Это свойство следует из того, что каждый член определителя умножается на число k.

Следствие. Общий множитель элементов строки определителя можно вынести за знак определителя.

Свойство 2.6. Определитель, содержащий пропорциональные строки, равен нулю.

Действительно, пусть элементы i-й строки определителя отличаются от соответствующих элементов j-й строки одним и тем же множителем. Вынося этот множитель за знак определителя, получим определитель с двумя одинаковыми строками, который по свойству 2.4 равен нулю.

Свойство 2.7. Если все элементы a_{ij} i-й строки определителя n-го порядка представлены в виде суммы двух слагаемых, т.е.

$$a_{ij} = b_{ij} + c_{ij}, \quad j = 1, 2, \ldots, n,$$

то определитель равен сумме двух определителей, у которых все строки, кроме i-й, такие же, как и у исходного определителя, а i-я строка в первом слагаемом состоит из элементов b_{ij}, во втором — из элементов c_{ij}, $j = 1, 2, \ldots, n$. Таким образом,

$$\begin{vmatrix} a_{11} & a_{21} & \ldots & a_{n1} \\ \ldots & \ldots & \ldots & \ldots \\ b_{12}+c_{12} & b_{22}+c_{22} & \ldots & b_{n2}+c_{n2} \\ \ldots & \ldots & \ldots & \ldots \\ a_{1n} & a_{2n} & \ldots & a_{nn} \end{vmatrix} =$$

$$= \begin{vmatrix} a_{11} & a_{21} & \ldots & a_{n1} \\ \ldots & \ldots & \ldots & \ldots \\ b_{12} & b_{22} & \ldots & c_{n2} \\ \ldots & \ldots & \ldots & \ldots \\ a_{1n} & a_{2n} & \ldots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{21} & \ldots & a_{n1} \\ \ldots & \ldots & \ldots & \ldots \\ c_{12} & c_{22} & \ldots & c_{n2} \\ \ldots & \ldots & \ldots & \ldots \\ a_{1n} & a_{2n} & \ldots & a_{nn} \end{vmatrix}.$$

Это свойство вытекает из того, что любой член

$$a_{1\alpha_1} a_{2\alpha_2} \ldots (b_{i\alpha_i} + c_{i\alpha_i}) \ldots a_{n\alpha_n}$$

исходного определителя можно представить в виде суммы

$$a_{1\alpha_1} a_{2\alpha_2} \ldots b_{i\alpha_i} \ldots a_{n\alpha_n} + a_{1\alpha_1} a_{2\alpha_2} \ldots c_{i\alpha_i} \ldots a_{n\alpha_n}.$$

Замечание. Свойство 2.7 распространяется на случай, когда каждый элемент некоторой строки определителя является суммой m слагаемых, $m > 2$.

Свойство 2.8. Определитель не меняется, если к элементам одной его строки прибавить соответствующие элементы другой строки, умноженные на одно и то же число.

Действительно, пусть к i-й строке определителя Δ прибавлена его k-я строка, умноженная на число α. Тогда в новом определителе элементы i-й строки имеют вид:

$$a_{ij} + \alpha a_{kj}, \quad j = 1, 2, \ldots, n,$$

где a_{ij} и a_{kj} — элементы i-й и k-й строк исходного определителя. По предыдущему свойству этот определитель равен сумме двух определителей, из которых первый — это исходный определитель Δ, а второй содержит две пропорциональные строки и потому равен нулю.

Замечание. Так как множитель α может быть отрицательным, то определитель не меняется и при вычитании из одной строки его другой строки, умноженной на некоторое число.

Говорят, что i-я строка определителя является линейной комбинацией других его строк, если она представлена в виде суммы этих строк, умноженных на некоторые числа.

Свойство 2.9. Если одна из строк определителя является линейной комбинацией других его строк, то этот определитель равен нулю.

Действительно, пусть, например, i-я строка определителя является линейной комбинацией его других строк. Тогда каждый элемент i-й строки будет суммой m слагаемых. По свойству 2.7 определитель может быть представлен в виде суммы m определителей, но при этом у каждого из этих определителей i-я строка будет пропорциональна одной из других строк. По свойству 2.6 все эти определители равны нулю. Следовательно, исходный определитель также равен нулю.

2.4. Миноры и алгебраические дополнения

Пусть дан определитель n-го порядка. Выделим в нем произвольно k строк и k столбцов. Определитель k-го порядка, составленный из элементов, стоящих на пересечениях выделенных строк и столбцов, называют ***минором*** k-го порядка определителя. В частности, минорами первого порядка являются элементы определителя. Миноры, расположенные симметрично относительно главной диагонали определителя, называют ***главными***.

Если в определителе n-го порядка выделить k строк и k столбцов и из элементов, стоящих на их пересечении, составить минор M k-го порядка, затем вычеркнуть выделенные k строк и k столбцов, то из оставшихся элементов можно составить определитель M' $(n-k)$-го порядка. Этот определитель называют ***дополнительным минором*** к минору M.

Очевидно, что если минор M' является дополнительным к минору M, то и, наоборот, минор M является дополнительным к минору M'. Дополнительный минор элемента a_{ij} обозначают M'_{ij}. ***Алгебраическим дополнением*** минора M называют его дополнительный минор M', взятый со знаком $(-1)^{s_M}$, т.е. $(-1)^{s_M}M'$, где s_M — сумма номеров всех строк и столбцов, в которых располагается минор M. В частности, для алгебраического дополнения A_{ij} элемента a_{ij} получается формула

$$A_{ij} = (-1)^{i+j}M'_{ij}.$$

Теорема 2.1. *Произведение любого минора M k-го порядка на его алгебраическое дополнение $A = (-1)^{s_M}M'$ представляет собой сумму $k!(n-k)!$ различных членов определителя Δ с теми же знаками, с какими они входят в определитель Δ.*

$\triangleright$ Сначала рассмотрим случай, когда минор M расположен в верхнем левом углу определителя:

$$\Delta = \left|\begin{array}{ccc|ccc} a_{11} & \ldots & a_{1k} & a_{1,k+1} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{k1} & \ldots & a_{kk} & a_{k,k+1} & \ldots & a_{kn} \\ \hline a_{k+1,1} & \ldots & a_{k+1,k} & a_{k+1,k+1} & \ldots & a_{k+1,n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{nk} & a_{n,k+1} & \ldots & a_{nn} \end{array}\right|,$$

т.е. в строках и столбцах с номерами 1, 2, ..., n. Тогда дополнительный минор M' будет расположен в правом нижнем углу определителя Δ, т.е. в строках и столбцах с номерами $k+1$, $k+2$, ..., n. Алгебраическое дополнение $A = (-1)^{s_M}M'$

в этом случае совпадает с дополнительным минором M', так как число

$$s_M = (1+2+\ldots+k)+(1+2+\ldots+k) = 2(1+2+\ldots+k)$$

является четным.

Произвольные члены миноров M и M' имеют вид соответственно

$$(-1)^t a_{1\alpha_1} a_{2\alpha_2} \ldots a_{k\alpha_k} \tag{2.7}$$

и

$$(-1)^{t'} a_{k+1,\beta_{k+1}} a_{k+2,\beta_{k+2}} \ldots a_{n\beta_n}, \tag{2.8}$$

где t — число инверсий в перестановке $\alpha_1\ \alpha_2\ \ldots\ \alpha_k$, а t' — число инверсий в перестановке $\beta_{k+1}\ \beta_{k+2}\ \ldots\ \beta_n$. Общее число членов вида (2.7) равно $k!$, а число членов вида (2.8) равно $(n-k)!$ Перемножив произведения (2.7) и (2.8), получим произведение

$$(-1)^{t+t'} a_{1\alpha_1} a_{2\alpha_2} \ldots a_{k\alpha_k} a_{k+1,\beta_{k+1}} a_{k+2,\beta_{k+2}} \ldots a_{n\beta_n}. \tag{2.9}$$

Это произведение является членом определителя Δ, так как его сомножители взяты из разных строк и разных столбцов этого определителя и их количество равно n. Кроме того, число $t+t'$ совпадает с числом инверсий в перестановке

$$\alpha_1\ \alpha_2\ \ldots\ \alpha_k\ \beta_{k+1}\ \beta_{k+2}\ \ldots\ \beta_n$$

в силу того, что все $\alpha_i \leqslant k$, а все $\beta_j \geqslant k+1$ и никакой символ α_i не может образовать инверсию ни с каким символом β_j. Число членов вида (2.9) равно $k!(n-k)!$ Этим доказана теорема в рассматриваемом частном случае.

Теперь рассмотрим общий случай, когда минор M расположен в произвольных строках с номерами $i_1, i_2, \ldots, i_k$ и произвольных столбцах с номерами $j_1, j_2, \ldots, j_k$, где можно считать, что $i_1 < i_2 < \ldots < i_k$ и $j_1 < j_2 < \ldots < j_k$. Последовательно переставляя строку с номером i_1 с предыдущими строками с номерами $i_1 - 1$, $i_1 - 2$ и т.д. (всего таких перестановок $i_1 - 1$), переведем строку с номером i_1 на первое место. Затем аналогичным образом, совершив $i_2 - 2$ перестановок, переведем строку с номером i_2 на второе место. Продолжая так и далее, получим определитель, в котором минор M окажется расположенным в первых k строках, причем общее число перестановок строк равно

$$(i_1 - 1) + (i_2 - 2) + \ldots + (i_k - k) = \\ = (i_1 + i_2 + \ldots + i_k) - (1 + 2 + \ldots + k).$$

Описанную процедуру повторим и для столбцов минора M, совершив всего

$$(j_1 - 1) + (j_2 - 2) + \ldots + (j_k - k) = \\ = (j_1 + j_2 + \ldots + j_k) - (1 + 2 + \ldots + k)$$

перестановок столбцов. В результате всех перестановок (и строк, и столбцов) минор M окажется в левом верхнем углу нового определителя Δ'. Дополнительный минор M' в результате этих перестановок не изменится и окажется в правом нижнем углу определителя Δ'. Определитель Δ' получен из определителя Δ в результате

$$\big[(i_1 - 1) + (i_2 - 2) + \ldots + (i_k - k)\big] + \\ + \big[(j_1 - 1) + (j_2 - 2) + \ldots + (j_k - k)\big] = \\ = s_M - 2(1 + 2 + \ldots + k)$$

перестановок его строк и столбцов, при каждой из которых меняется знак определителя. Поэтому

$$\Delta' = (-1)^{s_M} \Delta, \tag{2.10}$$

так как число $2(1 + 2 + \ldots + k)$ четное.

По доказанному выше произведение MM' является суммой $k!(n-k)!$ членов определителя Δ' с теми же знаками, с какими они входят в определитель Δ'. В то же время из соотношения

(2.10) следует, что члены определителей Δ' и Δ отличаются лишь множителем $(-1)^{s_M}$. Поэтому произведение

$$(-1)^{s_M} M M' = M(-1)^{s_M} M' = MA$$

состоит из $k!(n-k)!$ членов определителя Δ с теми же знаками, какие они имеют в определителе Δ. Тем самым теорема доказана полностью. ►

2.5. Разложение определителя по строке или столбцу. Теорема Лапласа

Теорема 2.2. *Определитель Δ n-го порядка равен сумме произведений всех элементов любой его строки (столбца) на их алгебраические дополнения, т.е.*

$$\Delta = a_{i1}A_{i1} + a_{i2}A_{i2} + \ldots + a_{in}A_{in}, \tag{2.11}$$

$$\Delta = a_{1j}A_{1j} + a_{2j}A_{2j} + \ldots + a_{nj}A_{nj}. \tag{2.12}$$

В частности, если в какой-либо строке (столбце) все элементы, кроме одного, нули, то определитель равен произведению этого не равного нулю элемента на его алгебраическое дополнение.

◁ Выберем какую-либо, например i-ю, строку определителя Δ и составим произведения

$$a_{i1}A_{i1}, \quad a_{i2}A_{i2}, \quad \ldots, \quad a_{in}A_{in} \tag{2.13}$$

каждого ее элемента a_{ij} на его алгебраическое дополнение A_{ij}. По теореме 2.1 каждое такое произведение является суммой $(n-1)!$ различных членов определителя Δ с такими же знаками, с какими они входят в определитель Δ. При этом разные произведения $a_{ij}A_{ij}$ и $a_{ik}A_{ik}$ не имеют общих членов определителя Δ, так как члены произведения $a_{ij}A_{ij}$ из i-й строки имеют множитель a_{ij}, а члены произведения

$a_{ik}A_{ik}$ — множитель a_{ik}. Общее число членов, входящих в произведения (2.13), равно $(n-1)!n = n!$ и исчерпывает все члены определителя Δ. Этим доказано равенство (2.11). Равенство (2.12) доказывается аналогично. ▶

Теорема 2.3. *Сумма произведений элементов какой-либо строки (столбца) определителя на алгебраические дополнения соответствующих элементов другой строки (столбца) равна нулю, т.е.*

$$a_{i1}A_{k1} + a_{i2}A_{k2} + \ldots + a_{in}A_{kn} = 0, \quad i \neq k; \tag{2.14}$$

$$a_{1j}A_{1k} + a_{2j}A_{2k} + \ldots + a_{nj}A_{nk} = 0, \quad j \neq k. \tag{2.15}$$

▷ В определителе Δ n-го порядка выберем какую-либо, например k-ю, строку и вычислим разложение

$$\Delta = a_{k1}A_{k1} + a_{k2}A_{k2} + \ldots + a_{kn}A_{kn}$$

определителя Δ по этой строке. Заменив в правой части этого разложения множители $a_{k1}, a_{k2}, \ldots, a_{kn}$ элементами i-й строки $a_{i1}, a_{i2}, \ldots, a_{in}$ определителя Δ, получим выражение

$$a_{i1}A_{k1} + a_{i2}A_{k2} + \ldots + a_{in}A_{kn}\,,$$

представляющее собой разложение по k-й строке нового определителя Δ', который получен из определителя Δ заменой k-й строки на i-ю строку. Определитель Δ' равен нулю как определитель с одинаковыми i-й и k-й строками. Поэтому

$$a_{i1}A_{k1} + a_{i2}A_{k2} + \ldots + a_{in}A_{kn} = 0,$$

т.е. равенство (2.14) доказано. Равенство (2.15) доказывается аналогично. ▶

Теорема 2.4 (теорема Лапласа). *Пусть в определителе Δ n-го порядка произвольно выбраны k строк (или k столбцов), $1 \leqslant k \leqslant n-1$. Тогда сумма произведений всех миноров k-го порядка, расположенных в выбранных* *строках (столбцах), на их алгебраические дополнения равна определителю* Δ.

$\triangleright$ Пусть в определителе Δ n-го порядка произвольно выбраны k строк с номерами i_1, i_2, ..., i_k, где $i_1 < i_2 < \ldots < i_k$. Обозначим всевозможные миноры k-го порядка, расположенные в выбранных строках, через M_1, M_2, ..., M_t, а их алгебраические дополнения — через A_1, A_2, ..., A_t. Докажем, что

$$M_1A_1 + M_2A_2 + \ldots + M_tA_t = \Delta. \tag{2.16}$$

По теореме 2.1 каждое произведение M_iA_i $(i = 1, 2, \ldots, t)$ состоит из членов определителя Δ с теми же знаками, с какими они входят в определитель Δ.

Произведения M_iA_i и M_jA_j при $i \neq j$ не имеют общих членов определителя Δ, так как миноры M_i и M_j отличаются друг от друга по крайней мере одним столбцом. Следовательно, левая часть равенства (2.16) состоит из разных членов определителя Δ. Покажем, что в левую часть этого равенства входят все члены определителя Δ. Произвольный член определителя Δ имеет вид:

$$a_{1\alpha_1}a_{1\alpha_1}\ldots a_{n\alpha_n}. \tag{2.17}$$

Возьмем отдельно произведение тех множителей из этого члена определителя, которые входят в выбранные строки с номерами i_1, i_2, ..., i_k. Это будет произведение

$$a_{i_1\alpha_{i_1}}a_{i_2\alpha_{i_2}}\ldots a_{i_k\alpha_{i_k}}.$$

Очевидно, что оно является членом минора M, расположенного в выбранных строках и в столбцах с номерами α_{i_1}, α_{i_2}, ..., α_{i_k}. Произведение оставшихся сомножителей члена (2.17) определителя будет членом алгебраического дополнения A минора M, а весь член (2.17) определителя содержится в произведении MA. Следовательно, произвольно взятый член (2.17) определителя Δ содержится в левой части равенства (2.16), а само это равенство выполняется.

Таким образом, теорема доказана для случая, когда в определителе выбраны строки. В случае выбора в нем столбцов теорема доказывается аналогично. ►

2.6. Вычисление определителей

Разложение определителя по строке или столбцу позволяет сводить вычисление определителей больших порядков к вычислению определителей меньших порядков, но с каждым понижением порядка количество составляющих определителей возрастает, причем очень быстро. В связи с этим целесообразно, предваряя разложение определителя, преобразовывать его так, чтобы среди его элементов оказалось как можно больше нулей.

Пример 2.4. Вычислить определители:

$$1)\ \begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{vmatrix}; \qquad 2)\ \begin{vmatrix} 2 & 4 & 3 & 2 \\ 2 & 6 & 4 & 5 \\ 4 & 10 & 9 & 7 \\ 2 & 9 & 4 & 9 \end{vmatrix}.$$

Решение. 1) Разлагая первый определитель по первому столбцу, затем полученный определитель 3-го порядка по его первому столбцу и так далее, получим цепочку следующих равенств:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{vmatrix} = a_{11}(-1)^{1+1} \begin{vmatrix} a_{22} & a_{23} & a_{24} \\ 0 & a_{33} & a_{34} \\ 0 & 0 & a_{44} \end{vmatrix} =$$

$$= a_{11}a_{22}(-1)^{1+1} \begin{vmatrix} a_{33} & a_{34} \\ 0 & a_{44} \end{vmatrix} =$$

$$= a_{11}a_{22}a_{33}(-1)^{1+1}a_{44} = a_{11}a_{22}a_{33}a_{44}.$$

Замечание. Пользуясь определением определителя n-го порядка, легко установить, что если все элементы определителя, расположенные по одну или по обе стороны главной диагонали, равны нулю, то такой определитель равен произведению диагональных элементов (элементов, расположенных на главной диагонали). Аналогично если все элементы определителя, расположенные по одну или по обе стороны его побочной диагонали, равны нулю, то этот определитель равен произведению всех элементов побочной диагонали, взятому со знаком $(-1)^{\frac{n(n-1)}{2}}$.

2) В определителе

$$\begin{vmatrix} 2 & 4 & 3 & 2 \\ 2 & 6 & 4 & 5 \\ 4 & 10 & 9 & 7 \\ 2 & 9 & 4 & 9 \end{vmatrix}$$

сначала, как при методе Гаусса, первую строку, умноженную на минус единицу, прибавим ко второй строке. Затем первую строку, умноженную на минус два, прибавим к третьей строке. И наконец, первую строку, умноженную на минус единицу, прибавим к четвертой строке. Полученный определитель, у которого первый столбец будет иметь лишь один ненулевой элемент, разложим по этому столбцу. В результате придем к следующей цепочке равенств:

$$\begin{vmatrix} 2 & 4 & 3 & 2 \\ 2 & 6 & 4 & 5 \\ 4 & 10 & 9 & 7 \\ 2 & 9 & 4 & 9 \end{vmatrix} = \begin{vmatrix} 2 & 4 & 3 & 2 \\ 0 & 2 & 1 & 3 \\ 0 & 2 & 3 & 3 \\ 0 & 5 & 1 & 7 \end{vmatrix} = 2(-1)^{1+1} \begin{vmatrix} 2 & 1 & 3 \\ 2 & 3 & 3 \\ 5 & 1 & 7 \end{vmatrix}.$$

Последний определитель в этой цепочке можно уже вычислить непосредственно по формуле треугольников для определителей третьего порядка. Однако будет проще, если его сначала преобразовать, вычтя из второй строки первую, а затем разложить по второй строке. Тогда получим следующее продолжение цепочки равенств:

$$2(-1)^{1+1}\begin{vmatrix} 2 & 1 & 3 \\ 2 & 3 & 3 \\ 5 & 1 & 7 \end{vmatrix} = 2\begin{vmatrix} 2 & 1 & 3 \\ 0 & 2 & 0 \\ 5 & 1 & 7 \end{vmatrix} =$$

$$= 2\cdot 2\cdot(-1)^{2+2}\begin{vmatrix} 2 & 3 \\ 5 & 7 \end{vmatrix} = 4(14-15) = -4.$$

Таким образом, рассматриваемый определитель равен -4.

При вычислении определителей наиболее эффективен способ, в основе которого лежит метод Гаусса. Пусть дан определитель

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

Выберем ведущим элементом первого шага преобразований элемент a_{11}. Такой возможности всегда можно добиться с помощью перестановок строк и столбцов. После преобразований в соответствии с первым шагом метода Гаусса определитель преобразуется в определитель вида

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & b_{22} & b_{23} & \dots & b_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & b_{n2} & b_{n3} & \dots & b_{nn} \end{vmatrix}.$$

В этом определителе проведем преобразования согласно второму шагу метода Гаусса, затем согласно третьему шагу и т.д. В результате придем к определителю треугольного вида

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & 0 & b_{23} & \dots & b_{2n} \\ 0 & 0 & c_{33} & \dots & c_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & d_{nn} \end{vmatrix},$$

равному произведению $a_{11}b_{22}c_{33}\dots d_{nn}$. Так, для определителя

$$\Delta = \begin{vmatrix} 2 & 3 & 4 & 5 \\ 2 & 6 & 4 & 3 \\ 4 & 9 & 13 & 7 \\ 6 & 3 & 7 & 2 \end{vmatrix}$$

получаем следующую цепочку равенств:

$$\Delta = \begin{vmatrix} 2 & 3 & 4 & 5 \\ 2 & 6 & 4 & 3 \\ 4 & 9 & 13 & 7 \\ 6 & 3 & 7 & 2 \end{vmatrix} = \begin{vmatrix} 2 & 3 & 4 & 5 \\ 0 & 3 & 0 & -2 \\ 0 & 3 & 5 & -3 \\ 0 & -6 & -5 & -13 \end{vmatrix} = \begin{vmatrix} 2 & 3 & 4 & 5 \\ 0 & 3 & 0 & -2 \\ 0 & 0 & 5 & -1 \\ 0 & 0 & -5 & -17 \end{vmatrix} =$$

$$= \begin{vmatrix} 2 & 3 & 4 & 5 \\ 0 & 3 & 0 & -2 \\ 0 & 0 & 5 & -1 \\ 0 & 0 & 0 & -18 \end{vmatrix} = 2\cdot 3\cdot 5\cdot(-18) = -540.$$

2.7. Крамеровские системы

Систему

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \dots\dots\dots\dots\dots\dots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n \end{cases} \qquad (2.18)$$

n линейных уравнений с n неизвестными, определитель

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

которой отличен от нуля, будем называть ***крамеровской системой***. Для таких систем верно следующее утверждение.

Теорема 2.5. Крамеровская система (2.18) линейных уравнений всегда совместна и имеет единственное решение, *причем это решение может быть вычислено по* ***формулам Крамера***

$$x_1 = \frac{\Delta_1}{\Delta}, \quad x_2 = \frac{\Delta_2}{\Delta}, \quad \ldots, \quad x_n = \frac{\Delta_n}{\Delta}, \tag{2.19}$$

где определитель Δ_j, $j = 1, 2, \ldots, n$, получается из определителя Δ системы заменой в нем j-го столбца столбцом свободных членов системы.

▷ Предположим, что система (2.18) совместна, и докажем, что она имеет единственное решение. Пусть x_1, x_2, …, x_n — какое-либо решение системы (2.18). Умножим обе части первого уравнения системы на алгебраическое дополнение A_{1j} элемента a_{1j} определителя Δ, обе части второго уравнения системы на алгебраическое дополнение A_{2j} элемента a_{2j} определителя Δ и так далее до последнего уравнения, обе части которого умножим на алгебраическое дополнение A_{nj} элемента a_{nj} определителя Δ. Сложив полученные уравнения, придем к равенству

$$\begin{aligned}
&(a_{11}A_{1j} + a_{21}A_{2j} + \ldots + a_{n1}A_{nj})x_1 + \\
&\qquad + (a_{12}A_{1j} + a_{22}A_{2j} + \ldots + a_{n2}A_{nj})x_2 + \ldots \\
&\qquad \ldots + (a_{1j}A_{1j} + a_{2j}A_{2j} + \ldots + a_{nj}A_{nj})x_j + \ldots \\
&\qquad \ldots + (a_{1n}A_{1j} + a_{2n}A_{2j} + \ldots + a_{nn}A_{nj})x_n = \\
&\qquad\qquad\qquad = b_1A_{1j} + b_2A_{2j} + \ldots + b_nA_{nj}.
\end{aligned}$$

В этом равенстве в силу соотношений (2.12) и (2.15) выражение, стоящее множителем при x_k, $k = 1, 2, \ldots, n$, равно нулю при $k \neq j$ и равно Δ при $k = j$. Выражение в правой части равенства представляет собой разложение по j-му столбцу определителя Δ_j, получаемого из определителя Δ заменой в нем j-го столбца столбцом свободных членов системы. Следовательно, $\Delta x_j = \Delta_j$, $j = 1, 2, \ldots, n$, и, поскольку $\Delta \neq 0$,

$$x_j = \frac{\Delta_j}{\Delta} = \frac{b_1A_{1j} + b_2A_{2j} + \ldots + b_nA_{nj}}{\Delta} \tag{2.20}$$

т.е. каждое неизвестное x_j определяется однозначно. Таким образом, если система (2.18) совместная, то она имеет единственное решение.

Докажем, что система (2.18) на самом деле совместная. В этом можно убедиться непосредственной подстановкой значений $x_1, x_2, \ldots, x_n$, вычисленных по формулам Крамера (2.19), в уравнения системы (2.18). Такая подстановка в i-е уравнение системы ($i = 1, 2, \ldots, n$) дает

$$
\begin{aligned}
&a_{i1}\frac{\Delta_1}{\Delta} + a_{i2}\frac{\Delta_2}{\Delta} + \ldots + a_{in}\frac{\Delta_n}{\Delta} = \\
&\quad = \frac{1}{\Delta}\Big[a_{i1}\big(b_1A_{11} + b_2A_{21} + \ldots + b_iA_{i1} + \ldots + b_nA_{n1}\big) + \\
&\quad + a_{i2}\big(b_1A_{12} + b_2A_{22} + \ldots + b_iA_{i2} + \ldots + b_nA_{n2}\big) + \ldots \\
&\quad \ldots + a_{in}\big(b_1A_{1n} + b_2A_{2n} + \ldots + b_iA_{in} + \ldots + b_nA_{nn}\big)\Big] = \\
&\qquad = \frac{1}{\Delta}\Big[b_1\big(a_{i1}A_{11} + a_{i2}A_{12} + \ldots + a_{in}A_{1n}\big) + \\
&\qquad + b_2\big(a_{i1}A_{21} + a_{i2}A_{22} + \ldots + a_{in}A_{2n}\big) + \ldots \\
&\qquad \ldots + b_i\big(a_{i1}A_{i1} + a_{i2}A_{i2} + \ldots + a_{in}A_{in}\big) + \ldots \\
&\quad \ldots + b_n\big(a_{i1}A_{n1} + a_{i2}A_{n2} + \ldots + a_{in}A_{nn}\big)\Big] = \frac{1}{\Delta}\cdot b_i\cdot\Delta = b_i,
\end{aligned}
$$

так как выражение, стоящее множителем при b_k, $k = 1, 2, \ldots, n$, в силу равенств (2.12) и (2.15) равно нулю при $k \neq i$ и равно Δ при $k = i$. ▶

С л е д с т в и е. Однородная система n линейных уравнений с n неизвестными, у которой определитель отличен от нуля, имеет единственное, а именно нулевое, решение.

Вычисление решения системы n линейных уравнений с n неизвестными с ненулевым определителем с помощью формул Крамера (2.19) называют ***правилом Крамера***. Формулы Крамера позволяют выразить решение крамеровской системы через коэффициенты этой системы, что бывает удобно, особенно в теоретических рассуждениях. Однако практическое применение этих формул для решения систем линейных уравнений приводит к трудоемким вычислениям, так как для системы n-го порядка требуются вычисления $n+1$ определителя n-го порядка. Например, для решения системы

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 5, \\ x_1 + 2x_2 - x_3 + 4x_4 = -2, \\ 2x_1 - 3x_2 - x_3 - 5x_4 = -2, \\ 3x_1 + x_2 + 2x_3 + 11x_4 = 0 \end{cases}$$

по формулам Крамера необходимо вычислить пять определителей четвертого порядка, а именно определителей

$$\Delta = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & -1 & 4 \\ 2 & -3 & -1 & -5 \\ 3 & 1 & 2 & 11 \end{vmatrix} = -142,$$

$$\Delta_1 = \begin{vmatrix} 5 & 1 & 1 & 1 \\ -2 & 2 & -1 & 4 \\ -2 & -3 & -1 & -5 \\ 0 & 1 & 2 & 11 \end{vmatrix} = -142, \qquad \Delta_2 = \begin{vmatrix} 1 & 5 & 1 & 1 \\ 1 & -2 & -1 & 4 \\ 2 & -2 & -1 & -5 \\ 3 & 0 & 2 & 11 \end{vmatrix} = -284,$$

$$\Delta_3 = \begin{vmatrix} 1 & 1 & 5 & 1 \\ 1 & 2 & -2 & 4 \\ 2 & -3 & -2 & -5 \\ 3 & 1 & 0 & 11 \end{vmatrix} = -426, \qquad \Delta_4 = \begin{vmatrix} 1 & 1 & 1 & 5 \\ 1 & 2 & -1 & -2 \\ 2 & -3 & -1 & -2 \\ 3 & 1 & 2 & 0 \end{vmatrix} = 142,$$

и лишь затем по формулам Крамера найти решение

$$x_1 = \frac{\Delta_1}{\Delta} = 1, \quad x_2 = \frac{\Delta_2}{\Delta} = 2, \quad x_3 = \frac{\Delta_3}{\Delta} = 3, \quad x_4 = \frac{\Delta_4}{\Delta} = -1.$$

В то же время решение этой системы методом Гаусса потребовало бы примерно столько же операций, сколько затрачено для вычисления лишь одного определителя четвертого порядка.

Упражнения

2.1. Решить методом Гаусса системы уравнений:

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

2) $\begin{cases} x_1 - 2x_2 + 3x_3 - 4x_4 = 0, \\ 2x_1 - 4x_2 + 5x_3 + 7x_4 = 0, \\ 6x_1 - 12x_2 + 17x_3 - 9x_4 = 0, \\ 7x_1 - 14x_2 + 18x_3 + 17x_4 = 0; \end{cases}$

3) $\begin{cases} 3x_1 + 4x_2 + 5x_3 + 7x_4 = 1, \\ 2x_1 + 6x_2 - 3x_3 + 4x_4 = 2, \\ 4x_1 + 2x_2 + 13x_3 + 10x_4 = 0, \\ 5x_1 + 21x_3 + 13x_4 = 3; \end{cases}$

4) $\begin{cases} x_1 + 2x_2 - x_3 - 3x_4 + 4x_5 = 1, \\ 2x_1 - x_2 + 3x_3 + 2x_4 - x_5 = 2, \\ x_1 + 4x_2 + 2x_3 - 5x_4 + 3x_5 = 3, \\ x_1 + 15x_2 + 6x_3 - 19x_4 + 9x_5 = 1; \end{cases}$

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

6) $\begin{cases} 2x_1 - 3x_2 + 4x_3 - 7x_4 + 3x_5 = 1, \\ 3x_1 + x_2 - 5x_3 + 6x_4 + x_5 = 0, \\ x_1 - 7x_2 + 13x_3 - 20x_4 + 5x_5 = 2, \\ 4x_1 + 5x_2 - 14x_3 + 19x_4 - x_5 = -1, \\ 5x_1 - 13x_2 + 21x_3 - 34x_4 + 11x_5 = 4. \end{cases}$

2.2. Вычислить определители:

1) $\begin{vmatrix} 1 & 2 & 3 & 4 \\ 3 & 6 & 8 & 11 \\ 7 & 13 & 20 & 26 \\ 31 & 23 & 55 & 42 \end{vmatrix}$; 2) $\begin{vmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 5 & 6 \\ 5 & 6 & 7 & 9 \\ 31 & 23 & 55 & 42 \end{vmatrix}$;

3) $\begin{vmatrix} 3 & 4 & 5 & 6 \\ 1 & -2 & 3 & 4 \\ 5 & -1 & 2 & 4 \\ 8 & 7 & 1 & 5 \end{vmatrix}$; 4) $\begin{vmatrix} 1 & 2 & 3 & 4 \\ -2 & 1 & 5 & 6 \\ -3 & -5 & 1 & 7 \\ -4 & -6 & -7 & 1 \end{vmatrix}$;

5) $\begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{vmatrix}$; 6) $\begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{vmatrix}$.

2.3. Применяя теорему Лапласа, вычислить определители:

1) $\begin{vmatrix} 7 & 6 & 5 & 4 & 3 & 2 \\ 9 & 7 & 8 & 9 & 4 & 3 \\ 7 & 4 & 9 & 7 & 0 & 0 \\ 5 & 3 & 6 & 1 & 0 & 0 \\ 0 & 0 & 5 & 6 & 0 & 0 \\ 0 & 0 & 6 & 8 & 0 & 0 \end{vmatrix}$; 2) $\begin{vmatrix} 1 & 2 & 0 & 0 & 0 & 0 \\ 3 & 4 & 0 & 0 & 0 & 0 \\ 7 & 6 & 5 & 4 & 0 & 0 \\ 2 & 3 & 4 & 5 & 0 & 0 \\ 5 & 1 & 2 & 6 & 7 & 3 \\ 2 & 7 & 5 & 3 & 4 & 1 \end{vmatrix}$;

3) $\begin{vmatrix} 5 & -5 & -3 & 4 & 2 \\ -4 & 4 & 3 & 6 & 3 \\ 3 & -1 & 5 & -9 & -5 \\ -7 & 7 & 6 & 8 & 4 \\ 5 & -3 & 2 & -1 & -2 \end{vmatrix}$; 4) $\begin{vmatrix} 2 & 1 & 2 & 3 & 2 \\ 3 & -2 & 7 & 5 & -1 \\ 3 & -1 & -5 & -3 & -2 \\ 5 & -6 & 4 & 2 & -4 \\ 2 & -3 & 3 & 1 & -2 \end{vmatrix}$.

2.4. Представить произведения определителей в виде одного определителя:

1) $\begin{vmatrix} 3 & 4 \\ 5 & 6 \end{vmatrix} \cdot \begin{vmatrix} 1 & 2 & 3 \\ 3 & 1 & 4 \\ 2 & 3 & 1 \end{vmatrix}$; 2) $\begin{vmatrix} 1 & 3 & 5 \\ 2 & 6 & 1 \\ 4 & 3 & 1 \end{vmatrix} \cdot \begin{vmatrix} 2 & 1 \\ 3 & 5 \end{vmatrix}$;

3) $\begin{vmatrix} 2 & 3 \\ 1 & 4 \end{vmatrix} \cdot \begin{vmatrix} 1 & 2 & 1 \\ 2 & 3 & 4 \\ 1 & 2 & 5 \end{vmatrix}$.

2.5. По формулам Крамера решить системы:

1) $\begin{cases} x_1 + 2x_2 + 3x_3 - 2x_4 = 6, \\ 2x_1 - x_2 - 2x_3 - 3x_4 = 8, \\ 3x_1 + 2x_2 - x_3 + 2x_4 = 4, \\ 2x_1 - 3x_2 + 2x_3 + x_4 = -8; \end{cases}$

2) $\begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 = 5, \\ 2x_1 + x_2 + 2x_3 + 3x_4 = 1, \\ 3x_1 + 2x_2 + x_3 + 2x_4 = 1, \\ 4x_1 + 3x_2 + 2x_3 + x_4 = -5; \end{cases}$

3) $\begin{cases} x_2 - 3x_3 + 4x_4 = -5, \\ x_1 - 2x_3 + 3x_4 = -4, \\ 3x_1 + 2x_2 - 5x_4 = 12, \\ 4x_1 + 3x_2 - 5x_3 = 5; \end{cases}$

4) $\begin{cases} 2x_1 - x_2 + 3x_3 + 2x_4 = 4, \\ 3x_1 + 3x_2 + 3x_3 + 2x_4 = 6, \\ 3x_1 - x_2 - x_3 + 2x_4 = 6, \\ 3x_1 - x_2 + 3x_3 - x_4 = 6; \end{cases}$

5) $\begin{cases} 3x_1 + 3x_2 + 4x_3 - 5x_4 = 9, \\ 5x_1 - 7x_2 + 8x_3 + 2x_4 = 18, \\ 4x_1 + 5x_2 - 7x_3 - 3x_4 = -5, \\ 7x_1 + 8x_2 + 3x_3 + 4x_4 = -2; \end{cases}$

6) $\begin{cases} 3x_1 + 5x_2 - 3x_3 + 2x_4 = 12, \\ 4x_1 - 2x_2 + 5x_3 + 3x_4 = 27, \\ 7x_1 + 8x_2 - x_3 + 5x_4 = 40, \\ 6x_1 + 4x_2 + 5x_3 + 3x_4 = 41. \end{cases}$

Глава 3 МАТРИЦЫ И ДЕЙСТВИЯ НАД НИМИ

3.1. Первоначальные сведения о матрицах

Прямоугольной, или ***$(m \times n)$-матрицей*** называют совокупность mn чисел a_{ij}, расположенных в виде прямоугольной таблицы

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

содержащей m строк и n столбцов. Кратко матрицу A записывают в виде $A = \left(a_{ij}\right)$, $i = 1, 2, \dots, m$; $j = 1, 2, \dots, n$. Числа a_{ij}, составляющие данную матрицу, называют ее ***элементами***. Первый индекс у элемента указывает номер строки, второй — номер столбца, на пересечении которых находится этот элемент в матрице.

Матрицу называют ***комплексной***, если хотя бы один ее элемент является комплексным числом, и ***действительной*** (***вещественной***), если все ее элементы — действительные (вещественные) числа.

Две матрицы одинакового размера $m \times n$ считают ***равными***, если попарно равны их соответствующие элементы, т.е. элементы, стоящие на одинаковых местах в этих матрицах.

Матрицу, состоящую из одной строки или одного столбца, называют соответственно ***строкой*** или ***столбцом***. Элементы строк (столбцов) называют их ***компонентами***. Матрица, состоящая из одного числа, отождествляется с этим числом, поэтому любое число можно рассматривать как матрицу размера 1×1. Матрицу, состоящую из нулей, называют ***нулевой*** и обозначают символом 0.

Если число m строк матрицы равно числу n ее столбцов, то матрицу называют ***квадратной*** порядка n (n-го порядка). В квадратной матрице совокупность элементов на линии, соединяющей верхний левый угол с правым нижним, называют ***главной диагональю***. У элементов главной диагонали номер строки совпадает с номером столбца.

Квадратные матрицы, у которых отличны от нуля лишь элементы главной диагонали, называют ***диагональными***. Диагональная матрица, у которой все элементы главной диагонали одинаковые, называется ***скалярной***. Частным случаем скалярных матриц является ***единичная матрица***

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

Квадратную матрицу называют ***треугольной***, если все ее элементы, стоящие выше (или ниже) главной диагонали, равны нулю. При этом матрицу вида

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ 0 & a_{12} & \ldots & a_{2n} \\ \multicolumn{4}{c}{\dotfill} \\ 0 & 0 & \ldots & a_{nn} \end{pmatrix}$$

называют ***правой*** (или ***верхней***) ***треугольной***, а матрицу вида

$$B = \begin{pmatrix} b_{11} & 0 & \ldots & 0 \\ b_{21} & b_{22} & \ldots & 0 \\ \multicolumn{4}{c}{\dotfill} \\ b_{n1} & b_{n2} & \ldots & b_{nn} \end{pmatrix}$$

— ***левой*** (или ***нижней***) ***треугольной***. Аналогично, рассмотрев побочную диагональ (совокупность элементов на линии, соединяющей правый верхний угол с левым нижним), можно ввести понятия ***левой нижней*** и ***правой верхней*** матриц.

Транспонированием матрицы A называют операцию, при которой в матрице строки заменены на соответствующие столбцы. Полученную в результате этой операции матрицу называют ***транспонированной*** к матрице A и обозначают через A' или A^{T}. Если A — матрица размера $m \times n$, то A^{T} — матрица размера $n \times m$. При транспонировании строки получается столбец, и, наоборот, при транспонировании столбца получается строка:

$$(\alpha_1, \alpha_2, \ldots, \alpha_n)^{\mathrm{T}} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \ldots \\ \alpha_n \end{pmatrix}, \qquad \begin{pmatrix} \beta_1 \\ \beta_2 \\ \ldots \\ \beta_n \end{pmatrix}^{\mathrm{T}} = (\beta_1, \beta_2, \ldots, \beta_n).$$

В дальнейшем мы будем этим пользоваться, записывая столбцы как транспонированные строки.

Квадратную матрицу $A = \left(a_{ij}\right)$, $i, j = 1, 2, \ldots, n$, называют ***симметрической***, если в ней элементы, расположенные симметрично относительно главной диагонали, равны, т.е. если $a_{ij} = a_{ji}$, $i, j = 1, 2, \ldots, n$. Для симметрических матриц, и только для них, верно тождество $A^{\mathrm{T}} = A$.

Квадратную матрицу $A = \left(a_{ij}\right)$, $i, j = 1, 2, \ldots, n$, называют ***кососимметрической***, если элементы, расположенные в ней симметрично относительно главной диагонали, равны по абсолютной величине и противоположны по знаку, т.е. если $a_{ij} = -a_{ji}$, $i, j = 1, 2, \ldots, n$. Для кососимметрических матриц, и только для них, верно тождество $A^{\mathrm{T}} = -A$. Непосредственно из определения следует, что на главной диагонали кососимметрической матрицы все элементы равны нулю.

Комплексно сопряженной с матрицей A называют матрицу $\overline{A}$, которая получается из матрицы A заменой в ней элементов на комплексно сопряженные, т.е. $\overline{A} = \overline{(a_{ij})} = (\overline{a_{ij}})$. Матрицу $A^* = \overline{A}^{\mathrm{T}}$ называют ***эрмитово-сопряженной*** с матрицей A. Квадратную матрицу A называют ***эрмитовой***, если в ней элементы, симметричные относительно главной диагонали, являются комплексно сопряженными числами, т.е. $\overline{a_{ij}} = a_{ji}$. Для эрмитовых матриц, и только для них, выполняется равенство $A^* = A$. Элементы главной диагонали эрмитовой матрицы являются действительными числами.

Понятие минора, введенное для определителей, без изменений переносится на матрицы.

3.2. Сложение матриц и умножение матрицы на число

Произведением матрицы $A = (a_{ij})$ ***на число*** λ называют матрицу

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

элементы которой получаются из соответствующих элементов матрицы A умножением на число λ.

Отметим, что при умножении квадратной матрицы A порядка n на число λ соответствующий определитель (определитель матрицы A) умножается на число λ^n, т.е. $|\lambda A| = \lambda^n |A|$.

Матрицу $(-1)A$ называют противоположной матрице A и обозначают через $-A$.

Суммой матриц $A = (a_{ij})$ и $B = (b_{ij})$ одинакового размера $m \times n$ называют матрицу $C = (c_{ij})$ того же размера, элементы которой равны суммам $a_{ij} + b_{ij}$ соответствующих элементов матриц A и B, т.е.

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

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

Операцию нахождения суммы матриц называют ***сложением матриц***. Эта операция распространяется на случай любого конечного числа матриц одинакового размера.

Разность матриц определяется аналогично:

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

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

Операцию нахождения разности матриц называют ***вычитанием матриц***.

Сложение матриц и умножение матрицы на число обладают следующими свойствами:

1) $A+B=B+A$,
2) $A+(B+C)=(A+B)+C$,
3) $A+0=A$,
4) $A+(-A)=0$,
5) $1\cdot A=A$,
6) $(\alpha+\beta)A=\alpha A+\beta A$,
7) $\alpha(A+B)=\alpha A+\alpha B$,
8) $\alpha(\beta A)=(\alpha\beta)A$

для любых матриц A, B, C и любых чисел α, β.

3.3. Линейные комбинации столбцов (строк)

Столбцы (или строки) являются частными случаями матриц, поэтому их можно сравнивать, складывать и умножать на числа. Рассуждения будем проводить для столбцов. Для строк рассуждения аналогичны.

Два столбца

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{и} \quad y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

одинаковой высоты n считаются равными, если равны их соответствующие компоненты, т.е. если $x_i = y_i$, $i = 1, 2, \ldots, n$.

Суммой столбцов x одинаковой высоты называют столбец z той же высоты, компоненты которого равны суммам соответствующих компонент слагаемых столбцов, т.е.

$$x + y = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix} = z.$$

Произведением столбца x ***на число*** λ называют столбец, компонентами которого являются соответствующие компоненты столбца x, умноженные на число λ, т.е.

$$\lambda x = \lambda \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_n \end{pmatrix}.$$

Пусть даны столбцы

$$a_1 = \begin{pmatrix} \alpha_{11} \\ \alpha_{21} \\ \vdots \\ \alpha_{n1} \end{pmatrix}, \quad a_2 = \begin{pmatrix} \alpha_{12} \\ \alpha_{22} \\ \vdots \\ \alpha_{n2} \end{pmatrix}, \quad \ldots, \quad a_k = \begin{pmatrix} \alpha_{1k} \\ \alpha_{2k} \\ \vdots \\ \alpha_{nk} \end{pmatrix}, \quad b = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix}.$$

Столбец b называют ***линейной комбинацией столбцов*** a_1, a_2, ..., a_k, если существует набор чисел λ_1, λ_2, ..., λ_k, для которых выполняется равенство

$$b = \lambda_1 a_1 + \lambda_2 a_2 + \ldots + \lambda_k a_k. \qquad (3.1)$$

В этом случае говорят также, что столбец b линейно выражается через столбцы a_1, a_2, ..., a_k. Числа λ_1, λ_2, ..., λ_k называют ***коэффициентами*** рассматриваемой ***линейной комбинации*** столбцов. Равенство (3.1) означает, что компоненты рассматриваемых столбцов связаны линейными соотношениями

$$\begin{aligned} \beta_1 &= \lambda_1\alpha_{11} + \lambda_2\alpha_{12} + \ldots + \lambda_k\alpha_{1k}, \\ \beta_2 &= \lambda_1\alpha_{21} + \lambda_2\alpha_{22} + \ldots + \lambda_k\alpha_{2k}, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \beta_n &= \lambda_1\alpha_{n1} + \lambda_2\alpha_{n2} + \ldots + \lambda_k\alpha_{nk}, \end{aligned} \tag{3.2}$$

т.е. компоненты столбца b являются линейными комбинациями соответствующих компонентов столбцов a_1, a_2, ..., a_k с теми же коэффициентами.

Для того чтобы найти линейное выражение заданного столбца b через столбцы a_1, a_2, ..., a_k, необходимо записать равенство (3.1) с неизвестными коэффициентами λ_1, λ_2, ..., λ_k и от него перейти к покомпонентной записи (3.2). В результате получится система линейных уравнений с неизвестными λ_1, λ_2, ..., λ_k, любое решение которой и даст значения коэффициентов линейного выражения столбца b через столбцы a_1, a_2, ..., a_k.

Пример 3.1. Найти линейное выражение столбца $b = (1,\ -1,\ 4,\ -1)^{\mathrm{T}}$ через два столбца $a_1 = (1,\ -1,\ 2,\ 1)^{\mathrm{T}}$ и $a_2 = (1,\ -1,\ 1,\ 2)^{\mathrm{T}}$.

Решение. Составим уравнение $b = \lambda_1 a_1 + \lambda_2 a_2$ с неизвестными λ_1 и λ_2, т.е. уравнение

$$b = \begin{pmatrix} 1 \\ -1 \\ 4 \\ -1 \end{pmatrix} = \lambda_1 \begin{pmatrix} 1 \\ -1 \\ 2 \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ -1 \\ 1 \\ 2 \end{pmatrix},$$

и запишем его в покомпонентной форме. Тогда получим систему линейных уравнений

$$\begin{cases} 1 = \lambda_1 + \lambda_2, \\ -1 = -\lambda_1 - \lambda_2, \\ 4 = 2\lambda_1 + \lambda_2, \\ -1 = \lambda_1 + 2\lambda_2. \end{cases}$$

Из этой системы находим, что $\lambda_1 = 3$, $\lambda_2 = -2$, поэтому $b = 3a_1 - 2a_2$.

Замечание. Если столбец b линейно выражается через столбцы $a_1, a_2, \ldots, a_k$, то он будет линейно выражаться через любую систему столбцов, содержащую столбцы $a_1, a_2, \ldots, a_k$. Действительно, если

$$b = \alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_k a_k,$$

то для любой системы столбцов $a_1, a_2, \ldots, a_k, a_{k+1}, \ldots, a_r$ верно равенство

$$b = \alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_k a_k + 0 \cdot a_{k+1} + \ldots + 0 \cdot a_r.$$

Если для системы линейных уравнений

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{cases}$$

ввести столбцы коэффициентов при отдельных переменных, то эту систему можно записать в виде

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} x_1 + \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} x_2 + \ldots + \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} x_n = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Такая запись позволяет интерпретировать систему линейных уравнений как представление столбца свободных членов системы в виде линейной комбинации столбцов матрицы системы, коэффициентами которой являются неизвестные системы.

3.4. Умножение матриц

Умножение матриц определяется лишь для случая, когда число столбцов первого множителя равно числу строк второго множителя. Пусть даны матрицы

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \quad \text{и} \quad B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1p} \\ b_{21} & b_{22} & \dots & b_{2p} \\ \dots & \dots & \dots & \dots \\ b_{n1} & b_{n2} & \dots & b_{np} \end{pmatrix}.$$

Произведением матриц A и B, взятых в указанном порядке, называют матрицу C, элементы c_{ij} которой определяются по следующему правилу:

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

Другими словами, элемент c_{ij} матрицы C равен сумме произведений элементов i-й строки матрицы A на соответствующие элементы j-го столбца матрицы B.

Произведение матриц A и B, взятых в указанном порядке, обозначают через AB. Операцию нахождения произведения матриц называют ***умножением матриц***. Произведение прямоугольных матриц есть прямоугольная матрица, число строк которой равно числу строк первого множителя, а число столбцов — числу столбцов второго множителя. Например,

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} a & d \\ b & e \\ c & f \end{pmatrix} = \begin{pmatrix} 1\cdot a + 2\cdot b + 3\cdot c & 1\cdot d + 2\cdot e + 3\cdot f \\ 4\cdot a + 5\cdot b + 6\cdot c & 4\cdot d + 5\cdot e + 6\cdot f \\ 7\cdot a + 8\cdot b + 9\cdot c & 7\cdot d + 8\cdot e + 9\cdot f \end{pmatrix},$$

$$\begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = (1\cdot 4 + 2\cdot 5 + 3\cdot 6) = (32),$$

$$\begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 4\cdot 1 & 4\cdot 2 & 4\cdot 3 \\ 5\cdot 1 & 5\cdot 2 & 5\cdot 3 \\ 6\cdot 1 & 6\cdot 2 & 6\cdot 3 \end{pmatrix} = \begin{pmatrix} 4 & 8 & 12 \\ 5 & 10 & 15 \\ 6 & 12 & 18 \end{pmatrix}.$$

Возможна ситуация, когда произведение матриц A и B определено, а произведение тех же матриц B и A, взятых в другом порядке, не определено. Так, для матриц

$$A = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} \quad \text{и} \quad B = \begin{pmatrix} 5 & 6 & 7 \end{pmatrix}$$

определено произведение AB, а произведение BA не определено.

Оба произведения AB и BA определены, если A — матрица размера $m \times n$, а B — матрица размера $n \times m$. В частности, это условие выполняется для квадратных матриц одного порядка.

Из приведенных примеров видно, что если даже оба произведения AB и BA имеют смысл, то эти произведения могут оказаться неодинаковыми, т.е. умножение матриц не обладает свойством коммутативности. Если произведения AB и BA определены и $AB = BA$, то матрицы A и B называют ***перестановочными***. Например, скалярная матрица n-го порядка перестановочна со всеми квадратными матрицами того же порядка, причем единичная матрица n-го порядка играет среди квадратных матриц n-го порядка ту же роль, что и единица при умножении чисел, т.е. выполняются равенства $EA = AE = A$.

Непосредственной проверкой можно доказать, что умножение матриц обладает следующими свойствами:

1) $A(BC) = (AB)C$;

2) $\alpha(AB) = (\alpha A)B = A(\alpha B)$;

3) $C(A+B) = CA + CB$;

4) $(A+B)C = AC + BC$.

Докажем, например, первое равенство. Пусть $A = (a_{ij})$, $B = (b_{jk})$, $C = (c_{kl})$. По определению произведения матриц элементами произведений $U = AB$ и $V = BC$ будут элементы

$$u_{ik} = \sum_j a_{ij}b_{jk} \qquad \text{и} \qquad v_{jl} = \sum_k b_{jk}c_{kl},$$

а элементами двойных произведений $S=(AB)C$ и $T=A(BC)$ — соответственно элементы

$$s_{il} = \sum_k u_{ik}c_{kl} = \sum_k \Big(\sum_j a_{ij}b_{jk}\Big) b_{kl} = \sum_k \sum_j a_{ij}b_{jk}c_{kl}$$

и

$$t_{il} = \sum_j a_{ij}v_{jl} = \sum_j a_{ij} \Big(\sum_k b_{jk}c_{kl}\Big) = \sum_j \sum_k a_{ij}b_{jk}c_{kl}.$$

Таким образом, соответствующие элементы матриц $(AB)C$ и $A(BC)$ равны. Следовательно, сами эти матрицы равны.

Непосредственной проверкой можно также доказать следующие свойства:

1) $\overline{A+B} = \overline{A} + \overline{B}$,
2) $(A+B)^{\mathrm{T}} = A^{\mathrm{T}} + B^{\mathrm{T}}$,
3) $(\alpha A)^{\mathrm{T}} = \alpha A^{\mathrm{T}}$,
4) $(AB)^{\mathrm{T}} = B^{\mathrm{T}} A^{\mathrm{T}}$,
5) $(A^{\mathrm{T}})^{\mathrm{T}} = A$,
6) $\overline{\alpha A} = \overline{\alpha}\,\overline{A}$,
7) $\overline{AB} = \overline{A}\,\overline{B}$,
8) $(A+B)^* = A^* + B^*$,
9) $(\alpha A)^* = \overline{\alpha} A^*$,
10) $(AB)^* = B^* A^*$,
11) $(A^*)^* = A$.

В этих равенствах $\overline{A}$ — матрица, комплексно сопряженная с матрицей A; A^{T} — матрица, транспонированная к матрице A; A^* — матрица, эрмитово сопряженная с матрицей A.

Произведение матриц позволяет систему линейных уравнений

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{cases}$$

записать в матричной форме

$$Ax = b,$$

где

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, \qquad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \qquad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Важное значение имеет следующая теорема.

Теорема 3.1. *Определитель произведения конечного числа матриц n-го порядка равен произведению определителей этих матриц.*

$\triangleright$ Утверждение теоремы достаточно доказать для случая двух матриц $A = (a_{ij})$ и $B = (b_{ij})$.

Рассмотрим вспомогательный определитель

$$\Delta = \left| \begin{array}{cccc|cccc} a_{11} & a_{12} & \dots & a_{1n} & 0 & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 & 0 & \dots & 0 \\ \hline -1 & 0 & \dots & 0 & a_{11} & a_{12} & \dots & a_{1n} \\ 0 & -1 & \dots & 0 & a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & -1 & a_{n1} & a_{n2} & \dots & a_{nn} \end{array} \right|.$$

Разлагая этот определитель с помощью теоремы Лагранжа по первым n строкам, получим равенство $\Delta = |A|\,|B|$. Покажем далее, что $\Delta = |AB|$. Для этого преобразуем определитель следующим образом. Сначала первые n столбцов, умноженных соответственно на $b_{11}, b_{21}, \dots, b_{n1}$, прибавим к $(n+1)$-му столбцу. Затем первые n столбцов, умноженных соответственно на $b_{12}, b_{22}, \dots, b_{n2}$, прибавим к $(n+2)$-му столбц и т.д. На последнем шаге к $(2n)$-му столбцу будут прибавлены первые n столбцов, умноженных соответственно на b_{1n}, $b_{2n}, \ldots, b_{nn}$. В результате получим определитель

$$\Delta = \left|\begin{array}{cccc|cccc} a_{11} & a_{12} & \ldots & a_{1n} & c_{11} & c_{12} & \ldots & c_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} & c_{21} & c_{22} & \ldots & c_{2n} \\ \multicolumn{4}{c|}{\dotfill} & \multicolumn{4}{c}{\dotfill} \\ a_{n1} & a_{n2} & \ldots & a_{nn} & c_{n1} & c_{n2} & \ldots & c_{nn} \\ \hline -1 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 \\ 0 & -1 & \ldots & 0 & 0 & 0 & \ldots & 0 \\ \multicolumn{4}{c|}{\dotfill} & \multicolumn{4}{c}{\dotfill} \\ 0 & 0 & \ldots & -1 & 0 & 0 & \ldots & 0 \end{array}\right|,$$

в котором

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj}, \quad i, j = 1, 2, \ldots, n,$$

т.е. c_{ij} — элементы матрицы $C = AB$.

Разлагая полученный определитель с помощью теоремы Лапласа по последним n столбцам, находим:

$$\Delta = |C| \cdot (-1)^{(1+2+\ldots+n)+[(n+1)+(n+2)+\ldots+2n]} \begin{vmatrix} -1 & & 0 \\ & \ddots & \\ 0 & & -1 \end{vmatrix} =$$

$$= (-1)^{(2n+1)n} \cdot (-1)^n |C| = (-1)^{2n(n+1)} |C| = |C| = |AB|.$$

Итак, доказаны равенства $\Delta = |A|\,|B|$ и $\Delta = |AB|$, из которых следует, что $|AB| = |A|\,|B|$. ►

Квадратную матрицу называют ***невырожденной***, или ***неособенной***, если ее определитель отличен от нуля, и ***вырожденной***, или ***особенной***, если ее определитель равен нулю. Из доказанной теоремы следует, что произведение нескольких квадратных матриц является невырожденной матрицей тогда и только тогда, когда все сомножители являются невырожденными матрицами.

3.5. Элементарные преобразования над матрицами и элементарные матрицы

Под ***элементарными преобразованиями*** над $(m \times n)$-матрицей понимают следующие действия.

1. Умножение любой строки матрицы на любое ненулевое число.

2. Прибавление к любой i-й строке матрицы любой ее j-й строки, умноженной на произвольное число α.

3. Умножение любого столбца матрицы на любое ненулевое число.

4. Прибавление к любому i-му столбцу матрицы любого ее j-го столбца, умноженного на произвольное число α.

5. Перестановку строк (столбцов) матрицы.

Матрицы A и B будем называть ***эквивалентными***, если одна из них может быть преобразована в другую с помощью элементарных преобразований, и будем писать $A \sim B$. Эквивалентность матриц обладает следующими свойствами:

1) рефлексивностью, т.е. $A \sim A$;

2) симметричностью, т.е. если $A \sim B$, то $B \sim A$;

3) транзитивностью, т.е. если $A \sim B$, $B \sim C$, то $A \sim C$.

Любую $(m \times n)$-матрицу элементарными преобразованиями над строками и столбцами можно привести к виду

$$\begin{pmatrix} 1 & & & & & \\ & \ddots & & & \text{\Large 0} & \\ & & 1 & & & \\ & & & 0 & & \\ & \text{\Large 0} & & & \ddots & \\ & & & & & 0 \end{pmatrix}.$$

В частности, невырожденную квадратную матрицу таким способом можно привести к единичной матрице.

Непосредственная проверка позволяет установить, что первое и второе элементарные преобразования над $(m \times n)$-матрицей равносильны умножению этой матрицы слева на ***элементарные матрицы***

$$
\begin{array}{cc}
 & \begin{array}{ccccccc} & & & i & & & \end{array} \\
\begin{array}{c} \\ \\ \\ i \\ \\ \\ \\ \end{array} &
\begin{pmatrix}
1 & & & & & & \\
 & \ddots & & & & 0 & \\
 & & 1 & & & & \\
 & & & \alpha & & & \\
 & & & & 1 & & \\
 & 0 & & & & \ddots & \\
 & & & & & & 1
\end{pmatrix}
\end{array},
\qquad
\begin{array}{cc}
 & \begin{array}{ccccccc} & & j & & & & \end{array} \\
\begin{array}{c} \\ \\ \\ \\ i \\ \\ \\ \end{array} &
\begin{pmatrix}
1 & & & & & & \\
 & \ddots & & & & 0 & \\
 & & 1 & & & & \\
 & & & \ddots & & & \\
 & & \alpha & & 1 & & \\
 & 0 & & & & \ddots & \\
 & & & & & & 1
\end{pmatrix}
\end{array},
$$

которые получаются такими же преобразованиями над строками единичной матрицы m-го порядка. Третье и четвертое преобразования над $(m \times n)$-матрицей равносильны умножению этой матрицы справа на элементарные матрицы n-го порядка, которые получаются с помощью тех же преобразований над столбцами единичной матрицы n-го порядка.

Перестановка строк $(m \times n)$-матрицы равносильна ее умножению слева на матрицу P, получаемую из единичной матрицы m-го порядка той же перестановкой строк. Аналогично перестановка столбцов $(m \times n)$-матрицы равносильна ее умножению справа на матрицу Q, получаемую из единичной матрицы n-го порядка той же перестановкой столбцов. Матрицы P и Q называют ***матрицами перестановок***.

Если над строками $(m \times n)$-матрицы A последовательно проведено k элементарных преобразований, которым соответствуют элементарные матрицы R_1, R_2, ..., R_k, то в результате получается матрица $R_k R_{k-1} \ldots R_1 A$. Очевидно, что матрица $R = R_k R_{k-1} \ldots R_1 = R_k R_{k-1} \ldots R_1 E$ получается теми же элементарными преобразованиями и в том же порядке над строками единичной матрицы m-го порядка. Аналогично если над столбцами $(m \times n)$-матрицы A последовательно проведено l элементарных преобразований, которым соответствуют элементарные матрицы S_1, S_2, ..., S_l, то в результате получается матрица $AS_1 S_2 \ldots S_l$. Результирующая матрица преобразований $S = S_1 S_2 \ldots S_l = ES_1 S_2 \ldots S_l$ получается теми же преобразованиями в том же порядке над столбцами единичной матрицы n-го порядка.

Например, если в (3×2)-матрице

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}$$

первую строку, умноженную на α, прибавим ко второй строке, затем в полученной матрице вторую строку, умноженную на β, прибавим к третьей строке, то получим следующую цепочку эквивалентных матриц:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \sim \begin{pmatrix} a_{11} & a_{12} \\ \alpha a_{11} + a_{21} & \alpha a_{12} + a_{22} \\ a_{31} & a_{32} \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} a_{11} & a_{12} \\ \alpha a_{11} + a_{21} & \alpha a_{12} + a_{22} \\ \beta(\alpha a_{11} + a_{21}) + a_{31} & \beta(\alpha a_{12} + a_{22}) + a_{32} \end{pmatrix} = B,$$

что с применением элементарных матриц записывается так:

$$A \sim R_2 R_1 A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \beta & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ \alpha & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} =$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \beta & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ \alpha a_{11} + a_{21} & \alpha a_{12} + a_{22} \\ a_{31} & a_{32} \end{pmatrix} =$$

$$= \begin{pmatrix} a_{11} & a_{12} \\ \alpha a_{11} + a_{21} & \alpha a_{12} + a_{22} \\ \beta(\alpha a_{11} + a_{21}) + a_{31} & \beta(\alpha a_{12} + a_{22}) a_{32} \end{pmatrix} = B.$$

Если над строками матрицы A проведены преобразования, которым соответствуют элеметарные матрицы $R_1, R_2, \ldots, R_k$, а над столбцами — преобразования, которым соответствуют матрицы $S_1, S_2, \ldots, S_l$, то получится матрица $B = RAS$, где $R = R_k R_{k-1} \ldots R_1$ и $S = S_1 S_2 \ldots S_l$. Поэтому мы приходим к следующему второму определению эквивалентных матриц.

Матрицы A и B называют ***эквивалентными***, если они связаны соотношением $B = RAS$ при некоторых невырожденных матрицах R и S.

3.6. Обратная матрица

Пусть дана квадратная матрица A n-го порядка. Квадратную матрицу A^{-1} того же порядка называют ***обратной*** к матрице A, если

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

где E — единичная матрица n-го порядка.

В соответствии с этим определением матрица A является обратной к матрице A^{-1}, поэтому можно говорить, что матрицы A и A^{-1} взаимно обратны.

Если существует матрица, обратная к матрице A, то такая матрица единственная.

Допустим, что матрица A^{-1} является не неединственной матрицей, обратной к матрице A. Возьмем другую обратную матрицу B. Тогда выполняются условия

$$AA^{-1} = A^{-1}A = E, \quad AB = BA = E.$$

Рассмотрим произведение BAA^{-1}. Для него имеют место равенства

$$BAA^{-1} = B(AA^{-1}) = BE = B,$$

$$BAA^{-1} = (BA)A^{-1}) = EA^{-1} = A^{-1},$$

из которых вытекает, что $B = A^{-1}$. Тем самым единственность обратной матрицы доказана.

При доказательстве теоремы о существовании обратной матрицы нам потребуется понятие „присоединенная матрица“. Пусть дана матрица

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

Матрица

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

элементами которой являются алгебраические дополнения A_{ij} элементов a_{ij} матрицы A, называется ***присоединенной*** (или ***союзной***) матрицей к матрице A. Обратим внимание на то, что для построения присоединенной матрицы C элементы матрицы C нужно заменить их алгебраическими дополнениями, а затем полученную матрицу транспонировать.

Теорема 3.2. *Для того чтобы матрица A имела обратную матрицу A^{-1}, необходимо и достаточно, чтобы матрица A была невырожденной. При этом матрица A^{-1} определяется формулой*

$$A^{-1} = \frac{1}{|A|} \begin{pmatrix} A_{11} & A_{21} & \dots & a_{n1} \\ A_{12} & A_{22} & \dots & a_{n2} \\ \dots & \dots & \dots & \dots \\ A_{1m} & A_{2m} & \dots & a_{nn} \end{pmatrix}, \qquad (3.3)$$

где A_{ij} — алгебраические дополнения элементов матрицы A.

$\triangleright$ Пусть матрица A имеет обратную матрицу A^{-1}. Тогда выполняются условия $AA^{-1} = A^{-1}A = E$, из которых по теореме 3.1 получаем равенства $|A|\,|A^{-1}| = |A^{-1}|\,|A| = 1$. Отсюда следует, что определители $|A|$ и $|A^{-1}|$ матриц A и A^{-1} отличны от нуля и связаны соотношением

$$|A^{-1}| = \frac{1}{|A|}. \qquad (3.4)$$

Матрицы A и A^{-1} невырожденные, поскольку их определители отличны от нуля.

Пусть теперь матрица A невырожденная. Докажем, что матрица A имеет обратную матрицу A^{-1} и она определяется формулой (3.3). Для этого рассмотрим произведение

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

матрицы A и присоединенной к ней матрицы C.

По правилу умножения матриц элемент u_{ij} произведения AC матриц A и C имеет вид:

$$u_{ij} = a_{i1}A_{j1} + a_{i2}A_{j2} + \dots + a_{in}A_{jn}, \quad i, j = 1, 2, \dots, n,$$

но согласно формулам (2.11) и (2.14) сумма произведений элементов i-й строки на алгебраические дополнения соответствующих элементов j-й строки равна нулю при $i \neq j$ и определителю при $i = j$. Следовательно,

$$u_{ij} = \begin{cases} |A| & \text{при } i = j, \\ 0 & \text{при } i \neq j. \end{cases}$$

Поэтому

$$AC = \begin{pmatrix} |A| & & 0 \\ & \ddots & \\ 0 & & |A| \end{pmatrix} = |A| \cdot E,$$

где E — единичная матрица n-го порядка. Аналогично доказывается равенство $CA = |A| \cdot E$. Таким образом,

$$AC = |A| \cdot E, \qquad CA = |A| \cdot E, \tag{3.5}$$

а это означает, что

$$A \cdot \frac{1}{|A|}C = \frac{1}{|A|}C \cdot A = E$$

и матрица $\frac{1}{|A|} \cdot C$ является обратной к матрице A. Следовательно, невырожденная матрица A имеет обратную матрицу, которая определяется формулой (3.3). ▶

Следствие 3.1. Определители матриц A и A^{-1} связаны соотношением (3.4).

Следствие 3.2. Основное свойство присоединенной матрицы C к матрице A выражается равенствами (3.5).

Следствие 3.3. Определители невырожденной матрицы A и присоединенной к ней матрицы C связаны равенством

$$|C| = |A|^{n-1}.$$

Следствия 3.1 и 3.2 установлены в ходе доказательства теоремы. Следствие 3.3 вытекает из равенства $AC = |A| \cdot E$ и свойства определителей, согласно которому при умножении матрицы на число определитель умножается на n-ю степень этого числа. В данном случае

$$|A|\,|C| = |AC| = \bigl||A| \cdot E\bigr| = |A|^n |E| = |A|^n,$$

откуда следует, что $|C| = |A|^{n-1}$.

Обратная матрица обладает следующими свойствами:

1. $(A^{-1})^{-1} = A$;
2. $(A^{\mathrm{T}})^{-1} = (A^{-1})^{\mathrm{T}}$;
3. $(A^*)^{-1} = (A^{-1})^*$;
4. $(A_1A_2)^{-1} = A_2^{-1}A_1^{-1}$.

Для проверки, например, четвертого свойства достаточно рассмотреть произведения $(A_1A_2)(A_2^{-1}A_1^{-1})$ и $(A_2^{-1}A_1^{-1})(A_1A_2)$. Из свойств произведения матриц вытекает, что

$$(A_1A_2)(A_2^{-1}A_1^{-1}) = A_1(A_2A_2^{-1})A_1^{-1} = A_1EA_1^{-1} = A_1A_1^{-1} = E,$$

$$(A_2^{-1}A_1^{-1})(A_1A_2) = A_2^{-1}(A_1^{-1}A_1)A_2 = A_2^{-1}EA_2 = A_2^{-1}A_2 = E.$$

Для того чтобы проверить второе свойство, достаточно рассмотреть произведения $A^{\mathrm{T}}(A^{-1})^{\mathrm{T}}$ и $(A^{-1})^{\mathrm{T}}A^{\mathrm{T}}$:

$$A^{\mathrm{T}}(A^{-1})^{\mathrm{T}} = (A^{-1}A)^{\mathrm{T}} = E^{\mathrm{T}} = E,$$
$$(A^{-1})^{\mathrm{T}}A^{\mathrm{T}} = (AA^{-1})^{\mathrm{T}} = E^{\mathrm{T}} = E.$$

Формула (3.3) позволяет находить обратную матрицу. Она удобна в случае матриц небольших размеров. В частности, для матриц второго порядка обратная матрица находится по этой формуле практически без вычислений:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Пример 3.2. Найти матрицу, обратную к матрице

$$A = \begin{pmatrix} 5 & -2 & 2 \\ 3 & -2 & 3 \\ 2 & -3 & 4 \end{pmatrix}.$$

Решение. Определитель матрицы

$$|A| = \begin{vmatrix} 5 & -2 & 2 \\ 3 & -2 & 3 \\ 2 & -3 & 4 \end{vmatrix} = 7$$

отличен от нуля. Поэтому матрица A имеет обратную. Чтобы ее найти, сначала вычислим алгебраические дополнения:

$$A_{11} = (-1)^{1+1}\begin{vmatrix} -2 & 3 \\ -3 & 4 \end{vmatrix} = 1, \qquad A_{21} = (-1)^{2+1}\begin{vmatrix} -2 & 2 \\ -3 & 4 \end{vmatrix} = 2,$$

$$A_{12} = (-1)^{1+2}\begin{vmatrix} 3 & 3 \\ 2 & 4 \end{vmatrix} = -6, \qquad A_{22} = (-1)^{2+2}\begin{vmatrix} 5 & 2 \\ 2 & 4 \end{vmatrix} = 16,$$

$$A_{13} = (-1)^{1+3}\begin{vmatrix} 3 & -2 \\ 2 & -3 \end{vmatrix} = -5, \qquad A_{23} = (-1)^{2+3}\begin{vmatrix} 5 & -2 \\ 2 & -3 \end{vmatrix} = 11,$$

$$A_{31} = (-1)^{3+1}\begin{vmatrix} -2 & 2 \\ -2 & 3 \end{vmatrix} = -2, \qquad A_{32} = (-1)^{3+2}\begin{vmatrix} 5 & 2 \\ 3 & 3 \end{vmatrix} = -9,$$

$$A_{33} = (-1)^{3+3}\begin{vmatrix} 5 & -2 \\ 3 & -2 \end{vmatrix} = -4.$$

Теперь по формуле (3.3) запишем обратную матрицу

$$A^{-1} = \frac{1}{|A|}\begin{pmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{pmatrix} = \frac{1}{7}\begin{pmatrix} 1 & 2 & -2 \\ -6 & 16 & -9 \\ -5 & 11 & -4 \end{pmatrix}.$$

Для матриц больших размеров отыскание обратной матрицы удобно проводить с помощью элементарных преобразований над матрицами. Этот метод состоит в следующем. Выписывают составную матрицу $(A\,|\,E)$ и по схеме метода Гаусса выполняют над строками этой матрицы (т.е. одновременно и в матрице A, и в матрице E) элементарные преобразования. В результате матрица A преобразуется в единичную матрицу, а матрица E — в матрицу A^{-1}.

Действительно, пусть одними и теми же элементарными преобразованиями над строками матриц A и E матрица A преобразована в единичную, а матрица E — в некоторую матрицу R. Это означает (см. разд. 3.5), что матрица R является результирующей матрицей преобразований, выполнение которых равносильно умножению слева на матрицу R. Поэтому имеем равенство $RA = E$, откуда, умножая справа на матрицу A^{-1}, получаем $RAA^{-1} = A^{-1}$, или $R = A^{-1}$.

Пример 3.3. Найти матрицу, обратную к матрице

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 6 \end{pmatrix}.$$

Решение. Запишем составную матрицу $(A\,|\,E)$ и преобразуем ее с помощью элементарных преобразований строк в соответствии с методом Гаусса. В результате получим:

$$(A\,|\,E) = \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & 3 & 0 & 1 & 0 \\ 1 & 3 & 6 & 0 & 0 & 1 \end{array}\right) \sim$$

$$\sim \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & -1 & 1 & 0 \\ 0 & 2 & 5 & -1 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & -1 & 1 & 0 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array}\right) \sim$$

$$\sim \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 0 & 2 & -1 \\ 0 & 1 & 0 & -3 & 5 & -2 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 3 & -3 & 1 \\ 0 & 1 & 0 & -3 & 5 & -2 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array}\right).$$

Из этих преобразований заключаем, что

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

По схеме проведения операций рассмотренный способ совпадает со следующим. Для вычисления матрицы A^{-1}, обратной к матрице $A = (a_{ij})$, составим систему линейных уравнений

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

и методом Гаусса разрешим ее относительно переменных x_1, x_2, ..., x_n. В результате получим соотношения

$$\begin{cases} x_1 = b_{11}y_1 + b_{12}y_2 + \ldots + b_{1n}y_n, \\ x_2 = b_{21}y_1 + b_{22}y_2 + \ldots + b_{2n}y_n, \\ \dots\dots\dots\dots\dots\dots \\ x_n = b_{n1}y_1 + b_{n2}y_2 + \ldots + b_{nn}y_n. \end{cases}$$

Матрица B этой системы, составленная из коэффициентов при переменных y_1, y_2, ..., y_n, и будет обратной матрицей A^{-1}.

Для иллюстрации сходства этого способа с предыдущим найдем матрицу, обратную к матрице

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 6 \end{pmatrix}$$

из предыдущего примера. Для этого поставим систему

$$\begin{cases} x_1 + x_2 + x_3 = y_1, \\ x_1 + 2x_2 + 3x_3 = y_2, \\ x_1 + 3x_2 + 6x_3 = y_3 \end{cases}$$

и решим ее методом Гаусса, проводя его в матричной форме записи:

$$(A \mid E) = \left(\begin{array}{ccc|c} 1 & 1 & 1 & y_1 \\ 1 & 2 & 3 & y_2 \\ 1 & 3 & 6 & y_3 \end{array}\right) \sim$$

$$\sim \left(\begin{array}{ccc|l} 1 & 1 & 1 & y_1 \\ 0 & 1 & 2 & -y_1+y_2 \\ 0 & 2 & 5 & -y_1+y_3 \end{array}\right) \sim \left(\begin{array}{ccc|l} 1 & 1 & 1 & y_1 \\ 0 & 1 & 2 & -y_1+y_2 \\ 0 & 0 & 1 & y_1-2y_2+y_3 \end{array}\right) \sim$$

$$= \left(\begin{array}{ccc|l} 1 & 1 & 0 & 2y_2-y_3 \\ 0 & 1 & 0 & -3y_1+5y_2-2y_3 \\ 0 & 0 & 1 & y_1-2y_2+y_3 \end{array}\right) \sim \left(\begin{array}{ccc|l} 1 & 0 & 0 & 3y_1-3y_2+y_3 \\ 0 & 1 & 0 & -3y_1+5y_2-2y_3 \\ 0 & 0 & 1 & y_1-2y_2+y_3 \end{array}\right).$$

Отсюда:

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

В этих преобразованиях коэффициенты при переменных y_1, y_2, y_3 образуют правую часть составной матрицы, преобразование которой было проведено в примере 3.3.

В заключение заметим, что для крамеровской системы $AX = b$ решение с помощью обратной матрицы A^{-1} можно записать в виде

$$X = A^{-1}b. \tag{3.6}$$

Эта формула получается, если равенство $AX = b$ умножить слева на матрицу A^{-1}. Формула оказывается удобной во многих случаях. Например, из нее легко получаются формулы Крамера. Формула полезна в теоретических рассуждениях. Если обратная матрица A^{-1} известна, то с помощью формулы (3.6) легко решить систему $AX = b$. Например, рассмотрим систему

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

Для матрицы

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

в примере 3.2 найдена обратная матрица

$$A^{-1} = \frac{1}{7}\begin{pmatrix} 1 & 2 & -2 \\ -6 & 16 & -9 \\ -5 & 11 & -4 \end{pmatrix}.$$

Поэтому решение системы можно вычислить так:

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \frac{1}{7}\begin{pmatrix} 1 & 2 & -2 \\ -6 & 16 & -9 \\ -5 & 11 & -4 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{7}\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}.$$

3.7. Простейшие матричные уравнения

К простейшим матричным уравнениям относят уравнения

$$AX = B, \quad XC = B, \quad AXC = B, \quad AX = XB, \quad AX + XB = C,$$

в которых матрицы A, B, C заданы, а X — неизвестная матрица. Если матрицы A и C имеют обратные матрицы, то первые три уравнения можно решить по соответствующим формулам:

$$X = A^{-1}B, \quad X = BC^{-1}, \quad X = A^{-1}BC^{-1}.$$

Для решения последних двух уравнений такой метод не подходит, так как решить подобные уравнения, опираясь только на формальное преобразование матричных выражений, нельзя. Во всех пяти случаях решение нельзя получить матричными преобразованиями и тогда, когда хотя бы один из сомножителей при неизвестной матрице X является вырожденной матрицей. Приведем общий подход к решению таких уравнений.

Матрицу X выписываем поэлементно (т.е. неизвестными будут элементы матрицы X, а не матрица в целом), проводим указанные в уравнении действия над матрицами и равенство двух частей уравнения записываем поэлементно. В результате получим систему линейных уравнений, решив которую, мы найдем возможные значения элементов неизвестной матрицы X. Если система линейных уравнений окажется несовместной, то и исходное матричное уравнение не имеет решений.

Пример 3.4. Решить матричное уравнение $AX + XB = C$ при

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} -3 & -2 \\ -1 & -4 \end{pmatrix}, \quad C = \begin{pmatrix} -1 & -5 \\ 5 & 15 \end{pmatrix}.$$

Решение. Запишем матрицу X поэлементно:

$$X = \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix}.$$

Тогда в подробной записи матричное уравнение примет вид:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix} + \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix}\begin{pmatrix} -3 & -2 \\ -1 & -4 \end{pmatrix} = \begin{pmatrix} -1 & -5 \\ 5 & 15 \end{pmatrix}.$$

Вычислив произведения матриц в левой части уравнения и сложив эти произведения, придем к уравнению

$$\begin{pmatrix} -2x_1 + 2x_2 - y_1 & -2x_1 - 3y_1 + 2y_2 \\ 3x_1 + x_2 - y_2 & -2x_2 + 3y_1 \end{pmatrix} = \begin{pmatrix} -1 & -5 \\ 5 & 15 \end{pmatrix}.$$

Записывая это матричное равенство по элементам, получим следующую систему линейных уравнений:

$$\begin{cases} -2x_1 + 2x_2 - y_1 = -1, \\ 3x_1 + x_2 - y_2 = 5, \\ -2x_1 - 3y_1 + 2y_2 = -5, \\ -2x_2 + 3y_1 = 15. \end{cases}$$

Решив эту систему, найдем

$$x_1 = 2, \quad x_2 = 3, \quad y_1 = 3, \quad y_2 = 4.$$

Следовательно, искомая матрица имеет вид:

$$X = \begin{pmatrix} 2 & 3 \\ 3 & 4 \end{pmatrix}.$$

3.8. Разложение квадратной матрицы на треугольные множители

Любую квадратную матрицу

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

у которой все главные диагональные (угловые) миноры

$$a_{11}, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix},$$

отличны от нуля, можно представить в виде произведения

$$A = LU \tag{3.7}$$

левой треугольной матрицы L с единицами на главной диагонали и правой треугольной матрицы U. Разложение (3.7) матрицы A называют ее ***LU-разложением***. Оно находит широкое применение в вычислительной практике. Такое разложение можно получить, например, с помощью метода Гаусса. Поясним это на примере.

Пример 3.5. Построить LU-разложение матрицы

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

Решение. Первую строку матрицы A, умноженную на -2, прибавим ко второй ее строке, затем первую строку, умноженную на 1, прибавим к третьей строке. В результате получим матрицу

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

Вторую строку этой матрицы, умноженную на 3, прибавим к ее третьей строке. Тогда получим матрицу

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

Проведенные элементарные операции над строками равносильны умножению матрицы A слева последовательно на элементарные матрицы

$$E_{21}=\begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},\quad E_{31}=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix},\quad E_{32}=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{pmatrix}.$$

В итоге имеем: $E_{32}E_{31}E_{21}A=U$, т.е. $SA=U$, где

$$S=E_{32}E_{31}E_{21}=\begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -5 & 3 & 1 \end{pmatrix}.$$

Умножая обе части равенства $SA=U$ слева на S^{-1}, получим LU-разложение

$$A=S^{-1}U=LU=\begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & -3 & 1 \end{pmatrix}\begin{pmatrix} 2 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & -4 \end{pmatrix},$$

в котором $L=S^{-1}$.

Для вычисления неизвестных элементов матриц L и U разложения (3.7) можно также провести умножение матриц в правой части этого разложения, сравнить элементы полученной матрицы с соответствующими элементами матрицы A и решить полученную систему относительно неизвестных элементов матриц L и U. Решение в общем виде дает формулы для вычисления нужных элементов (см. [6], [7]). Детальная разработка таких вопросов относится к курсу методов вычислений. Поясним такой подход на примере.

Пример 3.6. Построить LU-разложение матрицы

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

Решение. Запишем подробно LU-разложение матрицы A пока с неизвестными элементами матриц L и U:

$$\begin{pmatrix} 1 & 2 & -1 \\ 2 & 1 & -1 \\ 1 & -7 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{pmatrix}.$$

Проведя умножение матриц в правой части этого равенства, получим:

$$\begin{pmatrix} 1 & 2 & -1 \\ 2 & 1 & -1 \\ 1 & -7 & 3 \end{pmatrix} = \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ l_{21}u_{11} & l_{21}u_{12}+u_{22} & l_{21}u_{13}+u_{23} \\ l_{31}u_{11} & l_{31}u_{12}+l_{32}u_{22} & l_{31}u_{13}+l_{32}u_{23}+u_{33} \end{pmatrix}.$$

Сравнивая соответствующие элементы этих матриц, приходим к системе уравнений

$$\begin{array}{lll} u_{11} = 1, & u_{12} = 2, & u_{13} = -1, \\ l_{21}u_{11} = 2, & l_{21}u_{12} + u_{22} = 1, & l_{21}u_{13} + u_{23} = -1, \\ l_{31}u_{11} = 1, & l_{31}u_{12} + l_{32}u_{22} = -7, & l_{31}u_{13} + l_{32}u_{23} + u_{33} = 3. \end{array}$$

Решив эту систему, найдем:

$$\begin{array}{lll} u_{11} = 1, & u_{12} = 2, & u_{13} = -1, \\ l_{21} = 2, & u_{22} = -3, & u_{23} = 1, \\ l_{31} = 1, & l_{32} = 3, & u_{33} = 1. \end{array}$$

Следовательно,

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

— искомое LU-разложение матрицы A.

Выделив главную диагональ матрицы U в разложении (3.7) в диагональную матрицу, матрицу U можно представить в виде произведения $U = DU_1$ с правой треугольной матрицей U_1 с единицами на главной диагонали и диагональной матрицей D. Тогда разложение (3.7) превратится в разложение

$$A = LDU_1. \tag{3.8}$$

Если матрица A симметрическая, то разложение (3.8) превратится в разложение

$$A = U_1^{\mathrm{T}} DU_1. \tag{3.9}$$

Если, кроме того, в матрице D все диагональные элементы $d_1, d_2, \ldots, d_n$ положительные, то можно положить $V = D^{1/2}U_1$, где $D^{1/2}$ — диагональная матрица с элементами $\sqrt{d_1}$, $\sqrt{d_2}, \ldots,$ $\sqrt{d_n}$ по диагонали. Тогда разложение (3.9) превратится в разложение

$$A = V^{\mathrm{T}}V. \tag{3.10}$$

Разложения (3.8)–(3.10), как и разложение (3.7), широко применяются в вычислительной практике.

Упражнения

3.1. Вычислить суммы матриц:

1) $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} + \begin{pmatrix} 9 & 8 & 7 \\ 6 & 5 & 4 \\ 3 & 2 & 1 \end{pmatrix}$; 2) $\begin{pmatrix} 4 & 2 & 3 \\ 3 & 3 & 3 \\ 5 & 1 & 6 \end{pmatrix} + \begin{pmatrix} 3 & 2 & 4 \\ 4 & 1 & 5 \\ 6 & 1 & 3 \end{pmatrix}$.

3.2. Вычислить разности матриц:

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

3.3. Найти произведение матриц:

1) $\begin{pmatrix} 3 & 1 & 2 \\ 2 & 1 & 3 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 2 & 2 \\ 3 & 4 \end{pmatrix}$; 2) $\begin{pmatrix} 1 & 2 & 3 \\ -2 & 1 & 4 \\ -3 & -4 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & -2 \\ 3 & 2 & -1 \\ 2 & -3 & 1 \end{pmatrix}$;

3) $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 3 & 2 & 1 \\ 2 & 3 & 5 \\ 3 & 4 & 1 \end{pmatrix}$; 4) $\begin{pmatrix} 1 & 2 & 1 \\ 1 & 3 & 1 \\ 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} 2 & 3 & 1 \\ 3 & 1 & 2 \\ 1 & 2 & 1 \end{pmatrix}$;

5) $\begin{pmatrix} 3 & 5 & -6 \\ 2 & 4 & 3 \\ 3 & 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & 8 \\ -3 & -1 & -2 \\ 4 & 5 & -3 \end{pmatrix}$; 6) $\begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} \begin{pmatrix} \mu & \nu & \delta \end{pmatrix}$;

7) $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 3 & 1 \end{pmatrix} \begin{pmatrix} 4 & 3 & 1 \\ 0 & 1 & 2 \\ 4 & 0 & 1 \end{pmatrix}$; 8) $\begin{pmatrix} 1 & 1 & 3 \\ 2 & 1 & 1 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 2 \\ -2 & 1 & 2 \\ 1 & 0 & -1 \end{pmatrix}$.

3.4. Найти обратные матрицы к следующим матрицам:

1) $\begin{pmatrix} 1 & 2 & -3 \\ 0 & 1 & 2 \\ 1 & 0 & 4 \end{pmatrix}$; 2) $\begin{pmatrix} 1 & 2 & 1 \\ 4 & 3 & -2 \\ -5 & -4 & -1 \end{pmatrix}$; 3) $\begin{pmatrix} 1 & 1 & 1 \\ 2 & -3 & 1 \\ 4 & 1 & -5 \end{pmatrix}$;

4) $\begin{pmatrix} 4 & -8 & -5 \\ -4 & 7 & -1 \\ -3 & 5 & 1 \end{pmatrix}$; 5) $\begin{pmatrix} 2 & 1 & 0 & 0 \\ 3 & 2 & 0 & 0 \\ 1 & 1 & 3 & 4 \\ 2 & -1 & 2 & 3 \end{pmatrix}$; 6) $\begin{pmatrix} 0 & 1 & 1 & 1 \\ -1 & 0 & 1 & 1 \\ -1 & -1 & 0 & 1 \\ -1 & -1 & -1 & 0 \end{pmatrix}$;

7) $\begin{pmatrix} 1 & 1 & 3 & 2 \\ 2 & 2 & 1 & 2 \\ 1 & 2 & 3 & 4 \\ 1 & 1 & 2 & 3 \end{pmatrix}$; 8) $\begin{pmatrix} 4 & 7 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 7 & -4 \\ 0 & 0 & -5 & 3 \end{pmatrix}$; 9) $\begin{pmatrix} 1 & 1 & 1 & 0 \\ -1 & 2 & 1 & 0 \\ 1 & 4 & 1 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$.

3.5. Решить систему $AX = b$ формуле $X = A^{-1}b$, если:

1) $A = \begin{pmatrix} 1 & -1 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix}, \quad b = \begin{pmatrix} 5 \\ 10 \\ 5 \end{pmatrix};$

2) $A = \begin{pmatrix} 2 & 5 & 7 \\ 6 & 3 & 4 \\ 5 & -2 & -3 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$

3.6. Решить матричные уравнения:

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

3) $\begin{pmatrix} 2 & -3 \\ 8 & -12 \end{pmatrix} X = \begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix};$ 4) $\begin{pmatrix} 3 & 2 \\ 6 & 4 \end{pmatrix} X = \begin{pmatrix} 1 & -1 \\ 2 & -2 \end{pmatrix};$

5) $X \begin{pmatrix} 5 & -3 \\ 10 & -6 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix};$ 6) $X \begin{pmatrix} 3 & -5 \\ 9 & -15 \end{pmatrix} = \begin{pmatrix} 9 & -15 \\ 6 & -10 \end{pmatrix};$

7) $\begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix} X + X \begin{pmatrix} 2 & 3 \\ 5 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$

3.7. Для матриц из упражнения 3.4 построить LU-разложение $A = LU$ и, пользуясь соотношением $A^{-1} = (LU)^{-1} = U^{-1}L^{-1}$, найти обратные матрицы к данным.

Глава 4 ЛИНЕЙНЫЕ ПРОСТРАНСТВА

4.1. Определение линейного пространства

Пусть дано произвольное числовое поле P, например, поле рациональных, действительных или комплексных чисел, и множество X элементов a, b, c, ... В множестве X определены операции сложения элементов и умножения их на числа из поля P: операция сложения элементов каждой паре a и b из X ставит во взаимно однозначное соответствие элемент $a+b$ из X, называемый ***суммой элементов*** a и b; операция умножения элементов на числа каждому элементу a из X и каждому числу α из P ставит во взаимно однозначное соответствие элемент αa из X, называемый ***произведением элемента*** a ***на число*** α.

Элементы множества X называют ***векторами***, а само множество X — ***линейным пространством над полем*** P, если операции сложения элементов из X и умножения их на числа α из P обладают следующими свойствами:

1) сложение коммутативно, т.е. $a+b=b+a$ для любых a, b из X;

2) сложение ассоциативно, т.е. $(a+b)+c=a+(b+c)$ для любых a, b, c из X;

3) в множестве X существует ***нулевой*** элемент 0, такой, что $a+0=a$ при любом a из X;

4) в множестве X для любого элемента a существует противоположный элемент $-a$, такой, что $a+(-a)=0$;

5) $\alpha(a+b)=\alpha a+\alpha b$ при любых a и b из X и любом $\alpha \in P$;

6) $(\alpha+\beta)a=\alpha a+\beta a$ при любом $a \in X$ и любых $\alpha, \beta \in P$;

7) $(\alpha\beta)a = \alpha(\beta a)$ при любом $a \in X$ и любых $\alpha, \beta \in P$;

8) $1 \cdot a = a$ при любом $a \in X$.

Свойства 1–8 называют ***аксиомами линейного пространства X над полем P***, поле P — ***основным полем***. В дальнейшем нас будет интересовать линейное пространство над полем действительных чисел, называемое ***действительным линейным пространством***, и линейное пространство над полем комплексных чисел, называемое ***комплексным линейным пространством***.

В линейном пространстве существует ***операция вычитания***, ставящая во взаимно однозначное соответствие каждой паре элементов a и b элемент $a - b = a + (-b)$, называемый ***разностью*** элементов a и b.

Примерами линейных пространств являются:

1) множество геометрических векторов, выходящих из начала координат на плоскости или в пространстве с обычными правилами сложения векторов и умножения их на действительные числа;

2) множество всех функций действительного переменного, определенных и непрерывных на отрезке $[a, b]$, с обычными правилами сложения функций и умножения их на действительные числа;

3) множество $P_n[x]$ многочленов степени не выше n с коэффициентами из поля P с обычными правилами сложения многочленов и умножения их на числа из поля P;

4) множество M_{mn} прямоугольных матриц размера $m \times n$ с элементами из поля P с обычными операциями сложения матриц и умножения их на числа из поля P;

5) множество всех векторов — решений однородной системы линейных уравнений с коэффициентами из поля P относительно сложения векторов и умножения их на числа из поля P;

6) ***арифметическое (координатное) пространство*** K_m векторов-столбцов $(\alpha_1, \alpha_2, \ldots, \alpha_m)^{\mathrm{T}}$ высоты m с компонентами $\alpha_1, \alpha_2, \ldots, \alpha_m$ из поля P, в котором операция сложения векторов-столбцов и умножения их на числа из поля P

6-4350

осуществляется по правилам:

$$(\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_m)^{\mathrm{T}} + (\beta_1,\ \beta_2,\ \ldots,\ \beta_m)^{\mathrm{T}} = \\ = (\alpha_1+\beta_1,\ \alpha_2+\beta_2,\ \ldots,\ \alpha_m+\beta_m)^{\mathrm{T}},$$

$$k(\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_m)^{\mathrm{T}} = (k\alpha_1,\ k\alpha_2,\ \ldots,\ k\alpha_m)^{\mathrm{T}}.$$

Из аксиом линейного пространства вытекают следующие простейшие следствия.

1. Значение суммы конечного числа векторов не зависит от порядка суммирования (например, для трех векторов $a + b + c = (a + b) + c = a + (b + c)$ в силу аксиомы 2). Поэтому операцию сложения векторов можно распространить на любое конечное число векторов.

2. В линейном пространстве уравнение $a + x = b$ относительно неизвестного x всегда имеет решение и притом единственное. Поэтому в линейном пространстве определена операция ***вычитания векторов*** $b - a$, обратная к операции сложения векторов. При этом $b - a = b + (-a)$.

3. В линейном пространстве нулевой вектор единственный. Это утверждение является частным случаем предыдущего, так как в соответствии с определением нулевой вектор — это решение уравнения $a + x = a$.

4. В линейном пространстве каждый вектор имеет единственный противоположный вектор. Это утверждение вытекает из единственности решения уравнения $a + x = 0$.

5. Произведение нулевого вектора на любое число α из основного поля равно нулевому вектору, т.е. $\alpha \cdot 0 = 0$. Действительно, $\alpha \cdot 0 = \alpha(0 + 0) = \alpha \cdot 0 + \alpha \cdot 0$. Следовательно, $\alpha \cdot 0 = \alpha \cdot 0 - \alpha \cdot 0 = 0$.

6. Произведение любого вектора a на число 0 равно нулевому вектору, т.е. $0 \cdot a = 0$. Действительно, $0 \cdot a = (0 + 0)a = 0 \cdot a + 0 \cdot a$, откуда: $0 \cdot a = 0 \cdot a - 0 \cdot a = 0$.

7. Если $\alpha a = 0$, то либо $\alpha = 0$, либо $a = 0$. В самом деле, если $\alpha \neq 0$, то

$$a = 1 \cdot a = \left(\frac{1}{\alpha} \cdot \alpha\right) a = \frac{1}{\alpha}(\alpha a) = \frac{1}{\alpha} 0 = 0.$$

8. Для каждого вектора a противоположный вектор равен произведению a на число -1, т.е. $(-a) = (-1)a$. В самом деле, $a + (-1)a = 1 \cdot a + (-1)a = (1-1)a = 0 \cdot a = 0$, а равенство $a + (-1)a = 0$ как раз и означает, что вектор $(-a)$ является противоположным к вектору a.

9. Для любого вектора a и любого числа α выполняется равенство $\alpha(-a) = (-\alpha a)$. Действительно,

$$\alpha(-a) = \alpha\big[(-1)a\big] = \big[\alpha(-1)\big]a = (-1)(\alpha a) = (-\alpha a).$$

10. Для любых векторов a и b и любого числа α выполняется равенство $\alpha(a-b) = \alpha a - \alpha b$. В самом деле,

$$\alpha(a-b) = \alpha\big[a + (-b)\big] = \alpha a + \alpha(-b) = \alpha a - \alpha b.$$

11. Для любого вектора a и любых чисел α, β выполняется равенство $(\alpha - \beta)a = \alpha a - \beta a$, так как

$$(\alpha - \beta)a = \big[\alpha + (-\beta)\big]a = \alpha a + (-\beta)a = \alpha a - \beta a.$$

В дальнейшем аксиомы линейного пространства и следствия, вытекающие из этих аксиом, будем использовать без специальных оговорок.

4.2. Линейная зависимость векторов

Вектор b называют ***пропорциональным*** вектору a, если $b = ka$ для некоторого числа k. В аналитической геометрии такие векторы называют *коллинеарными*. Вектор b называют ***линейной комбинацией*** (конечной) ***системы векторов*** $a_1, a_2, \ldots, a_s$, если существуют такие числа $\alpha_1, \alpha_2, \ldots, \alpha_s$, что

$$b = \alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_s a_s. \tag{4.1}$$

При этом говорят также, что ***вектор b линейно выражается через векторы*** $a_1, a_2, \ldots, a_s$.

Если вектор b линейно выражается через систему векторов $a_1, a_2, \ldots, a_s$, то он будет линейно выражаться и через любую конечную систему векторов, включающую в себя систему $a_1, a_2, \ldots, a_s$.

Действительно, если выполняется равенство

$$b = \alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_s a_s,$$

а система векторов $a_1, a_2, \ldots, a_s$ содержится в системе $a_1, a_2, \ldots, a_r$, $r > s$, то верно и равенство

$$b = \alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_s a_s + 0 \cdot a_{s+1} + \ldots + 0 \cdot a_r.$$

Это равенство означает, что вектор b линейно выражается через систему векторов $a_1, a_2, \ldots, a_r$.

Конечная система векторов $a_1, a_2, \ldots, a_r$ называется ***линейно зависимой***, если существуют такие числа $\alpha_1, \alpha_2, \ldots, \alpha_r$, не все равные нулю, что

$$\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_r a_r = 0. \tag{4.2}$$

В противном случае система векторов $a_1, a_2, \ldots, a_r$ ***линейно независима***.

Система векторов $a_1, a_2, \ldots, a_s$ линейно зависима тогда и только тогда, когда хотя бы один из ее векторов линейно выражается через остальные векторы.

Действительно, если система векторов $a_1, a_2, \ldots, a_s$ линейно зависима, то выполняется равенство

$$\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_s a_s = 0, \tag{4.3}$$

в котором, например, $\alpha_s \neq 0$. Тогда из этого равенства получаем:

$$a_s = -\frac{\alpha_1}{\alpha_s} a_1 - \frac{\alpha_2}{\alpha_s} a_2 - \ldots - \frac{\alpha_{s-1}}{\alpha_s} a_{s-1}.$$

Это означает, что вектор a_s линейно выражается через систему векторов $a_1, a_2, \ldots, a_{s-1}$.

Наоборот, пусть вектор a_s линейно выражается через систему векторов $a_1, a_2, \ldots, a_{s-1}$, т.е.

$$a_s = \alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_{s-1} a_{s-1}.$$

Тогда верно и равенство (4.3), в котором $\alpha_s = -1 \neq 0$. Значит, система векторов $a_1, a_2, \ldots, a_s$ линейно зависима.

Понятие линейной комбинации системы векторов в случае линейного арифметического пространства сводится к понятию линейной комбинации столбцов (см. разд. 3.3). Поэтому для определения линейного выражения вектора через заданную систему векторов, а также для проверки линейной зависимости векторов можно руководствоваться следующими правилами.

Правило 1. *Чтобы найти линейное выражение вектора b в линейном арифметическом пространстве K_m через систему векторов $a_1, a_2, \ldots, a_s$, следует записать равенство (4.1) и перейти к покомпонентным равенствам. В результате получается система линейных уравнений относительно неизвестных $\alpha_1, \alpha_2, \ldots, \alpha_s$, решив которую найдем коэффициенты линейного выражения вектора b через систему векторов $a_1, a_2, \ldots, a_s$.*

Правило 2. *Для того чтобы в линейном арифметическом пространстве проверить, является ли система векторов $a_1, a_2, \ldots, a_s$ линейно зависимой, необходимо записать векторное равенство (4.3) и перейти к покомпонентным равенствам. В результате получится система линейных однородных уравнений относительно неизвестных $\alpha_1, \alpha_2, \ldots, \alpha_s$. Если эта система имеет ненулевые решения, то система векторов $a_1, a_2, \ldots, a_s$ линейно зависима. Если же эта система имеет лишь нулевое решение $\alpha_1 = \alpha_2 = \ldots = \alpha_s = 0$, то система векторов $a_1, a_2, \ldots, a_r$ линейно независима.*

Пример 4.1. Выяснить вопрос о линейной зависимости векторов $a_1 = (1,\ -1,\ 2,\ 1)^{\mathrm{T}}$, $a_2 = (1,\ -1,\ 1,\ 2)^{\mathrm{T}}$, $a_3 = (1,\ -1,\ 4,\ -1)^{\mathrm{T}}$.

Решение. Составим векторное равенство $\alpha_1 a_1 + \alpha_2 a_2 + \alpha_3 a_3 = 0$ и от него перейдем к покомпонентным равенствам. Тогда получим систему

$$\begin{cases} \alpha_1 + \alpha_2 + \alpha_3 = 0, \\ -\alpha_1 - \alpha_2 - \alpha_3 = 0, \\ 2\alpha_1 + \alpha_2 + 4\alpha_3 = 0, \\ \alpha_1 + 2\alpha_2 - \alpha_3 = 0. \end{cases} \tag{4.4}$$

Эта система имеет ненулевое решение $\alpha_1 = 3$, $\alpha_2 = -2$, $\alpha_3 = -1$. Поэтому система векторов a_1, a_2, a_3 линейно зависимая, причем $3a_1 - 2a_2 - a_3 = 0$.

Теорема 4.1. *Если некоторая подсистема заданной системы векторов линейно зависима, то и вся система векторов линейно зависима.*

▷ Пусть дана система векторов $a_1, a_2, \ldots, a_r$ и ее подсистема* a_1, a_2, ..., a_s, $s < r$, линейно зависима, т.е. выполняется равенство

$$\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_s a_s = 0,$$

в котором не все коэффициенты α_1, α_2, ..., α_s равны нулю. Тогда будет выполняться и равенство

$$\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_s a_s + 0 \cdot a_{s+1} + \ldots + 0 \cdot a_r = 0.$$

Это доказывает линейную зависимость рассматриваемой системы векторов. ▶

Следствие 4.1. Конечная система векторов, содержащая нулевой вектор, линейно зависима. Конечная система векторов, содержащая пропорциональные векторы, линейно зависима.

Утверждение следует из того, что нулевой вектор и пропорциональные векторы составляют линейно зависимые системы векторов.

---

*Порядок векторов в системе в данном случае не является существенным. Поэтому можно считать, что векторы подсистемы предшествуют всем остальным векторам.

Следствие 4.2. Если система векторов линейно независима, то и всякая ее подсистема линейно независима.

Это утверждение, по существу, — лишь переформулировка теоремы: существование линейно зависимой подсистемы согласно теореме 4.1 противоречит исходному предположению следствия.

Пусть дана некоторая система векторов S (конечная или бесконечная) в линейном пространстве X над полем P. Конечная подсистема системы векторов S называется ***максимальной линейно независимой подсистемой***, если эта подсистема векторов линейно независима, а добавление к ней хотя бы одного вектора из S делает подсистему линейно зависимой.

Приведенное определение можно переформулировать так: конечная подсистема системы векторов S — максимальная линейно независимая подсистема, если она линейно независима и не содержится ни в какой линейно независимой подсистеме в S из большего числа векторов.

Теорема 4.2. Каждый вектор заданной системы векторов S линейно выражается через векторы любой ее максимальной линейно независимой подсистемы.

▷ Пусть $a_1, a_2, \ldots, a_n$ — максимально линейно независимая подсистема в системе векторов S и x — произвольный вектор в S. Тогда подсистема $a_1, a_2, \ldots, a_n, x$ уже линейно зависима, т.е. существуют числа $\alpha_1, \alpha_2, \ldots, \alpha_n, \alpha$ из P, одновременно не равные нулю, для которых

$$\alpha_1 a_1 + \ldots + \alpha_n a_n + \alpha x = 0. \tag{4.5}$$

В этом равенстве $\alpha \neq 0$, так как иначе последнее слагаемое можно опустить, и мы приходим к заключению, что система $a_1, a_2, \ldots, a_n$ линейно зависима. А такое заключение противоречит исходным предположениям. Так как $\alpha \neq 0$, неравенство (4.5) можно преобразовать к виду

$$x = -\frac{\alpha_1}{\alpha}a_1 - \frac{\alpha_2}{\alpha}a_2 - \ldots - \frac{\alpha_n}{\alpha}a_n,$$

т.е. получить **линейное выражение** вектора x через векторы $a_1, a_2, \ldots, a_n$. ▶

Говорят, что система векторов $a_1, a_2, \ldots, a_r$ линейно выражается через систему векторов $b_1, b_2, \ldots, b_s$, если каждый вектор системы $a_1, a_2, \ldots, a_r$ линейно выражается через векторы $b_1, b_2, \ldots, b_s$. Легко убедиться, что это понятие обладает свойством транзитивности, т.е. если даны три конечные системы векторов и вторая система линейно выражается через первую, а третья — через вторую, то третья система векторов линейно выражается через первую.

Две конечные системы векторов называют ***эквивалентными***, если они линейно выражаются одна через другую.

Непосредственно легко проверить, что:

• эквивалентность систем векторов обладает свойством транзитивности, т.е. если первая система векторов эквивалентна второй, а вторая — третьей, то первая система векторов эквивалентна третьей;

• если вектор линейно выражается через данную систему векторов, то он линейно выражается через любую другую систему векторов, эквивалентную данной.

Теорема 4.3 (основная теорема о линейной зависимости векторов). *Пусть даны две системы векторов $a_1, a_2, \ldots, a_r$ и $b_1, b_2, \ldots, b_s$, причем первая линейно независима и линейно выражается через вторую. Тогда число векторов в первой системе не превышает числа векторов во второй, т.е. $r \leqslant s$.*

Замечание. Утверждение теоремы, по существу, означает, что из s векторов нельзя создать систему линейных комбинаций этих векторов, которая, с одной стороны, линейно независима, а с другой — содержит более s векторов.

▷ По условию теоремы система векторов $a_1, a_2, \ldots, a_r$ линейно независима и линейно выражается через векторы системы $b_1, b_2, \ldots, b_s$. Следовательно, существуют такие числа α_{ij}, что выполняются равенства

$$\begin{cases} a_1 = \alpha_{11}b_1 + \alpha_{12}b_2 + \ldots + \alpha_{1s}b_s, \\ a_2 = \alpha_{21}b_1 + \alpha_{22}b_2 + \ldots + \alpha_{2s}b_s, \\ \dots\dots\dots\dots\dots\dots\dots \\ a_r = \alpha_{r1}b_1 + \alpha_{r2}b_2 + \ldots + \alpha_{rs}b_s. \end{cases} \tag{4.6}$$

Допустим, что $r > s$, и рассмотрим линейную комбинацию векторов

$$\lambda_1 a_1 + \lambda_2 a_2 + \ldots + \lambda_r a_r.$$

В силу равенств (4.6) эту линейную комбинацию можно представить следующим образом:

$$\lambda_1 a_1 + \lambda_2 a_2 + \ldots + \lambda_r a_r = \sum_{i=1}^{r} \lambda_i a_i =$$
$$= \sum_{i=1}^{r} \lambda_i \left(\sum_{j=1}^{s} \alpha_{ij} b_j \right) = \sum_{j=1}^{s} \left(\sum_{i=1}^{r} \alpha_{ij} \lambda_i \right) b_j.$$

В рассматриваемой линейной комбинации векторов попытаемся подобрать числа λ_1, λ_2, ..., λ_r так, что они одновременно не равны нулю, но при этом все коэффициенты при векторах b_1, b_2, ..., b_s обнуляются. Это означает, что набор чисел λ_1, λ_2, ..., λ_r является решением системы линейных однородных уравнений

$$\begin{cases} \alpha_{11}\lambda_1 + \alpha_{21}\lambda_2 + \ldots + \alpha_{r1}\lambda_r = 0, \\ \alpha_{12}\lambda_1 + \alpha_{22}\lambda_2 + \ldots + \alpha_{r2}\lambda_r = 0, \\ \dots\dots\dots\dots\dots\dots\dots \\ \alpha_{1s}\lambda_1 + \alpha_{2s}\lambda_2 + \ldots + \alpha_{rs}\lambda_r = 0. \end{cases}$$

При $r > s$ число неизвестных в системе превышает число уравнений, поэтому она имеет ненулевое решение. Любое ненулевое решение системы дает такой набор коэффициентов λ_1, λ_2, ..., λ_r, одновременно не обращающихся в нуль, для которых

$$\lambda_1 a_1 + \lambda_2 a_2 + \ldots + \lambda_r a_r = 0.$$

Существование таких коэффициентов равносильно линейной зависимости системы векторов $a_1, a_2, \ldots, a_r$, что противоречит условиям теоремы. Значит, предположение $r > s$ неверно и на самом деле $r \leqslant s$. ▶

Следствие 4.3. Любые две эквивалентные линейно независимые системы векторов имеют одинаковое число векторов.

Действительно, по доказанной теореме для двух линейно независимых эквивалентных систем векторов количество векторов в первой системе не превышает количества векторов во второй. Но системы в этом утверждении можно поменять местами, поэтому в первой системе не меньше векторов, чем во второй.

Заметим, что любые две максимальные линейно независимые подсистемы данной системы векторов эквивалентны. Значит, согласно доказанному следствию они имеют одно и то же число векторов.

Базисом (конечной) ***системы векторов*** называют любую ее максимальную линейно независимую подсистему. Через такую подсистему линейно выражается любой вектор системы. Из сказанного выше следует, что в данной системе векторов все базисы имеют одно и то же число векторов. Число векторов в базисе данной системы векторов называют ***рангом системы векторов***.

Теорема 4.4 (теорема о рангах двух систем векторов). Пусть даны две системы векторов, причем ранг первой равен k, а ранг второй равен l. Если первая система линейно выражается через вторую, то $k \leqslant l$. Если две системы эквивалентны, то $k = l$.

▷ Если первая система линейно выражается через вторую, то она линейно выражается и через любой базис второй системы. Следовательно, произвольный базис первой системы линейно выражается через базис второй системы, а количество векторов в базисе первой системы согласно теореме 4.3 не превышает числа векторов в базисе второй системы. Другими словами, $k \leqslant l$.

Если две системы векторов эквивалентны, то первая линейно выражается через вторую, откуда в силу только что доказанного $k \leqslant l$. В то же время вторая система линейно выражается через первую, поэтому $k \geqslant l$ и, следовательно, $k = l$. ▶

4.3. Ранг матрицы. Скелетное разложение матрицы

Пусть дана матрица

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

Столбцы матрицы A будем рассматривать как векторы линейного арифметического пространства K_m, а саму матрицу A — как конечную систему векторов.

Рангом матрицы называют ранг системы столбцов этой матрицы. Из определения ранга системы векторов вытекает, что ранг матрицы равен максимальному количеству линейно независимых столбцов матрицы (более точно: количеству столбцов в максимальной линейно независимой подсистеме системы столбцов матрицы). Ранг матрицы обозначают через $r(A)$.

Напомним, что если в матрице A выбраны какие-либо k строк и k столбцов, то определитель k-го порядка, составленный из элементов матрицы A, расположенных на пересечении выбранных строк и столбцов, называют ***минором k-го порядка*** матрицы A.

Теорема 4.5. *Если все миноры k-го порядка матрицы A равны нулю, то равны нулю и все миноры более высокого порядка.*

▷ Пусть M — произвольный минор матрицы A порядка, большего k. Разлагая этот минор по любым его k строкам в соответствии с теоремой Лапласа, получим:

$$M = \sum_i M_i A_i,$$

где M_i — миноры k-го порядка матрицы A. По условию теоремы все такие миноры равны нулю. Поэтому равен нулю и минор M. ▶

Теорема 4.6 (теорема о ранге матрицы). *Ранг матрицы равен наивысшему порядку среди отличных от нуля ее миноров. Каждый столбец матрицы линейно выражается через ее столбцы, проходящие через какой-либо из таких миноров.*

▷ Пусть дана матрица $A = (a_{ij})$ и пусть наивысший порядок среди ее ненулевых миноров равен r. Для определенности предположим, что один из ненулевых миноров наивысшего порядка — это минор D, расположенный в верхнем левом углу матрицы, т.е. в первых r строках и первых r столбцах матрицы:

$$A = \left(\begin{array}{ccc|ccc} a_{11} & \ldots & a_{1r} & a_{1,r+1} & \ldots & a_{1n} \\ \ldots & D & \ldots & \ldots & \ldots & \ldots \\ a_{r1} & \ldots & a_{rr} & a_{r,r+1} & \ldots & a_{rn} \\ \cline{1-3} a_{r+1,1} & \ldots & a_{r+1,r} & a_{r+1,r+1} & \ldots & a_{r+1,n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{m1} & \ldots & a_{mr} & a_{m,r+1} & \ldots & a_{mn} \end{array}\right).$$

Тогда первые r столбцов матрицы A будут линейно независимыми. Действительно, если бы эти столбцы были линейно зависимы, то и в миноре D столбцы были бы линейно зависимы. Но это возможно только для нулевого минора, а минор D согласно выбору не равен нулю.

Докажем, что столбцы, входящие в минор D, образуют максимальную линейно независимую подсистему в системе столбцов матрицы A. Для этого надо показать, что любой столбец линейно выражается через столбцы $a_1, a_2, \ldots, a_r$, входящие в минор D. Выберем произвольный l-й столбец. При $l \leqslant r$ утверждение очевидно, так как в этом случае

$$a_l = 0 \cdot a_1 + \ldots + 0 \cdot a_{l-1} + 1 \cdot a_l + 0 \cdot a_{l+1} + \ldots + 0 \cdot a_r.$$

Пусть $l > r$. Выберем произвольную i-ю строку, $1 \leqslant i \leqslant m$, и составим определитель

$$\Delta_i = \begin{vmatrix} a_{11} & \ldots & a_{1r} & a_{1l} \\ \cdot & \cdot & \cdot & \cdot \\ a_{r1} & \ldots & a_{rr} & a_{rl} \\ a_{i1} & \ldots & a_{ir} & a_{il} \end{vmatrix}.$$

Если $i \leqslant r$, то в определителе Δ_i две одинаковые строки и он равен нулю. Если же $i > r$, то этот определитель есть минор $(r+1)$-го порядка матрицы A и потому равен нулю, поскольку наивысший порядок среди ненулевых миноров равен r. Таким образом, определитель Δ_i нулевой при любом i.

Разлагая определитель Δ_i по последней строке, получим равенство

$$a_{i1}A_1 + a_{i2}A_2 + \ldots + a_{ir}A_r + a_{il}A_{r+1} = 0, \tag{4.7}$$

где A_1, A_2, ..., A_{r+1} — алгебраические дополнения элементов последней строки определителя Δ_i, причем $A_{r+1} = D \neq 0$. Отметим, что эти алгебраические дополнения не зависят от элементов последней строки определителя Δ_i и, следовательно, во всех определителях Δ_i для всех значений индекса $i = 1, 2, \ldots, m$ они одни и те же. Из равенства (4.7) находим:

$$a_{il} = -\frac{A_1}{A_{r+1}}a_{i1} - \frac{A_2}{A_{r+1}}a_{i2} - \ldots - \frac{A_r}{A_{r+1}}a_{ir}, \quad i = 1, 2, \ldots, m.$$

Записанные m равенств представляют собой покомпонентную запись векторного равенства

$$a_l = -\frac{A_1}{A_{r+1}}a_1 - \frac{A_2}{A_{r+1}}a_2 - \ldots - \frac{A_r}{A_{r+1}}a_r,$$

в подробной записи имеющего следующий вид:

$$\begin{pmatrix} a_{1l} \\ a_{2l} \\ \vdots \\ a_{ml} \end{pmatrix} = -\frac{A_1}{A_{r+1}} \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} - \frac{A_2}{A_{r+1}} \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} - \ldots - \frac{A_r}{A_{r+1}} \begin{pmatrix} a_{1r} \\ a_{2r} \\ \vdots \\ a_{mr} \end{pmatrix}.$$

Данное равенство означает, что l-й столбец матрицы A линейно выражается через столбцы, входящие в минор D. Таким образом, доказано, что если минор D имеет наивысший порядок среди всех ненулевых миноров матрицы A, то столбцы, входящие в этот минор, образуют максимальную линейно независимую подсистему (иными словами, базис) в системе всех столбцов матрицы A. В частности, порядок такого минора совпадает с рангом матрицы. ▶

Поскольку столбцы ненулевого минора матрицы A, имеющего наивысший порядок среди ненулевых миноров, образуют базис в системе всех столбцов матрицы, такие столбцы называют ***базисными столбцами***, а сам минор — ***базисным минором*** матрицы A. Доказанную теорему о ранге матрицы называют также ***теоремой о базисном миноре*** и часто формулируют так: ранг матрицы равен порядку любого ее базисного минора; все столбцы матрицы линейно выражаются через базисные столбцы.

Следствие 4.4. Ранг матрицы равен максимальному числу линейно независимых строк этой матрицы.

◁ При транспонировании матрицы ее миноры не изменяются, а базисные миноры матрицы будут базисными и в транспонированной матрице. Поэтому ранг матрицы, равный порядку базисного минора, при транспонировании не изменяется. Отсюда вытекает сформулированное утверждение. ▶

Доказанное следствие показывает, что ранг матрицы можно определить как ранг системы ее строк. Фактически доказано, что ранг системы строк матрицы совпадает с рангом системы ее столбцов. Аналогично понятию базисных столбцов вводится понятие „*базисные строки*“. Базисные строки — это строки матрицы, проходящие через какой-либо базисный минор матрицы.

Следствие 4.5. Определитель n-го порядка тогда и только тогда равен нулю, когда его строки (столбцы) линейно зависимы.

▷ Если строки (столбцы) определителя линейно зависимы, то по свойству 2.9 (см. разд. 2.3) он равен нулю. Предположим, что определитель равен нулю. Рассмотрим матрицу, соответствующую этому определителю. Единственный ее минор n-го порядка, т.е. исходный определитель, равен нулю. Поэтому ранг матрицы, который совпадает с максимальным количеством линейно независимых строк (столбцов), меньше n. Значит, строки (столбцы) матрицы линейно зависимы. ▶

Ранг матрицы можно вычислять разными способами. Один из способов, известный как ***метод окаймления***, вытекает из доказательства теоремы. По этому способу находят какой-либо минор первого или второго порядка, отличный от нуля, и вычисляют окаймляющие его миноры следующего порядка. Если среди них найдется отличный от нуля, то далее окаймляют его. Пусть уже найден таким способом минор r-го порядка, отличный от нуля. Тогда вычисляют его окаймляющие миноры $(r+1)$-го порядка. Если все они окажутся равными нулю или таких миноров вообще нет, то ранг матрицы равен r.

Пользуясь этим методом, вычислим ранг матрицы

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

В матрице минор, расположенный на пересечении первых двух строк и первых двух столбцов, отличен от нуля. У него всего лишь два окаймляющих минора третьего порядка, а именно:

$$\begin{vmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \\ 1 & 1 & 3 \end{vmatrix}, \qquad \begin{vmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \\ 3 & 1 & 4 \end{vmatrix}.$$

Оба этих минора равны нулю. Поэтому ранг $r(A)$ матрицы A равен 2. Базисным, например, является минор второго порядка в верхнем левом углу. Отсюда следует, что первые два столбца линейно независимые. В действительности линейно независимы любые два столбца матрицы. Например, второй и третий столбцы проходят через базисный минор $\begin{vmatrix} 2 & 3 \\ 1 & 3 \end{vmatrix}$, первый и третий столбцы проходят через базисный минор $\begin{vmatrix} 2 & 3 \\ 1 & 3 \end{vmatrix}$.

Для того чтобы вычислить ранг системы векторов $a_1, a_2, \ldots, a_s$ в линейном арифметическом пространстве K_m, из этих векторов, как из столбцов, следует записать матрицу и вычислить ее ранг.

Записав систему векторов арифметического пространства в матрицу, по базисным минорам легко выделить все максимальные линейно независимые подсистемы данной системы векторов. Поясним это правило на примере.

Пример 4.2. Найти ранг системы векторов $a_1 = (1, -1, 2, 1)^{\mathrm{T}}$, $a_2 = (1, -1, 1, 2)^{\mathrm{T}}$, $a_3 = (1, -1, 4, -1)^{\mathrm{T}}$ и выделить в ней все максимальные линейно независимые подсистемы векторов.

Решение. Составим матрицу

$$A = (a_1, a_2, a_3) = \begin{pmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 2 & 1 & 4 \\ 1 & 2 & -1 \end{pmatrix}.$$

Ее ранг равен двум. Следовательно, ранг системы векторов a_1, a_2, a_3 также равен двум. При этом каждая пара этих векторов составляет максимальную линейно независимую подсистему, так как каждой паре векторов соответствует базисный минор. Например, парам векторов a_1 и a_2, a_2 и a_3, a_1 и a_3 можно поставить в соответствие базисные миноры

$$\begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix}, \qquad \begin{vmatrix} 1 & 4 \\ 2 & -1 \end{vmatrix}, \qquad \begin{vmatrix} 2 & 4 \\ 1 & -1 \end{vmatrix},$$

расположенные в последних двух строках.

Пример 4.3. В системе векторов $a_1 = (1,\ 1,\ 1)^{\mathrm{T}}$, $a_2 = (1,\ 2,\ 3)^{\mathrm{T}}$, $a_3 = (1,\ 0,\ 1)^{\mathrm{T}}$, $a_4 = (1,\ 3,\ 3)^{\mathrm{T}}$, $a_5 = (1,\ -1,\ 1)^{\mathrm{T}}$ выделить какую-либо максимальную линейно независимую подсистему и через нее линейно выразить остальные векторы системы.

Решение. Составив матрицу

$$A = (a_1,\ a_2,\ a_3,\ a_4,\ a_5) = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 0 & 3 & -1 \\ 1 & 3 & 1 & 3 & 1 \end{pmatrix},$$

непосредственной проверкой убеждаемся в том, что в ней отличен от нуля минор

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 3 & 1 \end{vmatrix}.$$

Поэтому ранг матрицы A равен трем, и векторы a_1, a_2, a_3 составляют одну из максимальных линейно независимых подсистем данной системы векторов. Чтобы через эту подсистему линейно выразить вектор a_4, составим, как в примере 3.1 (см. разд. 3.3), векторное равенство

$$a_4 = \alpha_1 a_1 + \alpha_2 a_2 + \alpha_3 a_3$$

и от него перейдем к покомпонентным равенствам. Тогда придем к системе

$$\begin{cases} \alpha_1 + \ \alpha_2 + \alpha_3 = 1, \\ \alpha_1 + 2\alpha_2 = 3, \\ \alpha_1 + 3\alpha_2 + \alpha_3 = 3, \end{cases}$$

из которой найдем: $\alpha_1 = \alpha_2 = 1$, $\alpha_3 = -1$. Поэтому $a_4 = a_1 + a_2 - a_3$. Точно так же найдем линейное выражение $a_5 = -a_1 + 2a_3$.

Приведем некоторые свойства ранга матрицы.

Свойство 4.1. При транспонировании матрицы ее ранг не изменяется.

Свойство 4.2. При перестановке строк (столбцов) матрицы ее ранг не изменяется.

Свойство 4.3. Ранг матрицы не изменяется при умножении всех элементов какой-либо строки (столбца) на ненулевое число.

Свойство 4.4. Ранг матрицы не изменится, если из нее удалить или в нее добавить строку (столбец), состоящую из нулей.

Свойство 4.5. Ранг матрицы не изменится, если к одной ее строке (столбцу) прибавить другую строку (столбец), умноженную на некоторое число.

Свойство 4.6. Ранг матрицы не изменяется, если из нее удалить или добавить строку (столбец), являющуюся линейной комбинацией других ее строк (столбцов).

Все эти свойства почти очевидны. Докажем, например, свойство 4.6. Добавление к матрице линейной комбинации ее строк можно рассматривать как добавление к ней нулевой строки, к которой затем добавлены другие строки матрицы, умноженные на соответствующие коэффициенты. По свойствам 4.4 и 4.5 такая последовательность преобразований не изменяет ранга матрицы. Отсюда также следует, что удаление из матрицы строки, являющейся линейной комбинацией остальных строк матрицы, также не изменяет ранга матрицы, поскольку получение второй матрицы из первой удалением такой строки можно интерпретировать как получение первой матрицы из второй добавлением той же строки.

Свойства 4.2, 4.3, 4.5 можно объединить одним утверждением: ***ранг матрицы не изменяется при элементарных преобразованиях ее строк или столбцов.***

Сформулированное утверждение указывает и возможный способ вычисления ранга матрицы. Чтобы вычислить ранг матрицы, целесообразно упростить матрицу с помощью элементарных преобразований настолько, что заключение о ранге матрицы становится очевидным. Поясним это на примере.

Пример 4.4. С помощью элементарных преобразований вычислить ранг матрицы

$$A = \begin{pmatrix} 1 & 1 & -2 & 2 & -1 \\ 2 & 3 & 5 & -3 & -2 \\ 5 & 6 & -1 & 3 & -5 \end{pmatrix}.$$

Решение. Выполним в соответствии с методом Гаусса цепочку элементарных преобразований. В результате получим цепочку эквивалентных матриц:

$$A = \begin{pmatrix} 1 & 1 & -2 & 2 & -1 \\ 2 & 3 & 5 & -3 & -2 \\ 5 & 6 & -1 & 3 & -5 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & -2 & 2 & -1 \\ 0 & 1 & 9 & -7 & 0 \\ 0 & 1 & 9 & -7 & 0 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} \boxed{\begin{matrix} 1 & 1 \\ 0 & 1 \end{matrix}} & \begin{matrix} -2 & 2 & -1 \\ 9 & -7 & 0 \end{matrix} \\ \begin{matrix} 0 & 0 \end{matrix} & \begin{matrix} 0 & 0 & 0 \end{matrix} \end{pmatrix}.$$

Ранг последней матрицы этой цепочки матриц равен двум (отличен от нуля выделенный минор второго порядка). Следовательно, ранг матрицы A также равен двум.

Отметим, что при этих преобразованиях строки матрицы не переставлялись, поэтому выделенному базисному минору в преобразованной матрице в исходной матрице соответствует минор, также являющийся базисным. Более того, элементарные преобразования строк матрицы не нарушают линейных соотношений между столбцами этой матрицы. Анализируя такие соотношения в упрощенной матрице, можно затем переносить эти соотношения и на исходную матрицу. В нашем примере продолжим элементарные преобразования строк матрицы, добившись, чтобы базисный минор стал единичной матрицей:

$$A \sim \begin{pmatrix} 1 & 1 & -2 & 2 & -1 \\ 0 & 1 & 9 & -7 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} \boxed{\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}} & \begin{matrix} -11 & 9 & -1 \\ 9 & -7 & 0 \end{matrix} \\ \begin{matrix} 0 & 0 \end{matrix} & \begin{matrix} 0 & 0 & 0 \end{matrix} \end{pmatrix}.$$

Для столбцов $b_1, \ldots, b_5$ последней матрицы очевидны соотношения $b_3 = -11b_1 + 9b_2$, $b_4 = 9b_1 - 7b_2$, $b_5 = -b_1$. Точно такие же соотношения имеют место и между столбцами исходной матрицы A:

$$a_3 = -11a_1 + 9a_2, \quad a_4 = 9a_1 - 7a_2, \quad a_5 = -a_1.$$

Теорема 4.7. *Ранг произведения нескольких матриц не превышает рангов каждого из сомножителей.*

$\triangleright$ Рассмотрим произведение AB матриц

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \quad \text{и} \quad B = \begin{pmatrix} b_{11} & b_{12} & \ldots & b_{1s} \\ b_{21} & b_{22} & \ldots & b_{2s} \\ \ldots & \ldots & \ldots & \ldots \\ b_{n1} & b_{n2} & \ldots & b_{ns} \end{pmatrix}.$$

Обозначим столбцы матрицы A через $A_1, A_2, \ldots, A_n$ и запишем произведение AB в следующем виде:

$$AB = (A_1, A_2, \ldots, A_n) \begin{pmatrix} b_{11} & b_{12} & \ldots & b_{1s} \\ b_{21} & b_{22} & \ldots & b_{2s} \\ \ldots & \ldots & \ldots & \ldots \\ b_{n1} & b_{n2} & \ldots & b_{ns} \end{pmatrix} =$$

$$= (A_1 b_{11} + \ldots + A_n b_{n1}, \ \ldots, \ A_1 b_{1s} + \ldots + A_n b_{ns}).$$

Тогда становится очевидным, что столбцы матрицы AB являются линейными комбинациями столбцов матрицы A. Поэтому $r(AB) \leqslant r(A)$.

Отметим, что неравенство $r(AB) \leqslant r(A)$ доказано для произвольных матриц A и B, для которых определено произведение. Неравенство $r(AB) \leqslant r(B)$ доказывается аналогично, если использовать краткую запись $B = (B_1, B_2, \ldots, B_n)^{\mathrm{T}}$, где $B_1, B_2, \ldots, B_n$ — строки матрицы B, или следующим образом:

$$r(AB) = r((AB)^{\mathrm{T}}) = r(B^{\mathrm{T}}A^{\mathrm{T}}) \leqslant r(B^{\mathrm{T}}) = r(B). \qquad \blacktriangleright$$

Теорема 4.8. *Если A — матрица размера $m \times n$, Q_1 и Q_2 — квадратные невырожденные матрицы порядков m и n, то $r(Q_1A) = r(AQ_2) = r(A)$.*

$\vartriangleright$ Согласно теореме 4.7 для произведения $C = Q_1A$ имеем: $r(C) \leqslant r(A)$ и в то же время, поскольку $A = Q_1^{-1}C$, заключаем, что $r(A) \leqslant r(C)$. Из полученных неравенств вытекает равенство $r(C) = r(A)$. Аналогично доказывается утверждение теоремы для второго произведения AQ_2. $\blacktriangleright$

Представление действительной или комплексной $(m \times n)$-матрицы A ранга r в виде произведения

$$A = BC \tag{4.8}$$

$(m \times r)$-матрицы B ранга r и $(r \times n)$-матрицы C ранга r называют ***скелетным разложением*** матрицы A.

Существование скелетного разложения у любой матрицы вытекает из следующих соображений. Произведение BC в равенстве (4.8) справа можно интерпретировать как матричную запись представления столбцов матрицы A в виде линейных комбинаций столбцов матрицы B (см. доказательство теоремы 4.7). Условие $r(B) = r$ означает, что столбцы матрицы B линейно независимы. Значит, скелетное разложение мы получим, если матрицу B составим из столбцов какой-либо максимальной линейно независимой подсистемы в системе столбцов матрицы A, а матрицу C — из столбцов коэффициентов линейных разложений столбцов матрицы A по столбцам матрицы B.

Эти рассуждения показывают, что скелетное разложение матрицы A не является единственным, так как в произведении

$A = BC$ матрицу B можно определить с помощью любой максимальной линейно независимой подсистемы в системе столбцов матрицы A. Более того, на самом деле матрица B может включать столбцы, не входящие в матрицу A. Это обстоятельство увеличивает свободу выбора матрицы B. Но при этом выбор одного из сомножителей скелетного разложения однозначно определяет выбор второго сомножителя.

Отметим два важных случая.

1. Если ранг $(m \times n)$-матрицы совпадает с числом ее столбцов, то в скелетном разложении (4.8) можно положить $B = A$ и $C = E$, где E — единичная матрица n-го порядка.

2. Если ранг матрицы A совпадает с числом ее строк, то в скелетном разложении (4.8) можно положить $C = A$, а в качестве матрицы B взять единичную матрицу m-го порядка.

Пример 4.5. Построить скелетное разложение для матрицы

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

Решение. Здесь ранг матрицы A равен двум и первые ее два столбца линейно независимы. Поэтому положим

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

и будем искать элементы j-го столбца матрицы C, $j = 1, 2, 3$, из равенства

$$c_{1j} \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} + c_{2j} \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} = A_j, \tag{4.9}$$

где A_j — j-й столбец матрицы A.

При $j=1$ имеем $A_j = A_1 = (2,\ -1,\ 0)^{\mathrm{T}}$, и равенство (4.9) приводит к системе

$$\begin{cases} 2c_{11} - c_{21} = 2, \\ -c_{11} + c_{21} = -1, \\ c_{21} = 0, \end{cases}$$

из которой находим: $c_{11} = 1$, $c_{21} = 0$. Поэтому первым столбцом матрицы C будет столбец $(1,\ 0)^{\mathrm{T}}$. Аналогично при $j = 2, 3$ находим второй и третий столбцы $(0,\ 1)^{\mathrm{T}}$ и $(1,\ 2)^{\mathrm{T}}$ матрицы C. В результате

$$C = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \end{pmatrix},$$

и одним из скелетных разложений матрицы A является

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

На практике при построении скелетного разложения матрицы A удобно пользоваться методом элементарных преобразований, который состоит в следующем. Матрицу с помощью элементарных преобразований строк приводят к ступенчатому виду (последовательность таких преобразований целесообразно проводить по схеме метода Гаусса). В матрице ступенчатого вида удаляют все нулевые строки, выбирают базисный минор и с помощью элементарных преобразований строк всей матрицы преобразуют этот минор в единичную матрицу. Как следует из примера 4.4, столбцы полученной матрицы связаны теми же линейными соотношениями, что и соответствующие столбцы матрицы A. Составив матрицу B из столбцов матрицы A, на которых расположен выбранный базисный минор, а в качестве матрицы C выбрав ту, которая получена из A элементарными преобразованиями строк, придем к равенству $A = BC$, дающему скелетное разложение матрицы A.

Пример 4.6. Для матрицы

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

из предыдущего примера получаем:

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

$$\sim \begin{pmatrix} \boxed{\begin{matrix} 1 & -1/2 \\ 0 & 1/2 \end{matrix}} & \begin{matrix} 0 \\ 1 \end{matrix} \\ \begin{matrix} 0 & 0 \end{matrix} & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1/2 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \end{pmatrix} = C.$$

В данном случае базисный минор выбран в первых двух столбцах матрицы ступенчатого вида. Значит, матрицу B составляем из первых двух столбцов матрицы A, а матрицей C будет конечная матрица в цепочке преобразований. В результате

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

4.4. Базис и размерность пространства

Линейное пространство называют ***конечномерным***, если в нем есть линейно независимая система, состоящая из n векторов, а любая конечная система из большего числа векторов линейно зависима. Если такого числа нет, т.е. если для любого числа n в линейном пространстве существует линейно независимая система из n векторов, то такое линейное пространство называют ***бесконечномерным***.

Конечномерным линейным пространством является линейное пространство геометрических векторов на прямой, на плоскости или в пространстве с обычными операциями сложения векторов и умножения вектора на действительные числа. Линейное пространство всех функций действительного переменного, определенных и непрерывных на отрезке $[a, b]$, с обычными операциями сложения функций и умножения функции на действительные числа является бесконечномерным. Чтобы убедиться в этом, рассмотрим функции 1, x, x^2, ..., x^N. Любая линейная комбинация этих функций представляет собой многочлен

$$p(x) = \alpha_0 \cdot 1 + \alpha_1 x + \alpha_2 x^2 + \ldots + \alpha_N x^N$$

степени не выше N. У всякого многочлена с ненулевыми коэффициентами есть лишь конечное число нулей. Поэтому $p(x) \equiv 0$ тогда и только тогда, когда $\alpha_0 = \alpha_1 = \ldots = \alpha_N = 0$. Следовательно, система функций 1, x, x^2, ..., x^N линейно независима, а линейное пространство бесконечномерно, поскольку число N может быть сколь угодно большим.

В линейной алгебре изучаются лишь конечномерные пространства.

Всякую конечную упорядоченную систему векторов линейного пространства X называют ***базисом*** или ***базой*** линейного пространства, если эта система векторов линейно независима и любой вектор линейного пространства линейно выражается через векторы системы. В определении сказано, что базис — упорядоченная система векторов. Это означает, что каждому вектору в базисе приписан определенный номер. Из одной и той же системы векторов можно получить разные базисы, по-разному нумеруя векторы. Как следует из определения, базис является максимальной линейно независимой системой векторов в линейном пространстве. Наоборот, любая максимальная линейно независимая система векторов в линейном пространстве является базисом (конечно, после фиксации определенного порядка векторов в системе). Отметим, что любую линейно независимую систему векторов в линейном пространстве можно дополнить до максимальной, т.е. до базиса в этом линейном пространстве. Отсюда, в частности, следует, что в конечномерном линейном пространстве существует бесконечное число базисов.

Как вытекает из основной теоремы о линейной зависимости векторов (см. теорему 4.3), в данном линейном пространстве все базисы имеют одно и то же число векторов. Если в линейном пространстве X существует базис из n векторов, то это пространство конечномерно, причем n есть максимальное число линейно независимых векторов в этом линейном пространстве. Это число называют ***размерностью линейного пространства*** и пишут $\dim X = n$, само линейное пространство называют ***n-мерным***.

Пример 4.7. Линейное арифметическое (координатное) пространство K_n над полем P, элементами которого являются векторы-столбцы $(x_1, x_2, \ldots, x_n)^{\mathrm{T}}$, составленные из элементов x_i поля P, является n-мерным пространством. Действительно, в этом линейном пространстве базисом является система векторов

$$
\begin{aligned}
e_1 &= (1,\ 0,\ 0,\ \ldots,\ 0)^{\mathrm{T}},\\
e_2 &= (0,\ 1,\ 0,\ \ldots,\ 0)^{\mathrm{T}},\\
&\ldots\ldots\ldots\ldots\\
e_1 &= (0,\ 0,\ 0,\ \ldots,\ 1)^{\mathrm{T}}.
\end{aligned}
$$

Любой вектор $x = (x_1, x_2, \ldots, x_n)^{\mathrm{T}} \in K_n$ может быть представлен в виде

$$x = x_1e_1 + x_2e_2 + \ldots + x_ne_n,$$

т.е. компоненты вектора-столбца в то же время являются коэффициентами его линейного выражения через векторы базиса. Ясно, что равенство нулю этого вектора равносильно равенству нулю всех коэффициентов линейной комбинации, т.е. указанная система векторов линейно независима. Количество n векторов системы и определяет размерность арифметического пространства. Базис $e_1, e_2, \ldots, e_n$ называют ***естественным базисом арифметического пространства***.

В качестве базиса арифметического пространства можно взять любые n векторов-столбцов $f_i = (f_{i1}, f_{i2}, \ldots, f_{in})^{\mathrm{T}}$, $i = 1, 2, \ldots, n$, образующих линейно независимую систему. Матрица, составленная из таких столбцов, будет невырожденной.

Пример 4.8. Линейное пространство $P_n[x]$ многочленов степени не выше n с коэффициентами из поля P является $(n+1)$-мерным. В качестве базиса в этом линейном пространстве можно выбрать систему векторов

$$1, \quad x, \quad x^2, \quad \ldots, \quad x^n. \tag{4.10}$$

Любой вектор

$$f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_m x^m, \quad m \leqslant n, \tag{4.11}$$

линейно выражается через систему векторов (4.10), причем коэффициенты многочлена в то же самое время являются коэффициентами его разложения по системе (4.10).

В качестве базиса в линейном пространстве $P_n[x]$ можно также взять систему векторов

$$1, \quad x - \alpha, \quad (x-\alpha)^2, \quad \ldots, \quad (x-\alpha)^n, \tag{4.12}$$

где $\alpha \neq 0$ фиксировано. Чтобы разложить вектор (4.11) по системе векторов (4.12), необходимо многочлен представить в виде

$$f(x) = b_0 + b_1(x-\alpha) + b_2(x-\alpha)^2 + \ldots + b_m(x-\alpha)^m.$$

Коэффициенты b_i этого представления могут быть найдены в соответствии с теоремой Тейлора по формулам

$$b_0 = f(\alpha), \ b_1 = f'(\alpha), \ b_2 = \frac{f''(\alpha)}{2!}, \ \ldots, \ b_m = \frac{f^{(m)}(\alpha)}{m!}, \ 0, \ \ldots, \ 0.$$

Теорема 4.9. *Пусть линейное пространство X_n обладает базисом $e_1, e_2, \ldots, e_n$. Тогда любой вектор x из X_n единственным образом представляется в виде*

$$x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n. \tag{4.13}$$

Замечание. Линейное выражение (4.13) вектора x через векторы базиса, единственное в силу сформулированной теоремы, называют ***разложением вектора*** x ***по базису*** $e_1, e_2, \ldots, e_n$. Это разложение удобно записывать в матричной форме

$$x = x_1e_1 + x_2e_2 + \ldots + x_ne_n = (e_1,\ e_2,\ \ldots,\ e_n)\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = e\cdot[x],$$

где $e = (e_1, e_2, \ldots, e_n)$ — заданный базис, записанный в виде матрицы-строки, а $[x]$ — столбец коэффициентов разложения вектора x по базису, называемых ***координатами вектора*** x ***в базисе*** e.

$\triangleright$ В силу определения базиса линейного пространства X любой вектор $x \in X$ имеет хотя бы одно представление вида (4.13). Предположим, что наряду с разложением (4.13) есть и другое разложение

$$x = x'_1e_1 + x'_2e_2 + \ldots + x'_ne_n$$

вектора x. Тогда будет выполняться равенство

$$x = x_1e_1 + x_2e_2 + \ldots + x_ne_n = x'_1e_1 + x'_2e_2 + \ldots + x'_ne_n,$$

которое приводит к равенству

$$(x_1 - x'_1)e_1 + (x_2 - x'_2)e_2 + \ldots + (x_n - x'_n)e_n = 0.$$

Последнее равенство в силу линейной независимости системы векторов $e_1, e_2, \ldots, e_n$ возможно лишь в случае, когда

$$x_1 - x'_1 = 0, \quad x_2 - x'_2 = 0, \quad \ldots, \quad x_n - x'_n = 0,$$

т.е. при $x_1 = x'_1$, $x_2 = x'_2$, $\ldots$, $x_n = x'_n$. Это доказывает, что разложение вектора по базису единственно. ▶

Как уже сказано, коэффициенты $x_1, x_2, \ldots, x_n$ в разложении (4.13) вектора x по базису $e_1, e_2, \ldots, e_n$ называют координатами вектора x. Эти координаты будем записывать в виде $x(x_1, x_2, \ldots, x_n)$. Столбец $[x]_e = (x_1, x_2, \ldots, x_n)^{\mathrm{T}}$ будем называть ***координатным столбцом вектора** x **в базисе** e*. Индекс в записи координатного столбца будем опускать, если ясно, о каком базисе идет речь и путаницы не возникает. Отметим, что часто (когда это не приводит к путанице) вектор отождествляют с его координатным столбцом.

Векторы линейного пространства X полностью определяются своими координатами в данном базисе. При этом операции над векторами сводятся к аналогичным операциям над их координатами. Так, для векторов

$$x = x_1e_1 + x_2e_2 + \ldots + x_ne_n \quad \text{и} \quad y = y_1e_1 + y_2e_2 + \ldots + y_ne_n$$

условие их равенства $x = y$ равносильно условиям

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

а равенства

$$x + y = (x_1e_1 + x_2e_2 + \ldots + x_ne_n) + (y_1e_1 + y_2e_2 + \ldots + y_ne_n) =$$
$$= (x_1 + y_1)e_1 + (x_2 + y_2)e_2 + \ldots + (x_n + y_n)e_n$$

и

$$\lambda x = \lambda(x_1e_1 + x_2e_2 + \ldots + x_ne_n) = (\lambda x_1)e_1 + (\lambda x_2)e_2 + \ldots + (\lambda x_n)e_n$$

показывают, что при сложении векторов их соответствующие координаты складываются, при умножении вектора на число все его координаты умножаются на это число. Из этих заключений вытекает следующее правило.

Координатный столбец линейной комбинации векторов является линейной комбинацией их координатных столбцов с теми же коэффициентами и наоборот. Система векторов линейно зависима тогда и только тогда, когда линейно зависима система их координатных столбцов.

Это правило указывает способ, позволяющий определять линейные соотношения между векторами.

Чтобы найти линейное выражение вектора b через векторы a_1, a_2, …, a_s, следует записать векторное равенство $b = \alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_s a_s$ и от него перейти к покоординатным равенствам. В результате получается система линейных уравнений относительно неизвестных α_1, α_2, …, α_s, любое решение которой дает записанное векторное равенство.

4.5. Связь между базисами линейного пространства

Пусть в линейном пространстве X_n заданы базисы

$$e = (e_1,\, e_2,\, \ldots,\, e_n) \quad \text{и} \quad e' = \bigl(e'_1,\, e'_2,\, \ldots,\, e'_n\bigr). \qquad (4.14)$$

Разложим векторы базиса e' по базису e:

$$\begin{cases} e'_1 = t_{11}e_1 + t_{21}e_2 + \ldots + t_{n1}e_n, \\ e'_2 = t_{12}e_1 + t_{22}e_2 + \ldots + t_{n2}e_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ e'_n = t_{1n}e_1 + t_{2n}e_2 + \ldots + t_{nn}e_n. \end{cases} \qquad (4.15)$$

Матрицу

$$T = \begin{pmatrix} t_{11} & t_{12} & \ldots & t_{1n} \\ t_{21} & t_{22} & \ldots & t_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ t_{n1} & t_{n2} & \ldots & t_{nn} \end{pmatrix}$$

называют ***матрицей перехода*** от базиса e к базису e'. Заметим, что столбцами матрицы T являются столбцы координат соответствующих векторов базиса e' в базисе e.

Соотношения (4.15) устанавливают связь между базисами e и e'. Эти соотношения удобно записать в матричной форме

$$\left(e_1', e_2', \ldots, e_n'\right) = (e_1, e_2, \ldots, e_n)\begin{pmatrix} t_{11} & t_{12} & \ldots & t_{1n} \\ t_{21} & t_{22} & \ldots & t_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ t_{n1} & t_{n2} & \ldots & t_{nn} \end{pmatrix}, \quad (4.16)$$

или кратко

$$e' = eT. \quad (4.17)$$

Точно так же векторы базиса e можно разложить по базису e', и тогда придем к соотношению

$$e = e'T', \quad (4.18)$$

где T' — матрица перехода от базиса e' к базису e. Столбцами матрицы T' служат координатные столбцы соответствующих векторов базиса e в базисе e'.

Из соотношений $e = e'T'$ и $e' = eT$ следуют соотношения

$$e = e'T' = (eT)T' = e(TT'),$$
$$e' = ET = (e'T')T = e'(T'T),$$

из которых вытекают соотношения

$$e(E - TT') = 0 \quad \text{и} \quad e'(E - T'T) = 0.$$

Из этих соотношений в силу линейной независимости векторов базисов e и e' получаем: $TT' = T'T = E$. Следовательно, $T' = T^{-1}$.

Таким образом, матрица перехода от одного базиса n-мерного линейного пространства к другому является невырожденной матрицей n-го порядка с элементами из основного поля P. Верно и противоположное утверждение, т.е. верна следующая теорема.

Теорема 4.10. *Любая невырожденная квадратная матрица n-го порядка с элементами из поля P служит матрицей перехода от данного базиса n-мерного линейного пространства X над полем P к некоторому другому базису в X.*

▷ Пусть даны базис $e = (e_1, e_2, \ldots, e_n)$ линейного пространства X и невырожденная квадратная матрица

$$T = \begin{pmatrix} t_{11} & t_{12} & \ldots & t_{1n} \\ t_{21} & t_{22} & \ldots & t_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ t_{n1} & t_{n2} & \ldots & t_{nn} \end{pmatrix}$$

n-го порядка с элементами из поля P. В пространстве X выберем упорядоченную систему векторов $e' = (e'_1, e'_2, \ldots, e'_n)$, для которых столбцы матрицы T являются координатными столбцами в базисе e.

Система векторов e' состоит из n векторов и является линейно независимой, так как у невырожденной матрицы T столбцы линейно независимы. Поэтому эта система — базис в линейном пространстве X, причем в силу выбора векторов системы выполняется равенство $e' = eT$. Это означает, что матрица T представляет собой матрицу перехода от базиса e к базису e'. ▶

Из доказанной теоремы вытекает, что в n-мерном линейном пространстве X над полем P существует столько различных базисов, сколько существует различных невырожденных квадратных матриц n-го порядка с элементами из поля P. При этом учтено, что различны базисы, состоящие из одних векторов, но по-разному упорядоченных.

Практическое правило. *Для построения матрицы T перехода от базиса e к базису e' нужно для каждого вектора e'_j базиса e' найти координаты в базисе e и из них, как из столбцов, построить матрицу T.*

Если векторы базисов e и e' заданы координатами в некотором базисе e°, то для отыскания координат вектора e'_j в базисе e следует составить векторное равенство

$$e'_j = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n,$$

от него перейти к покоординатным равенствам и из полученной системы уравнений найти искомый столбец координат $(\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_n)^{\mathrm{T}}$ вектора e'_j в базисе e.

При отыскании матрицы T можно также пользоваться формулой

$$T = T_1^{-1} T_2, \tag{4.19}$$

где T_1 — матрица перехода от базиса e° к базису e, T_2 — матрица перехода от базиса e° к базису e'.

Чтобы доказать формулу (4.19), замечаем, что выполняются соотношения

$$e = e^\circ T_1, \quad e' = e^\circ T_2, \quad e' = eT.$$

Из этих соотношений получаем: $e' = e^\circ T_2 = (eT_1^{-1})T_2 = e(T_1^{-1}T_2)$, откуда следует, что $T_1^{-1}T_2$ — матрица перехода от базиса e к базису e', т.е. верно равенство (4.19).

Пример 4.9. Найти матрицу перехода от базиса $e = (e_1,\ e_2)$ к базису $e' = (e'_1,\ e'_2)$, где $[e'_1]_e = (2,\ 1)^{\mathrm{T}}$, $[e'_2]_e = (3,\ 2)^{\mathrm{T}}$.

Решение. Здесь векторы нового базиса заданы координатами в старом базисе. Поэтому сразу можно составить искомую матрицу T из координатных столбцов векторов e'_1 и e'_2:

$e'_1 = 2e_1 + 1\cdot e$
$e'_2 = 3\cdot e_1 + 2\cdot$

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

Пример 4.10. Найти матрицу перехода от базиса $e = (e_1,\ e_2,\ e_3)$ к базису $e' = (e'_1,\ e'_2,\ e'_3)$, где векторы заданы своими координатами в некотором базисе: $e_1 = (1,\ 1,\ 1)^{\mathrm{T}}$, $e_2 = (1,\ 2,\ 3)^{\mathrm{T}}$, $e_3 = (1,\ 0,\ 1)^{\mathrm{T}}$, $e'_1 = (-1,\ 0,\ 1)^{\mathrm{T}}$, $e'_2 = (1,\ 3,\ 3)^{\mathrm{T}}$, $e'_3 = (1,\ -1,\ -1)^{\mathrm{T}}$.

Решение. Составим векторное равенство

$$e'_j = \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3, \quad j = 1, 2, 3.$$

При $j = 1$ это равенство принимает вид:

$$\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} = \alpha_1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \alpha_3 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Это равенство приводит к системе

$$\begin{cases} -1 = \alpha_1 + \alpha_2 + \alpha_3, \\ 0 = \alpha_1 + 2\alpha_2, \\ 1 = \alpha_1 + 3\alpha_2 + \alpha_3, \end{cases}$$

из которой находим: $\alpha_1 = -2$, $\alpha_2 = 1$, $\alpha_3 = 0$. Следовательно, $e_1' = -2e_1 + e_2$. Аналогично при $j = 2$ и $j = 3$ получаем: $e_2' = e_1 + e_2 - e_3$, $e_3' = e_1 - e_2 + e_3$. Из коэффициентов полученных разложений записываем матрицу перехода:

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

Можно также этот ответ получить по формуле (4.19). Для этого, пользуясь координатами векторов, записываем матрицы перехода

$$T_1 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 3 & 1 \end{pmatrix}, \qquad T_2 = \begin{pmatrix} -1 & 1 & 1 \\ 0 & 3 & -1 \\ 1 & 3 & -1 \end{pmatrix}.$$

Далее:

$$T = T_1^{-1}T_2 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 3 & 1 \end{pmatrix}^{-1} \begin{pmatrix} -1 & 1 & 1 \\ 0 & 3 & -1 \\ 1 & 3 & -1 \end{pmatrix} =$$

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

Пример 4.11. В линейном пространстве $P_2[x]$ многочленов не выше второй степени с действительными коэффициентами даны два базиса: $e = (e_1, e_2, e_3)$, где $e_1 = 1$, $e_2 = x$, $e_3 = x^2$, и $e' = (e'_1, e'_2, e'_3)$, где $e'_1 = 1$, $e'_2 = x - 1$, $e'_3 = (x-1)^2$. Найти матрицу перехода T от базиса e к базису e'.

Решение. Так как

$$\begin{aligned} e'_1 &= 1 = e_1, \\ e'_2 &= x - 1 = -1 + x = -e_1 + e_2, \\ e'_3 &= (x-1)^2 = 1 - 2x + x^2 = e_1 - 2e_2 + e_3, \end{aligned}$$

то

$$[e'_1]_e = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad [e'_2]_e = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \qquad [e'_3]_e = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}.$$

Следовательно,

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

4.6. Преобразование координат вектора при переходе от базиса к базису

Пусть в линейном пространстве X даны два базиса $e = (e_1, e_2, \ldots, e_n)$ и $e' = (e'_1, e'_2, \ldots, e'_n)$ с матрицей перехода T от базиса e к базису e', т.е. пусть $e' = eT$ и пусть вектор x имеет в базисах e и e' координаты $[x]_e = (x_1, x_2, \ldots, x_n)^{\mathrm{T}}$ и $[x]_{e'} = (x'_1, x'_2, \ldots, x'_n)^{\mathrm{T}}$, т.е. пусть $x = e[x]_e$ и $x = e'[x]_{e'}$. Тогда, с одной стороны,

$$x = e[x]_e,$$

а с другой стороны,

$$x = e'[x]_{e'} = (eT)[x]_{e'} = e(T[x]_{e'}).$$

Сравнив правые части этих равенств, получим:

$$e[x]_e = e(T[x]_{e'}),$$

откуда в силу единственности разложения вектора по базису e вытекает равенство

$$[x]_e = T[x]_{e'}, \tag{4.20}$$

или

$$\begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} = T \begin{pmatrix} x'_1 \\ x'_2 \\ \dots \\ x'_n \end{pmatrix}. \tag{4.21}$$

В развернутом виде записанное равенство имеет следующий вид:

$$\begin{cases} x_1 = t_{11}x'_1 + t_{12}x'_2 + \ldots + t_{1n}x'_n, \\ x_2 = t_{21}x'_1 + t_{22}x'_2 + \ldots + t_{2n}x'_n, \\ \dots\dots\dots\dots\dots\dots\dots \\ x_n = t_{n1}x'_1 + t_{n2}x'_2 + \ldots + t_{nn}x'_n. \end{cases} \tag{4.22}$$

Соотношения (4.20)–(4.22) называют ***формулами преобразования координат*** при изменении базиса линейного пространства. Они выражают старые координаты вектора через его новые координаты. Эти формулы можно разрешить относительно новых координат вектора. Тогда получим формулы:

$$[x]_{e'} = T^{-1}[x]_e, \tag{4.23}$$

или в развернутом виде

$$\begin{cases} x'_1 = c_{11}x_1 + c_{12}x_2 + \ldots + c_{1n}x_n, \\ x'_2 = c_{21}x_1 + c_{22}x_2 + \ldots + c_{2n}x_n, \\ \dots\dots\dots\dots\dots\dots\dots \\ x'_n = c_{n1}x_1 + c_{n2}x_2 + \ldots + c_{nn}x_n, \end{cases} \tag{4.24}$$

где c_{ij} — элементы матрицы T^{-1}.

Поясним применение этих формул на примере.

Пример 4.12. В пространстве X_3 заданы вектор x и векторы e'_1, e'_2, e'_3 базиса e' координатами в базисе e: $[x]_e = (1,\ 4,\ -1)^{\mathrm{T}}$, $[e'_1]_e = (5,\ -1,\ 2)^{\mathrm{T}}$, $[e'_2]_e = (2,\ 3,\ 0)^{\mathrm{T}}$, $[e'_3]_e = (-2,\ 1,\ 1)^{\mathrm{T}}$. Задан также вектор y своими координатами в базисе e': $[y]_{e'} = (1,\ 2,\ 3)^{\mathrm{T}}$. Найти координаты вектора x в базисе e' и координаты вектора y в базисе e.

Решение. Ввиду того, что векторы базиса e' заданы столбцами координат в базисе e, можно составить матрицу перехода T от базиса e к базису e', сгруппировав координатные столбцы векторов базиса e':

$$T = \begin{pmatrix} 5 & 2 & -2 \\ -1 & 3 & 1 \\ 2 & 0 & 1 \end{pmatrix}.$$

По формуле (4.20) получаем:

$$[y]_e = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = T \begin{pmatrix} y'_1 \\ y'_2 \\ y'_3 \end{pmatrix} = \begin{pmatrix} 5 & 2 & -2 \\ -1 & 3 & 1 \\ 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 \\ 8 \\ 1 \end{pmatrix}.$$

Далее вычисляем обратную матрицу перехода

$$T^{-1} = \begin{pmatrix} 3 & 2 & 8 \\ -1 & 1 & -3 \\ 6 & -4 & 17 \end{pmatrix}$$

и по формуле (4.23) находим:

$$[x]_{e'} = \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = T^{-1} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 & 2 & 8 \\ -1 & 1 & -3 \\ 6 & -4 & 17 \end{pmatrix} \begin{pmatrix} 1 \\ 4 \\ -1 \end{pmatrix} = \begin{pmatrix} -13 \\ 6 \\ -27 \end{pmatrix}.$$

Формулой (4.23) удобно пользоваться при отыскании матрицы перехода от базиса e к базису e', когда векторы этих базисов заданы своими координатами в некотором третьем базисе e°.

Пример 4.13. Найти матрицу перехода от базиса $e = (e_1, e_2, e_3)$ к базису $e' = (e'_1, e'_2, e'_3)$, если $[e_1]_{e^\circ} = (1, 1, 1)^{\mathrm{T}}$, $[e_2]_{e^\circ} = (1, 2, 3)^{\mathrm{T}}$, $[e_3]_{e^\circ} = (1, 0, 1)^{\mathrm{T}}$ и $[e'_1]_{e^\circ} = (1, 1, 1)^{\mathrm{T}}$, $[e'_2]_{e^\circ} = (1, 2, 1)^{\mathrm{T}}$, $[e'_3]_{e^\circ} = (1, 1, 3)^{\mathrm{T}}$.

Решение. Перейдем от базиса e° к базису e и найдем координаты векторов базиса e' в базисе e. Матрицей перехода от базиса e° к базису e является матрица

$$T_1 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 3 & 1 \end{pmatrix},$$

а обратной к ней — матрица

$$T_1^{-1} = \frac{1}{2}\begin{pmatrix} 2 & 2 & -2 \\ -1 & 0 & 1 \\ 1 & -2 & 1 \end{pmatrix}.$$

Поэтому для векторов e'_1, e'_2, e'_3 получаем:

$$[e'_1]_e = T_1^{-1}[e'_1]_{e^\circ} = \frac{1}{2}\begin{pmatrix} 2 & 2 & -2 \\ -1 & 0 & 1 \\ 1 & -2 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix};$$

$$[e'_2]_e = T_1^{-1}[e'_2]_{e^\circ} = \frac{1}{2}\begin{pmatrix} 2 & 2 & -2 \\ -1 & 0 & 1 \\ 1 & -2 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 4 \\ 0 \\ -2 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix};$$

$$[e'_3]_e = T_1^{-1}[e'_3]_{e^\circ} = \frac{1}{2}\begin{pmatrix} 2 & 2 & -2 \\ -1 & 0 & 1 \\ 1 & -2 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} -2 \\ 2 \\ 2 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}.$$

Из полученных координатных столбцов составляем матрицу перехода от базиса e к базису e':

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

Замечание. Изложенный в примере способ построения матрицы перехода равносилен применению формулы (4.19).

4.7. Изоморфизм линейных пространств

Векторами линейных пространств могут быть объекты различной природы: направленные отрезки, матрицы, многочлены, функции, решения систем линейных уравнений и т.п. При изучении линейных пространств интерес представляют не сами векторы, а операции над ними и свойства этих операций. Может случиться, что, хотя векторы каких-либо двух линейных пространств по своей природе совершенно различные, с точки зрения свойств операций над векторами эти пространства неразличимы. В линейной алгебре такая ситуация описывается следующим образом. Пусть X и X' — линейные пространства над одним и тем же полем P. Пространства X и X' называют ***изоморфными***, если между этими пространствами существует взаимно однозначное соответствие, удовлетворяющее двум условиям:

1) если векторам x и y линейного пространства X соответствуют векторы x' и y' линейного пространства X', то вектору $x+y\in X$ соответствует вектор $x'+y'\in X'$;

2) если вектору $x\in X$ соответствует вектор $x'\in X'$, то для любого элемента $\lambda\in P$ вектору $\lambda x\in X$ соответствует вектор $\lambda x'\in X'$.

Кратко такое соответствие можно записать так:

$$x\leftrightarrow x',\ y\leftrightarrow y' \quad\Longrightarrow\quad x+y\leftrightarrow x'+y' \quad (x,y\in X,\ x',y'\in X');$$
$$x\leftrightarrow x' \quad\Longrightarrow\quad \lambda x\leftrightarrow \lambda x' \quad (x\in X,\ x'\in X',\ \lambda\in P).$$

Приведем примеры изоморфных линейных пространств.

Пример 4.14. 1. Линейное пространство X геометрических векторов, выходящих из начала координат трехмерного пространства, с обычными операциями сложения векторов и умножения вектора на число изоморфно действительному арифметическому пространству K_3, так как каждому вектору

$x \in X$ можно поставить во взаимно однозначное соответствие вектор-столбец $x' = (x_1,\ x_2,\ x_3)^{\mathrm{T}}$ его координат в некотором фиксированном базисе. При таком соответствии будут выполняться соотношения

$$x + y \leftrightarrow x' + y' = (x_1 + y_1,\ x_2 + y_2,\ x_3 + y_3)^{\mathrm{T}};$$
$$\lambda x \leftrightarrow \lambda x' = (\lambda x_1,\ \lambda x_2,\ \lambda x_3)^{\mathrm{T}}.$$

2. Линейное пространство M_{22} квадратных матриц второго порядка над полем P с обычными операциями сложения матриц и умножения матрицы на элементы поля P и арифметическое пространство K_4 над полем P изоморфны, так как каждой матрице

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

из M_{22} во взаимно однозначное соответствие можно поставить вектор-столбец $\alpha = (a_{11},\ a_{12},\ a_{21},\ a_{22})^{\mathrm{T}}$ из K_4 и при этом будут выполняться соотношения

$$A + B \leftrightarrow \alpha + \beta, \qquad \lambda A \leftrightarrow \lambda\alpha.$$

3. Линейное пространство $P_2[x]$ многочленов степени $n \leqslant 2$ с действительными коэффициентами с обычными операциями сложения многочленов и умножения многочлена на действительное число изоморфно действительному арифметическому пространству K_3, так как многочлену $f(x) = a_0 + a_1 x + a_2 x^2$ можно поставить в соответствие вектор-столбец $(a_0,\ a_1,\ a_2)^{\mathrm{T}}$, при этом будут выполняться соотношения

$$(a_0 + a_1 x + a_2 x^2) + (b_0 + b_1 x + b_2 x^2) \leftrightarrow (a_0,\ a_1,\ a_2)^{\mathrm{T}} + (b_0,\ b_1,\ b_2)^{\mathrm{T}},$$
$$\lambda(a_0 + a_1 x + a_2 x^2) \leftrightarrow \lambda(a_0,\ a_1,\ a_2)^{\mathrm{T}}.$$

Непосредственно из определения изоморфизма линейных пространств вытекают следующие утверждения.

1. *При изоморфизме линейных пространств X и X' нулевому вектору в X соответствует нулевой вектор в X'.*

2. Если вектору $x \in X$ соответствует вектор $x \in X'$, то противоположному вектору $-x \in X$ соответствует противоположный вектор $-x' \in X'$.

3. Если векторам $x_1, x_2, \ldots, x_k$ в X соответствуют векторы $x_1', x_2', \ldots, x_k'$ в X', то линейной комбинации $\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_k x_k \in X$ с произвольными коэффициентами $\alpha_1, \ldots, \alpha_k$ соответствует линейная комбинация $\alpha_1 x_1' + \alpha_2 x_2' + \ldots \alpha_k x_k' \in X'$ с теми же коэффициентами.

4. Для изоморфных линейных пространств X и X' каждой линейно независимой системе векторов в X соответствует линейно независимая система векторов в X' и наоборот.

5. Для изоморфных линейных пространств X и X' каждому базису в X соответствует базис в X' и наоборот.

Действительно, первое утверждение следует из соотношений

$$0 = 0 \cdot x \leftrightarrow 0 \cdot x' = 0'$$

(здесь 0 и $0'$ — нулевые векторы в X и X'), второе — из соотношений

$$-x = (-1)x \leftrightarrow (-1)x' = -x'.$$

Если $x_1 \leftrightarrow x_1', \ldots, x_k \leftrightarrow x_k'$, то $\alpha_1 x_1 \leftrightarrow \alpha_1 x_1', \ldots, \alpha_k x_k \leftrightarrow \alpha_k x_k'$ и $\alpha_1 x_1 + \alpha_2 x_2 \leftrightarrow \alpha_1 x_1' + \alpha_2 x_2'$. Продолжая последовательно добавлять соответствующие друг другу слагаемые, получим:

$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_k x_k \leftrightarrow \alpha_1 x_1' + \alpha_2 x_2' + \ldots + \alpha_k' x_k'.$$

Чтобы доказать четвертое свойство, выберем в X произвольную линейно независимую систему векторов $a_1, a_2, \ldots, a_k$, которой соответствует система векторов $a_1', a_2', \ldots, a_k'$ в X'. Если

$$\alpha_1 a_1' + \alpha_2 a_2' + \ldots + \alpha_k a_k' = 0' \in X',$$

то с учетом соотношений

$$\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_k a_k \leftrightarrow \alpha_1 a_1' + \alpha_2 a_2' + \ldots + \alpha_k a_k' \quad \text{и} \quad 0 \leftrightarrow 0'$$

заключаем, что

$$\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_k a_k = 0,$$

откуда: $\alpha_1 = \alpha_2 = \ldots = \alpha_k = 0$. Следовательно, система векторов a'_1, a'_2, $\ldots$, a'_k линейно независима. Аналогично доказательство в обратную сторону.

Для доказательства пятого свойства отметим, что базис в линейном пространстве — это любая максимальная линейно независимая система. Но если систему e_1, $\ldots$, e_n в X можно расширить добавлением некоторого вектора f, сохраняя свойство линейной независимости, то и соответствующую систему e'_1, $\ldots$, e'_n можно также расширить, добавив к ней вектор f', соответствующий вектору f. Это рассуждение показывает, что при изоморфизме линейных пространств максимальные линейно независимые системы переходят в максимальные линейно независимые системы, а это фактически и утверждается пятым свойством.

Теорема 4.11. *Линейное пространство X', изоморфное конечномерному линейному пространству X, является конечномерным и имеет ту же размерность, что и X. Любые два конечномерных линейных пространства X и X' одинаковой размерности изоморфны.*

$\triangleright$ Как уже доказано, любому базису линейного пространства X соответствует базис изоморфного линейного пространства X'. Поэтому размерности этих линейных пространств, определяемые количеством векторов в базисе, совпадают.

Пусть X и X' — n-мерные линейные пространства над одним и тем же полем P. Выберем в них базисы соответственно $e = (e_1, e_2, \ldots, e_n)$ и $e' = (e'_1, e'_2, \ldots, e'_n)$. Каждому вектору $x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n$ линейного пространства X поставим в соответствие вектор $x' = x_1 e'_1 + x_2 e'_2 + \ldots + x_n e'_n$ линейного пространства X', имеющий в базисе e' те же координаты, что и вектор x в базисе e. Тогда сумме векторов в $x + y \in X$ будет соответствовать сумма векторов $x' + y' \in X'$:

$$x + y = (x_1 + y_1)e_1 + (x_2 + y_2)e_2 + \ldots + (x_n + y_n)e_n \leftrightarrow$$
$$\leftrightarrow (x_1 + y_1)e'_1 + (x_2 + y_2)e'_2 + \ldots + (x_n + y_n)e'_n = x' + y'.$$

Аналогично проверяем, что соответствие сохраняется при умножении векторов на числа:

$$\lambda x = \lambda x_1 e_1 + \lambda x_2 e_2 + \ldots + \lambda x_n e_n \leftrightarrow$$
$$\leftrightarrow \lambda x_1 e_1' + \lambda x_2 e_2' + \ldots + \lambda x_n e_n' = \lambda x'.$$

Это означает, что линейные пространства X и X' изоморфны. ▶

Следствие 4.6. Все линейные пространства над одним и тем же полем P одинаковой размерности n изоморфны n-мерному арифметическому линейному пространству K_n над полем P.

4.8. Системы линейных уравнений (общая теория)

Пусть дана система линейных уравнений

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{cases} \tag{4.25}$$

с матрицей

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

и расширенной матрицей

$$B = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{pmatrix}.$$

Такую систему можно записать в виде

$$Ax = b. \tag{4.26}$$

Напомним (см. разд. 2.1), что ***решением системы*** (4.25) называют упорядоченный набор из n чисел x_1, x_2, ..., x_n, при подстановке которых в уравнения системы вместо соответствующих неизвестных каждое уравнение превращается в тождество. Система, имеющая хотя бы одно решение, называется ***совместной***, а если у системы решений нет, она называется ***несовместной***, или ***противоречивой***. Две системы с одними и теми же неизвестными, имеющие одно и то же множество решений, называются ***эквивалентными*** (***равносильными***).

Теорема 4.12 (теорема Кронекера — Капелли). Система линейных уравнений (4.25) совместна тогда и только тогда, когда ранг матрицы A системы равен рангу ее расширенной матрицы B.

▷ Пусть система совместна. Докажем, что $r(A) = r(B)$. Для этого возьмем какое-либо решение k_1, k_2, ..., k_n системы и поставим его в каждое уравнение системы. Тогда получим систему тождеств

$$\begin{cases} a_{11}k_1 + a_{12}k_2 + \ldots + a_{1n}k_n = b_1, \\ a_{21}k_1 + a_{22}k_2 + \ldots + a_{2n}k_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}k_1 + a_{m2}k_2 + \ldots + a_{mn}k_n = b_m. \end{cases} \tag{4.27}$$

Эту систему тождеств перепишем в виде

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} k_1 + \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} k_2 + \ldots + \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} k_n = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Такая запись показывает, что столбец свободных членов системы является линейной комбинацией столбцов матрицы системы. Следовательно, при вычислении ранга матрицы B этот столбец в соответствии со свойством 4.6 (см. разд. 4.3) ранга матрицы можно из матрицы B удалить. Это означает выполнение равенства $r(A) = r(B)$.

Пусть теперь выполняется равенство $r(A) = r(B)$. Докажем, что система совместна. Равенство $r(A) = r(B)$ означает, что базис системы столбцов матрицы A является и базисом системы столбцов матрицы B, а столбец свободных членов линейно выражается через столбцы матрицы A, т.е. существует набор таких чисел k_1, k_2, ..., k_n, что выполняется равенство (4.27). Но это означает, что набор чисел $k_1, k_2, \ldots, k_n$ является решением рассматриваемой системы. Следовательно, система совместна. ▶

Если для данной системы выполняется равенство $r(A) = r(B) = r$, то эту систему будем называть ***системой ранга r***.

Пусть система (4.25) совместна и ее ранг равен r. Выберем какой-либо базисный минор матрицы. Уравнения системы, которым соответствуют строки матрицы A, входящие в базисный минор, назовем ***базисными уравнениями***, а подсистему всех базисных уравнений — ***базисной подсистемой***. Верно следующее утверждение.

Теорема 4.13. Совместная система линейных уравнений эквивалентна любой своей базисной подсистеме.

▷ По условию теоремы система совместна. Поэтому выполняется равенство $r(A) = r(B)$. Из теоремы 4.6 о ранге матрицы вытекает, что все строки матрицы B, не входящие в базисный минор, являются линейными комбинациями базисных строк. Это означает, что уравнения, не входящие в базисную подсистему, являются следствием базисных уравнений, а любое решение базисной подсистемы является и решением всей системы уравнений. Наоборот, очевидно, что любое решение системы является решением базисной подсистемы. Таким образом, множество решений системы уравнений совпадает с множеством решений базисной подсистемы. ▶

Пусть для системы (4.25) базисную подсистему образуют первые r уравнений, т.е. базисной является подсистема

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{r1}x_1 + a_{r2}x_2 + \ldots + a_{rn}x_n = b_r. \end{cases} \tag{4.28}$$

Очевидно, что ранг системы не превышает количества неизвестных, т.е. всегда выполняется неравенство $r \leqslant n$. Рассмотрим два случая: $r = n$ и $r < n$.

При $r = n$ базисная подсистема (4.28) является крамеровской, и мы приходим к следующему утверждению: ***если ранг совместной системы равен числу ее неизвестных, то система имеет единственное решение***.

Пусть теперь $r < n$. Предположим, что базисный минор матрицы системы расположен в первых r столбцах. Значит, и в базисной подсистеме базисный минор составляют первые r столбцов. Неизвестные, соответствующие столбцам базисного минора, т.е. при нашем предположении x_1, x_2, $\ldots$, x_r, назовем ***главными неизвестными***, а остальные неизвестные — ***свободными***. Перенесем все слагаемые в уравнениях базисной подсистемы, содержащие свободные неизвестные x_{r+1}, x_{r+2}, $\ldots$, x_n, из левых частей в правые и дадим этим неизвестным произвольные значения c_{r+1}, c_{r+2}, $\ldots$, c_n. Тогда получим систему относительно главных неизвестных, являющуюся крамеровской. Она имеет единственное решение c_1, $c_2, \ldots, c_r$. Очевидно, что набор чисел $c_1, c_2, \ldots, c_r, c_{r+1}, \ldots, c_n$ будет решением системы (4.28), а потому и решением системы (4.25). Значения c_{r+1}, c_{r+2}, $\ldots$, c_n можно задавать произвольно, и мы, подбирая их соответствующим образом, можем указанным путем получить любое решение системы (4.25). В самом деле, если $c_1, c_2, \ldots, c_r, c_{r+1}, \ldots, c_n$ — какое-либо решение системы (4.25), то, положив в системе (4.28) $x_{r+1} = c_{r+1}$, $x_{r+2} = c_{r+2}$, ..., $x_n = c_n$, получим систему относительно главных неизвестных, решением которой является набор чисел $x_1 = c_1$, $x_2 = c_2$, ..., $x_r = c_r$. Можно заключить, что, варьируя произвольным образом значения свободных неизвестных и определяя соответствующие значения главных неизвестных, мы получаем все множество решений системы (4.25). Попутно установлено следующее утверждение: ***если ранг совместной системы меньше числа неизвестных, то система имеет бесконечное множество решений.***

На практике нет необходимости специально выяснять, совместна система или нет, сравнивая ранги двух матриц. Несовместность системы естественным образом обнаружится в процессе решения системы. Для решения системы можно воспользоваться методом Гаусса в следующей модификации. Расширенную матрицу системы с помощью элементарных преобразований строк приводим к ступенчатому виду. Ранг матрицы ступенчатого вида равен количеству ее ненулевых строк. Поэтому после приведения системы к ступенчатому виду выяснение вопроса о совместности системы не вызывает труда. Если система совместна, то в матрице системы выбирают базисную подсистему, состоящую из всех ненулевых строк, а в ней разделяют неизвестные на главные и свободные. В системе ступенчатого вида выбрать главные неизвестные можно так, что минор в столбцах, соответствующих этим неизвестным, будет представлять собой верхний треугольный определитель. Решение системы относительно главных неизвестных — это решение крамеровской системы с верхней треугольной матрицей. Чтобы решить такую систему, достаточно расширенную матрицу базисной подсистемы с помощью элементарных преобразований строк преобразовать так, что выбранный базисный минор окажется единичной матрицей (это преобразование — обратный ход метода Гаусса). После всех преобразований мы получаем систему, эквивалентную исходной, но уже разрешенную относительно главных неизвестных. Поясним описанную процедуру на конкретном примере.

Пример 4.15. Решить систему

$$\begin{cases} x_1 + x_2 - x_3 - x_4 = 1, \\ x_1 + x_2 - 2x_3 + 3x_4 = 2, \\ x_1 + x_2 + x_3 - 9x_4 = -1. \end{cases}$$

Решение. Решение проведем в матричной записи. Для этого запишем расширенную матрицу системы и с помощью элементарных преобразований приведем к ступенчатому виду:

$$\left(\begin{array}{rrrr|r} 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & -2 & 3 & 2 \\ 1 & 1 & 1 & -9 & -1 \end{array}\right) \sim \left(\begin{array}{rrrr|r} 1 & 1 & -1 & -1 & 1 \\ 0 & 0 & -1 & 4 & 1 \\ 0 & 0 & 2 & -8 & -2 \end{array}\right) \sim$$

$$\sim \left(\begin{array}{rrrr|r} 1 & 1 & -1 & -1 & 1 \\ 0 & 0 & -1 & 4 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Последняя матрица соответствует системе, эквивалентной исходной. В этой матрице отбросим третью нулевую строку и выберем базисный минор в первом и третьем столбце, имеющий вид:

$$\begin{vmatrix} 1 & -1 \\ 0 & -1 \end{vmatrix}.$$

При таком выборе главными неизвестными являются x_1 и x_3, а свободными x_2 и x_4. Продолжим преобразование матрицы системы, превращая выбранный базисный минор в единичную матрицу:

$$\left(\begin{array}{rrrr|r} 1 & 1 & -1 & -1 & 1 \\ 0 & 0 & -1 & 4 & 1 \end{array}\right) \sim \left(\begin{array}{rrrr|r} 1 & 1 & 0 & -5 & 0 \\ 0 & 0 & -1 & 4 & 1 \end{array}\right) \sim$$

$$\sim \left(\begin{array}{rrrr|r} 1 & 1 & 0 & -5 & 0 \\ 0 & 0 & 1 & -4 & -1 \end{array}\right).$$

Полученная в результате всех преобразований матрица

$$\left(\begin{array}{rrrr|r} 1 & 1 & 0 & -5 & 0 \\ 0 & 0 & 1 & -4 & -1 \end{array}\right)$$

соответствует системе

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

Из этой системы находим: $x_1 = -x_2 + 5x_4$, $x_3 = -1 + 4x_4$.

4.9. Однородные системы линейных уравнений

Напомним, что однородной системой линейных уравнений называют систему

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = 0, \\ \dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = 0, \end{cases} \qquad (4.29)$$

у которой все свободные члены уравнений равны нулю. Такая система в матричной форме имеет вид:

$$Ax = 0. \qquad (4.30)$$

Любая однородная система совместна, поскольку имеет нулевое решение $x_1 = x_2 = \ldots = x_n = 0$. Для такой системы применимы результаты предыдущего параграфа. Эти результаты приводят к следующему утверждению.

Теорема 4.14. *Для того чтобы однородная система линейных уравнений имела единственное решение (это будет нулевое решение), необходимо и достаточно, чтобы ранг ее матрицы совпадал с количеством неизвестных системы.*

Если ранг матрицы однородной системы меньше числа неизвестных, то, как следует из предыдущего параграфа, система имеет бесконечное множество решений.

В частном случае, когда однородная система является квадратной (т.е. количество уравнений равно количеству неизвестных), совпадение ранга матрицы системы с числом неизвестных равносильно условию невырожденности матрицы системы. Если определитель однородной квадратной системы равен нулю, то система имеет бесконечное множество решений.

Если количество неизвестных системы превышает количество уравнений, то ранг матрицы меньше количества неизвестных, а система имеет бесконечное множество решений.

Множество решений однородной системы, которые естественно записывать как векторы-столбцы, обладает следующими свойствами.

1. Если вектор-столбец x — решение однородной системы $Ax = 0$, то для любого числа λ вектор-столбец λx также является решением этой системы.

2. Если векторы-столбцы x и y — решения однородной системы $Ax = 0$, то вектор-столбец $x + y$ также является решением системы.

Эти свойства вытекают из матричных соотношений

$$\begin{gathered}
Ax = 0 \implies A(\lambda x) = \lambda Ax = \lambda \cdot 0 = 0, \\
Ax = Ay = 0 \implies A(x + y) = Ax + Ay = 0 + 0 = 0.
\end{gathered}$$

Два приведенных свойства можно объединить в одно утверждение: линейная комбинация решений однородной системы является решением этой системы. Другими словами, множество решений однородной системы с n неизвестными является линейным пространством*. Это линейное пространство конечномерно, базис этого пространства называют ***фундаментальной системой решений*** (ФСР). Существование ФСР, а следовательно, и конечномерность линейного пространства решений вытекают из следующего утверждения.

Теорема 4.15. *Если ранг матрицы A однородной системы $Ax = 0$ меньше числа неизвестных, т.е. $r(A) = r < n$, то*

*На самом деле доказано несколько иное утверждение: множество решений однородной системы с n неизвестными является подпространством в n-мерном линейном арифметическом пространстве K_n (см. разд. 4.11). То, что множество решений — линейное пространство, можно установить непосредственной проверкой аксиом линейного пространства.

размерность линейного пространства решений этой системы равна $n - r$.

▷ Если ранг матрицы A равен $r < n$, то из n неизвестных r являются главными, а $n - r$ — свободными. Предположим, что в качестве свободных выбраны неизвестные x_{r+1}, x_{r+2}, ..., x_n. Возьмем произвольный ненулевой определитель d порядка $n - r$:

$$d = \begin{vmatrix} c_{1,r+1} & c_{1,r+2} & \dots & c_{1n} \\ c_{2,r+1} & c_{2,r+2} & \dots & c_{2n} \\ \dots & \dots & \dots & \dots \\ c_{n-r,r+1} & c_{n-r,r+2} & \dots & c_{n-r,n} \end{vmatrix} \neq 0.$$

Свободным неизвестным x_{r+1}, x_{r+2}, ..., x_n будем давать значения i-й строки этого определителя, $i = 1, 2, \dots, n - r$ и для заданных значений свободных неизвестных определять из системы $Ax = 0$ значения главных неизвестных x_1, x_2, ..., x_r. В результате получим $n - r$ решений однородной системы $Ax = 0$:

$$x_1 = \begin{pmatrix} c_{11} \\ \vdots \\ c_{1r} \\ c_{1,r+1} \\ \vdots \\ c_{1n} \end{pmatrix}, \qquad \dots, \qquad x_{n-r} = \begin{pmatrix} c_{n-r,1} \\ \vdots \\ c_{n-r,r} \\ c_{n-r,r+1} \\ \vdots \\ c_{n-r,n} \end{pmatrix}.$$

Покажем, что этот набор решений представляет собой фундаментальную систему решений рассматриваемой однородной системы $Ax = 0$. Из столбцов решений составим матрицу

$$M = \begin{pmatrix} c_{11} & c_{21} & \dots & c_{n-r,1} \\ \dots & \dots & \dots & \dots \\ c_{1r} & c_{2r} & \dots & c_{n-r,r} \\ c_{1,r+1} & c_{2,r+1} & \dots & c_{n-r,r+1} \\ \dots & \dots & \dots & \dots \\ c_{1n} & c_{2n} & \dots & c_{n-r,n} \end{pmatrix}.$$

Ранг этой матрицы равен $n-r$, т.е. количеству столбцов, поскольку в последних $n-r$ строках матрицы расположен ненулевой минор порядка $n-r$ (это определитель, транспонированный к тому, который был выбран нами). Следовательно, все столбцы матрицы, входящие в ненулевой минор, линейно независимы.

Покажем, что любое решение

$$x = (b_1, \ldots, b_r, b_{r+1}, \ldots, b_n)^{\mathrm{T}}$$

системы $Ax = 0$ линейно выражается через решения $x_1, \ldots, x_{n-r}$. Рассмотрим столбцы

$$x'_1 = \begin{pmatrix} c_{1,r+1} \\ \vdots \\ c_{1n} \end{pmatrix}, \qquad \ldots, \qquad x'_{n-r} = \begin{pmatrix} c_{n-r,r+1} \\ \vdots \\ c_{n-r,n} \end{pmatrix},$$

которые получаются, если в столбцах $x_1, \ldots, x_{n-r}$ вычеркнуть первые r компонент, соответствующих главным неизвестным. Поскольку квадратная матрица, составленная из этих столбцов, является невырожденной, столбцы линейно независимы. А так как количество столбцов равно размерности $n-r$ арифметического пространства K_{n-r}, элементами которого они являются, система столбцов является базисом в K_{n-r}, а столбец $x' = (b_{r+1}, \ldots, b_n)^{\mathrm{T}}$ линейно выражается через столбцы $x'_1, \ldots, x'_{n-r}$. Пусть

$$x' = k_1x'_1 + k_2x'_2 + \ldots + k_{n-r}x'_{n-r}. \tag{4.31}$$

Рассмотрим n-мерный вектор

$$y = k_1x_1 + k_2x_2 + \ldots + k_{n-r}x_{n-r} - x. \tag{4.32}$$

Вектор y является решением системы $Ax = 0$, поскольку он представляет собой линейную комбинацию решений системы. В силу равенства (4.31) в решении y все свободные неизвестные имеют нулевые значения. Поскольку значения свободных неизвестных однозначно определяют значения главных неизвестных, а у системы $Ax = 0$ есть решение с нулевыми значениями свободных неизвестных — нулевое решение, заключаем, что и главные неизвестные решения y имеют нулевые значения, т.е. $y = 0$. Но тогда из равенства (4.32) получаем:

$$x = k_1x_1 + k_2x_2 + \ldots + k_{n-r}x_{n-r},$$

т.е. произвольно взятое решение x системы $Ax = 0$ линейно выражается через решения $x_1, x_2, \ldots, x_{n-r}$. ▶

Теорема 4.16. *Пусть $x_1, x_2, \ldots, x_{n-r}$ — фундаментальная система решений однородной системы $Ax = 0$. Вектор x является решением системы $Ax = 0$ тогда и только тогда, когда он может быть представлен в виде*

$$x = c_1x_1 + c_2x_2 + \ldots + c_{n-r}x_{n-r}, \tag{4.33}$$

где $c_1, c_2, \ldots, c_{n-r}$ — некоторые постоянные.

▷ Любой вектор, представимый в виде (4.33), является решением системы $Ax = 0$, поскольку является линейной комбинацией решений системы. В то же время любое решение x системы $Ax = 0$ линейно выражается через фундаментальную систему решений, поскольку по определению фундаментальная система решений — базис в линейном пространстве всех решений системы. Значит, имеет место представление (4.33). ▶

Формула (4.33), описывающая все множество решений однородной системы, представляет собой общее решение системы.

Из доказательства теоремы 4.15 вытекает следующее ***правило построения фундаментальной системы решений.***

1. В матрице системы находят базисный минор, например, приведением матрицы к ступенчатому виду. Выбор базисного минора позволяет в системе уравнений выделить базисную подсистему, а неизвестные разделить на главные и свободные.

2. Выбирают произвольный ненулевой определитель порядка, равного количеству $n-r$ свободных неизвестных.

3. По каждой строке выбранного определителя находят соответствующее решение системы. Для этого в качестве значений свободных неизвестных берут элементы строки и, решая базисную подсистему, находят значения главных неизвестных.

4. Полученные таким образом $n-r$ решений системы составляют фундаментальную систему решений.

Поясним это правило на примере.

Пример 4.16. Найти какую-либо фундаментальную систему решений системы уравнений

$$\begin{cases} x_1 - 2x_2 + x_3 + x_4 = 0, \\ 5x_1 - x_2 - 13x_3 - 4x_4 = 0, \\ 2x_1 + 5x_2 - 16x_3 - 7x_4 = 0. \end{cases}$$

Решение. Записав матрицу системы, приводим ее элементарными преобразованиями строк к ступенчатому виду:

$$\begin{pmatrix} 1 & -2 & 1 & 1 \\ 5 & -1 & -13 & -4 \\ 2 & 5 & -16 & -7 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & 1 & 1 \\ 0 & 9 & -18 & -9 \\ 0 & 9 & -18 & -9 \end{pmatrix} \sim$$
$$\sim \begin{pmatrix} 1 & -2 & 1 & 1 \\ 0 & 1 & -2 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -3 & -1 \\ 0 & 1 & -2 & -1 \end{pmatrix}.$$

Здесь базисный минор расположен в первых двух столбцах, главные неизвестные x_1, x_2, свободные неизвестные x_3, x_4. Систему, эквивалентную исходной, можно записать в виде

$$\begin{cases} x_1 = 3x_3 + x_4, \\ x_2 = 2x_3 + x_4. \end{cases} \tag{4.34}$$

В качестве ненулевого определителя второго порядка проще всего взять единичный определитель $\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}$. Выбор первой строки определителя означает, что свободные неизвестные принимают значения $x_3 = 1$, $x_4 = 0$. В этом случае $x_1 = 3$, $x_2 = 2$. Для второй строки определителя $x_3 = 0$, $x_4 = 1$ и $x_1 = 1$, $x_2 = 1$. Итак, получены два решения $(3, 2, 1, 0)^{\text{т}}$, $(1, 1, 0, 1)^{\text{т}}$, которые составляют фундаментальную систему решений.

Если выбрать другой определитель $\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}$, то имеем: для первой строки $x_3 = 1$, $x_4 = 2$ и $x_1 = 5$, $x_2 = 4$; для второй строки $x_3 = 3$, $x_4 = 4$ и $x_1 = 13$, $x_2 = 10$. В результате получаем другую фундаментальную систему решений $(5, 4, 1, 2)^{\text{т}}$ и $(13, 10, 3, 4)^{\text{т}}$.

Из рассмотренного примера видно, что если при построении фундаментальной системы решений в качестве ненулевого определителя выбрать определитель единичной матрицы, то вычисление столбцов фундаментальной системы решений наиболее просто. Предположим, что найдено общее решение однородной системы в виде

$$\begin{cases} x_1 = \alpha_{1,r+1}x_{r+1} + \ldots + \alpha_{1n}x_n, \\ x_2 = \alpha_{2,r+1}x_{r+1} + \ldots + \alpha_{2n}x_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_r = \alpha_{r,r+1}x_{r+1} + \ldots + \alpha_{rn}x_n, \end{cases}$$

где $x_1, \ldots, x_r$ — главные неизвестные, а $x_{r+1}, \ldots, x_n$ — свободные неизвестные. Допишем к этой системе формальные уравнения $x_{r+1} = x_{r+1}$, $x_{r+2} = x_{r+2}, \ldots, x_n = x_n$ и расширенную систему запишем в векторной форме:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \\ x_{r+1} \\ x_{r+2} \\ \vdots \\ x_n \end{pmatrix} = x_{r+1}\begin{pmatrix} \alpha_{1,r+1} \\ \alpha_{2,r+1} \\ \vdots \\ \alpha_{r,r+1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_{r+2}\begin{pmatrix} \alpha_{1,r+2} \\ \alpha_{2,r+2} \\ \vdots \\ \alpha_{r,r+2} \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \ldots + x_n\begin{pmatrix} \alpha_{1n} \\ \alpha_{2n} \\ \vdots \\ \alpha_{rn} \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Правая часть векторного равенства представляет собой линейную комбинацию $n-r$ вектор-столбцов, которые и составляют фундаментальную систему решений рассматриваемой однородной системы.

В рассмотренном примере систему (4.34) дополним формальными уравнениями $x_3 = x_3$, $x_4 = x_4$ и запишем в векторной форме:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = x_3 \begin{pmatrix} 3 \\ 2 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}.$$

Отсюда получаем фундаментальную систему решений

$$(3,\ 2,\ 1,\ 0)^{\mathrm{T}}, \quad (1,\ 1,\ 0,\ 1)^{\mathrm{T}}.$$

4.10. Связь между решениями однородной и неоднородной систем

Пусть дана неоднородная система $Ax = b$. Однородную систему линейных уравнений $Ax = 0$, получающуюся из неоднородной системы заменой в ней свободных членов нулями, называют ***приведенной однородной системой*** для системы $Ax = b$.

Между решениями неоднородной и приведенной однородной систем существует тесная связь, которая описывается следующими утверждениями.

Теорема 4.17. Сумма любого решения x_1 неоднородной системы $Ax = b$ и любого решения x_2 приведенной однородной системы $Ax = 0$ является решением неоднородной системы $Ax = b$.

$\triangleright$ В силу равенств

$$Ax_1 = b \quad \text{и} \quad Ax_2 = 0$$

заключаем, что

$$A(x_1 + x_2) = Ax_1 + Ax_2 = b + 0 = b,$$

а это означает, что столбец $x_1 + x_2$ является решением неоднородной системы $Ax = b$. ▶

Теорема 4.18. *Разность $x_1 - x_2$ любых двух решений x_1 и x_2 неоднородной системы $Ax = b$ является решением ее приведенной системы $Ax = 0$.*

▷ Из равенств

$$Ax_1 = b \quad \text{и} \quad Ax_2 = b$$

получаем:

$$A(x_1 - x_2) = Ax_1 - Ax_2 = b - b = 0,$$

что и означает, что столбец $x_1 - x_2$ является решением однородной системы $Ax = 0$. ▶

Теорема 4.19. *Общее решение неоднородной системы $Ax = b$ можно представить формулой*

$$x = x_{\text{одн}} + x_0, \tag{4.35}$$

где $x_{\text{одн}}$ — общее решение приведенной однородной системы $Ax = 0$, а x_0 — какое-либо частное решение неоднородной системы $Ax = b$.

▷ В силу теоремы 4.17 любой вектор x, имеющий представление (4.35), является решением неоднородной системы $Ax = b$. Пусть x — произвольное решение системы $Ax = b$. Тогда по теореме 4.18 вектор $y = x - x_0$ является решением однородной системы $Ax = 0$. Следовательно, $x = y + x_0$ и вектор x содержится в множестве решений, определяемых формулой (4.35). ▶

Формула (4.35) позволяет находить общее решение неоднородной системы при известном ее частном решении, решая приведенную однородную систему.

Пример 4.17. Зная решение $x_0 = (2,\ 0,\ -1,\ 0)^{\mathrm{T}}$ системы

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 1, \\ x_1 + x_2 + 2x_3 + x_4 = 0, \\ x_1 + x_2 - x_3 + x_4 = 3, \end{cases}$$

найти общее решение этой системы.

Решение. Поскольку известно частное решение системы, можно ограничиться определением общего решения приведенной однородной системы. Решая эту систему, получаем общее решение в виде

$$x_{\text{одн}} = \begin{pmatrix} -x_2 - x_4 \\ x_2 \\ 0 \\ x_4 \end{pmatrix} = x_2 \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Теперь, используя формулу (4.35), можем записать общее решение неоднородной системы

$$x = x_{\text{одн}} + x_0 = x_2 \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 2 \\ 0 \\ -1 \\ 0 \end{pmatrix}.$$

4.11. Линейные подпространства

Подмножество L линейного пространства X над полем P называют ***линейным подпространством*** этого пространства, если оно само является линейным пространством относительно введенных в X операций сложения векторов и умножения векторов на числа из поля P.

Для того чтобы подмножество L линейного пространства X было его линейным подпространством, необходимо и достаточно, чтобы оно было замкнуто относительно операций сложения векторов и умножения векторов на числа из поля P, т.е. чтобы выполнялись следующие условия:

1) для любых векторов a и b из L их сумма $a+b$ также принадлежит L;

2) для любого вектора a из L и любого числа $\alpha \in P$ вектор αa принадлежит L.

Действительно, эти условия означают, что операции в линейном пространстве X можно рассматривать и как операции на множестве L. При этом будут верны все аксиомы, кроме третьей и четвертой. Выполнение третьей аксиомы равносильно утверждению, что при выполнении условий 1 и 2 нулевой вектор принадлежит L, а выполнение четвертой означает, что для любого вектора $x \in L$ противоположный вектор $-x$ также принадлежит L. Первое из этих утверждений следует из равенства $0 = 0 \cdot x$, где в качестве x можно взять любой вектор в множестве L. Второе утверждение — следствие равенства $-x = (-1)x$.

Условия 1 и 2 можно объединить в одно условие: для любых векторов множества L их линейная комбинация с произвольными коэффициентами принадлежит L.

Приведем примеры линейных подпространств.

1. Множество векторов на прямой или на плоскости является подпространством в обычном трехмерном пространстве.

2. Множество многочленов степени не выше второй является подпространством в линейном пространстве многочленов степени не выше третьей.

3. Множество решений однородной системы линейных уравнений с n неизвестными является подпространством в линейном n-мерном арифметическом пространстве K_n.

Пусть в линейном пространстве X над полем P дана система векторов $a_1, a_2, \ldots, a_k$. Множество всевозможных линейных комбинаций $\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_k a_k$ этой системы называют ***линейной оболочкой*** системы векторов $a_1, a_2, \ldots, a_k$.

Линейная оболочка L системы векторов $a_1, a_2, \ldots, a_k$ является подпространством в X. Действительно, если векторы a

и b принадлежат L, т.е. имеют представления

$$a = \alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_k a_k, \quad b = \beta_1 a_1 + \beta_2 a_2 + \ldots + \beta_k a_k,$$

то и векторы $a + b$ и λa имеют такие представления:

$$a + b = (\alpha_1 + \beta_1)a_1 + (\alpha_2 + \beta_2)a_2 + \ldots + (\alpha_k + \beta_k)a_k,$$
$$\lambda a = (\lambda\alpha_1)a_1 + (\lambda\alpha_2)a_2 + \ldots + (\lambda\alpha_k)a_k.$$

Следовательно, они принадлежат L.

Линейную оболочку L системы векторов $a_1, a_2, \ldots, a_k$ также называют подпространством, порожденным этой системой векторов, или подпространством, натянутым на эту систему векторов, и обозначают

$$L = \langle a_1,\, a_2,\, \ldots,\, a_n \rangle.$$

Теорема 4.20. *Линейное подпространство конечномерного линейного пространства является конечномерным, и размерность подпространства не превышает размерности всего линейного пространства.*

$\triangleright$ Действительно, размерность конечномерного линейного пространства X может быть определена как максимальное количество линейно независимых векторов в этом пространстве. Очевидно, что максимальное количество линейно независимых векторов в любом подмножестве L в X не превышает максимального количества линейно независимых векторов в X. Отсюда утверждение теоремы. $\blacktriangleright$

Любое конечномерное линейное пространство порождается конечной системой векторов, например любым своим базисом. Согласно доказанной теореме это верно и для всякого линейного подпространства линейного пространства.

Теорема 4.21. *Пусть L — подпространство n-мерного линейного пространства X. Любой базис $a_1, a_2, \ldots, a_k$ в L можно дополнить до базиса $a_1, a_2, \ldots, a_k, a_{k+1}, \ldots,$*

a_n всего линейного пространства X, причем линейному подпространству L принадлежат те и только те векторы, которые в указанном базисе имеют столбцы координат вида

$$\left(x_1', x_2', \ldots, x_k', 0, \ldots, 0\right)^{\mathrm{T}}. \tag{4.36}$$

$\triangleright$ Действительно, любой базис в L, как и вообще любую линейно независимую систему векторов в X, можно дополнить до базиса линейного пространства X. Если вектор x имеет столбец координат (4.36), то он имеет представление

$$x = x_1'a_1 + x_2'a_2 + \ldots x_k'a_k + 0 \cdot a_{k+1} + \ldots + 0 \cdot a_n = \\ = x_1'a_1 + x_2'a_2 + \ldots x_k'a_k$$

и, следовательно, принадлежит L как линейная комбинация векторов, принадлежащих L. Если вектор x принадлежит L, то он может быть разложен по базису $a_1, a_2, \ldots, a_k$:

$$x = x_1'a_1 + x_2'a_2 + \ldots x_k'a_k.$$

Это разложение в то же время является разложением вектора x в базисе $a_1, a_2, \ldots a_n$ линейного пространства X, дающим столбец координат вида (4.36). $\blacktriangleright$

Теорема 4.22. *Пусть X — конечномерное линейное пространство и в нем задан базис. Тогда для любого линейного подпространства L в X можно указать такую однородную систему линейных уравнений $Ax = 0$, что множество координатных столбцов всех векторов L будет совпадать с множеством решений системы $Ax = 0$.*

$\triangleright$ Пусть $e_1, e_2, \ldots, e_n$ — заданный базис в X. Выберем в L некоторый базис в $a_1, a_2, \ldots, a_k$ и дополним его векторами $a_{k+1}, a_{k+2}, \ldots, a_n$ до базиса в X. Из теоремы 4.21 вытекает, что множество векторов линейного подпространства L в базисе $a_1, \ldots, a_n$ описывается системой уравнений $x_{k+1}' = \ldots \ldots = x_n' = 0$.

Пусть T — матрица перехода от базиса $e_1, e_2, \ldots, e_n$ к базису $a_1, a_2, \ldots, a_n$. Тогда столбец координат x_a вектора в базисе $a_1, a_2, \ldots, a_n$ связан со столбцом координат x_e этого же вектора в базисе $e_1, e_2, \ldots, e_n$ равенством $x_e = Tx_a$. Если столбец координат x_a является решением однородной системы $Ax = 0$, то в силу соотношения $x_a = T^{-1}x_e$ столбец координат x_e является решением однородной системы $AT^{-1}x = 0$. Верно и обратное: если x_e есть решение системы $AT^{-1}x = 0$, т.е. $AT^{-1}x_e = 0$, то $Ax_a = 0$, т.е. столбец x_a является решением системы $Ax = 0$. Из этих рассуждений вытекает следующее: если некоторое множество в заданном базисе описывается однородной системой линейных уравнений (иными словами, совокупность всех столбцов координат векторов множества совпадает с множеством решений системы), то и в любом другом базисе это множество описывается некоторой системой линейных уравнений.

Так как линейное подпространство L в базисе $a_1, a_2, \ldots, a_n$ описывается однородной системой линейных уравнений, то это подпространство в соответствии с только что доказанным и в базисе $e_1, e_2, \ldots, e_n$ описывается системой линейных уравнений. ▶

Однородную систему линейных уравнений, описывающую данное линейное подпространство L, называют ***общими уравнениями*** этого подпространства.

Пример 4.18. Составить общие уравнения линейного подпространства $L = \langle a_1, a_2 \rangle$ в четырехмерном линейном пространстве X, если векторы a_1 и a_2 заданы своими координатами в некотором базисе e: $[a_1]_e = (1, 1, 2, 0)^{\mathrm{T}}$, $[a_2]_e = (1, -1, 0, 2)^{\mathrm{T}}$.

Решение. Нетрудно убедиться в том, что система векторов a_1 и a_2 линейно независима и потому составляет базис в линейном подпространстве L. Дополним эту систему до базиса в линейном пространстве X векторами a_3 и a_4, в качестве которых можно взять пару векторов из базиса e: $[a_3]_e = (0, 0, 1, 0)^{\mathrm{T}}$, $[a_4]_e = (0, 0, 0, 1)^{\mathrm{T}}$. Матрицей перехода от базиса e к новому базису $a = (a_1,\ a_2,\ a_3,\ a_4)$ является матрица

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

обратной к которой является матрица

$$T^{-1} = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & -1/2 & 0 & 0 \\ -1 & -1 & 1 & 0 \\ -1 & 1 & 0 & 1 \end{pmatrix}.$$

В базисе a линейное подпространство описывается однородной системой из двух уравнений $x'_3 = 0,\ x'_4 = 0$, которая в матричной форме имеет вид:

$$\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \\ x'_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Применив формулу $x_a = T^{-1}x_e$, полученную систему преобразуем в систему

$$\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & -1/2 & 0 & 0 \\ -1 & -1 & 1 & 0 \\ -1 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

а после умножения матриц — в систему

$$\begin{pmatrix} -1 & -1 & 1 & 0 \\ -1 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

или

$$\begin{cases} -x_1 - x_2 + x_3 = 0, \\ -x_1 + x_2 + x_4 = 0, \end{cases}$$

описывающую линейное подпространство L в базисе e.

Общие уравнения линейного подпространства определяются неоднозначно: достаточно систему линейных уравнений заменить любой эквивалентной системой, чтобы получить другие общие уравнения того же линейного подпространства. В рассмотренном примере ответ зависит от того, какими векторами a_3 и a_4 мы дополняем систему a_1, a_2 до базиса. Например, если положить $[a_3]_e = (0,\ 0,\ -1,\ -1)^{\text{т}}$, $[a_4]_e = (0,\ 0,\ -1,\ 1)^{\text{т}}$, то, повторив все вычисления, получим систему линейных уравнений

$$\begin{cases} x_1 - \dfrac{1}{2}x_3 - \dfrac{1}{2}x_4 = 0, \\ x_2 - \dfrac{1}{2}x_3 + \dfrac{1}{2}x_4 = 0. \end{cases}$$

Если подпространство задано общими уравнениями, то для построения базиса этого подпространства следует построить фундаментальную систему решений для общих уравнений подпространства.

Пример 4.19. Найти какой-либо базис подпространства L, заданного системой уравнений

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

Решение. Выбрав в качестве главных неизвестные x_1, x_2, а свободных — x_3, x_4, решим систему. В результате получим:

$$x = \begin{pmatrix} -x_3 \\ -x_4 \\ x_3 \\ x_4 \end{pmatrix} = x_3 \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 0 \\ -1 \\ 0 \\ 1 \end{pmatrix}.$$

Фундаментальную систему решений рассматриваемой однородной системы линейных уравнений составляют столбцы $(-1,\ 0,\ 1,\ 0)^{\text{т}}$, $(0,\ -1,\ 0,\ 1)^{\text{т}}$, а базис линейного пространства — векторы, которые в заданном базисе имеют указанные столбцы координат.

Если $a_1, a_2, \ldots, a_k$ — базис линейного подпространства L в линейном пространстве X, то L можно задать уравнением

$$x = t_1 a_1 + t_2 a_2 + \ldots + t_n a_n, \tag{4.37}$$

которое называют ***параметрическим уравнением подпространства** L **в векторной форме.***

Пусть векторы $a_1, a_2, \ldots, a_k$ заданы своими координатами в некотором базисе пространства X:

$$a_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}, \quad a_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix}, \quad \ldots, \quad a_k = \begin{pmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{nk} \end{pmatrix}.$$

Тогда векторное уравнение (4.37) в координатах можно записать следующим образом:

$$\begin{cases} x_1 = a_{11}t_1 + a_{12}t_2 + \ldots + a_{1k}t_k, \\ x_2 = a_{21}t_1 + a_{22}t_2 + \ldots + a_{2k}t_k, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n = a_{n1}t_1 + a_{n2}t_2 + \ldots + a_{nk}t_k. \end{cases} \tag{4.38}$$

Систему (4.38) называют ***параметрическими уравнениями подпространства** L **в координатной форме.***

Если из параметрических уравнений (4.38) подпространства L исключить параметры $t_1, t_2, \ldots, t_k$, получим общие уравнения подпространства L. Таким образом, мы пришли еще к одному способу получения общих уравнений подпространства.

Пример 4.20. Подпространство $L = \langle a_1, a_2 \rangle$, где $a_1 = (1, 1, 2, 0)^{\mathrm{T}}$, $a_2 = (1, -1, 0, 2)^{\mathrm{T}}$, задать параметрическими уравнениями (4.37) и (4.38) и общими уравнениями.

Решение. Векторное уравнение (4.37) в данном случае имеет вид:

$$x = t_1 \begin{pmatrix} 1 \\ 1 \\ 2 \\ 0 \end{pmatrix} + t_2 \begin{pmatrix} 1 \\ -1 \\ 0 \\ 2 \end{pmatrix}.$$

Переходя к координатам, получаем координатную форму параметрических уравнений

$$\begin{cases} x_1 = t_1 + t_2, \\ x_2 = t_1 - t_2, \\ x_3 = 2t_1, \\ x_4 = 2t_2. \end{cases}$$

Исключив параметры t_1 и t_2, получим общие уравнения подпространства L:

$$\begin{cases} 2x_1 - x_3 - x_4 = 0, \\ 2x_2 - x_3 + x_4 = 0. \end{cases}$$

Пусть в линейном пространстве X даны подпространства L_1 и L_2. Множество $L_1 \cap L_2$ векторов, принадлежащих как L_1, так и L_2, является подпространством в X. Его называют ***пересечением подпространств*** L_1 и L_2.

Множество всех векторов x вида $x = a + b$, где $a \in L_1$, $b \in L_2$, называют ***суммой подпространств*** L_1 и L_2 и обозначают через $L_1 + L_2$. Если при этом пересечение $L_1 \cap L_2$ — нулевое подпространство, то сумму $L_1 + L_2$ называют ***прямой суммой*** и обозначают через $L_1 \oplus L_2$.

Сумма подпространств является подпространством. Действительно, пусть $x = a + b$, $y = c + d$, где $a, c \in L_1$, $b, d \in L_2$. Тогда $x + y = (a + c) + (b + d) \in L_1 + L_2$, поскольку $a + c \in L_1$ и $b + d \in L_2$. Аналогично для любого числа α имеем: $\alpha x = \alpha a + \alpha b \in L_1 + L_2$, так как $\alpha a \in L_1$ и $\alpha b \in L_2$.

Понятия пересечения и суммы подпространств распространяются на любое число подпространств.

Если сумма $L_1 + L_2$ подпространств L_1 и L_2 в X является прямой, то представление любого вектора x в виде $x = a + b$, где $a \in L_1$, $b \in L_2$, единственно. В частном случае $X = L_1 \oplus L_2$ каждый вектор $x \in X$ имеет представление $x = a + b$, причем единственное. В этом случае подпространства L_1 и L_2 называют ***прямыми дополнениями*** друг друга, а слагаемое $a \in L_1$ — ***проекцией вектора*** x ***на подпространство*** L_1 ***параллельно подпространству*** L_2.

Пример 4.21. В пространстве $X = K_4$ построить какое-либо прямое дополнение L_2 к подпространству $L_1 = \langle a_1, a_2\rangle$, где $a_1 = (1, 1, 1, 0)^{\mathrm{T}}$, $a_2 = (1, 0, 1, 0)^{\mathrm{T}}$. Найти проекцию вектора $x = (2, 1, 5, 5)^{\mathrm{T}}$ на L_1 параллельно L_2.

Решение. Векторы a_1 и a_2 линейно независимы и поэтому составляют базис в подпространстве L_1. Дополним систему векторов a_1, a_2 до базиса во всем пространстве X, например, векторами $b_1 = (0, 0, 1, 0)^{\mathrm{T}}$ и $b_2 = (0, 0, 0, 1)^{\mathrm{T}}$ и положим $L_2 = \langle b_1, b_2\rangle$. Очевидно, что L_2 является искомым подпространством. Далее запишем векторное равенство

$$x = (\alpha_1 a_1 + \alpha_2 a_2) + (\beta_1 b_1 + \beta_2 b_2),$$

перейдем от него к покоординатным уравнениям (см. разд. 4.2) и, решив систему этих уравнений, найдем: $\alpha_1 = \alpha_2 = 1$, $\beta_1 = 3$, $\beta_2 = 5$. Поэтому

$$x = (a_1 + a_2) + (3b_1 + 5b_2) = \begin{pmatrix} 2 \\ 1 \\ 2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 3 \\ 5 \end{pmatrix},$$

где $(2, 1, 2, 0)^{\mathrm{T}} \in L_1$, $(0, 0, 3, 5)^{\mathrm{T}} \in L_2$. Следовательно, проекцией вектора $x = (2, 1, 5, 5)^{\mathrm{T}}$ на подпространство L_1 параллельно подпространству L_2 является вектор $x_1 = (2, 1, 2, 0)^{\mathrm{T}}$.

Пусть $L_1 = \langle a_1, a_2, \ldots, a_k\rangle$, $L_2 = \langle b_1, b_2, \ldots, b_l\rangle$ — подпространства в линейном пространстве X. Чтобы найти какой-либо базис в подпространстве $L_1 + L_2$, следует выделить какую-либо максимальную линейно независимую подсистему системы векторов $a_1, a_2, \ldots, a_k, b_1, b_2, \ldots, b_l$. Для этого достаточно составить матрицу из координатных столбцов этих векторов и в этой матрице выделить какой-либо базисный минор. Векторы, на координатных столбцах которых находится базисный минор, образуют базис в подпространстве $L_1 + L_2$. Отметим, что базисный минор можно выбирать не в исходной, а преобразованной матрице (после выполнения последовательности элементарных преобразований строк).

Пример 4.22. Найти базис суммы $L_1 + L + 2$ подпространств $L_1 = \langle a_1, a_2, a_3 \rangle$ и $L_2 = \langle b_1, b_2, b_3 \rangle$, если $a_1 = (1, 1, 1, 1)^{\text{т}}$, $a_2 = (1, 1, -1, -1)^{\text{т}}$, $a_3 = (1, -1, 1, -1)^{\text{т}}$ и $b_1 = (1, -1, -1, 1)^{\text{т}}$, $b_2 = (2, -2, 0, 0)^{\text{т}}$, $b_3 = (3, -1, 1, 1)^{\text{т}}$.

Решение. Составим матрицу

$$(a_1, a_2, a_3, b_1, b_2, b_3) = \begin{pmatrix} 1 & 1 & 1 & 1 & 2 & 3 \\ 1 & 1 & -1 & -1 & -2 & -1 \\ 1 & -1 & 1 & -1 & 0 & 1 \\ 1 & -1 & -1 & 1 & 0 & 1 \end{pmatrix}.$$

Проводя элементарные преобразования строк матрицы, приведем ее к ступенчатому виду:

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

Видим, что ранг матрицы равен четырем, а один из ее базисных миноров располагается на векторах a_1, a_2, a_3, b_1. Следовательно, эти векторы составляют базис суммы $L_1 + L_2$.

Если пространства L_1 и L_2 заданы однородными системами уравнений, то пересечение $L_1 \cap L_2$ будет определяться системой, получаемой объединением всех уравнений двух систем. Любая фундаментальная система решений такой системы уравнений дает базис пересечения $L_1 \cap L_2$.

Пример 4.23. Найти базис пересечения подпространства L_1, заданного системой уравнений

$$\begin{cases} x_1 - x_3 + x_5 = 0, \\ x_2 - x_4 + x_6 = 0, \\ x_1 - x_2 + x_5 - x_6 = 0, \end{cases}$$

и подпространства L_2, заданного системой уравнений

$$\begin{cases} x_2 - x - 3 + x_6 = 0, \\ x_1 - x_4 + x_5 = 0. \end{cases}$$

Решение. Составим объединенную систему уравнений

$$\begin{cases} x_1 - x_3 + x_5 = 0, \\ x_2 - x_4 + x_6 = 0, \\ x_1 - x_2 + x_5 - x_6 = 0, \\ x_2 - x - 3 + x_6 = 0, \\ x_1 - x_4 + x_5 = 0 \end{cases}$$

и найдем ее общее решение

$$x = (x_4 - x_5,\ x_4 - x_6,\ x_4,\ x_4,\ x_5,\ x_6)^{\mathrm{T}}.$$

Здесь три свободных неизвестных: x_4, x_5, x_6. Поэтому каждая фундаментальная система решений этой системы состоит из трех решений. Одну из фундаментальных систем решений составляют столбцы

$$\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ -1 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Они и представляют собой один из базисов подпространства $L_1 \cap L_2$.

Если подпространства L_1 и L_2 заданы как линейные оболочки систем векторов

$$L_1 = \langle a_1,\ a_2,\ \ldots,\ a_k \rangle, \quad L_2 = \langle b_1,\ b_2,\ \ldots,\ b_l \rangle,$$

то при построении базиса пересечения $L_1 \cap L_2$ этих подпространств достаточно перейти к описанию этих подпространств общими уравнениями, а затем действовать, как в последнем примере: объединяя две однородные системы в одну, искать фундаментальную систему решений объединенной системы.

Существуют и другие способы построения базиса пересечения. Например (см. [21], решение задачи 1319), можно составить векторное уравнение

$$\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_k a_k = \beta_1 b_1 + \beta_2 b_2 + \ldots + \beta_l b_l \tag{4.39}$$

с неизвестными $\alpha_1, \alpha_2, \ldots, \alpha_k, \beta_1, \beta_2, \ldots, \beta_l$ и от него перейти к системе покоординатных уравнений. Это линейная однородная система. Построив фундаментальную систему решений этой системы, для каждого решения из ФСР вычислим, например, левую часть векторного уравнения. Получим систему векторов, порождающую линейное пространство $L_1 \cap L_2$. Теперь базис в $L_1 \cap L_2$ можно построить, выделив в этой системе максимальную линейно независимую подсистему. Отметим, что если система векторов $a_1, a_2, \ldots, a_k$ линейно независима, то и построенная, как описано выше, система векторов, порождающая $L_1 \cap L_2$, будет линейно независимой. В этом случае дополнительно выделять максимальную линейно независимую подсистему не нужно.

Пример 4.24. Найти базис пересечения подпространств $L_1 = \langle a_1, a_2, a_3 \rangle$ и $L_2 = \langle b_1, b_2, b_3 \rangle$, где $a_1 = (1,\ 1,\ 1,\ 1)^{\mathrm{T}}$, $a_2 = (1,\ 1,\ -1,\ -1)^{\mathrm{T}}$, $a_3 = (1,\ -1,\ 1,\ -1)^{\mathrm{T}}$ и $b_1 = (1,\ -1,\ -1,\ 1)^{\mathrm{T}}$, $b_2 = (2,\ -2,\ 0,\ 0)^{\mathrm{T}}$, $b_3 = (3,\ -1,\ 1,\ 1)^{\mathrm{T}}$.

Решение. Сначала используем первый способ, переходя к общим уравнениям подпространств. Подпространство L_1 описывается параметрическим уравнением $x = t_1 a_1 + t_2 a_2 + t_3 a_3$, которое в координатной форме имеет вид:

$$\begin{cases} x_1 = t_1 + t_2 + t_3, \\ x_2 = t_1 + t_2 - t_3, \\ x_3 = t_1 - t_2 + t_3, \\ x_4 = t_1 - t_2 - t_3. \end{cases}$$

Исключив параметры t_1, t_2, t_3, придем к общему уравнению $x_1 - x_2 - x_3 + x_4 = 0$ подпространства L_1. Аналогично получаем общее уравнение $x_1 - x_2 - x_3 - x_4 = 0$ подпространства L_2. Общие уравнения подпространства $L_1 \cap L_2$ имеют вид:

$$\begin{cases} x_1 - x - 2 - x_3 + x_4 = 0, \\ x_1 - x - 2 - x_3 - x_4 = 0. \end{cases}$$

Фундаментальную систему решений этой системы составляют, например, векторы $x_1 = (1,\ 0,\ 1,\ 0)^{\text{т}}$, $x_2 = (0,\ 1,\ 0,\ 1)^{\text{т}}$. Эти векторы образуют базис в подпространстве $L_1 \cap L_2$.

Теперь применим второй способ решения примера. Составим векторное уравнение

$$\alpha_1 a_1 + \alpha_2 a_2 + \alpha_3 a_3 = \beta_1 b_1 + \beta_2 b_2 + \beta_3 b_3,$$

в подробной записи имеющее вид:

$$\alpha_1 \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix} + \alpha_3 \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} =$$

$$= \beta_1 \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} + \beta_2 \begin{pmatrix} 2 \\ -2 \\ 0 \\ 0 \end{pmatrix} + \beta_3 \begin{pmatrix} 3 \\ -1 \\ 1 \\ 1 \end{pmatrix}.$$

Переходя к покоординатным уравнениям, получим однородную систему

$$\begin{cases} \alpha_1 + \alpha_2 + \alpha_3 = \beta_1 + 2\beta_2 + 3\beta_3, \\ \alpha_1 + \alpha_2 - \alpha_3 = -\beta_1 - 2\beta_2 - \beta_3, \\ \alpha_1 - \alpha_2 + \alpha_3 = -\beta_1 + \beta_3, \\ \alpha_1 - \alpha_2 - \alpha_3 = \beta_1 + \beta_3 \end{cases}$$

с шестью неизвестными. Ее общее решение таково:

$$\begin{cases} \alpha_1 = \beta_1, \\ \alpha_2 = 0, \\ \alpha_3 = \beta_2 + \beta_3, \\ \beta_1 = -\beta_2 - \beta_3. \end{cases}$$

Свободных неизвестных два, и фундаментальная система решений состоит из двух столбцов

$$(0,\ 0,\ 1,\ -1,\ 1,\ 0)^{\mathrm{T}} \quad \text{и} \quad (1,\ 0,\ 1,\ -1,\ 0,\ 1)^{\mathrm{T}}.$$

В этих двух столбцах выбираем первые три компоненты и принимаем в качестве значений α_1, α_2, α_3 в выражении $\alpha_1 a_1 + \alpha_2 a_2 + \alpha_3 a_3$. Получаем два вектора

$$x_1 = 0\cdot a_1 + 0\cdot a_2 + 1\cdot a_3 = a_3 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix},$$

$$x_2 = 1\cdot a_1 + 0\cdot a_2 + 1\cdot a_3 = a_1 + a_3 = \begin{pmatrix} 2 \\ 0 \\ 2 \\ 0 \end{pmatrix},$$

составляющих базис подпространства $L_1 \cap L_2$.

Теорема 4.23. *В конечномерном линейном пространстве X размерность суммы $L_1 + L_2$ подпространств L_1 и L_2 равна сумме размерностей этих подпространств минус размерность их пересечения, т.е.*

$$\dim(L_1 + L_2) = \dim L_1 + \dim L_2 - \dim(L_1 \cap L_2).$$

$\triangleright$ В подпространстве $L_1 \cap L_2$ выберем какой-либо базис $e = (e_1,\ e_2,\ \ldots,\ e_k)$. Эта система векторов линейно независима и одновременно принадлежит и L_1, и L_2. Дополним ее до базиса в L_1 системой векторов $f = (f_1,\ f_2,\ \ldots,\ f_l)$ и до базиса в L_2 системой векторов $g = (g_1,\ g_2,\ \ldots,\ g_m)$. Из трех систем векторов составим объединенную систему

$$(e,\ f,\ g) = (e_1,\ e_2,\ \ldots,\ e_k,\ f_1,\ f_2,\ \ldots,\ f_l,\ g_1,\ g_2,\ \ldots,\ g_m) \qquad (4.40)$$

и докажем, что она является базисом в $L_1 + L_2$.

Через систему векторов (e, f, g) линейно выражается любой вектор $z \in L_1 + L_2$. Действительно, для вектора z имеет место представление $z = x + y$, где $x \in L_1$, $y \in L_2$. Вектор x линейно выражается через систему (e, f), а вектор y — через систему (e, g). Поэтому их сумма z линейно выражается через объединенную систему (e, f, g).

Система векторов (e, f, g) линейно независима. Чтобы доказать это, запишем равенство

$$\alpha_1 e_1 + \ldots + \alpha_k e_k + \beta_1 f_1 + \ldots + \beta_l f_l + \gamma_1 g_1 + \ldots + \gamma_m g_m = 0 \qquad (4.41)$$

и покажем, что оно возможно только при нулевых значениях всех коэффициентов. В равенстве (4.41) объединим слагаемые, относящиеся к векторам систем e и f:

$$a = \alpha_1 e_1 + \ldots + \alpha_k e_k + \beta_1 f_1 + \ldots + \beta_k f_l. \qquad (4.42)$$

Вектор a принадлежит подпространству L_1. Но из равенства (4.41) следует, что

$$a = -\gamma_1 g_1 - \ldots - \gamma_m g_m$$

и вектор a принадлежит подпространству L_2. Значит, $a \in \in L_1 \cap L_2$. Но из этого условия вытекает, что вектор a линейно выражается через систему векторов e, т.е. имеет место представление $a = \mu_1 e_1 + \ldots + \mu_k e_k$. Это представление можно рассматривать как разложение вектора $a \in L_1$ по базису (e, f). В силу единственности разложения по базису заключаем, что оба разложения совпадают, т.е. $\mu_1 = \alpha_1$, ..., $\mu_k = \alpha_k$, $\beta_1 = \ldots = \beta_l = 0$. С учетом полученных соотношений равенство (4.41) принимает вид:

$$\alpha_1 e_1 + \ldots + \alpha_k e_k + \gamma_1 g_1 + \ldots + \gamma_m g_m = 0.$$

Поскольку система векторов (e, g) линейно независима (как базис в L_2), это равенство возможно лишь при нулевых значениях всех коэффициентов:

$$\alpha_1 = \ldots = \alpha_k = \gamma_1 = \ldots = \gamma_k = 0.$$

Таким образом, доказано, что равенство (4.41) выполняется лишь при нулевых значениях всех коэффициентов, а система $(t,\ f,\ g)$ линейно независима и является базисом в подпространстве $L_1 + L_2$. Число векторов в этом базисе, а потому и размерность пространства $L_1 + L_2$, равна $k + l + m$. Поскольку $\dim L_1 = k + l$, $\dim L_2 = k + m$, $\dim(L_1 \cap L_2) = k$, то

$$\dim(L_1 + L_2) = k + l + m = (k + l) + (k + m) - k = \dim L_1 + \dim L_2 - \dim(L_1 \cap L_2). \quad \blacktriangleright$$

Упражнения

4.1. Найти линейную комбинацию векторов:

1) $3a_1 - 2a_2 + 8a_3$, если

$$a_1 = (1,\ 2,\ 1,\ 2)^{\mathrm{T}}, \quad a_2 = (-1,\ -3,\ 4,\ 5)^{\mathrm{T}}, \quad a_3 = (-5,\ 0,\ 2,\ 3)^{\mathrm{T}};$$

2) $2a_1 + 3a_2 - 8a_3 + 4a_4$, если

$$a_1 = (1,\ -1,\ 2,\ -1,\ 1)^{\mathrm{T}}, \quad a_2 = (3,\ 1,\ 1,\ -3,\ 4)^{\mathrm{T}},$$
$$a_3 = (3,\ 1,\ -1,\ 2,\ 4)^{\mathrm{T}}, \quad a_4 = (-5,\ -2,\ -3,\ 1,\ 2)^{\mathrm{T}};$$

3) $5a_1 - 6a_2 + 7a_3$, если

$$a_1 = (1,\ -1,\ 2,\ -2)^{\mathrm{T}}, \quad a_2 = (1,\ 1,\ -1,\ -1)^{\mathrm{T}}, \quad a_3 = (3,\ 0,\ -1,\ 2)^{\mathrm{T}}.$$

4.2. Решить уравнения:

1) $2a_1 + 3a_2 - a_3 - 7x = a_4$, где

$$a_1 = (-1,\ 2,\ -3,\ 4)^{\mathrm{T}}, \quad a_2 = (-1,\ -1,\ -1,\ 5)^{\mathrm{T}},$$
$$a_3 = (2,\ -5,\ -1,\ 3)^{\mathrm{T}}, \quad a_4 = (2,\ 1,\ -2,\ 1)^{\mathrm{T}};$$

2) $3(a_1 - 2x) + 5(a_2 + a_3 - 3x) = 2(a_3 - 4x)$, где

$$a_1 = (4,\ 3,\ 1,\ 2)^{\mathrm{T}}, \quad a_2 = (2,\ -1,\ -3,\ 4)^{\mathrm{T}}, \quad a_3 = (-1,\ 4,\ -5,\ 3)^{\mathrm{T}};$$

3) $2(x - a_1 + a_3) - 5(x - 2a_2 - a_3) + 3(2x + a_3 + a_4) = x - a_4$, где

$$a_1 = (1,\ 1,\ -1,\ -1)^{\mathrm{T}}, \quad a_2 = (3,\ 0,\ -5,\ 1)^{\mathrm{T}},$$
$$a_3 = (1,\ -3,\ 0,\ -4)^{\mathrm{T}}, \quad a_4 = (2,\ 3,\ 4,\ -5)^{\mathrm{T}}.$$

4.3. Используя правило из разд. 4.2, выяснить вопрос о линейной зависимости системы векторов:

1) $a_1 = (1,\ -1,\ 1,\ -1)^{\text{т}}$, $a_2 = (1,\ 0,\ 1,\ 0)^{\text{т}}$, $a_3 = (1,\ -3,\ 1,\ -3)^{\text{т}}$;
2) $a_1 = (1,\ 1,\ 1,\ 1)^{\text{т}}$, $a_2 = (1,\ -1,\ 1,\ -1)^{\text{т}}$, $a_3 = (2,\ 3,\ 1,\ 4)^{\text{т}}$, $a_4 = (2,\ 1,\ 1,\ 3)^{\text{т}}$;
3) $a_1 = (1,\ 2,\ 3)^{\text{т}}$, $a_2 = (2,\ 5,\ 7)^{\text{т}}$, $a_3 = (3,\ 7,\ 10)^{\text{т}}$;
4) $a_1 = (1,\ 2,\ 3)^{\text{т}}$, $a_2 = (2,\ 5,\ 7)^{\text{т}}$, $a_3 = (3,\ 7,\ 11)^{\text{т}}$;
5) $a_1 = (1,\ 1,\ 1,\ 1)^{\text{т}}$, $a_2 = (1,\ -1,\ 1,\ -1)^{\text{т}}$, $a_3 = (1,\ -1,\ 1,\ -1)^{\text{т}}$, $a_4 = (1,\ 1,\ -1,\ -1)^{\text{т}}$;
6) $a_1 = (1,\ 2,\ 3,\ 4)^{\text{т}}$, $a_2 = (4,\ 3,\ 2,\ 1)^{\text{т}}$, $a_3 = (5,\ 5,\ 5,\ 5)^{\text{т}}$;
7) $a_1 = (1,\ -1,\ 1,\ 1)^{\text{т}}$, $a_2 = (1,\ 0,\ 1,\ 0)^{\text{т}}$, $a_3 = (1,\ -3,\ 1,\ 3)^{\text{т}}$;
8) $a_1 = (1,\ 1,\ 1,\ 1)^{\text{т}}$, $a_2 = (1,\ -1,\ 1,\ -1)^{\text{т}}$, $a_3 = (2,\ 3,\ 1,\ 4)^{\text{т}}$, $a_4 = (2,\ 1,\ 1,\ 3)^{\text{т}}$;
9) $a_1 = (1,\ 2,\ 3,\ 4)^{\text{т}}$, $a_2 = (4,\ 1,\ 2,\ 3)^{\text{т}}$, $a_3 = (3,\ 4,\ 1,\ 2)^{\text{т}}$, $a_4 = (-1,\ -1,\ -1,\ -1)^{\text{т}}$.

4.4. Вычислить ранги матриц:

$$1)\ \begin{pmatrix} 2 & -4 & 3 & -3 & 5 \\ 1 & -2 & 1 & 5 & 3 \\ 1 & -2 & 4 & 34 & 0 \end{pmatrix};\quad 2)\ \begin{pmatrix} 2 & 0 & 2 & 0 & 2 \\ 0 & 1 & 0 & 1 & 0 \\ 2 & 1 & 0 & 2 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{pmatrix};$$

$$3)\ \begin{pmatrix} 1 & 2 & 3 & -2 \\ 2 & -3 & 1 & -4 \\ 1 & 9 & 8 & -2 \\ 1 & 12 & -7 & -2 \end{pmatrix};\quad 4)\ \begin{pmatrix} 2 & 1 & 11 & 2 \\ 1 & 0 & 4 & -1 \\ 11 & 0 & 56 & 5 \\ 2 & -1 & 5 & -6 \end{pmatrix};$$

$$5)\ \begin{pmatrix} 2 & 1 & 3 & -1 \\ 3 & -1 & 2 & 0 \\ 1 & 3 & 4 & -2 \\ 4 & -3 & 1 & 1 \end{pmatrix};\quad 6)\ \begin{pmatrix} 2 & 3 & 1 & -1 \\ 3 & 1 & 4 & 2 \\ 1 & 2 & 3 & -1 \\ 1 & -4 & -7 & 5 \end{pmatrix}.$$

4.5. Вычислить ранги систем векторов:

1) $a_1 = (1,\ 3,\ 1,\ -3)^{\text{т}}$, $a_2 = (2,\ 1,\ 1,\ 1)^{\text{т}}$, $a_3 = (3,\ -11,\ -1,\ 19)^{\text{т}}$, $a_4 = (1,\ 12,\ 2,\ -16)^{\text{т}}$;
2) $a_1 = (1,\ -2,\ 3,\ -1,\ -1)^{\text{т}}$, $a_2 = (2,\ -1,\ 1,\ 0,\ -2)^{\text{т}}$, $a_3 = (1,\ -1,\ -1,\ -1,\ 1)^{\text{т}}$, $a_4 = (1,\ 3,\ -10,\ 1,\ 3)^{\text{т}}$;

3) $a_1 = (3,\ 2,\ -1,\ 2,\ 0,\ 1)^{\mathrm{T}}$, $a_2 = (4,\ 1,\ 0,\ -3,\ 0,\ 2)^{\mathrm{T}}$,
$a_3 = (2,\ -1,\ -2,\ 1,\ 1,\ -3)^{\mathrm{T}}$, $a_4 = (3,\ 1,\ 3,\ -9,\ -1,\ 6)^{\mathrm{T}}$,
$a_5 = (3,\ -1,\ -5,\ 7,\ 2,\ -7)^{\mathrm{T}}$;

4) $a_1 = (2,\ 1,\ 1,\ 1)^{\mathrm{T}}$, $a_2 = (1,\ 3,\ 1,\ 1)^{\mathrm{T}}$, $a_3 = (1,\ 1,\ 5,\ 1)^{\mathrm{T}}$,
$a_4 = (1,\ -4,\ 4,\ 0)^{\mathrm{T}}$, $a_5 = (0,\ 1,\ -13,\ -1)^{\mathrm{T}}$, $a_6 = (2,\ 3,\ -3,\ 1)^{\mathrm{T}}$.

4.6. Вычислить ранги и указать всевозможные базы систем векторов:

1) $a_1 = (1,\ 2,\ 2)^{\mathrm{T}}$, $a_2 = (1,\ 2,\ 3)^{\mathrm{T}}$, $a_3 = (1,\ 2,\ -2)^{\mathrm{T}}$;

2) $a_1 = (1, -1, 1, -1)^{\mathrm{T}}$, $a_2 = (1, -2, 0, -3)^{\mathrm{T}}$, $a_3 = (1, 1, -2, 3)^{\mathrm{T}}$,
$a_4 = (2, 2, -4, 6)^{\mathrm{T}}$;

3) $a_1 = (1,\ 1,\ 1,\ 1)^{\mathrm{T}}$, $a_2 = (1,\ 2,\ 1,\ 1)^{\mathrm{T}}$, $a_3 = (1,\ 1,\ 3,\ 1)^{\mathrm{T}}$,
$a_4 = (1,\ 2,\ -1,\ 1)^{\mathrm{T}}$;

4) $a_1 = (1, -3, 5, 6)^{\mathrm{T}}$, $a_2 = (1, -3, 1, 1)^{\mathrm{T}}$, $a_3 = (-1, -3, 13, 16)^{\mathrm{T}}$,
$a_4 = (1, -3, 9, 11)^{\mathrm{T}}$;

5) $a_1 = (2, 1, 1)^{\mathrm{T}}$, $a_2 = (3, 1, 4)^{\mathrm{T}}$, $a_3 = (-3, 5, -34)^{\mathrm{T}}$, $a_4 = (5, 3, 0)^{\mathrm{T}}$.

4.7. Найти какую-либо базу системы векторов и через нее выразить остальные векторы системы:

1) $a_1 = (1,\ 2,\ 3,\ 1)^{\mathrm{T}}$, $a_2 = (2,\ 3,\ 1,\ 2)^{\mathrm{T}}$, $a_3 = (3,\ 1,\ 2,\ -2)^{\mathrm{T}}$,
$a_4 = (0,\ 4,\ 2,\ 5)^{\mathrm{T}}$;

2) $a_1 = (3,\ 1,\ -2,\ 4)^{\mathrm{T}}$, $a_2 = (1,\ 3,\ 1,\ 2)^{\mathrm{T}}$, $a_3 = (1,\ 5,\ 0,\ 1)^{\mathrm{T}}$,
$a_4 = (3,\ -5,\ 1,\ 7)^{\mathrm{T}}$, $a_5 = (7,\ -15,\ -11,\ 8)^{\mathrm{T}}$;

3) $a_1 = (1, 1, 1, 1, 1)^{\mathrm{T}}$, $a_2 = (1, 2, 3, 4, 5)^{\mathrm{T}}$, $a_3 = (5, 1, 2, 3, 4)^{\mathrm{T}}$,
$a_4 = (4, 5, 1, 2, 3)^{\mathrm{T}}$, $a_5 = (3, 4, 5, 1, 2)^{\mathrm{T}}$, $a_6 = (2, 3, 4, 5, 1)^{\mathrm{T}}$.

4.8. Найти какую-либо базу системы линейных форм и через нее выразить остальные линейные формы:

1) $f_1 = 4x_1 + 3x_2 - x_3 + x_4$,
$f_2 = 2x_1 + x_2 - 3x_3 + 3x_4$,
$f_3 = x_1 + x_2 + x_3 - x_4$,
$f_4 = 6x_1 + 5x_2 + x_3 - x_4$;

2) $f_1 = x_1 - x_2 + x_3 - x_4$,
$f_2 = x_1 + 2x_2 - 3x_3 + 2x_4$,
$f_3 = x_1 + x_2 - 2x_3 + 3x_4$,
$f_4 = 2x_1 - x_2 + x_3 - 3x_4$;

3) $f_1 = 2x_1 + 2x_2 + 7x_3 - x_4$,
$f_2 = 3x_1 - x_2 + 2x_3 + 4x_4$,
$f_3 = x_1 + x_2 + 3x_3 + x_4$;

4) $f_1 = 3x_1 + 2x_2 - 5x_3 + 4x_4$,
$f_2 = 3x_1 - x_2 + 3x_3 - 3x_4$,
$f_3 = 3x_1 + 5x_2 - 13x_3 + 11x_4$.

4.9. **Исследовать систему на совместность и решить ее, если она совместна:**

1) $\begin{cases} 2x_1 - 3x_2 + 4x_3 - x_4 = 1, \\ 2x_1 - 3x_2 + 2x_3 + 3x_4 = 2, \\ 2x_1 - 3x_2 + 2x_3 - 11x_4 = -4; \end{cases}$ 2) $\begin{cases} x_1 - x_2 + 2x_3 - x_4 = 1, \\ x_1 - x_2 + x_3 - x_4 = 0, \\ x_1 - x_2 + 5x_3 - x_4 = 4, \\ x_1 - x_2 + 6x_3 - x_4 = 5; \end{cases}$

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

4.10. Найти общее решение и построить какую-либо фундаментальную систему решений однородных систем уравнений:

1) $\begin{cases} x_1 + x_2 - x_3 - 2x_4 = 0, \\ x_1 + x_2 + 3x_3 + 4x_4 = 0, \\ x_1 + x_2 - 5x_3 - 8x_4 = 0, \\ x_1 + x_2 - 9x_3 - 14x_4 = 0; \end{cases}$ 2) $\begin{cases} 2x_1 - x_2 + 3x_3 + 4x_4 - x_5 = 0, \\ x_1 + 2x_2 - 3x_3 + x_4 + 2x_5 = 0, \\ 5x_1 - 5x_2 + 12x_3 + 11x_4 - 5x_5 = 0, \\ x_1 - 3x_2 + 6x_3 + 3x_4 - 3x_5 = 0; \end{cases}$

3) $\begin{cases} x_1 + 2x_2 + 3x_3 + x_4 = 0, \\ 2x_1 + 3x_2 + x_3 + 2x_4 = 0, \\ 3x_1 + x_2 + 2x_3 - 2x_4 = 0, \\ 4x_2 + 2x_3 + 5x_4 = 0; \end{cases}$ 4) $\begin{cases} 2x_1 + x_2 - 3x_3 + 5x_4 = 0, \\ 3x_1 - 2x_2 + x_3 - 3x_4 = 0, \\ x_1 + 4x_2 - 7x_3 + 13x_4 = 0, \\ 3x_1 + 5x_2 - 10x_3 + 18x_4 = 0. \end{cases}$

4.11. Зная частное решение x_0 неоднородной системы, найти ее общее решение, решая соответствующую однородную систему:

1) $\begin{cases} 3x_1 - 2x_2 + 5x_3 + 4x_4 = 2, \\ 6x_1 - 4x_2 + 4x_3 + 3x_4 = 3, \\ 9x_1 - 6x_2 + 3x_3 + 2x_4 = 4, \end{cases}$ $x_0 = (1,\ 1,\ 1,\ -1)^{\mathrm{T}}$;

2) $\begin{cases} 2x_1 - x_2 + 3x_3 - 7x_4 = 5, \\ 4x_1 - 3x_2 + x_3 - 4x_4 = 7, \\ 6x_1 - 2x_2 + 14x_3 - 31x_4 = 18, \end{cases}$ $x_0 = \left(2,\ 1,\ \frac{22}{5},\ \frac{8}{5}\right)^{\mathrm{T}}$;

3) $\begin{cases} 9x_1 - 3x_2 + 5x_3 + 6x_4 = 4, \\ 6x_1 - 2x_2 + 3x_3 + x_4 = 5, \\ 3x_1 - x_2 + 3x_3 + 14x_4 = -8, \end{cases}$ $\quad x_0 = \left(1,\ 1,\ \frac{8}{13},\ -\frac{11}{13}\right)^{\mathrm{T}};$

4) $\begin{cases} 3x_1 + 2x_2 + 2x_3 + 2x_4 = 2, \\ 2x_1 + 3x_2 + 2x_3 + 5x_4 = 3, \\ 9x_1 + x_2 + 4x_3 - 5x_4 = 1, \end{cases}$ $\quad x_0 = (-2,\ 2,\ 3,\ -1)^{\mathrm{T}};$

5) $\begin{cases} 2x_1 - 4x_2 + 3x_3 - 2x_4 = 3, \\ x_1 - 2x_2 - 2x_3 - x_4 = -2, \\ 3x_1 - 6x_2 + 5x_3 - 3x_4 = 5, \\ 4x_1 - 8x_2 - 3x_3 - 4x_4 = -3, \end{cases}$ $\quad x_0 = (3,\ 1,\ 1,\ 1)^{\mathrm{T}};$

6) $\begin{cases} x_1 - 2x_2 + 3x_3 - 4x_4 = 4, \\ x_2 - x_3 + x_4 = -3, \\ x_1 + 3x_2 - 3x_4 = 1, \\ 7x_2 + 3x_3 + x_4 = -3, \end{cases}$ $\quad x_0 = (-8,\ 4,\ 8,\ 1)^{\mathrm{T}}.$

4.12. Проверить, являются ли следующие системы векторов базисами в пространстве K_4:

1) $a_1 = (1,\ 2,\ -1,\ -2)^{\mathrm{T}}$, $a_2 = (2,\ 3,\ 0,\ -1)^{\mathrm{T}}$, $a_3 = (1,\ 2,\ 1,\ 3)^{\mathrm{T}}$, $a_4 = (1,\ 3,\ -1,\ 0)^{\mathrm{T}}$;

2) $a_1 = (1,\ 2,\ -1,\ -2)^{\mathrm{T}}$, $a_2 = (2,\ 3,\ 0,\ -1)^{\mathrm{T}}$, $a_3 = (1,\ 2,\ 1,\ 4)^{\mathrm{T}}$, $a_4 = (1,\ 3,\ -1,\ 0)^{\mathrm{T}}$.

4.13. Векторы базисов e, e' и вектор x даны координатами в некотором базисе e°: $[e_1] = (3,\ 2,\ 3)^{\mathrm{T}}$, $[e_2] = (-4,\ -3,\ -5)^{\mathrm{T}}$, $[e_3] = (5,\ 1,\ -1)^{\mathrm{T}}$, $[e'_1] = (2,\ 2,\ -1)^{\mathrm{T}}$, $[e'_2] = (2,\ -1,\ 2)^{\mathrm{T}}$, $[e'_3] = (-1,\ 2,\ 2)^{\mathrm{T}}$, $[x] = (1,\ 2,\ 1)^{\mathrm{T}}$. Найти матрицы перехода от базиса e° к базисам e, e' и от базиса e к базису e', а также координаты вектора x в базисах e и e'.

4.14. Найти размерности и базисы линейных подпространств, натянутых на системы векторов:

1) $a_1 = (1,\ 2,\ 0,\ 1)^{\mathrm{T}}$,	2) $a_1 = (1,\ 1,\ 1,\ 1)^{\mathrm{T}}$,
$a_2 = (1,\ 1,\ 1,\ 0)^{\mathrm{T}}$,	$a_2 = (1,\ -1,\ 1,\ -1)^{\mathrm{T}}$,
$a_3 = (1,\ 0,\ 1,\ 0)^{\mathrm{T}}$,	$a_3 = (1,\ 3,\ 1,\ 3)^{\mathrm{T}}$,
$a_4 = (1,\ 3,\ 0,\ 1)^{\mathrm{T}}$;	$a_4 = (1,\ 2,\ 0,\ 2)^{\mathrm{T}}$;

3) $a_1 = (1,\ 0,\ 0,\ -1)^{\text{т}}$,
$a_2 = (2,\ 1,\ 1,\ 0)^{\text{т}}$,
$a_3 = (1,\ 1,\ 1,\ 1)^{\text{т}}$,
$a_4 = (1,\ 2,\ 3,\ 4)^{\text{т}}$,
$a_5 = (0,\ 1,\ 2,\ 3)^{\text{т}}$;

4) $a_1 = (1,\ 1,\ 1,\ 1)^{\text{т}}$,
$a_2 = (1,\ -1,\ -1,\ -1)^{\text{т}}$,
$a_3 = (2,\ 2,\ 0,\ 0)^{\text{т}}$,
$a_4 = (1,\ 1,\ 5,\ 5)^{\text{т}}$,
$a_5 = (1,\ -1,\ -1,\ 0)^{\text{т}}$.

4.15. Найти системы линейных уравнений, задающие линейные подпространства, натянутые на системы векторов из упражнения 4.14.

4.16. Найти размерности и базисы подпространств, заданных системами уравнений из упражнения 4.10.

4.17. Найти размерности и базисы сумм и пересечений подпространств $L_1 = \langle a_1,\ a_2,\ a_3\rangle$, $L_2 = \langle b_1,\ b_2,\ b_3\rangle$, если:

1) $a_1 = (1,\ 2,\ 1)^{\text{т}}$, $a_2 = (1,\ 1,\ -1)^{\text{т}}$, $a_3 = (1,\ 3,\ 3)^{\text{т}}$,
$b_1 = (2,\ 3,\ -1)^{\text{т}}$, $b_2 = (1,\ 2,\ 2)^{\text{т}}$, $b_3 = (1,\ 1,\ -3)^{\text{т}}$;

2) $a_1 = (1,\ 1,\ 1)^{\text{т}}$, $a_2 = (1,\ 2,\ 0)^{\text{т}}$, $a_3 = (2,\ 3,\ 1)^{\text{т}}$,
$b_1 = (1,\ 2,\ 1)^{\text{т}}$, $b_2 = (1,\ 1,\ 0)^{\text{т}}$, $b_3 = (2,\ 3,\ 1)^{\text{т}}$;

3) $a_1 = (1,\ 1,\ 0,\ 0)^{\text{т}}$, $a_2 = (0,\ 1,\ 1,\ 0)^{\text{т}}$, $a_3 = (0,\ 0,\ 1,\ 1)^{\text{т}}$,
$b_1 = (1,\ 0,\ 1,\ 0)^{\text{т}}$, $b_2 = (0,\ 2,\ 1,\ 1)^{\text{т}}$, $b_3 = (1,\ 2,\ 1,\ 2)^{\text{т}}$.

4.18. Найти размерности и базисы сумм и пересечений подпространств, заданных системами уравнений:

1) $\begin{cases} x_1 + x_2 + x_3 + x_4 = 0, \\ x_1 - x_2 + x_3 - x_4 = 0, \end{cases}$ $\begin{cases} x_1 + x_2 - x_3 + x_4 = 0, \\ x_1 - 2x_2 + x_3 - 2x_4 = 0; \end{cases}$

2) $\begin{cases} x_1 + x_2 + x_3 - 2x_4 = 0, \\ x_1 - x_2 + x_3 + x_4 = 0, \end{cases}$ $\begin{cases} x_1 + 2x_2 - x_3 + x_4 = 0, \\ x_1 + x_2 - x_3 - x_4 = 0; \end{cases}$

3) $\begin{cases} x_1 + x_2 - x_3 + x_4 = 0, \\ x_1 - x_2 + x_3 - 2x_4 = 0, \end{cases}$ $\begin{cases} x_1 - x_2 + x_3 - x_4 = 0, \\ x_1 - 2x_2 + 2x_3 + 2x_4 = 0. \end{cases}$

4.19. Для подпространств L_1 из упражнения 4.14 построить какие-либо прямые дополнения L_2 и в каждом случае найти проекцию вектора $x = (1,\ 3,\ 2,\ 1)^{\text{т}}$ на подпространство L_1 параллельно подпространству L_2.

4.20. Построить скелетные разложения следующих матриц:

$$1)\ \begin{pmatrix} 7 & 1 & 2 \\ 5 & 5 & -2 \\ 3 & 3 & 6 \\ 1 & 7 & 2 \end{pmatrix}; \qquad 2)\ \begin{pmatrix} 10 & 4 & -1 \\ 10 & 4 & -1 \\ 2 & 8 & 7 \\ -2 & -8 & -7 \end{pmatrix};$$

$$3)\ \begin{pmatrix} 8 & 5 & 1 \\ 8 & 5 & 1 \\ 4 & 7 & 5 \\ -4 & -7 & -5 \end{pmatrix}; \qquad 4)\ \begin{pmatrix} 9 & 3 & 3 \\ 7 & 7 & -1 \\ 5 & 5 & 7 \\ 3 & 9 & 3 \end{pmatrix};$$

$$5)\ \begin{pmatrix} 6 & 3 & 0 \\ 6 & 3 & 0 \\ 2 & 5 & 4 \\ -2 & -5 & -4 \end{pmatrix}; \qquad 6)\ \begin{pmatrix} 15 & -3 & 3 \\ 9 & 9 & -9 \\ 3 & 3 & 15 \\ -3 & 15 & 3 \end{pmatrix}.$$

Глава 5 ЛИНЕЙНЫЕ ОПЕРАТОРЫ В ЛИНЕЙНЫХ ПРОСТРАНСТВАХ

5.1. Определение и примеры линейных операторов

Пусть даны линейные пространства X и Y над одним и тем же полем P. Говорят, что ***из пространства X в пространство Y действует оператор*** φ или, что то же самое, ***отображение*** φ, ***преобразование*** φ, ***функция*** φ, если каждому вектору a из X по какому-либо правилу ставится в соответствие определенный вектор $a' = \varphi(a) = \varphi a$ из Y. Вектор a' называют ***образом вектора*** a, вектор a — ***прообразом вектора*** a' при отображении φ.

Если пространства X и Y совпадают, то говорят, что оператор φ действует в пространстве X_n.

Наиболее простыми являются линейные операторы. Оператор φ, действующий из X в Y, называют ***линейным***, если он сумму любых векторов a и b из X переводит в сумму их образов a' и b', а произведение любого вектора a из X на любое число α из P — в произведение образа a' вектора a на то же число α, т.е. если

$$\varphi(a+b) = \varphi a + \varphi b = a' + b', \quad \varphi(\alpha a) = \alpha\varphi a = \alpha a'.$$

Непосредственно из определения линейного оператора вытекают следующие утверждения:

1) линейный оператор φ переводит линейную комбинацию векторов a_1, a_2, $\ldots$, a_k из X в линейную комбинацию образов

$a'_1, a'_2, \ldots, a'_k$ этих векторов с теми же коэффициентами, т.е.

$$\varphi(\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_k a_k) = \alpha_1 \varphi a_1 + \alpha_2 \varphi a_2 + \ldots + \alpha_k \varphi a_k = \\ = \alpha_1 a'_1 + \alpha_2 a'_2 + \ldots + \alpha_k a'_k;$$

2) линейный оператор φ переводит нулевой вектор 0 из X в нулевой вектор $0'$ и Y;

3) линейный оператор переводит вектор $-a$, противоположный вектору a, в вектор $-a'$, противоположный вектору $a' = \varphi a$.

Первое утверждение несложно доказать методом математической индукции по количеству векторов k. Второе утверждение доказывается следующим образом:

$$\varphi 0 = \varphi(0 \cdot a) = 0 \cdot \varphi a = 0.$$

Доказательство третьего утверждения аналогично:

$$\varphi(-a) = \varphi\big[(-1)a\big] = (-1)\varphi a = (-1)a' = -a'.$$

Пусть из линейного пространства X в линейное пространство Y действует линейный оператор φ. Множество φX образов всех векторов из X при действии оператора φ называют ***областью значений оператора*** φ.

Область значений оператора φ является подпространством в Y. Действительно, если $y_1, y_2 \in \varphi X$, то $y_1 = \varphi x_1$, $y_2 = \varphi x_2$, где $x_1, x_2 \in X$. Поэтому $y_1 + y_2 = \varphi x_1 + \varphi x_2 = \varphi(x_1 + x_2) \in \varphi X$. Аналогично $\lambda y_1 = \lambda \varphi x_1 = \varphi(\lambda x_1) \in \varphi X$. Следовательно, множество φX замкнуто относительно операций линейного пространства Y.

Размерность области значений линейного оператора φ называют ***рангом линейного оператора***.

Множество $\operatorname{Ker}\varphi$ всех векторов линейного пространства X, которые переводятся линейным оператором φ в нулевой вектор линейного пространства Y, называют ***ядром линейного оператора*** φ. Как и область значений, ***ядро линейного оператора является линейным подпространством*** в X. Действительно, условие $x_1, x_2 \in \operatorname{Ker}\varphi$ означает, что $\varphi x_1 = \varphi x_2 = 0$. Поэтому

$$\varphi(x_1 + x_2) = \varphi x_1 + \varphi x_2 = 0 + 0 = 0$$

и

$$\varphi(\lambda x_1) = \lambda\varphi x_1 = \lambda \cdot 0 = 0,$$

т.е. множество $\operatorname{Ker}\varphi$ замкнуто относительно операций линейного пространства X.

Размерность ядра линейного оператора называют ***дефектом линейного оператора***. Если ядро линейного оператора φ состоит только из нулевого вектора (т.е. дефект оператора равен нулю), то φ называют ***невырожденным линейным оператором***. В противном случае его называют ***вырожденным линейным оператором***.

Два линейных оператора φ и ψ, действующих из линейного пространства X в линейное пространство Y, называют ***равными***, если для любого $a \in X$ выполняется условие $\varphi a = \psi a$.

Приведем несколько примеров линейных операторов.

1. Оператор ε: $X \to X$, переводящий любой вектор a линейного пространства X в тот же вектор $a \in X$, является линейным оператором. Такой оператор называют ***тождественным***.

2. Оператор ω: $X \to Y$, переводящий любой вектор a линейного пространства X в нулевой вектор линейного пространства Y, является линейным оператором. Такой оператор называют ***нулевым***.

3. Умножение векторов линейного пространства X на одно и то же число α (растяжение линейного пространства X в α раз) является линейным оператором, действующим в X. Такой оператор называют ***оператором подобия***.

4. Оператор отражения φ: $X \to X$, определяемый равенством $\varphi x = -x$, является линейным оператором.

5. Пусть X — трехмерное пространство векторов, выходящих из начала системы координат $0xyz$, Y — одномерное

пространство векторов на оси $0y$. Ортогональное проектирование вектора на ось Oy определяет линейный оператор, действующий из X в Y. Этот оператор можно рассматривать также как оператор, действующий в пространстве X.

6. Пусть в трехмерном пространстве X_3 с базисом $e = (e_1, e_2, e_3)$ оператор φ переводит вектор $x = (x_1, x_2, x_3)^{\mathrm{T}}$ в вектор $\varphi x = (x_1 + x_2, x_2 + x_3, x_3 + x_1)^{\mathrm{T}}$. Непосредственной проверкой легко убедиться в том, что выполняются условия $\varphi(x+y) = \varphi x + \varphi y$ и $\varphi(\alpha x) = \alpha\varphi x$. Поэтому оператор φ линейный.

7. Пусть в линейном пространстве $M_{nn}(P)$ квадратных матриц над полем P оператор φ действует по правилу

$$\varphi A = \varphi\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} = \begin{pmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix}.$$

Этот оператор линейный, поскольку выполняются соотношения

$$\varphi(A+B) = \varphi A + \varphi B \quad \text{и} \quad \varphi(\lambda A) = \lambda\varphi A.$$

8. В линейном пространстве $C^{\infty}[a, b]$ бесконечно дифференцируемых функций на отрезке $[a, b]$ дифференцирование определяет линейный оператор, поскольку производная суммы равна сумме производных, а производная произведения функции на число равна произведению производной этой функции на то же число.

Теорема 5.1. *Пусть линейный оператор φ действует из линейного пространства X в линейное пространство Y и $(e_1, e_2, \dots, e_n)$ — базис в X. Тогда оператор φ однозначно определяется заданием образов $\varphi e_1, \varphi e_2, \dots, \varphi e_n$ векторов базиса $(e_1, e_2, \dots, e_n)$.*

$\triangleright$ Если известны образы $\varphi e_1, \varphi e_2, \dots, \varphi e_n$ векторов базиса $(e_1, e_2, \dots, e_n)$, то для любого вектора

$$x = x_1 e_1 + x_2 e_2 + \dots x_n e_n$$

мы однозначно определяем его образ при действии оператора φ:

$$\varphi x = \varphi(x_1e_1 + x_2e_2 + \ldots x_ne_n) = x_1\varphi e_1 + x_2\varphi e_2 + \ldots + x_n\varphi e_n.$$

Следовательно, оператор φ однозначно определяется образами векторов заданного базиса. ►

5.2. Линейные операторы и матрицы

Предположим, что в линейном пространстве X задан базис $e = (e_1, e_2, \ldots, e_n)$, а в линейном пространстве Y — базис $q = (q_1, q_2, \ldots, q_m)$. По теореме 5.1 линейный оператор φ, действующий из X в Y, однозначно определяется образами $\varphi e_1, \varphi e_2, \ldots, \varphi e_n$ векторов базиса e. Каждый такой вектор разложим по базису q:

$$\begin{cases} \varphi e_1 = a_{11}q_1 + a_{21}q_2 + \ldots + a_{m1}q_m, \\ \varphi e_2 = a_{12}q_1 + a_{22}q_2 + \ldots + a_{m2}q_m, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \varphi e_n = a_{1n}q_1 + a_{2n}q_2 + \ldots + a_{mn}q_m. \end{cases} \tag{5.1}$$

Из координатных столбцов этих векторов, т.е. из коэффициентов разложений (5.1), составим матрицу

$$A_{qe} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \tag{5.2}$$

и назовем ее ***матрицей линейного оператора*** φ в паре базисов e и q.

Итак, каждому линейному оператору, действующему из X в Y, в паре заданных базисов e и q соответствует $(m \times n)$-матрица. Наоборот, любая $(m \times n)$-матрица $A = (a_{ij})$ является матрицей некоторого линейного оператора φ, действующего из X в Y. Действительно, рассмотрим линейный оператор φ, который каждый вектор e_j базиса e переводит в вектор φe_j, координатным столбцом которого является j-й столбец матрицы A, т.е. положим:

$$\varphi e_j = a_{1j}q_1 + a_{2j}q_2 + \ldots + a_{mj}q_m, \quad j = 1, 2, \ldots, n.$$

Согласно теореме 5.1 оператор φ определен однозначно, а его матрицей в базисах e и q будет матрица A.

Мы приходим к выводу, что между множеством линейных операторов, действующих из n-мерного линейного пространства X в m-мерное линейное пространство Y, и множеством матриц размера $m \times n$ установлено взаимно однозначное соответствие. При фиксированных базисах e и q это соответствие позволяет отождествить линейные операторы с их матрицами.

Соотношения (5.1) устанавливают связь между линейным оператором φ, базисами $e = (e_1, \ldots, e_n)$ и $q = (q_1, \ldots, q_m)$ и матрицей A_{qe} оператора φ в паре этих базисов. В матричной форме эти соотношения записываются в виде

$$\varphi e = qA_{qe}, \tag{5.3}$$

где $\varphi e = (\varphi e_1, \varphi e_2, \ldots, \varphi e_n)$.

Если оператор φ действует в линейном пространстве X, то X выступает и в качестве области определения, и области значений. В этом случае естественно ограничиться лишь одним базисом, т.е. считать, что базисы e и q совпадают. При этом матрица A_e линейного оператора будет квадратной, а соотношение (5.3) примет вид:

$$\varphi e = eA_e. \tag{5.4}$$

Найдем матрицы некоторых линейных операторов, указанных в предыдущих примерах.

1. Тождественный оператор ε любой вектор $a \in X$ переводит в вектор a. В частности, каждый вектор e_i базиса e

оператором ε переводится в вектор e_i:

$$\begin{cases} \varepsilon e_1 = e_1 = (1,\ 0,\ \ldots,\ 0,\ 0)^{\text{т}}, \\ \varepsilon e_2 = e_2 = (0,\ 1,\ \ldots,\ 0,\ 0)^{\text{т}}, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ \varepsilon e_n = e_n = (0,\ 0,\ \ldots,\ 0,\ 1)^{\text{т}}. \end{cases}$$

Столбцы координат векторов εe_1, εe_2, ..., εe_n в базисе e составляют матрицу A_e оператора ε в базисе e, представляющую собой единичную матрицу:

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

2. Оператор подобия φ растягивает каждый вектор линейного пространства X в α раз (умножает вектор на α). Значит,

$$\begin{cases} \varepsilon e_1 = \alpha e_1 = (\alpha,\ 0,\ \ldots,\ 0,\ 0)^{\text{т}}, \\ \varepsilon e_2 = \alpha e_2 = (0,\ \alpha,\ \ldots,\ 0,\ 0)^{\text{т}}, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ \varepsilon e_n = \alpha e_n = (0,\ 0,\ \ldots,\ 0,\ \alpha)^{\text{т}}. \end{cases}$$

Из столбцов координат векторов αe_i составляем матрицу оператора φ в базисе e:

$$A_e = \begin{pmatrix} \alpha & & 0 \\ & \ddots & \\ 0 & & \alpha \end{pmatrix}.$$

3. Пусть в пространстве E_3 векторов, выходящих из начала координат, линейный оператор φ действует так, что векторы по оси Ox растягиваются в λ_1 раз, по оси Oy — в λ_2 раз, по оси Oz — в λ_3 раз. Тогда $\varphi e_1 = \lambda_1 e_1$, $\varphi e_2 = \lambda_2 e_2$, $\varphi e_3 = \lambda_3 e_3$, и матрица A_e линейного оператора в базисе e_1, e_2, e_3 имеет вид:

$$A_e = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}.$$

Теорема 5.2. *Пусть φ — линейный оператор, действующий из линейного пространства X в линейное пространство Y, имеющий в двух заданных базисах e в X и q в Y матрицу A_{qe}. Тогда:*

- *ранг r оператора φ совпадает с рангом его матрицы A_{qe};*
- *дефект оператора φ равен разности $n-r$ размерности n линейного пространства X и ранга r оператора φ;*
- *сумма ранга и дефекта оператора φ равна размерности линейного пространства X.*

$\triangleright$ Область значений φX оператора φ является линейной оболочкой системы векторов $\varphi e_1, \varphi e_2, \ldots, \varphi e_n$, представляющих собой образы векторов $e_1, e_2, \ldots, e_n$ базиса e. Ее размерность, т.е. ранг оператора φ, совпадает с рангом системы векторов $\varphi e_1, \varphi e_2, \ldots, \varphi e_n$. Ранг системы векторов совпадает с рангом матрицы, составленной из координатных столбцов векторов системы. В данном случае матрица, составленная из координатных столбцов, есть матрица линейного оператора. Значит, ранг линейного оператора совпадает с рангом любой его матрицы.

Пусть x — произвольный вектор из ядра линейного оператора φ. Тогда его координатный столбец $[x]_e$ в базисе e является решением однородной системы линейных уравнений $A_{qe}[x]_e = 0$. Верно и обратное утверждение: если координатный столбец $[x]_e$ вектора x в базисе e является решением однородной системы $A_{qe}[x]_e = 0$, то вектор x принадлежит ядру линейного оператора. Из общей теории линейных систем (см. разд. 4.8) известно, что размерность пространства решений однородной системы $A_{qe}[x]_e = 0$ равна $n-r$, где n — количество неизвестных системы, равное размерности линейного пространства X, а r — ранг матрицы системы. Но размерность пространства решений системы $A_{qe}[x]_e = 0$, т.е. размерность ядра линейного оператора, по определению есть дефект линейного оператора, а ранг матрицы линейного оператора, как уже доказано, совпадает с рангом линейного оператора. Тем самым доказано второе утверждение теоремы. Третье утверждение теоремы непосредственно следует из второго. ▶

5.3. Выражение координат вектора-образа через координаты вектора-прообраза

Теорема 5.3. *Если φ — линейный оператор, действующий из линейного пространства X в линейное пространство Y, который в паре базисов $e = (e_1, e_2, \ldots, e_n)$ в X и $q = (q_1, q_2, \ldots, q_m)$ в Y имеет матрицу A_{qe}, то столбец координат $[x']_q$ вектора $x' = \varphi x$ в базисе q равен произведению матрицы A_{qe} на столбец координат $[x]_e$ вектора x в базисе e, т.е.*

$$[x']_q = A_{qe}[x]_e. \tag{5.5}$$

В частности, если пространства $X = Y$ и $q = e$, то

$$[x']_e = A_e[x]_e. \tag{5.6}$$

◁ По условию линейный оператор φ в паре базисов e и q имеет матрицу A_{qe}. Поэтому $\varphi e = qA_{qe}$. Произвольный вектор $x \in X$ может быть представлен в виде $x = e[x]_e$, где $[x]_e$ — координатный столбец этого вектора. Поэтому

$$\varphi x = \varphi(e[x]_e) = (\varphi e)[x]_e = (qA_{qe})[x]_e = q(A_{qe}[x]_e).$$

Тем самым мы пришли к соотношению $x' = \varphi x = q(A_{qe}[x]_e)$, означающему, что столбец $A_{qe}[x]_e$ является координатным столбцом вектора x' в базисе q, т.е. $[x']_q = A_{qe}[x]_e$. ▶

Рассмотрим примеры использования формул (5.5) и (5.6).

Пример 5.1. Пусть линейный оператор, действующий из линейного пространства X в линейное пространство Y, в паре базисов $e = (e_1, e_2, e_3, e_4)$ в X и $q = (q_1, q_2, q_3)$ в Y задан матрицей

$$A_{qe} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \end{pmatrix}.$$

Найти столбец координат в базисе q образа вектора x, имеющего в базисе e столбец координат $(1, 1, 1, 1)^{\mathrm{T}}$, и столбец координат в базисе e прообраза вектора y, имеющего в базисе q столбец координат $(3, 2, 1)^{\mathrm{T}}$.

Р е ш е н и е. Столбец координат образа вектора x в базисе q находим непосредственно по формуле (5.5):

$$[x']_q = \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 10 \\ 6 \\ 3 \end{pmatrix}.$$

Для отыскания прообраза вектора y воспользуемся той же формулой (5.5). Если $[x]_e = (x_1, x_2, x_3, x_4)^{\mathrm{T}}$ — столбец координат искомого вектора, то

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}.$$

Мы имеем систему линейных уравнений относительно неизвестных x_1, x_2, x_3, x_4. Решив эту систему, находим все прообразы вектора y:

$$x = (0,\ x_4,\ 1 - 2x_4,\ x_4)^{\mathrm{T}}.$$

Здесь x_4 — свободное неизвестное, которое может принимать любые значения.

П р и м е р 5.2. В линейном пространстве X, в котором задан базис (e_1, e_2), действует линейный оператор φ, переводящий векторы a_1 и a_2 с координатами $[a_1]_e = (1, 1)^{\mathrm{T}}$ и

$[a_2]_e = (1,\ 0)^{\mathrm{T}}$ в векторы b_1 и b_2 с координатами $[b_1]_e = (2,\ 1)^{\mathrm{T}}$ и $[b_2]_e = (1,\ 2)^{\mathrm{T}}$. Найти матрицу оператора φ в базисе e.

Решение. Пусть A_e — искомая матрица линейного оператора φ. Согласно формуле (5.6) имеем два матричных равенства

$$A_e\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}2\\1\end{pmatrix}, \qquad A_e\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}1\\2\end{pmatrix}.$$

Эти равенства можно объединить:

$$A_e\begin{pmatrix}1 & 1\\1 & 0\end{pmatrix} = \begin{pmatrix}2 & 1\\1 & 2\end{pmatrix}.$$

Отсюда заключаем, что

$$A_e = \begin{pmatrix}1 & 1\\1 & 0\end{pmatrix}^{-1}\begin{pmatrix}2 & 1\\1 & 2\end{pmatrix} = \begin{pmatrix}0 & 1\\1 & -1\end{pmatrix}\begin{pmatrix}2 & 1\\1 & 2\end{pmatrix} = \begin{pmatrix}1 & 1\\2 & -1\end{pmatrix}.$$

Другие подходы к решению таких задач рассмотрены далее (см. разд. 5.4 и разд. 5.5).

5.4. Связь между матрицами линейного оператора в разных базисах

Теорема 5.4. *Пусть φ — линейный оператор, действующий из линейного пространства X в линейное пространство Y, A_{qe} — матрица этого оператора в паре базисов e и q, $A_{q'e'}$ — матрица оператора в паре базисов e' и q'. Тогда матрицы A_{qe} и $A_{q'e'}$ связаны соотношением*

$$A_{q'e'} = Q^{-1}A_{qe}P, \tag{5.7}$$

где Q — матрица перехода от базиса q к базису q' в пространстве Y, P — матрица перехода от базиса e к базису e' в пространстве X.

Замечание 5.1. Если $X = Y$, $e = q$, $e' = q'$, то формула (5.7) принимает вид:

$$A_{e'} = P^{-1} A_e P. \tag{5.8}$$

▷ Если вектор x в базисе e имеет столбец координат $[x]_e$, а в базисе e' — столбец координат $[x]_{e'}$, то

$$[x]_e = P[x]_{e'}.$$

Аналогично столбец координат $[y]_q$ вектора $y \in Y$ в базисе q и столбец координат $[y]_{q'}$ этого же вектора в базисе q' связаны равенством

$$[y]_q = Q[y]_{q'}.$$

Пусть y — образ вектора x при действии оператора φ. Тогда:

$$[y]_q = A_{qe}[x]_e \quad \text{и} \quad [y]_{q'} = A_{q'e'}[x]_{e'}.$$

Комбинируя записанные равенства и учитывая невырожденность матриц P и Q, получим, с одной стороны,

$$[y]_{q'} = Q^{-1}[y]_q = Q^{-1} A_{qe}[x]_e,$$

а с другой,

$$[y]_{q'} = A_{q'e'}[x]_{e'} = A_{q'e'} P^{-1}[x]_e.$$

Следовательно,

$$Q^{-1} A_{qe} = A_{q'e'} P^{-1},$$

что равносильно равенству (5.7). ▶

Напомним, что матрицы A и B называют ***эквивалентными***, если при некоторых невырожденных квадратных матрицах R и S выполняется соотношение $B = RAS$. Квадратные матрицы A и B называют ***подобными***, если они связаны соотношением $B = P^{-1}AP$ при некоторой невырожденной матрице P. Матрицу P называют ***трансформирующей*** и говорят, что матрица B получается из матрицы A ***трансформированием при помощи матрицы*** P.

С помощью этих понятий теорему 5.4 можно переформулировать следующим образом.

Теорема 5.5. *Если φ — линейный оператор, действующий из линейного пространства X в линейное пространство Y, то его матрицы, соответствующие разным парам базисов, эквивалентны. Если φ — линейный оператор, действующий в линейном пространстве X, то его матрицы, соответствующие разным базисам в X, подобны.*

Верно и обратное утверждение.

Теорема 5.6. *Если A — матрица линейного оператора φ в паре базисов линейных пространств X и Y (в базисе линейного пространства X), то любая эквивалентная (подобная) ей матрица B является матрицей этого же оператора в некоторой другой паре базисов (другом базисе).*

$\triangleright$ Если A — матрица линейного оператора φ в паре базисов e и q, матрица B связана с матрицей A равенством $B = RAS$, то существует такой базис e в линейном пространстве X, что матрицей перехода из e в e' является матрица S. Аналогично существует такой базис q' в линейном пространстве Y, что матрицей перехода от q к q' является матрица R^{-1}. Согласно формуле (5.7) оператор φ в паре базисов e', q' имеет матрицу RAS, т.е. матрицу B.

Если квадратная матрица A является матрицей линейного оператора φ в базисе e, а матрица B подобна A, т.е. для некоторой невырожденной матрицы P имеем: $B = P^{-1}AP$, то достаточно выбрать базис e' так, что матрицей перехода из e в e' является матрица P. Тогда в силу формулы (5.8) матрицей оператора φ в базисе e' будет матрица $P^{-1}AP = B$. $\blacktriangleright$

Остановимся на применении формулы (5.8).

Пример 5.3. В линейном пространстве X задан базис $e = (e_1, e_2)$. Линейный оператор φ, действующий в X, переводит векторы a_1 и a_2 с координатами $[a_1]_e = (3,\ 4)^{\mathrm{T}}$ и $[a_2]_e = (4,\ 5)^{\mathrm{T}}$ в векторы b_1 и b_2 с координатами $[b_1]_e = (1,\ 2)^{\mathrm{T}}$ и $[b_2]_e = (3,\ 3)^{\mathrm{T}}$. Найти матрицу оператора φ в базисе e.

Решение. Примем систему векторов a_1, a_2 за базис e'. Тогда матрицей перехода от базиса e к базису e' будет

$$T_1 = \begin{pmatrix} 3 & 4 \\ 4 & 5 \end{pmatrix}.$$

Используя формулу (5.8), найдем координаты векторов b_1 и b_2 в базисе e':

$$[b_1]_{e'} = T_1^{-1}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} -5 & 4 \\ 4 & -3 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ -2 \end{pmatrix},$$

$$[b_2]_{e'} = T_1^{-1}\begin{pmatrix} 3 \\ 3 \end{pmatrix} = \begin{pmatrix} -5 & 4 \\ 4 & -3 \end{pmatrix}\begin{pmatrix} 3 \\ 3 \end{pmatrix} = \begin{pmatrix} -3 \\ 3 \end{pmatrix}.$$

В базисе e' оператор φ имеет матрицу (см. разд. 5.2)

$$A'_{e'} = \begin{pmatrix} 3 & -3 \\ -2 & 3 \end{pmatrix}.$$

Заметим, что матрица $A'_{e'}$ является матрицей перехода от базиса a_1, a_2 к базису b_1, b_2. Ее можно было находить следующим образом:

$$A'_{e'} = T_1^{-1}T_2 = \begin{pmatrix} 3 & 4 \\ 4 & 5 \end{pmatrix}^{-1}\begin{pmatrix} 1 & 3 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} -5 & 4 \\ 4 & -3 \end{pmatrix} = \begin{pmatrix} 3 & -3 \\ -2 & 3 \end{pmatrix}.$$

Теперь, используя формулу (5.8), найдем матрицу оператора φ в базисе e:

$$A_e = T_1 A'_{e'} T_1^{-1} = \begin{pmatrix} 3 & 4 \\ 4 & 5 \end{pmatrix}\begin{pmatrix} 3 & -3 \\ -2 & 3 \end{pmatrix}\begin{pmatrix} -5 & 4 \\ 4 & -3 \end{pmatrix} = \begin{pmatrix} 7 & -5 \\ 2 & -1 \end{pmatrix}.$$

Еще один подход к решению таких задач приведен ниже (см. разд. 5.5).

5.5. Действия с линейными операторами

Пусть из линейного пространства X над полем P в линейное пространство Y над тем же полем действуют линейные операторы φ и ψ. Операторы φ и ψ считаются ***равными***, если для любого вектора $x \in X$ его образы при действии этих операторов равны, т.е. если $\varphi x = \psi x$, $x \in X$. Равные линейные операторы в одном и том же базисе имеют одинаковые матрицы.

Суммой операторов φ и ψ называют оператор $\varphi + \psi$, переводящий любой вектор $x \in X$ в сумму образов от действия на x порознь операторов φ и ψ, т.е.

$$(\varphi + \psi)x = \varphi x + \psi x.$$

Сумма линейных операторов является линейным оператором. Действительно, для любых векторов $x, y \in X$ имеем:

$$(\varphi + \psi)(x + y) = \varphi(x + y) + \psi(x + y) = \\ = \varphi x + \varphi y + \psi x + \psi y = (\varphi + \psi)x + (\varphi + \psi)y.$$

Аналогично для любого вектора $x \in X$ и числа $\lambda \in P$

$$(\varphi + \psi)(\lambda x) = \varphi(\lambda x) + \psi(\lambda x) = \lambda\varphi x + \lambda\psi x = \\ = \lambda(\varphi x + \psi x) = \lambda\big[(\varphi + \psi)x\big].$$

Матрицей суммы линейных операторов в фиксированных базисах является сумма матриц слагаемых операторов в тех же базисах. Действительно, пусть линейные операторы φ и ψ в паре базисов e и q линейных пространств X и Y имеют матрицы A и B. Тогда для любого вектора $x \in X$ имеем:

$$(\varphi + \psi)x = \varphi x + \psi x,$$

что в матричной записи имеет вид:

$$(A + B)[x] = A[x] + B[x],$$

где $[x]$ — столбец координат вектора x.

Операция сложения линейных операторов ассоциативна. В самом деле, пусть φ, ψ и χ — три линейных оператора, действующих из линейного пространства X в линейное пространство Y. Тогда для любого вектора $x \in X$

$$\begin{aligned}\big[(\varphi+\psi)+\chi\big]x = (\varphi+\psi)x + \chi x = \varphi x + \psi x + \chi x = \\ = \varphi x + (\psi+\chi)x = \big[\varphi + (\psi+\chi)\big]x,\end{aligned}$$

т.е. верно равенство

$$\big[(\varphi+\psi)+\chi\big]x = \big[\varphi+(\psi+\chi)\big]x.$$

Это равенство равносильно равенству линейных операторов $(\varphi+\psi)+\chi$ и $\varphi+(\psi+\chi)$.

Ассоциативность сложения линейных операторов означает, что сложение распространяется на любое конечное число слагаемых.

Операция сложения линейных операторов коммутативна, так как для любых операторов φ и ψ и любого вектора $x \in X$ выполняются равенства

$$(\varphi+\psi)x = \varphi x + \psi x = \psi x + \varphi x = (\psi+\varphi)x,$$

что равносильно равенству $\varphi+\psi = \psi+\varphi$.

Кроме этих свойств очевидными являются существование в множестве всех линейных операторов, действующих из линейного пространства X в линейное пространство Y, нулевого оператора, а также у каждого оператора противоположного ему. С учетом этих свойств заключаем, что множество всех линейных операторов, действующих из X в Y, есть абелева группа по сложению.

Произведением линейного оператора φ на число $\lambda \in P$ называют оператор $\lambda\varphi$, действие которого таково, что образ каждого вектора при действии оператора φ умножается на число λ, т.е.

$$(\lambda\varphi)x = \lambda(\varphi x), \quad x \in X.$$

Произведение линейного оператора на число является линейным оператором. Это вытекает из соотношений

$$(\lambda\varphi)(x+y) = \lambda\varphi(x+y) = \lambda\varphi x + \lambda\varphi y = (\lambda\varphi)x + (\lambda\varphi)y,$$

$$(\lambda\varphi)(\mu x) = \lambda\bigl[\varphi(\mu x)\bigr] = \lambda(\mu\varphi x) = (\lambda\mu)(\varphi x) = \mu(\lambda\varphi)x,$$

выполняющихся при любых $x, y \in X$ и $\mu \in P$.

При умножении линейного оператора на число его матрица умножается на это число, поскольку соотношению $(\lambda\varphi)x = \lambda(\varphi x)$ в матричной записи соответствует соотношение $(\lambda A)[x] = \lambda(A[x])$. Непосредственной проверкой легко убедиться в том, что при сложении операторов и умножении на числа выполняются все аксиомы линейного пространства. Поэтому множество всех линейных операторов, действующих из линейного пространства X над полем P в линейное пространство Y над тем же полем, с операциями сложения операторов и умножения оператора на число есть линейное пространство, причем это линейное пространство изоморфно линейному пространству $(m \times n)$-матриц с элементами из поля P. В частном случае $X = Y$ получаем линейное пространство линейных операторов, действующих в линейном пространстве X, которое изоморфно линейному пространству квадратных матриц порядка n с элементами из поля P.

Пусть даны линейные пространства X, Y, Z над одним и тем же полем P, оператор φ действует из X в Y, а оператор ψ — из Y в Z. Результат их последовательного применения, т.е. оператор $\psi\varphi$, действующий по правилу $(\psi\varphi)x = \psi(\varphi x)$, $x \in X$, называют ***произведением операторов*** φ и ψ (оператор, действующий первым, записывается справа). Оператор $\psi\varphi$ действует из линейного пространства X в линейное пространство Z и является линейным. Это вытекает из следующих соотношений:

$$(\psi\varphi)(x+y) = \psi\bigl[\varphi(x+y)\bigr] = \psi(\varphi x + \varphi y) = (\psi\varphi)x + (\psi\varphi)y,$$

$$(\psi\varphi)(\lambda x) = \psi\bigl[\varphi(\lambda x)\bigr] = \psi(\lambda\varphi x) = \lambda\psi(\varphi x) = \lambda(\psi\varphi)x.$$

Умножение операторов не является алгебраической операцией уже потому, что оно определено не для любой пары операторов. Умножение операторов, если оно выполнимо, обладает следующими свойствами:

1. $(\varphi\psi)\chi = \varphi(\psi\chi)$.
2. $\lambda(\varphi\psi) = (\lambda\varphi)\psi = \varphi(\lambda\psi)$.
3. $(\varphi + \psi)\chi = \varphi\chi + \psi\chi$.
4. $\varphi(\psi + \chi) = \varphi\psi + \varphi\chi$.

Эти свойства можно доказать непосредственной проверкой. Докажем, например, первое свойство. Пусть даны линейные пространства X, Y, Z, U, линейный оператор χ действует из X в Y, линейный оператор ψ действует из Y в Z, а линейный оператор φ действует из Z в U. Тогда для любого вектора $x \in X$ имеем:

$$\big[(\varphi\psi)\chi\big]x = (\varphi\psi)(\chi x) = \varphi\big[\psi(\varphi x)\big] = \varphi\big[(\psi\varphi)x\big] = \varphi(\psi\chi)x,$$

что и означает выполнение первого свойства.

В линейных пространствах X, Y, Z выберем базисы e, q, f. Пусть линейный оператор φ действует из X в Y, линейный оператор ψ действует из Y в Z. Если A — матрица оператора φ в паре базисов e и q, B — матрица оператора ψ в паре базисов q и f, то матрицей оператора $\psi\varphi$ в паре базисов e и f является матрица BA (матрица оператора, действующего первым, записывается справа). Это вытекает из того, что соотношению $(\psi\varphi)x = \psi(\varphi x)$ в матричной форме соответствует запись $(BA)[x] = B(A[x])$.

Поскольку матрица произведения операторов есть произведение матриц сомножителей, а ранг произведения матриц не превышает рангов сомножителей, делаем вывод, что ранг произведения операторов не превышает рангов перемножаемых операторов.

Так как произведение операторов обладает свойством ассоциативности, понятие произведения линейных операторов распространяется на любое конечное число сомножителей. При этом матрицей произведения k операторов является произведение матриц сомножителей. Рассмотрим пример на умножение операторов.

Пример 5.4. В линейном пространстве X задан базис $e = (e_1, e_2)$. Линейный оператор φ переводит векторы a_1 и a_2 с координатами $[a_1]_e = (2, -1)^{\mathrm{T}}$ и $[a_2]_e = (-1, 1)^{\mathrm{T}}$ соответственно в векторы b_1 и b_2 с координатами $[b_1]_e = (1, 2)^{\mathrm{T}}$, $[b_2]_e = (3, 3)^{\mathrm{T}}$. Найти матрицу оператора φ в базисе e.

Решение. Через ψ обозначим оператор, переводящий векторы e_1 и e_2 в векторы a_1 и a_2. Его матрицей в базисе e будет

$$T_1 = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}.$$

Оператор φ переводит векторы a_1 и a_2 в векторы b_1 и b_2. Значит, произведение $\varphi\psi$ операторов ψ и φ переводит векторы e_1, e_2 в векторы b_1, b_2. Поэтому в базисе e оператор $\varphi\psi$ имеет матрицу

$$T_2 = \begin{pmatrix} 1 & 3 \\ 2 & 3 \end{pmatrix}.$$

Матрица T_2 оператора $\varphi\psi$ равна произведению матриц операторов φ и ψ, т.е. если A_e — матрица оператора φ в базисе e, то $T_2 = A_e T_1$. Отсюда получаем:

$$A_e = T_2 T_1^{-1} = \begin{pmatrix} 1 & 3 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 4 & 7 \\ 5 & 8 \end{pmatrix}.$$

Пусть φ — линейный оператор, действующий из линейного пространства X в линейное пространство Y. Линейный оператор φ^{-1}, действующий из Y в X, называют ***обратным к оператору*** φ, если $\varphi^{-1}\varphi = \varepsilon$ и $\varphi\varphi^{-1} = \widetilde{\varepsilon}$, где ε и $\widetilde{\varepsilon}$ — тождественные операторы в линейных пространствах X и Y. Другими словами, линейный оператор φ^{-1} обратный к оператору φ, если для каждого $x \in X_n$ выполняется условие $\varphi^{-1}\varphi x = x$ и для каждого $y \in Y_m$ — условие $\varphi\varphi^{-1}y = y$.

Из определения обратного оператора непосредственно вытекает, что если оператор φ имеет обратный, то он осуществляет взаимно однозначное отображение линейного пространства X на линейное пространство Y, в то время как обратный оператор φ^{-1} — взаимно однозначное отображение Y на X. Это значит, что, во-первых, операторы φ и φ^{-1} невырожденные, а во-вторых, линейные пространства X и Y изоморфны и имеют одну и ту же размерность. Отсюда следует, что матрица оператора φ является квадратной невырожденной. Матрицы операторов φ и φ^{-1} являются взаимно обратными.

Если матрица линейного оператора является прямоугольной (размерности линейных пространств X и Y не совпадают) или квадратной, но вырожденной, то этот оператор не имеет обратного. В этих случаях вводят понятие „псевдообратный оператор", обобщающее понятие „обратный оператор". Псевдообратному оператору соответствует псевдообратная матрица (см. разд. 8.20).

5.6. Характеристический и минимальный многочлены

Пусть A — квадратная действительная или комплексная матрица n-го порядка. Матрицу

$$A-\lambda E=\begin{pmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{11} & a_{12}-\lambda & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{1n}-\lambda \end{pmatrix}$$

с переменной λ, принимающей любые числовые значения, называют ***характеристической матрицей*** матрицы A. Её определитель

$$f(\lambda)=|A-\lambda E|$$

представляет собой многочлен от переменной λ степени n. Этот многочлен называют ***характеристическим многочленом*** матрицы A.

То, что характеристический многочлен на самом деле является многочленом от переменной λ, непосредственно вытекает из определения определителя. Наивысшую степень, равную n, среди всех слагаемых определителя $|A-\lambda E|$ имеет произведение

$$(a_{11}-\lambda)(a_{22}-\lambda)\ldots(a_{nn}-\lambda). \tag{5.9}$$

Остальные слагаемые определителя не содержат по крайней мере двух элементов матрицы $A-\lambda E$ с переменным λ и потому имеют степень не выше $n-2$. Поэтому степень многочлена равна n. Отметим, что произведение (5.9) определяет не только степень характеристического многочлена, но и два его слагаемых со старшими степенями

$$(-1)^n\lambda^n \quad \text{и} \quad (-1)^{n-1}(a_{11}+a_{22}+\ldots+a_{nn})\lambda^{n-1}.$$

Свободный член характеристического многочлена совпадает с его значением при $\lambda=0$ и равен $|A-0\cdot E|=|A|$, т.е. определителю матрицы A.

Итак, характеристический многочлен матрицы A порядка n имеет вид:

$$p_0\lambda^n+p_1\lambda^{n-1}+\ldots+p_n, \tag{5.10}$$

где $p_0=(-1)^n$, $p_1=(-1)^{n-1}(a_{11}+a_{22}+\ldots+a_{nn})$, $p_n=|A|$. Заметим, что коэффициент p_1 с точностью до знака, зависящего от порядка n матрицы, совпадает со следом $\operatorname{Sp}A=$ $=a_{11}+a_{22}+\ldots+a_{nn}$ матрицы A.

Корни характеристического многочлена $|a-\lambda E|$ называют ***характеристическими корнями*** или ***характеристическими числами*** матрицы A. Кратность k_i характеристического корня λ_i в характеристическом многочлене называют ***алгебраической кратностью*** этого корня. Множество всех характеристических корней матрицы, в котором каждый характеристический корень повторяется столько раз, какова его кратность, называют ***спектром матрицы*** A. Если все характеристические корни матрицы простые (т.е. имеют единичную кратность), то спектр матрицы называют ***простым***.

В соответствии с формулами Виета коэффициенты характеристического многочлена связаны с характеристическими корнями следующим образом:

$$
\begin{aligned}
p_1 &= \lambda_1 + \lambda_2 + \ldots + \lambda_n, \\
p_2 &= \lambda_1\lambda_2 + \lambda_1\lambda_3 + \ldots + \lambda_{n-1}\lambda_n, \\
&\ldots\ldots\ldots\ldots\ldots\ldots \\
p_n &= \lambda_1\lambda_2\ldots\lambda_n.
\end{aligned}
$$

Из этих формул, в частности, вытекают часто применяемые соотношения

$$\lambda_1 + \lambda_2 + \ldots + \lambda_n = a_{11} + a_{22} + \ldots + a_{nn}, \qquad \lambda_1\lambda_2\ldots\lambda_n = |A|.$$

Согласно последнему равенству характеристический многочлен матрицы имеет нулевые характеристические корни тогда и только тогда, когда определитель этой матрицы равен нулю, т.е. матрица вырожденная.

Пример 5.5. Вычислить характеристический многочлен матрицы

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

Решение. В соответствии с определением характеристического многочлена получаем:

$$|A - \lambda E| = \begin{vmatrix} 1-\lambda & 2 & 0 \\ 0 & 2-\lambda & 0 \\ -2 & -2 & -1-\lambda \end{vmatrix} =$$

$$= (1-\lambda)(2-\lambda)(-1-\lambda) = -\lambda^3 + 2\lambda^2 + \lambda - 2.$$

О методах вычисления характеристического многочлена см. в приложении, помещенном в конце книги.

Теорема 5.7. *Характеристические многочлены подобных матриц совпадают.*

$\triangleright$ Если матрицы A и B подобные, то для некоторой невырожденной матрицы Q выполняется равенство $B = Q^{-1}AQ$.

Следовательно,

$$|B-\lambda E| = |Q^{-1}AQ-\lambda E| = |Q^{-1}(A-\lambda E)Q| =$$
$$= |Q^{-1}||A-\lambda E||Q| = |Q|^{-1}|A-\lambda E||Q| = |A-\lambda E|. \quad \blacktriangleright$$

В произвольный многочлен

$$P(\lambda) = a_0\lambda^n + a_1\lambda^{n-1} + \ldots + a_{n-1}\lambda + a_n$$

вместо переменной λ можно подставить квадратную матрицу A порядка n. В результате получим матрицу $P(A) = a_0A^n + a_1A^{n-1} + \ldots + a_{n-1}A + a_nE$, которую называют значением многочлена $P(\lambda)$ при $\lambda = A$. Если для данной матрицы A верно равенство $P(A) = 0$ (значением многочлена $P(\lambda)$ при $\lambda = A$ является нулевая матрица), то A называют ***матричным корнем многочлена*** $P(\lambda)$, а сам многочлен $P(\lambda)$ — ***многочленом, аннулируемым матрицей*** A.

Теорема 5.8. Всякая квадратная матрица является корнем некоторого ненулевого многочлена.

$\triangleright$ Множество всех квадратных матриц порядка n с элементами из поля P есть линейное пространство над P размерности n^2. В этом линейном пространстве любая система, в которой не менее n^2+1 элементов, является линейно зависимой. Следовательно, система A^{n^2}, A^{n^2-1}, ..., A, E из n^2+1 матриц линейно зависима, т.е. существует такой набор чисел α_0, α_1, ..., α_{n^2}, одновременно не обращающихся в нуль, что выполняется равенство

$$\alpha_0A^{n^2} + \alpha_1A^{n^2-1} + \ldots + \alpha_{n^2-1}A + \alpha_{n^2} = 0.$$

Это равенство означает, что матрица A является корнем многочлена

$$P(\lambda) = \alpha_0\lambda^{n^2} + \alpha_1\lambda^{n^2-1} + \ldots + \alpha_{n^2-1}\lambda + \alpha_{n^2}. \quad \blacktriangleright$$

Доказанная теорема на самом деле вытекает из следующего утверждения.

Теорема 5.9 (теорема Гамильтона — Кэли). *Любая квадратная матрица является корнем своего характеристического многочлена.*

Прежде чем доказывать эту теорему, введем понятие ***λ-матрицы*** — матрицы, элементами которой являются многочлены от переменной λ. Любую λ-матрицу можно представить как многочлен переменной λ, коэффициентами которого являются квадратные матрицы соответствующего порядка. Например,

$$\begin{pmatrix} \lambda^2+2l+3 & 4\lambda \\ 6\lambda & 7\lambda^2+8\lambda+9 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 7 \end{pmatrix}\lambda^2 + \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix}\lambda + \begin{pmatrix} 3 & 5 \\ 0 & 9 \end{pmatrix}.$$

$\triangleright$ Пусть A — квадратная матрица n-го порядка. Рассмотрим присоединенную матрицу C к матрице $A-\lambda E$. Ее элементами являются алгебраические дополнения элементов определителя $|A-\lambda E|$, представляющие собой многочлены от λ степени не выше $n-1$. Как отмечено выше, матрицу C можно представить в виде

$$C = C_1\lambda^{n-1} + C_2\lambda^{n-2} + \ldots + C_{n-1}\lambda + C_n, \tag{5.11}$$

где C_1, C_2, $\ldots$, C_n — некоторые числовые матрицы. По основному свойству присоединенной матрицы (см. следствие 3.2) имеем:

$$C(A-\lambda E) = |A-\lambda E|\,E.$$

В этом равенстве заменим матрицу C суммой (5.11), а характеристический многочлен — суммой (5.10). Тогда получим равенство

$$(C_1\lambda^{n-1} + C_2\lambda^{n-2} + \ldots + C_{n-1}\lambda + C_n)(A-\lambda E) =$$
$$= (-1)^n(\lambda^n - p_1\lambda^{n-1} + p_2\lambda^{n-2} - \ldots \pm p_n)E.$$

Раскрывая скобки в обеих частях равенства и приравнивая коэффициенты при одинаковых степенях, получим систему из

$n+1$ равенства:

$$
\begin{aligned}
&-C_1 = (-1)^n E,\\
&C_1 A - C_2 = (-1)^{n-1} p_1 E,\\
&\ldots\ldots\ldots\ldots\ldots\\
&C_{n-1} A - C_n = (-1) p_{n-1} E,\\
&C_n A = p_n E.
\end{aligned}
$$

Умножим первое равенство системы на A^n, второе — на A^{n-1} и т.д., n-е равенство — на A, $(n+1)$-е равенство — на $A^0 = E$:

$$
\begin{aligned}
&-C_1 A^n = (-1)^n A^n,\\
&C_1 A^n - C_2 A^{n-1} = (-1)^{n-1} p_1 A^{n-1},\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&C_{n-1} A^2 - C_n A = (-1) p_{n-1} A^2,\\
&C_n A = p_n E.
\end{aligned}
$$

При сложении этих равенств в левой части получим нулевую матрицу, а в правой части — выражение

$$(-1)^n (A^n - p_1 A^{n-1} + p_2 A^{n-2} - \ldots \pm p_n) = f(A).$$

Поэтому $f(A) = 0$. ▶

Многочлен $\varphi(\lambda)$ минимальной степени, имеющий старший коэффициент, равный единице, и аннулируемый матрицей A, называют ***минимальным многочленом*** этой матрицы.

Теорема 5.10. *Любой многочлен, аннулируемый матрицей A, нацело делится на минимальный многочлен этой матрицы. В частности, характеристический многочлен матрицы делится на ее минимальный многочлен.*

▷ Разделим многочлен $P(\lambda)$ на минимальный многочлен $\varphi(\lambda)$ с остатком: $P(\lambda) = \varphi(\lambda) q(\lambda) + r(\lambda)$, где многочлен $r(\lambda)$ имеет степень меньше степени $\varphi(\lambda)$. Заменив переменную λ матрицей A, получим:

$$P(A) = \varphi(A) q(A) + r(A).$$

Так как $P(A) = \varphi(A) = 0$, то и $r(A) = 0$. Но это равенство возможно только в том случае, когда многочлен $r(\lambda)$ нулевой. Иначе возникает противоречие с определением минимального многочлена. Равенство $r = 0$ означает, что многочлен $P(\lambda)$ нацело делится на $\varphi(\lambda)$. ▶

Следствие 5.1. Любой корень минимального многочлена матрицы является корнем ее характеристического многочлена.

▷ Как установлено при доказательстве теоремы, характеристический многочлен $f(\lambda)$ связан с минимальным многочленом $\varphi(\lambda)$ равенством $f(\lambda) = \varphi(\lambda)q(\lambda)$. Из этого равенства вытекает утверждение следствия. ▶

Отметим еще несколько полезных фактов.

Характеристический многочлен $|A - \lambda E|$ матрицы A и ее минимальный многочлен $\varphi(\lambda)$ связаны соотношением

$$\varphi(\lambda) = \frac{(-1)^n |A - \lambda E|}{D_{n-1}}, \tag{5.12}$$

где D_{n-1} — наибольший общий делитель всех миноров матрицы $A - \lambda E$, имеющих $(n-1)$-й порядок.

Корнями минимального многочлена $\varphi(\lambda)$ являются все различные корни характеристического многочлена $|A - \lambda E|$, причем если

$$|A - \lambda E| = (-1)^n (\lambda - \lambda_1)^{m_1} (\lambda - \lambda_2)^{m_2} \ldots (\lambda - \lambda_s)^{m_s},$$

то

$$\varphi(\lambda) = (\lambda - \lambda_1)^{n_1} (\lambda - \lambda_2)^{n_2} \ldots (\lambda - \lambda_s)^{n_s}, \tag{5.13}$$

где $1 \leqslant n_k \leqslant m_k$, $k = 1, 2, \ldots, s$.

Формула (5.12) позволяет находить минимальный многочлен матрицы. О другом способе построения минимального многочлена матрицы сказано ниже (см. разд. 6.5).

Пример 5.6. Найти минимальный многочлен матрицы

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

Решение. В предыдущих примерах для матрицы A найден характеристический многочлен $|A - \lambda E| = -\lambda^3 + 2\lambda^2 + \lambda - 2$. Общий наибольший делитель D_2 всех миноров второго порядка матрицы

$$A - \lambda E = \begin{pmatrix} 1-\lambda & 2 & 0 \\ 0 & 2-\lambda & 0 \\ -2 & -2 & -1-\lambda \end{pmatrix}$$

равен единице, так как ее миноры

$$\begin{vmatrix} 1-\lambda & 2 \\ -2 & -2 \end{vmatrix} = 2(\lambda+1), \qquad \begin{vmatrix} 0 & 2-\lambda \\ -2 & -2 \end{vmatrix} = 2(2-\lambda)$$

взаимно простые. Поэтому

$$\varphi(\lambda) = \frac{(-1)^3|A-E|}{D_2} = \lambda^3 - 2\lambda^2 - \lambda + 2.$$

Пример 5.7. Найти характеристические и минимальные многочлены матриц

$$A_1 = \begin{pmatrix} 6 & 2 & -2 \\ -2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 6 & 2 & 2 \\ -2 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Решение. Для матрицы A_1 непосредственным вычислением определителя находим характеристический многочлен

$$|A_1 - \lambda E| = \begin{vmatrix} 6-\lambda & 2 & -2 \\ -2 & 2-\lambda & 2 \\ 2 & 2 & 2-\lambda \end{vmatrix} = (\lambda-4)^2(2-\lambda).$$

Выпишем все миноры второго порядка матрицы $A_1 - \lambda E$:

$$\begin{vmatrix} 6-\lambda & 2 \\ -2 & 2-\lambda \end{vmatrix} = (\lambda-4)^2, \qquad \begin{vmatrix} 6-\lambda & -2 \\ -2 & 2 \end{vmatrix} = 2(\lambda-4),$$

$$\begin{vmatrix} 2 & -2 \\ 2-\lambda & 2 \end{vmatrix} = 2(4-\lambda), \qquad \begin{vmatrix} 6-\lambda & 2 \\ 2 & 2 \end{vmatrix} = 2(4-\lambda),$$

$$\begin{vmatrix} 6-\lambda & -2 \\ 2 & 2-\lambda \end{vmatrix} = (\lambda-4)^2, \qquad \begin{vmatrix} 2 & -2 \\ 2 & 2-\lambda \end{vmatrix} = 2(4-\lambda),$$

$$\begin{vmatrix} -2 & 2-\lambda \\ 2 & 2 \end{vmatrix} = 2(\lambda-4), \qquad \begin{vmatrix} -2 & -2 \\ 2 & 2-\lambda \end{vmatrix} = 2(\lambda-4),$$

$$\begin{vmatrix} 2-\lambda & 2 \\ 2 & 2-\lambda \end{vmatrix} = \lambda(\lambda-4).$$

Общий наибольший делитель D_2 всех этих миноров есть $\lambda - 4$. Поэтому минимальный многочлен матрицы A_1 имеет вид:

$$\varphi(\lambda) = \frac{(-1)^3|A_1 - E|}{D_2} = \frac{(-1)^3(\lambda-4)^2(2-\lambda)}{\lambda-4} = (\lambda-4)(\lambda-2).$$

Заметим, что D_2 можно найти иначе. Действительно, если в матрицу $A_1 - \lambda E$ подставить $\lambda = 4$, то получим матрицу

$$\begin{pmatrix} 2 & 2 & -2 \\ -2 & -2 & 2 \\ 2 & 2 & -2 \end{pmatrix}$$

ранга $r = 1$. Следовательно, все миноры второго порядка этой матрицы равны нулю. Это означает, что все миноры второго порядка матрицы $A_1 - \lambda E$ делятся на $\lambda - 4$, причем все эти миноры не могут делиться на большую степень двучлена $\lambda - 4$, так как, например, минор $\begin{vmatrix} 2 & -2 \\ 2-\lambda & 2 \end{vmatrix} = 2(4-\lambda)$ делится лишь на первую степень этого двучлена. Следовательно, в D_2 входит множитель $\lambda - 4$ в первой степени. Другие множители из $|A_1 - \lambda E|$ в D_2 не входят, так как на них не делится, например, выписанный только что минор второго порядка. Поэтому $D_2 = \lambda - 4$.

Для матрицы A_2 также непосредственным вычислением определителя находим характеристический многочлен

$$|A_2 - \lambda E| = \begin{vmatrix} 6-\lambda & 2 & 2 \\ -2 & 2-\lambda & 0 \\ 0 & 0 & 2-\lambda \end{vmatrix} = (\lambda-4)^2(2-\lambda).$$

Далее замечаем, что в матрице

$$A_2 - \lambda E = \begin{pmatrix} 6-\lambda & 2 & 2 \\ -2 & 2-\lambda & 0 \\ 0 & 0 & 2-\lambda \end{pmatrix}$$

миноры второго порядка

$$\begin{vmatrix} 6-\lambda & 2 \\ -2 & 0 \end{vmatrix}, \qquad \begin{vmatrix} 2 & 2 \\ 2-\lambda & 0 \end{vmatrix}$$

взаимно простые. Поэтому $D_2 = 1$ и

$$\varphi_2(\lambda) = (-1)^3|A_2 - \lambda E| = (\lambda-4)^2(\lambda-2).$$

Рассмотренный пример показывает, что разные матрицы могут иметь одинаковые характеристические, но разные минимальные многочлены.

Учитывая, что матрицы данного линейного оператора в разных базисах подобны и имеют один и тот же характеристический многочлен, логично этот многочлен назвать ***характеристическим многочленом линейного оператора***, а его корни — ***характеристическими корнями линейного оператора***.

Заметим также, что транспонированная матрица A^{T} имеет одинаковые с матрицей A характеристические многочлены и характеристические числа.

5.7. Собственные векторы и собственные значения линейного оператора

Пусть линейный оператор φ, действующий в линейном пространстве X над полем P, имеет в некотором базисе e матрицу A. Ненулевой вектор $x \in X$ называют ***собственным вектором оператора*** φ, если этим оператором он переводится в вектор λx, т.е. $\varphi x = \lambda x$, где λ — некоторое число из поля P, называемое ***собственным значением оператора*** φ. При этом говорят, что ***собственный вектор*** x ***принадлежит собственному значению*** λ.

Теорема 5.11. *Собственными значениями линейного оператора φ, действующего в линейном пространстве X над полем P, являются характеристические корни этого оператора, принадлежащие P, и только они.*

$\triangleright$ Пусть линейный оператор φ в некотором базисе e имеет матрицу A. Вектор x является собственным вектором линейного оператора φ, если выполняется равенство $\varphi x = \lambda x$, где λ — собственное значение оператора φ, отвечающее собственному вектору x. Это равенство можно записать в виде $(\varphi - \lambda\varepsilon)x = 0$, где ε — тождественный оператор. Переходя к представлению векторов и оператора в заданном базисе, мы приходим к записи этого равенства в матричном виде

$$(A - \lambda E)[x] = 0, \tag{5.14}$$

где $[x]$ — столбец координат вектора x.

Матричное равенство (5.14) представляет собой однородную систему линейных уравнений относительно элементов столбца $[x]$. Так как вектор x ненулевой, система (5.14) имеет ненулевые решения, а для этого необходимо и достаточно, чтобы определитель системы был равен нулю, т.е. $\det(A - \lambda E) = 0$. Это значит, что число λ является корнем характеристического многочлена.

Если число λ — корень характеристического многочлена, то выполняется равенство $\det(A - \lambda E) = 0$. Следовательно, в однородной системе (5.14) определитель равен нулю и эта система имеет ненулевые решения. Каждому такому решению соответствует вектор x, удовлетворяющий условию $(\varphi-\lambda\varepsilon)x=0$, а потому по определению *являющийся* собственным вектором, которому отвечает собственное значение λ. ►

Следствие 5.2. В конечномерном комплексном линейном пространстве собственными значениями линейного оператора являются все его характеристические корни, и только они. В конечномерном действительном линейном пространстве собственными значениями линейного оператора являются все его действительные характеристические корни, и только они.

Из следствия вытекает следующее правило. *Для отыскания всех собственных значений оператора с матрицей A нужно найти все характеристические числа матрицы A и из них выбрать лишь те, которые принадлежат основному полю, а для отыскания всех собственных векторов оператора с матрицей A нужно для каждого собственного значения λ найти все ненулевые решения системы $(A-\lambda E)x=0$.*

Алгебраической кратностью собственного значения линейного оператора называют кратность соответствующего корня характеристического многочлена.

Пример 5.8. Для оператора в действительном линейном пространстве, имеющего в заданном базисе матрицу

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

найти собственные значения и собственные векторы.

Решение. Характеристический многочлен матрицы A имеет вид:

$$|A-\lambda E|=\begin{vmatrix}3-\lambda & 1 & -3\\ 3 & 1-\lambda & -1\\ 2 & -2 & -\lambda\end{vmatrix}=$$

$$=-\lambda(3-\lambda)(1-\lambda)+18+6(1-\lambda)-2(3-\lambda)+3\lambda=(\lambda^2+4)(4-\lambda).$$

Его корни $\lambda_1 = 4$ и $\lambda_{2,3} = \pm 2i$. Так как рассматриваемый оператор действует в действительном линейном пространстве, то его собственным значением будет лишь $\lambda = 4$. При этом значении λ система $(A - \lambda E)X = 0$ принимает вид:

$$\begin{cases} -x_1 + x_2 - 3x_3 = 0, \\ 3x_1 - 3x_2 - x_3 = 0, \\ 2x_1 - 2x_2 - 4x_3 = 0. \end{cases}$$

Ее общее решение можно записать в виде $X = (x_1,\ x_1,\ 0)^{\mathrm{т}}$, где x_1 — произвольная постоянная. Если x_1 пробегает все действительные значения, то общее решение дает все собственные векторы оператора с матрицей A, принадлежащие собственному значению $\lambda = 4$. Других собственных векторов оператор с матрицей A не имеет, так как у него нет других собственных значений.

Теорема 5.12. *Множество L_λ собственных векторов линейного оператора φ, отвечающих собственному значению λ, вместе с нулевым вектором линейного пространства является линейным подпространством.*

$\triangleright$ Если $x, y \in L_\lambda$, то $\varphi x = \lambda x$, $\varphi y = \lambda y$. Поэтому

$$\varphi(x + y) = \varphi x + \varphi y = \lambda x + \lambda y = \lambda(x + y),$$

т.е. вектор $x + y$ является собственным вектором, отвечающим собственному значению λ. Аналогично для любого вектора $x \in L_\lambda$ и любого числа α

$$\varphi(\alpha x) = \alpha\varphi x = \alpha\lambda x = \lambda(\alpha x),$$

т.е. вектор αx является собственным вектором, отвечающим собственному значению λ. ▶

Следствие 5.3. Любая линейная комбинация собственных векторов линейного оператора φ, отвечающих одному и тому же собственному значению λ, является собственным вектором этого оператора, отвечающим тому же собственному значению.

Подпространство L_λ, которое образовано собственными векторами, отвечающими собственному значению λ, называют ***собственным подпространством оператора***. Оно представляет собой ядро линейного оператора $\varphi - \lambda\varepsilon$. Размерность этого подпространства равна $n - r_\lambda$, где n — размерность линейного пространства, в котором действует линейный оператор, а r_λ — ранг линейного оператора $\varphi - \lambda\varepsilon$.

Собственное подпространство является частным случаем ***инвариантного подпространства линейного оператора***, т.е. такого подпространства L, что образом каждого вектора в L является вектор, также принадлежащий L. Так как любой многочлен имеет комплексные или действительные корни, то любой линейный оператор в комплексном линейном пространстве имеет собственные векторы. Значит, у каждого такого оператора есть одномерное инвариантное подпространство. В случае действительного линейного пространства это уже не так: линейный оператор поворота в $\mathbb{R}^2$ на угол $\vartheta < \pi$ не имеет собственных векторов, а поэтому не имеет и одномерных инвариантных (собственных) подпространств.

Теорема 5.13. *Любой линейный оператор, действующий в действительном линейном пространстве X, имеет одно- или двумерное инвариантное подпространство.*

$\triangleright$ Если действительное линейное пространство X одномерно, то утверждение теоремы тривиально, так как все линейное пространство можно рассматривать как инвариантное подпространство. Поэтому остановимся на случае, когда размерность n линейного пространства не меньше двух.

Если характеристический многочлен линейного оператора φ, действующего в X, имеет действительные корни, то каждый такой корень является собственным значением. Любой собственный вектор, отвечающий собственному значению, порождает одномерное инвариантное подпространство. Таким образом, в этом случае теорема верна.

Пусть все характеристические корни линейного оператора φ комплексные и $\lambda = \alpha + i\beta$ — один из них. Рассмотрим систему $(A-\lambda E)x=0$, где A — матрица линейного оператора φ в некотором базисе. Эта система имеет комплексные коэффициенты и имеет комплексные решения. Представим столбец x в виде $x=u+iv$, где столбцы u и v состоят из действительных чисел. Подставив это равенство в систему, получим

$$A(u+iv)=(\alpha+i\beta)(u+iv),$$

или

$$Au+iAv=(\alpha u-\beta v)+i(\beta u+\alpha v).$$

Комплексное равенство распадается на два действительных:

$$Au=\alpha u-\beta v \quad \text{и} \quad Av=\beta u+\alpha v. \tag{5.15}$$

Эти два матричных равенства в совокупности составляют однородную систему линейных уравнений. Поскольку система комплексных уравнений $(A-\lambda E)x=0$ имеет ненулевые решения (определитель этой системы равен нулю), то и система действительных уравнений (5.15) имеет ненулевые решения. Пусть столбцы u и v составляют одно из ненулевых решений. Рассмотрим векторы, координатными столбцами которых являются столбцы u и v. Эти векторы также обозначим через u и v. Покажем, что подпространство $L=\langle u,\ v\rangle$ инвариантно относительно оператора φ. В силу системы (5.15) выполняются равенства $\varphi u=\alpha u-\beta v$, $\varphi v=\beta u+\alpha v$. Следовательно, образы векторов u и v при действии φ принадлежат L. Но тогда и образ любой линейной комбинации векторов u и v принадлежит L. Убедимся в этом непосредственно. Пусть $x=k_1u+k_2v\in L$. Тогда:

$$\varphi x=\varphi(k_1u+k_2v)=k_1\varphi u+k_2\varphi v,$$

т.е. вектор φx является линейной комбинацией двух векторов φu и φv, принадлежащих подпространству L. Значит, $\varphi x\in L$.

В принципе теорема доказана, поскольку подпространство, натянутое на два вектора, имеет размерность не более двух. Однако отметим, что подпространство L на самом деле двумерное и не может быть одномерным. Последнее возможно в случае, когда векторы u и v линейно зависимы, т.е. выполняется равенство $u = kv$. Но тогда равенство верно и для столбцов координат этих векторов. При этом $u + iv = (k + i)v$, откуда следует, что столбец v с действительными элементами является решением системы $(A - \lambda E)v = 0$, равносильной системе $Av = \lambda v$. Но система $Av = \lambda v$ не может иметь действительных решений при комплексном λ, так как для действительного столбца v матричное произведение Av дает столбец с действительными элементами, а произведение λv имеет комплексные элементы. ▶

Заметим, что в силу соотношений (5.15), а также (5.1) и (5.2) матрицей сужения линейного оператора φ на подпространство $L = \langle u,\ v\rangle$ в базисе $e = (u,\ v)$ будет матрица $\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$.

Теорема 5.14. *Пусть линейное пространство X распадается в прямую сумму $X = L_1 \oplus L_2$ инвариантных подпространств линейного оператора φ, действующего в X. Если $e_1,\ e_2,\ \ldots,\ e_k$ — базис в L_1, а $e_{k+1},\ e_{k+2},\ \ldots,\ e_n$ — базис в L_2, то в базисе $e_1, e_2,\ \ldots,\ e_n$ матрица оператора φ имеет блочно-диагональный вид:*

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}. \tag{5.16}$$

При этом блок A_2 — это матрица сужения оператора φ на подпространство L_1 в базисе $e_1,\ e_2,\ \ldots,\ e_k$, а A_2 — это матрица сужения оператора φ на подпространство L_2 в базисе $e_{k+1},\ e_{k+2},\ \ldots,\ e_n$.

▷ В силу инвариантности подпространства L_1 имеем:

$$\begin{aligned} \varphi e_1 &= a_{11}e_1 + \ldots + a_{k1}e_k, \\ \varphi e_2 &= a_{12}e_1 + \ldots + a_{k2}e_k, \\ &\ldots\ldots\ldots\ldots\ldots \\ \varphi e_k &= a_{1k}e_1 + \ldots + a_{kk}e_k. \end{aligned}$$

Эти разложения означают, что в первых k столбцах матрицы A линейного оператора φ элементы, находящиеся в строках ниже k-й, равны нулю. Аналогично из инвариантности подпространства L_2 вытекает, что

$$\begin{aligned}
\varphi e_{k+1} &= a_{k+1,k+1}e_{k+1} + \ldots + a_{n,k+1}e_n,\\
\varphi e_{k+2} &= a_{k+1,k+2}e_{k+1} + \ldots + a_{n,k+2}e_n,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
\varphi e_n &= a_{k+1,n}e_{k+1} + \ldots + a_{nn}e_n,\\
\varphi e_k &= a_{1k}e_1 + \ldots + a_{kk}e_k.
\end{aligned}$$

В соответствующих столбцах матрицы A равны нулю элементы, расположенные в строках от первой до k-й. Из этих представлений запишем матрицу

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1k} & 0 & \ldots & 0\\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots\\ a_{k1} & \ldots & a_{kk} & 0 & \ldots & 0\\ 0 & \ldots & 0 & a_{k+1,k+1} & \ldots & a_{k+1,n}\\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots\\ 0 & \ldots & 0 & a_{n,k+1} & \ldots & a_{nn} \end{pmatrix} = \begin{pmatrix} A_1 & 0\\ 0 & A_2 \end{pmatrix}. \quad \blacktriangleright$$

Отметим, что верно и утверждение, обратное доказанной теореме: если матрица A оператора φ имеет блочно-диагональный вид (5.16), то подпространства, натянутые на две группы векторов базиса, соответствующие блокам A_1 и A_2, являются инвариантными, а линейное пространство распадается в прямую сумму этих подпространств.

Геометрической кратностью собственного значения λ линейного оператора φ называют размерность собственного подпространства этого оператора, отвечающего собственному значению λ. Можно также сказать, что геометрическая кратность собственного значения — это максимальное число линейно независимых собственных векторов, отвечающих этому собственному значению. Геометрическая кратность l_λ собственного значения λ может быть выражена через ранг r_λ линейного оператора $\varphi - \lambda\varepsilon$: $l_\lambda = n - r_\lambda$, где n — размерность линейного пространства.

Теорема 5.15. *Геометрическая кратность собственного значения линейного оператора не превышает его алгебраической кратности.*

▷ Пусть в линейном пространстве X действует линейный оператор φ, для которого λ_0 — собственное значение с геометрической кратностью l. В соответствии с определением геометрической кратности существует l линейно независимых собственных векторов $e_1, e_2, \ldots, e_l$, отвечающих собственному значению λ_0. Систему этих векторов можно дополнить до базиса $e_1, e_2, \ldots, e_n$ в X. В этом базисе матрица линейного оператора имеет вид:

$$B = \left(\begin{array}{c|c} \begin{matrix} \lambda_0 & & 0 \\ & \ddots & \\ 0 & & \lambda_0 \end{matrix} & R \\ \hline 0 & Q \end{array}\right).$$

Характеристический многочлен $|B - \lambda E|$ по теореме Лапласа можно представить в виде

$$|B - \lambda E| = \left|\begin{array}{c|c} \begin{matrix} \lambda_0 - \lambda & & 0 \\ & \ddots & \\ 0 & & \lambda_0 - \lambda \end{matrix} & R \\ \hline 0 & Q - \lambda E \end{array}\right| = (\lambda_0 - \lambda)^l |Q - \lambda E|.$$

Отсюда следует, что характеристический многочлен $|B - \lambda E|$ имеет корень λ_0 кратности, не меньшей l. Значит, геометрическая кратность l собственного значения λ_0 не превышает его алгебраической кратности. ▶

Если отождествлять оператор с его матрицей, то естественно говорить о собственных значениях и собственных векторах матрицы. На практике так обычно и поступают.

Квадратную матрицу называют ***простой***, если для каждого собственного значения матрицы его геометрическая кратность совпадает с алгебраической кратностью. В противном случае матрицу называют ***дефектной***.

Теорема 5.16. *Собственные векторы $b_1, b_2, \ldots, b_k$ линейного оператора, отвечающие различным собственным значениям $\lambda_1, \lambda_2, \ldots, \lambda_k$, линейно независимы.*

$\triangleright$ Доказательство проводится методом математической индукции по количеству k собственных векторов. При $k=1$ утверждение теоремы очевидно, так как $b_1 \neq 0$. Допустим, что утверждение верно для систем из $k-1$ собственных векторов. Рассмотрим систему $b_1, b_2, \ldots, b_k$ и ее линейную комбинацию, равную нулю:

$$\alpha_1 b_1 + \alpha_2 b_2 + \ldots + \alpha_k b_k = 0. \tag{5.17}$$

Применим к обеим частям этого равенства линейный оператор φ и учтем, что все векторы для оператора собственные:

$$\lambda_1\alpha_1 b_1 + \lambda_2\alpha_2 b_2 + \ldots + \lambda_k\alpha_k b_k = 0. \tag{5.18}$$

Из равенства (5.18) вычтем равенство (5.17), умноженное на λ_k:

$$(\lambda_1 - \lambda_k)\alpha_1 b_1 + (\lambda_2 - \lambda_k)\alpha_2 b_2 + \ldots \\ \ldots + (\lambda_{k-1} - \lambda_k)\alpha_{k-1} b_{k-1} = 0. \tag{5.19}$$

Согласно предположению математической индукции векторы $b_1, \ldots, b_{k-1}$ линейно независимы. Поэтому из равенства (5.19) вытекает, что все коэффициенты в левой части равны нулю, т.е. $(\lambda_i - \lambda_k)\alpha_i = 0$ для $i = 1, 2, \ldots, k-1$. С учетом различия всех собственных значений λ_i делаем вывод, что $\alpha_1 = \alpha_2 = \ldots = \alpha_{k-1} = 0$. Но тогда в равенстве (5.17) можно опустить все слагаемые, кроме последнего, т.е. $\alpha_k b_k = 0$. Поскольку все собственные векторы ненулевые, заключаем, что $\alpha_k = 0$. Итак, из равенства (5.17) вытекает, что все коэффициенты линейной комбинации в левой части равенства равны нулю. Это означает, что векторы b_1, b_2, ..., b_k линейно независимы. Мы тем самым доказали утверждение теоремы для k векторов в предположении, что оно верно для $k-1$ вектора. Согласно методу математической индукции утверждение теоремы верно для любого числа собственных векторов. ▶

Собственный вектор x, определенный условием $Ax = \lambda x$, также называют ***правым собственным вектором матрицы A***, принадлежащим собственному значению λ. Рассматривают и левые собственные векторы матрицы. Ненулевой вектор y называют ***левым собственным вектором матрицы A***, принадлежащим собственному значению λ, если $y^*A = \lambda y^*$. Напомним, что здесь символ $*$ означает транспонирование и комплексное сопряжение. Если это равенство подвергнуть транспонированию, то придем к равенству $A^{\mathrm{T}}\overline{y} = \lambda\overline{y}$. Если провести еще и комплексное сопряжение, то получим: $A^*y = \overline{\lambda}y$. Это означает, что левый собственный вектор y матрицы A, принадлежащий собственному значению λ, является правым собственным вектором матрицы A^* по $\overline{\lambda}$, а вектор $\overline{y}$ — правым собственным вектором матрицы A^{T} по λ. В дальнейшем, говоря о собственных векторах без указаний „правый“, „левый“, будем понимать, что речь идет о правых собственных векторах.

5.8. Линейные операторы простой структуры

Линейный оператор, действующий в линейном пространстве X, называют ***оператором простой структуры***, если в X существует базис, состоящий из собственных векторов этого оператора.

Теорема 5.17. *Линейный оператор φ в заданном базисе e_1, e_2, ..., e_n имеет диагональную матрицу*

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

тогда и только тогда, когда все векторы этого базиса являются собственными векторами, отвечающими собственным значениям λ_1, λ_2, ..., λ_n.

$\triangleright$ Пусть все векторы e_1, e_2, ..., e_n базиса e являются собственными векторами линейного оператора φ, отвечающими собственным значениям λ_1, λ_2, ..., λ_n. Тогда

$$\begin{cases} \varphi e_1 = \lambda_1 e_1 = \lambda_1 e_1 + 0 \cdot e_2 + \ldots + 0 \cdot e_n, \\ \varphi e_2 = \lambda_2 e_2 = 0 \cdot e_1 + \lambda_2 e_2 + \ldots + 0 \cdot e_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \varphi e_n = \lambda_n e_n = 0 \cdot e_1 + 0 \cdot e_2 + \ldots + \lambda_n e_n. \end{cases} \tag{5.20}$$

Столбцами матрицы оператора в данном базисе e являются столбцы координат векторов φe_1, φe_2, ..., φe_n в этом базисе, т.е. коэффициенты из разложений (5.20). Поэтому в базисе e линейный оператор φ имеет матрицу

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}.$$

Наоборот, пусть в базисе $e_1, e_2, \ldots, e_n$ линейный оператор φ имеет матрицу Λ. Тогда для вектора e_i, $i = 1, 2, \ldots, n$, с координатным столбцом $(0, \ldots, 0, 1, 0, \ldots, 0)^{\mathrm{T}}$ образ φe_i будет иметь столбец координат

$$\Lambda \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ \lambda_i \\ \vdots \\ 0 \end{pmatrix}.$$

Поэтому $\varphi e_i = \lambda_i e_i$, $i = 1, 2, \ldots, n$. Это означает, что векторы $e_1, e_2, \ldots, e_n$ являются собственными векторами оператора φ, отвечающими собственным значениям λ_1, λ_2, ..., λ_n. $\blacktriangleright$

С л е д с т в и е 5.4. Линейный оператор простой структуры в базисе, состоящем из собственных векторов, имеет диагональную матрицу, в которой по диагонали стоят собственные значения этого оператора.

Если матрица оператора φ в некотором базисе имеет диагональную матрицу, то говорят, что матрица этого оператора приводится к диагональному виду.

Т е о р е м а 5.18. *Для того чтобы в линейном пространстве X линейный оператор φ имел базис из собственных векторов, необходимо и достаточно, чтобы все характеристические числа λ_i оператора φ принадлежали основному полю и чтобы каждому числу λ_i соответствовало столько линейно независимых собственных векторов оператора φ, какова алгебраическая кратность корня λ_i характеристического многочлена оператора φ.*

$\triangleright$ Утверждение о том, что числа λ_i должны принадлежать основному полю, обеспечивает наличие в линейном пространстве X собственных векторов оператора φ, отвечающих собственному числу λ_i. Докажем утверждение о совпадении алгебраической и геометрической кратностей числа λ_i.

Пусть $\lambda_1, \lambda_2, \ldots, \lambda_s$ — все различные характеристические числа матрицы A, а $k_1, k_2, \ldots, k_s$ — их алгебраические кратности. Тогда $k_1 + k_2 + \ldots + k_s = n$. Предположим, что существует базис из собственных векторов оператора φ и пусть числу λ_i отвечает l_i линейно независимых собственных векторов оператора φ. Тогда $l_i \leqslant k_i$, $i = 1, 2, \ldots, s$ (см. теорему 5.15). В то же время выполняются равенства

$$l_1 + l_2 + \ldots + l_s = n = k_1 + k_2 + \ldots + k_s.$$

Это возможно лишь в случае, когда каждое неравенство $l_i \leqslant k_i$ фактически является равенством, т.е. $l_i \leqslant k_i$, $i = 1, 2, \ldots, s$. $\blacktriangleright$

Доказанную теорему можно переформулировать для матриц следующим образом. ***Для того чтобы матрица A была матрицей простой структуры, т.е. чтобы она***

приводилась к диагональному виду, необходимо и достаточно, чтобы все характеристические корни λ_i матрицы A принадлежали основному полю и чтобы для каждого корня λ_i его геометрическая кратность $l_i = n - r_i$ (r_i — ранг матрицы $A - \lambda_i E$) совпадала с алгебраической кратностью k_i, т.е. с кратностью корня λ_i характеристического многочлена $|A - \lambda E|$ матрицы A.

Линейный оператор называют ***оператором простого спектра***, если все его характеристические числа различные и принадлежат основному полю.

Теорема 5.19. *Линейный оператор простого спектра является оператором простой структуры.*

$\triangleright$ Действительно, по условию все характеристические корни оператора φ различные и принадлежат основному полю. Следовательно, они являются собственными значениями этого оператора. Тогда в линейном пространстве существует базис, состоящий из собственных векторов оператора φ. Поэтому оператор φ является оператором простой структуры. $\blacktriangleright$

Теорему 5.19 для матриц можно переформулировать следующим образом. ***Квадратная матрица с элементами из поля P, все характеристические числа которой различны и принадлежат полю P, подобна диагональной матрице, т.е. приводится к диагональному виду.***

Пусть линейный оператор φ в базисе e имеет матрицу A, а в базисе e', состоящем из его собственных векторов, — матрицу Λ. Тогда, учитывая связь (5.8) между матрицами оператора в разных базисах, имеем

$$T^{-1}AT = \Lambda, \tag{5.21}$$

где T — матрица перехода от базиса e к базису e'. В данном базисе столбцами матрицы T являются столбцы координат собственных векторов e'_1, e'_2, ..., e'_n в базисе e. Из соотношения (5.21) получается соотношение

$$A = T\Lambda T^{-1}, \tag{5.22}$$

которое называют ***каноническим*** или ***спектральным разложением*** матрицы A. Таким образом, матрица оператора простой структуры имеет каноническое (спектральное) разложение.

При построении матрицы T для соотношений (5.21) и (5.22) нужно найти все собственные значения матрицы A и при каждом собственном значении λ_i построить фундаментальную систему решений однородной системы уравнений $(A - \lambda_i E)X = 0$; из решений всех построенных ФСР, как из столбцов, составить матрицу T. При этом в матрице T столбцами записываются решения по каждому λ_i в порядке нумерации собственных значений $\lambda_1, \lambda_2, \ldots, \lambda_n$ (одинаковые λ_i считаются столько раз, каковы их крайности; все λ_i можно занумеровать так, что будет $\lambda_1, \lambda_2, \ldots, \lambda_r \neq 0$, $\lambda_{r+1} = \ldots = \lambda_n = 0$). Если матрица T окажется квадратной, то она будет удовлетворять соотношениям (5.21) и (5.22). Если же матрица T окажется неквадратной, то соотношения (5.21) и (5.22) для матрицы A будут невозможны, т.е. матрица A не приводится к диагональному виду и, следовательно, не имеет канонического разложения. Этот способ равносилен нахождению невырожденной матрицы T из матричного уравнения $AT = T\Lambda$. Из правила построения матрицы T видно, что она будет квадратной лишь в случае, когда каждое характеристическое число λ_i матрицы A является ее собственным значением и для каждого λ_i совпадает его алгебраическая кратность с геометрической кратностью, т.е. с максимальным числом линейно независимых собственных векторов матрицы A по λ_i, равным $l_i = n - r_i$, где r_i — ранг матрицы $A - \lambda_i E$. Лишь в таком случае оператор с матрицей A является оператором простой структуры, а матрица A приводится к диагональному виду.

Пример 5.9. Привести, если возможно, действительные матрицы

$$A_1 = \begin{pmatrix} -1 & 3 & -1 \\ -3 & 5 & -1 \\ -3 & 3 & 1 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 0 & 3 & 3 \\ -1 & 8 & 6 \\ 2 & -14 & -10 \end{pmatrix}$$

к диагональному виду и построить для них канонические разложения.

Решение. а) Корнями характеристического многочлена

$$|A_1-\lambda E|=\begin{vmatrix}1-\lambda & 3 & -1\\ -3 & 5-\lambda & -1\\ -3 & 3 & 1-\lambda\end{vmatrix}=-(\lambda-2)^2(\lambda-1)$$

матрицы A_1 являются числа $\lambda_1=2$, $\lambda_2=1$ соответственно кратности $k_1=2$, $k_2=1$. Все они действительные и потому являются собственными значениями матрицы A_1. При $\lambda_1=2$ матрица

$$A_1-\lambda_1E=\begin{pmatrix}-3 & 3 & -1\\ -3 & 3 & -1\\ -3 & 3 & -1\end{pmatrix}$$

имеет ранг $r_1=1$, и потому $l_1=n-r_1=3-1=2=k_1$.

При $\lambda_2=1$ матрица

$$A_1-\lambda_2E=\begin{pmatrix}-2 & 3 & -1\\ -3 & 4 & -1\\ -3 & 3 & 0\end{pmatrix}$$

имеет ранг $r_2=2$, и потому $l_2=n-r_2=3-2=1=k_2$.

Таким образом, у матрицы A_1 геометрическая кратность каждого λ_i совпадает с его алгебраической кратностью. Поэтому матрица A_1 приводится к диагональному виду

$$\Lambda=\begin{pmatrix}2 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 1\end{pmatrix}.$$

В этом можно убедиться и непосредственным конструированием матрицы T, удовлетворяющей соотношению $T^{-1}A_1T=\Lambda$. Действительно, при $\lambda=2$ система $(A_1-\lambda E)X=0$, т.е. система

$$\begin{cases}-3x_1+3x_2-x_3=0,\\ -3x_1+3x_2-x_3=0,\\ -3x_1+3x_2-x_3=0\end{cases}$$

имеет общее решение $X=(x_1,\ x_2,\ -3x_1+3x_2)^{\mathrm{T}}$, в котором два ($l_1=k_1$) свободных неизвестных. Из общего решения при $x_1=1$, $x_2=0$ и при $x_1=0$, $x_2=0$ получаем фундаментальную систему решений $X_1=(1,\ 0,\ -3)^{\mathrm{T}}$, $X_2=(0,\ 1,\ 3)^{\mathrm{T}}$.

При $\lambda=1$ система $(A_1-\lambda E)X=0$, т.е. система

$$\begin{cases} -2x_1+3x_2-x_3=0,\\ -3x_1+4x_2-x_3=0,\\ -3x_1+3x_2=0 \end{cases}$$

имеет общее решение $X=(x_1,\ x_1,\ x_1)^{\mathrm{T}}$, в котором одно ($l_2=$ $=k_2$) свободное неизвестное. Поэтому ФСР этой системы состоит из одного решения, например из решения $X_3=(1,\ 1,\ 1)^{\mathrm{T}}$. Из решений X_1, X_2, X_3, как из столбцов, составляется невырожденная матрица

$$T=(X_1,\ X_2,\ X_3)=\begin{pmatrix} 1 & 0 & 1\\ 0 & 1 & 1\\ -3 & 3 & 1 \end{pmatrix}.$$

Поэтому матрица A_1 приводится к диагональному виду

$$\Lambda=T^{-1}A_1T=\begin{pmatrix} 2 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 1 \end{pmatrix}$$

и имеет каноническое разложение

$$A_1=T\Lambda T^{-1}=T\begin{pmatrix} 2 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 1 \end{pmatrix}T^{-1}.$$

б) Корнями характеристического многочлена

$$|A_2-\lambda E|=\begin{vmatrix} -\lambda & 3 & 3\\ -1 & 8-\lambda & 6\\ 2 & -14 & -10-\lambda \end{vmatrix}=-\lambda(\lambda+1)^2$$

являются $\lambda_{1,2} = -1$, $\lambda_3 = 0$ соответственно кратности $k_1 = 2$, $k_2 = 1$. Все они действительные и поэтому являются собственными значениями матрицы A_2. При $\lambda_1 = -1$ матрица

$$A_2 - \lambda_1 E = \begin{pmatrix} 1 & 3 & 3 \\ -1 & 9 & 6 \\ 2 & -14 & -9 \end{pmatrix}$$

имеет ранг $r_1 = 2$, и потому $l_1 = n - r_1 = 3 - 2 = 1 \neq k_1$, т.е. геометрическая кратность $l_1 = 1$ характеристического числа $\lambda_1 = -1$ не равна его алгебраической кратности $k_1 = 2$. Поэтому матрица A_2 не приводится к диагональному виду. Если бы строили для матрицы A_2 матрицу T, удовлетворяющую соотношениям (5.21) и (5.22), то она получилась бы неквадратной.

Действительно, при $\lambda = -1$ система $(A_2 - \lambda E)X = 0$, т.е. система

$$\begin{cases} x_1 + 3x_2 + 3x_3 = 0, \\ -x_1 + 9x_2 + 6x_3 = 0, \\ 2x_1 - 14x_2 - 9x_3 = 0 \end{cases}$$

имеет ФСР, состоящую из одного решения ($l_1 \neq k_1$), например из решения $X_1 = (3,\ 3,\ -4)^{\mathrm{T}}$.

При $\lambda = 0$ система $(A_2 - \lambda E)X = 0$, т.е. система

$$\begin{cases} 3x_2 + 3x_3 = 0, \\ -x_1 + 8x_2 + 6x_3 = 0, \\ 2x_1 - 14x_2 - 10x_3 = 0, \end{cases}$$

также имеет ФСР, состоящую из одного решения ($l_2 = k_2$), например из решения $X_2 = (2,\ 1,\ -1)^{\mathrm{T}}$. Решений X_1 и X_2 недостаточно для конструирования квадратной невырожденной матрицы T третьего порядка. Поэтому матрица A_2 не приводится к диагональному виду и не имеет канонического разложения.

Приведение матриц к диагональному виду и каноническое разложение матриц широко используется в теории и вычислительной практике. Например, если известно каноническое разложение $A = T\Lambda T^{-1}$ матрицы A, то ее m-я степень при натуральном числе m легко находится по формуле

$$A^m = T\Lambda^m T^{-1} = T\begin{pmatrix} \lambda_1^m & & 0 \\ & \ddots & \\ 0 & & \lambda_n^m \end{pmatrix} T^{-1}, \qquad (5.23)$$

так как $A^m = T\Lambda T^{-1} \cdot T\Lambda T^{-1} \ldots T\Lambda T^{-1} = T\Lambda^m T^{-1}$. Формула (5.23) сохраняется при m целом отрицательном для невырожденной матрицы A. В частности,

$$A^{-1} = T\begin{pmatrix} \lambda_1^{-1} & & 0 \\ & \ddots & \\ 0 & & \lambda_n^{-1} \end{pmatrix} T^{-1}. \qquad (5.24)$$

Один из корней m-й степени из матрицы A определяется формулой

$$\sqrt[m]{A} = T\begin{pmatrix} \sqrt[m]{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt[m]{\lambda_n} \end{pmatrix} T^{-1}. \qquad (5.25)$$

Действительно, возведя правую часть равенства (5.25) по формуле (5.23) в m-ю степень, получим A. Если в формуле (5.25) все $\lambda_i > 0$, то, беря арифметические значения корней m-й степени из каждого λ_i, получим единственный корень m-й степени из матрицы A, у которого все характеристические числа положительные. О всех корнях из матрицы см. [7], гл. VIII.

Решение системы $AX = b$ линейных уравнений также значительно упрощается, если известно каноническое разложение $A = T\Lambda T^{-1}$. В этом случае от системы $T\Lambda T^{-1}X = b$ переходят к системе $\Lambda T^{-1}X = T^{-1}b$. Затем вводят обозначение $Z = T^{-1}X$ и решают систему $\Lambda Z = T^{-1}b$. Причем неизвестным $z_{r+1}, \ldots, z_n$, при которых множителями стоят $\lambda_{r+1}, \ldots, \lambda_n$, равные нулю, придают соответственно произвольные значения $C_1, C_2, \ldots, C_{n-r}$. В результате получают:

$$Z = (z_1, z_2, \ldots, z_r, C_1, C_2, \ldots, C_{n-r})^{\mathrm{T}}.$$

По найденному Z находят:

$$X = TZ = (X_1,\, X_2,\, \ldots,\, X_r,\, X_{r+1},\, \ldots,\, X_n)\begin{pmatrix} z_1 \\ z_2 \\ \ldots \\ z_r \\ C_1 \\ C_2 \\ \ldots \\ C_{n-r} \end{pmatrix} =$$

$$= (z_1X_1 + \ldots + z_rX_r) + (C_1X_{r+1} + \ldots + C_{n-r}X_n) =$$
$$= x^0 + X_{\text{одн}}, \quad (5.26)$$

где $X_1,\, X_2,\, \ldots,\, X_n$ — столбцы матрицы T.

Пример 5.10. Решить систему $AX = (12,\ 12,\ -8)^{\text{т}}$, если известно каноническое разложение:

$$A = T\Lambda T^{-1} = \begin{pmatrix} 3 & 2 & 1 \\ 0 & 1 & 1 \\ 2 & 0 & -1 \end{pmatrix}\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & -2 & -1 \\ -2 & 5 & 3 \\ 2 & -4 & -3 \end{pmatrix}.$$

Решение. От системы $AX = (12,\ 12,\ -8)^{\text{т}}$ перейдем к системе $\Lambda T^{-1}X = T^{-1}(12,\ 12,\ -8)^{\text{т}}$, т.е. к системе

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} T^{-1}X = \begin{pmatrix} -4 \\ 12 \\ 0 \end{pmatrix}.$$

Полагая здесь $T^{-1}X = Z$, получим систему

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} -4 \\ 12 \\ 0 \end{pmatrix},$$

или в подробной записи

$$\begin{cases} 2z_1 = -4, \\ 2z_2 = 12, \\ 0z_3 = 0. \end{cases}$$

Отсюда находим $Z=(-2,\ 6,\ C)^{\text{т}}$. Поэтому

$$X=TZ=\begin{pmatrix}3&2&1\\0&1&1\\2&0&-1\end{pmatrix}\begin{pmatrix}-2\\6\\C\end{pmatrix}=\begin{pmatrix}6\\6\\-4\end{pmatrix}+C\begin{pmatrix}1\\1\\-1\end{pmatrix}.$$

Пример 5.11. Найти квадратный корень из матрицы A при известном ее каноническом разложении

$$A=\begin{pmatrix}17&16&16\\16&41&32\\16&32&41\end{pmatrix}=$$
$$=\begin{pmatrix}1&-2&-2\\2&-1&0\\2&0&1\end{pmatrix}\begin{pmatrix}81&0&0\\0&9&0\\0&0&9\end{pmatrix}\frac{1}{9}\begin{pmatrix}1&2&2\\-2&5&-4\\-2&-4&5\end{pmatrix}.$$

Решение. По формуле (5.25) получаем:

$$\sqrt{A}=\begin{pmatrix}1&-2&-2\\2&-1&0\\2&0&1\end{pmatrix}\begin{pmatrix}\sqrt{81}&0&0\\0&\sqrt{9}&0\\0&0&\sqrt{9}\end{pmatrix}\times$$
$$\times\frac{1}{9}\begin{pmatrix}1&2&2\\-2&5&-4\\-2&-4&5\end{pmatrix}=\frac{1}{9}\begin{pmatrix}33&12&12\\12&51&24\\12&24&51\end{pmatrix}.$$

Упражнения

5.1. Оператор φ переводит вектор $x=(x_1,\ x_2,\ x_3)^{\text{т}}$ в вектор $\varphi(x)$. В каких случаях оператор φ является линейным? В случае линейности оператора указать его матрицу в том же базисе, в каком заданы координаты векторов x и $\varphi(x)$:

1) $\varphi(x)=(x_1+2x_2+3x_3,\ x_2-2x_3,\ x_1-2x_2-3x_3)^{\text{т}}$;
2) $\varphi(x)=(x_1+1,\ x_2+2,\ x_3+3)^{\text{т}}$;
3) $\varphi(x)=(2x_1+3x_2+x_3,\ x_1-x_2-x_3,\ x_1+2x_2+3x_3)^{\text{т}}$;
4) $\varphi(x)=(x_1+x_2,\ x_2+x_3,\ x_1+x_2+x_3)^{\text{т}}$;
5) $\varphi(x)=\left(2x_1+x_2,\ x_1^3+x_2,\ x_3^2\right)^{\text{т}}$;
6) $\varphi(x)=(x_1+2x_2+x_3,\ x_1+x_2,\ x_1+x_2+2x_3)^{\text{т}}$.

5.2. Линейный оператор φ переводит векторы a_1, a_2, a_3 соответственно в векторы b_1, b_2, b_3. Найти матрицу оператора φ в том же базисе, в каком заданы координатами все векторы:

1) $a_1 = (1,\ 2,\ -3)^{\mathrm{T}}$, $a_2 = (0,\ 1,\ 2)^{\mathrm{T}}$, $a_3 = (1,\ 0,\ 4)^{\mathrm{T}}$,
$b_1 = (1,\ 1,\ 1)^{\mathrm{T}}$, $b_2 = (1,\ 2,\ 1)^{\mathrm{T}}$, $b_3 = (0,\ 1,\ 1)^{\mathrm{T}}$;

2) $a_1 = (1,\ 2,\ 1)^{\mathrm{T}}$, $a_2 = (4,\ 3,\ -2)^{\mathrm{T}}$, $a_3 = (-5,\ -4,\ -1)^{\mathrm{T}}$,
$b_1 = (1,\ 1,\ 1)^{\mathrm{T}}$, $b_2 = (1,\ 0,\ 1)^{\mathrm{T}}$, $b_3 = (0,\ -1,\ 1)^{\mathrm{T}}$;

3) $a_1 = (1,\ 1,\ 1)^{\mathrm{T}}$, $a_2 = (2,\ -3,\ 1)^{\mathrm{T}}$, $a_3 = (4,\ 1,\ -5)^{\mathrm{T}}$,
$b_1 = (0,\ 1,\ 0)^{\mathrm{T}}$, $b_2 = (0,\ 1,\ 1)^{\mathrm{T}}$, $b_3 = (1,\ 1,\ 0)^{\mathrm{T}}$;

4) $a_1 = (4,\ -8,\ -5)^{\mathrm{T}}$, $a_2 = (-4,\ 7,\ -1)^{\mathrm{T}}$, $a_3 = (-3,\ 5,\ 1)^{\mathrm{T}}$,
$b_1 = (1,\ 1,\ 0)^{\mathrm{T}}$, $b_2 = (0,\ 2,\ 1)^{\mathrm{T}}$, $b_3 = (0,\ 1,\ 3)^{\mathrm{T}}$;

5) $a_1 = (1,\ 2,\ 4)^{\mathrm{T}}$, $a_2 = (1,\ -3,\ 1)^{\mathrm{T}}$, $a_3 = (1,\ 1,\ -5)^{\mathrm{T}}$,
$b_1 = (1,\ 1,\ 1)^{\mathrm{T}}$, $b_2 = (0,\ 1,\ 2)^{\mathrm{T}}$, $b_3 = (0,\ 1,\ 3)^{\mathrm{T}}$;

6) $a_1 = (1,\ 2,\ 4)^{\mathrm{T}}$, $a_2 = (1,\ 1,\ -5)^{\mathrm{T}}$, $a_3 = (1,\ -3,\ 1)^{\mathrm{T}}$,
$b_1 = (1,\ 2,\ 0)^{\mathrm{T}}$, $b_2 = (1,\ 3,\ 0)^{\mathrm{T}}$, $b_3 = (0,\ 0,\ 1)^{\mathrm{T}}$.

5.3. Линейный оператор φ в базисе e_1, e_2, e_3 имеет матрицу A. Найти матрицу этого оператора в базисе e'_1, e'_2, e'_3, если:

1) $A = \begin{pmatrix} 1 & -1 & 1 \\ 2 & 1 & -1 \\ 3 & -1 & 1 \end{pmatrix}$, $\begin{array}{ll} e_1 = (1,\ 0,\ 1)^{\mathrm{T}}, & e'_1 = (1,\ -1,\ 1)^{\mathrm{T}}, \\ e_2 = (2,\ 1,\ 0)^{\mathrm{T}}, & e'_2 = (0,\ 1,\ -1)^{\mathrm{T}}, \\ e_3 = (-3,\ 2,\ 4)^{\mathrm{T}}, & e'_3 = (0,\ 1,\ 1)^{\mathrm{T}}; \end{array}$

2) $A = \begin{pmatrix} 1 & -2 & 3 \\ 0 & 1 & 2 \\ -1 & 2 & 1 \end{pmatrix}$, $\begin{array}{ll} e_1 = (1,\ 4,\ -5)^{\mathrm{T}}, & e'_1 = (2,\ 1,\ 2)^{\mathrm{T}}, \\ e_2 = (2,\ 3,\ -4)^{\mathrm{T}}, & e'_2 = (1,\ -1,\ 2)^{\mathrm{T}}, \\ e_3 = (1,\ -2,\ -1)^{\mathrm{T}}, & e'_3 = (0,\ 1,\ -1)^{\mathrm{T}}; \end{array}$

3) $A = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ -1 & 2 & 1 \end{pmatrix}$, $\begin{array}{ll} e_1 = (1,\ 2,\ 4)^{\mathrm{T}}, & e'_1 = (1,\ 3,\ 2)^{\mathrm{T}}, \\ e_2 = (1,\ -3,\ 1)^{\mathrm{T}}, & e'_2 = (1,\ -2,\ -1)^{\mathrm{T}}, \\ e_3 = (1,\ 1,\ -5)^{\mathrm{T}}, & e'_3 = (0,\ 1,\ 1)^{\mathrm{T}}; \end{array}$

4) $A = \begin{pmatrix} 2 & 3 & -1 \\ 1 & -1 & 2 \\ 1 & 0 & 1 \end{pmatrix}$, $\begin{array}{ll} e_1 = (4,\ -4,\ -3)^{\mathrm{T}}, & e'_1 = (3,\ 1,\ 2)^{\mathrm{T}}, \\ e_2 = (-8,\ 7,\ 5)^{\mathrm{T}}, & e'_2 = (1,\ 2,\ 1)^{\mathrm{T}}, \\ e_3 = (-5,\ -1,\ 1)^{\mathrm{T}}, & e'_3 = (0,\ 0,\ 1)^{\mathrm{T}}; \end{array}$

5) $A = \begin{pmatrix} 1 & 2 & -3 \\ 3 & 1 & -2 \\ 1 & 3 & -2 \end{pmatrix}$, $\begin{matrix} e_1 = (1,\ 2,\ 4)^{\mathrm{T}}, & e'_1 = (1,\ 1,\ 1)^{\mathrm{T}}, \\ e_2 = (1,\ -3,\ 1)^{\mathrm{T}}, & e'_2 = (0,\ 1,\ 2)^{\mathrm{T}}; \\ e_3 = (1,\ 1,\ -5)^{\mathrm{T}}, & e'_3 = (0,\ 1,\ 3)^{\mathrm{T}}, \end{matrix}$

6) $A = \begin{pmatrix} 3 & 1 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 2 \end{pmatrix}$, $\begin{matrix} e_1 = (1,\ 0,\ 1)^{\mathrm{T}}, & e'_1 = (2,\ 1,\ 1)^{\mathrm{T}}, \\ e_2 = (2,\ 1,\ 0)^{\mathrm{T}}, & e'_2 = (3,\ 2,\ 1)^{\mathrm{T}}, \\ e_3 = (-3,\ 2,\ 4)^{\mathrm{T}}, & e'_3 = (0,\ 0,\ 1)^{\mathrm{T}}. \end{matrix}$

5.4. Задан линейный оператор φ, который переводит вектор $x = (x_1,\ x_2,\ x_3)^{\mathrm{T}}$ в вектор φx. Найти образ вектора a и прообраз вектора y, если:

1) $\varphi x = (x_1 + 2x_2 + x_3,\ x_2 + 2x_2 + 2x_3,\ x_1 + x_3)^{\mathrm{T}}$,
$a = (1,\ 1,\ 1)^{\mathrm{T}}$, $y = (1,\ 2,\ 3)^{\mathrm{T}}$;

2) $\varphi x = (2x_1,\ 3x_1 + x_2,\ x_1 + 2x_2 + x_3)^{\mathrm{T}}$,
$a = (2,\ 1,\ 1)^{\mathrm{T}}$, $y = (1,\ 1,\ 0)^{\mathrm{T}}$;

3) $\varphi x = (x_1 + x_2,\ 4x_1 + 2x_2,\ x_1 + 3x_3)^{\mathrm{T}}$,
$a = (1,\ 2,\ 1)^{\mathrm{T}}$, $y = (1,\ 3,\ 1)^{\mathrm{T}}$.

5.5. Линейный оператор φ в паре базисов e и q имеет матрицу A. Найти прообраз вектора y, если:

1) $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 1 & 2 \\ 1 & 1 & 1 & 1 \end{pmatrix}$, $[y]_q = \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}$;

2) $A = \begin{pmatrix} 1 & 2 & 1 & 1 \\ 3 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \end{pmatrix}$, $[y]_q = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}$;

3) $A = \begin{pmatrix} 2 & 1 & 2 & 1 \\ 1 & 2 & 1 & 2 \\ 2 & 1 & 1 & 2 \end{pmatrix}$, $[y]_q = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$;

4) $A = \begin{pmatrix} 2 & 1 & 1 & 1 \\ 2 & 1 & 2 & 1 \\ 2 & 2 & 1 & 1 \end{pmatrix}$, $[y]_q = \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix}$;

5) $A = \begin{pmatrix} 1 & 2 & 1 & 1 \\ 2 & 3 & 3 & 1 \\ 3 & 4 & 2 & 2 \end{pmatrix}$, $[y]_q = \begin{pmatrix} 3 \\ 1 \\ -2 \end{pmatrix}$;

6) $A = \begin{pmatrix} 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 2 & 3 & 1 & 2 \end{pmatrix}$, $[y]_q = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$.

5.6. Для каждой из матриц непосредственным вычислением определителя найти ее характеристический многочлен и вычислить его корни:

1) $A = \begin{pmatrix} 0 & 3 & 3 \\ -1 & 8 & 6 \\ 2 & -14 & -10 \end{pmatrix}$; 2) $A = \begin{pmatrix} 1 & -3 & 4 \\ 4 & -7 & 8 \\ 6 & -7 & 7 \end{pmatrix}$;

3) $A = \begin{pmatrix} 4 & -1 & -2 \\ 2 & 1 & -2 \\ 1 & -1 & 1 \end{pmatrix}$; 4) $A = \begin{pmatrix} 7 & -12 & 6 \\ 10 & -19 & 10 \\ 12 & -24 & 13 \end{pmatrix}$;

5) $A = \begin{pmatrix} -1 & 3 & -1 \\ -3 & 5 & -1 \\ -3 & 3 & 1 \end{pmatrix}$; 6) $A = \begin{pmatrix} 6 & -5 & -3 \\ 3 & -2 & -2 \\ 2 & -2 & 0 \end{pmatrix}$;

7) $A = \begin{pmatrix} 4 & -2 & 2 \\ -5 & 7 & -5 \\ -6 & 6 & -4 \end{pmatrix}$; 8) $A = \begin{pmatrix} 3 & 1 & 0 \\ -4 & -1 & 0 \\ 4 & -8 & -2 \end{pmatrix}$;

9) $A = \begin{pmatrix} 4 & 6 & 0 \\ -3 & -5 & 0 \\ -3 & -6 & 1 \end{pmatrix}$.

5.7. Пользуясь формулой (5.13), найти минимальные многочлены для матриц из упражнения 5.6.

5.8. Для матриц из упражнения 5.6 построить, если возможно, канонические разложения.

5.9. Построить канонические разложения следующих матриц и, пользуясь ими, вычислить сотую степень и корень квадратный из каждой матрицы:

1) $A=\begin{pmatrix} 7 & -12 & 6 \\ 10 & -19 & 10 \\ 12 & -24 & 13 \end{pmatrix}$; 2) $A=\begin{pmatrix} -1 & 3 & -1 \\ -3 & 5 & -1 \\ -3 & 3 & 1 \end{pmatrix}$;

3) $A=\begin{pmatrix} 7 & -12 & -2 \\ 3 & -4 & 0 \\ -2 & 0 & -2 \end{pmatrix}$; 4) $A=\begin{pmatrix} -2 & 8 & 6 \\ -4 & 10 & 6 \\ 4 & -8 & -4 \end{pmatrix}$;

5) $A=\begin{pmatrix} 2 & -2 & -1 \\ -2 & 5 & 2 \\ -1 & 2 & 2 \end{pmatrix}$; 6) $A=\begin{pmatrix} 3 & 2 & 0 \\ 2 & 4 & -2 \\ 0 & -2 & 5 \end{pmatrix}$;

7) $A=\begin{pmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{pmatrix}$; 8) $A=\begin{pmatrix} 5 & -2 & -2 \\ -2 & 6 & 0 \\ -2 & 0 & 4 \end{pmatrix}$;

9) $A=\begin{pmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{pmatrix}$.

5.10. Решить систему $AX=b$, если известны каноническое разложение $A=T\Lambda T^{-1}$ матрицы A и столбец свободных членов b:

1) $A=\begin{pmatrix} 2 & 5 & 7 \\ 6 & 3 & 4 \\ 5 & -2 & -3 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & -1 & 1 \\ -38 & 41 & -34 \\ 27 & -29 & 24 \end{pmatrix}$,

$b=(1,\ 0,\ -1)^{\mathrm{T}}$;

2) $A=\begin{pmatrix} 3 & -4 & 5 \\ 2 & -3 & 1 \\ 3 & -5 & 1 \end{pmatrix}\begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} -8 & 29 & -11 \\ -5 & 18 & -7 \\ 1 & -3 & 1 \end{pmatrix}$,

$b=(3,\ 1,\ 0)^{\mathrm{T}}$;

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

$b=(3,\ 0,\ 0)^{\mathrm{T}}$;

4) $A = \begin{pmatrix} 2 & 2 & 3 \\ 1 & -1 & 0 \\ -1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & -4 & -3 \\ 1 & -5 & -3 \\ -1 & 6 & 4 \end{pmatrix}$,

$b = (4,\ 0,\ 1)^{\mathrm{T}}$;

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

$b = (-6,\ 3,\ 0)^{\mathrm{T}}$;

6) $A = \begin{pmatrix} 1 & -4 & -3 \\ 1 & -5 & -3 \\ -1 & 6 & 4 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 & 2 & 3 \\ 1 & -1 & 0 \\ -1 & 2 & 1 \end{pmatrix}$,

$b = (22,\ 22,\ -22)^{\mathrm{T}}$.

Глава 6 КАНОНИЧЕСКАЯ ЖОРДАНОВА ФОРМА МАТРИЦЫ

6.1. Жорданов базис

Было показано, что матрица оператора простой структуры приводится к диагональному виду. Оказывается, что в общем случае комплексная квадратная матрица приводится к квазидиагональной, так называемой ***жордановой форме***. Ниже мы выясним, как практически получается такая форма матриц.

Пусть дана квадратная матрица A порядка n и пусть

$$f(\lambda)=|A-\lambda E|=(-1)^n(\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\ldots(\lambda-\lambda_s)^{m_s} \qquad (6.1)$$

— характеристический многочлен матрицы A. Матрица A в фиксированном базисе n-мерного комплексного пространства X определяет линейный оператор $\boldsymbol{A}$. Напомним, что множество образов всех векторов из X, которые получаются при действии оператора $\boldsymbol{A}$, составляет *область значений* этого оператора. Ее будем обозначать через $\boldsymbol{A}(X)$ или $\operatorname{Im}\boldsymbol{A}$. Область значений оператора $\boldsymbol{A}$ совпадает с пространством $L(A)$ столбцов матрицы A. *Ядро оператора* $\boldsymbol{A}$ — это множество всех векторов из X, которые оператор $\boldsymbol{A}$ переводит в нулевой вектор. Ядро оператора $\boldsymbol{A}$ также является подпространством в X. Его мы будем обозначать через $\operatorname{Ker}\boldsymbol{A}$. Размерность области значений оператора называют его *рангом*, а размерность ядра оператора — его *дефектом*. Характеристический многочлен (6.1) матрицы A называют также характеристическим многочленом оператора $\boldsymbol{A}$.

Корневым подпространством оператора $\boldsymbol{A}$ по характеристическому числу λ_i характеристического многочлена

(6.1) называют подпространство

$$K_i = \operatorname{Ker}(\boldsymbol{A} - \lambda_i \boldsymbol{E})^{m_i}, \tag{6.2}$$

где $\boldsymbol{E}$ — тождественный оператор, m_i — кратность корня λ_i в характеристическом многочлене (6.1).

Известно, что размерность корневого подпространства K_i равна m_i. Каждое корневое подпространство $K_1, \ldots, K_s$ инвариантно относительно оператора $\boldsymbol{A}$, и пространство X разлагается в прямую сумму корневых подпространств

$$X = K_1 \oplus K_2 \oplus \ldots \oplus K_s. \tag{6.3}$$

Пусть K_i — одно из корневых подпространств и $\boldsymbol{B}_i$ — оператор, индуцируемый на нем оператором $\boldsymbol{A} - \lambda_i \boldsymbol{E}$. Найдутся векторы $e^{(2)}, \ldots, e^{(h)}$ корневого подпространства K_i, удовлетворяющие условиям

$$\begin{aligned} (\boldsymbol{A} - \lambda_i \boldsymbol{E})(e^{(2)}) &= \boldsymbol{B}_i(e^{(2)}) = e^{(1)}, \\ (\boldsymbol{A} - \lambda_i \boldsymbol{E})(e^{(3)}) &= \boldsymbol{B}_i(e^{(3)}) = e^{(2)}, \\ &\ldots\ldots\ldots\ldots \\ (\boldsymbol{A} - \lambda_i \boldsymbol{E})(e^{(h)}) &= \boldsymbol{B}_i(e^{(h)}) = e^{(h-1)}, \end{aligned} \tag{6.4}$$

где $e^{(1)}$ — собственный вектор оператора $\boldsymbol{A}$ по λ_i. Векторы $e^{(2)}, \ldots, e^{(h)}$ называют соответственно ***первым***, ***вторым*** и т.д. ***присоединенными*** к $e^{(1)}$ ***векторами***. Говорят также, что векторы $e^{(1)}, e^{(2)}, \ldots, e^{(h)}$ образуют в K_i ***жорданову цепочку*** длины h с началом в $e^{(1)}$.

Если для вектора x из K_i построить систему векторов x, $\boldsymbol{B}_i x$, $\boldsymbol{B}_i^2 x, \ldots$, то в силу определения корневого подпространства K_i в этой системе векторов когда-то первый раз встретится вектор $\boldsymbol{B}_i^h(x) = 0$. В этом случае говорят, что ***вектор x имеет высоту h в корневом подпространстве K_i***.

Подпространство, порожденное векторами

$$x, \quad \boldsymbol{B}_i x, \quad \boldsymbol{B}_i^2 x, \quad \ldots, \quad \boldsymbol{B}_i^{h-1}$$

называют **циклическим подпространством, порожденным вектором** x. Его размерность равна h. Каждое циклическое подпространство в K_i инвариантно относительно оператора $\boldsymbol{B}_i$.

Очевидно, что жорданова цепочка векторов $e^{(1)}$, $e^{(2)}$, ..., $e^{(h)}$ порождает циклическое подпространство размерности h. Корневое подпространство распадается в прямую сумму циклических подпространств, порождаемых жордановыми цепочками векторов.

Векторы всех жордановых цепочек, образующих базис такой прямой суммы, составляют базис в K_i. Этот базис называют ***корневым***, или ***жордановым***, ***базисом*** в K_i. Чтобы построить жорданов базис в K_i, необходимо учесть следующее [3, с. 66]. Собственный вектор $e^{(1)} \in K_i$ имеет точно $h-1$ присоединенных векторов $e^{(2)}$, $e^{(3)}$, ..., $e^{(h)}$, или, что то же самое, является началом жордановой цепочки наибольшей из возможных для него длин h, тогда и только тогда, когда

$$\begin{aligned} &e^{(1)} \in R_{h-1} = \boldsymbol{B}_i^{h-1}(K_i) \cap \operatorname{Ker} \boldsymbol{B}_i, \\ &e^{(1)} \notin \boldsymbol{B}_i^{h}(K_i). \end{aligned} \tag{6.5}$$

Для подпространств R_0, R_1, ..., R_k, где k — наибольшая длина жордановых цепочек в K_i, выполняются соотношения [3, с. 67]

$$0 = R_k \subseteq R_{k-1} \subseteq \ldots \subseteq R_2 \subseteq R_1 \subseteq R_0 = \operatorname{Ker} \boldsymbol{B}_i = E_i(\lambda_i), \tag{6.6}$$

где $E(\lambda_i)$ — собственное подпространство оператора $\boldsymbol{A}$, отвечающее собственному значению λ_i.

Наибольшая длина k_i жордановых цепочек в K_i совпадает с кратностью корня λ_i в минимальном многочлене матрицы A или, что то же самое, с наименьшим числом $m = 1, 2, \ldots$, при котором выполняется соотношение

$$r(A - \lambda_i E)^m = n - m_i, \tag{6.7}$$

где $r(A - \lambda_i E)^m$ — ранг матрицы $(A - \lambda_i E)^m$; n — порядок матрицы A; m_i — кратность корня λ_i в характеристическом многочлене матрицы A.

Для удобства в дальнейшем там, где это не будет вносить путаницы, мы станем опускать индекс i числа k_i.

Если $k = 1$, то корневое подпространство K_i совпадает с собственным подпространством $E(\lambda_i)$ и в K_i существует базис, состоящий из собственных векторов оператора $\boldsymbol{A}$. Это будет жорданов базис в K_i в рассматриваемом случае. Матрица оператора, индуцируемого в K_i, в этом базисе будет диагональной матрицей порядка m_i с λ_i по главной диагонали.

Если же $k > 1$, то для построения жорданова базиса в K_i недостаточно собственных векторов оператора $\boldsymbol{A}$. В этом случае сначала в пространстве $E(\lambda_i) = \operatorname{Ker} \boldsymbol{B}_i$ собственных векторов по λ_i строят базис, связанный с системой подпространств (6.6). Для этого выбирают какой-либо базис $e_1^{(1)}, \ldots, e_{p_1}^{(1)}$ в подпространстве R_{k-1}. Это будут собственные векторы, с которых начинаются жордановы цепочки наибольшей длины k в K_i. Затем эти векторы дополняют какими-либо векторами $e_{p_1+1}^{(1)}, \ldots, e_{p_2}^{(1)}$ до базиса в R_{k-2}. Векторы $e_{p_1+1}^{(1)}, \ldots, e_{p_2}^{(1)}$ будут собственными векторами, с которых начинаются жордановы цепочки длины $k-1$. Такой процесс продолжают до тех пор, пока не получится базис в $E(\lambda_i) = \operatorname{Ker} \boldsymbol{B}_i$.

Далее для каждого вектора $e_1^{(1)}, \ldots, e_{p_1}^{(1)}$ строят по формулам (6.4) при $h = k$ жордановы цепочки длины k, для каждого вектора $e_{p_1+1}^{(1)}, \ldots, e_{p_2}^{(1)}$ по формулам (6.4) при $h = k-1$ строят жордановы цепочки длины $k-1$ и т.д. Векторы всех таких жордановых цепочек образуют жорданов базис в K_i. Выпишем его в виде табл. 6.1.

Мы предполагаем, что в жордановом базисе пространства K_i векторы пронумерованы так, что сперва идут векторы первой строки табл. 6.1 слева направо, затем второй строки и т.д. В первом столбце табл. 6.1 располагаются собственные векторы оператора. В первых p_1 строках записаны p_1 жордановых цепочек базиса наибольшей в K_i длины k, затем идут $p_2 - p_1$ жордановых цепочек длины $k-1$ и т.д. При этом выполняют-

Таблица 6.1

$e_1^{(1)}$	$e_1^{(2)}$	$\dots$	$e_1^{(k-1)}$	$e_1^{(k)}$
. . .	. . .	. . .	. . .	. . .
$e_{p_1}^{(1)}$	$e_{p_1}^{(2)}$	$\dots$	$e_{p_1}^{(k-1)}$	$e_{p_1}^{(k)}$
$e_{p_1+1}^{(1)}$	$e_{p_1+1}^{(2)}$	$\dots$	$e_{p_1+1}^{(k-1)}$	
. . .	. . .	. . .	. . .	. . .
$e_{p_2}^{(1)}$	$e_{p_2}^{(2)}$	$\dots$	$e_{p_2}^{(k-1)}$	
. . .	. . .	. . .	. . .	. . .
$e_{p_k}^{(1)}$				

ся соотношения

$$\begin{aligned} q_h &= 2m_h - m_{h-1} - m_{h+1}, \quad h = 1, 2, \dots, k, \\ q_h &= r_{h-1} - 2r_h + r_{h+1}, \quad h = 1, 2, \dots, k, \end{aligned} \tag{6.8}$$

где q_h — число жордановых цепочек длины h; $m_0 = 0$, m_ρ — дефект оператора с матрицей $(A - \lambda_i E)^\rho$; $r_0 = n$, r_ρ — ранг матрицы $(A - \lambda_i E)^\rho$; k — наибольшая длина жордановых цепочек в K_i.

В общем случае может оказаться, что в табл. 6.1 отсутствуют некоторые жордановы цепочки меньших длин.

Заметим, что в каждой l-й строке табл. 6.1 только один вектор $e_l^{(1)}$ является собственным вектором оператора $\boldsymbol{A}$ по λ_i, а остальные векторы присоединенные к $e_l^{(1)}$. Векторы, стоящие в правом столбце табл. 6.1, имеют высоту k, векторы второго справа столбца — высоту $k-1$ и т.д. Векторы первого слева столбца имеют высоту 1. Если для любого $l \leqslant k$ обозначить через H_l множество векторов из K_i, высоты которых не превосходят l, то окажется, что H_l является подпространством в K_i и порождается векторами, которые в табл. 6.1 расположены в первых слева l столбцах. Для подпространств $H_0, H_1, H_2, \dots, H_k$ выполняются соотношения

$$0 = H_0 \subset H_1 \subset H_2 \subset \dots \subset H_{k-1} \subset H_k = K_i. \tag{6.9}$$

Этими соотношениями можно также пользоваться при построении жорданова базиса в K_i [4, § 74]. Для этого выбирают линейно независимые векторы $e_1^{(k)}, \ldots, e_{p_1}^{(k)}$ наибольшей в K_i высоты k, принадлежащие H_k, но не лежащие в H_{k-1} и порождающие вместе с H_{k-1} подпространство K_i. Для каждого из этих векторов $e_j^{(k)}$, $j = 1, 2, \ldots, p_1$, строят жорданову цепочку $e_j^{(k)}$, $e_j^{(k-1)} = (A - \lambda_i E)e_j^{(k)}$, $\ldots$, $e_j^{(1)} = (A - \lambda_i E)^{k-1}e_j^{(k)}$. Затем к уже построенным векторам $e_1^{(k-1)}, \ldots, e_{p_1}^{(k-1)}$ высоты $k-1$ добавляют линейно независимые векторы $e_{p_1+1}^{(k-1)}, \ldots, e_{p_2}^{(k-1)}$, принадлежащие H_{k-1}, но не лежащие в H_{k-2}, и такие, которые вместе с векторами $e_1^{(k-1)}, \ldots, e_{p_1}^{(k-1)}$ и вместе с H_{k-2} порождают H_{k-1}. Для каждого из векторов $e_j^{(k-1)}$, $j = p_1 + 1, \ldots, p_2$, строят жорданову цепочку $e_j^{(k-1)}$, $e_j^{(k-2)} = (A - \lambda_i E)e_j^{(k-1)}$, $\ldots$, $e_j^{(1)} = (A - \lambda_i E)^{k-2}e_j^{(k-1)}$. Далее так же поступают с подпространствами H_{k-2} и H_{k-3} и т.д. В результате приходят к системе векторов, образующих жорданов базис в K_i.

Итак, исходя из системы подпространств (6.6) мы строили векторы, представленные в табл. 6.1, начиная с векторов, располагающихся в левом верхнем углу. Если же эти векторы строить исходя из системы подпространств (6.9), то нужно начинать с векторов, располагающихся в правом верхнем углу.

На наш взгляд, практически легче строить базис, используя систему подпространств (6.6).

Каждая строка в табл. 6.1, представляющая собой жорданову цепочку векторов, определяет циклическое подпространство, инвариантное относительно оператора $\boldsymbol{A} - \lambda_i \boldsymbol{E}$, а следовательно, и относительно оператора $\boldsymbol{A}$. Первые p_1 циклических подпространств имеют размерность k, следующие $p_2 - p_1$ циклических подпространств имеют размерность $k-1$ и т.д.

Укажем матрицу оператора, индуцируемого, например, в первом циклическом подпространстве. Предположим, что в этом подпространстве в качестве базисных взяты векторы $e_1^{(1)}, e_1^{(2)}, \ldots, e_1^{(k)}$. В соответствии с построением для этих векторов выполняются соотношения

$$
\begin{aligned}
&(\boldsymbol{A}-\lambda_i\boldsymbol{E})e_1^{(1)}=0,\\
&(\boldsymbol{A}-\lambda_i\boldsymbol{E})e_1^{(2)}=e_1^{(1)},\\
&\dots\dots\dots\dots\\
&(\boldsymbol{A}-\lambda_i\boldsymbol{E})e_1^{(k)}=e_1^{(k-1)}.
\end{aligned}
$$

Перепишем эти соотношения в виде

$$
\begin{aligned}
&\boldsymbol{A}e_1^{(1)}=\lambda_i e_1^{(1)},\\
&\boldsymbol{A}e_1^{(2)}=\lambda_i e_1^{(2)}+e_1^{(1)},\\
&\dots\dots\dots\dots\\
&\boldsymbol{A}e_1^{(k)}=\lambda_i e_1^{(k)}+e_1^{(k-1)}.
\end{aligned}
$$

Отсюда следует (см. разд. 5.2), что матрицей оператора, индуцируемого в первом циклическом подпространстве размерности k при выбранном базисе, будет матрица

$$
\begin{pmatrix}
\lambda_i & 1 & & 0\\
 & \ddots & \ddots & \\
 & & \ddots & 1\\
0 & & & \lambda_i
\end{pmatrix} \tag{6.10}
$$

порядка k. Такую матрицу называют ***жордановой клеткой*** по λ_i ***порядка*** k.

Матрица оператора, индуцируемого в корневом пространстве K_i, будет состоять из жордановых клеток вида (6.10), расположенных по главной диагонали в таком же порядке, в каком в табл. 6.1 расположены соответствующие жордановы цепочки векторов жорданова базиса в K_i. При этом следует помнить, что длины h жордановых цепочек в K_i совпадают с порядками h соответствующих этим цепочкам жордановых клеток; числа q_h, определяющие в K_i число жордановых цепочек длины h, определяют также числа жордановых клеток порядка h по λ_i. Поэтому для подсчета чисел q_i клеток порядка h применимы формулы (6.8).

Жордановым базисом во всем пространстве X будет объединение жордановых базисов всех его корневых подпространств K_1, K_2, ..., K_S.

В жордановом базисе матрица оператора $\boldsymbol{A}$ имеет ***каноническую жорданову форму***, т.е. является ***квазидиагональной матрицей*** J, состоящей из жордановых клеток по главной диагонали. Первыми располагаются жордановы клетки по λ_1, затем жордановы клетки по λ_2 и т.д. При этом жордановы клетки располагаются в матрице J по главной диагонали в том же порядке, в каком расположены в жордановом базисе соответствующие им жордановы цепочки.

Таким образом, матрица J имеет вид:

$$J = \begin{pmatrix} \boxed{\begin{matrix}\lambda_1 & 1 & & 0\\ & \ddots & \ddots & \\ & & \ddots & 1\\ 0 & & & \lambda_1\end{matrix}} & & & & & & \\ & \ddots & & & & & \\ & & \boxed{\lambda_1} & & & & \\ & & & \ddots & & & \\ & & & & \boxed{\begin{matrix}\lambda_s & 1 & & 0\\ & \ddots & \ddots & \\ & & \ddots & 1\\ 0 & & & \lambda_s\end{matrix}} & & \\ & & & & & \ddots & \\ & & & & & & \boxed{\lambda_s} \end{pmatrix}$$

В жордановой матрице J по каждому λ_i жордановы клетки располагаются по убыванию их порядков. Некоторые из жордановых клеток могут повторяться, а некоторые из жордановых клеток низших порядков могут отсутствовать. Частным случаем жордановой матрицы является диагональная матрица.

Заметим, что если в построенном жордановом базисе изменить нумерацию жордановых цепочек, то в жордановой матрице на главной диагонали изменится положение соответствующих жордановых клеток.

В силу (5.8) матрица A и ее жорданова форма J связаны соотношением

$$J = T^{-1}AT, \tag{6.11}$$

где T — матрица перехода от исходного базиса пространства X к его жорданову базису.

Матрица T *состоит из столбцов координат векторов жорданова базиса в исходном базисе.* Эти столбцы располагаются в матрице T в том же порядке, в каком соответствующие векторы располагаются в жордановом базисе пространства X. Матрицу T называют ***трансформирующей***, или ***приводящей***, матрицу A к ее жордановой форме J.

6.2. Построение жорданова базиса и жордановой матрицы

Для практического построения жорданова базиса матрицы A, т.е. для вычисления жордановой формы данной матрицы и трансформирующей матрицы, необходимо выполнить для каждого характеристического корня λ_i матрицы A, имеющего кратность m_i, следующие действия.

1. Составить матрицу $A - \lambda_i E$ и возводить ее последовательно в степени $m = 1, 2, \ldots$ до тех пор, пока не получится равенство

$$r(A - \lambda_i E)^m = n - m_i, \tag{6.12}$$

где $r(A - \lambda_i E)^m$ — ранг матрицы $(A - \lambda_i E)^m$; n — порядок матрицы A; m_i — кратность характеристического корня λ_i матрицы A. Наименьшее натуральное число m, при котором выполняется равенство (6.12), даст максимальную длину k_i жордановых цепочек в корневом подпространстве K_i.

2. Построить подпространство $E(\lambda_i)$ собственных векторов оператора $\boldsymbol{A}$ по λ_i. Для этого следует найти какую-либо фундаментальную систему решений b_1, b_2, …, b_l системы $(A - \lambda_i E)X = 0$ и положить

$$E(\lambda_i) = \langle b_1,\, b_2,\, \ldots,\, b_l \rangle. \tag{6.13}$$

3. Найти пересечение

$$R_{k_i-1} = L_{k_i-1} \cap E(\lambda_i), \tag{6.14}$$

где $L_{k_i-1} = \langle a_{l_1}, \ldots, a_{l_r} \rangle$ — пространство столбцов матрицы $(A - \lambda_i E)^{k_i-1}$ (через $a_{l_1}, \ldots, a_{l_r}$ обозначены базисные столбцы матрицы $(A - \lambda_i E)^{k_r-1}$). Для этого полагают

$$e^{(1)} = \alpha_1 a_{l_1} + \ldots + \alpha_r a_{l_r} = \beta_1 b_1 + \ldots + \beta_l b_l, \tag{6.15}$$

затем переходят к покоординатным равенствам и для полученной при этом однородной системы относительно α_1, …, α_r, β_1, …, β_l находят какую-либо фундаментальную систему решений. Подставляя эти решения поочередно в (6.15), получают систему линейно независимых векторов

$$e_1^{(1)}, \quad e_2^{(1)}, \quad \ldots, \quad e_{p_1}^{(1)}, \tag{6.16}$$

которые составляют базис в R_{k_i-1}. Эти векторы являются собственными векторами, с которых начинаются жордановы цепочки максимальной длины k_i в корневом подпространстве K_i.

4. Для каждого вектора $e_j^{(1)}$, $j = 1, 2, \ldots, p_1$, из системы векторов (6.16) найти присоединенные векторы $e_j^{(2)}, \ldots, e_j^{(k_i)}$, составляющие j-ю жорданову цепочку длины k_i, для чего решить систему

$$(A - \lambda_i E)e_j^{(l)} = e_j^{(l-1)}, \qquad l = 2, 3, \ldots, k_i. \tag{6.17}$$

5. Если общее число векторов

$$\begin{array}{llll} e_1^{(1)}, & e_1^{(2)}, & \ldots, & e_1^{(k_i)}, \\ e_2^{(1)}, & e_2^{(2)}, & \ldots, & e_2^{(k_i)}, \\ \ldots & \ldots & \ldots & \ldots \\ e_{p_1}^{(1)}, & e_{p_1}^{(2)}, & \ldots, & e_{p_1}^{(k_i)}, \end{array} \tag{6.18}$$

построенных в предыдущих пунктах, меньше m_i, то следует перейти к построению подпространства

$$R_{k_i-2} = L_{k_i-2} \cap E(\lambda_i), \tag{6.19}$$

где L_{k_i-2} — подпространство столбцов матрицы $(A-\lambda_i E)^{k_i-2}$. В подпространстве R_{k_i-2} построить базис, содержащий использованные в предыдущих пунктах векторы $e_1^{(1)}$, $e_2^{(1)}$, …, $e_{p_1}^{(1)}$ системы (6.16). Пусть таким базисом в R_{k_i-1} является система собственных векторов

$$e_1^{(1)}, \quad e_2^{(1)}, \quad \ldots, \quad e_{p_1}^{(1)}, \quad e_{p_1+1}^{(1)}, \quad \ldots, \quad e_{p_2}^{(1)}. \tag{6.20}$$

Тогда векторы

$$e_{p_1+1}^{(1)}, \quad \ldots, \quad e_{p_2}^{(1)} \tag{6.21}$$

будут собственными векторами, с которых начинаются жордановы цепочки длины $k_i - 1$ в корневом подпространстве K_i.

6. Для каждого вектора $e_j^{(1)}$, $j = p_1+1, \ldots, p_2$, найти присоединенные векторы $e_j^{(2)}, \ldots, e_j^{(k_i-1)}$, образующие j-ю жорданову цепочку длины $k_i - 1$, для чего решить систему

$$(A-\lambda_i E)e_j^{(l)} = e_j^{(l-1)}, \qquad l = 2, 3, \ldots, k_i - 1. \tag{6.22}$$

Так следует переходить от подпространства R_{k_i-l} к подпространству $R_{k_i-(l+1)}$, $l = 1, 2, \ldots,$ и поступать, как описано в предыдущих пунктах, до тех пор, пока общее число векторов во всех построенных жордановых цепочках не окажется равным m_i. Затем выписать все эти цепочки векторов одну за другой и получить жорданов базис оператора $\boldsymbol{A}$ в корневом подпространстве K_i. Проделав эти операции для каждого характеристического корня λ_i, $i = 1, 2, \ldots, s$, и объединив полученные при этом жордановы базисы всех корневых подпространств $K_1, \ldots, K_s$, получить жорданов базис оператора $\boldsymbol{A}$ во всем пространстве X.

7. Выписать жорданову матрицу J в соответствии с построенным жордановым базисом оператора $\boldsymbol{A}$ в пространстве X.

8. Выписать трансформирующую матрицу T из столбцов координат векторов построенного жорданова базиса оператора $\boldsymbol{A}$ в пространстве X.

Замечание. Жорданов базис оператора $\boldsymbol{A}$ в пространстве X и трансформирующая матрица T находятся неоднозначно.

Пример 6.1. Для матрицы

$$A = \begin{pmatrix} 1 & -3 & 0 & 3 \\ -2 & -6 & 0 & 13 \\ 0 & -3 & 1 & 3 \\ -1 & -4 & 0 & 8 \end{pmatrix}$$

построить жорданову форму J и трансформирующую матрицу T.

Решение. Будем считать, что в четырехмерном пространстве X в некотором фиксированном базисе матрица A определяет линейный оператор $\boldsymbol{A}$, и построим в пространстве X жорданов базис этого оператора. Для этого сначала найдем характеристический многочлен

$$|A - \lambda E| = \begin{vmatrix} 1-\lambda & -3 & 0 & 3 \\ -2 & -6-\lambda & 0 & 13 \\ 0 & -3 & 1-\lambda & 3 \\ -1 & -4 & 0 & 8-\lambda \end{vmatrix} = (\lambda - 1)^4.$$

Он имеет лишь один корень $\lambda_1 = 1$ кратности $m_1 = 4$. Следовательно, пространство X является корневым по $\lambda_1 = 1$.

Составим матрицу $(A - \lambda_1 E) = A - E$ и будем возводить ее в степени $m = 1, 2, \ldots$ до тех пор, пока не получится равенство

$$(A - E)^m = n - m_1 = 4 - 4 = 0.$$

При $m = 1$

$$(A-E)^m = A - E = \begin{pmatrix} 0 & -3 & 0 & 3 \\ -2 & -7 & 0 & 13 \\ 0 & -3 & 0 & 3 \\ -1 & -4 & 0 & 7 \end{pmatrix},$$

$$r(A - E) = 2 \neq n - m_1 = 0.$$

При $m = 2$

$$(A-E)^m = (A-E)^2 = \begin{pmatrix} 3 & 9 & 0 & -18 \\ 1 & 3 & 0 & -6 \\ 3 & 9 & 0 & -18 \\ 1 & 3 & 0 & -6 \end{pmatrix},$$

$$r(A - E)^2 = 1 \neq n - m_1 = 0.$$

При $m = 3$

$$(A-E)^m = (A-E)^3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$r(A - E)^3 = 0 = n - m_1.$$

Следовательно, в пространстве X жордановы цепочки имеют наибольшую длину $k = 3$.

Найдем подпространство $E(1)$ собственных векторов матрицы A по $\lambda_1 = 1$. Для этого рассмотрим систему $(A - E)X = 0$, т.е. систему

$$\begin{cases} -3x_2 + 3x_4 = 0, \\ 2x_1 - 7x_2 + 13x_3 = 0, \\ -3x_2 + 3x_4 = 0, \\ -x_1 - 4x_2 + 7x_4 = 0. \end{cases}$$

Одну из фундаментальных систем решений этой системы составляют векторы $b_1 = (0,\ 0,\ 1,\ 0)^{\text{т}}$, $b_2 = (3,\ 1,\ 0,\ 1)^{\text{т}}$. Поэтому $E(1) = \langle b_1,\ b_2 \rangle$.

Далее, найдем пересечение $R_{k-1} = R_2 = L_{k-1} \cap E(1) = L_2 \cap \cap E(1)$, где L_2 — пространство столбцов матрицы $(A - E)^2$. Пространство L_2 порождается вектором $(3,\ 1,\ 3,\ 1)^{\text{т}}$, поскольку ранг матрицы $(A - E)^2$ равен единице.

Замечаем, что $(3,\ 1,\ 3,\ 1)^{\text{т}} = 3b_1 + b_2$. Поэтому $R_2 = L_2 \cap \cap E(1) = L_2 = \langle (3,\ 1,\ 3,\ 1)^{\text{т}} \rangle$. Положим $e_1^{(1)} = (3,\ 1,\ 3,\ 1)^{\text{т}}$. Это собственный вектор, с которого начинается жорданова цепочка длины $k = 3$. Присоединенные векторы

$$e_1^{(2)} = \left(x_1^{(2)},\ x_2^{(2)},\ x_3^{(2)},\ x_4^{(2)}\right)^{\text{т}}, \quad e_1^{(3)} = \left(x_1^{(3)},\ x_2^{(3)},\ x_3^{(3)},\ x_4^{(3)}\right)^{\text{т}}$$

найдем из систем $(A - E)e_1^{(l)} = e_1^{(l-1)}$, $l = 2, 3$, т.е. из систем

$$\begin{pmatrix} 0 & -3 & 0 & 3 \\ -2 & -7 & 0 & 13 \\ 0 & -3 & 0 & 3 \\ -1 & -4 & 0 & 7 \end{pmatrix} \begin{pmatrix} x_1^{(2)} \\ x_2^{(2)} \\ x_3^{(2)} \\ x_4^{(2)} \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 3 \\ 1 \end{pmatrix},$$

$$\begin{pmatrix} 0 & -3 & 0 & 3 \\ -2 & -7 & 0 & 13 \\ 0 & -3 & 0 & 3 \\ -1 & -4 & 0 & 7 \end{pmatrix} \begin{pmatrix} x_1^{(3)} \\ x_2^{(3)} \\ x_3^{(3)} \\ x_4^{(3)} \end{pmatrix} = \begin{pmatrix} x_1^{(2)} \\ x_2^{(2)} \\ x_3^{(2)} \\ x_4^{(2)} \end{pmatrix}.$$

Решениями этих систем, например, являются

$$x_1^{(2)} = 0, \quad x_2^{(2)} = -2, \quad x_3^{(2)} = 0, \quad x_4^{(2)} = -1,$$
$$x_1^{(3)} = 1, \quad x_2^{(3)} = x_3^{(3)} = x_4^{(3)} = 0.$$

Следовательно,

$$e_1^{(2)} = (0,\ -2,\ 0,\ -1)^{\text{т}}, \quad e_1^{(3)} = (1,\ 0,\ 0,\ 0)^{\text{т}}.$$

Жорданова цепочка векторов $e_1^{(1)}$, $e_1^{(2)}$, $e_1^{(3)}$ еще не дает базиса в корневом пространстве X. Поэтому перейдем к рассмотрению подпространства

$$R_{k-2} = R_1 = L_{k-2} \cap E(1) = L_1 \cap E(1),$$

где L_1 — подпространство столбцов матрицы $A - E$. Так как ранг матрицы $A - E$ равен двум, то L_1 порождается двумя его линейно независимыми столбцами, например столбцами

$$a_1 = (0,\ -2,\ 0,\ -1)^{\text{т}} \quad \text{и} \quad a_2 = (-3,\ -7,\ -3,\ -4)^{\text{т}}.$$

Таким образом, $L_1 = \langle a_1,\ a_2\rangle$, $E(1) = \langle b_1,\ b_2\rangle$, где

$$a_1 = (0,\ -2,\ 0,\ -1)^{\text{т}}, \quad a_2 = (-3,\ -7,\ -3,\ -4)^{\text{т}},$$
$$b_1 = (0,\ 0,\ 1,\ 0)^{\text{т}}, \quad b_2 = (3,\ 1,\ 0,\ 1)^{\text{т}}.$$

Для отыскания базиса в $R_1 = L_1 \cap E(1)$ составим векторное равенство $c = \alpha_1 a_1 + \alpha_2 a_2 = \beta_1 b_1 + \beta_2 b_2$ и перейдем от него к покоординатным равенствам. Тогда получим систему

$$\begin{cases} -3\alpha_2 = 3\beta_2, \\ -2\alpha_1 - 7\alpha_2 = \beta_2, \\ -3\alpha_2 = \beta_1, \\ -\alpha_1 - 4\alpha_2 = \beta_2. \end{cases}$$

Фундаментальную систему решений этой системы уравнений составляет, например, решение $\alpha_1 = 3$, $\alpha_2 = -1$, $\beta_1 = 3$, $\beta_2 = 1$. Поэтому $c = 3a_1 - a_2 = (3,\ 1,\ 3,\ 1)^{\text{т}}$ и $R_1 = \langle c\rangle$.

Подпространство R_1 совпало с уже рассмотренным подпространством R_2. Поэтому следует перейти к подпространству

$$R_{k-3} = R_0 = E(1) = \langle b_1,\ b_2\rangle = \langle (0,\ 0,\ 1,\ 0)^{\text{т}},\ (3,\ 1,\ 0,\ 1)^{\text{т}}\rangle.$$

Выберем в R_0 базис, начинающийся с $e_1^{(1)} = (3,\ 1,\ 3,\ 1)^{\text{т}}$. Очевидно, что $R_0 = \langle e_1^{(1)},\ b_1\rangle = \langle e_1^{(1)},\ (0,\ 0,\ 1,\ 0)^{\text{т}}\rangle$. Поэтому за $e_2^{(1)}$

следует принять $e_2^{(1)} = (0,\ 0,\ 1,\ 0)^{\text{т}}$. В итоге построен жорданов базис

$$e_1^{(1)} = (3,\ 1,\ 3,\ 1)^{\text{т}}, \qquad e_1^{(2)} = (0,\ -2,\ 0,\ -1)^{\text{т}},$$
$$e_1^{(3)} = (1,\ 0,\ 0,\ 0)^{\text{т}}, \qquad e_2^{(1)} = (0,\ 0,\ 1,\ 0)^{\text{т}}.$$

В этом базисе матрица A имеет жорданову форму

$$J = \left(\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & \boxed{1} \end{array}\right),$$

а матрицей T, трансформирующей A в J, является матрица

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

составленная из столбцов координат векторов построенного жорданова базиса.

Пример 6.2. Для матрицы

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

построить ее жорданову форму J и трансформирующую матрицу T.

Решение. В пятимерном пространстве X в некотором базисе матрица A определяет линейный оператор $\boldsymbol{A}$. Построим в пространстве X жорданов базис этого оператора. Для этого сначала найдем характеристический многочлен

$$|A-\lambda E| = \begin{vmatrix} -2-\lambda & -1 & -1 & 3 & 2 \\ -4 & 1-\lambda & -1 & 3 & 2 \\ 1 & 1 & -\lambda & -3 & -2 \\ -4 & -2 & -1 & 5-\lambda & 1 \\ 4 & 1 & 1 & -3 & -\lambda \end{vmatrix} = -(\lambda-2)^3(\lambda+1)^2.$$

Он имеет корень $\lambda_1 = 2$ кратности $m_1 = 3$ и корень $\lambda_2 = -1$ кратности $m_2 = 2$. Поэтому пространство X распадается в прямую сумму корневых подпространств K_1 и K_2 соответственно по $\lambda_1 = 2$ и $\lambda_2 = -1$. Сначала построим жорданов базис оператора $\boldsymbol{A}$ в K_1. Для этого составим матрицу $A-2E$ и будем возводить ее в степени $m = 1, 2, \ldots$ до тех пор, пока не получим равенство

$$r(A-2E)^m = n - m_1 = 5 - 3 = 2.$$

При $m = 1$

$$(A-2E)^m = A-2E = \begin{pmatrix} -4 & -1 & -1 & 3 & 2 \\ -4 & -1 & -1 & 3 & 2 \\ 1 & 1 & -2 & -3 & -2 \\ -4 & -2 & -1 & 3 & 1 \\ 4 & 1 & 1 & -3 & -2 \end{pmatrix},$$

$$r(A-2E) = 3 \neq n - m_1 = 2.$$

При $m = 2$

$$(A-2E)^m = (A-2E)^2 = \begin{pmatrix} 15 & 0 & 6 & -9 & -9 \\ 15 & 0 & 6 & -9 & -9 \\ 6 & 0 & 3 & 9 & 9 \\ 15 & 0 & 6 & -9 & -9 \\ -15 & 0 & -6 & 9 & 9 \end{pmatrix},$$

$$r(A-2E)^2 = 2 = n - m_1.$$

Следовательно, в K_1 жордановы цепочки имеют максимальную длину $k_1 = 2$.

Найдем пространство $E(2)$ собственных векторов по $\lambda=2$. Для этого рассмотрим систему $(A-2E)X=0$, т.е. систему

$$\begin{cases} -4x_1-x_2-x_3+3x_4+2x_5=0,\\ -4x_1-x_2-x_3+3x_4+2x_5=0,\\ x_1+x_2-2x_3-3x_4-2x_5=0,\\ -4x_1-2x_2-x_3+3x_4+x_5=0,\\ 4x_1+x_2+x_3-3x_4-2x_5=0. \end{cases}$$

Ее фундаментальная система решений состоит из двух векторов, например из векторов

$$b_1=(1,\ 0,\ -1,\ 1,\ 0)^{\mathrm{T}},\qquad b_2=(1,\ -1,\ -1,\ 0,\ -1)^{\mathrm{T}}.$$

Поэтому $E(2)=\langle b_1,\ b_2\rangle$. Найдем пересечение

$$R_{k_1-1}=R_1=L_1\cap E(2)=\langle a_1,\ a_2,\ a_3\rangle\cap\langle b_1,\ b_2\rangle,$$

где $L_1=L_{k_1-1}=\langle a_1,\ a_2,\ a_3,\ a_4,\ a_5\rangle=\langle a_1,\ a_2,\ a_3\rangle$ — пространство столбцов матрицы $(A-2E)^{k_1-1}=A-2E$ (через a_1, a_2, a_3, a_4, a_5 обозначены столбцы матрицы $A-2E$). Положим

$$e^{(1)}=\alpha_1a_1+\alpha_2a_2+\alpha_3a_3=\beta_1b_1+\beta_2b_2$$

и от полученного векторного равенства перейдем к покоординатным равенствам

$$\begin{cases} -4\alpha_1-\alpha_2-\alpha_3=\beta_1+\beta_2,\\ -4\alpha_1-\alpha_2-\alpha_3=-\beta_2,\\ \alpha_1+\alpha_2-2\alpha_3=\beta_1-\beta_2,\\ -4\alpha_1-2\alpha_2-\alpha_3=\beta_1,\\ 4\alpha_1+\alpha_2+\alpha_3=\beta_2. \end{cases}$$

Эта система имеет фундаментальную систему решений $\alpha_1=0$, $\alpha_2=-1$, $\alpha_3=0$, $\beta_1=2$, $\beta_2=-1$. Она определяет лишь один вектор

$$e_1^{(1)}=-a_2=2b_1-b_2=(1,\ 1,\ -1,\ 2,\ -1)^{\mathrm{T}}.$$

Поэтому $R_1 = \langle e_1^{(1)} \rangle = \langle (1,\ 1,\ -1,\ 2,\ -1)^{\mathrm{T}} \rangle$. Вектор $e_1^{(1)}$ является собственным вектором в K_1, с которого начинается жорданова цепочка длины 2. Присоединенный вектор $e_1^{(2)}$ жордановой цепочки, связанной с $e_1^{(1)}$, найдем из системы $(A-2E)e_1^{(2)} = e_1^{(1)}$, т.е. из системы

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

Решая эту систему, находим, например, $e_1^{(2)} = (0,\ -1,\ 0,\ 0,\ 0)^{\mathrm{T}}$.

Построенная жорданова цепочка векторов $e_1^{(1)}$, $e_1^{(2)}$ еще не дает базиса в K_1. Поэтому перейдем к рассмотрению пересечения

$$R_{K_1-2} = R_0 = L_{k_1-2} \cap E(2) = L_0 \cap E(2).$$

Здесь L_0 — подпространство столбцов единичной матрицы $(A-2E)^0$. Поэтому L_0 совпадает со всем пространством пятимерных векторов, и, следовательно, R_0 будет совпадать с $E(2)$, т.е. $R_0 = E(2) = \langle b_1,\ b_2 \rangle$. В подпространстве R_0 построим базис, содержащий уже использованный ранее собственный вектор $e_1^{(1)}$. Таким базисом в R_0 является, например, базис, состоящий из векторов $e_1^{(1)}$ и $e_2^{(1)} = b_1 = (1,\ 0,\ -1,\ 1,\ 0)^{\mathrm{T}}$. Тогда вектор $e_2^{(1)} = b_1$ является собственным вектором в K_1, составляющим жорданову цепочку единичной длины.

Число построенных векторов $e_1^{(1)}$, $e_1^{(2)}$, $e_2^{(1)}$ совпадает с размерностью $m_1 = 3$ подпространства K_1. Следовательно, эти векторы составляют жорданов базис в K_1.

Перейдем к построению жорданова базиса оператора $\boldsymbol{A}$ в корневом подпространстве K_2 по $\lambda = -1$. Для этого составим матрицу $A+E$ и будем ее возводить в степени $m = 1, 2, \ldots$ до тех пор, пока не получим равенство $r(A+E)^m = n - m_2 = 5 - 2 = 3$.

При $m=1$

$$(A+E)^m = A+E = \begin{pmatrix} -1 & -1 & -1 & 3 & 2 \\ -4 & 2 & -1 & 3 & 2 \\ 1 & 1 & 1 & 3 & 2 \\ -4 & -2 & -1 & 6 & 1 \\ 4 & 1 & 1 & -3 & 1 \end{pmatrix},$$

$$r(A+E) = 4 \neq n - m_2 = 3.$$

При $m=2$

$$(A+E)^m = (A+E)^2 = \begin{pmatrix} 0 & -6 & 0 & 9 & 3 \\ -9 & 3 & 0 & 9 & 3 \\ 0 & 6 & 0 & -9 & -3 \\ -9 & -12 & 0 & 18 & -3 \\ 9 & 6 & 0 & -9 & 6 \end{pmatrix},$$

$$r(A+E)^2 = 3 = n - m_2.$$

Следовательно, в K_2 жордановы цепочки имеют максимальную длину $k_2 = 2$.

Найдем подпространство $E(-1)$ собственных векторов пс $\lambda = -1$. Для этого рассмотрим систему $(A+E)X = 0$, т.е. систему

$$\begin{cases} -x_1 - x_2 - x_3 + 3x_4 + 3x_5 = 0, \\ -4x_1 + 2x_2 - x_3 + 3x_4 + 2x_5 = 0, \\ x_1 + x_2 + x_3 - 3x_4 - 2x_5 = 0, \\ -4x_1 - 2x_2 - x_3 + 6x_4 + x_5 = 0, \\ 4x_1 + x_2 + x_3 - 3x_4 + x_5 = 0. \end{cases}$$

Ее фундаментальная система решений состоит из одногс вектора, например, $b = (-1,\ -1,\ 1,\ -1,\ 1)^{\mathrm{T}}$. Поэтому $E(-1) = \langle b \rangle = \langle (-1,\ -1,\ 1,\ -1,\ 1)^{\mathrm{T}} \rangle$. Пересечение $R_{k_2-1} = R_1 = L_{k_2-1} \cap E(-1) = L_1 \cap E(-1)$, где $L_{k_2-1} = L_1$ — подпространство столбцов матрицы $A+E$, совпадает с $E(-1)$, так как вектор b совпадает с третьим столбцом матрицы $A+E$.

Итак, $R_1 = E(-1) = \langle b \rangle$. Поэтому полагаем: $e_3^{(1)} = b = (-1,\ -1,\ 1,\ -1,\ 1)^{\mathrm{T}}$. Присоединенный вектор $e_3^{(2)}$ найдем из системы $(A+E)e_3^{(2)} = e_3^{(1)}$, т.е. из системы

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

Одним из решений указанной системы является, например, вектор $e_3^{(2)} = (-1,\ -1,\ 2,\ -1,\ 1)^{\text{т}}$. Число векторов $e_3^{(1)}$, $e_3^{(2)}$ построенной жордановой цепочки совпадает с размерностью $k_2 = 2$ корневого подпространства K_2. Следовательно, эти векторы составляют базис в K_2. В итоге для оператора $\boldsymbol{A}$ в пространстве X построен жорданов базис $e_1^{(1)}$, $e_1^{(2)}$, $e_2^{(1)}$, $e_3^{(1)}$, $e_3^{(2)}$. Он состоит из трех жордановых цепочек

$$e_1^{(1)} = (1,\ 1,\ -1,\ 2,\ -1)^{\text{т}}, \quad e_1^{(2)} = (0,\ -1,\ 0,\ 0,\ 0)^{\text{т}},$$

$$e_2^{(1)} = (1,\ 0,\ -1,\ 1,\ 0)^{\text{т}},$$

$$e_3^{(1)} = (-1,\ -1,\ 1,\ -1,\ 1)^{\text{т}}, \quad e_3^{(2)} = (-1,\ -1,\ 2,\ -1,\ 1)^{\text{т}},$$

имеющих соответственно длины 2, 1 и 2. Первые две относятся к корневому подпространству K_1 по $\lambda = 2$, третья — к корневому подпространству K_2 по $\lambda = -1$. Поэтому жорданова форма J матрицы A имеет вид:

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

Матрица T, приводящая матрицу A к жордановой форме J, составляется из столбцов координат векторов $e_1^{(1)}$, $e_1^{(2)}$, $e_2^{(1)}$,

$e_3^{(1)}$, $e_3^{(2)}$ и имеет вид:

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

6.3. Второй способ построения жордановой и трансформирующей матриц

В некоторых случаях не требуется построение жорданова базиса, а нужно лишь определить жорданову форму матрицы. Тогда можно сократить объем вычислений. Для этого по каждому характеристическому корню λ_i матрицы A необходимо выполнить следующие действия.

1. Составить матрицу $A - \lambda_i E$ и возводить ее последовательно в степени $m = 1, 2, \ldots$ до тех пор, пока не получится равенство

$$r(()A - \lambda_i E)^m = n - m_i, \tag{6.23}$$

где $r(A - \lambda_i E)^m$ — ранг матрицы $(A - \lambda_i E)^m$; n — порядок матрицы A; m_i — кратность характеристического корня λ_i матрицы A. Наименьшее натуральное число m, при котором выполняется равенство (6.23), даст максимальный порядок k_i жордановых клеток по λ_i в матрице J.

2. По формуле

$$q_h = 2m_h - m_{h-1} - m_{h+1}, \quad h = 1, 2, \ldots, k_i, \tag{6.24}$$

или по формуле

$$q_h = r_{h-1} - 2r_h + r_{h+1}, \quad h = 1, 2, \ldots, k_i, \tag{6.25}$$

определить число q_h жордановых клеток по λ_i порядка h, $h=1,2,\ldots,k_i$, $i=1,2,\ldots,s$. Здесь q_h — число жордановых клеток по λ_i порядка h в жордановой форме матрицы A; $m_0=0$, m_ρ — дефект оператора с матрицей $(A-\lambda_i E)^\rho$; $r_0=n$, r_ρ — ранг матрицы $(A-\lambda_i E)^\rho$.

3. По найденным числам q_h, $h=1,2,\ldots,k_i$, для всех λ_i матрицы A составить матрицу J.

Пример 6.3. Для матрицы

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

построить жорданову форму J.

Решение. Характеристический многочлен

$$|A-\lambda E|=\begin{vmatrix}-\lambda & -4 & -2\\ 1 & 4-\lambda & 1\\ 0 & 0 & 1-\lambda\end{vmatrix}=-(\lambda-2)^2(\lambda-1)$$

имеет корень $\lambda_1=2$ кратности $m_1=2$ и корень $\lambda_2=1$ кратности $m_2=1$.

При $\lambda_1=2$ имеем:

$$A-2E=\begin{pmatrix}-2 & -4 & -2\\ 1 & 2 & 1\\ 0 & 0 & -1\end{pmatrix},\ r(A-2E)=r_1=2\neq n-m_1=3-2=1,$$

$$(A-2E)^2=\begin{pmatrix}0 & 0 & 2\\ 0 & 0 & -1\\ 0 & 0 & 1\end{pmatrix},\quad r(A-2E)^2=r_2=1=n-m_1,$$

$$(A-2E)^3=\begin{pmatrix}0 & 0 & -2\\ 0 & 0 & -1\\ 0 & 0 & 1\end{pmatrix},\quad r(A-2E)^3=r_3=1.$$

Следовательно, наибольший порядок жордановых клеток по $\lambda_1=2$ равен $k_1=2$ и по формуле (6.25) находим:

$$q_1=r_0-2r_1+r_2=3-2\cdot 2+1=0,$$

$$q_2=r_1-2r_2+r_3=2-2\cdot 1+1=1.$$

Поэтому в жордановой форме J матрицы A по характеристическому числу $\lambda_1 = 2$ будет всего лишь одна жорданова клетка $\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$.

При $\lambda_2 = 1$ имеем:

$$A - 2E = \begin{pmatrix} -1 & -4 & -2 \\ 1 & 3 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad r(A - 2E) = r_1 = 2 = n - m_2,$$

$$(A - 2E)^2 = \begin{pmatrix} -3 & -8 & -2 \\ 2 & 5 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad r(A - 2E)^2 = r_2 = 2.$$

Поэтому в жордановой форме по характеристическому числу $\lambda_2 = 1$ будет $q_1 = r_0 - 2r_1 + r_2 = 3 - 2 \cdot 2 + 2 = 1$ жордановых клеток порядка 1. Из полученных жордановых клеток составляем жорданову форму матрицы A:

$$J = \left(\begin{array}{|cc|c} \hline 2 & 1 & 0 \\ 0 & 2 & 0 \\ \hline 0 & 0 & \multicolumn{1}{|c|}{1} \\ \cline{3-3} \end{array}\right).$$

Если для матрицы A известна ее жорданова матрица J, то для отыскания трансформирующей матрицы T достаточно решить матричное уравнение

$$\boldsymbol{TJ = AT.} \tag{6.26}$$

По столбцам матрицы T легко выписать векторы жорданова базиса оператора A в пространстве X.

Пример 6.4. Для матрицы A, имеющей жорданову форму J, найти трансформирующую матрицу T, если

$$A = \begin{pmatrix} 0 & -4 & -2 \\ 1 & 4 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \qquad J = \left(\begin{array}{|cc|c} \hline 2 & 1 & 0 \\ 0 & 2 & 0 \\ \hline 0 & 0 & \multicolumn{1}{|c|}{1} \\ \cline{3-3} \end{array}\right).$$

Решение. Для решения задачи следует решить матричное уравнение $TJ = AT$, т.е. уравнение

z_1 z_2 z_3

$$\begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -4 & -2 \\ 1 & 4 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{pmatrix}.$$

Проведя умножение матриц в обеих частях этого уравнения и сравнив соответствующие элементы левой и правой частей этого равенства, получим систему

$$\begin{cases} 2x_1 = -4x_2 - 2x_3, \\ x_1 + 2y_1 = -4y_2 - 2y_3, \\ z_1 = -4z_2 - 2z_3, \\ 2x_2 = x_1 + 4x_2 + x_3, \\ x_2 + 2y_2 = y_1 + 4y_2 + y_3, \\ z_2 = z_1 + 4z_2 + z_3, \\ 2x_3 = x_3, \\ x_3 + 2y_3 = y_3, \\ z_3 = z_3. \end{cases}$$

Одним из решений этой системы является

$$\begin{array}{lll} x_1 = 2, & y_1 = -1, & z_1 = 2, \\ x_2 = -1, & y_2 = 0, & z_2 = -1, \\ x_3 = 0, & y_3 = 0, & z_3 = 1. \end{array}$$

Поэтому

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

По столбцам этой матрицы можно выписать векторы жорданова базиса оператора $\boldsymbol{A}$:

$$e_1^{(1)} = (2,\ -1,\ 0)^{\mathrm{T}}, \quad e_1^{(2)} = (-1,\ 0,\ 0)^{\mathrm{T}}, \quad e_2^{(1)} = (2,\ -1,\ 1)^{\mathrm{T}}.$$

6.4. Третий способ построения жордановой и трансформирующей матриц

Приведем еще один способ построения жордановой и трансформирующей матриц. Предварительно сделаем несколько замечаний.

Под элементарными преобразованиями над λ-матрицей $A(\lambda)$ понимают:

1) умножение любой строки матрицы $A(\lambda)$ на любое число $\alpha \neq 0$;

2) умножение любого столбца матрицы $A(\lambda)$ на любое число $\alpha \neq 0$;

3) прибавление к любой i-й строке любой j-й строки, $j \neq i$, умноженной на любой многочлен $\varphi(\lambda)$;

4) прибавление к любому i-му столбцу любого j-го столбца, умноженного на любой многочлен $\varphi(\lambda)$.

Проверкой легко убедиться, что первое и третье элементарные преобразования над матрицей $A(\lambda)$ равносильны умножению ее слева соответственно на элементарные матрицы

$$\begin{array}{cc} & \begin{array}{ccccccc} & & & i & & & \end{array} \\ \begin{array}{c} \\ \\ \\ i \\ \\ \\ \\ \end{array} & \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & 0 & \\ & & 1 & & & & \\ & & & \alpha & & & \\ & & & & 1 & & \\ & 0 & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \end{array}, \quad \begin{array}{cc} & \begin{array}{ccccccc} & & j & & & & \end{array} \\ \begin{array}{c} \\ \\ \\ \\ i \\ \\ \\ \end{array} & \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & 0 & \\ & & 1 & & & & \\ & & & \ddots & & & \\ & & \alpha & & 1 & & \\ & 0 & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \end{array}. \qquad (6.27)$$

Аналогично второе и четвертое элементарные преобразования над матрицей $A(\lambda)$ равносильны умножению ее справа на элементарные матрицы вида (6.27).

Каноническим видом матрицы $A(\lambda)$ называют диагональную матрицу

$$\begin{pmatrix} e_1(\lambda) & & & \\ & e_2(\lambda) & & \\ & & \ddots & \\ & & & e_n(\lambda) \end{pmatrix},$$

в которой каждый многочлен $e_i(\lambda)$, $i = 2, 3, \ldots, n$, нацело делится на многочлен $e_{i-1}(\lambda)$ и старший коэффициент каждого многочлена $e_i(\lambda)$, $i = 1, 2, \ldots, n$, равен единице. Многочлены $e_1(\lambda), \ldots, e_n(\lambda)$ называют ***инвариантными множителями матрицы*** $A(\lambda)$. Многочлен $e_n(\lambda)$ совпадает с минимальным многочленом матрицы A.

Для построения жордановой формы J матрицы A порядка n можно использовать метод, который состоит в выполнении следующих действий.

1. Составить матрицу $A - \lambda E$ и для каждого k, $1 \leqslant k \leqslant n$, вычислить наибольший общий делитель $d_k(\lambda)$ всех миноров k-го порядка этой матрицы, взятый со старшим коэффициентом, равным единице.

2. Вычислить инвариантные множители $e_1(\lambda)$, ..., $e_n(\lambda)$ матрицы $A - \lambda E$ по формулам

$$e_k(\lambda) = \frac{d_k(\lambda)}{d_{k-1}(\lambda)}, \quad k = 1, 2, \ldots, n. \tag{6.28}$$

3. Разложить каждый инвариантный множитель $e_k(\lambda)$ в произведение элементарных делителей, т.е. в произведение вида

$$e_k(\lambda) = (\lambda - \lambda_1)^{n_{k_1}} \ldots (\lambda - \lambda_s)^{n_{k_s}}, \quad k = 1, 2, \ldots, n. \tag{6.29}$$

4. Составить жорданову форму J матрицы A, ставя по ее диагонали жорданову клетку по λ_j порядка n_{k_j} для каждого элементарного делителя $(\lambda - \lambda_j)^{n_{k_j}}$ в разложениях (6.29).

Замечание. Пункты 1 и 2 в описанной последовательности действий можно заменить следующим: инвариантные множители $e_1(\lambda), \ldots, e_n(\lambda)$ матрицы $A - \lambda E$ получить путем приведения ее к каноническому виду с помощью элементарных преобразований над нею.

Пример 6.5. Найти жорданову форму J для матрицы

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

Решение. Составим матрицу

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

Очевидно, что общий наибольший делитель всех ее элементов $d_1(\lambda) = 1$. Для отыскания $d_2(\lambda)$ выпишем все миноры второго порядка матрицы $A - \lambda E$:

$$\begin{vmatrix} -\lambda & 1 \\ -4 & 4-\lambda \end{vmatrix} = (\lambda-2)^2, \quad \begin{vmatrix} -\lambda & 0 \\ -4 & 0 \end{vmatrix} = 0, \quad \begin{vmatrix} 1 & 0 \\ 4-\lambda & 0 \end{vmatrix} = 0,$$

$$\begin{vmatrix} -\lambda & 1 \\ -2 & 1 \end{vmatrix} = -\lambda + 2, \quad \begin{vmatrix} -\lambda & 0 \\ -2 & 2-\lambda \end{vmatrix} = \lambda(\lambda-2),$$

$$\begin{vmatrix} 1 & 0 \\ 1 & 2-\lambda \end{vmatrix} = 2-\lambda, \quad \begin{vmatrix} -4 & 1-\lambda \\ -2 & 1 \end{vmatrix} = -2(\lambda-2),$$

$$\begin{vmatrix} -4 & 0 \\ -2 & 2-\lambda \end{vmatrix} = 4(\lambda-2), \quad \begin{vmatrix} 4-\lambda & 0 \\ 1 & 2-\lambda \end{vmatrix} = (2-\lambda)(4-\lambda).$$

Отсюда видно, что $d_2(\lambda) = \lambda - 2$. Для отыскания $d_3(\lambda)$ замечаем, что

$$|A - \lambda E| = \begin{vmatrix} -\lambda & 1 & 0 \\ -4 & 4-\lambda & 0 \\ -2 & 1 & 2-\lambda \end{vmatrix} = -(\lambda-2)^3.$$

Следовательно, $d_3(\lambda) = (\lambda-2)^3$. Теперь по формулам (6.28) получаем:

$$e_1(\lambda)=d_1(\lambda)=1,\ e_2(\lambda)=\frac{d_2(\lambda)}{d_1(\lambda)}=\lambda-2,\ e_3(\lambda)=\frac{d_3(\lambda)}{d_2(\lambda)}=(\lambda-2)^2.$$

Поэтому

$$J=\left(\begin{array}{cc|c} 2 & 1 & 0 \\ 0 & 2 & 0 \\ \hline 0 & 0 & \boxed{1} \end{array}\right).$$

Если действовать в соответствии с замечанием, то матрицу $A-\lambda E$ нужно элементарными преобразованиями привести к каноническому виду:

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

$$\sim\begin{pmatrix} -\lambda & 1 & 0 \\ -4 & 4-\lambda & 0 \\ 0 & 0 & 2-\lambda \end{pmatrix}\sim\begin{pmatrix} 1 & 1 & 0 \\ -\lambda & 4-\lambda & 0 \\ 0 & 0 & 2 \end{pmatrix}\sim$$

$$\sim\begin{pmatrix} 1 & \lambda-3 & 0 \\ -\lambda & 4-\lambda & 0 \\ 0 & 0 & 2-\lambda \end{pmatrix}\sim\begin{pmatrix} 1 & \lambda-3 & 0 \\ 0 & (\lambda-2)^2 & 0 \\ 0 & 0 & 2-\lambda \end{pmatrix}\sim$$

$$\sim\begin{pmatrix} 1 & 0 & 0 \\ 0 & (\lambda-2)^2 & 0 \\ 0 & 0 & 2-\lambda \end{pmatrix}\sim\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 2-\lambda \\ 0 & (\lambda-2)^2 & 0 \end{pmatrix}\sim$$

$$\sim\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2-\lambda & 0 \\ 0 & 0 & (\lambda-2)^2 \end{pmatrix}\sim\begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda-2 & 0 \\ 0 & 0 & (\lambda-2)^2 \end{pmatrix}.$$

Отсюда следует, что $e_1(\lambda)=1$, $e_2(\lambda)=\lambda-2$, $e_3(\lambda)=(\lambda-1)^2$. Поэтому

$$J=\left(\begin{array}{cc|c} 2 & 1 & 0 \\ 0 & 2 & 0 \\ \hline 0 & 0 & \boxed{1} \end{array}\right).$$

Чтобы определить матрицу T, приводящую матрицу A к жордановой форме J, можно использовать следующий метод.

1. Привести матрицу $A-\lambda E$ элементарными преобразованиями к виду $J-\lambda E$.

2. Выбрать по порядку все проведенные элементарные преобразования над столбцами, найти соответствующие им элементарные матрицы вида (6.27) и составить матрицу $V(\lambda)$, равную произведению этих элементарных матриц слева направо в том же порядке, в каком они использовались.

3. Матрицу $V(\lambda)$ записать в виде

$$V(\lambda)=V_0\lambda^k+V_1\lambda^{k-1}+\ldots+V_k\lambda^0 \tag{6.30}$$

и подсчитать матрицу

$$T=V(A)=V_0A^k+V_1A^{k-1}+\ldots+V_kA^0. \tag{6.31}$$

Если матрица $V(\lambda)$ окажется матрицей нулевой степени от λ, то она сразу дает матрицу T.

Запись λ-матрицы в виде многочлена от λ не представляет затруднений. Например,

$$\begin{pmatrix}\lambda^2+5\lambda+1 & 6\lambda+2\\ 7\lambda & \lambda^2-2\lambda+3\end{pmatrix}=\begin{pmatrix}1&0\\0&1\end{pmatrix}\lambda^2+\begin{pmatrix}5&6\\7&-2\end{pmatrix}\lambda+\begin{pmatrix}1&2\\0&3\end{pmatrix}.$$

Замечание. Этот метод можно видоизменить, если вместо пунктов 2 и 3 выполнить следующее. Выбрать по порядку все проведенные элементарные преобразования над строками, заменить их элементарными матрицами вида (6.27) и составить матрицу $U(\lambda)$, равную произведению этих элементарных матриц справа налево в том порядке, в каком они использовались. Матрицу $U(\lambda)$ записать в виде

$$U(\lambda)=\lambda^mU_0+\lambda^{m-1}U_1+\ldots+\lambda^0U_m \tag{6.32}$$

и подсчитать матрицу

$$T^{-1}=U(A)=A^mU_0+A^{m-1}U_1+\ldots+A^0U_m. \tag{6.33}$$

Пример 6.6. Матрица

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

имеет жорданову форму

$$J = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Найти трансформирующую матрицу T, приводящую матрицу A к матрице J.

Решение. Сначала элементарными преобразованиями приведем матрицу $A - \lambda E$ к матрице $J - \lambda E$:

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

$$\sim \begin{pmatrix} -\lambda & 1 & 0 \\ -4 & 4-\lambda & 0 \\ 0 & 0 & 2-\lambda \end{pmatrix} \sim \begin{pmatrix} -\lambda & 1 & 0 \\ -4+2\lambda & 2-\lambda & 0 \\ 0 & 0 & 2-\lambda \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 2-\lambda & 1 & 0 \\ 0 & 2-\lambda & 0 \\ 0 & 0 & 2-\lambda \end{pmatrix} = J - \lambda E.$$

Над столбцами матрицы $A - \lambda E$ проведены элементарные преобразования, равносильные умножению ее справа последовательно на матрицы

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Поэтому

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

Поскольку матрица $V(\lambda)$ оказалась матрицей нулевой степени от λ, то она сразу дает матрицу T. Следовательно,

$$T = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

Точно так же элементарные преобразования, проведенные над строками матрицы $A - \lambda E$, равносильны умножению ее слева последовательно на матрицы

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Поэтому

$$U(\lambda) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} = T^{-1}.$$

6.5. К построению минимального многочлена

При построении минимального многочлена матрицы A порядка n может оказаться полезным следующий метод. Для каждого характеристического корня λ_i, $i = 1, 2, \ldots, s$, матрицы A составить матрицу $A - \lambda_i E$ и возводить ее в степени $m = 1, 2, \ldots$ до тех пор, пока не будет выполнено равенство

$$r(A - \lambda_i E)^m = n - m_i, \tag{6.34}$$

где $r(A - \lambda_i E)^m$ — ранг матрицы $(A - \lambda_i E)^m$; n — порядок матрицы A; m_i — кратность характеристического корня λ_i матрицы A.

Наименьшее натуральное число m, при котором выполняется равенство (6.34), дает кратность k_i корня λ_i в минимальном многочлене $\varphi(\lambda)$ матрицы A. Поэтому минимальным многочленом матрицы A будет многочлен

$$\varphi(\lambda) = (\lambda - \lambda_1)^{k_1}(\lambda - \lambda_2)^{k_2}\ldots(\lambda - \lambda_s)^{k_s}. \tag{6.35}$$

Заметим, что каждое k_i совпадает с наибольшим порядком жордановой клетки по λ_i в жордановой форме матрицы A. Поэтому все числа $k_1, \ldots, k_s$, а следовательно, и минимальный многочлен матрицы A легко выписать, если известна жорданова форма матрицы A. Полезно также помнить, что минимальный многочлен матрицы A совпадает с инвариантным множителем $e_n(\lambda)$ матрицы $A - \lambda E$.

Пример 6.7. Для матрицы

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

характеристический многочлен

$$|A - \lambda E| = \begin{vmatrix} -\lambda & 1 & 0 \\ -4 & 4-\lambda & 0 \\ -2 & 1 & 2-\lambda \end{vmatrix} = -(\lambda - 2)^3$$

имеет единственный корень $\lambda_1 = 2$ кратности $m_1 = 3$. При этом

$$A - 2E = \begin{pmatrix} -2 & 1 & 0 \\ -4 & 2 & 0 \\ -2 & 1 & 0 \end{pmatrix}, \quad r(A - 2E) = 1 \neq n - m_1 = 0,$$

$$(A - 2E)^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad r(A - 2E)^2 = 0 = n - m_1.$$

Следовательно, $k_1 = 2$ и $\varphi(\lambda) = (\lambda - 2)^2$. Минимальный многочлен матрицы A можно было выписать по ее жордановой форме J или по инвариантному множителю $e_3(\lambda)$ матрицы $A - \lambda E$, полученным в примере 6.5.

Упражнения

6.1. Пользуясь первым методом (см. разд. 6.2), построить жорданов базис оператора $\mathcal{A}$ с матрицей A, жорданову форму J матрицы A и трансформирующую матрицу T для следующих матриц A:

1) $\begin{pmatrix} 9 & -6 & -2 \\ 18 & -12 & -3 \\ 18 & -9 & -6 \end{pmatrix}$; 2) $\begin{pmatrix} 4 & 6 & -15 \\ 1 & 3 & -5 \\ 1 & 2 & -4 \end{pmatrix}$; 3) $\begin{pmatrix} 0 & -4 & 0 \\ 1 & -4 & 0 \\ 1 & -2 & -2 \end{pmatrix}$;

4) $\begin{pmatrix} 1 & -3 & 3 \\ -2 & -6 & 13 \\ -1 & -4 & 8 \end{pmatrix}$; 5) $\begin{pmatrix} -4 & 2 & 10 \\ -4 & 3 & 7 \\ -3 & 1 & 7 \end{pmatrix}$; 6) $\begin{pmatrix} 1 & -3 & 4 \\ 4 & -7 & 8 \\ 6 & -7 & 7 \end{pmatrix}$;

7) $\begin{pmatrix} 1 & -3 & 0 & 3 \\ -2 & -6 & 0 & 13 \\ 0 & -3 & 1 & 3 \\ -1 & -4 & 0 & 8 \end{pmatrix}$; 8) $\begin{pmatrix} 3 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 3 & 0 & 5 & -3 \\ 4 & -1 & 3 & -1 \end{pmatrix}$;

9) $\begin{pmatrix} 3 & 1 & 0 & 0 \\ -4 & -1 & 0 & 0 \\ 7 & 1 & 2 & 1 \\ -17 & -6 & -1 & 0 \end{pmatrix}$; 10) $\begin{pmatrix} 3 & -4 & 0 & 2 \\ 4 & -5 & -2 & 4 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & 2 & -1 \end{pmatrix}$;

11) $\begin{pmatrix} 1 & 1 & -2 & 0 \\ 2 & 1 & 0 & 2 \\ 1 & 0 & 1 & 1 \\ 0 & -1 & 2 & 1 \end{pmatrix}$; 12) $\begin{pmatrix} 3 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 1 \\ -1 & -1 & -1 & 1 \end{pmatrix}$.

6.2. Пользуясь вторым методом построения жордановой матрицы (см. разд. 6.3), найти для каждой матрицы A из упражнения 6.1 жорданову форму J и трансформирующую матрицу T.

6.3. Пользуясь третьим методом построения жордановой матрицы (см. разд. 6.4), найти для каждой матрицы A из упражнения 6.1 жорданову форму J и трансформирующую матрицу T.

6.4. Для каждой матрицы A из упражнения 6.1 найти ее минимальный многочлен.

Глава 7 Функции от матриц

7.1. Интерполяционный многочлен Лагранжа — Сильвестра

Пусть даны различные значения $\lambda_1, \lambda_2, \ldots, \lambda_s$ аргумента λ и значения

$$f(\lambda_k), \quad f'(\lambda_k), \quad f''(\lambda_k), \quad \ldots, \quad f^{(m_k-1)}(\lambda_k), \quad k=1,2,\ldots,s,$$

функции $f(\lambda)$ и ее производных до (m_k-1)-го порядка включительно при этих значениях аргумента λ. Будем искать многочлен $P(\lambda)$ степени $n-1$ при $n = m_1+m_2+\ldots+m_s$, удовлетворяющий условиям

$$\begin{cases} P(\lambda_k) = f(\lambda_k), \\ P'(\lambda_k) = f'(\lambda_k), \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ P^{(m_k-1)}(\lambda_k) = f^{(m_k-1)}(\lambda_k), \end{cases} \qquad k = 1,2,\ldots,s. \tag{7.1}$$

Такой многочлен называют ***интерполяционным многочленом Лагранжа — Сильвестра*** для функции $f(\lambda)$ при интерполяционных условиях (7.1). Для его построения составляют ***определяющий многочлен***

$$\psi(\lambda) = (\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\ldots(\lambda-\lambda_s)^{m_s}, \tag{7.2}$$

где $m_1+m_2+\ldots+m_s = n$, и многочлены

$$\psi_k(\lambda) = \frac{\psi(\lambda)}{(\lambda-\lambda_k)^{m_k}}. \tag{7.3}$$

Затем находят коэффициенты

$$\alpha_{kj} = \frac{1}{(j-1)!}\left[\frac{f(\lambda)}{\psi_k(\lambda)}\right]^{(j-1)}_{\lambda=\lambda_k}, \tag{7.4}$$
$$j = 1, 2, \ldots, m_k, \quad k = 1, 2, \ldots, s,$$

и выписывают искомый многочлен

$$P(\lambda)=\sum_{k=1}^{s}\left[\alpha_{k1}+\alpha_{k2}(\lambda-\lambda_k)+\ldots+\alpha_{km_k}(\lambda-\lambda_k)^{m_k-1}\right]\psi_k(\lambda). \tag{7.5}$$

Пример 7.1. Для функции $f(\lambda)=e^\lambda$ построить интерполяционный многочлен $P(\lambda)$, удовлетворяющий интерполяционным условиям:

$$P(1)=e, \quad P'(1)=e, \quad P(2)=e^2, \quad P(3)=e^3.$$

Решение. Здесь узлами интерполяции являются $\lambda_1=1$, $\lambda_2=2$, $\lambda_3=3$, для этих узлов $m_1=2$, $m_2=1$, $m_3=1$. По формуле (7.2) составляем определяющий многочлен

$$\psi(\lambda)=(\lambda-1)^2(\lambda-2)(\lambda-3),$$

а по формулам (7.3) — многочлены

$$\psi_1(\lambda)=\frac{\psi(\lambda)}{(\lambda-1)^2}=(\lambda-2)(\lambda-3), \quad \psi_2(\lambda)=\frac{\psi(\lambda)}{\lambda-2}=(\lambda-1)^2(\lambda-3),$$
$$\psi_3(\lambda)=\frac{\psi(\lambda)}{\lambda-3}=(\lambda-1)^2(\lambda-2).$$

Поэтому формула (7.5) дает

$$P(\lambda)=\left[\alpha_{11}+\alpha_{12}(\lambda-1)\right](\lambda-2)(\lambda-3)+$$
$$+\alpha_{21}(\lambda-1)^2(\lambda-3)+\alpha_{31}(\lambda-1)^2(\lambda-2)$$

где коэффициенты

$$\alpha_{11}=\left.\frac{e^\lambda}{(\lambda-2)(\lambda-3)}\right|_{\lambda=1}=\frac{e}{2}, \quad \alpha_{12}=\left[\frac{e^\lambda}{(\lambda-2)(\lambda-3)}\right]'_{\lambda=1}=\frac{5e}{4},$$
$$\alpha_{21}=\left.\frac{e^\lambda}{(\lambda-1)^2(\lambda-3)}\right|_{\lambda=2}=-e^2, \quad \alpha_{31}=\left.\frac{e^\lambda}{(\lambda-1)^2(\lambda-2)}\right|_{\lambda=3}=\frac{e^3}{4}$$

находятся по формулам (7.4). Окончательно имеем:

$$P(\lambda)=\left[\frac{e}{2}+\frac{5e}{4}(\lambda-1)\right](\lambda-2)(\lambda-3)-e^2(\lambda-1)^2(\lambda-3)+$$
$$+\frac{e^3}{4}(\lambda-1)^2(\lambda-2)=\left(\frac{e^3}{4}-e^2+\frac{5e}{4}\right)\lambda^3+(5e^2-e^3-7e)\lambda^2+$$
$$+\left(\frac{5e^3}{4}-7e^2+\frac{45e}{4}\right)\lambda+\left(3e^2-\frac{9e}{4}-\frac{e^3}{4}\right).$$

7.2. Функции от матриц

Пусть

$$\varphi(\lambda)=(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\ldots(\lambda-\lambda_s)^{n_s} \tag{7.6}$$

— минимальный многочлен матрицы A. Говорят, что ***функция*** $f(\lambda)$ ***определена на спектре матрицы*** A, если существуют значения

$$f(\lambda_k),\quad f'(\lambda_k),\quad \ldots,\quad f^{(m_k-1)}(\lambda_k) \tag{7.7}$$

при $m_k \geqslant n_k$ для всех характеристических чисел λ_k, $k = 1, 2, \ldots, s$, матрицы A.

За значение функции $f(\lambda)$ при $\lambda = A$ принимают значение интерполяционного многочлена Лагранжа — Сильвестра $P(\lambda)$ при $\lambda = A$, построенного для функции $f(\lambda)$ при определяющем многочлене $\psi(\lambda)=\varphi(\lambda)$ и интерполяционных условиях (7.7) с $m_k = n_k$, т.е. полагают $f(A) = P(A)$. Определяющий многочлен $\psi(\lambda)$ можно брать с показателями $m_k \geqslant n_k$. При этом степень многочлена $P(\lambda)$ соответственно увеличится, значение же $f(A) = P(A)$ не изменится. Еще раз подчеркнем, что интерполяционный многочлен $P(\lambda)$ для функции $f(\lambda)$ получается наименьшей степени при $\psi(\lambda)=\varphi(\lambda)$.

Пример 7.2. Найти $f(\lambda)=\sin\frac{\pi}{2}\lambda$ при

$$\lambda = A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Решение. Характеристический многочлен матрицы A

$$|A-\lambda E| = \begin{vmatrix} 2-\lambda & 0 & 0 \\ 0 & 1-\lambda & 1 \\ 0 & 0 & 1-\lambda \end{vmatrix} = (2-\lambda)(1-\lambda)^2$$

имеет корни $\lambda_1 = \lambda_2 = 1$, $\lambda_3 = 2$. Общий наибольший делитель D_2 миноров второго порядка матрицы $A - \lambda E$ равен единице, так как ее миноры

$$\begin{vmatrix} 2-\lambda & 0 \\ 0 & 1 \end{vmatrix} = 2-\lambda, \qquad \begin{vmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{vmatrix} = (1-\lambda)^2$$

взаимно простые. Поэтому минимальный многочлен матрицы A

$$\varphi(\lambda) = \frac{(-1)^3|A-\lambda E|}{D_2} = (\lambda-1)^2(\lambda-2).$$

За определяющий многочлен $\psi(\lambda)$ интерполяционного многочлена $P(\lambda)$ примем

$$\psi(\lambda) = \varphi(\lambda) = (\lambda-1)^2(\lambda-2).$$

Тогда

$$\psi_1(\lambda) = \frac{\psi(\lambda)}{(\lambda-1)^2} = \lambda-2, \qquad \psi_2(\lambda) = \frac{\psi(\lambda)}{\lambda-2} = (\lambda-1)^2$$

и

$$P(\lambda) = \left[\alpha_{11} + \alpha_{12}(\lambda-1)\right](\lambda-2) + \alpha_{21}(\lambda-1)^2,$$

где

$$\alpha_{11} = \left.\frac{\sin\frac{\pi}{2}\lambda}{\lambda-2}\right|_{\lambda=1} = -1, \qquad \alpha_{12} = \left[\frac{\sin\frac{\pi}{2}\lambda}{\lambda-2}\right]'_{\lambda=1} = -1,$$

$$\alpha_{21} = \left.\frac{\sin\frac{\pi}{2}\lambda}{(\lambda-1)^2}\right|_{\lambda=2} = 0.$$

Поэтому окончательно имеем

$$P(\lambda) = \left[-1-(\lambda-1)\right](\lambda-2) = 2\lambda-\lambda^2.$$

Следовательно,

$$\sin\frac{\pi}{2}A = P(A) = 2A - A^2 =$$

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

Если бы за определяющий многочлен интерполяционного многочлена взяли, например,

$$\psi(\lambda) = (\lambda - 1)^2(\lambda - 2)^2,$$

то по формуле (7.5) для функции $f(\lambda) = \sin\frac{\pi}{2}\lambda$ получили бы:

$$P_1(\lambda) = \left[\alpha_{11} + \alpha_{12}(\lambda - 1)\right](\lambda - 2)^2 + \left[\alpha_{21} + \alpha_{22}(\lambda - 2)^2\right](\lambda - 1)^2 =$$

$$= \left[1 + 2(\lambda - 1)\right](\lambda - 2)^2 - \frac{\pi}{2}(\lambda - 2)(\lambda - 1)^2 =$$

$$= (2\lambda - 1)(\lambda - 2)^2 - \frac{\pi}{2}(\lambda - 2)(\lambda - 1)^2,$$

так как по формулам (7.4) в этом случае получаем:

$$\alpha_{11} = \left.\frac{f(\lambda)}{(\lambda - 2)^2}\right|_{\lambda=1} = 1, \qquad \alpha_{12} = \left[\frac{f(\lambda)}{(\lambda - 2)^2}\right]'_{\lambda=1} = 2,$$

$$\alpha_{21} = \left.\frac{f(\lambda)}{(\lambda - 1)^2}\right|_{\lambda=2} = 0, \qquad \alpha_{22} = \left[\frac{f(\lambda)}{(lm - 1)^2}\right]'_{\lambda=2} = -\frac{\pi}{2}.$$

Естественно, что для $\sin\frac{\pi}{2}A$ получим тот же результат, хотя его вычисление несколько усложнится:

$$\sin\frac{\pi}{2}A = P_1(A) = (2A - E)(A - 2E)^2 -$$

$$- \frac{\pi}{2}(A - 2E)(A - E)^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Заметим, что $P(\lambda)$ получается как остаток от деления многочлена $P_1(\lambda)$ на минимальный многочлен $\varphi(\lambda)$ матрицы A.

7.3. Спектральное разложение матрицы $f(A)$

Пусть для функции $f(\lambda)$, определенной на спектре матрицы A, построен интерполяционный многочлен

$$P(\lambda) = \sum_{k=1}^{s} \left[\alpha_{k1} + \alpha_{k2}(\lambda - \lambda_k) + \ldots + \alpha_{kn_k}(\lambda - \lambda_k)^{n_k - 1}\right]\psi_k(\lambda)$$

наименьшей степени, т.е. при $\psi(\lambda) = \varphi(\lambda)$. Если в формулу $f(\lambda) = P(\lambda)$ вместо коэффициентов α_{kj} подставить их выражения из формул (7.4) в развернутой форме, раскрыть скобки, объединить члены, содержащие $f^{(i-1)}(\lambda_k)$, $i = 1, 2, \ldots, n_k$, $k = 1, 2, \ldots, s$, то ее можно записать в виде

$$f(\lambda) = P(\lambda) = \sum_{k=1}^{s} \left[f(\lambda_k)\varphi_{k1}(\lambda) + f'(\lambda_k)\varphi_{k2}(\lambda) + \ldots \right.$$
$$\left. \ldots + f^{(n_k - 1)}(\lambda_k)\varphi_{kn_k}(\lambda)\right].$$

Поэтому получаем:

$$f(A) = P(A) = \sum_{k=1}^{s} \left[f(\lambda_k)Z_{k1} + f'(\lambda_k)Z_{k2} + \ldots \right.$$
$$\left. \ldots + f^{(n_k - 1)}(\lambda_k)Z_{kn_k}\right], \quad (7.8)$$

где

$$Z_{kj} = \varphi_{ki}(A), \quad j = 1, 2, \ldots, n_k, \quad k = 1, 2, \ldots, s. \quad (7.9)$$

Разложение (7.8) называют ***спектральным разложением матрицы*** $f(A)$, а матрицы Z_{kj} — ***компонентами матрицы*** A.

Разложение (7.8) особенно удобно, если требуется вычислить несколько функций от одной и той же матрицы. Если при построении спектрального разложения берется интерполяционный многочлен не наименьшей степени, то в разложении (7.8) появятся дополнительные компоненты $Z_{k,j}$, но все они окажутся нулевыми матрицами. На практике компоненты Z_{kj} матрицы A проще находить не по формулам (7.9), а из системы, которая получится в результате последовательной подстановки в (7.8) вместо $f(\lambda)$ простейших линейно независимых многочленов столько раз, сколько содержится в (7.8) компонент Z_{kj}. Поясним этот способ на примере.

Пример 7.3. Найти компоненты матрицы

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

и вычислить значения функций $\sin\frac{\pi}{2}\lambda$, $\cos\frac{\pi}{2}\lambda$, $\operatorname{tg}\pi\lambda$, $e^{\lambda t}$, λ^{-1}, $\sqrt{\lambda}$, λ^{100} при $\lambda = A$.

Решение. В предыдущем примере для матрицы A найден минимальный многочлен

$$\varphi(\lambda) = (\lambda - 1)^2(\lambda - 2).$$

Он имеет корни $\lambda_1 = \lambda_2 = 1$, $\lambda_3 = 2$. Поэтому для любой функции $f(\lambda)$, определенной на спектре матрицы A, будет

$$f(A) = f(1)Z_{11} + f'(1)Z_{12} + f(2)Z_{21}. \tag{7.10}$$

Отсюда получаем:

$$\begin{aligned} &\text{при } f(\lambda) = 1 && E = Z_{11} + Z_{21}, \\ &\text{при } f(\lambda) = 1 - \lambda && E - A = -Z_{12} - Z_{21}, \\ &\text{при } f(\lambda) = (1 - \lambda)^2 && (E - A)^2 = Z_{21}. \end{aligned}$$

Из этой системы находим:

$$Z_{11} = 2A - A^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad Z_{21} = (E - A)^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$Z_{12} = 3A - 2E - A^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Теперь, полагая в (7.10) вместо $f(\lambda)$ последовательно $\sin\frac{\pi}{2}\lambda$, $\cos\frac{\pi}{2}\lambda$, $\operatorname{tg}\pi\lambda$, $e^{\lambda t}$, λ^{-1}, $\sqrt{\lambda}$, λ^{100}, получаем:

$$\sin\frac{\pi}{2}A = Z_{11} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix};$$

$$\cos\frac{\pi}{2}A = -\frac{\pi}{2}Z_{12} - Z_{21} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & -\frac{\pi}{2} \\ 0 & 0 & 0 \end{pmatrix};$$

$$\operatorname{tg}\pi A = \pi Z_{12} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \pi \\ 0 & 0 & 0 \end{pmatrix};$$

$$e^{At} = e^t Z_{11} + te^t Z_{12} + e^{2t} Z_{21} = \begin{pmatrix} e^{2t} & 0 & 0 \\ 0 & e^t & te^t \\ 0 & 0 & e^t \end{pmatrix};$$

$$A^{-1} = Z_{11} - Z_{12} + \frac{1}{2}Z_{21} = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix};$$

$$\sqrt{A} = Z_{11} + \frac{1}{2}Z_{12} + \sqrt{2}Z_{21} = \begin{pmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix};$$

$$A^{100} = Z_{11} + 100Z_{12} + 2^{100}Z_{21} = \begin{pmatrix} 2^{100} & 0 & 0 \\ 0 & 1 & 100 \\ 0 & 0 & 1 \end{pmatrix}.$$

Полезным может оказаться следующее компактное выражение компонент матрицы с помощью криволинейных интегралов:

$$Z_{kj} = \frac{1}{(j-1)!2\pi i} \int\limits_{L_k} (\lambda - \lambda_k)^{j-1} (\lambda E - A)^{-1} d\lambda, \tag{7.11}$$
$$k = 1, 2, \ldots, s, \quad j = 1, 2, \ldots, n_k,$$

где L_k — окружность с центром в точке $\lambda = \lambda_k$, не содержащая других характеристических корней матрицы A внутри и на L_k; λ_1, λ_2, ..., λ_s — различные характеристические корни матрицы A; n_1, n_2, ..., n_s — кратности корней λ_1, λ_2, ..., λ_s в минимальном многочлене матрицы A.

Криволинейные интегралы применяют и для представления матрицы $f(A)$. Если квадратная матрица A порядка n имеет различные характеристические числа λ_1, λ_2, ..., λ_s, L — замкнутый контур, ограничивающий область, содержащую внутри точки λ_1, λ_2, ..., λ_s, и $f(\lambda)$ — функция, определенная на спектре матрицы A, непрерывная на контуре L и аналитическая внутри области, ограниченной контуром L, то

$$f(A) = \frac{1}{2\pi i} \int\limits_{L} f(\lambda)(\lambda E - A)^{-1} d\lambda. \tag{7.12}$$

Замечание. В формулах (7.11) и (7.12) под интегралом от матрицы понимается матрица, элементами которой являются интегралы от соответствующих элементов подынтегральной матрицы.

При применении формул (7.11) и (7.12) вычисление криволинейных интегралов удобно проводить, опираясь на теорему о вычетах, по которой для функции $\varphi(\lambda)$, непрерывной на замкнутом контуре L и аналитической внутри области, ограниченной контуром L, кроме конечного числа изолированных особых точек λ_1, λ_2, ..., λ_s, имеет место равенство

$$\int\limits_{L} \varphi(\lambda) d\lambda = 2\pi i \sum_{k=1}^{s} \text{Выч}[\varphi(\lambda), \lambda_k],$$

где $\text{Выч}[\varphi(\lambda), \lambda_k]$ — вычет аналитической функции $\varphi(\lambda)$ в точке λ_k.

7.4. Представление функций от матриц рядами

Пусть дана квадратная матрица A порядка n с минимальным многочленом

$$\varphi(\lambda) = (\lambda - \lambda_1)^{n_1} \dots (\lambda - \lambda_s)^{n_s}.$$

Говорят, что ряд $\sum\limits_{p=0}^{\infty} u_p(\lambda)$ сходится на спектре матрицы A к функции $f(\lambda)$, если функции $f(\lambda)$, $u_0(\lambda)$, ..., $u_p(\lambda)$, ... определены на спектре матрицы A и выполняются соотношения

$$\begin{cases} \sum\limits_{p=0}^{\infty} u_p(\lambda_k) = f(\lambda_k), \\ \sum\limits_{p=0}^{\infty} u_p'(\lambda_k) = f'(\lambda_k), \\ \dots\dots\dots\dots\dots\dots \\ \sum\limits_{p=0}^{\infty} u_p^{(n_k-1)}(\lambda_k) = f^{(n_k-1)}(\lambda_k), \end{cases} \qquad k = 1, 2, \dots, s.$$

Для того чтобы ряд $\sum\limits_{p=0}^{\infty} u_p(A)$ функций от матрицы A сходился к матрице $f(A)$, т.е. чтобы

$$f(A) = \sum_{p=0}^{\infty} u_p(A),$$

необходимо и достаточно, чтобы ряд $\sum\limits_{p=0}^{\infty} u_p(\lambda)$ сходился на спектре матрицы A к функции $f(\lambda)$. Отсюда вытекают, например, следующие разложения:

$$e^A = \sum_{p=0}^{\infty} \frac{A^p}{p!}, \quad \cos A = \sum_{p=0}^{\infty} (-1)^p \frac{A^{2p}}{(2p)!}, \quad \sin A = \sum_{p=0}^{\infty} (-1)^p \frac{A^{2p+1}}{(2p+1)!}.$$

7.5. Некоторые приложения функций от матриц

Приведем три случая использования функций от матриц в решении конкретных задач.

1. Для вычисления m-й степени матрицы A при любом действительном m следует рассмотреть функцию $f(\lambda) = \lambda^m$ и вычислить ее значение при $\lambda = A$ (см. разд. 7.3).

2. Система линейных однородных дифференциальных уравнений с постоянными коэффициентами

$$\begin{cases} \dfrac{dx_1}{dt} = a_{11}x_1 + \ldots + a_{1n}x_n, \\ \ldots\ldots\ldots\ldots\ldots \\ \dfrac{dx_n}{dt} = a_{n1}x_1 + \ldots + a_{nn}x_n \end{cases} \tag{7.13}$$

в матричной форме записывается в виде

$$\frac{dX}{dt} = AX, \tag{7.14}$$

где $X = (x_1, \ldots, x_n)^{\mathrm{T}}$, $A = (a_{ij})$. Ее решение, удовлетворяющее начальным условиям

$$X\Big|_{t=0} = (x_{10}, \ldots, x_{n0})^{\mathrm{T}} = X_0, \tag{7.15}$$

находится по формуле

$$X = e^{At}X_0. \tag{7.16}$$

Если матрицу e^{At} представить по формуле (7.8) в виде спектрального разложения

$$e^{At} = \sum_{k=1}^{s}\left[Z_{k1} + tZ_{k2} + \ldots + t^{n_k-1}Z_{kn_k}\right]e^{\lambda_k t}, \tag{7.17}$$

то решение (7.16) системы (7.13), удовлетворяющее начальным условиям (7.15), примет вид:

$$X = \left[\sum_{k=1}^{s}(Z_{k1} + tZ_{k2} + \ldots + t^{n_k-1}Z_{kn_k})e^{\lambda_k t}\right]X_0. \tag{7.18}$$

Заметим, что если $X_0 = (c_1, c_2, \ldots, c_n)^{\mathrm{T}}$, где $c_1, c_2, \ldots, c_n$ — произвольные постоянные, то формулы (7.16), (7.18) дают общее решение системы (7.13).

Пример 7.4. Найти решение системы

$$\begin{cases} \dfrac{dx_1}{dt} = 6x_1 + 2x_2 + 2x_3, \\ \dfrac{dx_2}{dt} = -2x_1 + 2x_2, \\ \dfrac{dx_3}{dt} = 2x_3, \end{cases}$$

удовлетворяющее начальным условиям

$$x_1\big|_{t=0} = 0, \quad x_2\big|_{t=0} = 1, \quad x_3\big|_{t=0} = 1.$$

Решение. В матричной форме данная система записывается в виде

$$\frac{dX}{dt} = AX,$$

где

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad A = \begin{pmatrix} 6 & 2 & 2 \\ -2 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix},$$

а начальные условия — в виде $X\big|_{t=0} = X_0 = (0,\ 1,\ 1)^{\mathrm{T}}$.

Для того чтобы воспользоваться формулой (7.16), нужно знать матрицу e^{At}. Вычислим ее. В примере 5.7 был найден минимальный многочлен $\varphi(\lambda) = (\lambda-4)^2(\lambda-2)$ матрицы A. За определяющий многочлен $\psi(\lambda)$ интерполяционного многочлена $P(\lambda)$ примем $\psi(\lambda) = \varphi(\lambda) = (\lambda-4)^2(\lambda-2)$. Тогда по формулам (7.3) находим:

$$\psi_1(\lambda) = \frac{\psi(\lambda)}{(\lambda-4)^2} = \lambda - 2, \qquad \psi_2(\lambda) = \frac{\psi(\lambda)}{\lambda-2} = (\lambda-4)^2,$$

а по формуле (7.5) для функции $f(\lambda) = e^{\lambda t}$ имеем:

$$P(\lambda) = \left[\alpha_{11} + \alpha_{12}(\lambda - 4)\right](\lambda - 2) + \alpha_{21}(\lambda - 4)^2 =$$

$$= \left[\frac{1}{2}e^{4t} + \frac{1}{4}e^{4t}(2t-1)(\lambda - 4)\right](\lambda - 2) + \frac{1}{4}e^{2t}(\lambda - 4)^2,$$

так как в соответствии с (7.4)

$$\alpha_{11} = \left.\frac{e^{\lambda t}}{\lambda - 2}\right|_{\lambda=4} = \frac{e^{4t}}{2}, \qquad \alpha_{12} = \left[\frac{e^{\lambda t}}{\lambda - 2}\right])'_{\lambda=4} = \frac{e^{4t}}{4}(2t-1),$$

$$\alpha_{21} = \left.\frac{e^{\lambda t}}{(\lambda - 4)^2}\right|_{\lambda=2} = \frac{e^{2t}}{4}.$$

Полагая $e^{At} = P(A)$, получаем:

$$e^{At} = \frac{e^{4t}}{2}\left[E + \frac{1}{2}(2t-1)(A - 4E)\right](A - 2E) + \frac{e^{2t}}{4}(A - 4E)^2 =$$

$$= \begin{pmatrix} (2t+1)e^{4t} & 2te^{4t} & 2te^{4t} \\ -2te^{4t} & (1-2t)e^{4t} & (1-2t)e^{4t} - e^{2t} \\ 0 & 0 & e^{2t} \end{pmatrix}.$$

Теперь по формуле (7.16) находим:

$$X = e^{At}X_0 = \begin{pmatrix} 4te^{4t} \\ (2-4t)e^{4t} - e^{2t} \\ e^{2t} \end{pmatrix}.$$

Если воспользоваться формулой (7.18), то сначала следует найти компоненты матрицы A. Для этого по формуле (7.8) для любой функции $f(\lambda)$, определенной на спектре матрицы A, получаем:

$$f(A) = f(4)Z_{11} + f'(4)Z_{12} + f(2)Z_{21}. \qquad (7.19)$$

Полагая в этом разложении поочередно $f(\lambda) = 1$, $f(\lambda) = \lambda - 4$, $f(\lambda) = (\lambda - 4)^2$, приходим к системе

$$\begin{cases} Z_{11} + Z_{21} = E, \\ Z_{12} - 2Z_{21} = A - 4E, \\ 4Z_{21} = (A - 4E)^2, \end{cases}$$

из которой находим:

$$Z_{11} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad Z_{12} = \begin{pmatrix} 2 & 2 & 2 \\ -2 & -2 & -2 \\ 0 & 0 & 0 \end{pmatrix}, \quad Z_{21} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Теперь по формуле (7.18) имеем:

$$X = \left[(Z_{11} + tZ_{12})e^{4t} + Z_{21}e^{2t}\right] X_0 = \begin{pmatrix} 4te^{4t} \\ (2 - 4t)e^{4t} - e^{2t} \\ e^{2t} \end{pmatrix}.$$

3. Для системы неоднородных дифференциальных уравнений с постоянными коэффициентами

$$\frac{dX}{dt} = AX + f, \tag{7.20}$$

где

$$X = (x_1, \ldots, x_n)^{\mathrm{T}}, \quad A = (a_{ij}), \quad f = (f_1(t), \ldots, f_n(t))^{\mathrm{T}},$$

решение, удовлетворяющее начальным условиям

$$X\big|_{t=0} = X_0, \tag{7.21}$$

находится по любой из следующих формул:

$$X = e^{At} X_0 + \int\limits_0^t e^{A(t-\tau)} f(\tau) d\tau, \tag{7.22}$$

$$X = \left[\sum_{k=1}^{s} (Z_{k1} + tZ_{k2} + \ldots + t^{n_k - 1} Z_{kn_k}) e^{\lambda_k t}\right] X_0 + \\ + \int\limits_0^t e^{A(t-\tau)} f(\tau) d\tau. \tag{7.23}$$

Пример 7.5. Найти решение системы

$$\begin{cases} \dfrac{dx_1}{dt} = 6x_1 + 2x_2 + 2x_3 + ae^4, \\ \dfrac{dx_2}{dt} = -2x_1 + 2x_2, \\ \dfrac{dx_3}{dt} = 2x_3, \end{cases}$$

удовлетворяющее начальным условиям $X\big|_{t=0} = X_0 = (0,\ 1,\ 1)^{\mathrm{T}}$.

Решение. Выражение

$$e^{At}X_0 = \left[\sum_{k=1}^{s}\left(Z_{k1} + tZ_{k2} + \ldots + t^{n_k-1}Z_{kn_k}\right)e^{\lambda_k t}\right]X_0 =$$

$$= \begin{pmatrix} 4te^{4t} \\ (2-4t)e^{4t} - e^{2t} \\ e^{2t} \end{pmatrix},$$

входящее в формулы (7.22), (7.23), найдено в предыдущем примере. Чтобы вычислить интеграл, входящий в эти формулы, составим сначала спектральное разложение для матрицы $e^{A(t-\tau)}$. Для этого в формуле (7.19) положим $f(\lambda) = e^{\lambda(t-\tau)}$ и получим

$$e^{A(t-\tau)} = e^{4(t-\tau)}Z_{11} + (t-\tau)e^{4(t-\tau)}Z_{12} + e^{2(t-\tau)}Z_{21}. \qquad (7.24)$$

Матрицы Z_{kj} также найдены в предыдущем примере. Теперь нужный интеграл распишется в виде

$$\int_0^t e^{A(t-\tau)}f(\tau)d\tau = Z_{11}\int_0^t e^{4(t-\tau)}f(\tau)d\tau +$$

$$+ Z_{12}\int_0^t (t-\tau)e^{4(t-\tau)}f(\tau)d\tau + Z_{21}\int_0^t e^{2(t-\tau)}f(\tau)d\tau.$$

Вычислим интегралы, стоящие в правой части этого равенства:

$$\int_0^t e^{4(t-\tau)} f(\tau) d\tau = \int_0^t e^{4(t-\tau)} \begin{pmatrix} ae^{4\tau} \\ 0 \\ 0 \end{pmatrix} d\tau = \begin{pmatrix} ate^{4t} \\ 0 \\ 0 \end{pmatrix},$$

$$\int_0^t (t-\tau) e^{4(t-\tau)} f(\tau) d\tau = \int_0^t (t-\tau) e^{4(t-\tau)} \begin{pmatrix} ae^{4\tau} \\ 0 \\ 0 \end{pmatrix} d\tau = \begin{pmatrix} \frac{at^2}{2} e^{4t} \\ 0 \\ 0 \end{pmatrix},$$

$$\int_0^t e^{2(t-\tau)} f(\tau) d\tau = \int_0^t e^{2(t-\tau)} \begin{pmatrix} ae^{4\tau} \\ 0 \\ 0 \end{pmatrix} d\tau = \begin{pmatrix} \frac{a}{2} e^{2t}(e^{2t}-1) \\ 0 \\ 0 \end{pmatrix}.$$

Таким образом,

$$\int_0^t e^{A(t-\tau)} f(\tau) d\tau = Z_{11} \begin{pmatrix} ate^{4t} \\ 0 \\ 0 \end{pmatrix} + Z_{12} \begin{pmatrix} \frac{at^2}{2} e^{4t} \\ 0 \\ 0 \end{pmatrix} +$$

$$+ Z_{21} \begin{pmatrix} \frac{a}{2} e^{2t}(e^{2t}-1) \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} ate^{4t}(1+t) \\ -at^2 e^{4t} \\ 0 \end{pmatrix}.$$

Следовательно, искомое решение имеет вид:

$$X = \begin{pmatrix} 4te^{4t} \\ (2-4t)e^{4t} - e^{2t} \\ e^{2t} \end{pmatrix} + \begin{pmatrix} ate^{4t}(1+t) \\ -at^2 e^{4t} \\ 0 \end{pmatrix}.$$

Упражнения

7.1. Для функции $f(\lambda) = \lambda^{-1}$ построить интерполяционный многочлен Лагранжа — Сильвестра, удовлетворяющий следующим интерполяционным условиям: $P(1) = f(1)$, $P'(1) = f'(1)$, $P(2) = f(2)$, $P'(2) = f'(2)$.

7.2. Используя интерполяционный многочлен Лагранжа — Сильвестра, вычислить следующие значения функций от матриц:

1) A^{100}, где $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$; 2) $\sqrt{A}$, где $A = \begin{pmatrix} 5 & 16 \\ 1 & 5 \end{pmatrix}$;

3) e^A, где $A = \begin{pmatrix} 73 & -48 \\ 96 & -63 \end{pmatrix}$; 4) $\ln A$, где $A = \begin{pmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{pmatrix}$.

7.3. Найти компоненты матрицы

$$A = \begin{pmatrix} 1 & -3 & 3 \\ -2 & -6 & 13 \\ -1 & -4 & 8 \end{pmatrix}$$

и вычислить матрицы $\sin\frac{\pi}{2}A$, $\operatorname{tg}\frac{\pi}{4}A$, A^{-1}, $\sqrt{A}$, A^{100}, e^{At}.

7.4. Найти решение системы дифференциальных уравнений

$$\begin{cases} \dfrac{dx_1}{dt} = 2x_1, \\ \dfrac{dx_2}{dt} = x_2 + x_3, \\ \dfrac{dx_3}{dt} = x_3, \end{cases}$$

удовлетворяющее начальным условиям $x_1\big|_{t=0} = 1$, $x_2\big|_{t=0} = 1$, $x_3\big|_{t=0} = 1$.

Глава 8 ЕВКЛИДОВЫ И УНИТАРНЫЕ ПРОСТРАНСТВА

8.1. Определение евклидова пространства. Матрица Грама

Говорят, что в действительном линейном пространстве X определена операция ***скалярного умножения векторов***, если любой паре векторов x и y из X поставлено в соответствие действительное число, которое называют ***скалярным произведением*** векторов x и y и обозначают символом (x, y), и если для любых $x, y, z \in X$ и любого действительного числа α выполняются следующие ***аксиомы скалярного произведения***:

1. $(x, y) = (y, x)$.
2. $(x + y, z) = (x, z) + (y, z)$.
3. $(\alpha x, y) = \alpha(x, y)$.
4. $(x, x) > 0$ при $x \neq 0$ и $(x, x) = 0$ при $x = 0$.

Пример 8.1. Пусть X — пространство геометрических векторов, изучаемых в векторной алгебре. Скалярное произведение, определяемое как произведение длин двух векторов на косинус угла между ними, удовлетворяет всем аксиомам скалярного произведения.

Пример 8.2. В арифметическом пространстве K_n столбцов высоты n скалярное произведение векторов

$$x = (x_1,\ x_2,\ \ldots,\ x_n)^{\mathrm{T}} \quad \text{и} \quad y = (y_1,\ y_2,\ \ldots,\ y_n)^{\mathrm{T}}$$

можно определить формулой

$$(x, y) = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n.$$

Нетрудно проверить выполнимость аксиом скалярного произведения. Например, проверим выполнимость аксиомы 4. Заметим, что

$$(x, x) = x_1^2 + x_2^2 + \ldots + x_n^2.$$

Но сумма квадратов положительна, если хотя бы одно из чисел x_i ненулевое (или $x \neq 0$), и равна нулю, если все x_i равны нулю (т.е. $x = 0$).

Пример 8.3. В линейном пространстве многочленов с действительными коэффициентами степени не выше $n-1$ скалярное произведение можно ввести формулой

$$(f, g) = \int\limits_{-1}^{1} f(x)\, g(x)\, dx.$$

Проверка аксиом скалярного произведения опирается на свойства определенного интеграла и не составляет труда.

Пример 8.4. В линейном пространстве $C[a, b]$ функций действительного переменного, непрерывных на отрезке $[a, b]$, скалярное произведение можно ввести таким же образом, как и в линейном пространстве многочленов — с помощью определенного интеграла:

$$(f, g) = \int\limits_{a}^{b} f(x)\, g(x)\, dx.$$

Проверка аксиом скалярного произведения проводится так же, как и в предыдущем примере.

Из аксиом 2 и 3 следует, что ***любую конечную линейную комбинацию векторов можно умножать скалярно на другую линейную комбинацию векторов по правилу, аналогичному правилу умножения многочлена на многочлен, т.е. по формуле***

$$\left(\sum_{i=1}^{k} \alpha_i a_i,\ \sum_{j=1}^{l} \beta_j b_j\right) = \sum_{i=1}^{k} \sum_{j=1}^{l} \alpha_i \beta_j \,(a_i, b_j). \qquad (8.1)$$

Действительное линейное пространство, в котором определено скалярное умножение векторов, называют ***евклидовым пространством***. Конечномерное линейное пространство можно превратить в евклидово многими способами. Если в n-мерном евклидовом пространстве X фиксирован базис e_1, $e_2, \ldots, e_n$, то любые векторы x и y имеют в нем разложения

$$x = \sum_{i=1}^{n} x_i e_i, \qquad y = \sum_{j=1}^{n} y_j e_j$$

и формула (8.1) для векторов x и y дает

$$(x, y) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j (e_i, e_j), \tag{8.2}$$

или в матричном виде

$$(x, y) = x^{\mathrm{T}} \Gamma y, \tag{8.3}$$

где положено

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix}, \qquad y = \begin{pmatrix} y_1 \\ y_2 \\ \ldots \\ y_n \end{pmatrix},$$

$$\Gamma = (g_{ij}) = \begin{pmatrix} (e_1, e_1) & (e_1, e_2) & \ldots & (e_1, e_n) \\ (e_2, e_1) & (e_2, e_2) & \ldots & (e_2, e_n) \\ \ldots & \ldots & \ldots & \ldots \\ (e_n, e_1) & (e_n, e_2) & \ldots & (e_n, e_n) \end{pmatrix}.$$

Таким образом, ***скалярное произведение в евклидовом пространстве X полностью определяется матрицей Г***. Не всякая квадратная матрица может появиться в формуле (8.3). Но если одно скалярное произведение в заданном базисе определяется некоторой матрицей Г, то нетрудно понять, что та же матрица, только в другом базисе также определяет скалярное произведение. Сохраняя матрицу Γ и меняя базисы, мы получим бесконечное множество скалярных произведений в данном n-мерном линейном пространстве.

Матрицу Γ, участвующую в формуле (8.3), называют ***матрицей Грама*** базиса $e = (e_1, \ldots, e_n)$. Матрицу Грама (матрицу скалярных произведений) можно определить не только для базисов, но и для произвольных упорядоченных конечных систем векторов.

Отметим некоторые свойства матрицы Грама базиса в n-мерном евклидовом пространстве.

1. ***Матрица Грама* $\mathbf{\Gamma}$ *симметрическая и для любого n-мерного столбца $x \neq 0$ удовлетворяет условию* $x^{\mathrm{T}}\Gamma x > 0$.**

Симметричность матрицы Грама вытекает из аксиомы 1 скалярного произведения, согласно которой $(e_i, e_j) = (e_j, e_i)$ для любых двух векторов базиса, а условие $x^{\mathrm{T}}\Gamma x > 0$, $x \neq 0$, равносильно аксиоме 4 скалярного произведения.

Симметрическую матрицу A, удовлетворяющую условию $x^{\mathrm{T}}Ax > 0$, $x \neq 0$, называют ***положительно определенной***. С учетом этого термина доказанное свойство звучит так: ***матрица Грама является положительно определенной.***

2. ***Матрицы Грама* $\mathbf{\Gamma}$ *и* $\mathbf{\Gamma'}$ *двух базисов* e *и* e' *евклидова пространства связаны соотношением***

$$\Gamma' = T^{\mathrm{T}}\Gamma T, \tag{8.4}$$

***где* T *— матрица перехода от базиса* e *к базису* e'.**

Действительно, при переходе от базиса e к базису e' координаты x и y двух векторов x и y преобразуются в координаты x' и y' по формулам (см. разд. 4.6)

$$x = Tx', \quad y = Ty'.$$

Поэтому

$$(x, y) = x^{\mathrm{T}}\Gamma y = (Tx')^{\mathrm{T}}\Gamma(Ty) = (x')^{\mathrm{T}}T^{\mathrm{T}}\Gamma Ty'.$$

Следовательно, матрица $T^{\mathrm{T}}\Gamma T$ есть матрица Грама для базиса e

3. ***Определитель матрицы Грама любого базиса положителен.***

Действительно, из формулы (8.4) вытекает, что при замене базиса определитель матрицы Грама сохраняет знак (или остается равным нулю), так как определитель матрицы перехода ненулевой:

$$|\Gamma'| = |T^{\mathrm{T}}||\Gamma||T| = |\Gamma||T|^2.$$

Остается учесть, что в качестве матрицы Грама можно взять единичную матрицу (см. замечание ниже), которая имеет определитель, равный единице.

4. ***Все угловые диагональные миноры***

$$\boldsymbol{\Delta_k} = \begin{vmatrix} (e_1, e_1) & (e_1, e_2) & \dots & (e_1, e_k) \\ (e_2, e_1) & (e_2, e_2) & \dots & (e_2, e_k) \\ \dots & \dots & \dots & \dots \\ (e_k, e_1) & (e_k, e_2) & \dots & (e_k, e_k) \end{vmatrix}, \quad \boldsymbol{k = 1, 2, \dots, n,}$$

матрицы Грама базиса $e_1, e_2, \dots, e_n$ положительны.

Действительно, для любого k можно рассмотреть подпространство $L_k = \langle e_1, \dots, e_k \rangle$ как самостоятельное евклидово пространство. Тогда определитель матрицы Грама для базиса $e_1, \dots, e_k$ будет совпадать с Δ_k. Согласно предыдущему свойству этот определитель положителен.

Замечание. В разд. 9.6 установлено, что свойство 4 — необходимое и достаточное условие положительной определенности квадратной матрицы. Поэтому свойство 4 вытекает из свойства 1. Любая положительно определенная матрица является матрицей Грама некоторого базиса в данном евклидовом пространстве. Действительно, скалярное произведение можно определить формулой (8.3), в которой в качестве Γ можно взять любую положительно определенную матрицу. Тогда аксиома 1 скалярного произведения будет вытекать из симметричности матрицы Г, аксиомы 2 и 3 — из свойства дистрибутивности матричного произведения, а аксиома 4 — из условия положительной определенности Г. Следовательно, любая матрица, обладающая свойством 4, может рассматриваться как матрица Грама. В частности, в качестве матрицы Грама можно выбрать единичную матрицу, т.е. в заданном базисе $e_1, e_2, \ldots, e_n$ определить скалярное произведение формулой

$$(x, y) = x^{\mathrm{T}} y = (x_1,\ x_2,\ \ldots,\ x_n)\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} =$$

$$= x_1 y_1 + x_2 y_2 + \ldots x_n y_n. \quad (8.5)$$

Как уже отмечено, понятие матрицы Грама можно ввести для произвольной упорядоченной конечной системы векторов. При этом и в общем случае матрица Грама остается симметричной, но остальные свойства (положительная определенность, положительность определителя) утрачиваются. Имеет место следующее утверждение.

Теорема 8.1. *Матрица Грама системы векторов является невырожденной тогда и только тогда, когда эта система линейно независима. Матрица Грама линейно независимой системы векторов положительно определенная и, в частности, имеет положительный определитель.*

▷ Любая линейно независимая система векторов может рассматриваться как базис в некотором евклидовом пространстве, а именно в своей линейной оболочке. Поэтому второе утверждение теоремы вытекает из свойств матрицы Грама для базиса. Остается показать, что если матрица Грама невырождена, то система векторов линейно независима.

Рассмотрим произвольную линейную комбинацию системы векторов $a_1, a_2, \ldots, a_k$, равную нулю:

$$\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_k a_k = 0.$$

Умножая это векторное равенство скалярно на векторы a_1, a_2, ..., a_k, получим однородную систему линейных уравнений

$$\begin{cases} \alpha_1(a_1,a_1)+\alpha_2(a_1,a_2)+\ldots+\alpha_k(a_1,a_k)=0,\\ \alpha_1(a_2,a_1)+\alpha_2(a_2,a_2)+\ldots+\alpha_k(a_2,a_k)=0,\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ \alpha_1(a_k,a_1)+\alpha_2(a_k,a_2)+\ldots+\alpha_k(a_k,a_k)=0 \end{cases}$$

относительно коэффициентов α_1, ..., α_k рассматриваемой линейной комбинации. Матрицей этой системы является матрица Грама Γ системы векторов a_1, ..., a_k. Если матрица Γ невырождена, то однородная система имеет только нулевое решение. Это означает, что рассматриваемая линейная комбинация векторов a_1, ..., a_k равна нулю только в случае, когда все коэффициенты линейной комбинации равны нулю. Следовательно, система векторов a_1, ..., a_k линейно независима. ▶

8.2. Длины и углы. Ортогональность. Процесс ортогонализации

Длиной $|x|$ ***вектора*** x евклидова пространства E называют величину

$$|x| = \sqrt{(x,x)}. \tag{8.6}$$

Нормировать вектор x — значит заменить его вектором

$$x^0 = \frac{x}{|x|}. \tag{8.7}$$

Вектор x^0 называют ***единичным вектором***, или ***ортом*** вектора x.

Углом между ненулевыми векторами x и y евклидова пространства E называют угол φ, определяемый соотношениями

$$\cos\varphi = \frac{(x,y)}{|x|\,|y|}, \quad 0 \leqslant \varphi \leqslant \pi. \tag{8.8}$$

Корректность определения угла вытекает из неравенств

$$-1 \leqslant \frac{(x, y)}{|x|\,|y|} \leqslant 1,$$

равносильных ***неравенству Коши — Буняковского***

$$(x, y)^2 \leqslant (x, x)\,(y, y). \tag{8.9}$$

Докажем неравенство Коши — Буняковского. Для любых векторов x и y и любого числа α выполняется неравенство $(x - \alpha y, x - \alpha y) \geqslant 0$. Из этого неравенства получаем:

$$(x, x) - 2\alpha (x, y) + \alpha^2 (y, y) \geqslant 0.$$

Левая часть последнего неравенства представляет собой квадратный трехчлен относительно α. Поскольку этот трехчлен неотрицательный, его дискриминант не превышает нуля, т.е.

$$(x, y)^2 - (x, x)\,(y, y) \leqslant 0.$$

Отсюда следует неравенство (8.9).

Определим в n-мерном линейном пространстве скалярное произведение векторов

$$x = \sum_{i=1}^{n} x_i e_i, \qquad y = \sum_{i=1}^{n} y_i e_i$$

формулой

$$(x, y) = \sum_{i=1}^{n} x_i y_i.$$

Тогда неравенство Коши — Буняковского, выраженное через координаты векторов, примет вид:

$$\left(\sum_{i=1}^{n} x_i y_i\right)^2 \leqslant \left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i^2\right).$$

В общем случае скалярное произведение в координатах выражается через матрицу Грама $\Gamma = (g_{ij})$ формулой

$$(x, y) = \sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij}x_iy_j,$$

а неравенство Коши — Буняковского принимает вид:

$$\left(\sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij}x_iy_j\right)^2 \leqslant \left(\sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij}x_ix_j\right)\left(\sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij}y_iy_j\right).$$

Из неравенства Коши — Буняковского вытекает другое важное неравенство

$$|x+y| \leqslant |x| + |y|,$$

называемое ***неравенством треугольника***. Действительно, раскрывая величину $|x+y|^2$ как скалярный квадрат и учитывая, что в силу неравенства Коши — Буняковского $(x, y) \leqslant \leqslant |x|\,|y|$, находим:

$$\begin{aligned}|x+y|^2 &= (x+y, x+y) = (x, y) + 2(x, y) + (y, y) = \\ &= |x|^2 + 2(x, y) + |y|^2 \leqslant |x|^2 + 2|x|\,|y| + |y|^2 = \bigl(|x| + |y|\bigr)^2. \end{aligned}$$

Векторы x и y евклидова пространства E называют ***ортогональными***, если их скалярное произведение равно нулю, т.е. $(x, y) = 0$. Непосредственно из определения ортогональности векторов вытекает, что нулевой вектор ортогонален любому другому вектору.

Система ненулевых векторов называется ***ортогональной системой***, если любые два вектора этой системы ортогональны. Под ***ортонормированной системой*** понимают ортогональную систему, все векторы которой имеют единичную длину (т.е. нормированы). Отметим, что любую ортогональную систему можно превратить в ортонормированную простой нормировкой, так как нормирование векторов, как и вообще умножение на произвольные ненулевые числа, не нарушает условия их взаимной ортогональности. Например, если векторы x и y ортогональны, то в силу равенства

$$(\alpha x, \beta y) = \alpha\beta(x, y)$$

векторы αx и βy также ортогональны.

Теорема 8.2. *Любая ортогональная система линейно независима.*

$\triangleright$ Пусть a_1, a_2, ..., a_k — произвольная ортогональная система векторов. Это означает, что все векторы a_i ненулевые, т.е. $(a_i, a_i) \neq 0$, и попарно ортогональны, т.е. $(a_i, a_j) = 0$ при $i \neq j$. Чтобы доказать линейную независимость рассматриваемой системы векторов, приравняем к нулю произвольную линейную комбинацию

$$\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_k a_k = 0$$

и докажем, что из этого равенства вытекает равенство нулю всех коэффициентов α_i линейной комбинации. Умножим векторное равенство скалярно на произвольный вектор a_j рассматриваемой ортогональной системы. Получим:

$$\alpha_1 (a_1, a_j) + \alpha_2 (a_2, a_j) + \ldots + \alpha_k (a_k, a_j) = 0.$$

В силу соотношений $(a_i, a_j) = 0$, $i \neq j$, равенство упрощается:

$$\alpha_j (a_j, a_j) = 0.$$

Поскольку $(a_j, a_j) \neq 0$, то $\alpha_j = 0$. Так как номер $j = 1, 2, \ldots, k$ выбирался произвольно, заключаем, что $\alpha_1 = \alpha_2 = \ldots = \alpha_k = 0$. $\blacktriangleright$

Существуют ли ортогональные системы векторов в произвольном евклидовом пространстве? Ответ на этот естественный вопрос положительный. Существует специальная процедура, которая позволяет преобразовать произвольную линейно независимую систему из k векторов в ортогональную систему, также имеющую k векторов. Эту процедуру называют ***процессом ортогонализации***, и состоит она в следующем. Предположим, что система $a_1, a_2, \ldots, a_k$ линейно независима.

1. Полагаем $b_1 = a_1$.

2. Если векторы $b_1, b_2, \ldots, b_{i-1}$ ($i \leqslant k$) найдены, ищем ненулевой вектор $b_i = a_i + \beta_{i1}b_1 + \ldots \beta_{i,i-1}b_{i-1}$, выбирая коэффициенты $\beta_{i1}, \ldots, \beta_{i,i-1}$ так, что вектор b_i будет ортогонален каждому из векторов $b_1, \ldots, b_{i-1}$.

Описанная процедура состоит из k шагов, на каждом из которых определяется очередной вектор b_i. Ясно, что если условия выбора этих векторов будут обеспечены, мы получим ортогональную систему $b_1, b_2, \ldots, b_k$. Условий выбора два: ортогональность очередного вектора всем предыдущим и необращение в нуль очередного вектора. Если обеспечить первое условие, то второе будет выполняться в силу линейной независимости векторов $a_1, \ldots, a_k$. Действительно, если

$$b_i = a_i + \beta_{i1}b_1 + \ldots \beta_{i,i-1}b_{i-1} = 0,$$

то вектор a_i можно представить в виде линейной комбинации векторов $b_1, \ldots, b_{i-1}$. Но каждый вектор b_j, $j = 1, 2, i-1$, в свою очередь является линейной комбинацией векторов $a_1, \ldots, a_j$, поэтому вектор a_i есть линейная комбинация векторов $a_1, \ldots, a_{i-1}$. А такой вывод противоречит линейной независимости системы $a_1, \ldots, a_k$.

Итак, необходимо выбрать вектор

$$b_i = a_i + \beta_{i1}b_1 + \ldots \beta_{i,i-1}b_{i-1} \tag{8.10}$$

так, чтобы он был ортогонален всем предыдущим. Предположим, что мы уже выбрали векторы $b_1, \ldots, b_{i-1}$, удовлетворяющие этому условию. Это значит, что все эти векторы попарно ортогональны. Умножим равенство (8.10) скалярно на вектор b_j, $j = 1, 2, \ldots, i-1$. С учетом попарной ортогональности векторов $b_1, \ldots, b_{i-1}$ и условия ортогональности им вектора b_i получим:

$$0 = (a_i, b_j) + \beta_j (b_j, b_j).$$

Отсюда находим:

$$\beta_j = -\frac{(a_i, b_j)}{(b_j, b_j)} = -\frac{(a_i, b_j)}{|b_j|^2}.$$

Мы пришли к следующему. В процессе ортогонализации очередной вектор b_i нужно выбирать согласно формуле

$$b_i = a_i - \frac{(a_i, b_1)}{|b_1|^2} - \frac{(a_i, b_2)}{|b_2|^2} - \ldots - \frac{(a_i, b_{i-1})}{|b_{i-1}|^2}. \tag{8.11}$$

Процесс ортогонализации рассчитан на линейно независимые системы векторов. Но этот процесс можно модифицировать так, что станет возможным его применение и к линейно зависимым системам векторов. Если система $a_1, \ldots, a_k$ линейно зависима, то один из векторов a_i является линейной комбинацией предыдущих векторов $a_1, \ldots, a_{i-1}$. В результате процесса ортогонализации на i-м шаге получим нулевой вектор b_i. Мы опускаем этот вектор и начинаем следующий шаг. В результате мы придем к ортогональной системе векторов $b_1, \ldots, b_s$, но в этой системе будет меньше векторов, чем в исходной системе $a_1, \ldots, a_k$, т.е. $s < k$. Число s на самом деле есть ранг системы векторов $a_1, \ldots, a_k$.

Пример 8.5. Применяя процесс ортогонализации и нормирование векторов, ортонормировать систему векторов

$$a_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \qquad a_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \qquad a_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \qquad a_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix},$$

считая, что в четырехмерном евклидовом пространстве E_4 скалярное произведение определено формулой (8.5).

Решение. Положим $b_1 = a_1 = (1, 1, 0, 0)^{\mathrm{T}}$. В соответствии с формулой (8.10) находим:

$$b_2 = a_2 - \frac{(a_2, b_1)}{|b_1|^2} b_1 = a_2 - \frac{1}{2} b_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/2 \\ -1/2 \\ 1 \\ 0 \end{pmatrix}.$$

Далее находим b_3:

$$b_3 = a_3 - \frac{(a_3, b_1)}{|b_1|^2} b_1 - \frac{(a_3, b_2)}{|b_2|^2} b_2 =$$

$$= \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} - \frac{1}{3} \begin{pmatrix} 1/2 \\ -1/2 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/3 \\ -1/3 \\ -1/3 \\ 1 \end{pmatrix}.$$

Наконец, вычисляем b_4:

$$b_4 = a_4 - \frac{(a_4, b_1)}{|b_1|^2} b_1 - \frac{(a_4, b_2)}{|b_2|^2} b_2 - \frac{(a_4, b_3)}{|b_3|^2} b_3 =$$

$$= a_4 - 0 \cdot b_1 - 0 \cdot b_2 - \frac{3}{4} b_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} - \frac{3}{4} \begin{pmatrix} 1/3 \\ -1/3 \\ -1/3 \\ 1 \end{pmatrix} = \begin{pmatrix} -1/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{pmatrix}$$

Нормируя векторы b_1, b_2, b_3, b_4, придем к ортонормированной системе векторов

$$q_1 = \frac{b_1}{|b_1|} = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0 \right)^{\mathrm{T}},$$

$$q_2 = \frac{b_2}{|b_2|} = \left(\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, 0 \right)^{\mathrm{T}},$$

$$q_3 = \frac{b_3}{|b_3|} = \left(\frac{\sqrt{3}}{6}, -\frac{\sqrt{3}}{6}, -\frac{\sqrt{3}}{6}, \frac{\sqrt{3}}{2} \right)^{\mathrm{T}},$$

$$q_4 = \frac{b_4}{|b_4|} = \left(-\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2} \right)^{\mathrm{T}}.$$

Пример 8.6. В линейном пространстве многочленов степени не выше 3 ортонормировать систему векторов $f_1(x) = x$, $f_2(x) = x$, $f_3(x) = x^2$, $f_4(x) = x^3$, если скалярное произведение в этом линейном пространстве определено формулой

$$(f, g) = \int_{-1}^{1} f(x) g(x) \, dx.$$

Решение. Положим $\varphi_1 = f_1$ и вычислим φ_2 в соответствии с формулой (8.10):

$$\varphi_2(x) = f_2(x) - \frac{\int\limits_{-1}^{1} f_2(x)\varphi_1(x)\,dx}{\int\limits_{-1}^{1} \varphi_1^2(x)\,dx}\varphi_1(x) =$$

$$= f_2(x) - \frac{\int\limits_{-1}^{1} x\,dx}{\int\limits_{-1}^{1} dx}\varphi_1(x) = f_2(x) - 0\cdot\varphi_1(x) = x.$$

Далее по той же формуле вычисляем $\varphi_3(x)$:

$$\varphi_3(x) = f_3(x) - \beta_{31}\varphi_1(x) - \beta_{32}\varphi_2(x),$$

где

$$\beta_{31} = \frac{\int\limits_{-1}^{1} f_3(x)\varphi_1(x)\,dx}{\int\limits_{-1}^{1} \varphi_1^2(x)\,dx} = \frac{\int\limits_{-1}^{1} x^2\,dx}{\int\limits_{-1}^{1} dx} = \frac{1}{3},$$

$$\beta_{32} = \frac{\int\limits_{-1}^{1} f_3(x)\varphi_2(x)\,dx}{\int\limits_{-1}^{1} \varphi_2^2(x)\,dx} = \frac{\int\limits_{-1}^{1} x^3\,dx}{\int\limits_{-1}^{1} x^2 dx} = 0.$$

В результате

$$\varphi_3(x) = f_3(x) - \frac{1}{3}\varphi_1(x) = x^2 - \frac{1}{3}.$$

Наконец,

$$\varphi_4(x) = f_4(x) - \beta_{41}\varphi_1(x) - \beta_{42}\varphi_2(x) - \beta_{43}\varphi_3(x),$$

где

$$\beta_{41} = \frac{\int\limits_{-1}^{1} f_4(x)\,\varphi_1(x)\,dx}{\int\limits_{-1}^{1} \varphi_1^2(x)\,dx} = \frac{\int\limits_{-1}^{1} x^3\,dx}{\int\limits_{-1}^{1} dx} = 0,$$

$$\beta_{42} = \frac{\int\limits_{-1}^{1} f_4(x)\,\varphi_2(x)\,dx}{\int\limits_{-1}^{1} \varphi_2^2(x)\,dx} = \frac{\int\limits_{-1}^{1} x^4\,dx}{\int\limits_{-1}^{1} x^2 dx} = \frac{3}{5},$$

$$\beta_{43} = \frac{\int\limits_{-1}^{1} f_4(x)\,\varphi_3(x)\,dx}{\int\limits_{-1}^{1} \varphi_3^2(x)\,dx} = \frac{\int\limits_{-1}^{1} x^3(x^2-1/3)\,dx}{\int\limits_{-1}^{1} (x^2-1/3)^2 dx} = 0.$$

Таким образом,

$$\varphi_4(x) = f_4(x) - \frac{2}{5}\varphi_2(x) = x^3 - \frac{3}{5}x.$$

Нормируя полученные векторы φ_1, φ_2, φ_3, φ_4, придем к нормированной системе

$$\psi_1(x) = \frac{\varphi_1(x)}{\left(\int\limits_{-1}^{1} dx\right)^{1/2}} = \frac{1}{\sqrt{2}},$$

$$\psi_2(x) = \frac{\varphi_2(x)}{\left(\int\limits_{-1}^{1} x^2\,dx\right)^{1/2}} = \sqrt{\frac{3}{2}}x,$$

$$\psi_3(x) = \frac{\varphi_3(x)}{\left(\int\limits_{-1}^{1} (x^2-1/3)^2\,dx\right)^{1/2}} = \frac{3}{2}\sqrt{\frac{5}{2}}x^3 - \frac{1}{2}\sqrt{\frac{5}{2}},$$

$$\psi_4(x) = \frac{\varphi_3(x)}{\left(\int\limits_{-1}^{1} (x^3-3x/5)^2\,dx\right)^{1/2}} = \frac{5}{2}\sqrt{\frac{7}{2}}x^3 - \frac{3}{2}\sqrt{\frac{7}{2}}x.$$

О другом подходе к ортонормированным системам см. разд. 8.17.

8.3. Ортонормированные базисы

Базис $e_1, e_2, \ldots, e_n$ евклидова пространства E называют ***ортогональным базисом***, если его векторы попарно ортогональны. Если, кроме того, векторы этого базиса имеют единичную длину (т.е. нормированы), то он называется ***ортонормированным базисом***. В ортонормированном базисе $e_1, e_2, \ldots, e_n$ выполняются условия

$$(e_i, e_j) = \begin{cases} 0 & \text{при } i \neq j, \\ 1 & \text{при } i = j. \end{cases} \tag{8.12}$$

Теорема 8.3. *В любом конечномерном евклидовом пространстве E существуют ортогональные и ортонормированные базисы. При этом любой вектор в E входит в состав какого-либо ортогонального базиса, а любой единичный вектор — в состав какого-либо ортонормированного базиса.*

◁ Чтобы получить ортогональный базис, достаточно применить процесс ортогонализации к какой-либо максимальной линейно независимой системе векторов в E. Поскольку любой вектор евклидова пространства можно взять в качестве первого вектора максимальной линейно независимой системы векторов, заключаем, что этот вектор входит в какой-либо ортогональный базис, так как в процессе ортогонализации первый вектор исходной (ортогонализируемой) системы не изменяется. Если выбран единичный вектор, то он будет входить в состав ортогонального базиса. Нормировка этого базиса не изменяет единичные векторы. Значит, любой единичный вектор входит в состав того или иного ортонормированного базиса. ▶

Примеры построения ортонормированных базисов с помощью процесса ортогонализации приведены выше (см. примеры 8.5 и 8.6). Ортогональные (ортонормированные) базисы можно также строить, дополняя подходящими векторами данный вектор или данную ортогональную (ортонормированную) систему векторов.

Пример 8.7. В трехмерном арифметическом пространстве K_3 со скалярным произведением

$$(x, y) = x_1y_1 + x_2y_2 + x_3y_3, \quad x = (x_1,\ x_2,\ x_3)^{\mathrm{T}}, \quad y = (y_1,\ y_2,\ y_3)^{\mathrm{T}},$$

построить ортонормированный базис, содержащий вектор $e_1 = (1,\ 1,\ 1)^{\mathrm{T}}$.

Решение. Добавим к вектору e_1 вектор $e_2 = (y_1,\ y_2,\ y_3)^{\mathrm{T}}$, удовлетворяющий условию $(e_1, e_2) = 0$, которое в координатах имеет вид $y_1 + y_2 + y_3 = 0$. Одним из решений этого уравнения является вектор $e_2 = (2,\ -1,\ -1)^{\mathrm{T}}$.

Далее, к векторам e_1 и e_2 добавим вектор $e_3 = (z_1,\ z_2,\ z_3)^{\mathrm{T}}$, удовлетворяющий условиям $(e_3, e_1) = 0$, $(e_3, e_2) = 0$, которые в координатной форме имеют вид $z_1 + z_2 + z_3 = 0$, $2z_1 - z_2 - z_3 = 0$. Решая однородную систему из этих двух уравнений, получим, например, решение $e_3 = (0,\ -1,\ 1)^{\mathrm{T}}$. Система векторов e_1, e_2, e_3 является одним из ортогональных базисов в K_3. Нормируя эти векторы, получим в K_3 ортонормированный базис

$$f_1 = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad f_2 = \frac{1}{\sqrt{6}}\begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix}, \quad f_3 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}.$$

Теорема 8.4. *В любом ортонормированном базисе e n-мерного евклидова пространства E скалярное произведение векторов $x = (x_1,\ x_2,\ \ldots,\ x_n)^{\mathrm{T}}$ и $y = (y_1,\ y_2,\ \ldots,\ y_n)^{\mathrm{T}}$, заданных координатами в этом базисе, определяется формулой*

$$(x, y) = x^{\mathrm{T}}y = (x_1,\ x_2,\ \ldots,\ x_n)\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \\ = x_1y_1 + x_2y_2 + \ldots x_ny_n. \qquad (8.13)$$

Наоборот, если в базисе e скалярное произведение определяется формулой (8.13), то этот базис ортонормированный.

▷ Пусть базис $e = (e_1, e_2, \ldots, e_n)$ ортонормированный. Тогда выполняются равенства (8.12). Поэтому матрицей Грама базиса e является матрица

$$\Gamma = \begin{pmatrix} (e_1, e_1) & (e_1, e_2) & \ldots & (e_1, e_n) \\ (e_2, e_1) & (e_2, e_2) & \ldots & (e_2, e_n) \\ \ldots & \ldots & \ldots & \ldots \\ (e_n, e_1) & (e_n, e_2) & \ldots & (e_n, e_n) \end{pmatrix} = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} = E,$$

а скалярное произведение вычисляется по формуле

$$(x, y) = x^{\mathrm{T}} \Gamma y = x^{\mathrm{T}} y = x_1 y_1 + x_2 y_2 + \ldots x_n y_n.$$

Наоборот, если скалярное произведение в данном базисе e евклидова пространства вычисляется по формуле (8.13), то легко убедиться в том, что скалярные произведения базисных векторов удовлетворяют условию (8.12). А это означает, что базис e ортонормированный. ▶

Следствие 8.1. В n-мерном линейном пространстве X с заданным базисом e можно задать скалярное произведение так, что заданный базис будет ортонормированным.

▷ Достаточно определить скалярное произведение через координаты векторов в базисе e формулой (8.13). Тогда по доказанной теореме базис e будет ортонормированным. ▶

8.4. Ортогональные матрицы

Квадратная матрица Q, для которой транспонированная матрица Q^{T} совпадает с обратной матрицей Q^{-1}, называется ***ортогональной матрицей***. Квадратная матрица Q является ортогональной, если и только если $Q^{\mathrm{T}}Q = QQ^{\mathrm{T}} = E$.

Приведем основные свойства ортогональных матриц.

1. ***Квадратная матрица Q ортогональная тогда и только тогда, когда сумма квадратов всех элементов любого ее столбца (строки) равна единице, а***

сумма попарных произведений элементов двух любых столбцов (строк) равна нулю.

Действительно, диагональные элементы матрицы $Q^{\mathrm{T}}Q$ равны сумме квадратов элементов соответствующих столбцов матрицы Q, а недиагональные элементы равны сумме попарных произведений элементов двух столбцов. Поэтому сформулированное утверждение означает, что $Q^{\mathrm{T}}Q = E$. Утверждение для строк вытекает из рассмотрения произведения QQ^{T}.

2. ***Определитель ортогональной матрицы равен 1 или −1.***

Действительно,

$$|Q|^2 = |Q|\,|Q| = |Q^{\mathrm{T}}|\,|Q| = |Q^{\mathrm{T}}Q| = |E| = 1.$$

Поэтому $Q = \pm 1$.

3. ***Матрица, обратная к ортогональной матрице, тоже ортогональная.***

Действительно,

$$(Q^{-1})^{\mathrm{T}} = (Q^{\mathrm{T}})^{\mathrm{T}} = Q = (Q^{-1})^{-1},$$

т.е. транспонированная к матрице Q^{-1} совпадает с обратной к этой матрице. А это по определению и означает, что Q^{-1} ортогональная.

4. ***Произведение ортогональных матриц является ортогональной матрицей.***

Действительно, для двух ортогональных матриц Q и R имеем:

$$(QR)^{\mathrm{T}} = R^{\mathrm{T}}Q^{\mathrm{T}} = R^{-1}Q^{-1} = (QR)^{-1},$$

что и означает ортогональность матрицы QR. Доказательство для произведения большего числа ортогональных матриц можно провести по методу математической индукции.

Роль ортогональных матриц в теории евклидовых пространств проясняет следующая теорема.

Теорема 8.5. *В евклидовом пространстве матрица перехода от одного ортонормированного базиса к другому ортонормированному базису является ортогональной. Если* *матрица перехода от ортонормированного базиса ко второму базису является ортогональной, то этот второй базис тоже ортонормированный.*

$\triangleright$ Пусть e — ортонормированный базис, а T — матрица перехода от базиса e к базису e'. Тогда матрица Грама Γ' для базиса e' равна: $\Gamma' = T^{\mathrm{T}}\Gamma T = T^{\mathrm{T}}T$, где $\Gamma = E$ — матрица Грама для базиса e (см. разд. 8.1). Если базис e' ортонормированный, то $\Gamma' = E$ и $T^{\mathrm{T}}T = E$, т.е. матрица T ортогональная. Если T ортогональная, то $T^{\mathrm{T}}T = E$, матрица Грама $\Gamma' = T^{\mathrm{T}}T$ базиса e' оказывается единичной, а сам базис — ортонормированным. ▶

Пример 8.8. В двумерном евклидовом пространстве E_2 рассмотрим базис e', который получается поворотом ортонормированного базиса e на угол α. Тогда $[e'_1]_e = (\cos\alpha,\ \sin\alpha)^{\mathrm{T}}$ и $[e'_2]_e = (-\sin\alpha,\ \cos\alpha)^{\mathrm{T}}$. Следовательно, матрица перехода T от базиса к базису e' имеет вид:

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

Поворот векторов сохраняет их длины и углы между ними. Поэтому базис e' ортонормированный. Непосредственной проверкой легко убедиться, что

$$T^{\mathrm{T}}T = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}\begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = E,$$

т.е. матрица T ортогональная.

8.5. Ортогональное дополнение. Ортогональная проекция вектора на подпространство

Пусть E — евклидово пространство, а L — его подпространство. Множество $L^{\perp}$ векторов в E, ортогональных к каждому вектору подпространства L, называют ***ортогональным дополнением*** к подпространству L.

Теорема 8.6. *Ортогональное дополнение $L^{\perp}$ к подпространству L евклидова пространства E является подпространством в E.*

$\triangleright$ Пусть $y_1, y_2 \in L^{\perp}$. Тогда для любого вектора $x \in L$ имеем: $(x, y_1) = 0$ и $(x, y_2) = 0$. Следовательно,

$$(x, y_1 + y_2) = (x, y_1) + (x, y_2) = 0,$$

т.е. вектор $y_1 + y_2$ ортогонален любому вектору $x \in L$. Это означает, что $y_1 + y_2 \in L^{\perp}$. Мы доказали, что сумма любых двух векторов множества $L^{\perp}$ принадлежит $L^{\perp}$. Аналогично для любого действительного числа λ и любого $x \in L$ имеем:

$$(x, \lambda y_1) = \lambda (x, y_1) = \lambda \cdot 0 = 0,$$

т.е. вектор λy_1 ортогонален любому вектору $x \in L$, а значит, принадлежит $L^{\perp}$. Таким образом, множество $L^{\perp}$ замкнуто относительно сложения векторов и умножения векторов на числа и, следовательно, является подпространством. $\blacktriangleright$

Теорема 8.7. *Конечномерное евклидово пространство E является прямой суммой любого своего подпространства L и его ортогонального дополнения $L^{\perp}$, т.е. ортогональное дополнение к подпространству является прямым дополнением.*

$\triangleright$ Покажем, что подпространства L и $L^{\perp}$ составляют прямую сумму. Для этого достаточно доказать, что подпространство $L \cap L^{\perp}$ является нулевым (т.е. содержит лишь нулевой вектор). Если $x \in L \cap L^{\perp}$, то вектор x, как принадлежащий $L^{\perp}$, ортогонален каждому вектору из L, в том числе и самому себе, поскольку $x \in L$. Но единственным вектором, ортогональным самому себе, т.е. вектором с нулевым скалярным квадратом, является нулевой вектор. Следовательно, $L \cap L^{\perp} = \{0\}$.

Покажем, что подпространство $H = L \oplus L^{\perp}$ совпадает с E. Выберем в H какой-либо ортонормированный базис $a_1, a_2, \ldots, a_k$ и произвольный вектор $x \in E$. Нетрудно убедиться в том, что вектор

$$x_0 = x - (x, a_1)\, a_1 - \ldots - (x, a_k)\, a_k,$$

получающийся в процессе ортогонализации системы $a_1, \ldots, a_k, x$, ортогонален каждому вектору a_i, $i = 1, \ldots, k$. Действительно,

$$(x_0, a_i) = (x, a_i) - (x, a_1)(a_1, a_i) - \ldots - (x, a_k)(a_k, a_i) = \\ = (x, a_i) - (x, a_i)(a_i, a_i) = 0,$$

поскольку система $a_1, \ldots, a_k$ ортонормированная. Так как вектор x_0 ортогонален каждому вектору базиса в подпространстве H, то он ортогонален всему подпространству и, в частности, подпространству L. Значит, $x_0 \in L^\perp$ и $x \in L \cap L^\perp$. Но уже доказано, что в пересечении L и $L^\perp$ только один вектор — нулевой. Тем самым мы доказали, что $x_0 = 0$ и что

$$x = (x, a_1)a_1 + \ldots + (x, a_k).$$

Последнее равенство показывает, что $x \in H$. Но вектор x выбирался в E произвольно. Следовательно, $H = E$ и $E = L \oplus L^\perp$. ▶

Следствие 8.2. Если подпространство L в n-мерном евклидовом пространстве E имеет размерность k, то его ортогональное дополнение $L^\perp$ имеет размерность $n - k$.

▷ Это очевидное следствие теоремы о размерности суммы подпространств (см. теорему 4.23). ▶

Следствие 8.3. Ортогональным дополнением к подпространству $L^\perp$ является подпространство L.

▷ Так как каждый вектор L ортогонален каждому вектору $L^\perp$, то подпространство L содержится в $(L^\perp)^\perp$. В силу предыдущего следствия подпространства L и $(L^\perp)^\perp$ имеют одинаковую размерность, а потому совпадают. ▶

Следствие 8.4. Если L — подпространство в евклидовом пространстве E, то любой вектор $x \in E$ имеет разложение

$$x = x_0 + x^\perp,$$

где $x_0 \in L$, $x^\perp \in L^\perp$. Такое разложение единственно.

$\triangleright$ Это утверждение — фактически расшифровка утверждения, что $E = L \oplus L^{\perp}$. ▶

Пример 8.9. В четырехмерном пространстве E_4 скалярное произведение в заданном базисе определено формулой (8.5). Построить ортогональное дополнение $L^{\perp}$ для подпространства $L = \langle a_1, a_2 \rangle$, где $a_1 = (1, 1, 1, 1)^{\mathrm{T}}$, $a_2 = (1, -1, 1, 1)^{\mathrm{T}}$.

Решение. Векторы a_1 и a_2 составляют базис в L. Дополним эту систему до базиса в E_4 векторами b_1 и b_2, удовлетворяющими условиям

$$(a_1, b_i) = 0, \quad (a_2, b_i) = 0,$$

и положим $L_1 = \langle b_1, b_2 \rangle$. Векторы b_1, b_2 являются решениями системы из двух уравнений $(a_1, x) = 0$, $(a_2, x) = 0$, и в качестве их можно взять любую фундаментальную систему решений, например, $b_1 = (-1, 0, 1, 0)^{\mathrm{T}}$, $b_2 = (-1, 0, 0, 1)^{\mathrm{T}}$. Из выбора векторов b_1 и b_2 следует, что они составляют базис в $L^{\perp}$, т.е. $L_1 = L^{\perp}$.

Пусть L — подпространство евклидова пространства E. Каждый вектор $y \in E$ может быть единственным способом представлен в виде

$$y = y_0 + y^{\perp}, \tag{8.14}$$

где $y_0 \in L$, а вектор $y^{\perp}$ ортогонален к каждому вектору из L, т.е. $y^{\perp} \in L^{\perp}$. Вектор y_0 называют ***ортогональной проекцией вектора y на пространство L*** и обозначают $\text{пр}_L\, y$, а вектор $y^{\perp}$ называют ***ортогональной составляющей вектора y***. Очевидно, что если $y \in L$, то $\text{пр}_L\, y = y$, и, наоборот, если $\text{пр}_L\, y = y$, то $y \in L$. Отметим, что в представлении (8.14) вектор $y^{\perp}$ есть ортогональная проекция вектора y на подпространство $L^{\perp}$.

Ортогональная проекция вектора y на подпространство L является частным случаем проекции вектора на подпространство параллельно подпространству L_2, являющемуся прямым дополнением к L (см. разд. 4.11). В случае ортогональной проекции $L_2 = L^{\perp}$.

На практике при отыскании ортогональной проекции вектора x на подпространство $L = \langle a_1,\ a_2,\ \ldots,\ a_k \rangle$, где система a_1, a_2, ..., a_k линейно независима, поступают следующим образом. В разложении

$$x = x_0 + x^{\perp}$$

вектора x на ортогональную проекцию $x_0 = \text{пр}_L x$ и ортогональную составляющую $x^{\perp}$ вектор x_0 можно представить в виде линейной комбинации

$$x_0 = \alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_k a_k. \tag{8.15}$$

Тогда равенство $x = x_0 + x^{\perp}$ принимает вид:

$$x = \alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_k a_k + x^{\perp}. \tag{8.16}$$

Для отыскания коэффициентов $\alpha_1, \alpha_2, \ldots \alpha_k$ достаточно умножить равенство (8.16) скалярно на векторы a_1, a_2, ..., a_k. Учитывая, что $(a_1, x^{\perp}) = (a_2, x^{\perp}) = \ldots = (a_k, x^{\perp}) = 0$, получаем систему линейных уравнений

$$\begin{cases} (a_1, x) = \alpha_1 (a_1, a_1) + \ldots + \alpha_k (a_1, a_k), \\ (a_2, x) = \alpha_1 (a_2, a_1) + \ldots + \alpha_k (a_2, a_k), \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ (a_k, x) = \alpha_1 (a_k, a_1) + \ldots + \alpha_k (a_k, a_k) \end{cases} \tag{8.17}$$

относительно неизвестных α_1, α_2, $\ldots \alpha_k$. Из этой системы находят коэффициенты α_1, α_2, ..., α_k. В матричной форме равенство (8.15) и система (8.17) записываются в виде

$$x_0 = Aa, \tag{8.18}$$

$$A^{\text{т}} x = A^{\text{т}} Aa, \tag{8.19}$$

где $A = (a_1,\ a_2,\ \ldots,\ a_k)$ — матрица, для которой столбцами являются столбцы координат векторов a_1, a_2, ..., a_k; $a = (\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_k)^{\text{т}}$ — столбец высоты k. В уравнении (8.19) матрица $A^{\text{т}} A$ невырожденная, так как она представляет собой матрицу Грама линейно независимой системы векторов a_1, a_2, ..., a_k (см. теорему 8.1).

Поскольку матрица $A^{\mathrm{T}}A$ невырожденная, из уравнения (8.19) однозначно определяется столбец a:

$$a = (A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}x.$$

Учитывая равенство (8.18), заключаем, что

$$x_0 = \text{пр}_L x = Aa = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}x. \tag{8.20}$$

Пример 8.10. Для вектора $x = (3,\ 6,\ 0)^{\mathrm{T}}$ найти ортогональную проекцию x_0 на подпространство $L = \langle a_1,\ a_2\rangle$ и ортогональную составляющую $x^{\perp}$, если $a_1 = (1,\ -1,\ 0)^{\mathrm{T}}$, $a_2 = (-1,\ 2,\ 1)^{\mathrm{T}}$.

Решение. Отметим, что векторы a_1 и a_2 линейно независимы. Запишем $x_0 = \text{пр}_L x$ в виде $x_0 = \alpha_1 a_1 + \alpha_2 a_2$. Коэффициенты α_1 и α_2 можно найти, решив систему (8.17), которая в данном случае имеет вид:

$$\begin{cases} (a_1, x) = \alpha_1 (a_1, a_1) + \alpha_2 (a_1, a_2), \\ (a_2, x) = \alpha_1 (a_2, a_1) + \alpha_2 (a_2, a_2). \end{cases}$$

Вычислим все скалярные произведения. В результате получим:

$$\begin{cases} -3 = 2\alpha_1 - 3\alpha_2, \\ 9 = -3\alpha_1 + 6\alpha_2. \end{cases}$$

Решая систему относительно неизвестных α_1 и α_2, находим: $\alpha_1 = \alpha_2 = 3$. Таким образом, $\text{пр}_L x = 3a_1 + 3a_2 = (0,\ 3,\ 3)^{\mathrm{T}}$ и $x^{\perp} = x - \text{пр}_L x = (-1,\ -1,\ -2)^{\mathrm{T}}$. Можно было также воспользоваться формулой (8.20). Вычислив

$$A^{\mathrm{T}}A = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ -1 & 6 \end{pmatrix},$$

получаем:

$$\text{пр}_L x = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}x = \begin{pmatrix} 1 & 1 \\ -1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -1 & 6 \end{pmatrix}^{-1} \times$$

$$\times \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 6 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \\ 3 \end{pmatrix}.$$

Заметим, что если $e = (e_1, e_2, \ldots, e_n)$ — ортонормированный базис в евклидовом пространстве E, а подпространство L является линейной оболочкой части базисных векторов, например $L = \langle e_1, e_2, \ldots, e_k \rangle$, то для любого вектора

$$x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n$$

ортогональная проекция $\text{пр}_L x$ совпадает с суммой слагаемых в разложении x по базису, соответствующих векторам, порождающим L, а ортогональная проекция — с суммой всех остальных слагаемых, т.е.

$$\text{пр}_L x = x_1 e_1 + x_2 e_2 + \ldots + x_k e_k,$$

$$x^{\perp} = x_{k+1} e_{k+1} + x_{k+2} e_{k+2} + \ldots + x_n e_n.$$

Например, для вектора $x = (1,\ 2,\ 3,\ 4,\ 5)^{\text{т}}$ проекция на подпространство $L = \langle e_1, e_2, e_3 \rangle$ равна $x_0 = (1,\ 2,\ 3,\ 0,\ 0)^{\text{т}}$.

8.6. Изоморфизм евклидовых пространств

Пусть E и E' — евклидовы пространства со скалярными произведениями $(x, y)_E$ и $(x', y')_{E'}$ соответственно. Евклидовы пространства E и E' называют ***изоморфными***, если между их векторами существует взаимно однозначное соответствие, которое:

1) осуществляет изоморфизм E и E' как линейных пространств;

2) сохраняет скалярное произведение, т.е. если векторам x и y в E соответствуют векторы x' и y' в E', то $(x, y)_E = (x', y')_{E'}$.

Теорема 8.8. *Два конечномерных евклидовых пространства изоморфны тогда и только тогда, когда они имеют одинаковую размерность.*

▷ Если конечномерные евклидовы пространства E и E' изоморфны, то они изоморфны как линейные пространства. Следовательно, они имеют одинаковую размерность (см. теорему 4.11).

Пусть евклидовы пространства E и E' имеют одинаковую размерность n. Выберем в них ортонормированные базисы

$$e = (e_1,\, e_2,\, \ldots,\, e_n) \quad \text{и} \quad e' = \left(e'_1,\, e'_2,\, \ldots,\, e'_n\right).$$

Между векторами двух евклидовых пространств установим взаимно однозначное соответствие, при котором вектору $x = x_1e_1 + \ldots x_ne_n$ в E соответствует вектор $x' = x_1e'_1 + \ldots x_ne'_n$ в E'. Из доказательства теоремы 4.11 следует, что это соответствие устанавливает изоморфизм линейных пространств E и E', и нам остается доказать, что при рассматриваемом соответствии сохраняется скалярное произведение. Произвольным векторам $x = x_1e_1 + \ldots x_ne_n$ и $y = y_1e_1 + \ldots y_ne_n$ в E соответствуют векторы $x' = x_1e'_1 + \ldots x_ne'_n$ и $y' = y_1e'_1 + \ldots y_ne'_n$ в E. Для этих векторов имеем:

$$(x, y)_E = x_1y_1 + x_2y_2 + \ldots + x_ny_n = \left(x', y'\right)_{E'}.$$

Тем самым доказано, что рассматриваемое соответствие сохраняет скалярное произведение. ▶

Если два евклидовых пространства изоморфны, то они неразличимы по своей внутренней структуре. Среди всех n-мерных евклидовых пространств можно выделить арифметическое пространство столбцов высоты n со стандартным скалярным произведением

$$(x, y) = x_1y_1 + x_2y_2 + \ldots + x_ny_n,$$

$$x = (x_1,\, x_2,\, \ldots,\, x_n)^{\mathrm{T}}, \quad y = (y_1,\, y_2,\, \ldots,\, y_n)^{\mathrm{T}}.$$

Это пространство удобно рассматривать как своего рода эталон, поскольку, с одной стороны, его элементы имеют простую числовую природу, а с другой стороны, ему изоморфно любое другое n-мерное евклидово пространство.

8.7. Понятие об унитарном пространстве

Пусть дано комплексное линейное пространство X. Говорят, что в X определена операция ***скалярного умножения векторов***, если любой паре векторов x и y из X_n поставлено в соответствие комплексное число, называемое ***скалярным произведением векторов*** x и y и обозначаемое символом (x, y), и если для любых $x, y, z \in X$ и любого комплексного числа α выполняются следующие ***аксиомы скалярного произведения***.

1. $(x, y) = \overline{(y, x)}$.
2. $(x + y, z) = (x, z) + (y, z)$.
3. $(\alpha x, y) = \alpha (x, y)$.
4. $(x, x) > 0$ при $x \neq 0$ и $(x, x) = 0$ при $x = 0$.

Например, в комплексном арифметическом пространстве, элементами которого являются столбцы высоты n с комплексными компонентами, скалярное произведение можно ввести формулой

$$(x, y) = x^{\mathrm{T}} \overline{y} = x_1 \overline{y_1} + x_2 \overline{y_2} + \ldots + x_n \overline{y_n},$$

где $x = (x_1,\ x_2,\ \ldots,\ x_n)^{\mathrm{T}}$, $y = (x_1,\ x_2,\ \ldots,\ x_n)^{\mathrm{T}}$. Нетрудно проверить, что при таком задании скалярного произведения выполняются все четыре аксиомы.

Комплексное n-мерное линейное пространство, в котором определена операция скалярного умножения векторов, называют ***унитарным пространством***.

Из аксиом 1–3 скалярного произведения следует, что

$$(x, \alpha y) = \overline{(\alpha y, x)} = \overline{\alpha (y, x)} = \overline{\alpha}\overline{(y, x)} = \overline{\alpha} (x, y), \tag{8.21}$$

$$(x, y + z) = (x, y) + (x, z), \tag{8.22}$$

$$\left(\sum_{i=1}^{k} \alpha_i a_i,\ \sum_{j=1}^{l} \beta_j b_j \right) = \sum_{i=1}^{k} \sum_{j=1}^{l} \alpha_i \overline{\beta_j} (a_i, b_j). \tag{8.23}$$

Пусть в унитарном пространстве U задан базис $e_1, e_2, \ldots, e_n$. Любые векторы x и y имеют в этом базисе разложения

$$x = \sum_{i=1}^{n} x_i e_i, \quad y = \sum_{j=1}^{n} y_j e_j,$$

и формула (8.23) для векторов x и y дает:

$$(x, y) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i \overline{y_j}\,(e_i, e_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i \overline{y_j} g_{ij}, \tag{8.24}$$

где $g_{ij} = (e_i, e_j)$. Формулу (8.24) можно представить в матричной записи

$$(x, y) = x^{\mathrm{T}} \Gamma \overline{y}, \tag{8.25}$$

где $x = (x_1, x_2, \ldots, x_n)^{\mathrm{T}}$, $y = (y_1, y_2, \ldots, y_n)^{\mathrm{T}}$, а

$$\Gamma = \begin{pmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ g_{n1} & g_{n2} & \cdots & g_{nn} \end{pmatrix} = \begin{pmatrix} (e_1, e_1) & (e_1, e_2) & \ldots & (e_1, e_n) \\ (e_2, e_1) & (e_2, e_2) & \ldots & (e_2, e_n) \\ \cdot & \cdot & \cdot & \cdot \\ (e_n, e_1) & (e_n, e_2) & \ldots & (e_n, e_n) \end{pmatrix}$$

— матрица Грама базиса $e_1, e_2, \ldots, e_n$.

Поскольку $(e_i, e_j) = \overline{(e_j, e_i)}$, матрица Грама Γ удовлетворяет условию

$$\Gamma = \overline{\Gamma}^{\mathrm{T}} = \Gamma^*. \tag{8.26}$$

Напомним, что знак $*$ означает транспонирование матрицы последующей заменой в ней элементов на комплексно сопряженные.

Матрицу A^* называют ***сопряженной к матрице*** A. Если $A = A^*$, то A называют ***эрмитовой матрицей***. Так, в силу условия (8.26) матрица Грама является эрмитовой. Если матрица A действительная, то $A^* = A^{\mathrm{T}}$.

В унитарном пространстве, как и в евклидовом, длину вектора определяют формулой

$$|x| = \sqrt{(x, x)}. \tag{8.27}$$

Понятие „угол“ между векторами в унитарном пространстве, как правило, не вводят. Рассматривают лишь случай ортогональности векторов. При этом, как и в евклидовом пространстве, ортогональными считают векторы x и y, удовлетворяющие условию $(x, y) = 0$.

Процесс ортогонализации системы векторов, понятия „ортогональный базис“ и „ортонормированный базис“, „ортогональное дополнение“, „ортогональная проекция вектора на подпространство“ и вообще многие факты евклидова пространства распространяются на унитарное пространство без изменения определений и общих схем рассуждений. Однако каждый раз следует быть внимательным при применении скалярного произведения, так как в унитарном пространстве скалярное произведение существенно отличается от скалярного произведения в евклидовом пространстве. В унитарном пространстве в ортонормированном базисе для векторов

$$x = (x_1,\ x_2,\ \ldots,\ x_n)^{\mathrm{T}} \quad \text{и} \quad y = (y_1,\ y_2,\ \ldots,\ y_n)^{\mathrm{T}}$$

формула (8.24) принимает вид:

$$(x, y) = x^{\mathrm{T}}\overline{y} = y^{*}x = x_1\overline{y_1} + x_2\overline{y_2} + \ldots + x_n\overline{y_n}, \tag{8.28}$$

а для скалярного квадрата она превращается в формулу

$$(x, x) = x^{\mathrm{T}}\overline{x} = x^{*}x = x_1\overline{x_1} + x_2\overline{x_2} + \ldots + x_n\overline{x_n} =$$
$$= |x_1|^2 + |x_2|^2 + \ldots + |x_n|^2. \tag{8.29}$$

Эти формулы постоянно применяются при решении задач в унитарном пространстве.

Пример 8.11. Ортонормировать систему векторов

$$a_1 = (1,\ i,\ i)^{\mathrm{T}}, \quad a_2 = (i,\ i,\ i)^{\mathrm{T}}, \quad a_3 = (i,\ 0,\ i)^{\mathrm{T}},$$

считая, что векторы заданы координатами в ортонормированном базисе.

Решение. Сначала проведем процесс ортогонализации данной системы векторов. Положим $b_1 = a_1$, $b_2 = \alpha_1 b_1 + a_2$ и найдем α_1 из условия $(b_2, b_1) = 0$. Так как

$$(b_2, b_1) = (\alpha_1 b_1 + a_2, b_1) = \alpha_1 (b_1, b_1) + (a_2, b_1),$$

то из условия $(b_2, b_1) = 0$ находим, что

$$\alpha_1 = -\frac{(a_2, b_1)}{(b_1, b_1)} = -\frac{i\cdot\bar{1} + i\cdot\bar{i} + i\cdot\bar{i}}{|1|^2 + |i|^2 + |i|^2} = \frac{-2-i}{3}.$$

Следовательно,

$$b_2 = \frac{-2-i}{3}\begin{pmatrix} 1 \\ i \\ i \end{pmatrix} + \begin{pmatrix} i \\ i \\ i \end{pmatrix} = \frac{1}{3}\begin{pmatrix} -2+2i \\ 1+i \\ 1+i \end{pmatrix}.$$

Аналогично находим: $b_3 = a_3 - \beta_1 b_1 - \beta_2 b_2$, где

$$\beta_1 = -\frac{(a_3, b_1)}{(b_1, b_1)} = -\frac{i\cdot\bar{1} + 0\cdot\bar{i} + i\cdot\bar{i}}{|1|^2 + |i|^2 + |i|^2} = \frac{-1-i}{3},$$

$$\beta_2 = -\frac{(a_3, b_2)}{(b_2, b_2)} = -\frac{i\cdot\frac{-2-2i}{3} + i\cdot\frac{1-i}{3}}{\left|\frac{-2+2i}{3}\right|^2 + \left|\frac{1+i}{3}\right|^2 + \left|\frac{1+i}{3}\right|^2} = \frac{-3+i}{4}.$$

Поэтому

$$b_3 = \frac{-1-i}{3}\begin{pmatrix} 1 \\ i \\ i \end{pmatrix} + \frac{-3+i}{12}\begin{pmatrix} -2+2i \\ 1+i \\ 1+i \end{pmatrix} + \begin{pmatrix} i \\ 0 \\ i \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 0 \\ -i \\ i \end{pmatrix}.$$

Система векторов b_1, b_2, b_3 ортогональная. Чтобы получить ортонормированную систему, нормируем каждый вектор этой системы:

$$b_1^0=\frac{b_1}{|b_1|}=\frac{b_1}{\sqrt{(b_1,b_1)}}=\frac{b_1}{\sqrt{|1|^2+|i|^2+|i|^2}}=\frac{1}{\sqrt{3}}b_1=\frac{1}{\sqrt{3}}(1,\ i,\ i)^{\mathrm{T}},$$

$$b_2^0=\frac{b_2}{|b_2|}=\frac{b_2}{\sqrt{(b_2,b_2)}}=\frac{b_2}{\sqrt{\left|\frac{-2+2i}{3}\right|^2+\left|\frac{1+i}{3}\right|^2+\left|\frac{1+i}{3}\right|^2}}=$$

$$=\frac{\sqrt{3}}{2}b_2=\frac{1}{2\sqrt{3}}(-2+2i,\ 1+i,\ 1+i)^{\mathrm{T}},$$

$$b_3^0=\frac{b_3}{|b_3|}=\frac{b_3}{\sqrt{(b_3,b_3)}}=\frac{b_3}{\sqrt{\left|\frac{-i}{2}\right|^2+\left|\frac{i}{2}\right|^2}}=\sqrt{2}b_3=\frac{1}{\sqrt{2}}(0,\ -i,\ i)^{\mathrm{T}}.$$

П р и м е р 8.12. Убедиться, что система векторов

$$a_1=(4+3i,\ 4+3i,\ 2)^{\mathrm{T}},\qquad a_2=(4-3i,\ -4+3i,\ 0)^{\mathrm{T}}$$

ортогональная и дополнить ее до ортогонального базиса пространства U_3, считая, что векторы a_1, a_2 заданы координатами в ортонормированном базисе.

Р е ш е н и е. Векторы a_1, a_2 ортогональны, так как

$$(a_1,a_2)=(4+3i)(4+3i)+(4+3i)(-4-3i)+2\cdot 0=0.$$

К системе векторов a_1, a_2 добавим вектор $x=(x_1,\ x_2,\ x_3)^{\mathrm{T}}$, удовлетворяющий условиям

$$(x,a_1)=x_1(4-3i)+x_2(4-3i)+2\cdot x_3=0,$$
$$(x,a_2)=x_1(4+3i)-x_2(4+3i)=0.$$

Первое уравнение этой системы умножим на $4+3i$, второе — на $4-3i$. Тогда получим систему уравнений

$$\begin{cases}25x_1+25x_2+2(4+3i)x_3=0,\\ 25x_1-25x_2=0.\end{cases}$$

Из второго уравнения системы получаем: $x_1 = x_2$. Если сложить первое уравнение со вторым, то придем к уравнению

$$50x_1 + 2(4+3i)x_3 = 0,$$

из которого находим:

$$x_1 = -\frac{4+3i}{25}x_3.$$

Выберем $x_3 = -25$. В результате получим: $x_1 = x_2 = 4+3i$ и $x = (4+3i,\ 4+3i,\ -25)^{\mathrm{T}}$. Итак, одним из ортогональных базисов пространства U_3, содержащим векторы a_1, a_2, является базис, состоящий из векторов a_1, a_2 и x.

Квадратную матрицу Q, обладающую свойством $Q^* = Q^{-1}$, называют ***унитарной матрицей***. Квадратная матрица Q унитарная, если и только если выполнены условия

$$Q^*Q = QQ^* = E.$$

Унитарные матрицы являются комплексным аналогом ортогональных матриц и обладают похожими свойствами. Так, столбцы (строки) унитарной матрицы составляют ортонормированную систему, если их рассматривать как элементы комплексного арифметического пространства. Произведение унитарных матриц является унитарной матрицей. Унитарная матрица невырожденная и имеет обратную матрицу, также являющуюся унитарной. Унитарные матрицы, и только они, являются матрицами перехода от ортонормированного базиса унитарного пространства к ортонормированному. Отметим, что действительная унитарная матрица является ортогональной.

8.8. Сопряженные операторы в евклидовом пространстве

Все сказанное в гл. 5 о линейных операторах в действительных линейных пространствах сохраняет силу и для линейных операторов в евклидовых пространствах. В то же время наличие в евклидовых пространствах скалярного произведения векторов позволяет выделить важные классы линейных операторов. Обычно здесь рассматривают сопряженные, симметрические (самосопряженные) и ортогональные операторы. Рассмотрим такие операторы.

Линейный оператор φ^*, действующий в евклидовом пространстве E, называют ***сопряженным с оператором*** φ, если для любых векторов $x, y \in E$ выполняется равенство

$$(\varphi x, y) = (x, \underset{A^{\mathrm{T}}}{\varphi^*} y). \qquad (8.30)$$

Замечание. Для линейного оператора φ, действующего из евклидова пространства X в евклидово пространство Y, сопряженным называют линейный оператор φ^*, действующий из Y в X, если для любых векторов $x \in X$ и $y \in Y$ выполняется равенство (8.30). Такие сопряженные операторы обладают в основном теми же свойствами, что и сопряженные операторы, действующие в евклидовом пространстве E.

Теорема 8.9. *Для любого линейного оператора φ, действующего в евклидовом пространстве E, существует, и притом единственный, сопряженный оператор φ^*.*

$\triangleright$ В евклидовом пространстве E выберем ортонормированный базис e. Пусть оператор φ в этом базисе имеет матрицу A. Рассмотрим оператор φ_1 с матрицей A^{T} в базисе e. Тогда

$$(\varphi x, y) = (Ax)^{\mathrm{T}} y = x^{\mathrm{T}} A^{\mathrm{T}} y \quad \text{и} \quad (x, \varphi_1 y) = x^{\mathrm{T}} A^{\mathrm{T}} y,$$

где через x и y обозначены векторы евклидова пространства E и их столбцы координат в базисе e. Из записанных равенств видно, что для операторов φ и φ_1 выполняется равенство $(\varphi x, y) = (x, \varphi_1 y)$, каковы бы ни были векторы $x, y \in E$. Это означает, что оператор φ_1 является сопряженным с оператором φ.

Предположим, что оба линейных оператора φ_1, φ_2 являются сопряженными с оператором φ. Тогда $(x, \varphi_1 y) = (\varphi x, y) = (x, \varphi_2 y)$ и

$$(x, (\varphi_1 - \varphi_2) y) = (x, \varphi_1 y) - (x, \varphi_2 y) = 0, \quad x, y \in E.$$

Положив $x = (\varphi_1 - \varphi_2)y$, где $y \in E$ выбран произвольно, заключаем, что

$$((\varphi_1 - \varphi_2)y,\ (\varphi_1 - \varphi_2)y) = 0$$

и $(\varphi_1 - \varphi_2)y = 0$, так как из равенства нулю скалярного квадрата следует равенство нулю вектора. Итак, для любого вектора $y \in E$ выполняется равенство $(\varphi_1 - \varphi_2)y = 0$. Это значит, что линейный оператор $\varphi_1 - \varphi_2$ нулевой, а линейные операторы φ_1 и φ_2 совпадают. ▶

Поскольку равенство $(\varphi x, y) = (x, \varphi^* y)$ равносильно равенству $(\varphi^* y, x) = (y, \varphi x)$, то оператор φ является сопряженным с оператором φ^*, т.е. $(\varphi^*)^* = \varphi$. Это свойство позволяет говорить о ***взаимно сопряженных операторах***, так как отношение „оператор — сопряженный оператор" оказывается симметричным.

Теорема 8.10. *Матрицы A и A_1 операторов φ и φ^* в произвольном базисе e евклидова пространства E, имеющем матрицу Грама Γ, связаны соотношением*

$$A_1 = \Gamma^{-1} A^{\mathrm{T}} \Gamma. \tag{8.31}$$

В частности, если базис e ортонормированный, то

$$A_1 = A^{\mathrm{T}}. \tag{8.32}$$

▷ В матричной форме равенство (8.30) записывается в виде

$$x^{\mathrm{T}} A^{\mathrm{T}} \Gamma y = x^{\mathrm{T}} \Gamma A_1 y.$$

Поэтому если матрица A_1 оператора φ_1 связана с матрицей A оператора φ равенством $A^{\mathrm{T}}\Gamma = \Gamma A_1$, то $(\varphi x, y) = (x, \varphi_1 y)$ и оператор φ_1 оказывается сопряженным с оператором φ. Остается заметить, что равенство $A^{\mathrm{T}}\Gamma = \Gamma A_1$ эквивалентно равенству (8.31). Если базис e ортонормированный, то $\Gamma = E$ и равенство (8.31) переходит в равенство (8.32). ▶

Пример 8.13. Линейный оператор φ в базисе

$$e_1 = (1,\ 1,\ 1)^{\mathrm{T}}, \quad e_2 = (0,\ 1,\ 1)^{\mathrm{T}}, \quad e_3 = (0,\ 0,\ 1)^{\mathrm{T}}$$

имеет матрицу

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

Найти матрицу A_1 сопряженного оператора φ^* в этом же базисе, если известно, что координаты векторов e_1, e_2, e_3 заданы в ортонормированном базисе.

Решение. Поскольку координаты векторов e_1, e_2, e_3 заданы в ортонормированном базисе, скалярные произведения (e_i, e_j), $i, j = 1, 2, 3$, можно вычислить по формуле (8.13). Из этих скалярных произведений составим матрицу Грама

$$\Gamma = \begin{pmatrix} (e_1, e_1) & (e_1, e_2) & (e_1, e_3) \\ (e_2, e_1) & (e_2, e_2) & (e_2, e_3) \\ (e_3, e_1) & (e_3, e_2) & (e_3, e_3) \end{pmatrix} = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

Теперь по формуле (8.31) можно найти матрицу A_1 оператора φ^* в базисе e_1, e_2, e_3:

$$A_1 = \Gamma^{-1} A^{\mathrm{T}} \Gamma = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix} =$$

$$= \begin{pmatrix} 6 & 5 & 3 \\ -3 & -2 & -1 \\ -2 & -2 & -1 \end{pmatrix}.$$

Поставленную задачу можно решить другим способом. Матрица

$$T = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

является матрицей перехода от ортонормированного базиса e°, в котором были заданы координаты векторов e_1, e_2, e_3, к базису $e=(e_1,\ e_2,\ e_3)$. Найдем матрицу линейного оператора φ в базисе e°:

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

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

Теперь согласно формуле (8.32) матрица A_1' линейного оператора φ^* в базисе e° получается транспонированием матрицы A':

$$A_1'=(A')^{\mathrm{T}}=\begin{pmatrix}1&2&3\\0&1&2\\0&0&1\end{pmatrix}.$$

По найденной матрице A_1' в базисе e° получим матрицу A_1 оператора φ^* в базисе e:

$$A_1=T^{-1}A_1'T=\begin{pmatrix}1&0&0\\-1&1&0\\0&-1&1\end{pmatrix}\begin{pmatrix}1&2&3\\0&1&2\\0&0&1\end{pmatrix}\begin{pmatrix}1&0&0\\1&1&0\\1&1&1\end{pmatrix}=$$

$$=\begin{pmatrix}6&5&3\\-3&-2&-1\\-2&-2&-1\end{pmatrix}.$$

Пример 8.14. В базисе e скалярное произведение задано формулой

$$(x,y)=3x_1y_1+x_2y_2+2x_3y_3+x_1y_2+x_2y_1-$$
$$-2x_1y_3-2x_3y_1-x_2y_3-x_3y_2, \quad (8.33)$$

а линейный оператор φ — матрицей

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

Найти матрицу A_1 сопряженного оператора φ^* в том же базисе e.

Решение. По коэффициентам формулы (8.33), определяющей скалярное произведение в базисе e, составляем матрицу Грама базиса e:

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

Теперь по формуле (8.31) находим:

$$A_1 = \Gamma^{-1} A^{\mathrm{T}} \Gamma = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 \\ 2 & 0 & 3 \\ 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} 3 & 1 & -2 \\ 1 & 1 & -1 \\ -2 & -1 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 3 & -1 \\ 5 & 2 & -1 \\ 3 & 4 & 0 \end{pmatrix}.$$

Для решения задачи можно также построить ортонормированный базис пространства и поступить так, как сделано в предыдущем примере при решении задачи вторым способом. Однако при заданном законе (8.33) скалярного умножения векторов вычисления будут более громоздкими.

Отметим некоторые свойства сопряженных операторов.

1. $(\varphi^*)^* = \varphi$.
2. $(\varphi + \psi)^* = \varphi^* + \psi^*$.
3. $(\alpha\varphi)^* = \alpha\varphi^*$, α — любое действительное.
4. $(\varphi\psi)^* = \psi^*\varphi^*$.

Эти свойства доказываются по единой схеме. Нужно показать, что оператор в правой части является сопряженным с оператором, записанным в скобках в левой части. Свойство 1, кстати, уже доказано выше. Рассмотрим для примера свойство 4. Оно утверждает, что оператор $\psi^*\varphi^*$ является сопряженным с оператором $\varphi\psi$. Проверим это:

$$(\varphi\psi x, y) = (\psi x, \varphi^* y) = (x, \psi^*\varphi^* y).$$

Мы получили равенство (8.30), записанное для операторов $\varphi\psi$ и $\psi^*\varphi^*$. Это и означает, что оператор $\psi^*\varphi^*$ является сопряженным с оператором $\varphi\psi$.

Теорема 8.11. *Областью значений сопряженного оператора φ^* является ортогональное дополнение к ядру оператора φ.*

▷ Это следует из того, что для любого вектора x из ядра оператора φ и любого вектора y выполняется соотношение

$$(x, \varphi^* y) = (\varphi x, y) = (0, y) = 0,$$

т.е. вектор $\varphi^* y$ ортогонален вектору x. Это доказывает, что область значений оператора φ^* содержится в ортогональном дополнении к ядру оператора φ. Для совпадения двух подпространств достаточно равенства их размерностей. Если ядро оператора φ имеет размерность k, размерность евклидова пространства равна n, то согласно следствию 8.2 размерность ортогонального дополнения к ядру оператора φ равна $n-k$. Ту же размерность согласно теореме 5.2 имеет и область значений оператора φ^*, так как ранг этого оператора совпадает с рангом его матрицы в ортонормированном базисе и, следовательно, с рангом оператора φ. ▶

Основное свойство сопряженного оператора состоит в следующем.

Теорема 8.12. *Если некоторое подпространство L инвариантно относительно оператора φ, то ортогональное дополнение $L^\perp$ этого подпространства инвариантно относительно сопряженного оператора φ^*.*

▷ Действительно, пусть $x \in L$ и $y \in L^\perp$. Тогда из условия $\varphi x \in L$ следует $(\varphi x, y) = 0$, а так как $(\varphi x, y) = (x, \varphi^* y)$, то и $(x, \varphi^* y) = 0$. Отсюда следует, что $\varphi^* y \in L^\perp$. ▶

Теорема 8.13. *Характеристические многочлены, а следовательно, и собственные значения сопряженных операторов одинаковы.*

$\triangleright$ Пусть A — матрица линейного оператора φ в ортонормированном базисе e. Тогда $A^{\text{т}}$ — матрица сопряженного оператора φ^* в том же базисе. Поскольку определитель не меняет значения при транспонировании, то $|A-\lambda E| = |A^{\text{т}}-\lambda E|$, т.е. характеристические многочлены матриц A и $A^{\text{т}}$ одинаковы. Следовательно, совпадают и характеристические многочлены операторов φ и φ^*. $\blacktriangleright$

Теорема 8.14. *Каждый собственный вектор сопряженного оператора φ^* ортогонален ко всем собственным векторам оператора φ, принадлежащим другим собственным значениям.*

$\triangleright$ Пусть $\varphi x = \lambda_1 x$, $\varphi^* y = \lambda_2 y$, где $\lambda_1 \neq \lambda_2$. По определению сопряженного оператора выполняется равенство $(\varphi x, y) = (x, \varphi^* y)$, которое в данном случае принимает вид: $\lambda_1 (x, y) = \lambda_2 (x, y)$. Отсюда следует, что $(\lambda_1 - \lambda_2)(x, y) = 0$. Поскольку $\lambda_1 \neq \lambda_2$, то $(x, y) = 0$. $\blacktriangleright$

Последнее утверждение можно переформулировать для матриц следующим образом. Каждый собственный вектор действительной матрицы $A^{\text{т}}$ ортогонален каждому собственному вектору матрицы A, принадлежащему другому собственному значению.

8.9. Симметрические (самосопряженные) операторы

Линейный оператор φ, действующий в евклидовом пространстве E, называют ***симметрическим***, или ***самосопряженным***, если он является сопряженным с самим собой, т.е. $\varphi^* = \varphi$. Оператор φ является самосопряженным тогда и только тогда, когда для любых векторов $x, y \in E$ выполняется равенство

$$(\varphi x, y) = (x, \varphi y). \tag{8.34}$$

Теорема 8.15. *Симметрический оператор в любом ортонормированном базисе евклидова пространства имеет симметрическую матрицу. Наоборот, если линейный оператор φ в каком-либо ортонормированном базисе имеет симметрическую матрицу, то этот оператор симметрический.*

▷ Пусть линейный оператор φ в некотором ортонормированном базисе e имеет матрицу A. Тогда матрицей сопряженного оператора φ^* в базисе e является матрица A^{T}. Равенство операторов $\varphi^* = \varphi$ равносильно равенству матриц $A^{\mathrm{T}} = A$, но первое означает, что оператор симметрический, а второе — что его матрица симметрическая. ▶

Отметим некоторые свойства симметрических операторов.

1. ***Тождественный оператор является симметрическим.***

Действительно, матрица тождественного оператора в любом базисе, в том числе и ортонормированном, — единичная матрица, а такая матрица симметрическая.

2. ***Сумма симметрических операторов — симметрический оператор.***

Это утверждение вытекает из свойства симметрических матриц, в соответствии с которым сумма симметрических матриц является симметрической матрицей.

3. ***Произведение симметрического оператора на действительное число — симметрический оператор.***

Это свойство вытекает из соответствующего свойства симметрических матриц.

4. ***Если симметрический оператор φ имеет обратный оператор φ^{-1}, то этот оператор симметрический.***

Свойство вытекает из свойства матриц, согласно которому матрица, обратная к симметрической матрице, симметрическая.

5. Для симметрических операторов φ и ψ произведение $\varphi\psi$ является симметрическим оператором в том и только в том случае, когда операторы φ и ψ перестановочны, т.е. $\varphi\psi = \psi\varphi$.

Действительно, для симметрических операторов φ и ψ имеем

$$(\varphi\psi)^* = \psi^*\varphi^* = \psi\varphi.$$

Поэтому равенство $(\varphi\psi)^* = \varphi\psi$, означающее симметричность произведения $\varphi\psi$, равносильно равенству $\psi\varphi = \varphi\psi$, означающему перестановочность φ и ψ.

Теорема 8.16. *Пусть φ — симметрический оператор в евклидовом пространстве E. Если L — инвариантное подпространство оператора φ, то и ортогональное дополнение $L^\perp$ — инвариантное подпространство оператора φ.*

▷ Утверждение непосредственно вытекает из теоремы 8.12: если L — инвариантное подпространство оператора φ, то $L^\perp$ — инвариантное подпространство оператора φ^*, который на самом деле и есть оператор φ. ▶

Теорема 8.17. *Все корни характеристического многочлена симметрического оператора действительные.*

▷ Пусть φ — симметрический оператор с матрицей A в каком-либо ортонормированном базисе. Рассмотрим произвольный корень λ_0 характеристического многочлена $|A - \lambda E|$ оператора φ. Однородная система линейных уравнений $(A - \lambda_0 E)x = 0$ имеет ненулевые решения (возможно, комплексные), поскольку матрица системы имеет нулевой определитель. Пусть $x = (x_1,\ x_2,\ \ldots,\ x_n)^{\mathrm{T}}$ — одно из таких решений. Тогда $Ax = \lambda_0 x$. Умножим обе части этого матричного равенства на строку $\overline{x}^{\mathrm{T}}$:

$$\overline{x}^{\mathrm{T}} A x = \lambda_0 \overline{x}^{\mathrm{T}} x. \tag{8.35}$$

Произведение $x^{\mathrm{T}}x$ представляет собой действительное число, не равное нулю, так как

$$x^{\mathrm{T}}x = x_1\overline{x_1} + x_2\overline{x_2} + \ldots + x_n\overline{x_n} = |x_1|^2 + |x_2|^2 + \ldots + |x_n|^2.$$

Покажем, что в случае симметрической матрицы A произведение $\overline{x}^{\mathrm{T}}Ax$ также является действительным числом. Это произведение есть квадратная матрица порядка 1, т.е. число, а потому не изменяется при транспонировании. Следовательно,

$$\overline{x}^{\mathrm{T}} A x = (\overline{x}^{\mathrm{T}} A x)^{\mathrm{T}} = x^{\mathrm{T}} A^{\mathrm{T}} \overline{x} = x^{\mathrm{T}} A \overline{x}.$$

В то же время

$$\overline{\overline{x}^{\mathrm{T}} A x} = x^{\mathrm{T}} \overline{A x} = x^{\mathrm{T}} A \overline{x}.$$

В этих равенствах правые части одинаковы. Значит, и левые части тоже равны, т.е. $\overline{x}^{\mathrm{T}} A x = \overline{\overline{x}^{\mathrm{T}} A x}$. Поэтому число $\overline{x}^{\mathrm{T}} A x$ является действительным, как совпадающее со своим комплексно сопряженным числом. Из равенства (8.35) вытекает, что число λ_0, равное отношению двух действительных чисел, является действительным. ►

Замечание. Из доказанной теоремы вытекает, что ***все характеристические корни симметрической матрицы A являются действительными.*** В самом деле, достаточно рассмотреть оператор, матрицей которого в ортонормированном базисе является матрица A. Этот оператор будет симметрическим, а его характеристический многочлен будет совпадать с характеристическим многочленом матрицы A.

Теорема 8.18. *Собственные векторы симметрического оператора, принадлежащие различным собственным значениям, ортогональны.*

◁ Пусть выполняются равенства $\varphi x = \lambda x$, $\varphi y = \mu y$, где $\lambda \neq$ $\neq \mu$. По условию φ — симметрический оператор. Поэтому $(\varphi x, y) = (x, \varphi y)$, которое для рассматриваемых векторов x и y принимает вид $\lambda (x, y) = \mu (x, y)$. Перенося правую часть в левую, заключаем, что $(\lambda - \mu)(x, y) = 0$. А так как $\lambda - \mu \neq 0$, то $(x, y) = 0$, т.е. векторы x и y ортогональны. ►

Основное свойство симметрических операторов состоит в следующем.

Теорема 8.19. *Линейный оператор φ, действующий в евклидовом пространстве E, является симметрическим тогда и только тогда, когда в E существует ортонормированный базис, состоящий из собственных векторов оператора φ.*

$\triangleright$ Если e — ортонормированный базис e в E, состоящий из собственных векторов линейного оператора φ, то матрица оператора φ в этом базисе диагональная (см. теорему 5.17). Диагональная матрица является симметрической. Поэтому оператор φ, имеющий симметрическую матрицу в ортонормированном базисе, является симметрическим (см. 8.15).

Доказательство обратного утверждения проведем методом математической индукции по числу n измерений евклидова пространства E. При $n = 1$ утверждение теоремы тривиально, так как в одномерном евклидовом пространстве любой вектор отображается в коллинеарный ему. Следовательно, любой базис состоит из собственных векторов.

Допустим, что утверждение верно для евклидовых пространств размерности $n-1$, и рассмотрим произвольное евклидово пространство E размерности n и в нем симметрический оператор φ. В силу теорем 5.11 и 8.17 любой симметрический линейный оператор φ имеет собственные значения, а следовательно, и собственные векторы. Пусть e_1 — собственный вектор оператора φ, принадлежащий собственному значению λ_1. Можно считать, что этот вектор единичный, так как иначе его можно нормировать. Одномерное пространство E_1, порожденное вектором e_1, является инвариантным для оператора φ. Ортогональное дополнение $E_1^{\perp}$ к подпространству E_1 также инвариантно относительно оператора φ (см. 8.16). Это подпространство является $(n-1)$-мерным, оператор φ в $E_1^{\perp}$ симметрический. Поэтому в соответствии с индуктивным предположением в $E_1^{\perp}$ можно выбрать ортонормированный базис $e_2, \ldots, e_n$ из собственных векторов оператора φ. Добавив к этому базису вектор e_1, получим базис в E. Этот базис является ортонормированным, так как, во-первых, все векторы в нем единичные, а во-вторых, они попарно ортогональны. Кроме того, каждый вектор e_i — это собственный вектор оператора φ. Значит, базис $e_1, e_2, \ldots, e_n$ есть ортонормированный базис в E из собственных векторов оператора φ. Согласно методу математической индукции каждое конечномерное евклидово пространство имеет ортонормированный базис из собственных векторов симметрического оператора. ▶

Следствие. Симметрический оператор является оператором простой структуры.

Непосредственно из доказанной теоремы вытекает следующее утверждение.

Теорема 8.20. *Любая симметрическая матрица A с помощью ортогональной матрицы T приводится к диагональному виду, т.е. существует такая ортогональная матрица T, что матрица*

$$\Lambda = T^{-1}AT \tag{8.36}$$

диагональная. Любая симметрическая матрица A имеет каноническое разложение

$$A = T\Lambda T^{-1} = T\Lambda T^{T} \tag{8.37}$$

т.к. $T^{-1}=T^{T}$

с ортогональной матрицей T.

▷ Симметрическую матрицу A можно рассматривать как матрицу некоторого симметрического оператора в ортонормированном базисе евклидова пространства. Такой оператор имеет ортонормированный базис из собственных векторов, в котором матрица оператора есть некоторая диагональная матрица Λ. Матрицы A и Λ, как матрицы одного оператора в разных базисах, подобны, т.е. имеет место равенство (8.36), в котором T, как матрица перехода из одного ортонормированного базиса в другой, ортогональна (см. теорему 8.5). Второе утверждение теоремы — очевидная переформулировка первого, поскольку равенства (8.36) и (8.37) равносильны. ▶

Правило построения ортогональной матрицы T, входящей в соотношения (8.36) и (8.37), сохраняется таким же, как и для любых операторов простой структуры (см. разд. 5.8) с единственным различием. Если оператор имеет кратные собственные значения, то каждому такому значению отвечают в базисе из собственных векторов несколько векторов. Необходимо эту систему векторов ортонормировать. Поясним это на примере.

Пример 8.15. Построить каноническое разложение симметрической матрицы

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

с ортогональной трансформирующей матрицей.

Решение. Характеристический многочлен матрицы A

$$|A - \lambda E| = \begin{vmatrix} -\lambda & 1 & 1 & -1 \\ 1 & -\lambda & -1 & 1 \\ 1 & -1 & -\lambda & 1 \\ -1 & 1 & 1 & -\lambda \end{vmatrix} = (\lambda - 1)^3(\lambda + 3)$$

имеет корни $\lambda_1 = \lambda_2 = \lambda_3 = 1$, $\lambda_4 = -3$. Все они являются собственными значениями матрицы A.

При $\lambda = 1$ система $(A - \lambda E)x = 0$ имеет фундаментальную систему из трех решений, например, $a_1 = (1,\ 1,\ 0,\ 0)^{\text{т}}$, $a_2 = (1,\ 0,\ 1,\ 0)^{\text{т}}$, $a_3 = (-1,\ 0,\ 0,\ 1)^{\text{т}}$. Так как мы строим каноническое разложение матрицы A с ортогональной трансформирующей матрицей, то дополнительно нужно обеспечить, чтобы ФСР была ортонормированной системой. Для этого систему собственных векторов a_1, a_2, a_3 ортонормируем. В результате получим ортонормированную систему векторов

$$e_1' = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad e_2' = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ -1 \\ 2 \\ 0 \end{pmatrix}, \quad e_3' = \frac{1}{2\sqrt{3}}\begin{pmatrix} -1 \\ 1 \\ 1 \\ 3 \end{pmatrix}.$$

При $\lambda = -3$ система $(A - \lambda E)x = 0$ имеет фундаментальную систему решений из одного вектора, например, $a_4 = (1, -1, -1, 1)^{\text{т}}$. Нормируя его, получим $e_4' = \frac{1}{2}(1, -1, -1, 1)^{\text{т}}$.

Собственные векторы e'_1, e'_2, e'_3, e'_4 матрицы A составляют ортонормированный базис в четырехмерном евклидовом пространстве. Из столбцов координат этих векторов составим матрицу T и запишем искомое каноническое разложение

$$A = T\Lambda T^{-1} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & -\frac{1}{2\sqrt{3}} & \frac{1}{2} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{2\sqrt{3}} & -\frac{1}{2} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{2\sqrt{3}} & -\frac{1}{2} \\ 0 & 0 & \frac{3}{2\sqrt{3}} & \frac{1}{2} \end{pmatrix} \times$$

$= \Lambda T^{T}$

$$\times \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -3 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & 0 \\ -\frac{1}{2\sqrt{3}} & \frac{1}{2\sqrt{3}} & \frac{1}{2\sqrt{3}} & \frac{3}{2\sqrt{3}} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Представление симметрических матриц в виде канонического разложения с ортогональной трансформирующей матрицей имеет самое широкое применение в теории и приложениях. Для построения такого канонического разложения применимы и косвенные способы, например, метод вращений (см. разд. 11.1).

Симметрический оператор φ называют ***неотрицательным*** (***положительно определенным***), если для любого ненулевого вектора x выполняется неравенство $(\varphi x, x) \geqslant 0$ ($(\varphi x, x) > 0$). Если симметрический оператор φ неотрицательный (положительно определенный), то пишут $\varphi \geqslant 0$ ($\varphi > 0$).

Теорема 8.21. Симметрический оператор является неотрицательным (положительно определенным) тогда и только тогда, когда все его собственные значения неотрицательные (положительные).

▷ Если симметрический оператор φ является неотрицательным, то для любого собственного значения λ и принадлежащего ему собственного вектора x имеем

$$(\varphi x, x) = (\lambda x, x) = \lambda\,(x, x) = \lambda|x|^2 \geqslant 0.$$

Поскольку $|x|^2 > 0$, то $\lambda \geqslant 0$. Аналогичен случай положительно определенного оператора: достаточно нестрогие неравенства заменить строгими.

Пусть все собственные значения $\lambda_1, \lambda_2, \ldots, \lambda_n$ симметрического оператора φ неотрицательные. В евклидовом пространстве выберем базис $e_1, e_2, \ldots, e_n$ из собственных векторов оператора φ, упорядочивая векторы базиса так, что $\varphi x_i = \lambda_i e_i$, $i = 1, 2, \ldots, n$. Тогда для любого вектора $x = x_1e_1 + x_2e_2 + \ldots$ $\ldots + x_ne_n$ с учетом равенств $(e_i, e_j) = 0$ при $i \neq j$ и $(e_i, e_i) = 1$ получим:

$$(\varphi x, x) = \left(\varphi\Bigl(\sum_{i=1}^{n} x_ie_i\Bigr),\ \sum_{j=1}^{n} x_je_j\right) =$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i x_i x_j\,(e_i, e_j) = \sum_{i=1}^{n} \lambda_i x_i^2 \geqslant 0.$$

Следовательно, оператор φ неотрицательный. В случае, когда все собственные значения положительные, неравенство $(\varphi x, x) \geqslant 0$ на самом деле будет строгим, а оператор — положительно определенным. ▶

Теорема 8.22. *Для любого неотрицательного оператора φ и любого натурального числа m существует, и притом единственный, неотрицательный оператор f, такой, что $f^m = \varphi$. Если φ положительно определенный, то и f положительно определенный.*

Неотрицательный оператор f, удовлетворяющий условию $f^m = \varphi$, называют ***арифметическим корнем m-й степени из неотрицательного оператора*** φ.

▷ Пусть φ — неотрицательный оператор в n-мерном евклидовом пространстве E. Так как неотрицательный оператор по определению симметрический, в E существует ортонормированный базис e из собственных векторов оператора φ. В этом базисе оператор φ имеет диагональную матрицу

$$\Lambda = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, \tag{8.38}$$

причем все диагональные элементы λ_i — собственные числа оператора φ все неотрицательны. Рассмотрим линейный оператор f, который в базисе e имеет матрицу

$$F = \begin{pmatrix} \lambda_1^{1/m} & & 0 \\ & \ddots & \\ 0 & & \lambda_n^{1/m} \end{pmatrix}. \tag{8.39}$$

Оператор f симметрический, так как его матрица в ортонормированном базисе e симметрическая. Все его собственные числа — это диагональные элементы матрицы F. Они неотрицательны, а следовательно, оператор f неотрицательный. Наконец, из очевидного равенства $F^m = \Lambda$ вытекает равенство $f^m = \varphi$.

Пусть g — другой неотрицательный линейный оператор, для которого $g^m = \varphi$. Существует ортонормированный базис e' из собственных векторов оператора g, в котором его матрица диагональная. В этом базисе матрица оператора φ также диагональная. Значит, базис e' состоит из собственных векторов оператора φ. Нетрудно заключить, что собственные подпространства оператора g совпадают с собственными подпространствами оператора f, построенного ранее, причем и собственные значения этих операторов одинаковые. Отсюда следует, что операторы f и g совпадают. ▶

Введенные понятия и доказанную теорему можно перенести на матрицы следующим образом (см. гл. 9).

Симметрическую матрицу A называют ***неотрицательной*** (***положительно определенной***), если для любого столбца $x \neq 0$ выполняется условие $x^{\mathrm{T}}Ax \geqslant 0$ ($x^{\mathrm{T}}Ax > 0$). В этом случае пишут соответственно $A \geqslant 0$ и $A > 0$. Симметрическая матрица является неотрицательной (положительно определенной) тогда и только тогда, когда все ее характеристические числа неотрицательные (положительные).

Арифметическим корнем m-й степени из неотрицательной матрицы A называют такую неотрицательную матрицу B, что $B^m = A$. Как следует из теоремы 8.22, любая неотрицательная матрица имеет единственный корень m-й степени, причем если матрица положительно определена, то и корень m-й степени — положительно определенная матрица.

При вычислении арифметического корня $\sqrt[m]{A}$ достаточно построить каноническое разложение $A = T\Lambda T^{-1}$. Тогда $\sqrt[m]{A} = T\sqrt[m]{\Lambda}T^{-1}$, где корень m-й степени из диагональной матрицы Λ вида (8.38) вычисляется согласно равенству (8.39).

Пример 8.16. Для матрицы

$$A = \begin{pmatrix} 17 & 16 & 16 \\ 16 & 41 & 32 \\ 16 & 32 & 41 \end{pmatrix}$$

построить каноническое разложение с ортогональной трансформирующей матрицей и, пользуясь им, найти $\sqrt{A}$.

Решение. Характеристический многочлен

$$|A - \lambda E| = \begin{vmatrix} 17-\lambda & 16 & 16 \\ 16 & 41-\lambda & 32 \\ 16 & 32 & 41-\lambda \end{vmatrix} = -\lambda^3 + 99\lambda^2 - 1539\lambda + 6561$$

имеет корни $\lambda_1 = 81$, $\lambda_2 = \lambda_3 = 9$. Поэтому матрицей Λ является

$$\Lambda = \begin{pmatrix} 81 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{pmatrix}.$$

Определим трансформирующую ортогональную матрицу T. Для этого для каждого собственного значения λ_i нужно построить фундаментальную систему решений однородной системы $(A-\lambda_i E)x=0$ и ортонормировать ее, а затем из полученных столбцов собрать матрицу T.

При $\lambda_1=81$ эта система имеет вид

$$\begin{cases} -64x_1+16x_2+16x_3=0,\\ 16x_1-40x_2+32x_3=0,\\ 16x_1+32x_2-40x_3=0. \end{cases}$$

Фундаментальная система решений содержит одно решение, например $b_1=(1,\ 2,\ 2)^{\mathrm{T}}$. Нормируя его, находим $e'_1=\dfrac{1}{3}(1,\ 2,\ 2)^{\mathrm{T}}$.

При $\lambda_2=\lambda_3=9$ рассматриваемая система имеет вид:

$$\begin{cases} 8x_1+16x_2+16x_3=0,\\ 16x_1+32x_2+32x_3=0,\\ 16x_1+32x_2+32x_3=0. \end{cases}$$

Ее общим решением является $x=(-2x_2-2x_3,\ x_2,\ x_3)^{\mathrm{T}}$, а фундаментальная система решений состоит из двух столбцов, например $b_2=(-2,\ -1,\ 2)^{\mathrm{T}}$ и $b_3=(2,\ -2,\ 1)^{\mathrm{T}}$. Эти векторы, оказывается, уже ортогональны. Поэтому остается их лишь нормировать. В результате получим ортонормированную систему векторов $e'_2=\dfrac{1}{3}(-2,\ -1,\ 2)^{\mathrm{T}}$, $e'_3=\dfrac{1}{3}(2,\ -2,\ 1)^{\mathrm{T}}$. Из столбцов координат векторов e'_1, e'_2, e'_3 составляем ортогональную матрицу

$$T=\frac{1}{3}\begin{pmatrix} 1 & -2 & 2\\ 2 & -1 & -2\\ 2 & 2 & 1 \end{pmatrix}.$$

Поэтому искомое каноническое разложение матрицы A имеет вид:

$$A=T\Lambda T^{-1}=\frac{1}{3}\begin{pmatrix} 1 & -2 & 2\\ 2 & -1 & -2\\ 2 & 2 & 1 \end{pmatrix}\begin{pmatrix} 81 & 0 & 0\\ 0 & 9 & 0\\ 0 & 0 & 9 \end{pmatrix}\frac{1}{3}\begin{pmatrix} 1 & 2 & 2\\ -2 & -1 & 2\\ 2 & -2 & 1 \end{pmatrix}.$$

Теперь находим:

$$\sqrt{A}=\frac{1}{3}\begin{pmatrix}1&-2&2\\2&-1&-2\\2&2&1\end{pmatrix}\begin{pmatrix}9&0&0\\0&3&0\\0&0&3\end{pmatrix}\frac{1}{3}\begin{pmatrix}1&2&2\\-2&-1&2\\2&-2&1\end{pmatrix}=$$

$$=\frac{1}{9}\begin{pmatrix}33&12&12\\12&51&24\\12&24&51\end{pmatrix}.$$

Замечания. 1. При вычислении корня m-й степени матрицы можно использовать канонические разложения, в которых трансформирующая матрица не является ортогональной.

2. При компьютерных вычислениях корней из неотрицательных матриц эффективным методом получения канонического разложения является метод вращений (см. разд. 11.1).

3. О всех корнях матрицы см. [7], гл. VIII.

Следующая теорема дает примеры неотрицательных (положительно определенных) симметрических операторов.

Теорема 8.23. *Для любого линейного оператора φ, действующего в евклидовом пространстве E, операторы $\varphi^*\varphi$ и $\varphi\varphi^*$ являются неотрицательными, т.е. эти операторы симметрические, а все их собственные числа неотрицательные.*

$\triangleright$ Симметричность оператора $\varphi^*\varphi$ следует из равенств

$$(\varphi^*\varphi)^*=\varphi^*(\varphi^*)^*=\varphi^*\varphi.$$

Далее, для любого вектора x имеем:

$$(\varphi^*\varphi x,x)=(\varphi x,\varphi x)=|\varphi x|^2\geqslant 0,$$

что и означает неотрицательность оператора $\varphi^*\varphi$. Если положить $\psi=\varphi^*$, то $\varphi\varphi^*=\psi^*\psi$, так что утверждение теоремы для оператора $\varphi\varphi^*$ уже доказано. $\blacktriangleright$

Теорема 8.24. *Для любого линейного оператора φ, действующего в евклидовом пространстве E, и для любого* *натурального числа m существуют, и притом единственные, операторы f_r и f_l, для которых*

$$f_r^m = \varphi^*\varphi, \qquad f_l^m = \varphi\varphi^*.$$

▷ Сформулированное утверждение является прямым следствием теорем 8.22 и 8.23. ▶

Согласно теореме 8.24 для любого оператора φ в евклидовом пространстве определены операторы $\sqrt{\varphi^*\varphi}$ и $\sqrt{\varphi\varphi^*}$, называемые соответственно *правым* и *левым модулями* оператора φ.

Утверждения теорем 8.23 и 8.24 можно распространить на случай оператора, действующего из одного евклидова пространства в другое.

Теоремы 8.23 и 8.24 несложно переформулировать в терминах матриц. ***Для любой (вообще говоря, комплексной) матрицы A матрицы A^*A и AA^* являются симметрическими неотрицательными, существуют однозначно определенные арифметические корни произвольной степени m из матриц A^*A и AA^*, являющиеся неотрицательными матрицами. В частности, однозначно определены корни $\sqrt{A^*A}$ и $\sqrt{AA^*}$.***

8.10. Ортогональные операторы

Линейный оператор φ, действующий в евклидовом пространстве E, называется ***ортогональным***, если он сохраняет скалярное произведение любых векторов $x, y \in E$, т.е. если выполняется равенство

$$(\varphi x, \varphi y) = (x, y)\,. \tag{8.40}$$

Полагая в равенстве (8.40) $x = y$, получаем $|\varphi x|^2 = |x|^2$. Это означает, что ортогональный оператор сохраняет длины векторов.

Поскольку угол между векторами определяется из соотношения

$$\cos\varphi = \frac{(x, y)}{|x|\,|y|}, \quad 0 \leqslant \varphi \leqslant \pi,$$

и поскольку числитель и знаменатель в этом соотношении не меняются при действии ортогональных операторов, то ортогональные операторы сохраняют углы между векторами.

Теорема 8.25. *Ортогональный оператор любую ортонормированную систему векторов переводит в ортонормированную, а ортонормированный базис — в ортонормированный базис. Если линейный оператор переводит какой-либо ортонормированный базис в ортонормированный, то этот оператор ортогональный.*

◁ Первое утверждение теоремы — следствие того, что ортогональный оператор сохраняет длины векторов и переводит ортогональные векторы в ортогональные. Предположим, что линейный оператор φ переводит ортонормированный базис $e = (e_1, e_2, \ldots, e_n)$ в ортонормированный базис $e' = (e_1', e_2', \ldots, e_n')$. Тогда для векторов

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

имеем

$$\varphi x = x_1e_1' + x_2e_2' + \ldots + x_ne_n', \quad \varphi y = y_1e_1' + y_2e_2' + \ldots + y_ne_n'.$$

Следовательно,

$$(\varphi x, \varphi y) = x_1y_1 + x_2y_2 + \ldots + x_ny_n = (x, y),$$

и φ — ортогональный оператор. ▶

Теорема 8.26. *Если φ — ортогональный оператор, то оператор φ^*, сопряженный с оператором φ, совпадает с обратным оператором φ^{-1}.* $\varphi^* = \varphi^T = \varphi^{-1}$

◁ Сначала отметим, что ортогональный оператор имеет нулевое ядро, так как нулевой вектор может быть образом лишь того вектора, который имеет нулевую длину, т.е. нулевого вектора. Далее, из равенств

$$(x, y) = (\varphi x, \varphi y) = (x, \varphi^* \varphi y)$$

вытекает, что $(x, y - \varphi^* \varphi y) = 0$ для любых двух векторов x и y. В частности, при $x = y - \varphi^* \varphi y$, заключаем, что $|y - \varphi^* \varphi y|^2 = 0$ и $y - \varphi^* \varphi y = 0$. Таким образом, оператор $\varphi^* \varphi$ совпадает с тождественным оператором ε, т.е. $\varphi^* \varphi = \varepsilon$. Умножив последнее равенство справа на оператор φ^{-1}, приходим к равенству $\varphi^* = \varphi^{-1}$. ▶

Замечание. Подчеркнем еще раз факт, использованный в доказательстве теоремы: ортогональный оператор всегда невырожденный.

Теорема 8.27. *У ортогонального оператора матрица в любом ортонормированном базисе является ортогональной. Наоборот, если в каком-либо ортонормированном базисе матрица линейного оператора ортогональна, то этот оператор ортогональный.*

▷ Для ортогонального оператора φ выполняется равенство $\varphi^* \varphi = \varepsilon$, которое для матрицы A этого оператора имеет вид $A^{\mathrm{T}} A = E$. А это равносильно ортогональности матрицы A. Пусть матрица A линейного оператора φ удовлетворяет равенству $A^T A = E$. Тогда для самого оператора φ верно равенство $\varphi^* \varphi = \varepsilon$. Применяя это равенство, для любых векторов x и y получаем:

$$(\varphi x, \varphi y) = (x, \varphi^* \varphi y) = (x, y).$$

Следовательно, φ — ортогональный оператор. ▶

Используя свойства ортогональных матриц, легко установить некоторые свойства ортогональных операторов:

1) тождественный оператор является ортогональным;

2) произведение ортогональных операторов — ортогональный оператор;

3) оператор, обратный ортогональному оператору, является ортогональным.

Первое из перечисленных свойств очевидно; второе следует из того, что произведение ортогональных матриц является ортогональной матрицей; третье также вытекает из соответствующего свойства ортогональных матриц.

Теорема 8.28. *Корни характеристического многочлена ортогонального оператора (включая и комплексные) по абсолютной величине равны единице.*

▷ Утверждение относительно ортогонального оператора можно переформулировать для ортогональных матриц. Если λ_0 — корень характеристического многочлена ортогональной матрицы A, то $|A-\lambda_0 E|=0$ и однородная система в общем случае комплексных уравнений $(A-\lambda_0 E)x=0$ имеет ненулевое решение x, т.е. существует такой ненулевой столбец x, что

$$Ax=\lambda_0 x. \tag{8.41}$$

Равенство (8.41) транспонируем и перейдем к комплексно сопряженным числам:

$$\overline{x}^{\mathrm{T}}\overline{A}^{\mathrm{T}}=\overline{\lambda_0}\,\overline{x}^{\mathrm{T}}. \tag{8.42}$$

Приравняв произведение левых частей равенств (8.41) и (8.42) к произведению их правых частей, получим:

$$\overline{x}^{\mathrm{T}}\overline{A}^{\mathrm{T}}Ax=\overline{\lambda_0}\lambda_0\overline{x}^{\mathrm{T}}x.$$

По условию A — действительная ортогональная матрица. Поэтому $\overline{A}^{\mathrm{T}}A=A^{\mathrm{T}}A=E$ и

$$\overline{x}^{\mathrm{T}}x=\overline{\lambda_0}\lambda_0\overline{x}^{\mathrm{T}}x.$$

Так как число $\overline{x}^{\mathrm{T}}x=x^*x$ положительное ненулевое, на него можно сократить. В результате придем к равенству $1=\overline{\lambda_0}\lambda$, из которого вытекает, что $|\lambda_0|=1$. ▶

Следствие. Собственными числами ортогонального оператора могут быть только числа 1 и -1.

Теорема 8.29. *Любые два собственных вектора, принадлежащие различным собственным значениям ортогонального оператора, ортогональны.*

▷ Согласно следствию из теоремы 8.28 речь идет о векторах x и y, из которых вектор x принадлежит собственному значению 1, а другой y — собственному значению -1. В этом случае

$$(x, y) = (\varphi x, \varphi y) = (x, -y) = -(x, y).$$

Поэтому $(x, y) = 0$. ▶

Теорема 8.30. *Если подпространство L инвариантно относительно ортогонального оператора φ, то его ортогональное дополнение $L^\perp$ также инвариантно относительно φ.*

▷ Так как подпространство L инвариантно относительно φ, линейный оператор φ можно рассмотреть как оператор, действующий на этом подпространстве, т.е. рассмотреть сужение $\varphi\big|_L$ оператора φ на подпространство L. Оператор φ_L является ортогональным оператором, а потому невырожденным. Следовательно, для любого вектора $y \in L$ существует вектор $z \in L$, для которого $\varphi z = y$.

Пусть $x \in L^\perp$ выбран произвольно. Условие $x \in L^\perp$ означает, что $(x, y) = 0$ для любого вектора $y \in L$. Выберем вектор $z \in L$, для которого $\varphi z = y$. Тогда

$$(\varphi x, y) = (\varphi x, \varphi z) = (x, z) = 0,$$

так как вектор x ортогонален любому вектору из L, в том числе и вектору z. Но условие $(\varphi x, y) = 0$, выполняющееся для любого вектора $y \in L$, означает, что $\varphi x \in L^\perp$. Тем самым доказано, что $L^\perp$ инвариантно относительно φ. ▶

Теорема 8.31. *Пусть φ — ортогональный оператор, действующий в евклидовом пространстве E. Евклидово пространство E может быть представлено в виде прямой суммы инвариантных подпространств оператора φ; размерность каждого из которых равна 1 или 2.*

▷ Доказательство теоремы проведем методом математической индукции по размерности пространства. Утверждение теоремы тривиально для одно- и двумерных евклидовых пространств. Предположим, что теорема верна для всех евклидовых пространств размерности, меньшей n, и докажем утверждение для всех евклидовых пространств размерности n. Выберем произвольное n-мерное евклидово пространство E и в нем произвольный ортогональный оператор φ. По теореме 5.13 в E существует одно- или двумерное подпространство L, инвариантное относительно φ. По теореме 8.30 подпространство $L^\perp$ тоже инвариантно относительно φ. На $L^\perp$ оператор φ остается ортогональным, так как для него выполняется условие $(\varphi x, \varphi y) = (x, y)$, которое сохраняется и для любых векторов из $L^\perp$. Подпространства L и $L^\perp$ согласно теореме 8.7 образуют прямую сумму: $E = L \oplus L^\perp$. Поскольку $L^\perp$ имеет размерность $n-1$ или $n-2$, то согласно предположению математической индукции $L^\perp$ можно представить в виде прямой суммы инвариантных подпространств оператора φ. Добавляя к прямой сумме подпространство L, получим разложение в прямую сумму всего пространства E. ►

Перейдем к исследованию структуры произвольного ортогонального оператора и его матрицы.

Если оператор φ действует в одномерном пространстве E, то для любого вектора $e \in E$ имеем $\varphi e = \pm e$. Следовательно, φ — либо тождественный оператор, либо оператор центральной симметрии с матрицей (-1).

Пусть ортогональный оператор φ действует в двумерном пространстве E. Выберем некоторый ортонормированный базис $e = (e_1, e_2)$. Вектор $e_1' = \varphi e_1$ является единичным, и его координаты можно записать в виде $(\cos\alpha, \sin\alpha)^{\mathrm{T}}$, где α — угол, который составляет вектор e_1' с вектором e_1. Вектор $e_2' = \varphi e_2$ ортогонален e_1' и имеет единичную длину. Таких векторов два. Первый имеет координаты $(-\sin\alpha, \cos\alpha)^{\mathrm{T}}$ и получается поворотом вектора e_1' на угол $\frac{\pi}{2}$ против часовой стрелки. Второй противоположен первому и получается поворотом вектора e_1' на угол $\frac{\pi}{2}$ по часовой стрелке. В первом случае матрица A оператора φ имеет вид:

$$A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}, \tag{8.43}$$

а действие оператора φ состоит в повороте вектора на угол α против часовой стрелки (точнее, в направлении от вектора e_1 к вектору e_2).

Во втором случае матрица оператора φ имеет вид:

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

Отметим, что эта матрица симметрическая и, следовательно, приводится к диагональному виду. Поскольку оператор φ отличается от оператора $\lambda\varepsilon$, отличающегося от тождественного оператора ε множителем, его собственные значения различны, а по теореме 8.28 одно собственное значение равно 1, а второе равно -1. Таким образом, в рассматриваемом случае матрица оператора φ приводится к виду

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{8.44}$$

а оператор φ есть симметрия относительно некоторой прямой.

Рассмотренные случаи в свете теоремы 8.31 оказываются решающими в вопросе выяснения структуры ортогонального оператора. В общем случае верно следующее утверждение.

Теорема 8.32. *Для любого ортогонального оператора φ, действующего в n-мерном евклидовом пространстве E, существует ортонормированный базис, в котором матрица A оператора имеет вид:*

$$\begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & 0 & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & \Phi_1 & & \\ & 0 & & & & & & \ddots & \\ & & & & & & & & \Phi_k \end{pmatrix}, \tag{8.45}$$

где $\Phi_1, \ldots, \Phi_k$ *— двумерные блоки вида*

$$\Phi_i = \begin{pmatrix} \cos\alpha_i & -\sin\alpha_i \\ \sin\alpha_i & \cos\alpha_i \end{pmatrix}.$$

$\triangleright$ Эта теорема напрямую связана с теоремой 8.31. В разложении евклидова пространства E в прямую сумму одно- и двумерных инвариантных подпространств одномерные подпространства порождаются собственными векторами, и их можно разделить на две группы: собственные векторы, принадлежащие собственному числу 1, и собственные векторы, принадлежащие собственному числу -1. Двумерные инвариантные подпространства по типу действия в них ортогонального могут быть двух видов. В инвариантных подпространствах первого вида действие оператора состоит в повороте вектора на некоторый угол. В инвариантных подпространствах второго вида действие оператора состоит в преобразовании симметрии относительно иpямой. Такие инвариантные подпространства содержат собственные векторы оператора и распадаются на пару однюмерных инвариантных подпространств. Таким образом, уточняя теорему 8.31, можем утверждать, что для ортогонального оператора φ евклидово пространство E может быть представлено в виде

$$E = P_1 \oplus \ldots \oplus P_m \oplus Q_1 \oplus \ldots \oplus Q_l \oplus R_1 \oplus \ldots \oplus R_k,$$

где $P_1, \ldots, P_m$ — одномерные инвариантные подпространства, порождаемые собственными векторами с собственным значением 1; $Q_1, \ldots, Q_l$ — аналогичные одномерные подпространства, но отвечающие собственному значению -1; $R_1, \ldots, R_k$ — двумерные инвариантные подпространства, в которых оператор осуществляет поворот вектора на некоторый угол α.

Выбрав в этих инвариантных подпространствах ортонормированные базисы и собрав их вместе, получим базис евклидова пространства E, в котором оператор φ имеет матрицу вида (8.45). Отметим, что матрица вида (8.45) определена с точностью до перестановки двумерных клеток. ▶

Построение ортонормированного базиса, в котором матрица ортогонального оператора имеет канонический вид (8.45), проиллюстрируем примером.

Пример 8.17. Ортогональный оператор φ в ортонормированном базисе e евклидова пространства имеет матрицу

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

Построить базис, в котором оператор φ имеет матрицу B канонического вида, и найти эту матрицу.

Решение. Характеристический многочлен $|A - \lambda E| = -\lambda^3 + 2\lambda^2 - 2\lambda + 1$ оператора φ имеет корни $\lambda_1 = 1$, $\lambda_{2,3} = \alpha \pm \beta i = \frac{1}{2} \pm i\frac{\sqrt{3}}{2}$.

При $\lambda = 1$ система $(A - \lambda E)X = 0$ имеет фундаментальную систему решений из одного вектора, например, $X = (1,\ 1,\ 1)^{\text{т}}$. Нормируя этот вектор, получаем:

$$e'_1 = \frac{X}{|X|} = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)^{\text{т}}.$$

При $\lambda = \alpha + \beta i = \frac{1}{2} + i\frac{\sqrt{3}}{2}$ система $(A - \lambda E)X = 0$ имеет фундаментальную систему решений из одного вектора, например,

$$X = u + iv = \begin{pmatrix} 2 \\ -1 - i\sqrt{3} \\ -1 + i\sqrt{3} \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix} + i\begin{pmatrix} 0 \\ -\sqrt{3} \\ \sqrt{3} \end{pmatrix}.$$

Нормируя векторы $u = (2,\ -1,\ -1)^{\text{т}}$ и $v = \left(0,\ -\sqrt{3},\ \sqrt{3}\right)^{\text{т}}$, получим еще два вектора ортонормированного базиса:

$$e'_2 = \frac{u}{|u|} = \left(\frac{\sqrt{6}}{3},\ -\frac{\sqrt{6}}{6},\ -\frac{\sqrt{6}}{6}\right)^{\text{т}}, \quad e'_3 = \frac{v}{|v|} = \left(0,\ -\frac{\sqrt{2}}{2},\ \frac{\sqrt{2}}{2}\right)^{\text{т}}.$$

Базис e' из векторов $e'_1, e'_2,\ e'_3$ является искомым. В нем матрица B оператора φ имеет канонический вид:

$$B = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \alpha & \beta \\ 0 & -\beta & \alpha \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \dfrac{1}{2} & \dfrac{\sqrt{3}}{2} \\ 0 & -\dfrac{\sqrt{3}}{2} & \dfrac{1}{2} \end{pmatrix},$$

а матрицей перехода из базиса e в базис e' является матрица

$$T = \begin{pmatrix} \dfrac{1}{\sqrt{3}} & \dfrac{\sqrt{6}}{3} & 0 \\ \dfrac{1}{\sqrt{3}} & -\dfrac{\sqrt{6}}{6} & -\dfrac{\sqrt{2}}{2} \\ \dfrac{1}{\sqrt{3}} & -\dfrac{\sqrt{6}}{6} & \dfrac{\sqrt{2}}{2} \end{pmatrix}.$$

Замечание. Для комплексного собственного значения $\lambda = \alpha + \beta i$ ортогональной матрицы A векторы u и v в решении $X = u + iv$ системы $(A - \lambda E)X = 0$ всегда ортогональны и имеют одинаковую длину. Эти векторы удовлетворяют условиям $\varphi u = \alpha u - \beta v$, $\varphi v = \beta u + \alpha v$. Простое доказательство этих фактов дано, например, в [22] (решение задачи 1569). Последние условия выведены в разд. 5.7. Из них, в частности, вытекает, что матрица сужения оператора φ на подпространство $L = \langle u,\ v\rangle$ имеет вид $\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$.

Простейшими ортогональными операторами являются операторы ***вращения*** и ***отражения***. ***Оператор простого вращения*** — это ортогональный оператор, для которого в каноническом виде (8.45) нет диагональных значений -1, а лишь одна двумерная клетка. Операторы простого вращения часто используют в разложениях в произведение других ортогональных операторов. Если двумерное инвариантное подпространство оператора простого вращения порождено парой базисных векторов e_i и e_j, то матрица этого оператора имеет вид:

$$
\left(\begin{array}{ccccccccccc}
1 & & & \vdots & & & & \vdots & & & \\
 & \ddots & & \vdots & & & & \vdots & & & \\
 & & 1 & \vdots & & & & \vdots & & & \\
\cdots & \cdots & \cdots & \cos\alpha & \cdots & \cdots & \cdots & -\sin\alpha & \cdots & \cdots & \cdots \\
 & & & \vdots & 1 & & & \vdots & & & \\
 & & & \vdots & & \ddots & & \vdots & & & \\
 & & & \vdots & & & 1 & \vdots & & & \\
\cdots & \cdots & \cdots & \sin\alpha & \cdots & \cdots & \cdots & \cos\alpha & \cdots & \cdots & \cdots \\
 & & & \vdots & & & & \vdots & 1 & & \\
 & & & \vdots & & & & \vdots & & \ddots & \\
 & & & \vdots & & & & \vdots & & & 1
\end{array}\right)
\begin{array}{l}
\\ \\ \\
i\text{-я строка} \\
\\ \\ \\
j\text{-я строка} \\
\\ \\ \\
\end{array}
$$

Такую матрицу называют ***матрицей простого вращения***, или ***матрицей Гивенса***.

Матрицы вращений часто применяются в вычислительных методах при упрощении матриц.

Пример 8.18. Привести матрицу

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

с помощью вращений к треугольному виду.

Решение. Сначала получим в матрице A нуль на пересечении второй строки с первым столбцом. После умножения матрицы A слева на матрицу

$$
T_1 = \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}
$$

элементом матрицы $A_1 = T_1 A$ во второй строке первом столбце будет $2\sin\alpha + \cos\alpha$. Приравнивая этот элемент к нулю, найдим: $\operatorname{tg}\alpha = -\frac{1}{2}$. Поэтому

$$\cos\alpha = \frac{1}{\sqrt{1+\operatorname{tg}^2\alpha}} = \frac{2}{\sqrt{5}}, \qquad \sin\varphi = \frac{\operatorname{tg}\alpha}{\sqrt{1+\operatorname{tg}^2\alpha}} = -\frac{1}{\sqrt{5}}.$$

Следовательно,

$$T_1 = \begin{pmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 \\ -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

и

$$A_1 = T_1A = \begin{pmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 \\ -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & \sqrt{5} & \sqrt{5} \\ 1 & \sqrt{5} & \sqrt{5} \\ \sqrt{5} & 3 & 2 \end{pmatrix} = \begin{pmatrix} \sqrt{5} & 3 & 3 \\ 0 & 1 & 1 \\ \sqrt{5} & 3 & 2 \end{pmatrix}.$$

Возьмем теперь матрицу

$$T_2 = \begin{pmatrix} \cos\alpha & 0 & -\sin\alpha \\ 0 & 1 & 0 \\ \sin\alpha & 0 & \cos\alpha \end{pmatrix}.$$

Как и в предыдущем случае, из равенства нулю элемента матрицы $A_2 = T_2A_1$ в третьей строке первом столбце, т.е. из равенства $\sqrt{5}\sin\alpha + \sqrt{5}\cos\alpha = 0$, находим:

$$\operatorname{tg}\alpha = -1, \qquad \cos\alpha = \frac{1}{\sqrt{2}}, \qquad \sin\alpha = -\frac{1}{\sqrt{2}}.$$

Поэтому

$$T_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix}$$

и

$$A_2 = T_2 A_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \sqrt{5} & 3 & 3 \\ 0 & 1 & 1 \\ \sqrt{5} & 3 & 2 \end{pmatrix} = \begin{pmatrix} \sqrt{10} & \frac{6}{\sqrt{2}} & \frac{5}{\sqrt{2}} \\ 0 & 1 & 1 \\ 0 & 0 & -\frac{1}{\sqrt{2}} \end{pmatrix} = R.$$

Матрица $A_2 = R$ уже является матрицей нужного вида. Иначе процесс преобразований с помощью вращений можно было бы продолжить.

Отражением называют оператор, переводящий каждый вектор евклидова пространства E в симметричный ему вектор относительно некоторого $(n-1)$-мерного пространства L. Ортогональное дополнение $L^{\perp}$ является одномерным пространством, а вектором, симметричным вектору $x = x_0 + x^{\perp}$, где x_0 — ортогональная проекция на L, $x^{\perp}$ — ортогональная составляющая, является вектор $x_0 - x^{\perp}$. Векторы ортогонального дополнения $L^{\perp}$ характеризуются тем, что они коллинеарны разности любого вектора и его образа. Такие векторы называют ***определяющими векторами*** данного отражения. Выбрав ортонормированные базисы в подпространствах L и $L^{\perp}$ и составив из них общий базис, получим матрицу оператора отражения канонического вида (8.45), в которой нет двумерных клеток, а среди диагональных элементов лишь один элемент равен -1, т.е. ***матрицу отражения***

$$\begin{array}{cl} \begin{pmatrix} 1 & & & \vdots & & & \\ & \ddots & & \vdots & & & \\ & & 1 & \vdots & & & \\ \cdots & \cdots & \cdots & -1 & \cdots & \cdots & \cdots \\ & & & \vdots & 1 & & \\ & & & \vdots & & \ddots & \\ & & & \vdots & & & 1 \end{pmatrix} & \begin{array}{l} k\text{-я} \\ \text{строка} \end{array} \end{array}.$$

Непосредственной проверкой нетрудно убедиться в том, что матрицу отражения можно представить в виде

$$H = E - 2\frac{vv^{\mathrm{T}}}{|v|^2}, \tag{8.46}$$

где E — единичная матрица порядка n, v — столбец координат в ортонормированном базисе одного из определяющих векторов отражения.

Отражение с матрицей (8.46) в иностранной литературе называют также ***преобразованием Хаусхолдера***. Матрица (8.46) является симметрической и ортогональной, т.е. удовлетворяет условиям $H = H^{\mathrm{T}} = H^{-1}$.

В вычислительной практике отражения применяют, например, для изменения координат векторов в нужных позициях. Так, в векторе $y = (1, 1, 1, 1, 1, 1, 1)^{\mathrm{T}}$ можно одним отражением обратить в нуль 3-ю, 5-ю и 7-ю координаты. При таком отражении можно отобразить данный вектор в любой вектор той же длины, взяв их разность в качестве определяющего вектора. В данном случае в качестве образа вектора y можно взять вектор $y' = (-2, 1, 0, 1, 0, 1, 0)^{\mathrm{T}}$, имеющий нулевые 3-ю, 5-ю и 7-ю координаты, а по длине равный вектору y. Тогда определяющим вектором будет вектор $v = y - y' = (3, 0, 1, 0, 1, 0, 1)^{\mathrm{T}}$.

Нетрудно увидеть, что в определении отражения участвовали только четыре координаты вектора y, три обнуляемые и еще одна, использованная для регулировки длины измененного вектора. Поэтому вместо вектора y можно было взять вектор $x = (1, 0, 1, 0, 1, 0, 1)$, у которого ненулевыми являются только изменяемые координаты вектора y и еще одна, в данном случае первая, для регулировки длины. Чтобы обнулить координаты вектора y, в качестве образа вектора x следует взять вектор $x' = \pm|x|z$, где $z = (1, 0, 0, 0, 0, 0, 0)^{\mathrm{T}}$. Такое преобразование оставляет неизменными все координаты вектора y, кроме указанных в векторе x. При выборе определяющего вектора $v = x - x' = x \pm |x|z$ есть две возможности, связанные с выбором знака. Обычно знак в этой формуле выбирают так, чтобы он совпадал со знаком дополнительной координаты (в данном случае первой) вектора x, используемой для регулировки длины образа.

Определив вектор $v = x \pm |x|z = (3,\ 0,\ 1,\ 0,\ 1,\ 0,\ 1)^{\text{T}}$, можно вычислить матрицу отражения (в предположении, что все координаты даны в ортонормированном базисе):

$$H = E - 2\frac{vv^{\text{T}}}{|v|^2} = E - \frac{1}{6}\begin{pmatrix}3\\0\\1\\0\\1\\0\\1\end{pmatrix}(3,\ 0,\ 1,\ 0,\ 1,\ 0,\ 1) =$$

$$= \frac{1}{6}\begin{pmatrix} -3 & 0 & -3 & 0 & -3 & 0 & -3 \\ 0 & 6 & 0 & 0 & 0 & 0 & 0 \\ -3 & 0 & 5 & 0 & -1 & 0 & -1 \\ 0 & 0 & 0 & 6 & 0 & 0 & 0 \\ -3 & 0 & -1 & 0 & 5 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 6 & 0 \\ -3 & 0 & -1 & 0 & -1 & 0 & 5 \end{pmatrix}$$

Такая матрица переводит вектор y в вектор

$$y' = Hy = (-2,\ 1,\ 0,\ 1,\ 0,\ 1,\ 0)^{\text{T}}.$$

Отражения особенно часто применяют для упрощения матриц, а именно при приведении их к треугольному виду, почти треугольному виду, к двух- и трехдиагональной форме и т.п.

Пример 8.19. Привести матрицу

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

с помощью отражений к треугольному виду.

Решение. Чтобы получить в первом столбце матрицы A нули ниже главной диагонали, отобразим вектор $x = (1,\ 2,\ 2)^{\text{T}}$

в вектор $x' = (\pm|x|,\ 0,\ 0)^{\mathrm{T}}$. Для этого, положив $z = (1,\ 0,\ 0)^{\mathrm{T}}$, $v = x - x' = x + 3z = (4,\ 2,\ 2)^{\mathrm{T}}$, составим матрицу

$$H_1 = E - 2\frac{vv^{\mathrm{T}}}{|v|^2} = E - \frac{1}{2}\begin{pmatrix} 4 \\ 2 \\ 2 \end{pmatrix}(4,\ 2,\ 2) = \frac{1}{3}\begin{pmatrix} -1 & -2 & -2 \\ -2 & 2 & -1 \\ -2 & -1 & -2 \end{pmatrix}$$

и найдем матрицу

$$A_1 = U_1 A = \frac{1}{3}\begin{pmatrix} -1 & -2 & -2 \\ -2 & 2 & -1 \\ -2 & -1 & -2 \end{pmatrix}\begin{pmatrix} 1 & 3 & 3 \\ 2 & 3 & 0 \\ 2 & 0 & 3 \end{pmatrix}\begin{pmatrix} -3 & -3 & -3 \\ 0 & 0 & -3 \\ 0 & -3 & 0 \end{pmatrix}.$$

Для того чтобы в матрице A_1 получить нули ниже главной диагонали во втором столбце, необходимо преобразовать подматрицу, получаемую из A_1 вычеркиванием первой строки и первого столбца. Для этого отобразим с помощью отражения двумерный вектор $x = (0,\ -3)^{\mathrm{T}}$ в вектор $x' = (-|x|,\ 0)^{\mathrm{T}}$. При этом $z = (1,\ 0)^{\mathrm{T}}$, $v = x - x' = (3,\ -3)^{\mathrm{T}}$, а матрица отражения имеет вид:

$$H_2 = E - 2\frac{vv^{\mathrm{T}}}{|v|^2} = E - \frac{2}{18}\begin{pmatrix} 3 \\ -3 \end{pmatrix}(3,\ -3) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Чтобы преобразовывать не отдельный блок, а всю матрицу A_1, расширим матрицу H_2, вложив ее как блок в единичную матрицу. Получим матрицу

$$U_2 = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array}\right),$$

с помощью которой вычислим матрицу

$$A_2 = U_2 A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}\begin{pmatrix} -3 & -3 & -3 \\ 0 & 0 & -3 \\ 0 & -3 & 0 \end{pmatrix} =$$

$$= \begin{pmatrix} -3 & -3 & -3 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \end{pmatrix} = R.$$

Матрица $A_2 = U_2A_1 = UA$ уже нужного вида, а матрица совокупного преобразования $U = U_2H_1$ из двух отражений имеет вид:

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \frac{1}{3} \begin{pmatrix} -1 & -2 & -2 \\ -2 & 2 & -1 \\ -2 & -1 & -2 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 1 & -2 & -2 \\ -2 & -1 & 2 \\ -2 & 2 & -1 \end{pmatrix}.$$

Отражения можно применять для модификации строк матрицы. В этом случае матрица H отражения может быть найдена по формуле

$$H = E - 2\frac{v^{\mathrm{T}}v}{|v|^2},$$

близкой к формуле (8.46). При этом преобразуемая матрица умножается на матрицу отражения не слева, а справа.

Часто при упрощении матрицы приходится применять отражения к ее столбцам и строкам либо одновременно, либо последовательно. Например, при приведении симметрической матрицы

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

с помощью отражений к трехдиагональному виду нужно получить нули на месте элементов $a_{31} = a_{13} = 4$. Для этого положим $x = (3,\ 4)^{\mathrm{T}}$, $z = (1,\ 0)^{\mathrm{T}}$, $v = x + |x|z = x + 5z = (8,\ 4)^{\mathrm{T}}$ и построим матрицу

$$H_1 = E - 2\frac{vv^{\mathrm{T}}}{|v|^2} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \frac{1}{40}\begin{pmatrix} 8 \\ 4 \end{pmatrix}(8,\ 4) = \frac{1}{5}\begin{pmatrix} -3 & -4 \\ -4 & 3 \end{pmatrix}.$$

Найденную матрицу вложим в правый нижний угол единичной матрицы третьего порядка:

$$U_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{3}{5} & -\frac{4}{5} \\ 0 & -\frac{4}{5} & \frac{3}{5} \end{pmatrix}.$$

Теперь найдем матрицу

$$A_1 = U_1 A U_1^{-1} = U_1 A U_1^{\mathrm{T}} = \frac{1}{25}\begin{pmatrix} 25 & -125 & 0 \\ -125 & 49 & 7 \\ 0 & 7 & 1 \end{pmatrix}.$$

Матрица A_1 уже трехдиагональная, причем подобна исходной матрице A.

Здесь мы применили одно и то же отражение одновременно и к первому столбцу, и к первой строке матрицы A, в результате чего получили трехдиагональную матрицу $A_1 = U_1 A U^{-1} = U_1 A U_1^{\mathrm{T}}$, подобную данной. Так поступают и при приведении произвольной квадратной матрицы A к подобной матрице почти треугольного вида [3, с. 182–184]. В этом случае на i-м шаге получают матрицу $A_i = H_i A_{i-1} H_i^{-1} = H_i A_{i-1} H_i^{\mathrm{T}}$, $i = 1, 2, \ldots, k$, подобную матрице $A_0 = A$. Процесс продолжают до получения матрицы нужного вида. Например, чтобы привести матрицу

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 2 & -32 & 1 & 1 \\ 2 & -35 & 1 & 1 \end{pmatrix}$$

к подобной матрице почти треугольного вида, сначала получим нули в позициях a_{31} и a_{41}. Для этого возьмем $x = (1,\ 2,\ 2)^{\mathrm{T}}$, $z = (1,\ 0,\ 0)^{\mathrm{T}}$, $v = x + |x|z = (4,\ 2,\ 2)^{\mathrm{T}}$ и построим матрицу

$$H_1 = E - 2\frac{vv^{\mathrm{T}}}{|v|^2} = E - \frac{1}{24}\begin{pmatrix} 4 \\ 2 \\ 2 \end{pmatrix}(4,\ 2,\ 2) = \frac{1}{3}\begin{pmatrix} -1 & -2 & -2 \\ -2 & 2 & -1 \\ -2 & -1 & 2 \end{pmatrix}.$$

Вложим матрицу H_1 в правый нижний угол единичной матрицы четвертого порядка. Тогда получим матрицу

$$U_1 = \frac{1}{3}\begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & -1 & -2 & -2 \\ 0 & -2 & 2 & -1 \\ 0 & -2 & -1 & 2 \end{pmatrix}.$$

Далее найдем матрицу

$$A_1 = U_1 A U_1^{-1} = \begin{pmatrix} 1 & -\frac{5}{3} & -1 & -1 \\ -3 & -13 & -\frac{89}{3} & -\frac{91}{3} \\ 0 & 3 & \frac{23}{3} & \frac{19}{3} \\ 0 & 4 & \frac{29}{3} & \frac{25}{3} \end{pmatrix}.$$

Теперь получим нуль в позиции a_{42}. Для этого возьмем $x = (3,\ 4)^{\text{т}}$, $z = (1,\ 0)^{\text{т}}$, $v = x + |x|z = (8,\ 4)^{\text{т}}$ и построим матрицу

$$H_2 = E - 2\frac{vv^{\text{т}}}{|v|^2} = \frac{1}{5}\begin{pmatrix} -3 & -4 \\ -4 & 3 \end{pmatrix}.$$

Вложим матрицу H_2 в правый нижний угол единичной матрицы четвертого порядка. Тогда получим матрицу

$$U_2 = \frac{1}{5}\begin{pmatrix} 5 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & -3 & -4 \\ 0 & 0 & -4 & 3 \end{pmatrix}.$$

Затем найдем матрицу

$$A_2 = U_2 A_1 U_2^{-1} = \begin{pmatrix} 1 & -\frac{5}{3} & \frac{7}{5} & \frac{7}{5} \\ -3 & -3 & \frac{631}{15} & \frac{83}{15} \\ 0 & -5 & \frac{1183}{75} & \frac{269}{75} \\ 0 & 0 & -27 & \frac{92}{75} \end{pmatrix}.$$

Матрица A_2 почти треугольная и подобна данной.

Если не требуется, чтобы упрощенная матрица была подобна исходной, то отражения можно применять последовательно к столбцам и строкам. Для примера приведем матрицу

$$A = \frac{1}{75}\begin{pmatrix} 50 & 60 & -45 \\ 50 & -99 & -57 \\ 25 & -57 & 24 \end{pmatrix}$$

к двухдиагональной форме. Сначала получим нули в позициях a_{21} и a_{31}. С этой целью возьмем $x = (50,\ 50,\ 25)^{\text{т}}$, $z = (1,\ 0,\ 0)^{\text{т}}$. Построим вектор $v = x - |x|z = (-25,\ 50,\ 25)^{\text{т}}$ и матрицу

$$H_1 = E - 2\frac{vv^{\text{т}}}{|v|^2} = \frac{1}{3}\begin{pmatrix} 2 & 2 & 1 \\ 2 & -1 & -2 \\ 1 & -2 & 2 \end{pmatrix}.$$

Затем найдем матрицу

$$A_1 = H_1 A = \begin{pmatrix} 1 & -\frac{3}{5} & -\frac{4}{5} \\ 0 & \frac{37}{25} & -\frac{9}{25} \\ 0 & \frac{16}{25} & \frac{13}{75} \end{pmatrix}.$$

Теперь получим нуль в позиции a_{13}. Для этого возьмем векторы $x = \left(-\frac{3}{5},\ -\frac{4}{5}\right)^{\text{т}}$ и $z = (1,\ 0)^{\text{т}}$. Построим вектор $v = x - |x|z = \left(-\frac{8}{5},\ -\frac{4}{5}\right)^{\text{т}}$ и матрицу

$$H_2 = E - 2\frac{vv^{\text{т}}}{|v|^2} = \begin{pmatrix} -\frac{3}{5} & -\frac{4}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{pmatrix}.$$

Вложим матрицу H_2 в правый нижний угол единичной матрицы третьего порядка. Тогда получим матрицу

$$V_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{3}{5} & -\frac{4}{5} \\ 0 & -\frac{4}{5} & \frac{3}{5} \end{pmatrix}$$

и найдем матрицу

$$A_2 = A_1 V_1 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & -\frac{3}{5} & -\frac{7}{5} \\ 0 & -\frac{4}{5} & -\frac{1}{5} \end{pmatrix}.$$

Осталось получить нуль в позиции a_{32}. Для этого возьмем $x = \left(-\frac{3}{5}, \frac{4}{5}\right)^{\mathrm{T}}$ и так же, как и матрицу V_1, построим матрицу

$$U_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{3}{5} & -\frac{4}{5} \\ 0 & -\frac{4}{5} & \frac{3}{5} \end{pmatrix}$$

и найдем матрицу

$$A_3 = U_2 A_2 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Матрица A_3 уже нужного вида — двухдиагональная.

8.11. Произвольные линейные операторы в евклидовом пространстве

Для произвольного линейного оператора φ, действующего в евклидовом пространстве E, как установлено выше, операторы $\varphi^*\varphi$ и $\varphi\varphi^*$ являются неотрицательными, из них извлекаются корни любой степени. Эти факты лежат в основе следующего утверждения.

Теорема 8.33. *Любой невырожденный линейный оператор φ, действующий в евклидовом пространстве, может быть представлен в виде $\varphi = \psi\eta$, где ψ — симметрический, а η — ортогональный операторы. Такое представление единственно.*

▷ Для невырожденного линейного оператора φ существует симметрический оператор $\psi = \sqrt{\varphi\varphi^*}$. Для этого оператора верно равенство $\psi^2 = \varphi\varphi^*$. Так как оператор φ невырожденный, то и оператор φ^* согласно теореме 8.11 невырожденный. Поэтому $\psi^2(\varphi^*)^{-1} = \varphi$. Положим $\eta = \psi(\varphi^*)^{-1}$. Тогда $\varphi = \psi\eta$, и остается доказать, что оператор η ортогональный.

Учитывая свойства сопряженных операторов и симметричность оператора ψ, заключаем, что

$$\eta^* = \left(\psi(\varphi^*)^{-1}\right)^* = \left((\varphi^*)^{-1}\right)^*\psi^* = \varphi^{-1}\psi.$$

Следовательно,

$$\eta^*\eta = \varphi^{-1}\psi\psi(\varphi^*)^{-1} = \varphi^{-1}\varphi\varphi^*(\varphi^*)^{-1} = \varepsilon,$$

поскольку $\psi\psi = \varphi\varphi^*$. Равенство $\eta^*\eta = \varepsilon$ означает, что оператор η ортогональный.

Докажем, что разложение $\varphi = \psi\eta$ на симметрический и ортогональный операторы единственно. Из такого представления получаем:

$$\varphi\varphi^* = (\psi\eta)(\psi\eta)^* = \psi\eta\eta^*\psi = \psi^2,$$

так как $\eta\eta^* = \varepsilon$ в силу ортогональности оператора η. Таким образом, если $\varphi = \psi\eta$, то симметрический оператор ψ является квадратным корнем оператора $\varphi\varphi^*$, а потому определен однозначно. Так как φ невырожден по условию, η невырожден в силу своей ортогональности, то и оператор ψ невырожден. Но тогда оператор $\eta = \psi^{-1}\varphi$ определен однозначно. ▶

Замечание. Аналогично можно доказать, что для невырожденного оператора φ существует единственное разложение $\varphi = \eta_1\psi_1$ в произведение ортогонального оператора η_1 и неотрицательного оператора ψ_1. Впрочем, это разложение можно получить из утверждения теоремы, так как представление $\varphi = \eta_1\psi_1$ равносильно разложению $\varphi^* = \psi_1\eta_1^*$, которое в силу доказанной теоремы для оператора φ^* существует и единственно.

Отметим также, что разложение $\varphi = \psi\eta$, равно как и разложение $\varphi = \eta_1\psi_1$, существует для любого оператора, а не только для невырожденного. Однако для вырожденного оператора такое разложение уже не будет единственным.

8.12. Сопряженные операторы в унитарном пространстве

Все сказанное о линейных операторах в линейном пространстве (см. гл. 5) сохраняется и в случае, когда линейное пространство является унитарным. Наличие в унитарном пространстве скалярного произведения позволяет выделить несколько важных классов линейных операторов, характеризующихся дополнительными свойствами. Остановимся на понятии „сопряженный оператор в унитарном пространстве“, а также опишем классы эрмитовых, унитарных и нормальных операторов.

Линейный оператор φ^*, действующий в унитарном пространстве U, называют ***сопряженным с линейным оператором*** φ, если для любых векторов $x, y \in U$ выполняется равенство

$$(\varphi x, y) = (x, \varphi^* y)\,. \tag{8.47}$$

Нетрудно увидеть, что данное определение дословно повторяет определение сопряженного оператора в евклидовом пространстве. В унитарные пространства переносится не только это определение, но и все сказанное о сопряженных операторах в евклидовом пространстве (см. разд. 8.8). Повторяются не только формулировки утверждений, но и их доказательства. Поэтому ограничимся краткими формулировками и приведем дополнительные примеры, иллюстрирующие свойства операторов в случае унитарного пространства.

Для любого линейного оператора φ, действующего в унитарном пространстве, существует, и притом единственный, сопряженный оператор φ^*. При этом оператор φ является сопряженным с оператором φ^*, так что можно говорить, что операторы φ и φ^* взаимно сопряженные.

Сопряженные операторы обладают следующими свойствами:

1. $(\varphi^*)^* = \varphi$.

2. $(\varphi+\psi)^* = \varphi^* + \psi^*$.
3. $(\alpha\varphi)^* = \alpha\varphi^*$, α — любое комплексное.
4. $(\varphi\psi)^* = \psi^*\varphi^*$.

Матрицы A и A_1 операторов φ и φ^* в произвольном базисе e унитарного пространства U связаны соотношением

$$\overline{A_1} = \overline{\Gamma}^{-1} A^{\mathrm{T}} \Gamma, \tag{8.48}$$

где Γ — матрица Грама базиса e. В частности, если базис e ортонормированный, то

$$A_1 = \overline{A}^{\mathrm{T}} = A^*, \tag{8.49}$$

т.е. в ортонормированном базисе матрицей сопряженного оператора является сопряженная матрица, а именно матрица, полученная в результате транспонирования и комплексного сопряжения.

Пример 8.20. Пусть линейный оператор φ в базисе $e_1 = (1,\ 1)^{\mathrm{T}}$, $e_2 = (0,\ 1)^{\mathrm{T}}$ имеет матрицу

$$A = \begin{pmatrix} 1 & i \\ 2i & 1 \end{pmatrix}.$$

Найти матрицу A_1 сопряженного оператора φ^* в этом же базисе, если координаты векторов e_1, e_2 заданы в ортонормированном базисе.

Решение. Так как векторы e_1, e_2 заданы координатами в ортонормированном базисе, можно вычислить скалярные произведения базисных векторов

$$(e_1, e_1) = 2, \quad (e_1, e_2) = (e_2, e_1) = (e_2, e_2) = 1$$

и составить матрицу Грама базиса e_1, e_2:

$$\Gamma = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

Зная матрицу Грама, по формуле (8.48) можем вычислить:

$$\overline{A_1} = \overline{\Gamma}^{-1} A^{\mathrm{T}} \Gamma = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 1 & i \\ 2i & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & i \\ 2i & 1 \end{pmatrix}.$$

Остается перейти к комплексно сопряженной матрице, заменяя в матрице $\overline{A_1}$ каждый элемент комплексно сопряженным:

$$A_1 = \begin{pmatrix} 1 & -i \\ -2i & 1 \end{pmatrix}.$$

Задачу можно решить другим способом — путем перехода к ортонормированному базису. Пусть f_1, f_2 — ортонормированный базис, в котором заданы координаты векторов e_1, e_2. Матрица T перехода от f_1, f_2 к e_1, e_2 имеет вид:

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

С помощью матрицы перехода можно найти матрицу $\widetilde{A}$ оператора φ в ортонормированном базисе f_1, f_2:

$$\widetilde{A} = TAT^{-1} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & i \\ 2i & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1-i & i \\ i & 1+i \end{pmatrix}.$$

Теперь по формуле (8.49) определяем матрицу $\widetilde{A}_1$ оператора φ^* в базисе f_1, f_2:

$$\widetilde{A}_1 = \begin{pmatrix} 1+i & -i \\ -i & 1-i \end{pmatrix}.$$

Наконец, с помощью матрицы перехода T возвращаемся в базис e_1, e_2:

$$A_1 = T^{-1}\widetilde{A}_1 T = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1+i & -i \\ -i & 1-i \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -i \\ -2i & 1 \end{pmatrix}.$$

Если $\lambda_1, \lambda_2, \ldots, \lambda_k$ — собственные значения оператора φ в унитарном пространстве, то $\overline{\lambda_1}, \overline{\lambda_2}, \ldots, \overline{\lambda_k}$ — собственные значения оператора φ^*. Как и в евклидовом пространстве, область значений оператора φ^*, сопряженного с оператором φ в унитарном пространстве, совпадает с ортогональным дополнением к ядру оператора φ.

Если некоторое подпространство L инвариантно относительно оператора φ, то ортогональное дополнение $L^\perp$ этого подпространства инвариантно относительно сопряженного оператора φ^*.

8.13. Эрмитовы операторы

Линейный оператор φ, действующий в унитарном пространстве, называют ***эрмитовым***, или ***самосопряженным***, если он совпадает со своим сопряженным оператором, т.е. если $\varphi^* = \varphi$. Линейный оператор φ является эрмитовым тогда и только тогда, когда для любых векторов x и y

$$(\varphi x, y) = (x, \varphi y).$$

Если φ — эрмитов оператор, то при любом векторе x скалярное произведение $(\varphi x, x)$ является действительным числом. В самом деле, для эрмитова оператора φ верно равенство $(\varphi x, x) = (x, \varphi x)$. В то же время, учитывая свойства скалярного произведения в комплексном линейном пространстве, имеем: $(x, \varphi x) = \overline{(\varphi x, x)}$. Поэтому

$$(\varphi x, x) = \overline{(\varphi x, x)},$$

что возможно лишь в случае, когда $(\varphi x, x)$ — действительное число.

Выясним, какие матрицы являются матрицами эрмитовых операторов в ортонормированном базисе.

Теорема 8.34. *Матрицей любого эрмитова оператора в ортонормированном базисе является эрмитова матрица. Если линейный оператор, действующий в унитарном пространстве, в каком-либо ортонормированном базисе имеет эрмитову матрицу, то этот оператор эрмитов.*

▷ Доказательство этой теоремы аналогично доказательству теоремы 8.15. Действительно, согласно формуле (8.49) в ортонормированном базисе матрица A_1 сопряженного оператора φ^* связана с матрицей A оператора φ соотношением $A_1 = A^*$. Если оператор эрмитов, то $\varphi^* = \varphi$, чему в координатной записи в ортонормированном базисе соответствует равенство $A^* = A$. Следовательно, условие, что оператор эрмитов, равносильно условию, что его матрица эрмитова. ▶

На эрмитовы операторы распространяются и другие свойства симметрических операторов (см. разд. 8.9). Приведем эти свойства.

1. Все собственные числа эрмитова оператора являются действительными числами. Собственные векторы, отвечающие различным собственным числам эрмитова оператора, ортогональны.

2. Любой линейный, в частности эрмитов, оператор имеет одномерное инвариантное подпространство.

3. Если L — инвариантное подпространство эрмитова оператора, то ортогональное дополнение $L^{\perp}$ — также инвариантное подпространство этого оператора.

4. Для любого эрмитова оператора существует ортонормированный базис, состоящий из собственных векторов этого оператора.

5. Произведение эрмитовых операторов φ и ψ является эрмитовым оператором тогда и только тогда, когда эти операторы перестановочны, т.е. $\varphi\psi = \psi\varphi$.

6. Эрмитов оператор является оператором простой структуры: в некотором ортонормированном базисе матрица эрмитова оператора является диагональной.

Последнее из перечисленных свойств равносильно следующему утверждению для матриц: ***любая эрмитова матрица A имеет каноническое разложение $A = T\Lambda T^{-1}$, в котором Λ — диагональная матрица, а T — унитарная матрица.*** Построить такое разложение можно так же, как и в случае симметрических матриц (см. разд. 8.9).

Пример 8.21. Построить каноническое разложение эрмитовой матрицы

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

с унитарной трансформирующей матрицей T.

Решение. Характеристический многочлен

$$|A-\lambda E|=\begin{vmatrix} 1-\lambda & 1+i \\ 1-i & -\lambda \end{vmatrix}=\lambda^2-\lambda-2$$

матрицы A имеет корни $\lambda_1=2$ и $\lambda_2=-1$. Поэтому диагональной матрицей в каноническом разложении является матрица

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

При $\lambda=2$ система $(A-\lambda E)x=0$ имеет вид:

$$\begin{cases} -x_1+(1+i)x_2=0, \\ (1-i)x_1-2x_2=0. \end{cases}$$

Ее фундаментальная система решений состоит из одного вектора, например, $b_1=(1+i,\ 1)^{\mathrm{T}}$. Нормируя его, получаем: $e_1=\left(\frac{1+i}{\sqrt{3}},\ \frac{1}{\sqrt{3}}\right)^{\mathrm{T}}$. При $\lambda=-1$ аналогичным образом определяем вектор $e_2=\left(\frac{1+i}{\sqrt{6}},\ -\frac{2}{\sqrt{6}}\right)^{\mathrm{T}}$. Из найденных векторов составляем трансформирующую матрицу

$$T=\begin{pmatrix} \frac{1+i}{\sqrt{3}} & \frac{1+i}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} \end{pmatrix}$$

и записываем искомое каноническое разложение:

$$A=T\Lambda T^{-1}=T\Lambda T^{*}=\begin{pmatrix} \frac{1+i}{\sqrt{3}} & \frac{1+i}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} \frac{1-i}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1-i}{\sqrt{6}} & -\frac{2}{\sqrt{6}} \end{pmatrix}.$$

Замечание. Численное построение канонического разложения можно проводить методом вращения (см. разд. 11.1).

Эрмитов оператор φ называют ***неотрицательным*** (***положительно определенным***), если для любого вектора $x\neq 0$ выполняется неравенство $(\varphi x,x)\geqslant 0$ ($(\varphi x,x)>0$). Если оператор φ неотрицательный (положительно определенный), то пишут $\varphi \geqslant 0$ ($\varphi > 0$).

Эрмитов оператор является неотрицательным (положительно определенным) тогда и только тогда, когда все его собственные значения неотрицательные (положительные). Для любого неотрицательного эрмитова оператора φ и любого натурального числа m существует единственный неотрицательный эрмитов оператор f, такой, что $f^m = \varphi$. Оператор f называют ***арифметическим корнем m-й степени из оператора*** φ.

Для любого оператора φ, действующего в унитарном пространстве U, операторы $\varphi^*\varphi$ и $\varphi\varphi^*$ являются эрмитовыми неотрицательными. Этот факт распространяется на общий случай линейного оператора, действующего из одного унитарного пространства U в другое унитарное пространство V.

Свойства эрмитовых операторов легко переформулировать для эрмитовых матриц. Так, эрмитову матрицу A называют неотрицательной (положительно определенной), если для любого столбца $x \neq 0$ выполняется неравенство $x^{\mathrm{T}}Ax \geqslant 0$ ($x^{\mathrm{T}}Ax > 0$). В этом случае пишут $A \geqslant 0$ ($A > 0$). Эрмитова матрица неотрицательна (положительно определена) тогда и только тогда, когда все ее характеристические числа неотрицательные (положительные).

Арифметическим корнем m-й степени $\sqrt[m]{A}$ ***из неотрицательной эрмитовой матрицы*** A называют такую матрицу B, что $B^m = A$. Вычислить $\sqrt[m]{A}$ можно, построив каноническое разложение $A = T\Lambda T^{-1}$ матрицы A и использовав формулу

$$\sqrt[m]{A} = T\,\sqrt[m]{\Lambda}T^{-1},$$

где

$$\sqrt[m]{\Lambda} = \begin{pmatrix} \lambda_1^{1/m} & & 0 \\ & \ddots & \\ 0 & & \lambda_n^{1/m} \end{pmatrix}.$$

Для любой, вообще говоря, прямоугольной матрицы A матрицы A^*A и AA^* являются эрмитовыми неотрицательными и имеют однозначно определенные арифметические корни любой натуральной степени. В частности, определены арифметические корни $\sqrt{A^*A}$ и $\sqrt{AA^*}$.

8.14. Унитарные операторы

Линейный оператор φ, действующий в унитарном пространстве, называют ***унитарным***, если он не изменяет скалярного произведения векторов, т.е. если для любых векторов x и y

$$(\varphi x, \varphi y) = (x, y).$$

Непосредственно из определения вытекает, что унитарный оператор сохраняет длины векторов. Унитарный оператор переводит любую ортонормированную систему векторов в ортонормированную и, в частности, ортонормированный базис в ортонормированный. Это свойство унитарного оператора критериальное: если линейный оператор φ переводит некоторый ортонормированный базис в ортонормированный, то этот оператор унитарный.

Унитарный оператор по своим свойствам близок ортогональному оператору в евклидовом пространстве. Утверждения, установленные для ортогонального оператора (см. разд. 8.10), переносятся на унитарные операторы. Приведем основные факты без доказательства:

1) линейный оператор φ является унитарным тогда и только тогда, когда $\varphi^*\varphi = \varepsilon$;

2) унитарный оператор φ невырожденный и $\varphi^{-1} = \varphi^*$;

3) собственные значения унитарного оператора по модулю равны единице;

4) собственные векторы унитарного оператора, принадлежащие различным собственным значениям, ортогональны;

5) матрица унитарного оператора в ортонормированном базисе является унитарной;

6) если матрица линейного оператора в некотором ортонормированном базисе унитарна, то этот оператор унитарный;

7) произведение унитарных операторов является унитарным оператором;

8) для любого унитарного оператора, действующего в унитарном пространстве, существует ортонормированный базис, состоящий из собственных векторов этого оператора;

9) любой унитарный оператор является оператором простой структуры.

Из последнего свойства вытекает, что унитарная матрица имеет каноническое разложение с унитарной трансформирующей матрицей и с диагональной матрицей, у которой все диагональные элементы по модулю равны единице. Каноническое разложение для матрицы унитарного оператора строится так же, как и для матрицы любого оператора простой структуры (см. разд. 5.8 и 8.9). Унитарность трансформирующей матрицы обеспечивается дополнительным ортонормированием полученного базиса из собственных векторов. Поясним это на примере.

Пример 8.22. Построить каноническое разложение унитарной матрицы

$$A=\begin{pmatrix} \frac{4}{5} & -\frac{3i}{5} \\ -\frac{3i}{5} & \frac{4}{5} \end{pmatrix}$$

с унитарной трансформирующей матрицей T.

Решение. Характеристический многочлен

$$|A-\lambda E|=\begin{vmatrix} \frac{4}{5}-\lambda & -\frac{3i}{5} \\ -\frac{3i}{5} & \frac{4}{5}-\lambda \end{vmatrix}=\lambda^2-\frac{8}{5}\lambda+1$$

матрицы A имеет корни $\lambda_1=\frac{4+3i}{5}$, $\lambda_2=\frac{4-3i}{5}$. Поэтому искомой диагональной матрицей Λ является матрица

$$\Lambda=\begin{pmatrix} \frac{4+3i}{5} & 0 \\ 0 & \frac{4-3i}{5} \end{pmatrix}.$$

При $\lambda = \dfrac{4+3i}{5}$ система $(A - \lambda E)x = 0$ имеет вид:

$$\begin{cases} -\dfrac{3i}{5}x_1 - \dfrac{3i}{5}x_2 = 0, \\ -\dfrac{3i}{5}x_1 - \dfrac{3i}{5}x_2 = 0. \end{cases}$$

Фундаментальная система решений состоит из одного решения, например, $b_1 = (1,\ -1)^{\mathrm{T}}$. Нормируя это решение, получим: $e_1 = \left(\dfrac{1}{\sqrt{2}},\ -\dfrac{1}{\sqrt{2}}\right)^{\mathrm{T}}$.

При $\lambda = \dfrac{4-3i}{5}$ аналогичным образом получаем вектор $e_2 = \left(\dfrac{1}{\sqrt{2}},\ \dfrac{1}{\sqrt{2}}\right)^{\mathrm{T}}$. Из столбцов координат векторов e_1 и e_2 составим матрицу

$$T = \begin{pmatrix} \dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{2}} \\ -\dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{2}} \end{pmatrix}$$

и запишем искомое каноническое разложение

$$A = T\lambda T^{-1} = \begin{pmatrix} \dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{2}} \\ -\dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \dfrac{4+3i}{5} & 0 \\ 0 & \dfrac{4-3i}{5} \end{pmatrix} \begin{pmatrix} \dfrac{1}{\sqrt{2}} & -\dfrac{1}{\sqrt{2}} \\ \dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{2}} \end{pmatrix}.$$

В приложениях наиболее часто используют ***элементарные унитарные матрицы***

$$T_{ij} = \begin{pmatrix} 1 & & & \vdots & & & & \vdots & & & \\ & \ddots & & \vdots & & & & \vdots & & & \\ & & 1 & \vdots & & & & \vdots & & & \\ \cdots & & & \cos\alpha & \cdots & & & -e^{i\psi}\sin\alpha & \cdots & & \\ & & & \vdots & 1 & & & \vdots & & & \\ & & & \vdots & & \ddots & & \vdots & & & \\ & & & \vdots & & & 1 & \vdots & & & \\ \cdots & & & e^{-i\psi}\sin\alpha & \cdots & & & \cos\alpha & \cdots & & \\ & & & \vdots & & & & \vdots & 1 & & \\ & & & \vdots & & & & \vdots & & \ddots & \\ & & & \vdots & & & & \vdots & & & 1 \end{pmatrix} \begin{array}{l} \\ \\ \\ i\text{-я строка} \\ \\ \\ \\ j\text{-я строка} \\ \\ \\ \\ \end{array},$$

являющиеся обобщением матриц простого вращения, а также матрицы отражения — аналоги матриц отражения в евклидовом пространстве. Как и в евклидовом пространстве, матрица отражения определяется равенством

$$H = E - \frac{2vv^*}{|v|^2},$$

где v — столбец координат определяющего вектора отражения. Матрица H не только унитарная, но и эрмитова, т.е. удовлетворяет условиям $H = H^{-1} = H^*$.

Матрицы T_{ij} и H применяются так же, как и матрицы вращений и отражений в евклидовом пространстве. Например, чтобы с помощью отражения обратить в нуль вторую и третью координаты вектора $x = (0,\ 3i,\ 4i)^{\mathrm{T}}$, найдем (см. разд. 8.10) $|x| = 5$ и положим $z = (1,\ 0,\ 0)^{\mathrm{T}}$, $v = x - |x|z = (-5,\ 3i,\ 4i)^{\mathrm{T}}$. Далее найдем $|v|^2 = 50$, составим матрицу отражения

$$H = E - \frac{2}{50}\begin{pmatrix} -5 \\ 3i \\ 4i \end{pmatrix}(-5,\ 3i,\ 4i) = \frac{1}{25}\begin{pmatrix} 0 & -15i & -20i \\ 15i & 16 & -12 \\ 20i & -12 & 9 \end{pmatrix}$$

и подсчитаем:

$$x' = Hx = \frac{1}{25}\begin{pmatrix} 0 & -15i & -20i \\ 15i & 16 & -12 \\ 20i & -12 & 9 \end{pmatrix}\begin{pmatrix} 0 \\ 3i \\ 4i \end{pmatrix} = \begin{pmatrix} 5 \\ 0 \\ 0 \end{pmatrix}.$$

Пример на применение матриц T_{ij} приведен в разд. 11.1.

8.15. Нормальные операторы

Линейный оператор φ, действующий в унитарном пространстве U, называют ***нормальным***, если он перестановочен со своим сопряженным оператором φ^*, т.е. если выполняется соотношение

$$\varphi\varphi^* = \varphi^*\varphi.$$

В ортонормированном базисе матрицей нормального оператора является матрица A, перестановочная со своей сопряженной матрицей A^*, т.е. удовлетворяющая условию $AA^* = A^*A$. Действительно, если оператор φ в ортонормированном базисе имеет матрицу A, то сопряженный оператор φ^* в том же базисе имеет матрицу A^*. Поэтому равенство $\varphi\varphi^* = \varphi^*\varphi$ в матричной записи имеет вид $AA^* = A^*A$.

Матрицы, перестановочные со своей сопряженной матрицей, называют ***нормальными***.

К нормальным операторам относятся эрмитовы и унитарные операторы. Если φ — эрмитов оператор, то $\varphi^* = \varphi$ и условие перестановочности тривиально. Если φ — унитарный оператор, то $\varphi^* = \varphi^{-1}$, а условие перестановочности трансформируется в условие перестановочности оператора со своим обратным.

Основным свойством нормального оператора является существование ортонормированного базиса из собственных векторов этого оператора. Чтобы прийти к этому заключению, докажем два вспомогательных результата.

Лемма 8.1. *Пусть φ — линейный оператор, действующий в унитарном пространстве U. Собственное подпространство L_λ линейного оператора φ, отвечающее собственному значению λ, инвариантно относительно любого оператора ψ, перестановочного с оператором φ.*

$\triangleright$ Надо показать, что если $x \in L_\lambda$, то $\psi x \in L_\lambda$. Условие $x \in L_\lambda$ означает, что $\varphi x = \lambda x$. Отсюда следует, что

$$\varphi(\psi x) = \psi\varphi x = \psi(\lambda x) = \lambda\psi x,$$

т.е. $\psi x \in L_\lambda$. $\blacktriangleright$

Лемма 8.2. *Любые два перестановочных линейных оператора φ и ψ в унитарном пространстве имеют общий собственный вектор.*

$\triangleright$ Рассмотрим собственное подпространство L_λ оператора φ, отвечающее какому-либо собственному значению λ. Согласно лемме 8.1 подпространство L_λ инвариатно относительно оператора ψ, перестановочного с φ. Поэтому на линейном пространстве L_λ определен оператор, являющийся сужением на L_λ оператора ψ. Этот оператор, как и всякий линейный оператор в комплексном линейном пространстве, имеет собственный вектор x_0. Этот вектор будет собственным и для оператора ψ, поскольку действие ψ на L_λ совпадает с действием его сужения. Он также будет собственным и для оператора φ, поскольку в L_λ каждый ненулевой вектор является собственным для φ. ▶

Теорема 8.35. *Пусть в унитарном пространстве U действует линейный оператор φ. Для того чтобы в U существовал ортонормированный базис из собственных векторов оператора φ, необходимо и достаточно, чтобы оператор φ был нормальным.*

▷ Пусть оператор φ имеет ортонормированный базис e из собственных векторов. В этом базисе оператор φ имеет диагональную матрицу

$$\Lambda = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Поскольку базис e ортонормированный, матрицей оператора φ^* в этом базисе является сопряженная матрица

$$\Lambda^* = \overline{\Lambda} = \begin{pmatrix} \overline{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \overline{\lambda_n} \end{pmatrix}.$$

Очевидно, что две диагональные матрицы Λ и Λ^* перестановочны. Следовательно, операторы φ и φ^* также перестановочны.

Пусть оператор φ нормальный, т.е. $\varphi\varphi^* = \varphi^*\varphi$. Согласно лемме 8.2 операторы φ и φ^* имеют общий собственный вектор, т.е. такой вектор e_1, что $\varphi e_1 = \lambda e_1$ и $\varphi^* e_1 = \mu e_1$. Можно считать, что вектор e_1 единичный. Подпространство L_1, порожденное вектором e_1, является инвариантным и относительно φ, и относительно φ^*. Поэтому инвариантным относительно каждого из этих двух операторов является ортогональное дополнение $L_1^\perp$ (см. разд. 8.12). Доказательство можно закончить по методу математической индукции по размерности унитарного пространства. Утверждение очевидно для одномерных пространств. Предполагая, что утверждение теоремы верно для пространств размерности $n-1$, делаем заключение, что в подпространстве $L_1^\perp$ существует ортонормированный базис из собственных векторов оператора φ. Добавив к этому базису вектор e_1, получим ортонормированный базис во всем унитарном пространстве U, состоящий из собственных векторов оператора φ. ►

Поскольку эрмитовы и унитарные операторы являются нормальными операторами, то из доказанной теоремы следует, что любой эрмитов или унитарный оператор имеет ортонормированный базис из собственных векторов. Кроме того, из существования ортонормированного базиса из собственных векторов вытекает, что собственный вектор нормального оператора φ является в то же время и собственным вектором сопряженного оператора φ^*, причем собственные значения этих операторов, отвечающие данному собственному вектору, являются комплексно сопряженными. В частности, ортонормированный базис из собственных векторов нормального оператора φ в то же время является базисом из собственных векторов оператора φ^*.

Из теоремы 8.35 следует, что нормальный оператор является оператором простой структуры. Обратное утверждение неверно, так как оператор простой структуры, хотя и имеет базис из собственных векторов, может не иметь среди таких базисов ортонормированного. Но можно утверждать, что если в комплексном линейном пространстве U действует линейный оператор простой структуры, то в U можно ввести скалярное произведение так, что этот оператор будет нормальным.

Теорему 8.35 можно переформулировать для матриц. В результате мы придем к выводу, что нормальные матрицы, и только они, приводятся к диагональному виду с помощью унитарной трансформирующей матрицы. Это утверждение можно также сформулировать следующим образом. Для того чтобы матрица A имела каноническое разложение $A = T\Lambda T^{-1}$ с унитарной матрицей T, необходимо и достаточно, чтобы эта матрица была нормальной. Такое разложение строится по общей схеме, но при этом полученный базис нужно дополнительно ортонормировать с тем, чтобы прийти к унитарной матрице T.

8.16. Произвольные линейные операторы в унитарном пространстве

На произвольные линейные операторы в унитарном пространстве переносится теорема 8.33 об операторах в евклидовом пространстве.

Теорема 8.36. *Для любого линейного оператора φ, действующего в унитарном пространстве U, существует разложение $\varphi = \psi\eta$ на эрмитов оператор ψ и унитарный оператор η. Если оператор φ невырожденный, то такое разложение единственное.*

Доказательство этой теоремы дословно повторяет доказательство теоремы 8.33 и поэтому не приводится.

8.17. QR-разложение матрицы

QR-разложением квадратной матрицы A называют ее представление в виде произведения $A = QR$ с ортогональной (унитарной) матрицей Q и верхней треугольной матрицей R. QR-разложение матрицы можно построить с помощью ортогонализации и ортонормирования ее столбцов, а также с помощью ортогональных (унитарных) преобразований.

Пусть в матрице $A = (a_1, a_2, \ldots, a_n)$ столбцы $a_1, a_2, \ldots, a_n$ линейно независимые. Ортогонализируем эту систему векторов, проводя процесс ортогонализации. Затем нормируем каждый вектор полученной ортогональной системы векторов. В результате придем к ортонормированной системе векторов

$$\begin{aligned} q_1 &= u_{11}a_1, \\ q_2 &= u_{12}a_1 + u_{22}a_2, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ q_n &= u_{1n}a_1 + u_{2n}a_2 + \ldots + u_{nn}a_n. \end{aligned}$$

В матричной записи это дает равенство

$$Q = AU,$$

где $Q = (q_1, q_2, \ldots, q_n)$ — ортогональная (унитарная) матрица,

$$U = \begin{pmatrix} u_{11} & u_{12} & \ldots & u_{1,n-1} & u_{1n} \\ 0 & u_{22} & \ldots & u_{2,n-1} & u_{2n} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 & u_{nn} \end{pmatrix}$$

— верхняя треугольная матрица.

Отсюда получается QR-разложение $A = QR$ с ортогональной (унитарной) матрицей Q и верхней треугольной матрицей $R = U^{-1}$.

QR-разложение можно строить с помощью вращений. Для этого сначала следует матрицу A с помощью вращений привести к треугольному виду. Пусть при этом использованы вращения с матрицами $T_1, T_2, \ldots, T_k$ и получена треугольная матрица $T_k \ldots T_2 T_1 A = R$. Тогда $A = QR$ будет искомым QR-разложением с ортогональной (унитарной) матрицей $Q = (T_k \ldots T_2 T_1)^{-1} = T_1^* T_2^* \ldots T_k^*$.

QR-разложение можно также строить с помощью отражений. Для этого матрицу A с помощью отражений необходимо привести к треугольному виду. Пусть при этом использованы отражения с матрицами U_1, U_2, ..., U_k и получена треугольная матрица $U_k \dots U_2 U_1 A = R$. Тогда $A = QR$ будет искомым QR-разложением с ортогональной (унитарной) матрицей $Q = (U_k \dots U_2 U_1)^{-1} = U_1^* U_2^* \dots U_k^*$.

Пример 8.23. Построить QR-разложение матрицы

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

Решение. К столбцам $a_1 = (3,\ 0,\ 4)^{\mathrm{T}}$, $a_2 = (0,\ 1,\ 0)^{\mathrm{T}}$, $a_3 = (1,\ 2,\ 3)^{\mathrm{T}}$ матрицы A применим процесс ортогонализации. В результате получим:

$$\begin{aligned}
b_1 &= a_1, \\
b_2 &= a_2 - \frac{(\overset{a}{\cancel{b}}_2, b_1)}{|b_1|^2} b_1 = a_2, \\
b_3 &= a_3 - \frac{(\overset{a}{\cancel{b}}_3, b_1)}{|b_1|^2} b_1 - \frac{(\overset{a}{\cancel{b}}_3, b_2)}{|b_2|^2} b_2 = a_3 - \frac{3}{5} b_1 - 2 b_2.
\end{aligned}$$

Таким образом,

$$b_1 = \begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix}, \qquad b_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad b_3 = \frac{1}{5} \begin{pmatrix} -4 \\ 0 \\ 3 \end{pmatrix}.$$

Нормируя векторы b_1, b_2, b_3, получим ортонормированную систему векторов

$$q_1 = \frac{1}{5} \begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix} = \frac{1}{5} a_1, \qquad q_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = a_2,$$

$$q_3 = \frac{1}{5} \begin{pmatrix} -4 \\ 0 \\ 3 \end{pmatrix} = -\frac{3}{5} a_1 - 2a_2 + a_3.$$

В матричной записи это дает равенство $Q = AU$, где

$$Q = \begin{pmatrix} \frac{3}{5} & 0 & -\frac{4}{5} \\ 0 & 1 & 0 \\ \frac{4}{5} & 0 & \frac{3}{5} \end{pmatrix}, \quad U = \begin{pmatrix} \frac{1}{5} & 0 & -\frac{3}{5} \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}.$$

Из равенства $Q = AU$ получается искомое QR-разложение $A = QR$, где $R = U^{-1}$.

Для построения QR-разложения матрицы A с помощью вращений приведем матрицу A к треугольному виду. В матрице A с помощью вращения

$$T_1 = \begin{pmatrix} \cos\varphi & 0 & -\sin\varphi \\ 0 & 1 & 0 \\ \sin\varphi & 0 & \cos\varphi \end{pmatrix}$$

получим нуль на месте элемента $a_{31} = 4$. В произведении $T_1 A$ в этой позиции вычисляем элемент $3\sin\varphi + 4\cos\varphi$ и приравниваем к нулю. Отсюда находим $\operatorname{tg}\varphi = -4/3$. Поэтому

$$\cos\varphi = \frac{1}{\sqrt{1+\operatorname{tg}^2\varphi}} = \frac{3}{5}, \quad \sin\varphi = \frac{\operatorname{tg}\varphi}{\sqrt{1+\operatorname{tg}^2\varphi}} = -\frac{4}{5}.$$

В результате

$$T_1 = \begin{pmatrix} \frac{3}{5} & 0 & \frac{4}{5} \\ 0 & 1 & 0 \\ -\frac{4}{5} & 0 & \frac{3}{5} \end{pmatrix}$$

и

$$T_1 A = \begin{pmatrix} \frac{3}{5} & 0 & \frac{4}{5} \\ 0 & 1 & 0 \\ -\frac{4}{5} & 0 & \frac{3}{5} \end{pmatrix} \begin{pmatrix} 3 & 0 & 1 \\ 0 & 1 & 2 \\ 4 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 5 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} = R.$$

Мы приходим к искомому QR-разложению $A = QR$, где

$$Q = T^{-1} = T_1^{\mathrm{T}} = \begin{pmatrix} \frac{3}{5} & 0 & -\frac{4}{5} \\ 0 & 1 & 0 \\ \frac{4}{5} & 0 & \frac{3}{5} \end{pmatrix}.$$

При построении QR-разложения матрицы A с помощью отражений ее необходимо привести с помощью отражений к треугольному виду. Возьмем $x = (3,\ 0,\ 4)^{\mathrm{T}}$, $z = (1,\ 0,\ 0)^{\mathrm{T}}$. Тогда $v = x + |x|z = (8,\ 0,\ 4)^{\mathrm{T}}$ и

$$H_1 = E - \frac{2vv^{\mathrm{T}}}{|v|^2} = E - \frac{2}{80}\begin{pmatrix} 8 \\ 0 \\ 4 \end{pmatrix}(8,\ 0,\ 4) = \frac{1}{5}\begin{pmatrix} -3 & 0 & -4 \\ 0 & 5 & 0 \\ -4 & 0 & 3 \end{pmatrix}.$$

Вычислив

$$H_1A = \frac{1}{5}\begin{pmatrix} -3 & 0 & -4 \\ 0 & 5 & 0 \\ -4 & 0 & 3 \end{pmatrix}\begin{pmatrix} 3 & 0 & 1 \\ 0 & 1 & 2 \\ 4 & 0 & 3 \end{pmatrix} = \begin{pmatrix} -5 & 0 & -3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} = R,$$

получим искомое QR-разложение $A = QR$, где

$$Q = H_1^{-1} = H_1^{\mathrm{T}} = \frac{1}{5}\begin{pmatrix} -3 & 0 & -4 \\ 0 & 5 & 0 \\ -4 & 0 & 3 \end{pmatrix}.$$

Пример 8.24. Построить QR-разложение комплексной матрицы

$$A = \begin{pmatrix} 0 & 6i & i \\ 3i & -10 & -i \\ 4i & 0 & 1 \end{pmatrix}.$$

Решение. QR-разложение построим, приводя матрицу A к треугольному виду умножением ее на последовательность унитарных матриц T_{ij} (см. разд. 8.14). Начнем с получения нуля во второй строке первом столбце (позиция a_{21}) с помощью матрицы

$$T_{21} = \begin{pmatrix} \cos\varphi & -e^{i\psi}\sin\varphi & 0 \\ e^{-i\psi}\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Вычислим элемент в этой позиции у произведения $T_{21}A$, равный $3i\cos\varphi$, и приравняем к нулю. В результате находим

$\cos\varphi=0$, $\sin\varphi=1$, ψ любое. Возьмем $\psi=0$. Тогда

$$T_{21}=\begin{pmatrix}0&-1&0\\1&0&0\\0&0&1\end{pmatrix}\quad\text{и}\quad T_{21}A=\begin{pmatrix}-3i&10&1\\0&6i&i\\4i&0&1\end{pmatrix}.$$

В матрице $T_{21}A$ с помощью матрицы

$$T_{31}=\begin{pmatrix}\cos\varphi&0&-e^{i\psi}\sin\varphi\\0&1&0\\e^{-i\psi}\sin\varphi&0&\cos\varphi\end{pmatrix}$$

получим нуль в третьей строке первом столбце. Для этого в матрице $T_{31}T_{21}A$ определяем элемент в соответствующей позиции $-3ie^{-i\psi}\sin\varphi+4i\cos\varphi$ и приравниваем к нулю. При $\psi=0$ получим $\operatorname{tg}\varphi=4/3$ и

$$\sin\varphi=\frac{\operatorname{tg}\varphi}{\sqrt{1+\operatorname{tg}^2\varphi}}=\frac{4}{5},\quad \cos\varphi=\frac{1}{\sqrt{1+\operatorname{tg}^2\varphi}}=\frac{3}{5}.$$

Составим матрицу T_{31} и вычислим произведение $T_{31}T_{21}A$:

$$T_{31}=\begin{pmatrix}\frac{3}{5}&0&-\frac{4}{5}\\0&1&0\\\frac{4}{5}&0&\frac{3}{5}\end{pmatrix}\quad\text{и}\quad T_{31}T_{21}A=\begin{pmatrix}-5i&6&-\frac{1}{5}\\0&6i&i\\0&8&\frac{7}{5}\end{pmatrix}.$$

В матрице $T_{31}T_{21}A$ с помощью матрицы

$$T_{32}=\begin{pmatrix}1&0&0\\0&\cos\varphi&-e^{i\psi}\sin\varphi\\0&e^{-i\psi}\sin\varphi&\cos\varphi\end{pmatrix}$$

получим нуль в третьей строке втором столбце. Для этого, вычислив соответствующий элемент $6ie^{-i\psi}\sin\varphi+8\cos\varphi$ произведения $T_{32}T_{31}T_{21}A$, приравняем его к нулю. При $\psi=\pi/2$ находим $\operatorname{tg}\varphi=-4/3$ и

$$\sin\varphi=\frac{\operatorname{tg}\varphi}{\sqrt{1+\operatorname{tg}^2\varphi}}=-\frac{4}{5},\quad \cos\varphi=\frac{1}{\sqrt{1+\operatorname{tg}^2\varphi}}=\frac{3}{5}.$$

Составим матрицу T_{32} и вычислим произведение $T_{32}T_{31}T_{21}A$:

$$T_{32} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{3}{5} & \frac{4i}{5} \\ 0 & \frac{4i}{5} & \frac{3}{5} \end{pmatrix} \quad \text{и} \quad T_{32}T_{31}T_{21}A = \begin{pmatrix} -5i & 6 & -\frac{1}{5} \\ 0 & 10i & \frac{43i}{25} \\ 0 & 0 & \frac{1}{25} \end{pmatrix} = R.$$

Из последнего равенства получим искомое QR-разложение $A = QR$, где

$$Q = T_{21}^* T_{31}^* T_{32}^* = \begin{pmatrix} 0 & \frac{3}{5} & -\frac{4i}{5} \\ -\frac{3}{5} & \frac{16i}{25} & -\frac{12i}{25} \\ -\frac{4}{5} & -\frac{12i}{25} & \frac{9}{25} \end{pmatrix}.$$

При построении QR-разложения с помощью отражений матрицу A необходимо последовательностью простых отражений привести к треугольному виду. Для этого возьмем $x = (0,\ 3i,\ 4i)^{\mathrm{T}}$, найдем $|x| = 5$ и положим $z = (1,\ 0,\ 0)^{\mathrm{T}}$, $v = x - |x|z = (-5,\ 3i,\ 4i)^{\mathrm{T}}$. Затем составим матрицу отражений

$$H = E - \frac{2vv^*}{|v|^2} = E - \frac{2}{50}\begin{pmatrix} -5 \\ -3i \\ -4i \end{pmatrix}(-5,\ -3i,\ -4i) =$$

$$= \frac{1}{25}\begin{pmatrix} 0 & -15i & -20i \\ 15i & 16 & -12 \\ 20i & -12 & 9 \end{pmatrix}$$

и найдем матрицу

$$HA = \frac{1}{25}\begin{pmatrix} 125 & 150i & -15-20i \\ 0 & -250 & -27-16i \\ 0 & 0 & -11+12i \end{pmatrix} = R.$$

Из последнего равенства получаем искомое QR-разложение $A = H^*R$.

Разложения, аналогичные QR-разложению, можно строить и для прямоугольных $(m \times n)$-матриц (см. [3, с. 214] и [34]). В этом случае множитель Q — это прямоугольная $(m \times n)$-матрица, у которой столбцы составляют ортонормированную систему. Такое разложение прямоугольной матрицы называют ее qR-разложением. Рассматривают и Qr-разложения (см. [3, с. 213] и [34]), в которых прямоугольной является матрица R.

Пример 8.25. Построить qR-разложение матрицы

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

Решение. В примере 8.5 в процессе ортогонализации векторов a_1, a_2, a_3, представляющих собой столбцы матрицы A, была получена ортонормированная система векторов

$$q_1 = \frac{1}{\sqrt{2}}a_1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0\right)^{\mathrm{T}},$$
$$q_2 = -\frac{1}{\sqrt{6}}a_1 + \frac{2}{\sqrt{6}}a_2 = \left(\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, 0\right)^{\mathrm{T}},$$
$$q_3 = -\frac{\sqrt{3}}{6}a_1 - \frac{\sqrt{3}}{6}a_2 + \frac{\sqrt{3}}{6}a_3 = \left(\frac{\sqrt{3}}{6}, -\frac{\sqrt{3}}{6}, -\frac{\sqrt{3}}{6}, \frac{\sqrt{3}}{2}\right)^{\mathrm{T}}.$$

В матричной записи связь q_1, q_2, q_3 с a_1, a_2, a_3 дает $Q = AU$, где

$$Q = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{\sqrt{3}}{6} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & -\frac{\sqrt{3}}{6} \\ 0 & \frac{2}{\sqrt{6}} & -\frac{\sqrt{3}}{6} \\ 0 & 0 & \frac{3}{\sqrt{2}} \end{pmatrix}, \quad U = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & -\frac{\sqrt{3}}{6} \\ 0 & \frac{2}{\sqrt{6}} & -\frac{\sqrt{3}}{6} \\ 0 & 0 & \frac{\sqrt{3}}{2} \end{pmatrix}.$$

Отсюда получается qR-разложение $A = QR$, где $R = U^{-1}$. Отметим, что qR-разложение матрицы A можно получить также с помощью вращений или отражений.

Если известно QR- или qR-разложение матрицы A, то по столбцам матрицы Q легко выписать ортонормированную систему векторов, полученную из системы столбцов матрицы A. Так как QR- и qR-разложения можно строить либо с помощью ортогонализации и нормирования, либо с помощью вращений, либо с помощью отражений, то это означает, что такими же способами можно осуществлять ортонормирование систем векторов.

QR-разложения матриц имеют широкое применение в численных методах линейной алгебры. Они являются основой QR-алгоритма (см. разд. 11.2).

Если известно QR-разложение $A = QR$ матрицы A системы линейных уравнений $AX = b$, то значительно упрощается решение этой системы, так как она сводится к системе $RX = Q^T b$ с треугольной матрицей R.

8.18. Сингулярное разложение матрицы

Существование для симметрической (эрмитовой) матрицы канонического разложения с ортогональной (унитарной) трансформирующей матрицей позволяет получить специальное разложение для произвольной $(m \times n)$-матрицы. Прежде всего заметим, что для произвольной $(m \times n)$-матрицы A ранга r матрицы A^*A и AA^* являются симметрическими (эрмитовыми) неотрицательными матрицами ранга r и порядков соответственно n и m. Характеристические числа этих матриц являются действительными неотрицательными числами.

Обозначим характеристические числа матрицы A^*A через $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$ и будем считать, что $\sigma_1^2 \geqslant \sigma_2^2 \geqslant \ldots \geqslant \sigma_n^2$ ($\sigma_i \neq 0$ при $i = 1, \ldots, r$). Оператор с симметрической (эрмитовой) матрицей A^*A имеет ортонормированную систему собственных векторов $e_1, e_2, \ldots, e_n$, отвечающих собственным значениям $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$. Векторы $e_1, e_2, \ldots, e_n$ удовлетворяют соотношениям

$$A^*Ae_i = \sigma_i^2 e_i, \quad (e_i, e_j) = \begin{cases} 1 \text{ при } i = j, \\ 0 \text{ при } i \neq j. \end{cases} \tag{8.50}$$

Система векторов $e_1, e_2, \ldots, e_n$ переводится оператором с матрицей A в некоторую ортогональную систему векторов $Ae_1, Ae_2, \ldots, Ae_n$, так как

$$(Ae_i, Ae_j) = (A^*Ae_i, e_j) = \sigma_i^2 (e_i, e_j) = 0, \quad i \neq j.$$

Кроме того, модуль вектора Ae_i равен σ_i, так как

$$|Ae_i| = \sqrt{(A^*Ae_i, e_i)} = \sqrt{\sigma_i^2 (e_i, e_i)} = \sigma_i.$$

Поэтому вектор Ae_i отличен от нулевого вектора тогда и только тогда, когда $\sigma_i \neq 0$, т.е. при $i = 1, 2, \ldots, r$. Ненулевой вектор Ae_i является собственным вектором оператора AA^* по собственному значению σ_i^2, так как

$$AA^*(Ae_i) = A(A^*Ae_i) = A(\sigma_i^2 e_i) = \sigma_i^2 Ae_i.$$

Верно и обратное. Следовательно, ненулевые характеристические числа матриц A^*A и AA^* совпадают с учетом их кратностей и их число равно r, а кратности нулевого характеристического числа этих матриц равны соответственно $n - r$ и $m - r$. Общих характеристических чисел у матриц A^*A и AA^* будет: $s = \min\{m, n\}$.

Арифметические значения $\sigma_1, \sigma_2, \ldots, \sigma_s$ квадратных корней из общих характеристических чисел матриц A^*A и AA^* называют ***сингулярными*** (или ***главными***) ***числами матрицы*** A.

В пространстве X примем за базис ортонормированную систему $e_1, e_2, \ldots, e_n$ собственных векторов оператора с матрицей A^*A и построим ортонормированную систему векторов

$$f_1 = \frac{Ae_1}{|Ae_1|} = \frac{Ae_1}{\sigma_1}, \quad \ldots, \quad f_r = \frac{Ae_r}{|Ae_r|} = \frac{Ae_r}{\sigma_r}.$$

Дополним эту систему любыми векторами $f_{r+1}, \ldots, f_m$ до ортонормированного базиса в Y. По построению векторы f_1, $f_2, \ldots, f_m$ удовлетворяют соотношениям

$$Ae_i = \begin{cases} \sigma_i f_i & \text{при } i \leqslant r, \\ 0 & \text{при } i > r. \end{cases} \tag{8.51}$$

Умножая эти равенства слева на A^* и учитывая, что $A^*Ae_i = \sigma_i^2 e_i$, получим соотношения

$$A^* f_i = \begin{cases} \sigma_i e_i & \text{при } i \leqslant r, \\ 0 & \text{при } i > r. \end{cases} \tag{8.52}$$

Ортонормированные базисы $e_1, e_2, \ldots, e_n$ и $f_1, f_2, \ldots, f_m$ пространств X и Y, связанные соотношениями (8.51) и (8.52), называют ***сингулярными базисами***. При этом векторы e_1, $e_2, \ldots, e_n$ называют ***правыми сингулярными векторами матрицы*** A, а векторы $f_1, f_2, \ldots, f_m$ — ее ***левыми сингулярными векторами***.

Оператор, имеющий в паре исходных базисов пространств X и Y матрицу A, в сингулярных базисах $e_1, e_2, \ldots, e_n$ и $f_1, f_2, \ldots, f_m$ этих пространств в силу определения матрицы оператора (см. разд. 5.2) и соотношений (8.51) имеет $(m \times n)$-матрицу

$$\Sigma = \begin{pmatrix} \sigma_1 & 0 & \ldots & 0 & \ldots & 0 \\ 0 & \sigma_2 & \ldots & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \sigma_r & \ldots & 0 \\ 0 & 0 & \ldots & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 & \ldots & 0 \end{pmatrix}. \tag{8.53}$$

Из формулы, устанавливающей связь между матрицами одного и того же оператора в разных базисах, получаем

$$A = Q\Sigma P^*, \tag{8.54}$$

где P — ортогональная (унитарная) матрица порядка n, столбцами которой служат столбцы координат векторов e_1,

$e_2, \ldots, e_n$ в исходном базисе пространства X, Q — ортогональная (унитарная) матрица порядка m, столбцами которой являются столбцы координат векторов $f_1, f_2, \ldots, f_m$ в исходном базисе пространства Y.

Разложение (8.54) называют ***сингулярным разложением матрицы*** A. Любая $(m \times n)$-матрица обладает многими различными сингулярными разложениями. Это следует из некоторого произвола при построении векторов $e_1, e_2, \ldots, e_m$ и $f_1, f_2, \ldots, f_m$.

Сингулярному разложению (8.54) можно придать вид

$$A = U\Sigma_r V^*, \tag{8.55}$$

где

$$\Sigma_r = \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{pmatrix}$$

— квадратная матрица порядка r, получающаяся из $(m\times n)$-матрицы Σ вычеркиванием $n-r$ нулевых столбцов справа и $m-r$ нулевых строк снизу, U — $(m \times r)$-матрица, состоящая из первых r столбцов матрицы Q, V^* — $(r \times n)$-матрица, состоящая из первых r строк матрицы P^*.

Разложение (8.55) называют ***второй формой сингулярного разложения*** матрицы A. В него входят матрицы меньших размеров, чем в первую форму, и, кроме того, в нем матрица Σ_r — квадратная невырожденная. Все это может оказаться существенным, особенно при компьютерных вычислениях.

При $m < n$ сингулярному разложению (8.54) иногда бывает удобно придать вид

$$A = Q\Sigma_m P_1^*, \tag{8.56}$$

где Σ_m — квадратная матрица m-го порядка, содержащаяся в первых m строках и столбцах матрицы Σ, матрица Q — та же, что и в разложении (8.54), а P_1^* — $(m \times n)$-матрица, состоящая из первых m строк матрицы P^*.

При $m > n$ сингулярному разложению (8.54) иногда также бывает удобно придать вид

$$A = Q_1 \Sigma_n P^*, \tag{8.57}$$

где Σ_n — квадратная матрица, содержащаяся в первых n строках и столбцах матрицы Σ, Q_1 — $(m \times n)$-матрица, состоящая из первых n столбцов матрицы Q, а матрица P^* — та же, что и в разложении (8.54).

При конструировании сингулярного разложения произвольной $(m \times n)$-матрицы A ранга r следует придерживаться следующего порядка действий.

1. Составить матрицу A^*A, найти ее характеристические числа $\lambda_1, \ldots, \lambda_n$ (не нарушая общности, их можно перенумеровать так, чтобы выполнялись условия $\lambda_1 \geqslant \ldots \geqslant \lambda_n$, $\lambda_i \neq 0$ при $i = 1, 2, \ldots, r$). Подсчитать ненулевые сингулярные числа $\sigma_i = \sqrt{\lambda_i}$, $i = 1, 2, \ldots, r$, и по формуле (8.53) составить $(m \times n)$-матрицу Σ.

2. Построить ортонормированную систему собственных векторов $e_1, e_2, \ldots, e_n$ оператора с матрицей A^*A, принадлежащих соответственно собственным значениям $\lambda_1, \lambda_2, \ldots, \lambda_n$. Из столбцов координат этих векторов составить матрицу P. Это будет матрица перехода от исходного базиса пространства X_n к базису $e_1, \ldots, e_n$.

3. Построить ортонормированную систему векторов

$$f_1 = \frac{Ae_1}{\sigma_1}, \quad \ldots, \quad f_r = \frac{Ae_r}{\sigma_r}$$

и дополнить ее любыми векторами $f_{r+1}, f_{r+2}, \ldots, f_m$ до ортонормированного базиса пространства Y. Из столбцов координат векторов $f_1, f_2, \ldots, f_m$ составить матрицу Q. Она будет матрицей перехода от исходного базиса пространства Y к базису $f_1, f_2, \ldots, f_m$.

4. Записать искомое сингулярное разложение $A = Q\Sigma P^*$.

Построенные базисы $e_1, e_2, \ldots, e_n$ и $f_1, f_2, \ldots, f_m$ являются сингулярными базисами пространств X и Y, причем векторы первого базиса — это правые сингулярные векторы матрицы A, а векторы второго — ее левые сингулярные векторы.

Если действительная (комплексная) матрица A симметрическая, то существует сингулярное разложение $A = Q\Sigma P^*$, в котором ортогональные (унитарные) матрицы P и Q совпадают.

Пример 8.26. Построить сингулярное разложение для матрицы

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

Решение. Для матрицы

$$A^*A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 3 \end{pmatrix}$$

характеристический многочлен $|A^*A - \lambda E| = \lambda(\lambda - 4)(1 - \lambda)^2$ имеет корни $\lambda_1 = 4$, $\lambda_2 = \lambda_3 = 1$, $\lambda_4 = 0$. Поэтому $\sigma_1 = \sqrt{\lambda_1} = 2$, $\sigma_2 = \sigma_3 = \sqrt{\lambda_{2,3}} = 1$ и

$$\Sigma = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

При $\lambda = 4$ система $(A^*A - \lambda E)x = 0$ имеет фундаментальную систему решений из одного вектора, например $(1,\ 1,\ 1,\ 3)^{\mathrm{T}}$. Нормируя этот вектор, получим

$$e_1 = \frac{1}{2\sqrt{3}}(1,\ 1,\ 1,\ 3)^{\mathrm{T}}.$$

При $\lambda = 1$ система $(A^*A - \lambda E)x = 0$ имеет фундаментальную систему решений из двух векторов, например, $(1,\ -1,\ 0,\ 0)^{\mathrm{T}}$

и $(1,\ 0,\ -1,\ 0)^{\mathrm{T}}$. Ортогонализируя и нормируя эту систему решений, получим

$$e_2 = \frac{1}{\sqrt{2}}(1,\ -1,\ 0,\ 0)^{\mathrm{T}}, \quad e_3 = \frac{1}{\sqrt{6}}(1,\ 1,\ -2,\ 0)^{\mathrm{T}}.$$

При $\lambda = 0$ система $(A^*A - \lambda E)x = 0$ имеет фундаментальную систему решений из одного вектора, например $(1,\ 1,\ 1,\ -1)^{\mathrm{T}}$. Нормируя этот вектор, получим

$$e_4 = \frac{1}{2}(1,\ 1,\ 1,\ -1)^{\mathrm{T}}.$$

Из столбцов координат векторов e_1, e_2, e_3, e_4 составим матрицу

$$P = \begin{pmatrix} \frac{1}{2\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{2} \\ \frac{1}{2\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{2} \\ \frac{1}{2\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} & \frac{1}{2} \\ \frac{3}{2\sqrt{3}} & 0 & 0 & -\frac{1}{2}. \end{pmatrix}.$$

Далее строим векторы

$$f_1 = \frac{Ae_1}{\sigma_1} = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \cdot \frac{1}{2\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 3 \end{pmatrix} = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

$$f_2 = \frac{Ae_2}{\sigma_2} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \cdot \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix},$$

$$f_3 = \frac{Ae_3}{\sigma_3} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \cdot \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 1 \\ -2 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}.$$

Число этих векторов $m = 3$ оказалось равным размерности пространства Y. Поэтому они составляют базис этого пространства.

Из столбцов координат векторов f_1, f_2, f_3 составим матрицу

$$Q = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \end{pmatrix}$$

и запишем искомое сингулярное разложение

$$A = Q\Sigma P^* = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \times$$

$$\times \begin{pmatrix} \frac{1}{2\sqrt{3}} & \frac{1}{2\sqrt{3}} & \frac{1}{2\sqrt{3}} & \frac{3}{2\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{pmatrix}.$$

Векторы e_1, e_2, e_3, e_4 и f_1, f_2, f_3 составляют сингулярные базисы пространств X и Y, причем e_1, e_2, e_3, e_4 — это правые сингулярные векторы матрицы A, а f_1, f_2, f_3 — левые.

Пример 8.27. Построить сингулярное разложение для матрицы

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

Решение. Составим матрицу

$$A^*A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & -1 & 0 \\ -1 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 5 & -3 & -1 \\ -3 & 3 & 3 \\ -1 & 3 & 5 \end{pmatrix}.$$

Ее характеристический многочлен

$$|A^*A-\lambda E| = \begin{vmatrix} 5-\lambda & -3 & -1 \\ -3 & 3-\lambda & 3 \\ -1 & 3 & 5-\lambda \end{vmatrix} = -\lambda(\lambda^2-13\lambda+36)$$

имеет корни $\lambda_1 = 9$, $\lambda_2 = 4$, $\lambda_3 = 0$. Поэтому $\sigma_1 = \sqrt{\lambda_1} = 3$, $\sigma_2 = \sqrt{\lambda_2} = 2$ и

$$\Sigma = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Из системы $(A^*A-\lambda E)x = 0$ при $\lambda = 9, 4, 0$ найдем соответственно векторы $b_1 = (1,\ -1,\ -1)^{\mathrm{T}}$, $b_2 = (1,\ 0,\ 1)^{\mathrm{T}}$, $b_3 = (1,\ 2,\ -1)^{\mathrm{T}}$. Они уже ортогональны, так как принадлежат различным собственным значениям. Нормируя их, получим векторы

$$e_1 = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}, \quad e_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}^{\mathrm{T}}, \quad e_3 = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}^{\mathrm{T}}.$$

Заметим, что в силу симметричности матрицы A эти векторы являются собственными векторами матрицы A.

Из столбцов координат векторов e_1, e_2, e_3 составим матрицу

$$P = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \end{pmatrix}.$$

Далее построим векторы

$$f_1 = \frac{Ae_1}{\sigma_1} = \frac{1}{3}\begin{pmatrix} 2 & -1 & 0 \\ -1 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix} \cdot \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix},$$

$$f_2 = \frac{Ae_2}{\sigma_2} = \frac{1}{2}\begin{pmatrix} 2 & -1 & 0 \\ -1 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix} \cdot \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Систему векторов f_1, f_2 дополним до ортонормированного базиса пространства Y, например, вектором $f_3 = \frac{1}{\sqrt{6}}(1, 2, -1)^{\mathrm{T}}$.

Из столбцов координат векторов f_1, f_2, f_3 построим матрицу Q. Она совпала с матрицей P.

Теперь запишем искомое сингулярное разложение

$$A = Q\Sigma P^* = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \times$$
$$\times \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{6}} \end{pmatrix}.$$

Пример 8.28. Построить сингулярное разложение матрицы

$$A = \begin{pmatrix} 4 & -3i \\ -3i & 4 \end{pmatrix}.$$

Решение. Характеристический многочлен

$$|A^*A - \lambda E| = \begin{vmatrix} 25-\lambda & 0 \\ 0 & 25-\lambda \end{vmatrix} = (25-\lambda)^2$$

матрицы A^*A имеет корни $\lambda_1 = \lambda_2 = 25$. Поэтому $\sigma_1 = \sigma_2 = \sqrt{25} = 5$. Следовательно, матрица Σ имеет вид

$$\Sigma = \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix}.$$

При $\lambda = 25$ система $(A^*A - \lambda E)x = 0$ имеет вид

$$\begin{cases} 0 \cdot x_1 + 0 \cdot x_2 = 0, \\ 0 \cdot x_1 + 0 \cdot x_2 = 0. \end{cases}$$

Ее фундаментальная система решений состоит из любых двух векторов. Возьмем ортонормированные векторы $e_1 = (1, 0)^{\mathrm{T}}$,

$e_2 = (0,\ 1)^{\text{т}}$. Из столбцов координат этих векторов построим матрицу

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

Далее строим векторы

$$f_1 = \frac{Ae_1}{\sigma_1} = \frac{1}{5}\begin{pmatrix} 4 & -3i \\ -3i & 4 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 4 \\ -3i \end{pmatrix},$$

$$f_2 = \frac{Ae_2}{\sigma_2} = \frac{1}{5}\begin{pmatrix} 4 & -3i \\ -3i & 4 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} -3i \\ 4 \end{pmatrix}.$$

Число этих векторов равно размерности пространства Y. Поэтому из столбцов их координат построим матрицу

$$Q = \begin{pmatrix} \frac{4}{5} & -\frac{3i}{5} \\ -\frac{3i}{5} & \frac{4}{5} \end{pmatrix}$$

и запишем искомое сингулярное разложение

$$A = Q\Sigma P^* = \begin{pmatrix} \frac{4}{5} & -\frac{3i}{5} \\ -\frac{3i}{5} & \frac{4}{5} \end{pmatrix}\begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

При конструировании сингулярного разложения матрицы на ЭВМ его обычно получают косвенным путем (см. [18]). Стандартную программу svd такого метода можно найти в [29], [33] и во многих пакетах математического обеспечения.

Процесс получения сингулярного разложения косвенным путем по аналогии с методом вращений (см. разд. 11.1), можно приспособить к отысканию абсолютных величин собственных значений (диагональных элементов матрицы Σ) и соответствующих им собственных векторов (столбцов матрицы P) произвольных симметрических (эрмитовых) матриц, а в случае положительно определенных матриц — к отысканию собственных значений и соответствующих им собственных векторов.

Сингулярное разложение находит самое широкое применение в теории (см. [34]) и приложениях, при конструировании псевдообратной матрицы (см. разд. 8.20), при отыскании псевдорешений системы линейных уравнений и их проекций на пространства правых сингулярных векторов (см. разд. 8.21, [5]), при отыскании устойчивого решения системы линейных уравнений (см. разд. 8.20, 8.26, [5]), при проведении сингулярного анализа модели выравнивающей функции по методу наименьших квадратов (см. разд. 8.25). Здесь приведем лишь один пример.

Пример 8.29. Используя сингулярное разложение матрицы, решить систему

$$\begin{cases} 2x_1 - x_2 = 1, \\ -x_1 + x_2 + x_3 = 0, \\ x_2 + 2x_3 = 1. \end{cases}$$

Решение. Сингулярное разложение матрицы данной системы получено в примере 8.27. Пользуясь этим разложением, запишем систему в виде

$$Q\Sigma P^* x = b.$$

Отсюда $\Sigma P^* x = Q^{-1} b = Q^{\mathrm{T}} b$. Полагая $P^* x = z$, приходим к системе уравнений $\Sigma z = Q^* b$, или в подробной записи

$$\begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \sqrt{2} \\ 0 \end{pmatrix}.$$

Мы имеем в матричной записи систему линейных уравнений простейшего вида

$$\begin{cases} 3z_1 = 0, \\ 2z_2 = \sqrt{2}, \\ 0 \cdot z_3 = 0, \end{cases}$$

общим решением которой является $z = \left(0,\ 1/\sqrt{2},\ C\right)^{\mathrm{T}}$, где C — произвольное постоянное. Из уравнения $P^*x = z$ находим

$$x = Pz = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ C \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \frac{C}{\sqrt{6}}\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}.$$

8.19. Полярное разложение матрицы

Любой линейный оператор, действующий в евклидовом (унитарном) пространстве, разлагается в произведение симметрического (эрмитова) и ортогонального (унитарного) операторов (см. разд. 8.11 и 8.16). Для матриц это означает, что любая квадратная действительная (комплексная) матрица разлагается в произведение

$$A = HU \tag{8.58}$$

симметрической (эрмитовой) матрицы H и ортогональной (унитарной) матрицы U. Такое разложение матрицы A называют ее ***полярным разложением***. Как показано в теоремах 8.33 и 8.36, в полярном разложении симметрическая составляющая H определена однозначно и совпадает с $\sqrt{AA^*}$. Если матрица A невырожденная, то и ортогональная составляющая U определена однозначно и совпадает с матрицей $A^{-1}H = H(A^*)^{-1}$.

Полярное разложение можно получить исходя из сингулярного разложения. Действительно, если $A = Q\Sigma P^*$, где Q и P — ортогональные (унитарные) матрицы, а Σ — диагональная матрица, то преобразуем сингулярное разложение следующим образом:

$$A = Q\Sigma Q^* QP^*.$$

Положив

$$H = Q\Sigma Q^*, \quad U = QP^*, \tag{8.59}$$

получим полярное разложение $A = HU$. Действительно, оператор $H = Q\Sigma Q^*$ симметрический (эрмитов), так как

$$H^* = (Q\Sigma Q^*)^* = Q\Sigma^* Q^* = Q\tilde{\Sigma} Q^*,$$

а оператор $U = QP^*$ ортогональный (унитарный) как произведение двух ортогональных (унитарных) операторов.

Пример 8.30. Построить полярное разложение для матрицы

$$A = \begin{pmatrix} -1 & -7 \\ 1 & 7 \end{pmatrix}.$$

Решение. Для матрицы

$$A^*A = \begin{pmatrix} -1 & 1 \\ -7 & 7 \end{pmatrix} \begin{pmatrix} -1 & -7 \\ 1 & 7 \end{pmatrix} = \begin{pmatrix} 2 & 14 \\ 14 & 98 \end{pmatrix}$$

характеристический многочлен $|A^{\mathrm{T}}A - \lambda E| = \lambda(\lambda - 100)$ имеет корни $\lambda_1 = 100$, $\lambda_2 = 0$. Поэтому $\sigma_1 = \sqrt{\lambda_1} = 10$, $\sigma_2 = 0$ и

$$\Sigma = \begin{pmatrix} 10 & 0 \\ 0 & 0 \end{pmatrix}.$$

При $\lambda = 100$ система $(A^*A - \lambda E)x = 0$ имеет фундаментальную систему решений из одного вектора, например $(1,\ 7)^{\mathrm{T}}$. Нормируя этот вектор, получим

$$e_1 = \frac{1}{5\sqrt{2}} \begin{pmatrix} 1 \\ 7 \end{pmatrix}.$$

Аналогично при $\lambda = 0$ находим

$$e_2 = \frac{1}{5\sqrt{2}} \begin{pmatrix} 7 \\ 1 \end{pmatrix}.$$

Далее строим вектор

$$f_1 = \frac{Ae_1}{\sigma_1} = \frac{1}{10} \begin{pmatrix} -1 & -7 \\ 1 & 7 \end{pmatrix} \cdot \frac{1}{5\sqrt{2}} \begin{pmatrix} 1 \\ 7 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

и дополняем его до ортонормированного базиса пространства, например, вектором

$$f_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Из столбцов координат векторов e_1 и e_2 составляем матрицу

$$P = \frac{1}{5\sqrt{2}}\begin{pmatrix} 1 & 7 \\ 7 & -1 \end{pmatrix},$$

а из столбцов координат векторов f_1 и f_2 — матрицу

$$Q = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}.$$

В результате для матрицы A получаем сингулярное разложение

$$A = Q\Sigma P^* = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 10 & 0 \\ 0 & 9 \end{pmatrix}\frac{1}{5\sqrt{2}}\begin{pmatrix} 1 & 7 \\ 7 & -1 \end{pmatrix}.$$

Полярное разложение матрицы получим из сингулярного, воспользовавшись формулами (8.59). По этим формулам находим

$$H = Q\Sigma Q^* = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 10 & 0 \\ 0 & 0 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 5 & -5 \\ -5 & 5 \end{pmatrix};$$

$$U = QP^* = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}\frac{1}{5\sqrt{2}}\begin{pmatrix} 1 & 7 \\ 7 & -1 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 3 & -4 \\ 4 & 3 \end{pmatrix}.$$

Таким образом, полярное разложение матрицы имеет вид:

$$A = HU = \begin{pmatrix} 5 & -5 \\ -5 & 5 \end{pmatrix}\frac{1}{5}\begin{pmatrix} 3 & -4 \\ 4 & 3 \end{pmatrix}.$$

Пример 8.31. Построить полярное разложение для матрицы

$$A = \begin{pmatrix} 4 & -3i \\ -3i & 4 \end{pmatrix}.$$

Решение. Вычисляем матрицу AA^*:

$$AA^* = \begin{pmatrix} 4 & -3i \\ -3i & 4 \end{pmatrix} \begin{pmatrix} 4 & 3i \\ 3i & 4 \end{pmatrix} = \begin{pmatrix} 25 & 0 \\ 0 & 25 \end{pmatrix}.$$

В данном случае матрицу H проще всего найти непосредственно по формуле $H = \sqrt{AA^*}$:

$$H = \sqrt{AA^*} = \sqrt{\begin{pmatrix} 25 & 0 \\ 0 & 25 \end{pmatrix}} = \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix}.$$

Поскольку матрица AA^* невырожденная, то и матрица A невырожденная. Следовательно, унитарную составляющую U полярного разложения можно найти по формуле $U = H^{-1}A$:

$$U = H^{-1}A = \frac{1}{5}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 4 & -3i \\ -3i & 4 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 4 & -3i \\ -3i & 4 \end{pmatrix}.$$

Таким образом, полярное разложение матрицы A имеет вид:

$$A = \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix} \cdot \frac{1}{5}\begin{pmatrix} 4 & -3i \\ -3i & 4 \end{pmatrix}.$$

Наряду с полярным разложением вида $A = HU$ используют и альтернативный вариант полярного разложения $A = U_1H_1$, в котором правый сомножитель H_1 является симметрической (эрмитовой) матрицей, а левый U_1 — ортогональной (унитарной) матрицей. Такое разложение можно получить из сингулярного разложения $A = Q\Sigma P^*$, представив его в виде $A = Q\Sigma P^* = (QP^*)(P\Sigma P^*)$. Тогда можно положить $U_1 = QP^*$, $H_1 = P\Sigma P^*$. При этом

$$H_1 = P\Sigma P^* = \sqrt{P\Sigma^2 P^*} = \sqrt{(P\Sigma Q^*)(Q\Sigma P^*)} = \sqrt{A^*A}.$$

Например, для матрицы

$$A = \begin{pmatrix} -1 & -7 \\ 1 & 7 \end{pmatrix}$$

в примере 8.30 найдены матрицы

$$Q = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} 10 & 0 \\ 0 & 0 \end{pmatrix}, \quad P = \frac{1}{5\sqrt{2}}\begin{pmatrix} 1 & 7 \\ 7 & -1 \end{pmatrix}.$$

Поэтому

$$U_1 = QP^* = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix} \cdot \frac{1}{5\sqrt{2}}\begin{pmatrix} 1 & 7 \\ 7 & -1 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 3 & -4 \\ 4 & 3 \end{pmatrix},$$

$$H_1 = P\Sigma P^* = \frac{1}{5\sqrt{2}}\begin{pmatrix} 1 & 7 \\ 7 & -1 \end{pmatrix}\begin{pmatrix} 10 & 0 \\ 0 & 0 \end{pmatrix} \times$$

$$\times \frac{1}{5\sqrt{2}}\begin{pmatrix} 1 & 7 \\ 7 & -1 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 1 & 7 \\ 7 & 49 \end{pmatrix}.$$

Отсюда находим альтернативное полярное разложение

$$A = U_1 H_1 = \frac{1}{5}\begin{pmatrix} 3 & -4 \\ 4 & 3 \end{pmatrix} \cdot \frac{1}{5}\begin{pmatrix} 1 & 7 \\ 7 & 49 \end{pmatrix}.$$

Альтернативное полярное разложение может быть также получено из полярного разложения транспонированной матрицы. Действительно, если $A^* = HU$, где H — симметрическая (эрмитова) матрица, U — ортогональная (унитарная) матрица, то $A = U^* H^*$, где матрица U^* ортогональная (унитарная), а матрица $H^* = H$ симметрическая (эрмитова).

Полярное разложение можно обобщить на произвольные, не обязательно квадратные, матрицы. Любую $(m \times n)$-матрицу можно представить в виде $A = HU$ и в виде $A = U_1 H_1$, где H и H_1 — симметрические (эрмитовы) матрицы порядков m и n, а Q и Q_1 — обобщенные ортогональные $(m \times n)$-матрицы, у которых при $m < n$ ортонормированные строки, а при $m > n$ ортонормированные столбцы. Это равносильно равенствам $UU^* = E$, $U_1 U_1^* = E$ при $m < n$ или $U^* U = E$, $U_1^* U_1 = E$ при $m < n$.

Указанные полярные разложения можно получить из сингулярного разложения (8.56) или (8.57). Пусть, например,

$m < n$. Тогда матрица A имеет сингулярное разложение $A = Q\Sigma_m P_1^*$. Преобразуем это разложение следующим образом:

$$A = (Q\Sigma_m Q^*)(QP_1^*).$$

Положив

$$H = Q\Sigma_m Q^*, \quad U = QP_1^*, \tag{8.60}$$

получим представление $A = HU$, в котором H — симметрическая (эрмитова) матрица, а U — матрица, удовлетворяющая условию $UU^* = QP_1^*P_1Q^* = QEQ^* = E$, т.е. обобщенная ортогональная. Аналогично получается представление $A = U_1H_1$, где

$$U_1 = QP_1^*, \quad H_1 = P_1\Sigma_m P_1^*. \tag{8.61}$$

Для матриц H и H_1 сохраняется связь с матрицей A:

$$H = Q\Sigma_m Q^* = \sqrt{Q\Sigma_m^2 Q^*} = \sqrt{Q\Sigma_m P_1^* P_1 \Sigma_m Q^*} = \sqrt{AA^*},$$

$$H_1 = P_1\Sigma_m P_1^* = \sqrt{P_1\Sigma_m^2 P_1^*} = \sqrt{P_1\Sigma_m Q^* Q\Sigma_m P_1^*} = \sqrt{A^*A}.$$

В случае $m > n$ необходимо использовать сингулярное разложение (8.57). Из этого представления получаем оба полярных разложения $A = HU$ и $A = U_1H_1$, в которых

$$H = Q_1\Sigma_n Q_1^*, \quad U = Q_1P^*, \tag{8.62}$$

$$U_1 = Q_1P, \quad H_1 = P\Sigma_n P^*. \tag{8.63}$$

Пример 8.32. Построить полярное разложение для матрицы

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

Решение. Для этой матрицы в примере 8.26 получено сингулярное разложение $A = Q\Sigma_m P_1^*$, в котором

$$Q = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \end{pmatrix}, \quad \Sigma = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

$$P = \begin{pmatrix} \frac{1}{2\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{2} \\ \frac{1}{2\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{2} \\ \frac{1}{2\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} & \frac{1}{2} \\ \frac{3}{2\sqrt{3}} & 0 & 0 & -\frac{1}{2} \end{pmatrix}.$$

Исходя из этого разложения, по формулам (8.60) и (8.61) находим:

$$H = Q\Sigma_m Q^* = \frac{1}{3}\begin{pmatrix} 4 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 4 \end{pmatrix}, \quad U = QP_1^* = \frac{1}{6}\begin{pmatrix} 5 & -1 & -1 & 3 \\ -1 & 5 & -1 & 3 \\ -1 & -1 & 5 & 3 \end{pmatrix}.$$

В результате получаем полярное разложение

$$A = HU = \frac{1}{3}\begin{pmatrix} 4 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 4 \end{pmatrix} \cdot \frac{1}{6}\begin{pmatrix} 5 & -1 & -1 & 3 \\ -1 & 5 & -1 & 3 \\ -1 & -1 & 5 & 3 \end{pmatrix}.$$

Аналогично

$$U_1 = U = QP_1^* = \frac{1}{6}\begin{pmatrix} 5 & -1 & -1 & 3 \\ -1 & 5 & -1 & 3 \\ -1 & -1 & 5 & 3 \end{pmatrix},$$

$$H_1 = P_1\Sigma_m P_1^* = \frac{1}{6}\begin{pmatrix} 5 & -1 & -1 & 3 \\ -1 & 5 & -1 & 3 \\ -1 & -1 & 5 & 3 \\ 3 & 3 & 3 & 9 \end{pmatrix}.$$

Отсюда

$$A = H_1U_1 = \frac{1}{6}\begin{pmatrix} 5 & -1 & -1 & 3 \\ -1 & 5 & -1 & 3 \\ -1 & -1 & 5 & 3 \\ 3 & 3 & 3 & 9 \end{pmatrix} \cdot \frac{1}{6}\begin{pmatrix} 5 & -1 & -1 & 3 \\ -1 & 5 & -1 & 3 \\ -1 & -1 & 5 & 3 \end{pmatrix}.$$

8.20. Псевдообратная матрица

Матрицу A^+ размера $n \times m$ называют ***псевдообратной*** к действительной или комплексной $(m \times n)$-матрице A, если выполняются соотношения

$$AA^+A = A, \qquad A^+ = UA^* = A^*V, \tag{8.64}$$

где U и V — некоторые матрицы.

Равенство $A^+ = UA^*$ означает, что строки матрицы A^+ являются линейными комбинациями строк матрицы A^*. Аналогично равенство $A^+ = A^*V$ означает, что столбцы матрицы A^+ являются линейными комбинациями столбцов матрицы A^*.

Теорема 8.37. *Для любой $(m \times n)$-матрицы A существует, и притом единственная, псевдообратная матрица A^+. Если ранг матрицы A совпадает с количеством ее столбцов (A имеет максимальный столбцовый ранг), то*

$$A^+ = (A^*A)^{-1}A^*. \tag{8.65}$$

Если ранг матрицы A совпадает с количеством строк (A имеет максимальный строчный ранг), то

$$A^+ = A^*(AA^*)^{-1}. \tag{8.66}$$

В общем случае, если $A = BC$ — скелетное разложение матрицы A, то

$$A^+ = C^+B^+ = C^*(CC^*)^{-1}(B^*B)^{-1}B^*. \tag{8.67}$$

▷ Единственность матрицы A^+. Пусть A_1^+ и A_2^+ — две матрицы, являющиеся псевдообратными к матрице A. Тогда в силу определения псевдообратной матрицы

$$AA_1^+A = A, \quad A_1^+ = U_1A^* = A^*V_1,$$
$$AA_2^+A = A, \quad A_2^+ = U_2A^* = A^*V_2.$$

Введем обозначения

$$D = A_1^+ - A_2^+, \quad U = U_1 - U_2, \quad V = V_1 - V_2.$$

Тогда:

$$ADA = A(A_1^+ - A_2^+)A = AA_1^+A - A_2^+A = A - A = 0,$$
$$D = A_1^+ - A_2^+ = U_1A^+ - U_2A^+ = (U_1 - U_2)A^+ = UA^+,$$
$$D = A_1^+ - A_2^+ = A^+V_1 - A^+V_2 = A^+(V_1 - V_2) = A^+V.$$

Отсюда следует, что

$$(DA)^*(DA) = A^*D^*DA = A^*(A^*V)^*DA = \\ = A^*V^*ADA = A^*V^*0 = 0,$$

т.е. матрица Грама системы столбцов матрицы DA равна нулю. Это равносильно утверждению, что ранг системы столбцов матрицы DA равен нулю, т.е. $DA = 0$. Теперь подсчитаем матрицу Грама системы столбцов матрицы D^*:

$$DD^* = D(UA^*)^* = DAU^* = 0U^* = 0.$$

Следовательно, $D^* = 0$ и $A_1^+ - A_2^+ = D = 0$. Мы пришли к выводу, что матрицы A_1^+ и A_2^+ равны.

Существование матрицы A^+. Доказательство существования псевдообратной матрицы сводится к проверке свойств (8.64) в предположении, что матрица A^+ вычислена в зависимости от ранга по одной из формул (8.65), (8.66) или (8.67).

Пусть матрица A имеет максимальный столбцовый ранг. В этом случае столбцы матрицы A линейно независимы, а матрица Грама A^*A системы столбцов матрицы A является невырожденной. Положим $A^+ = (A*A)^{-1}A^*$ и проверим выполнение соотношений (8.64). Имеем:

$$AA^+A = A(A*A)^{-1}A^*A = AE = A,$$

равенство $A^+ = UA^*$ выполняется при $U = (A^*A)^{-1}$, а равенство $A^+ = A^*V$ выполняется при $V = A(A^*A)^{-2}A^*$. Следовательно, матрица $A^+ = (A*A)^{-1}A^*$ является псевдообратной для матрицы A максимального столбцового ранга.

Аналогичным образом можно убедиться в том, что для матрицы A максимального строчного ранга матрица $A^+ = A^*(AA^*)^{-1}$ является псевдообратной. Рассмотрим произвольную матрицу A со скелетным разложением $A = BC$. Положим:

$$A^+ = C^+B^+ = C^*(CC^*)^{-1}(B^*B)^{-1}B^*.$$

Тогда

$$AA^+A = (BC)(C^*(CC^*)^{-1}(B^*B)^{-1}B^*)(BC) =$$
$$= B(CC^*)(CC^*)^{-1}(B^*B)^{-1}(B^*B)C = BC = A,$$

равенство $A = UA^*$ выполняется при

$$U = C^*(CC^*)^{-1}(B^*B)^{-1}(CC^*)^{-1}C,$$

а равенство $A = A^*V$ — при

$$V = B(B^*B)^{-1}(CC^*)^{-1}(B^*B)^{-1}B^*.$$

Поэтому указанная матрица A^+ действительно является псевдообратной к матрице A. ►

С л е д с т в и е. Для квадратной невырожденной матрицы A псевдообратная матрица A^+ совпадает с матрицей A^{-1}.

▷ В этом случае для вычисления псевдообратной матрицы можно воспользоваться одной из формул (8.65) или (8.66).

Учитывая, что матрицы A, A^* имеют обратные матрицы, из формулы (8.65) получаем

$$A^+ = (A^*A)^{-1}A^* = A^{-1}(A^*)^{-1}A^* = A^{-1}. \qquad \blacktriangleright$$

Лемма 8.3. *Если $(m \times n)$-матрица имеет разложение $A = RDS$ с ортогональными (унитарными) матрицами R, S и произвольной матрицей D, то*

$$A^+ = S^*D^+R^*. \tag{8.68}$$

$\triangleright$ По определению псевдообратной матрицы имеем

$$DD^+D = D, \quad D^+ = U_1D^* = D^*V_1.$$

Проверим выполнение аналогичных соотношений для матрицы $A = RDS$ в предположении, что $A^+ = S^*D^+R^*$. Из равенств $DD^+D = D$, $SS^* = E$, $R^*R = E$ вытекает, что

$$AA^+A = (RDS)(S^*D^+R^*)(RDS) = \\ = RD(SS^*)D^+(R^*R)DS = RDD^+DS = RDS = A.$$

Далее,

$$A^+ = S^*D^+R^* = S^*U_1D^*R^* = \\ = (S^*U_1S)(S^*D^*R^*) = (S^*U_1S)(RDS)^* = UA^*,$$

где $U = S^*U_1S$, и

$$A^+S^*D^+R^* = S^*D^*V_1R^* = \\ = (S^*D^*R^*)(RV_1R^*) = (RDS)^*(RV_1R^*) = A^*V,$$

где $V = RV_1R^*$. Так как для $A^+ = S^*D^+R^*$ выполнены все три равенства (8.64), эта матрица является псевдообратной к матрице $A = RDS$. $\blacktriangleright$

Лемма 8.4. Для матрицы

$$D = \begin{pmatrix} \lambda_1 & & & & & & & 0 \\ & \lambda_2 & & & & & & \\ & & \ddots & & & & & \\ & & & \lambda_r & & & & \\ & & & & 0 & & & \\ & & & & & 0 & & \\ & & & & & & \ddots & \\ 0 & & & & & & & 0 \end{pmatrix}$$

псевдообратной является матрица

$$D^+ = \begin{pmatrix} \lambda_1^{-1} & & & & & & & 0 \\ & \lambda_2^{-1} & & & & & & \\ & & \ddots & & & & & \\ & & & \lambda_r^{-1} & & & & \\ & & & & 0 & & & \\ & & & & & 0 & & \\ & & & & & & \ddots & \\ 0 & & & & & & & 0 \end{pmatrix}.$$

▷ Доказательство состоит в проверке равенств (8.64). Имеем:

$$DD^+D =$$

$$= \begin{pmatrix} \lambda_1 & & & & & 0 \\ & \ddots & & & & \\ & & \lambda_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ 0 & & & & & 0 \end{pmatrix} \begin{pmatrix} \lambda_1^{-1} & & & & & 0 \\ & \ddots & & & & \\ & & \lambda_r^{-1} & & & \\ & & & 0 & & \\ & & & & \ddots & \\ 0 & & & & & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 & & & & & 0 \\ & \ddots & & & & \\ & & \lambda_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ 0 & & & & & 0 \end{pmatrix} =$$

$$= \begin{pmatrix} 1 & & & & & 0 \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ 0 & & & & & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 & & & & & 0 \\ & \ddots & & & & \\ & & \lambda_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ 0 & & & & & 0 \end{pmatrix} = D.$$

Далее

$$D^+ = \begin{pmatrix} \lambda_1^{-1} & & & & \\ & \ddots & & & 0 \\ & & \lambda_r^{-1} & & \\ & & & 0 & \\ & 0 & & & \ddots & \\ & & & & & 0 \end{pmatrix} =$$

$$= \begin{pmatrix} \lambda_1^{-2} & & & & \\ & \ddots & & & 0 \\ & & \lambda_r^{-2} & & \\ & & & 0 & \\ & 0 & & & \ddots & \\ & & & & & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 & & & & \\ & \ddots & & & 0 \\ & & \lambda_r & & \\ & & & 0 & \\ & 0 & & & \ddots & \\ & & & & & 0 \end{pmatrix} = UD^*,$$

где

$$U = \begin{pmatrix} \lambda_1^{-2} & & & & \\ & \ddots & & & 0 \\ & & \lambda_r^{-2} & & \\ & & & 0 & \\ & 0 & & & \ddots & \\ & & & & & 0 \end{pmatrix}.$$

Аналогично

$$D^+ = \begin{pmatrix} \lambda_1^{-1} & & & & \\ & \ddots & & & 0 \\ & & \lambda_r^{-1} & & \\ & & & 0 & \\ & 0 & & & \ddots & \\ & & & & & 0 \end{pmatrix} =$$

$$= \begin{pmatrix} \lambda_1 & & & & \\ & \ddots & & & 0 \\ & & \lambda_r & & \\ & & & 0 & \\ & 0 & & & \ddots & \\ & & & & & 0 \end{pmatrix} \begin{pmatrix} \lambda_1^{-2} & & & & \\ & \ddots & & & 0 \\ & & \lambda_r^{-2} & & \\ & & & 0 & \\ & 0 & & & \ddots & \\ & & & & & 0 \end{pmatrix} = D^*V,$$

где

$$V = \begin{pmatrix} \lambda_1^{-2} & & & & 0 \\ & \ddots & & & \\ & & \lambda_r^{-2} & & \\ & & & 0 & \\ 0 & & & & \ddots \\ & & & & & 0 \end{pmatrix}. \quad \blacktriangleright$$

Из доказанных лемм вытекает следующее утверждение.

Теорема 8.38. *Если $A = U\Sigma V^*$ — сингулярное разложение матрицы A с диагональной матрицей*

$$\Sigma = \begin{pmatrix} \sigma_1 & & & & & & 0 \\ & \sigma_2 & & & & & \\ & & \ddots & & & & \\ & & & \sigma_r & & & \\ & & & & 0 & & \\ & & & & & 0 & \\ 0 & & & & & & \ddots \\ & & & & & & & 0 \end{pmatrix},$$

то $A^+ = V\Sigma^+U^*$, где

$$\Sigma^+ = \begin{pmatrix} \sigma_1^{-1} & & & & & & 0 \\ & \sigma_2^{-1} & & & & & \\ & & \ddots & & & & \\ & & & \sigma_r^{-1} & & & \\ & & & & 0 & & \\ & & & & & 0 & \\ 0 & & & & & & \ddots \\ & & & & & & & 0 \end{pmatrix}.$$

Если $A = P\Lambda P^{-1}$ — каноническое разложение матрицы A, в котором

$$\Lambda = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

то $A^+ = P\Lambda^+P^{-1}$, где

$$\Lambda^+ = \begin{pmatrix} \lambda_1^{-1} & & 0 \\ & \ddots & \\ 0 & & \lambda_n^{-1} \end{pmatrix}.$$

Эта теорема дает удобный способ вычисления псевдообратных матриц.

Пример 8.33. Найти матрицу, псевдообратную к матрице

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

Решение. Ранг матрицы A равен трем и совпадает с числом ее строк. Поэтому применима формула (8.65). Вычислив

$$AA^*=\begin{pmatrix}1&0&0&1\\0&1&0&1\\0&0&1&1\end{pmatrix}\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\\1&1&1\end{pmatrix}=\begin{pmatrix}2&1&1\\1&2&1\\1&1&2\end{pmatrix},$$

находим:

$$A^+=A^*(AA^*)^{-1}=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\\1&1&1\end{pmatrix}\begin{pmatrix}2&1&1\\1&2&1\\1&1&2\end{pmatrix}^{-1}=\frac{1}{4}\begin{pmatrix}3&-1&-1\\-1&3&-1\\-1&-1&3\\1&1&1\end{pmatrix}.$$

Для вычисления матрицы A^+ можно также воспользоваться сингулярным разложением матрицы A, найденным в примере 8.26. С помощью этого разложения в соответствии с теоремой 8.38 находим:

$$A^+=P\Sigma^+Q^*=\begin{pmatrix}\frac{1}{2\sqrt{3}}&\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{6}}&\frac{1}{2}\\\frac{1}{2\sqrt{3}}&-\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{6}}&\frac{1}{2}\\\frac{1}{2\sqrt{3}}&0&-\frac{2}{\sqrt{6}}&\frac{1}{2}\\\frac{3}{2\sqrt{3}}&0&0&-\frac{1}{2}.\end{pmatrix}\begin{pmatrix}\frac{1}{2}&0&0\\0&1&0\\0&0&1\\0&0&0\end{pmatrix}\times$$

$$\times\begin{pmatrix}\frac{1}{\sqrt{3}}&\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{6}}\\\frac{1}{\sqrt{3}}&-\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{6}}\\\frac{1}{\sqrt{3}}&0&-\frac{2}{\sqrt{6}}\end{pmatrix}=\frac{1}{4}\begin{pmatrix}3&-1&-1\\-1&3&-1\\-1&-1&3\\1&1&1\end{pmatrix}.$$

Пример 8.34. Найти матрицу, псевдообратную к матрице

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

Решение. В примере 4.5 найдено скелетное разложение

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

матрицы A.

Определив

$$CC^*=\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 2 \end{pmatrix}=\begin{pmatrix} 2 & 2 \\ 2 & 5 \end{pmatrix},$$

находим:

$$C^+=C^*(CC^*)^{-1}=\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 2 & 2 \\ 2 & 5 \end{pmatrix}^{-1}=\frac{1}{6}\begin{pmatrix} 5 & -2 \\ -2 & 2 \\ 1 & 2 \end{pmatrix}.$$

Аналогично, определив

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

находим:

$$B^+=(B^*B)^{-1}B^*=\begin{pmatrix} 5 & -3 \\ -3 & 3 \end{pmatrix}^{-1}\begin{pmatrix} 2 & -1 & 0 \\ -1 & 1 & 1 \end{pmatrix}=\frac{1}{6}\begin{pmatrix} 3 & 0 & 3 \\ 1 & 2 & 5 \end{pmatrix}.$$

Теперь по формуле (8.67) получаем:

$$A^+=C^+B^+=\frac{1}{36}\begin{pmatrix} 5 & -2 \\ -2 & 2 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 3 & 0 & 3 \\ 1 & 2 & 5 \end{pmatrix}=\frac{1}{36}\begin{pmatrix} 13 & -4 & 5 \\ -4 & 4 & 4 \\ 5 & 4 & 13 \end{pmatrix}.$$

Можно также использовать сингулярное разложение матрицы A, найденное в примере 8.27. С помощью этого разложения по теореме 8.38 получаем:

$$A^+ = P\Sigma^+Q^* = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{pmatrix} \times$$

$$\times \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{6}} \end{pmatrix} = \frac{1}{36}\begin{pmatrix} 13 & -4 & 5 \\ -4 & 4 & 4 \\ 5 & 4 & 13 \end{pmatrix}.$$

Пример 8.35. Найти матрицу, псевдообратную к матрице

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

Решение. Ранг матрицы A равен числу ее столбцов. Поэтому применима формула (8.65). Вычислив

$$A^*A = \begin{pmatrix} 1 & i & 1 \\ -i & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & i \\ -i & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 1+2i \\ 1-2i & 3 \end{pmatrix},$$

получаем:

$$A^+ = (A^*A)^{-1}A^* = \begin{pmatrix} 3 & 1+2i \\ 1-2i & 3 \end{pmatrix}^{-1} \begin{pmatrix} 1 & i & 1 \\ -i & 1 & 1 \end{pmatrix} =$$

$$= \frac{1}{4}\begin{pmatrix} 3 & -1-2i \\ -1+2i & 3 \end{pmatrix} \begin{pmatrix} 1 & i & 1 \\ -i & 1 & 1 \end{pmatrix} =$$

$$= \frac{1}{4}\begin{pmatrix} 1+i & -1+i & 2-2i \\ -1-i & 1-i & 2+2i \end{pmatrix}$$

Псевдообратная матрица обладает следующими свойствами.

1. $(A^*)^+ = (A^+)^*$.
2. $(A^+)^+ = A$.
3. $(AA^+)^* = AA^+$, $(AA^+)^2 = AA^+$.
4. $(A^+A)^* = A^+A$, $(A^+A)^2 = A^+A$.

$\triangleright$ Первое свойство вытекает из того факта, что если $A = BC$ — скелетное разложение матрицы A, то $A^* = C^*B^*$ — скелетное разложение матрицы A^*. Учитывая это, получаем

$$(A^*)^+ = (B^*)^+(C^*)^+ = B(B^*B)^{-1} \cdot (CC^*)^{-1}C = \\ = \bigl[C^*(CC^*)^{-1}(B^*B)^{-1}B^*\bigr]^* = (A^+)^*.$$

Второе свойство сначала докажем в частном случае, когда матрица A имеет максимальный столбцовый ранг. Тогда в соответствии с формулой (8.65) $A^+ = (A^*A)^{-1}A^*$, причем ранг матрицы A^+ совпадает с рангом матрицы A и матрица A^+ имеет максимальный строчный ранг. Поэтому

$$(A^+)^+ = (A^+)^*\bigl[A^+(A^+)^*)\bigr]^{-1} = \\ = A(A^*A)^{-1}\bigl[(A^*A)^{-1}A^*A(A^*A)^{-1}\bigr]^{-1} = A(A^*A)^{-1}A^*A = A.$$

Аналогично устанавливается равенство $(A^+)^+ = A$ для матриц максимального строчного ранга. Наконец, в общем случае, выбрав скелетное разложение $A = BC$ матрицы A, получаем скелетное разложение $A^+ = C^+B^+$ псевдообратной матрицы. Поэтому на основании разобранных частных случаев получаем

$$(A^+)^+ = (C^+B^+)^+ = (B^+)^+(C^+)^+ = BC = A.$$

Третье и четвертое свойства устанавливаются непосредственной проверкой с подстановкой в равенства выражения (8.67) для псевдообратной матрицы A^+. $\blacktriangleright$

Следующее утверждение фактически приводит к другому определению псевдообратной матрицы.

Теорема 8.39. *Матрица A^+ является псевдообратной к матрице A тогда и только тогда, когда выполняются равенства*

$$\begin{aligned} &AA^+A = A, && A^+AA^+ = A, \\ &(AA^+)^* = AA^+, && (A^+A)^* = A^+A. \end{aligned} \tag{8.69}$$

$\triangleright$ Необходимость условий (8.69) вытекает из определения псевдообратной матрицы и ее свойств 1–4. Первое условие вытекает из определения. Второе условие вытекает из того факта, что псевдообратной к матрице A^+ является матрица A (свойство 1). Третье и четвертое условия вытекают из свойств 3 и 4.

Установим достаточность условий (8.69). Тогда первое условие (8.64) выполнено. Применяя второе и третье условия (8.69), получаем

$$A^+ = A^+AA^+ = A^+(AA^+)^* = A^+(A^+)^*A^* = UA^*,$$

где $U = A^+(A^+)^*$. Следовательно, выполняется второе соотношение (8.64). Аналогично, используя второе и четвертое условия (8.69), приходим к третьему условию (8.64):

$$A^+ = A^+AA^+ = (A^+A)^*A^+ = A^*(A^+)^*A^* = A^*V,$$

где $V = (A^+)^*A^*$. Таким образом, условия (8.69) являются достаточными для того, чтобы матрица A^+ была псевдообратной к матрице A. $\blacktriangleright$

Возможны и другие способы введения псевдообратной матрицы. Сформулируем некоторые критерии. Матрица A^+ является псевдообратной к матрице A, если:

1) выполняются соотношения $A^*AA^+ = A^*$ и $A^+ = A^*V$, где V — некоторая матрица;

2) выполняется равенство $A^+ = \lim\limits_{t\to+0}(A^*A + tE)^{-1}A^*$;

3) выполняется равенство $A^+ = \lim\limits_{t\to+0} A^*(A^*A + tE)^{-1}$;

4) матрица A^+ является нормальным псевдорешением матричного уравнения $AX = E$, т.е. столбцы a_i^+, $i = 1, 2, \ldots, n$,

матрицы A^+ являются нормальными псевдорешениями линейных систем $Ax = e_i$, где e_i — i-й столбец единичной матрицы E;

5) для любого столбца b нормальное псевдорешение системы $Ax = b$ имеет вид: $x = A^+b$.

В связи с этими критериями отметим следующее. Нормальные псевдорешения обсуждаются ниже (см. разд. 8.21). Критерий 4 удобен при компьютерных вычислениях. Формула $x = A^+b$ представления нормального псевдорешения является естественным обобщением формулы $x = A^{-1}b$, представляющей решение системы с квадратной невырожденной матрицей.

Понятие „псевдообратная матрица“ было введено в 1920 г. В. Муром с помощью предела (критерии 2 и 3). Широкое применение псевдообратных матриц началось значительно позже, после того как Р. Пенроуз в 1955 г. ввел понятие псевдообратной матрицы через условия (8.69).

8.21. Решение систем линейных уравнений методом наименьших квадратов

Пусть дана действительная или комплексная система линейных уравнений

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = b_1, \\ \ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + \ldots + a_{mn}x_n = b_m, \end{cases} \tag{8.70}$$

в матричной форме имеющая вид:

$$Ax = b.$$

Если система (8.70) описывает реальный физический процесс, но оказывается несовместной, то естественно ввести коррективы в математическую модель и, в частности, изменить трактовку понятия „решение“. Например, часто возникает задача найти такие значения неизвестных $x_1, \ldots, x_n$, при которых функция

$$F(x) = |Ax - b|^2 = \sum_{i=1}^{m} \left| \sum_{j=1}^{n} a_{ij} x_j \right|^2, \tag{8.71}$$

где $x = (x_1,\ x_2,\ \ldots,\ x_n)^{\mathrm{T}}$, принимает наименьшее значение. Этот подход к понятию решения линейной системы составляет суть ***метода наименьших квадратов***. При этом найденный набор значений $x = (x_1,\ x_2,\ \ldots,\ x_n)^{\mathrm{T}}$ называют ***псевдорешением*** или ***обобщенным решением*** системы. Очевидно, что если система $Ax = b$ совместна, то ее псевдорешение представляет собой обычное решение, а любое решение системы можно искать как ее псевдорешение.

Описанный подход к понятию решения линейной системы объясняется следующими соображениями. Пусть некоторый процесс описывается некоторым набором параметров α_1, $\alpha_2, \ldots, \alpha_n, \beta$, причем из эмпирических исследований известно, что параметр β является линейной функцией от параметров $\alpha_1, \ldots, \alpha_n$, т.е.

$$\beta = \alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_n a_n.$$

Стремясь определить коэффициенты $a_1, \ldots, a_n$ этой зависимости, исследователи проводят ряд экспериментов, в результате которых измеряют в разных ситуациях значения параметров $\alpha_1, \ldots, \alpha_n, \beta$, а затем составляют следующую таблицу наблюдений (табл. 8.1).

Таблица 8.1

Номер эксперимента	Результаты измерений				
	α_1	α_2	$\ldots$	α_n	β
1	α_{11}	α_{12}	$\ldots$	α_{1n}	β_1
2	α_{21}	α_{22}	$\ldots$	α_{2n}	β_2
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
m	α_{m1}	α_{m2}	$\ldots$	α_{mn}	β_m

Подставив данные каждого наблюдения в рассматриваемую линейную функцию, заключаем, что неизвестные коэффициенты a_1, a_2, ..., a_n являются решением системы линейных уравнений

$$\begin{cases} \alpha_{11}a_1 + \alpha_{12}a_2 + \ldots + \alpha_{1n}a_n = \beta_1, \\ \alpha_{21}a_1 + \alpha_{22}a_2 + \ldots + \alpha_{2n}a_n = \beta_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \alpha_{m1}a_1 + \alpha_{m2}a_2 + \ldots + \alpha_{mn}a_n = \beta_m. \end{cases}$$

На практике количество измерений m значительно превышает количество неизвестных n, а записанная система, как правило, является несовместной вследствие погрешностей измерений. Естественно остановить свой выброр на таких значениях неизвестных коэффициентов a_1, a_2, ..., a_n, при которых уравнения системы наиболее близки к равенствам. Эта идея и приводит к методу наименьших квадратов.

Система может иметь неединственное псевдорешение. Среди всех псевдорешений системы ***нормальным псевдорешением*** называют то, которое имеет *наименьшую длину.* На таком псевдорешении функция

$$\Phi(x) = |x|^2 = \sum_{j=1}^{n} |x_j|^2 \tag{8.72}$$

принимает наименьшее значение.

Исходя из определения псевдорешений, укажем способ их нахождения. Для этого замечаем, что левая часть i-го уравнения системы (8.70) при любом наборе значений x_1, x_2, ..., x_n отличается от его правой части на некоторую невязку. Поэтому вместо системы (8.70) будем рассматривать систему

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n + z_1 = b_1, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n + z_m = b_m, \end{cases}$$

в матричной записи имеющей вид: $Ax + z = b$.

Для того чтобы функция $F(x) = |Ax - b|^2$ имела наименьшее значение, длина вектора z должна быть минимальной. Это возможно, лишь когда в равенстве $Ax + z = b$ вектор $Ax = A_1x_1 + A_2x_2 + \ldots + A_nx_n$ является ортогональной проекцией (см. разд. 9.5) вектора b на пространство $L(A)$ столбцов a_1, $a_2, \ldots, a_n$ матрицы A, а вектор z — ортогональной составляющей вектора b. Поэтому вектор z должен быть ортогональным к каждому вектору-столбцу a_j матрицы A. Следовательно, $(z, a_j) = 0$, $j = 1, 2, \ldots, n$. В матричной форме это означает, что $A^*z = 0$. Отсюда, умножив равенство $Ax + z = b$ слева на матрицу A^*, получим систему

$$A^*Ax = A^*b. \tag{8.73}$$

К системе (8.73) можно прийти также, используя методы математического анализа, а именно, приравняв к нулю дифференциал

$$dF(x) = 2dx^*(A^*Ax - A^*b)$$

функции

$$F(x) = |Ax - b|^2 = (Ax - b)^*(Ax - b).$$

Систему $A^*Ax = A^*b$ называют ***системой нормальных уравнений***. Она всегда совместна. Ее определителем является определитель матрицы Грама системы столбцов матрицы A.

Теорема 8.40. *Система $A^*Ax = A^*b$ всегда совместна, причем ее решения, и только они, являются псевдорешениями системы $Ax = b$.*

▷ Докажем, что вектор A^+b, где A^+ — псевдообратная к A матрица, является решением системы нормальных уравнений. Запишем какое-либо скелетное разложение $A = BC$ матрицы A. Тогда $A^+ = C^*(CC^*)^{-1}(B^*B)^{-1}B^*$ и

$$A^*A(A^+b) = C^*B^*BCC^*(CC^*)^{-1}(B^*B)^{-1}B^*b = \\ = C^*B^*B(B^*B)^{-1}B^*b = C^*B^*b = A^*b.$$

Равенство $A^*Ax = A^*b$ равносильно равенству $A^*(Ax - b) = 0$, означающему, что каждый столбец матрицы A ортогонален столбцу $Ax - b$. Следовательно, столбец Ax является проекцией вектора b на подпространство $L(A)$, порожденное столбцами матрицы A, а вектор x есть псевдорешение системы $Ax = b$. Эти рассуждения можно провести и в обратном порядке. Следовательно, псевдорешения являются решениями системы нормальных уравнений. ▶

Следствие. Система $A^*Ax = A^*b$ имеет единственное решение, если столбцы матрицы A линейно независимы, и бесконечное множество решений, если столбцы матрицы A линейно зависимы.

▷ Определитель системы $A^*Ax = A^*b$ есть определитель матрицы Грама системы столбцов матрицы A. Этот определитель отличен от нуля тогда и только тогда, когда столбцы матрицы A линейно независимы. В то же время это условие означает, что система $A^*Ax = A^*b$ имеет единственное решение. ▶

Итак, чтобы решить систему $Ax = b$ методом наименьших квадратов, нужно найти общее решение системы $A^*Ax = A^*b$ и, если требуется, из множества решений выделить нормальное решение, т.е. то, которое имеет наименьшую длину.

Пример 8.36. Для системы

$$\begin{cases} x_1 + x_4 = 4, \\ x_2 + x_4 = -4, \\ x_3 + x_4 = 4 \end{cases}$$

найти общее и нормальное псевдорешения.

Решение. Составим систему $A^*AX = A^*b$ нормальных уравнений, которая в данном случае имеет вид:

$$\begin{cases} x_1 + x_4 = 4, \\ x_2 + x_4 = -4, \\ x_3 + x_4 = 4, \\ x_1 + x_2 + x_3 + x_4 = 4. \end{cases}$$

Решив эту систему, найдем общее псевдорешение

$$x = (4 - x_4,\ -4 - x_4,\ 4 - x_4,\ x_4)^{\mathrm{T}}. \tag{8.74}$$

Теперь составим функцию

$$\Phi(X) = |X|^2 = (4 - x_4)^2 + (-4 - x_4)^2 + (4 - x_4)^2 + (x_4)^2.$$

Из равенства нулю ее производной по x_4 найдем $x_4 = 1$, при котором из общего псевдорешения X, заданного формулой (8.74), получается нормальное псевдорешение $x^0 = (3,\ -5,\ 3,\ 1)^{\mathrm{T}}$.

Пример 8.37. Найти многочлен $p(t) = x_1 + x_2 t + x_3 t^2$, с наименьшей квадратичной погрешностью приближающий функцию $f(t)$, которая задана таблицей

t	-2	-1	1	2
$f(t)$	$-5{,}5$	$0{,}5$	$3{,}5$	$3{,}0$

Решение. Подставляя в $p(t)$ данные из таблицы, придем к системе

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

Записанная система несовместная. Поэтому будем решать ее методом наименьших квадратов. Система $A^*Ax = A^*b$ нормальных уравнений в данном случае имеет вид:

$$\begin{cases} 4x_1 + 10x_3 = 1{,}5, \\ 10x_2 = 20 = 20, \\ 10x_1 + 34x_3 = -6. \end{cases}$$

Решив систему, получим

$$x_1 = \frac{55}{18}, \quad x_2 = 2, \quad x_3 = \frac{13}{12}.$$

Поэтому искомый многочлен $p(t)$ имеет вид:

$$p(t) = \frac{55}{18} + 2t - \frac{13}{12}t^2.$$

Пример 8.38. Найти нормальное псевдорешение системы

$$\begin{cases} x_1 + ix_2 = 1, \\ -ix_1 + x_2 = 1, \\ x_1 + x_2 = 1. \end{cases}$$

Решение. Следуя общему правилу, составим систему $A^*Ax = A^*b$ нормальных уравнений, которая в данном случае имеет вид:

$$\begin{cases} 3x_1 + (1+2i)x_2 = 2+i, \\ (1-2i)x_1 + 3x_2 = 2-i. \end{cases}$$

Решив эту систему, найдем ее единственное решение $x = \frac{1}{2}(1,\ 1)^{\mathrm{T}}$. Оно и является нормальным решением данной системы.

Искать псевдорешения линейной системы можно с помощью псевдообратной матрицы. Обсудим этот способ подробнее.

Лемма 8.5. Однородные системы $Ax = 0$ и $A^*Ax = 0$ эквивалентны.

▷ Если $Ax = 0$, то $A^*Ax = A^* \cdot 0 = 0$, т.е. все решения системы $Ax = 0$ являются и решениями системы $A^*Ax = 0$. Предположим, что столбец x является решением системы $A^*Ax = 0$. Тогда $x^*A^*Ax = x^* \cdot 0 = 0$. Но

$$x^*A^*Ax = (Ax)^*Ax = |Ax|^2.$$

Таким образом, из равенства $A^*Ax = 0$ вытекает, что столбец Ax имеет нулевую длину, а значит, он равен нулю, т.е. $Ax = 0$. ▶

Лемма 8.6. *Общее решение однородной системы $Ax=0$ представимо в виде*

$$x_{\text{одн}} = (E - A^{+}A)c, \tag{8.75}$$

где E — единичная матрица соответствующего порядка, $c=(c_1,\ c_2,\ c_n)^{\mathrm{T}}$ — вектор произвольных постоянных.

▷ При любом векторе c столбец $(E-A^{+}A)c$ является решением системы $Ax=0$, поскольку

$$A(E-A^{+}A)c = (A - AA^{+}A)c = (A-A)c = 0.$$

В то же время любое решение x_0 системы $Ax=0$ можно представить в виде $x_0 = (E-A^{+}A)x_0$, так как

$$(E-A^{+}A)x_0 = x_0 - A^{+}Ax_0 = x_0 - A^{+}\cdot 0 = x_0.$$

Это значит, что формула (8.75) содержит все решения системы $Ax=0$. ▶

Следствие. Общее решение системы $A^{*}Ax=0$ представимо в виде (8.75).

Теорема 8.41. *Общее псевдорешение системы $Ax=b$ представимо в виде*

$$x = A^{+}b + (E - A^{+}A)c, \tag{8.76}$$

где $c=(c_1,\ c_2,\ c_n)^{\mathrm{T}}$ — вектор произвольных постоянных.

▷ Общее псевдорешение системы $Ax=b$ совпадает с общим решением всегда совместной системы $A^{+}Ax=A^{*}b$. Одним из решений системы $A^{+}Ax=A^{*}b$ является столбец $A^{+}b$ (см. доказательство теоремы 8.40). Общее решение соответствующей однородной системы $A^{*}Ax=0$ может быть представлено в виде (8.75). Значит, общее решение неоднородной системы $A^{*}Ax=A^{*}b$ в соответствии с правилом „частное решение плюс общее решение однородной системы“ имеет вид (8.76). ▶

Теорема 8.42. *Нормальное псевдорешение системы $Ax = b$ есть столбец*

$$x^{\circ} = A^{+}b. \tag{8.77}$$

$\triangleright$ Воспользовавшись формулой (8.76) общего псевдорешения системы $Ax = b$, вычислим величину $|x|^2$ и выясним, для какого вектора c произвольных постоянных эта величина имеет наименьшее значение.

Полагая $x^{\circ} = A^{+}b$, $z = (E - A^{+}A)c$, получаем:

$$|x|^2 = |x^{\circ} + z|^2 = (x^{\circ} + z)^*(x^{\circ} + z) = |x^{\circ}|^2 + (x^{\circ})^* z + z^* x^{\circ} + |z|^2.$$

Поскольку столбец z является решением системы $Ax = 0$, а псевдообратная матрица A^{+} может быть представлена в виде $A^{+} = A^{*}V$, то

$$z^* x^{\circ} = z^* A^{+} b = z^* A^* V b = (Az)^{+} V b = 0 \cdot Vb = 0.$$

При этом и

$$(x^{\circ})^* z = (z^* x^{\circ})^* = 0^* = 0.$$

Значит,

$$|x|^2 = |x^{\circ}|^2 + |z|^2$$

и наименьшее значение псевдорешения $x = x^{\circ} + z$ достигается при $z = 0$, т.е. $x^{\circ} = A^{+}b$ и есть нормальное псевдорешение системы $Ax = b$. ▶

Пример 8.39. Вернемся к системе из примера 8.36. Для матрицы

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

этой системы в примере 8.33 найдена псевдообратная матрица

$$A^{+} = \frac{1}{4}\begin{pmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \\ 1 & 1 & 1 \end{pmatrix}.$$

По формуле (8.77) находим нормальное псевдорешение

$$x^\circ = A^+b = \frac{1}{4}\begin{pmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 4 \\ -4 \\ 4 \end{pmatrix} = \begin{pmatrix} 3 \\ -5 \\ 3 \\ 1 \end{pmatrix}.$$

Вычислив

$$E - A^+A = E - \frac{1}{4}\begin{pmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} =$$

$$= \frac{1}{4}\begin{pmatrix} 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & -1 \\ -1 & -1 & -1 & 1 \end{pmatrix},$$

затем по формуле (8.76) получаем общее псевдорешение

$$x_{\text{общ}} = A^+b + (E - A^+A)c =$$

$$= \begin{pmatrix} 3 \\ -5 \\ 3 \\ 1 \end{pmatrix} + \frac{1}{4}\begin{pmatrix} 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & -1 \\ -1 & -1 & -1 & 1 \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} =$$

$$= \frac{1}{4}\begin{pmatrix} 12 + (c_1 + c_2 + c_3 - c_4) \\ -15 + (c_1 + c_2 + c_3 - c_4) \\ 12 + (c_1 + c_2 + c_3 - c_4) \\ 4 - (c_1 + c_2 + c_3 - c_4) \end{pmatrix}.$$

Видно, что в найденном представлении общего решения произвольные постоянные связаны друг с другом. Положив $\gamma = x_4 = 1 - \frac{1}{4}(c_1 + c_2 + c_3 - c_4)$, придем к представлению общего решения в виде

$$x_{\text{общ}} = (4 - \gamma,\ -4 - \gamma,\ 4 - \gamma,\ \gamma)^{\mathrm{T}},$$

что эквивалентно (8.74).

Формулу $x^\circ = A^+b$ с помощью сингулярного разложения $A = Q\Sigma P^*$ матрицы A и с учетом теоремы 8.38 можно преобразовать следующим образом:

$$x^\circ = A^+b = P\Sigma^+Q^*b =$$
$$= (e_1, \ldots, e_n)\begin{pmatrix} \sigma_1^{-1} & & 0 \\ & \ddots & \\ 0 & & \sigma_r^{-1} \end{pmatrix}\begin{pmatrix} f_1^* \\ \vdots \\ f_m^* \end{pmatrix} b =$$
$$= \left(\frac{e_1}{\sigma_1}, \ldots, \frac{e_r}{\sigma_r}, 0, \ldots, 0\right)\begin{pmatrix} f_1^* \\ \vdots \\ f_m^* \end{pmatrix} b =$$
$$= \frac{1}{\sigma_1}e_1 f_1^* b + \ldots + \frac{1}{\sigma_r}e_r f_r^* b, \qquad (8.78)$$

где $e_1, e_2, \ldots, e_n$ — столбцы матрицы P, $f_1, f_2, \ldots, f_m$ — столбцы матрицы Q, $\sigma_1, \sigma_2, \ldots, \sigma_r$ — ненулевые сингулярные числа матрицы A. Поскольку (см. разд. 8.18)

$$f_i = \frac{Ae_i}{\sigma_i}, \quad e_i = \frac{A^*f_i}{\sigma_i}, \quad i = 1, 2, \ldots, r,$$

причем $e_1, e_2, \ldots, e_n$ и $f_1, f_2, \ldots, f_m$ — ортонормированные системы, получаем представления

$$x^\circ = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_r e_r, \qquad (8.79)$$

где

$$\alpha_i = \frac{(Ae_i, b)}{\sigma_i^2}, \quad i = 1, 2, \ldots, r, \qquad (8.80)$$

и

$$x^\circ = \beta_1 A^* f_1 + \beta_2 A^* f_2 + \ldots + \beta_r A^* f_r, \qquad (8.81)$$

где

$$\beta_i = \frac{(b, f_i)}{\sigma_i^2}, \quad i = 1, 2, \ldots, r.$$

Столбцы $e_{r+1}, e_{r+2}, \ldots, e_n$ матрицы P в сингулярном разложении $A = Q\Sigma P^*$, соответствующие нулевым сингулярным числам матрицы A, составляют фундаментальную систему решений системы уравнений $Ax = 0$.

Действительно, эти столбцы линейно независимы, а их количество равно $n - r$, где r — ранг матрицы A, равный общему числу ненулевых сингулярных чисел. Кроме того, каждый из этих столбцов является решением системы $Ax = 0$, поскольку

$$Ae_j = A\Sigma P^* e_j = Q\Sigma \begin{pmatrix} e_1^* \\ \vdots \\ e_n^* \end{pmatrix} e_j = Q\Sigma s_j = Q \cdot 0 = 0.$$

Здесь s_j — j-й столбец матрицы $P^*P = E$, а произведение Σs_j — j-й столбец матрицы Σ, при $j > r$ равный нулю. Поэтому ***общее псевдорешение системы $Ax = b$ можно записать в виде***

$$x_{\text{общ}} = x^\circ + \alpha_{r+1}e_{r+1} + \alpha_{r+2}e_{r+2} + \ldots + \alpha_n e_n,$$

или

$$x_{\text{общ}} = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n, \tag{8.82}$$

где первые r коэффициентов имеют фиксированные значения, определенные формулами (8.80), а остальные коэффициенты принимают произвольные значения.

Формула (8.82) удобна для получения проекций псевдорешений на подпространства, порожденные правыми сингулярными векторами $e_1, e_2, \ldots, e_n$ матрицы A. Проекции на подпространство, порожденное векторами $e_1, \ldots, e_k$, при $k \leqslant r$ можно также получить из соотношения (8.79) или, что то же самое, из системы

$$
Q\begin{pmatrix} \sigma_1 & & & & 0 \\ & \ddots & & & \\ & & \sigma_k & & \\ & & & 0 & \\ 0 & & & & \ddots \\ & & & & & 0 \end{pmatrix} P^* x = b. \tag{8.83}
$$

Пример 8.40. Найти проекции $x^{(k)}$, $k = 1, 2, 3, 4$, псевдорешений системы

$$
\begin{cases} x_1 + x_4 = 4, \\ x_2 + x_4 = -4, \\ x_3 + x_4 = 4 \end{cases}
$$

на подпространства, порожденные правыми сингулярными векторами матрицы системы.

Решение. Для матрицы данной системы в примере 8.26 найдены сингулярные числа $\sigma - 1 = 2$, $\sigma_2 = \sigma_3 = 1$, $\sigma_4 = 0$ и правые сингулярные векторы

$$
e_1 = \frac{1}{2\sqrt{3}}(1,\ 1,\ 1,\ 3)^{\mathrm{T}}, \qquad e_2 = \frac{1}{\sqrt{2}}(1,\ -1,\ 0,\ 0)^{\mathrm{T}},
$$

$$
e_3 = \frac{1}{\sqrt{6}}(1,\ 1,\ -2,\ 0)^{\mathrm{T}}, \qquad e_4 = \frac{1}{2}(1,\ 1,\ 1,\ -1)^{\mathrm{T}}.
$$

В соответствии с формулами (8.80) находим:

$$
\alpha_1 = \frac{(Ae_1, b)}{\sigma_1^2} = \frac{1}{2}(1,\ 1,\ 1)\begin{pmatrix} 4 \\ -4 \\ 4 \end{pmatrix} = \frac{2}{\sqrt{3}},
$$

$$
\alpha_2 = \frac{(Ae_2, b)}{\sigma_2^2} = \frac{1}{\sqrt{2}}(1,\ -1,\ 0)\begin{pmatrix} 4 \\ -4 \\ 4 \end{pmatrix} = \frac{8}{\sqrt{2}},
$$

$$
\alpha_3 = \frac{(Ae_3, b)}{\sigma_3^2} = \frac{1}{\sqrt{6}}(1,\ 1,\ 2)\begin{pmatrix} 4 \\ -4 \\ 4 \end{pmatrix} = -\frac{8}{\sqrt{6}},
$$

а α_4 — произвольное постоянное. Далее по формуле (8.82) получаем:

$$x^{(1)} = \alpha_1 e_1 = \frac{1}{3}(1,\ 1,\ 1,\ 3)^{\text{т}},$$

$$x^{(2)} = \alpha_1 e_1 + \alpha_2 e_2 = (13,\ -11,\ 1,\ 3)^{\text{т}},$$

$$x^{(3)} = \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3 = (3,\ -5,\ 3,\ 1)^{\text{т}},$$

$$x^{(4)} = \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3 + \alpha_4 e_4 = \\ = (3,\ -5,\ 3,\ 1)^{\text{т}} + \frac{\alpha_4}{2}(1,\ 1,\ 1,\ -1)^{\text{т}} = x_{\text{общ}}.$$

Эти результаты можно получить с помощью формулы (8.83), используя найденное в примере 8.26 сингулярное разложение:

$$x^{(1)} = P\begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} Q^* b = (1,\ 1,\ 1,\ 3)^{\text{т}},$$

$$x^{(2)} = P\begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} Q^* b = (13,\ -11,\ 1,\ 3)^{\text{т}},$$

$$x^{(3)} = P\begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} Q^* b = (3,\ -5,\ 3,\ 1)^{\text{т}}.$$

Отыскание нормального решения произвольной системы можно свести к решению одной или двух систем с невырожденными матрицами. Действительно, если дана совместная или несовместная система $Bx = b$ с $(m \times n)$-матрицей B ранга $r = n$, то, домножая ее на матрицу

B^* слева, придем к системе $B^*Bx = B^*b$ с невырожденной матрицей B^*B ранга n. Решая эту систему, получим: $x = (B^*B)^{-1}B^*b = B^+b = x^\circ$.

Если дана совместная или несовместная система $Cx = b$ с $(m \times n)$-матрицей C ранга $r = m$, то x будем искать в виде $X = C^*z$. Тогда придем к системе $CC^*z = b$ с невырожденной матрицей CC^* ранга m . Поэтому $x = C^*z = C^*(CC^*)^{-1}b = C^+b = x^\circ$.

Пусть, наконец, дана совместная или несовместная система $Ax = b$ с $(m \times n)$-матрицей A ранга $r \leqslant \min\{m, n\}$. Матрицу A представим скелетным разложением $A = BC$ с $(m \times n)$-матрицей B ранга r и $(r \times n)$-матрицей C ранга r. Тогда система $Ax = b$ примет вид $BCx = b$. Полагая $y = Cx$, придем к системе $By = b$, которая по только что рассмотренному первому случаю дает $y = Cx = B^+b$. Отсюда по второму случаю получаем

$$x = C^+y = C^+B^+b = A^+b = x^\circ.$$

Из приведенных рассуждений видно, что этот способ равносилен применению формулы $x^\circ = A^+b$ для отыскания нормального решения системы. Кроме того, отсюда становится понятной формула (8.67).

Пример 8.41. Найти нормальное решение системы

$$\begin{cases} 2x_1 - x_2 = 1, \\ -x_1 + x_2 + x_3 = 0, \\ x_2 + 2x_3 = 1 \end{cases}$$

путем сведения ее к решению систем с невырожденными матрицами.

Решение. В примере 4.5 для матрицы A рассматриваемой системы было найдено скелетное разложение

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

Используя это разложение, данную систему перепишем в виде

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

Полагая

$$y = Cx = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \end{pmatrix} x,$$

где $y = (y_1,\, y_2)^{\mathrm{T}}$, придем к системе

$$\begin{pmatrix} 2 & -1 \\ -1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Здесь имеем первый из рассмотренных выше случаев. Поэтому полученную систему домножаем слева на матрицу

$$B^* = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 1 & 1 \end{pmatrix}.$$

В результате придем к системе

$$\begin{pmatrix} 5 & -3 \\ -3 & 3 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}.$$

Эта система имеет невырожденную матрицу, следовательно, единственное решение. Находим это решение:

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \frac{1}{6} \begin{pmatrix} 3 & 3 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Теперь найденный столбец $y = (1,\, 1)^{\mathrm{T}}$ подставим в систему $Cx = y$. Получим систему

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \end{pmatrix} x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Здесь имеем второй из рассмотренных выше случаев. Поэтому будем искать столбец $X = (x_1, x_2, x_3)^{\mathrm{T}}$ в виде

$$x = C^* z = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}.$$

В результате придем к системе

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

или после перемножения матриц в левой части

$$\begin{pmatrix} 2 & 2 \\ 2 & 5 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Таким образом, снова получена система с невырожденной матрицей, решая которую находим:

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \frac{1}{6} \begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ 0 \end{pmatrix}.$$

Таким образом,

$$x = x^\circ = C^* z = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1/2 \\ 0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

8.22. Метод регуляризации для систем линейных уравнений

Пусть дана произвольная (совместная или несовместная) система $Ax = b$ из m линейных уравнений с действительными или комплексными коэффициентами и n неизвестными.

Для получения нормального псевдорешения этой системы, кроме указанных выше (см. разд. 8.21) способов, применим также ***метод регуляризации.*** Он состоит в следующем.

1. Составляют функцию

$$F_\alpha(x) = |Ax - b|^2 + \alpha|x|^2 = \\ = \sum_{i=1}^{m}\left|\sum_{j=1}^{n} a_{ij}x_j - b_j\right|^2 + \alpha\sum_{j=1}^{n}|x_j|^2, \quad (8.84)$$

где $\alpha > 0$, и находят вектор x^α, при котором эта функция достигает своего минимума. Таким вектором является решение системы

$$(A^*A + \alpha E)x = A^*b, \quad (8.85)$$

потому что для функции

$$F_\alpha(x) = |AX - b|^2 + \alpha|X|^2$$

дифференциал

$$dF_\alpha(x) = 2dx^*\left[A^*A + \alpha E)x - A^*b\right]$$

обращается в нуль лишь при таких x, которые являются решениями системы (8.85).

2. В полученном векторе x^α переходят к пределу при $\alpha \to \to +0$. Этот предел дает нормальное псевдорешение системы $Ax = b$.

При реализации обсуждаемого метода на ЭВМ функцию (8.84) рассматривают при конкретных значениях параметра α. При каждом таком значении α из системы (8.85) определяют конкретное приближение x^α к искомому нормальному псевдорешению системы $Ax = b$.

Для вычисления вектора x^α применимы также формулы

$$x^\alpha = (A^*A + \alpha E)^{-1}A^*b, \quad (8.86)$$

$$x^\alpha = \sum_{j=1}^{r}\frac{(Ae_j, b)}{\alpha + \sigma_j^2}e_j, \quad (8.87)$$

$$x^\alpha = \sum_{j=1}^{r}\frac{(f_j, b)}{\alpha + \sigma_j^2}A^*f, \quad (8.88)$$

где r — ранг $(m \times n)$-матрицы A; $\sigma_1, \sigma_2, \ldots, \sigma_r$ — ненулевые сингулярные числа матрицы A; $e_1, e_2, \ldots, e_n$ и $f_1, f_2, \ldots, f_m$ — сингулярные базисы.

Если матрица A порядка n и ранга r симметрическая, то применима формула

$$x^\alpha = \sum_{j=1}^{r} \frac{\lambda_j (e_j, b)}{\alpha + \lambda_j^2} e_j, \tag{8.89}$$

где $e_1, e_2, \ldots, e_n$ — ортонормированная система собственных векторов оператора с матрицей A, принадлежащих ненулевым собственным значениям $\lambda_1, \lambda_2, \ldots, \lambda_n$.

Пример 8.42. Методом регуляризации найти нормальное псевдорешение системы

$$\begin{cases} 2x_1 - x_2 = 1, \\ -x_1 + x_2 + x_3 = 0, \\ x_2 + 2x_3 = 1. \end{cases}$$

Решение. Составим систему $(A^*A + \alpha E) = A^*b$, которая в данном случае имеет вид:

$$\begin{cases} (5+\alpha)x_1 - 3x_2 - x_3 = 2, \\ 3x_1 + (3+\alpha)x_2 + 3x_3 = 0, \\ -x_1 + 3x_2 + (5+\alpha)x_3 = 2. \end{cases}$$

Решая эту систему, находим:

$$x^\alpha = \left(\frac{18 + 2\alpha}{36 + 13\alpha + \alpha^2}, \ 0, \ \frac{18 + 2\alpha}{36 + 13\alpha + \alpha^2} \right)^{\text{т}}.$$

Отсюда, переходя к пределу при $\alpha \to +0$, получим искомое нормальное псевдорешение

$$x^\circ = \frac{1}{2}(1, \ 0, \ 1)^{\text{т}}.$$

Для отыскания x^α по формуле (8.89) сначала найдем собственные значения $\lambda_1 = 3$, $\lambda_2 = 2$, $\lambda_3 = 0$ матрицы A и соответствующие им собственные векторы

$$e_1 = \frac{1}{\sqrt{3}}(1,\ -1,\ -1)^{\mathrm{T}}, \quad e_2 = \frac{1}{\sqrt{2}}(1,\ 0,\ 1)^{\mathrm{T}}, \quad e_3 = \frac{1}{\sqrt{6}}(1,\ 2,\ -1)^{\mathrm{T}}.$$

Затем по формуле (8.89) при $\alpha \to +0$ получаем:

$$x^\alpha = \frac{\lambda_1 (e_1, b)}{\alpha + \lambda_1^2} e_1 + \frac{\lambda_2 (e_2, b)}{\alpha + \lambda_2^2} e_2 = \frac{2}{\alpha + 4} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}^{\mathrm{T}} \to \frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}^{\mathrm{T}}.$$

Пример 8.43. Методом регуляризации найти нормальное псевдорешение системы

$$\begin{cases} x_1 + (1+i)x_2 = 1, \\ (1-i)x_1 + x_2 = 1, \\ x_1 + x_2 = 1. \end{cases}$$

Решение. Составим систему $(A^*A + \alpha E)x = A^*b$. В рассматриваемом случае она имеет вид

$$\begin{pmatrix} 4+\alpha & 3+2i \\ 3-2i & 4+\alpha \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3-i \\ 3+i \end{pmatrix}.$$

Решая эту систему, получим

$$x^\alpha = \left(\frac{1 + 3\alpha + (1+\alpha)i}{3 + 8\alpha + \alpha^2},\ \frac{1 + 3\alpha - (1+\alpha)i}{3 + 8\alpha + \alpha^2} \right)^{\mathrm{T}}.$$

Переходя к пределу в x^α при $\alpha \to +0$, находим искомое нормальное псевдорешение $x^\circ = \frac{1}{3}(1+i,\ 1-i)^{\mathrm{T}}$.

8.23. Нормы векторов и матриц

Пусть дано действительное (комплексное) линейное пространство X. Каждому вектору $x \in X$ поставим в соответствие действительное число $\|x\|$ и назовем его ***нормой***

вектора x, если для любых векторов $x, y \in X$ и любого действительного (комплексного) числа α выполняются следующие ***аксиомы нормы***:

1. $\|x\| > 0$ при $x \neq 0$ и $\|x\| = 0$ при $x = 0$.
2. $\|\alpha x\| = |\alpha| \, \|x\|$.
3. $\|x + y\| \leqslant \|x\| + \|y\|$.

Линейное пространство X при этом называется ***нормированным пространством***.

В арифметическом пространстве K_n наиболее употребительными являются:

1) ***октаэдрическая норма***

$$\|x\|_1 = \sum_{k=1}^{n} |x_k|;$$

2) ***евклидова***, или ***сферическая***, ***норма***

$$\|x\|_2 = \|x\|_E = \sqrt{\sum_{k=1}^{n} |x_k|^2};$$

3) ***кубическая норма***

$$\|x\|_\infty = \max_{1 \leqslant i \leqslant n} |x_i|.$$

Норму можно ввести в любом конечномерном пространстве. Если пространство евклидово (унитарное), то в нем можно ввести евклидову норму по формуле

$$\|x\| = \sqrt{(x, x)} = \|x\|_2 .$$

Как правило, такая норма подразумевается в евклидовом (унитарном) пространстве и поэтому евклидовы (унитарные) пространства относят к нормированным пространствам.

В линейном пространстве $(m \times n)$-матриц также рассматривают различные нормы. Наиболее употребительными являются:

1. $\|A\| = \sum\limits_{i=1}^{m}\sum\limits_{j=1}^{n}|a_{ij}|.$

2. $\|A\|_1 = \max\limits_{1\leqslant i\leqslant m}\sum\limits_{j=1}^{n}|a_{ij}|.$

3. $\|A\|_\infty = \max\limits_{1\leqslant j\leqslant n}\sum\limits_{i=1}^{m}|a_{ij}|.$

4. $\|A\|_E = \left(\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{n}|a_{ij}|^2\right)^{1/2}.$

Например, для матрицы

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 10 & 11 & 12 \end{pmatrix}$$

указанные нормы имеют следующие значения:

$$\|A\| = 1 + 2 + \ldots + 12 = 78,$$
$$\|A\|_1 = \max\{1+4+7+10,\ 2+5+8+11,\ 3+6+9+12\} = 30,$$
$$\|A\|_\infty = \max\{1+2+3,\ 4+5+6,\ 7+8+9,\ 10+11+12\} = 33,$$
$$\|A\|_E = \sqrt{1^2 + 2^2 + \ldots + 12^2} = \sqrt{650}.$$

В линейном пространстве квадратных матриц порядка n, кроме линейных операций, важную роль играет операция умножения матриц. В связи с этим в этом линейном пространстве предпочтение отдают нормам, согласованным с операцией умножения, а именно:

$$\|AB\| \leqslant \|A\|\,\|B\|.$$

При этом норму матриц, не подчиняющуюся этому неравенству, иногда называют ***обобщенной нормой матриц***.

Каждую $(m \times n)$-матрицу A можно интерпретировать как оператор, действующий из n-мерного арифметического пространства K_n в m-мерное арифметическое пространство K_m по формуле $y = Ax$, $x \in K_n$, $y \in K_m$. Если в K_n и K_m введены нормы, то желательно рассматривать ***норму матриц*** размера $m \times n$, ***согласованную с векторными нормами*** в K_n и K_m:

$$\|Ax\| \leqslant \|A\| \, \|x\|.$$

(Отметим, что это неравенство связывает сразу три нормы в трех разных линейных пространствах: в K_n, в K_m и в линейном пространстве $(m \times n)$-матриц.)

Примером такой нормы является ***матричная норма, индуцированная векторной нормой*** (или ***подчиненная векторной норме***)

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{\|x\| \leqslant 1} \|Ax\| = \sup_{\|x\| = 1} \|Ax\|.$$

Приведем примеры норм, индуцированных различными векторными нормами.

1. Для октаэдрической векторной нормы $\|x\|_1$ индуцированной является матричная норма $\|A\|_1$.

2. Для сферической векторной нормы $\|x\|_2$ индуцированной является ***спектральная норма***

$$\|A\|_c = \sqrt{\max\{|\lambda_1|, |\lambda_2|, \ldots, |\lambda_r|\}} = \sqrt{\max \lambda_{A^*A}},$$

где $\lambda_1, \lambda_2, \ldots, \lambda_r$ — собственные числа матрицы A^*A.

3. Для кубической векторной нормы $\|x\|_\infty$ индуцированной матричной нормой является норма $\|A\|_\infty$.

Между различными матричными нормами существуют определенные соотношения. Особенно много таких соотношений приведено в [34].

8.24. Оценка погрешности решения системы линейных уравнений

Для практических целей очень важно исследовать влияние малых изменений матрицы A системы $Ax = b$ (A — невырожденная матрица) и столбца b свободных членов на ее решение x. В общем случае система считается устойчивой, если малые возмущения в A и b приводят к малым возмущениям в x. Степень малости участвующих величин обычно измеряется в отношении к их значениям в невозмущенном состоянии, т.е. если

$$Ax = b \tag{8.90}$$

— точная система,

$$(A + \varepsilon_A)x = b + \varepsilon_b \tag{8.91}$$

— возмущенная система, то измеряются величины относительных возмущений

$$\delta A = \frac{\|\varepsilon_A\|}{\|A\|}, \qquad \delta A^{-1} = \frac{\left\|(A+\varepsilon_A)^{-1} - A^{-1}\right\|}{\|A^{-1}\|},$$

$$\delta b = \frac{\|\varepsilon_b\|}{\|b\|}, \qquad \delta x = \frac{\|\varepsilon_x\|}{\|x\|} = \frac{\|x_{\text{в}} - x_0\|}{\|x_0\|},$$

где $x_{\text{в}}$ — решение возмущенной системы (8.91), а x_0 — решение невозмущенной системы (8.90). При этом могут выбираться различные векторные и матричные нормы.

Предполагается, что не только матрица A, но и возмущенная матрица $A + \varepsilon_A$ невырожденная. Оказывается, что возмущенная матрица $A + \varepsilon_A$ будет невырожденной при всех возмущениях ε_A, удовлетворяющих условию

$$\|\varepsilon_A\| < \left\|A^{-1}\right\|^{-1}. \tag{8.92}$$

Если выполнено условие (8.92) и матричные нормы согласованы с векторными, то

$$\left\|(A+\varepsilon_A)^{-1}-A^{-1}\right\| \leqslant \frac{\left\|A^{-1}\right\|^2 \|\varepsilon_A\|}{1-\left\|A^{-1}\right\| \|\varepsilon_A\|}, \tag{8.93}$$

$$\delta A^{-1} \leqslant \frac{K_A \delta A}{1-K_A \delta A}, \tag{8.94}$$

где $K_A = \left\|A^{-1}\right\| \|A\|$ — ***число обусловленности матрицы*** A.

Если неравенство (8.92) выполнено при условии, что матричная норма подчинена векторной, то

$$\|\varepsilon_x\| = \|x_{\text{в}} - x_0\| \leqslant \frac{K_A \|x_0\|}{1-K_A \delta A}(\delta A + \delta b), \tag{8.95}$$

$$\delta x \leqslant \frac{K_A}{1-K_A \delta A}(\delta A + \delta b). \tag{8.96}$$

Неравенства (8.93)–(8.96) дают количественные оценки возмущения обратной матрицы и решения системы линейных алгебраических уравнений с невырожденной матрицей при возмущении системы. Они показывают, что обратная матрица и решение системы будут устойчивыми к возмущениям при не слишком большом числе обусловленности матрицы.

Пример 8.44. В системе

$$\begin{cases} x_1 + x_2 = 1, \\ x_1 - x_2 = 3 \end{cases}$$

элементы матрицы могут изменяться на ε, а свободные члены — на ε_1. Требуется оценить возможное изменение координат вектора-решения при $\varepsilon = \varepsilon_1 = 0{,}1$.

Решение. В данной системе

$$A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \quad A^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad x_0 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$

и по условию

$$\varepsilon_A = \begin{pmatrix} \varepsilon & \varepsilon \\ \varepsilon & \varepsilon \end{pmatrix}, \quad \varepsilon_b = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_1 \end{pmatrix}.$$

Для оценки возмущения координат вектора-решения используем кубическую норму $\|x\|_\infty = \max\limits_i |x_i|$. Подчиненной ей матричной нормой является

$$\|A\|_\infty = \max_i \sum_j |a_{ij}|.$$

В этом случае имеем:

$$\|A\|_\infty = 2, \quad \left\|A^{-1}\right\|_\infty = 1, \quad K_A = 2, \quad \|b\|_\infty = 3, \quad \|x_0\|_\infty = 2,$$

$$\|\varepsilon_A\|_\infty = 2\varepsilon < \left\|A^{-1}\right\|_\infty^{-1} = 1, \quad \|\varepsilon_b\|_\infty = \varepsilon_1,$$

$$\delta A = \frac{\|\varepsilon_A\|_\infty}{\|A\|_\infty} = \varepsilon, \quad \delta b = \frac{\|\varepsilon_b\|_\infty}{\|b\|_\infty} = \frac{\varepsilon_1}{3}.$$

Отсюда в соответствии с неравенствами (8.95), (8.96) получаем:

$$\|\varepsilon_x\|_\infty \leqslant \frac{1\cdot 2\cdot 2}{1-2\cdot 2\varepsilon}\left(\varepsilon + \frac{\varepsilon_1}{3}\right) = \frac{4}{1-2\varepsilon}\left(\varepsilon + \frac{\varepsilon_1}{3}\right),$$

$$\delta x \leqslant \frac{2}{1-2\varepsilon}\left(\varepsilon + \frac{\varepsilon_1}{3}\right).$$

В частности, если $\varepsilon = \varepsilon_1 = 0{,}1$, то

$$\|\varepsilon_x\|_\infty \leqslant \frac{4}{1-0{,}2}\left(0{,}1 + \frac{0{,}1}{3}\right) = \frac{2}{3}, \qquad \delta x \leqslant \frac{2}{1-0{,}2}\left(0{,}1 + \frac{0{,}1}{3}\right) = \frac{1}{3},$$

т.е. при $\varepsilon = \varepsilon_1 = 0{,}1$ координаты вектора-решения данной системы могут изменяться не более чем на $2/3$, а их относительное изменение не превосходит $1/3$.

Пример 8.45. В системе

$$\begin{cases} x_1 + x_2 = 3, \\ x_1 - x_2 = 4 \end{cases}$$

элементы главной диагонали матрицы могут изменяться на ε_1, а свободные члены — на ε_2. Оценить возможное изменение вектора-решения по длине при $\varepsilon = \varepsilon_1 = 0{,}01$.

Решение. Для оценки возмущения вектора-решения по длине естественно использовать евклидову векторную норму $\|x\|_2$. Этой векторной нормой индуцируется спектральная матричная норма $\|A\|_c = \sqrt{\max \lambda_{A^* A}}$.

В данном случае матрица

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

симметрическая. Поэтому спектральная норма матрицы A совпадает с максимальным из модулей ее собственных чисел, т.е. $\|A\|_c = \max |\lambda_A| = \sqrt{2}$, так как характеристический многочлен $|A - \lambda E| = \lambda^2 - 2$ имеет корни $\lambda_{1,2} = \pm\sqrt{2}$.

Поскольку собственные значения матриц A и A^{-1} взаимно обратные, получаем:

$$\|A^{-1}\|_c = \max \frac{1}{|\lambda_A|} = \frac{1}{\sqrt{2}}.$$

Матрица

$$\varepsilon_A = \begin{pmatrix} \varepsilon_1 & 0 \\ 0 & \varepsilon_1 \end{pmatrix}$$

является симметрической, а ее характеристический многочлен $|\varepsilon_A - \lambda E| = (\varepsilon_1 - \lambda)^2$ имеет корни $\lambda_{1,2} = \varepsilon_1$. Поэтому

$$\|\varepsilon_A\|_c = \varepsilon_1, \quad \delta A = \frac{\|\varepsilon_A\|_c}{\|A\|_c} = \frac{\varepsilon_1}{\sqrt{2}}.$$

Далее имеем:

$$\|b\|_2 = \sqrt{3^2 + 4^2} = 5, \qquad \|\varepsilon_b\|_2 = \sqrt{2}\varepsilon_2, \qquad \delta b = \frac{\|\varepsilon_b\|_2}{\|b\|_2} = \frac{\sqrt{2}}{5}\varepsilon_2,$$

$$x_0 = \left(\frac{7}{2}, -\frac{1}{2}\right)^{\mathrm{T}}, \qquad \|x_0\|_2 = \frac{1}{2}\sqrt{7^2 + 1} = \frac{5}{\sqrt{2}}.$$

Поэтому в соответствии с формулами (8.95) и (8.96) получаем

$$\|\varepsilon_x\|_2 \leqslant \frac{5\sqrt{2}}{\sqrt{2}-\varepsilon_1}\left(\frac{\varepsilon_1}{2}+\frac{\varepsilon_2}{5}\right), \qquad \delta x \leqslant \frac{2}{\sqrt{2}-\varepsilon_1}\left(\frac{\varepsilon_1}{2}+\frac{\varepsilon_2}{5}\right).$$

В частности, при $\varepsilon_1 = \varepsilon_2 = 0{,}01$ $\|\varepsilon_x\|_2 \leqslant 0{,}035$, $\delta x \leqslant 0{,}01$, т.е. длина вектора-решения может измениться не более чем на 0,035, а его относительное изменение по длине не превосходит 0,01.

8.25. Отыскание устойчивого решения системы линейных уравнений

Пусть даны точная система $Ax = b$ и ее возмущенная система $\widetilde{A}x = \widetilde{b}$.

Точная система с произвольной $(m \times n)$-матрицей A ранга r может быть как совместной (в частности, матрица может быть невырожденной), так и несовместной. Поэтому естественно вести речь о нормальном псевдорешении x° точной системы и о его возмущении при возмущении системы. В общем случае даже при малых возмущениях системы возможны большие возмущения нормального псевдорешения. Возникает вопрос, нельзя ли нормальное псевдорешение системы $Ax = b$ находить приближенно устойчивым образом и давать оценку погрешности такого приближения при возмущении системы. Оказывается, что эти вопросы можно решить положительно. Для этого следует в качестве приближения к нормальному псевдорешению брать определенную проекцию $\widetilde{x}^{(k)}$ псевдорешения возмущенной системы на подпространства правых сингулярных векторов (см. разд. 8.21), т.е. полагать $x^\circ \approx \widetilde{x}^{(k)}$ при определенным образом выбранном номере k. Можно также в качестве приближения нормального псевдорешения x° брать решение $\widetilde{x}^\alpha$, найденное методом регуляризации (см. разд. 8.22) для возмущенной системы, при некотором значении параметра α, т.е. полагать $X^\circ \approx \widetilde{x}^\alpha$ при определенным образом выбранном параметре α. Поясним такой подход на примере.

Пример 8.46. Пусть система

$$\begin{cases} 2x_1 - x_2 = 1, \\ -x_1 + x_2 + x_3 = 0, \\ x_2 + 2x_3 = 1 \end{cases}$$

с нормальным псевдорешением $x^\circ = \frac{1}{2}(1,\ 0,\ 1)^{\mathrm{T}}$ (см. пример 8.42) при возмущении перешла в систему

$$\begin{cases} 2x_1 - x_2 = 1, \\ -x_1 + x_2 + x_3 = 0, \\ x_2 + (2+\varepsilon)x_3 = 1. \end{cases}$$

Найдем устойчивым образом нормальное псевдорешение x° точной системы, решая возмущенную систему. С этой целью применим метод регуляризации (см. разд. 8.22) к возмущенной системе. Для этого составим систему $(\widetilde{A}^*\widetilde{A} + \alpha E)x = \widetilde{A}^*\widetilde{b}$, которая в нашем случае имеет вид:

$$\begin{cases} (5+\alpha)x_1 - 3x_2 - x_3 = 2, \\ -3x_1 + (3+\alpha)x_2 + (3+\varepsilon)x_3 = 0, \\ -x_1 + (3+\varepsilon)x_2 + (5+4\varepsilon+\varepsilon^2+\alpha)x_3 = 2+\varepsilon. \end{cases}$$

Решая эту систему, получим

$$\widetilde{x}^\alpha = \left(\frac{\Delta_1}{\Delta},\ \frac{\Delta_2}{\Delta},\ \frac{\Delta_3}{\Delta}\right)^{\mathrm{T}},$$

где

$$\Delta = 36\alpha + 13\alpha^2 + \alpha^3 + 26\alpha\varepsilon + 7\alpha\varepsilon^2 + 4\alpha^2\varepsilon^3 + \alpha^2\varepsilon^3 + \varepsilon^2,$$

$$\Delta_1 = 18\alpha + 2\alpha^2 + 9\alpha\varepsilon + 2\alpha\varepsilon^2 + \varepsilon^2,$$

$$\Delta_2 = \varepsilon^2 - 5\alpha\varepsilon - \alpha\varepsilon^2, \qquad \Delta_3 = 18\alpha + 2\alpha^2 + 8\alpha\varepsilon + \alpha^2\varepsilon.$$

При $\alpha \to +0$, $\widetilde{x}^\alpha \to \widetilde{x}^\circ = (1,\ 1,\ 0)^{\mathrm{T}}$. Этот вектор в два раза длиннее вектора нормального решения $x^\circ = \frac{1}{2}(1,\ 0,\ 1)^{\mathrm{T}}$ точной системы и составляет с x° угол $\frac{\pi}{3}$. Если же взять $\alpha = \varepsilon$, то $\widetilde{x}^\alpha$ уже достаточно точно совпадает с x°.

К такому же результату придем, если найдем нормальное псевдорешение $\widetilde{x}^{\circ} = (1,\ 1,\ 0)^{\mathrm{T}}$ возмущенной системы, вычислим его проекцию $\widetilde{x}^{(2)}$ и положим $x^{\circ} = \widetilde{x}^{(2)}$.

Данный подход обоснован в [5] (параграфы 16 и 17). Приведем оттуда основные результаты.

Обозначим через $\sigma_1,\ \sigma_2,\ \ldots,\ \sigma_s$ сингулярные числа матрицы A ранга r и будем считать, что $\sigma_1 \geqslant \sigma_2 \geqslant \ldots \geqslant \sigma_s$ ($\sigma_i \neq 0$ при $i = 1, 2, \ldots, r$). Соответствующие сингулярные числа матрицы $\widetilde{A}$ обозначим через $\widetilde{\sigma}_1,\ \widetilde{\sigma}_2,\ \ldots,\ \widetilde{\sigma}_s$. Предположим, что найдены псевдорешение $\widetilde{x}$ возмущенной системы и его проекции $\widetilde{x}^{(k)}$, $k = 1, 2, \ldots, r$ на пространства $L_k = \langle \widetilde{e}_1,\ \widetilde{e}_2,\ \ldots,\ \widetilde{e}_k \rangle$ правых сингулярных векторов $\widetilde{e}_1, \widetilde{e}_2, \ldots, \widetilde{e}_k$, соответствующих сингулярным числам $\widetilde{\sigma}_1,\ \widetilde{\sigma}_2,\ \ldots,\ \widetilde{\sigma}_s$. Пусть известно, что точная система совместная и возмущения ее малы по сравнению с наименьшим ненулевым сингулярным числом σ_r матрицы A, т.е. возмущение элементов матрицы A и возмущение свободных членов системы $Ax = b$ достаточно малы по сравнению с σ_r. Тогда для нормального псевдорешения x° точной системы имеет место приближенное равенство

$$x^{\circ} \approx \widetilde{x}^{(k)} \ \text{при}\ k = r, \tag{8.97}$$

причем

$$\delta x^{\circ} \leqslant K_A^{+}(\delta A + \delta b), \tag{8.98}$$

где

$$\delta x^{\circ} = \frac{\left\| x^{\circ} - \widetilde{x}^{(r)} \right\|}{\| x^{\circ} \|_E}, \qquad \delta A = \frac{\| \varepsilon_A \|}{\| A \|_E}, \qquad \delta b = \frac{\| \varepsilon_b \|}{\| b \|_E},$$

а $K_A^{+} = \left\| A^{+} \right\|_E \| A \|_E$ — обобщенное число обусловленности матрицы A.

Формулы (8.97) и (8.98) сохраняются и для почти совместной точной системы $Ax = b$, т.е. такой системы, в которой вектор

$$b = (b_r,\ 0)^{\mathrm{T}} + \left(0,\ b_r'\right)^{\mathrm{T}}$$

имеет достаточно малое слагаемое $(0,\ b_r')^{\mathrm{T}}$.

Заметим, что формула (8.98) аналогична формуле (8.96). При этом она показывает, что если точная система совместная или почти совместная и возмущение мало по сравнению с наименьшим ненулевым сингулярным числом матрицы системы, то нормальное псевдорешение можно определить по формуле (8.97) с той же точностью, что и для системы с невырожденной матрицей.

В случае несовместности точной системы $Ax = b$ и малого ее возмущения по сравнению с наименьшим ненулевым сингулярным числом также имеет место соотношение (8.97) и вместо (8.98) выполняется соотношение

$$\delta x^\circ \leqslant K_A^+(\delta A + \delta b) + K_A^+(K_A^+ \delta A + \delta b)\frac{\|b_r'\|_E}{\|b_r\|_E}. \tag{8.99}$$

Эта формула показывает, что точность приближения x° с помощью $\widetilde{x}^{(r)}$ в значительной мере зависит от отношений $\frac{\|b_r'\|_E}{\|b_r\|_E}$, т.е. от степени согласованности правой и левой частей точной системы $Ax = b$.

Формулы (8.97)–(8.99) верны при возмущениях, достаточно малых по сравнению с наименьшим ненулевым сингулярным числом σ_r матрицы системы. В общем случае, когда входные данные системы $Ax = b$ заданы с точностью порядка ε, одна из проекций $\widetilde{x}^{(k)}$ приближает нормальное псевдорешение x° с точностью порядка $(n\varepsilon)^{2/3}$, если точная система совместная, и с точностью порядка $(n\varepsilon)^{1/2}$, если точная система несовместная.

К аналогичному результату придем, если вместо приближенного равенства (8.97) пользоваться приближенным равенством

$$x^\circ \approx \widetilde{x}^\alpha \tag{8.100}$$

при некотором значении α, т.е. если нормальное псевдорешение x° системы $Ax = b$ приближать с помощью векторов $\widetilde{x}^\alpha$, которые находятся методом регуляризации для возмущенной системы $\widetilde{A}x = \widetilde{b}$, так как при этом оказывается верным следующее утверждение.

Если входные данные системы $Ax = b$ заданы с точностью порядка ε, то при некотором значении α вектор $\widetilde{x}^{\alpha}$ приближает нормальное псевдорешение x° системы $Ax = b$ с точностью порядка $\varepsilon^{2/3}$ в случае ее совместности и с точностью $\varepsilon^{1/2}$ в противном случае. При этом если $\alpha = \varepsilon^{2/3}$ в первом случае и $\alpha = \varepsilon^{1/2}$ во втором случае, то $\widetilde{x}^{\alpha}$ обеспечивает почти наилучшее приближение к x°.

Пример 8.47. В системах

$$\begin{cases} 2x_1 - x_2 = 1, \\ -x_1 + x_2 + x_3 = 0, \\ x_2 + 2x_3 = 1 \end{cases} \tag{8.101}$$

и

$$\begin{cases} 2x_1 - x_2 = 2, \\ -x_1 + x_2 + x_3 = 3, \\ x_2 + 2x_3 = -1 \end{cases} \tag{8.102}$$

элементы матрицы и свободные члены могут изменяться на $\varepsilon = 0{,}01$. Оценить возможную погрешность приближения нормального псевдорешения каждой из данных систем.

Решение. Здесь точная система (8.101) совместная, а система (8.102) несовместная, и наименьшее ненулевое сингулярное число матрицы этих систем $\sigma_2 = 2$ (см. пример 8.27). Оно значительно больше возмущений элементов матрицы и свободных членов. Поэтому для обеих систем применима формула (8.97) при $k = 2$, по которой $x^{\circ} \approx \widetilde{x}^{(2)}$, а погрешность этого приближенного равенства для системы (8.101) оценивается по формуле (8.98), для системы (8.102) — по формуле (8.99). Чтобы применить формулы (8.98) и (8.99), вычисляем для обеих систем

$$\|A\|_E = \sqrt{13}, \qquad \|A^+\|_E = \frac{\sqrt{13}}{6}, \qquad \|\varepsilon_A\|_E = 3\varepsilon,$$

$$\delta A = \frac{\|\varepsilon_A\|_E}{\|A\|_E} = \frac{3\varepsilon}{\sqrt{13}}, \quad K_A^+ = \|A^+\|_E \, \|A\|_E = \frac{13}{6}.$$

Кроме того, вычисляем для системы (8.102):

$$\|b\|_E = \sqrt{14}, \qquad \|\varepsilon_b\|_E = \varepsilon\sqrt{3}, \qquad \delta b = \frac{\|\varepsilon_b\|_E}{\|b\|_E} = \frac{\varepsilon\sqrt{3}}{\sqrt{14}},$$

$$\|b_2\|_E = \sqrt{13}, \qquad \|b_2'\|_E = 1$$

и для системы (8.101):

$$\|b\|_E = \sqrt{2}, \qquad \|\varepsilon_b\|_E = \varepsilon\sqrt{3}, \qquad \delta b = \frac{\|\varepsilon_b\|_E}{\|b\|_E} = \frac{\varepsilon\sqrt{3}}{\sqrt{2}}.$$

Теперь по формуле (8.98) получаем оценку возможной погрешности приближенного равенства $x^\circ \approx \widetilde{x}^{(2)}$ для системы (8.101):

$$\delta x^\circ \leqslant \left\|A^+\right\|_E \|A\|_E (\delta A + \delta b) = \frac{13}{6}\left(\frac{3\sqrt{13}\varepsilon}{13} + \frac{\varepsilon\sqrt{3}}{\sqrt{2}}\right).$$

В частности, при $\varepsilon = 0{,}01$ имеем: $\delta x^\circ \leqslant 0{,}02$, т.е. относительное изменение вектора-решения x° системы (8.101) по длине не превосходит 0,02 (ср. с результатами примера 1).

Для системы (8.102) формула (8.99) даст следующую оценку возможной погрешности приближенного равенства $x^\circ \approx \approx \widetilde{x}^{(2)}$:

$$\delta x^\circ \leqslant \frac{13}{6}\left(\frac{3\varepsilon}{\sqrt{13}} + \frac{\varepsilon\sqrt{3}}{\sqrt{14}}\right) + \frac{\sqrt{13}}{6}\left(\frac{13\varepsilon}{2\sqrt{13}} + \frac{\varepsilon\sqrt{3}}{\sqrt{14}}\right).$$

В частности, при $\varepsilon = 0{,}01$ имеем: $\delta x^\circ \leqslant 0{,}042$.

Если пользоваться приближенным равенством (8.100), то оно дает приближение нормального псевдорешения для системы (8.101) при $\alpha = \varepsilon^{2/3}$ с точностью порядка $\varepsilon^{2/3} \approx 0{,}047$, а для системы (8.102) при $\alpha = \varepsilon^{1/2}$ — с точностью порядка $\varepsilon^{1/2} \approx 0{,}1$.

Результаты, отмеченные в этом и предыдущем разделах, имеют исключительное значение для обоснования большинства численных методов решения систем линейных алгебраических уравнений, а также при выработке общей тактики решения таких систем на ЭВМ (см. [5, гл. 5]).

8.26. Рекомендации к решению систем линейных уравнений на ЭВМ

Существует много различных способов для отыскания решений системы линейных уравнений $Ax = b$ на ЭВМ. Общеизвестны, например, метод Гаусса и различные его модификации для отыскания решений совместных систем. Из других методов многие сводятся к составлению системы нормальных уравнений (см. разд. 8.21) $A^*Ax = A^*b$ и к отысканию общего или нормального решения этой совместной системы. Однако переход к системе нормальных уравнений часто бывает нежелательным, так как матрица A^*A системы нормальных уравнений, как правило, является во много раз хуже обусловленной, чем матрица A исходной системы $Ax = b$.

Классические вычислительные методы для совместных систем оказываются не всегда хорошими при применении их на ЭВМ. Они во многих случаях обладают способностью накапливать ошибки вычислений. Для системы $Ax = b$, у которой ранг матрицы A меньше, чем n, и для системы с большими размерами m и n матрицы A применение таких методов на ЭВМ может привести в результате накапливания ошибок к решениям, весьма далеким от истинных.

Существует надежный способ отыскания псевдорешений системы $Ax = b$ на ЭВМ, основанный на представлении ее матрицы A сингулярным разложением (см. разд. 8.18)

$$A = U\Sigma V^*, \tag{8.103}$$

где U — ортогональная (унитарная) матрица порядка m, V — ортогональная (унитарная) матрица порядка n,

$$\Sigma = \begin{pmatrix} \sigma_1 & & & & 0 \\ & \ddots & & & \\ & & \sigma_n & & \\ & & & 0 & \\ 0 & & & & \ddots \\ & & & & & 0 \end{pmatrix}$$

— диагональная $(m \times n)$-матрица.

Вместо сингулярного разложения (8.103) можно использовать любое разложение вида $A = R\Lambda S$ с ортогональными (унитарными) матрицами R и S и диагональной матрицей Λ. В частности, можно использовать каноническое разложение $A = P\lambda P^{-1}$ с ортогональной (унитарной) матрицей P (см. разд. 8.9 и 8.13), если для матрицы A существуют такие разложения.

Этот способ позволяет находить нормальное и общее псевдорешения совместной или несовместной системы $Ax = b$ без перехода к системе нормальных уравнений. Он одновременно приводит и к псевдообратной матрице и с ее помощью к *нормальному* и *общему* псевдорешениям системы $Ax = b$ (см. разд. 8.20 и 8.21). Кроме того, в этом способе предусмотрен учет неточностей во входных данных и погрешностей вычисления на ЭВМ. Указанный способ реализован в стандартных программах математического обеспечения (например, программа svd).

Поясним рассматриваемый метод. Используя сингулярное разложение (8.103) матрицы A, систему $Ax = b$ можно представить в виде

$$U\Sigma V^* x = b. \tag{8.104}$$

Полагая $z = V^* x$, $d = U^{-1}b = U^* b$, получаем систему

$$\Sigma z = d, \tag{8.105}$$

которая в подробной записи имеет вид

$$\begin{cases} \sigma_1 z_1 = d_1, \\ \sigma_2 z_2 = d_2, \\ \quad . \; . \; . \; . \; . \\ \sigma_n z_n = d_n, \\ 0 = d_{n+1}, \\ \quad . \; . \; . \; . \; . \\ 0 = d_m. \end{cases} \tag{8.106}$$

Для учета неточности входных данных и погрешности вычислений вводят границу τ. Например, если данные вводятся с тремя верными десятичными знаками, то подходящим будет значение $\tau = 10^{-3}$. Если данные введены точно, то значение для τ принимают равным порядку машинного „эпсилон“, умноженному на n. Выбор границы определяет, какую проекцию нормального решения возмущенной системы $U\Sigma V^*x = b$ на подпространства правых сингулярных векторов взять за основу, чтобы, используя ее, строить общее и нормальное решения системы $Ax = b$ (см. формулы (8.83) и (8.99)). Делают это следующим образом. Всякое σ_j, большее, чем τ, считают приемлемым, и из системы (8.106) находят соответствующее ему $z_j = d_j/\sigma_j$. Любое σ_j, меньшее, чем τ, считают пренебрежимо малым и соответствующему ему z_j придают произвольное значение.

По найденному вектору $z = (z_1, z_2, \ldots, z_r, c_1, c_2, \ldots, c_{n-r})^{\mathrm{T}}$ из равенства $z = V^*x$ находится общее псевдорешение $x_{\text{общ}} = (x_1, x_2, \ldots, x_n)^{\mathrm{T}}$ системы $Ax = b$ по формуле

$$x_{\text{общ}} = Vz_{\text{общ}}. \tag{8.107}$$

В частности, при $z = z^\circ = (z_1, \ldots, z_r, 0, \ldots, 0)^{\mathrm{T}}$ получается нормальное решение системы $Ax = b$:

$$x^\circ = Vz^\circ. \tag{8.108}$$

Поскольку $z_{\text{общ}} = z^\circ + (0, \ldots, 0, c_1, \ldots, c_{n-r})^{\mathrm{T}}$, то формулу (8.107) можно переписать в виде

$$\begin{aligned} x_{\text{общ}} = Vz^\circ + V(0, \ldots, 0, c_1, \ldots, c_{n-r})^{\mathrm{T}} = \\ = x^\circ + c_1v_{r+1} + c_2v_{r+2} + \ldots + c_{n-r}v_n, \end{aligned} \tag{8.109}$$

где $v_{r+1}, v_r + 2, \ldots, v_n$ — столбцы матрицы V, соответствующие нулевым сингулярным числам матрицы A. Столбцы v_{r+1}, $v_{r+2}, \ldots, v_n$ составляют фундаментальную систему решений однородной системы $Ax = 0$ (см. разд. 8.21).

Для практических целей обычно находят нормальное решение x° или какое-либо близкое к нему частное решение, получаемое из (8.109) при достаточно малых значениях постоянных c_1, c_2, ..., c_{n-r}. Но иногда приходится находить (см. формулу (8.83)) решения систем

$$U\begin{pmatrix} \Sigma_{kk} & 0 \\ 0 & 0 \end{pmatrix} V^* x = b,$$

где Σ_{kk} — матрица, построенная на k первых строках и k первых столбцах матрицы Σ, и из них выбирать наиболее подходящее решение. В этом состоит содержание сингулярного анализа (см. разд. 8.25 и [5]). Подробные разъяснения по сингулярному анализу можно найти в [18]. Там же даны рекомендации по практическому анализу задач методом наименьших квадратов.

На практике приведенный метод оказывается не намного более трудоемким, чем другие. Однако его применение не требует перехода к системе нормальных уравнений и гарантирует высокую надежность и точность получаемых решений (это следует из изложенных выше результатов, см. разд. 8.25). Кроме того, применение сингулярного разложения $A = U\Sigma V^*$ позволяет от системы $Ax = b$ перейти к системам $U\Sigma V^* x = b$ и $\Sigma z = d$. В этих системах легко представить себе структуру данных, уяснить смысл некорректности системы и увидеть пути ее реализации. Другими словами, применение сингулярного разложения позволяет наглядно проводить анализ данных и целенаправленно планировать эксперимент, приводящий к системе линейных уравнений. Действительно, по виду системы $\Sigma z = d$, имеющей вид (8.106), сразу замечаем, что для ее совместности необходимо и достаточно равенство нулю всех d_i при $i = r+1, \ldots, m$. Чем больше эти числа отличаются от нуля, тем будет большей некорректность системы $\Sigma z = d$, а следовательно, и системы $Ax = b$. Для исправления положения, очевидно, нужно попытаться так организовать эксперимент или так построить его математическую модель, приводящую к системе линейных уравнений, чтобы указанная некорректность линейных систем по крайней мере уменьшилась.

По виду системы $U\Sigma V^*x = b$ легко также сделать заключение о числе обусловленности K_A матрицы этой системы, так как это число равно σ_1/σ_s, где σ_1 и σ_s — наибольшее и наименьшее сингулярные числа матрицы A (при $\sigma_s = 0$ число обусловленности матрицы A считают равным бесконечности). Большая величина числа обусловленности матрицы A указывает на то, что математическая модель, приводящая к системе $Ax = b$, не является удачной. Если рассматриваемую модель сохраним без изменений, то при решении системы $Ax = b$ указанным способом придется более тщательно выбирать границу τ, которая отражает точность данных и точность машинной арифметики, или использовать проекции псевдорешений на подпространства правых сингулярных векторов, т.е. проводить сингулярный анализ. Однако надежнее изменить математическую модель, приводящую к системе линейных уравнений. Поясним это на примере (см. [33]) выбора функции, выравнивающей полученные результаты наблюдений (табл. 8.2).

Таблица 8.2

t	1900	1910	...	1970
$y(t)$	b_1	b_2	...	b_8

Если эти данные выравнивать функцией

$$y(t) = C_1 \cdot 1 + C_2 t + C_3 t^2, \tag{8.110}$$

то придем к системе уравнений для определения неизвестных C_1, C_2, C_3 с (8×3)-матрицей

$$A = \begin{pmatrix} 1 & 1900 & 1900^2 \\ 1 & 1910 & 1910^2 \\ \cdot & \cdots & \cdots \\ 1 & 1970 & 1970^2 \end{pmatrix}.$$

Ее сингулярными числами являются

$$\sigma_1 = 0{,}106 \cdot 10^8, \quad \sigma_2 = 0{,}648 \cdot 10^2, \quad \sigma_3 = 0{,}346 \cdot 10^{-3}.$$

Число обусловленности $\sigma_1/\sigma_3 = 0{,}306 \cdot 10^{11}$ матрицы A весьма велико. Следовательно, вид выравнивающей функции (8.110) выбран не совсем удачно. В ней базисные функции 1, t, t^2 для значений между 1900 и 1970 близки к линейной зависимости, так как σ_3 близко к нулю. Для исправления положения естественно изменить базисные функции и тем самым изменить вид выравнивающей функции. Вместо функции (8.110) можно взять, например, функцию

$$y(t) = C_1 \cdot 1 + C_2(t - 1900) + C_3(t - 1900)^2 \tag{8.111}$$

или функцию

$$y(t) = C_1 \cdot 1 + C_2 \frac{t - 1935}{10} + C_3 \left(\frac{t - 1935}{10}\right)^2 \tag{8.112}$$

соответственно с матрицами наблюдений

$$A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 10 & 10^2 \\ 1 & 20 & 20^2 \\ 1 & 30 & 30^2 \\ 1 & 40 & 40^2 \\ 1 & 50 & 50^2 \\ 1 & 60 & 60^2 \\ 1 & 70 & 70^2 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 1 & -3.5 & 3.5^2 \\ 1 & -2.5 & 2.5^2 \\ 1 & -1.5 & 1.5^2 \\ 1 & -0.5 & 0.5^2 \\ 1 & 0.5 & 0.5^2 \\ 1 & 1.5 & 1.5^2 \\ 1 & 2.5 & 2.5^2 \\ 1 & 3.5 & 3.5^2 \end{pmatrix}.$$

Сингулярными числами матрицы A_1 в модели (8.111) будут числа $\sigma_1 = 0{,}684 \cdot 10^4$, $\sigma_2 = 0{,}293 \cdot 10^2$, $\sigma_3 = 0{,}199 \cdot 10$, а матрицы A_2 в модели (8.112) — числа $\sigma_1 = 0{,}198 \cdot 10^2$, $\sigma_2 = 0{,}648 \cdot 10$, $\sigma_3 = 0{,}185 \cdot 10$. Числа обусловленности σ_1/σ_3 этих матриц соответственно равны $0{,}575 \cdot 10^4$ и 10,7. Последнее число обусловленности совсем невелико. Поэтому выравнивающая модель (8.112) более надежна и приемлема по сравнению с моделью (8.111) и особенно по сравнению с моделью (8.110).

При конструировании выравнивающей функции можно также увеличить степень многочлена, представляющего эту функцию, или выразить ее через какие-либо другие базисные функции, например, через $\sin t$, $\sin 2t$, $\ldots$, $\sin nt$, и с помощью сингулярного анализа выбирать наиболее приемлемую модель. Более подробное обсуждение сингулярного анализа приведено в [18].

Другим надежным методом вычисления на ЭВМ нормального решения системы является метод регуляризации (см. разд. 8.22 и 8.25).

Для решения систем линейных уравнений эффективными могут оказаться также итерационные методы. Они особенно удобны при больших размерах матрицы системы. (Об итерационных методах см. гл. 10.)

Таким образом, систему линейных уравнений на ЭВМ следует решать либо указанным выше приемом, основанным на применении сингулярного разложения, либо придерживаться следующей тактики.

Если известно, что исходная система устойчива к возмущениям (хорошо обусловлена), то ее можно решать любым обычным алгоритмом, например, методом Гаусса или какой-либо его модификацией. Если известно, что исходная система неустойчива к возмущениям (плохо обусловлена), то для отыскания ее нормального решения следует применять метод регуляризации, причем следует сразу полагать $\alpha = \varepsilon^{2/3}$ в случае совместности системы и $\alpha = \varepsilon^{1/2}$ — в случае ее несовместности (ε — граница точности входных данных).

Если об устойчивости системы ничего не известно, то систему упрощают и попутно выясняют вопрос об ее устойчивости к возмущениям. Если система окажется устойчивой к возмущениям, то продолжают ее решение обычными методами. Если же система окажется неустойчивой к возмущениям, то далее ищут ее нормальное решение методом регуляризации и в нем сразу полагают $\alpha = \varepsilon^{2/3}$ в случае совместности системы и $\alpha = \varepsilon^{1/2}$ — в случае ее несовместности. При таких значениях параметра α обеспечивается наилучшее приближение к искомому нормальному решению.

Об оценках погрешности на каждом этапе рассматриваемой схемы см. [5]. При больших размерах матрицы системы могут возникнуть трудности в возможностях памяти машины, и тогда более эффективными могут оказаться итерационные методы (см. гл. 10).

Упражнения

8.1. Пусть $x = (x_1,\ x_2)^{\mathrm{T}}$, $y = (y_1,\ y_2)^{\mathrm{T}}$ — произвольные векторы двумерного линейного пространства, заданные координатами в фиксированном базисе e_1, e_2. Убедиться, что скалярное произведение векторов в этом пространстве можно задать следующими способами:

1) $(x, y) = x_1y_1 + x_2y_2$;
2) $(x, y) = 2x_1y_1 + 3x_2y_2$;
3) $(x, y) = x_1y_1 + x_1y_2 + x_2y_1 + 2x_2x_2$;
4) $(x, y) = 2x_1y_1 - x_1y_2 - x_2y_1 + x_2y_2$;
5) $(x, y) = 2x_1y_1 + 2x_1y_2 + 2x_2y_1 + 3x_2y_2$;
6) $(x, y) = 3x_1y_1 - 2x_1y_2 - 2x_2y_1 + 5x_2y_2$.

Вычислить скалярное произведение векторов $x = (1,\ 1)^{\mathrm{T}}$ и $y = (2,\ -1)^{\mathrm{T}}$ при каждом из указанных способов его задания.

8.2. Пусть в двумерном линейном пространстве в базисе e_1, e_2 скалярное произведение задано формулой

$$(x, y) = \alpha_{11}x_1y_1 + \alpha_{12}x_2y_2 + \alpha_{21}x_2y_1 + \alpha_{22}x_2y_2,$$

где α_{ij} — произвольные фиксированные числа. Убедиться в том, что $\alpha_{ij} = (e_i, e_j)$, $i, j = 1, 2$, и записать матрицу Грама базиса e_1, e_2.

8.3. В трехмерном линейном пространстве задана матрица Грама

$$\Gamma = \begin{pmatrix} 5 & 2 & -4 \\ 2 & 1 & -2 \\ -4 & -2 & 5 \end{pmatrix}$$

базиса e_1, e_2, e_3. Записать формулу скалярного умножения векторов и, пользуясь ею, вычислить скалярные произведения

(e_i, e_j), $i, j = 1, 2, 3$, а также скалярное произведение векторов x и y, если:

1) $x = (1,\ 2,\ 3)^{\text{т}}$, $y = (0,\ 1,\ 2)^{\text{т}}$;
2) $x = (1,\ 1,\ 2)^{\text{т}}$, $y = (1,\ 0,\ 1)^{\text{т}}$;
3) $x = (2,\ 0,\ 1)^{\text{т}}$, $y = (1,\ 3,\ 1)^{\text{т}}$.

8.4. Считая, что векторы заданы координатами в ортонормированном базисе, с помощью процесса ортогонализации ортогонализировать следующие системы векторов:

1) $a_1 = (1,\ -2,\ 2)^{\text{т}}$, $a_2 = (-1,\ 0,\ -1)^{\text{т}}$, $a_3 = (5,\ -3,\ -7)^{\text{т}}$;
2) $a_1 = (1,\ 1,\ 1,\ 1)^{\text{т}}$, $a_2 = (3,\ 3,\ -1,\ -1)^{\text{т}}$, $a_3 = (-2,\ 0,\ 6,\ 8)^{\text{т}}$;
3) $a_1 = (1,\ 1,\ 1,\ 1)^{\text{т}}$, $a_2 = (0,\ 1,\ 1,\ 1)^{\text{т}}$, $a_3 = (0,\ 0,\ 1,\ 1)^{\text{т}}$.

8.5. Ортонормировать системы векторов предыдущего упражнения.

8.6. Применяя процесс ортогонализации, ортогонализировать систему векторов $a_1 = (1,\ 0,\ 1)^{\text{т}}$, $a_2 = (0,\ 1,\ 1)^{\text{т}}$, $a_3 = (0,\ 0,\ 1)^{\text{т}}$, заданных координатами в базисе e_1, e_2, e_3 с матрицей Грама

$$\Gamma = \begin{pmatrix} 2 & -2 & -1 \\ -2 & 5 & 2 \\ -1 & 2 & 2 \end{pmatrix}.$$

8.7. Считая, что векторы заданы координатами в ортонормированном базисе, показать, что следующие системы векторов ортогональны, и дополнить их до ортогональных базисов пространства:

1) $a_1 = (1,\ -2,\ 1,\ 3)^{\text{т}}$, $a_2 = (2,\ 1,\ -3,\ 1)^{\text{т}}$;
2) $a_1 = (1,\ -1,\ 1,\ -3)^{\text{т}}$, $a_2 = (-4,\ 1,\ 5,\ 0)^{\text{т}}$;
3) $a_1 = (1,\ 1,\ 1,\ 2)^{\text{т}}$, $a_2 = (1,\ 2,\ 3,\ -3)^{\text{т}}$.

8.8. Дополнить следующие ортонормированные системы векторов, заданных координатами в ортонормированном базисе, до ортонормированных базисов пространства:

1) $a_1 = \left(\frac{1}{2},\ -\frac{1}{2},\ \frac{1}{2}, -\frac{1}{2}\right)^{\text{т}}$, $a_2 = \left(-\frac{1}{2},\ \frac{1}{2},\ \frac{1}{2},\ -\frac{1}{2}\right)^{\text{т}}$;

2) $a_1 = \left(\frac{1}{2},\ \frac{1}{2},\ \frac{1}{2},\ \frac{1}{2}\right)^{\text{т}}$, $a_2 = \left(\frac{1}{2},\ \frac{1}{2},\ -\frac{1}{2},\ -\frac{1}{2}\right)^{\text{т}}$.

√ **8.9.** Применяя процесс ортогонализации, построить ортогональный базис пространства, натянутого на следующие *системы векторов, заданных координатами в ортонормированном базисе:*

1) $a_1 = (2,\ 3,\ -4,\ -6)^{\text{т}}$,
$a_2 = (1,\ 8,\ -2,\ -16)^{\text{т}}$,
$a_3 = (12,\ 5,\ -14,\ 5)^{\text{т}}$,
$a_4 = (3,\ 11,\ 4,\ 7)^{\text{т}}$;

2) $a_1 = (1,\ 1,\ -1,\ -2)^{\text{т}}$,
$a_2 = (-2,\ 1,\ 5,\ 11)^{\text{т}}$,
$a_3 = (0,\ 3,\ 3,\ 7)^{\text{т}}$,
$a_4 = (3,\ -3,\ -3,\ -9)^{\text{т}}$.

8.10. Найти базис ортогонального дополнения $L^{\perp}$ подпространства L, натянутого на векторы a_1, a_2, a_3, заданные координатами в ортонормированном базисе:

1) $a_1 = (1,\ 3,\ 0,\ 2)^{\text{т}}$, $a_2 = (3,\ 7,\ -1,\ 2)^{\text{т}}$, $a_3 = (2,\ 4,\ -1,\ 0)^{\text{т}}$;

2) $a_1 = (1,\ 2,\ 2,\ -1)^{\text{т}}$, $a_2 = (1,\ 1,\ -5,\ 3)^{\text{т}}$, $a_3 = (3,\ 2,\ 8,\ -7)^{\text{т}}$;

3) $a_1 = (1,\ 1,\ -1,\ -2)^{\text{т}}$, $a_2 = (5,\ 8,\ -2,\ -3)^{\text{т}}$, $a_3 = (3,\ 9,\ 3,\ 8)^{\text{т}}$.

8.11. Найти системы линейных уравнений, задающих подпространства L и $L^{\perp}$ предыдущего упражнения.

8.12. Найти систему уравнений, задающую ортогональное дополнение $L^{\perp}$, если подпространство L задано системой уравнений

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

8.13. Найти ортогональную проекцию x_0 и ортогональную составляющую $x^{\perp}$ вектора x на линейное пространство $L = \langle a_1,\ a_2,\ a_3 \rangle$, если:

1) $x = (14,\ -3,\ -6,\ -7)^{\text{т}}$,
$a_1 = (-3,\ 0,\ 7,\ 6)^{\text{т}}$,
$a_2 = (1,\ 4,\ 3,\ 2)^{\text{т}}$,
$a_3 = (2,\ 2,\ -2,\ -2)^{\text{т}}$;

2) $x = (2,\ -5,\ 3,\ 4)^{\text{т}}$,
$a_1 = (1,\ 3,\ 3,\ 5)^{\text{т}}$,
$a_2 = (1,\ 3,\ -5,\ -3)^{\text{т}}$,
$a_3 = (1,\ -5,\ 3,\ -3)^{\text{т}}$.

8.14. Найти ортогональную проекцию x_0 и ортогональную составляющую $x^{\perp}$ вектора $x = (-3,\ 0,\ -5,\ 9)^{\text{т}}$ на подпространство L, заданное системой уравнений

$$\begin{cases} 3x_1 + 2x_2 + x_3 - 2x_4 = 0, \\ 5x_1 + 4x_2 + 3x_3 + 2x_4 = 0, \\ x_1 - 2x_2 + 3x_3 + 10x_4 = 0. \end{cases}$$

8.15. Считая, что векторы заданы в ортонормированном базисе, ортонормировать систему векторов

1) $a_1 = (1,\ 1,\ i)^{\text{т}}$, $a_2 = (i,\ 1,\ 1)^{\text{т}}$, $a_3 = (1,\ i,\ 1)^{\text{т}}$;
2) $a_1 = (1,\ i,\ i)^{\text{т}}$, $a_2 = (1,\ i,\ 1)^{\text{т}}$, $a_3 = (i,\ i,\ i)^{\text{т}}$;
3) $a_1 = (1,\ 1,\ i)^{\text{т}}$, $a_2 = (i,\ i,\ 1)^{\text{т}}$, $a_3 = (i,\ 1,\ i)^{\text{т}}$.

8.16. Считая, что векторы заданы в ортонормированном базисе, убедиться в ортогональности системы векторов a_1, a_2 и дополнить эту систему векторов до ортогонального базиса пространства:

1) $a_1 = (1,\ i,\ 1,\ i)^{\text{т}}$, $a_2 = (1,\ i,\ -1,\ i)^{\text{т}}$;
2) $a_1 = (1,\ 1,\ i,\ i)^{\text{т}}$, $a_2 = (1,\ -1,\ i,\ -i)^{\text{т}}$;
3) $a_1 = (1+i,\ 1+i,\ 1,\ 1)^{\text{т}}$, $a_2 = (1-i,\ -1+i,\ 1,\ -1)^{\text{т}}$.

8.17. Найти матрицу перехода от базиса e_1, e_2, e_3 к базису

$$[e'_1]_e = \frac{1}{9}(4+3i,\ -4i,\ 6+2i)^{\text{т}}, \qquad [e'_2]_e = \frac{1}{9}(4i,\ 4-3i,\ -2-6i)^{\text{т}},$$

$$[e'_3]_e = \frac{1}{9}(-6-2i,\ -2-6i,\ 1)^{\text{т}}$$

и убедиться в том, что полученная матрица является унитарной.

8.18. Пусть e_1, e_2 — ортонормированный базис в X и линейный оператор φ в базисе $e'_1 = e_1$, $e'_2 = e_1 + e_2$ имеет матрицу

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

Найти матрицу сопряженного оператора φ^* в базисе e'_1, e'_2.

8.19. Линейный оператор φ в базисе $e_1' = (1,\ 2,\ 1)^{\mathrm{T}}$, $e_2' = (1,\ 1,\ 2)^{\mathrm{T}}$, $e_3' = (1,\ 1,\ 0)^{\mathrm{T}}$ имеет матрицу

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

Найти матрицу сопряженного оператора φ^* в том же базисе, считая, что векторы базиса e' даны координатами в ортонормированном базисе e.

8.20. Пусть скалярное произведение в базисе e задано формулой

$$(x, y) = 2x_1y_1 + 3x_2y_2 + x_3y_3 + 2x_1y_2 + \\ + 2x_2y_1 + x_1y_3 + x_3y_1 + x_2y_3 + x_3y_2,$$

а линейный оператор φ — матрицей

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

Найти матрицу сопряженного оператора φ^* в том же базисе.

8.21. Пусть даны матрица Грама Γ базиса e и матрица A линейного оператора φ в этом базисе:

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

Найти матрицу сопряженного оператора φ^* в том же базисе.

8.22. Построить канонические разложения с ортогональными трансформирующими матрицами для следующих симметрических матриц:

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

4) $\begin{pmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{pmatrix}$; 5) $\begin{pmatrix} 1 & -2 & 2 \\ -2 & -2 & 4 \\ 2 & 4 & -2 \end{pmatrix}$; 6) $\begin{pmatrix} 0 & 2 & 5 \\ 2 & -3 & -2 \\ 5 & -2 & 0 \end{pmatrix}$;

7) $\begin{pmatrix} 3 & -2 & -4 \\ -2 & 6 & -2 \\ -4 & -2 & 3 \end{pmatrix}$; 8) $\begin{pmatrix} 5 & -2 & -2 \\ -2 & 6 & 0 \\ -2 & 0 & 4 \end{pmatrix}$; 9) $\begin{pmatrix} 11 & 2 & -8 \\ 2 & 2 & 10 \\ -8 & 10 & -5 \end{pmatrix}$.

8.23. Убедиться в положительной определенности матрицы и найти квадратный корень из нее:

1) $\begin{pmatrix} 2 & -2 & -1 \\ -2 & 5 & 2 \\ -1 & 2 & 2 \end{pmatrix}$; 2) $\begin{pmatrix} 13 & 14 & -4 \\ 14 & 24 & -18 \\ -4 & -18 & 29 \end{pmatrix}$; 3) $\begin{pmatrix} 3 & 2 & 0 \\ 2 & 4 & -2 \\ 0 & -2 & 5 \end{pmatrix}$;

4) $\begin{pmatrix} 18 & 6 & -6 \\ 6 & 21 & 0 \\ -6 & 0 & 16 \end{pmatrix}$; 5) $\begin{pmatrix} 44 & -22 & 26 \\ -22 & 29 & -4 \\ 26 & -4 & 53 \end{pmatrix}$; 6) $\begin{pmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{pmatrix}$;

7) $\begin{pmatrix} 6 & 2 & -2 \\ 2 & 7 & 0 \\ -2 & 0 & 5 \end{pmatrix}$; 8) $\begin{pmatrix} 9 & -16 & -8 \\ -16 & 33 & 16 \\ -8 & 16 & 9 \end{pmatrix}$; 9) $\begin{pmatrix} 5 & 2 & -4 \\ 2 & 1 & -2 \\ -4 & -2 & 5 \end{pmatrix}$.

8.24. Применяя процесс ортогонализации, привести к треугольному виду матрицы из упражнений 8.22, 8.23 и следующие матрицы:

1) $\begin{pmatrix} 1 & -1 & 5 \\ -2 & 0 & -3 \\ 2 & -1 & -7 \end{pmatrix}$; 2) $\begin{pmatrix} 1 & 3 & -2 \\ 1 & 3 & 0 \\ 1 & -1 & 6 \end{pmatrix}$; 3) $\begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{pmatrix}$;

4) $\begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}$; 5) $\begin{pmatrix} 1 & i & -i \\ -i & 1 & i \\ i & -i & 1 \end{pmatrix}$; 6) $\begin{pmatrix} 2 & -i & i \\ i & 1 & -i \\ -i & i & 1 \end{pmatrix}$;

7) $\begin{pmatrix} 1 & -i & i \\ i & 2 & i \\ -i & -i & 3 \end{pmatrix}$; 8) $\begin{pmatrix} 3 & i & -i \\ -i & 2 & i \\ i & -i & 3 \end{pmatrix}$; 9) $\begin{pmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix}$.

8.25. Ортогональными (унитарными) преобразованиями привести матрицы из упражнений 8.22–8.24: 1) к треугольному виду; 2) к двухдиагональному виду.

8.26. Ортогональными (унитарными) преобразованиями привести к подобной матрице почти треугольного вида каждую из следующих матриц:

$$1)\ \frac{1}{9}\begin{pmatrix} 9 & -15 & -3 & -3 \\ -3 & 23 & 1 & 1 \\ -6 & 4 & -1 & -1 \\ -6 & 7 & 5 & 5 \end{pmatrix};\quad 2)\ \frac{1}{25}\begin{pmatrix} 25 & 25 & 0 & 0 \\ 25 & 25 & -15 & -20 \\ 0 & -15 & 49 & 7 \\ 0 & -20 & 7 & 1 \end{pmatrix};$$

$$3)\ \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 2 & -32 & 1 & 1 \\ 2 & -35 & 1 & 1 \end{pmatrix};\quad 4)\ \frac{1}{25}\begin{pmatrix} 25 & 25i & 0 & 0 \\ -25i & 25 & -15 & -20 \\ 0 & -15 & 49 & 7 \\ 0 & -20 & 7 & 1 \end{pmatrix};$$

$$5)\ \frac{1}{625}\begin{pmatrix} 625 & 0 & 375i & 500 \\ 0 & -600 & 225-140i & 300+105i \\ 375i & 225-140i & 384+480i & 288+140i \\ 500i & 300+105i & 288+140i & -216-480i \end{pmatrix}.$$

8.27. Найти канонический вид B ортогональной матрицы A и ортогональную матрицу T, такую, что $B = T^{-1}AT$, для следующих матриц A:

$$1)\ \begin{pmatrix} \frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \\ \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \\ -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \end{pmatrix};\quad 2)\ \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{\sqrt{2}}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \end{pmatrix};$$

$$3)\ \begin{pmatrix} \frac{3}{4} & \frac{1}{4} & -\frac{\sqrt{6}}{4} \\ \frac{1}{4} & \frac{3}{4} & \frac{\sqrt{6}}{4} \\ \frac{\sqrt{6}}{4} & -\frac{\sqrt{6}}{4} & \frac{1}{2} \end{pmatrix}.$$

8.28. Для следующих матриц построить каноническое разложение с унитарной трансформирующей матрицей:

$$1)\ \begin{pmatrix} 3 & -i & 0 \\ i & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix};\quad 2)\ \begin{pmatrix} 0 & i & 0 \\ -i & 0 & 0 \\ 0 & 0 & 4 \end{pmatrix};\quad 3)\ \begin{pmatrix} 3 & 2+2i & 0 \\ 2-2i & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

8.29. Убедиться в положительной определенности матрицы и найти квадратный корень из нее:

$$1)\ \begin{pmatrix} 3 & -i & 0 \\ i & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix}; \quad 2)\ \begin{pmatrix} 3 & 2-i & 0 \\ 2+i & 7 & 0 \\ 0 & 0 & 4 \end{pmatrix}; \quad 3)\ \begin{pmatrix} 3 & i & 0 \\ -i & 3 & 0 \\ 0 & 0 & 9 \end{pmatrix}$$

8.30. Убедиться в том, что матрица является нормальной и построить для нее каноническое разложение с унитарной трансформирующей матрицей:

$$1)\ \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}; \qquad 2)\ \begin{pmatrix} 0 & 2 & 1 \\ -2 & 0 & -2 \\ -1 & 2 & 0 \end{pmatrix}; \qquad 3)\ \begin{pmatrix} 2-i & -1 & 0 \\ -1 & 1-i & 1 \\ 0 & 1 & 2-i \end{pmatrix}.$$

8.31. Построить QR-разложения для матриц из упражнений 8.22 и 8.24.

8.32. Решить систему $Ax = b$ с матрицей A, заданной ее QR-разложением, и столбцом b свободных членов:

$$1)\ A = \frac{1}{3}\begin{pmatrix} 2 & 2 & 1 \\ 2 & -1 & -2 \\ 1 & -2 & 2 \end{pmatrix}\begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 6 \\ 6 \\ -6 \end{pmatrix};$$

$$2)\ A = \frac{1}{3}\begin{pmatrix} 1 & -2 & 2 \\ 2 & -1 & -2 \\ 2 & 2 & 1 \end{pmatrix}\begin{pmatrix} 3 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 18 \\ 18 \\ -18 \end{pmatrix};$$

$$3)\ A = \frac{1}{2}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}\begin{pmatrix} 2 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 4 \\ 8 \\ 12 \\ 4 \end{pmatrix};$$

$$4)\ A = \frac{1}{3}\begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 1 & 2 & 2 \\ 0 & 2 & 1 & -2 \\ 0 & 2 & -2 & 1 \end{pmatrix}\begin{pmatrix} 3 & 1 & 0 & 1 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 18 \\ 18 \\ 36 \\ 18 \end{pmatrix}.$$

8.33. Построить сингулярное разложение матриц:

$$1)\begin{pmatrix}7&1&2\\5&5&-2\\3&3&6\\1&7&2\end{pmatrix};\quad 2)\begin{pmatrix}10&4&-1\\10&4&-1\\2&8&7\\-2&-8&-7\end{pmatrix};\quad 3)\begin{pmatrix}8&5&1\\8&5&1\\4&7&5\\-4&-7&-5\end{pmatrix};$$

$$4)\begin{pmatrix}9&3&3\\7&7&-1\\5&5&7\\3&9&3\end{pmatrix};\quad 5)\begin{pmatrix}6&3&0\\6&3&0\\2&5&4\\-2&-5&-4\end{pmatrix};\quad 6)\begin{pmatrix}15&-3&3\\9&9&-9\\3&3&15\\-3&15&3\end{pmatrix};$$

$$7)\begin{pmatrix}4&1&-1\\4&1&-1\\0&3&3\\0&3&3\end{pmatrix};\quad 8)\begin{pmatrix}5&-1&1\\3&3&-3\\1&1&5\\-1&5&1\end{pmatrix};\quad 9)\begin{pmatrix}7&1&2\\5&5&-2\\3&3&6\\-1&-7&-2\end{pmatrix}.$$

Указание. Одно из ненулевых характеристических чисел матриц A^*A следует искать среди чисел 36, 144, 324.

8.34. Построить полярные разложения следующих матриц:

$$1)\begin{pmatrix}4&-2&2\\4&4&-1\\-2&4&2\end{pmatrix};\quad 2)\begin{pmatrix}2&-1&4&-2\\2&1&4&2\\4&2&2&-1\\-4&-2&2&1\end{pmatrix};\quad 3)\begin{pmatrix}1&1&1&1\\1&1&1&1\\1&1&1&1\\1&1&1&1\end{pmatrix}.$$

8.35. Построить скелетные разложения для матриц из упражнения 8.33.

8.36. Используя скелетные разложения, построить псевдообратные матрицы для следующих матриц:

$$1)\begin{pmatrix}0&1&2\\2&2&2\\5&3&1\\1&1&1\end{pmatrix};\quad 2)\begin{pmatrix}1&2&-3\\2&2&-4\\3&1&-4\\1&1&-2\end{pmatrix};\quad 3)\begin{pmatrix}1&2&3&1\\2&2&1&1\\-3&-4&-4&-2\end{pmatrix};$$

$$4)\begin{pmatrix}1&2&0\\1&1&-1\\2&2&-2\\3&1&-5\end{pmatrix};\quad 5)\begin{pmatrix}0&2&5&1\\1&2&3&1\\2&2&1&1\end{pmatrix};\quad 6)\begin{pmatrix}1&1&2&3\\2&1&1&2\\-3&-2&-3&-5\end{pmatrix};$$

7) $\begin{pmatrix} 1 & 2 & -3 \\ 1 & 1 & -2 \\ 2 & 1 & -3 \\ 3 & 2 & -5 \end{pmatrix}$; 8) $\begin{pmatrix} 1 & 2 & 3 & 1 \\ 1 & 1 & 2 & 0 \\ 1 & 3 & 4 & 1 \end{pmatrix}$; 9) $\begin{pmatrix} 1 & 1 & 2 & 3 \\ 2 & 1 & 2 & 1 \\ 0 & -1 & -2 & -5 \end{pmatrix}$.

8.37. Используя сингулярные разложения, построить псевдообратные матрицы для матриц из упражнения 8.33 и следующих матриц:

1) $\begin{pmatrix} 12 & 12 & 6 \\ 2 & -1 & -2 \\ 4 & -2 & -4 \\ 4 & -2 & -4 \end{pmatrix}$; 2) $\begin{pmatrix} 18 & 18 & 9 \\ 4 & -2 & -4 \\ 8 & -4 & -8 \\ 8 & -4 & -8 \end{pmatrix}$; 3) $\begin{pmatrix} 6 & 6 & 3 \\ 2 & -1 & -2 \\ 4 & -2 & -2 \\ 4 & -2 & -2 \end{pmatrix}$.

8.38. Методом наименьших квадратов решить системы линейных уравнений:

1) $\begin{cases} x_1 + x_2 + x_3 = 18, \\ x_1 + x_3 = 18, \\ 2x_1 + x_2 = 18, \\ 3x_1 + x_2 - x_3 = 36; \end{cases}$ 2) $\begin{cases} x_1 + x_2 + 2x_3 = 1, \\ 2x_1 - x_2 + x_3 = 0, \\ -x_1 + 2x_2 + x_3 = 0, \\ x_1 + 2x_2 + x_3 = 2; \end{cases}$

3) $\begin{cases} x_1 + x_2 + 2x_3 = 1, \\ x_1 + x_3 = -1, \\ -x_1 + x_2 = 1, \\ -x_1 - x_2 - 2x_3 = -1; \end{cases}$ 4) $\begin{cases} x_2 + x_3 = 1, \\ x_1 + x_3 = 0, \\ x_1 - x_2 = -1, \\ -x_1 - x_2 - 2x_3 = 1; \end{cases}$

5) $\begin{cases} x_1 - x_2 = 1, \\ x_1 + x_2 + 2x_3 = -1, \\ x_1 - x_2 + x_3 = 2, \\ 2x_1 + x_2 + x_3 = 3; \end{cases}$ 6) $\begin{cases} x_1 - x_2 = 66, \\ x_1 + x_3 = -33, \\ x_1 + x_2 + 2x_3 = -33, \\ x_1 + 2x_2 + 3x_3 = 66; \end{cases}$

7) $\begin{cases} x_1 + x_4 = 4, \\ x_2 + x_4 = 8, \\ x_3 + x_4 = 4; \end{cases}$ 8) $\begin{cases} x_1 + x_2 + x_3 + x_4 = 2, \\ x_1 + x_2 + x_3 + x_4 = 3, \\ x_1 + x_2 + x_3 + x_4 = 4. \end{cases}$

8.39. Для систем из предыдущего упражнения найти нормальные решения, решая соответствующие одну или две системы с невырожденными матрицами.

8.40. Для систем из упражнения 8.38 найти общие и нормальные псевдорешения, решая эти системы в матричном виде с применением псевдообратных матриц.

8.41. Найти общее и нормальное псевдорешения системы $Ax = b$ и их проекции $x^{(k)}$, $k = 1, \ldots, r$, на подпространства правых сингулярных векторов, если матрица A задана сингулярным разложением:

1) $$A = \frac{1}{2}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}\begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \\ 0 & 0 & 0 \end{pmatrix}\frac{1}{3}\begin{pmatrix} 2 & 1 & 1 \\ 2 & -1 & -2 \\ 1 & -2 & 2 \end{pmatrix},$$

$b = (216,\ 0,\ 0,\ -216)^{\mathrm{T}}$, где а) $\sigma_1 = 18$, $\sigma_2 = 12$, $\sigma_3 = 0$, б) $\sigma_1 = 18$, $\sigma_2 = 6$, $\sigma_3 = 0$, в) $\sigma_1 = \sigma_2 = 6$, $\sigma_3 = 0$;

2) $$A = \frac{1}{2}\begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 1 & 2 & 2 \\ 0 & 2 & 1 & -2 \\ 0 & 2 & -2 & 1 \end{pmatrix}\begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \\ 0 & 0 & 0 \end{pmatrix}\frac{1}{3}\begin{pmatrix} 1 & -2 & 2 \\ 2 & -1 & -2 \\ 2 & -2 & -1 \end{pmatrix},$$

$b = (162,\ 0,\ 0,\ -162)^{\mathrm{T}}$, где а) $\sigma_1 = 18$, $\sigma_2 = 9$, $\sigma_3 = 0$, б) $\sigma_1 = 9$, $\sigma_2 = \sigma_3 = 0$, в) $\sigma_1 = 18$, $\sigma_2 = \sigma_3 = 0$.

8.42. Методом регуляризации найти нормальные решения систем из упражнения 8.38 и следующих систем линейных уравнений:

1) $\begin{cases} x_1 + x_2 = 3, \\ x_1 - x_2 = 1, \\ x_1 + 2x_2 = 5; \end{cases}$ 2) $\begin{cases} x_1 + 2x_2 = 3, \\ 2x_1 - x_2 = 1, \\ x_1 + x_2 = 2; \end{cases}$

3) $\begin{cases} x_1 + x_2 = 3, \\ 2x_1 - x_2 = 3, \\ 2x_1 + x_2 = 5; \end{cases}$ 4) $\begin{cases} x_1 + x_3 = 0, \\ -x_1 + 8x_2 + 6x_3 = 1, \\ x_1 - 7x_2 - 5x_3 = -1; \end{cases}$

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

8.43. Пользуясь нормой $\|A\|_\infty$, найти число обусловленности K_A матриц:

$$1)\ \begin{pmatrix} 5 & -3{,}31 \\ 6 & -9{,}97 \end{pmatrix}; \qquad 2)\ \begin{pmatrix} 1 & -1 & 1 \\ 1 & \varepsilon & \varepsilon \\ 1 & \varepsilon & \varepsilon \end{pmatrix}, \quad \varepsilon \neq 0.$$

8.44. Оценить возможное изменение решений систем при изменении ε и ε_1 в пределах $0 \leqslant \varepsilon \leqslant 0{,}01$, $0 \leqslant \varepsilon_1 \leqslant 0{,}03$:

$$1)\ \begin{cases} x_1 - 2x_2 = -1 + \varepsilon, \\ 2x_1 + 4{,}01x_2 = 2 + \varepsilon; \end{cases} \qquad 2)\ \begin{cases} 5x_1 - 3{,}31x_2 = 1{,}69 + \varepsilon, \\ 6x_1 - 3{,}97x_2 = 2{,}07 + \varepsilon. \end{cases}$$

8.45. В эмпирической формуле $b = x_1a_1 + x_2a_2 + x_3a_3$ найти коэффициенты x_1, x_2, x_3 по результатам наблюдений:

1)

a_1	a_2	a_3	b
1	0	0	0
0	1	0	0
0	0	1	0
1	1	1	4

2)

a_1	a_2	a_3	b
1	1	−1	12
−1	1	1	12
1	−1	1	12
1	1	−1	12

3)

a_1	a_2	a_3	b
1	−1	0	12
2	−1	1	−12
0	1	1	12
1	−1	1	−24

4)

a_1	a_2	a_3	b
1	−1	0	9
−1	2	1	0
2	−3	−1	−9
0	1	1	0

8.46. В эмпирической формуле $p(t) = x_0 + x_1t + x_2t^2$ найти коэффициенты x_1, x_2, x_3 по результатам наблюдений:

1)

t	0	1	2	3
$p(t)$	1	4	9	16

2)

t	0	−1	−2	−3
$p(t)$	1	0	1	4

3)

t	−1	0	1	2
$p(t)$	5	1	−1	−1

4)

t	0	1	2	3
$p(t)$	1,1	3,9	9,1	16,1

5)

t	−1	0	1	2
$p(t)$	4	1	0	1

6)

t	0	−1	−2	−3
$p(t)$	1,1	−0,1	1,1	3,9

8.47. Найти многочлен второй степени $p(t) = x_0 + x_1 t + x_2 t^2$, приближающий с наименьшей квадратичной погрешностью функцию $f(t)$, заданную таблицей:

1)

t	0	1	2	3
$f(t)$	1	3	3	1

2)

t	−1	0	1	2
$f(t)$	4	2	2	4

8.48. Найти многочлен третьей степени $p(t) = x_0 + x_1 t + x_2 t^2 + x_3 t^3$, приближающий с наименьшей квадратичной погрешностью функцию $f(t)$, заданную таблицей:

1)

t	0	1	2
$f(t)$	1	1	−1

2)

t	−1	0	1
$f(t)$	4	2	−2

Глава 9 КВАДРАТИЧНЫЕ ФОРМЫ

9.1. Определение квадратичной формы

Действительной квадратичной формой от** n **действительных переменных x_1, x_2, ..., x_n называют многочлен с действительными коэффициентами a_{ij}, каждый член $a_{ij}x_ix_j$ которого имеет вторую степень, т.е. многочлен вида

$$f(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j. \tag{9.1}$$

Переменные x_1, x_2, ..., x_n можно рассматривать как координаты вектора $x = (x_1,\ x_2,\ \ldots,\ x_n)^{\mathrm{T}}$ в некотором фиксированном базисе действительного n-мерного пространства X. Поэтому квадратичную форму (9.1) можно представить как числовую функцию векторного аргумента $x \in X$ с действительными значениями.

Пусть $f(x_1, x_2, \ldots, x_n)$ — квадратичная форма (9.1). Коэффициент при x_i^2 в ней обозначен через a_{ii}, а коэффициент при произведении x_ix_j – через a_{ij}. Будем считать, что

$$a_{ij} = a_{ji}. \tag{9.2}$$

Этого всегда можно достичь, сложив члены $a_{ij}x_ix_j$ и $a_{ji}x_jx_i$ и разделив сумму на два. При такой договоренности квадратичную форму (9.1) можно записать следующим образом:

$$f(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j = a_{11}x_1^2 + a_{22}x_2^2 + \ldots + a_{nn}x_n^2 +$$

$$+ 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + \ldots + 2a_{1n}x_1x_n + \ldots + 2a_{n-1,n}x_{n-1}x_n.$$

Ее также можно записать в виде

$$\begin{aligned} f(x_1,x_2,\ldots,x_n) &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j = \\ &= \sum_{i=1}^{n}(a_{i1}x_ix_1 + a_{i2}x_ix_2 + \ldots + a_{in}x_ix_n) = \\ &= a_{11}x_1^2 \quad + a_{12}x_1x_2 + \ldots + a_{1n}x_1x_n + \\ &+ a_{21}x_2x_1 + a_{22}x_2^2 \quad + \ldots + a_{2n}x_2x_n + \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &+ a_{n1}x_nx_1 + a_{n2}x_nx_2 + \ldots + a_{nn}x_n^2. \end{aligned}$$

Такой вид записи квадратичной формы называют ее ***симметрическим видом.***

Матрицу

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

составленную из коэффициентов квадратичной формы, называют ***матрицей этой квадратичной формы***, а ранг $r(A)$ и определитель $|A|$ этой матрицы — соответственно ***рангом*** и ***дискриминантом квадратичной формы***.

В силу условия (9.2) матрица квадратичной формы симметрическая. Очевидно и обратное, а именно, любая действительная симметрическая матрица является матрицей некоторой квадратичной формы.

Если через x обозначить столбец переменных x_1, x_2, ..., x_n, то квадратичную форму $f(x_1,x_2,\ldots,x_n)$ в матричном виде можно записать следующим образом:

$$f(x) = x^{\mathrm{T}}Ax.$$

9.2. Линейное преобразование переменных

Пусть $x = (x_1, x_2, \ldots, x_n)^{\mathrm{T}}$ — координаты некоторого вектора в линейном пространстве X, записанные в базисе e, а $y = (y_1, y_2, \ldots, y_n)^{\mathrm{T}}$ — координаты того же вектора в другом базисе f. Тогда (см. разд. 4.5 и 4.6) столбцы x и y связаны соотношением

$$x = Qy, \tag{9.3}$$

где $Q = (q_{ij})$ — матрица перехода от первого базиса e ко второму базису f.

В подробной записи соотношение (9.3) принимает вид:

$$\begin{cases} x_1 = q_{11}y_1 + q_{12}y_2 + \ldots + q_{1n}y_n, \\ x_2 = q_{21}y_1 + q_{22}y_2 + \ldots + q_{2n}y_n, \\ \dots\dots\dots\dots\dots\dots \\ x_n = q_{n1}y_1 + q_{n2}y_2 + \ldots + q_{nn}y_n. \end{cases} \tag{9.4}$$

Эти формулы дают преобразование координат вектора при переходе от второго базиса к первому. В теории квадратичных форм их называют ***линейным преобразованием переменных*** $y_1, y_2, \ldots, y_n$ в переменные $x_1, x_2, \ldots, x_n$. Напомним, что матрица Q составляется из столбцов координат соответствующих векторов второго базиса f в первом базисе e, а коэффициентами в уравнениях (9.4) являются элементы из соответствующих строк матрицы Q.

Если рассматриваемый вектор в некотором третьем базисе g имеет координаты $z = (z_1, z_2, \ldots, z_n)^{\mathrm{T}}$ и если R — матрица перехода от второго базиса f к третьему базису g, то QR будет матрицей перехода от базиса e к базису g. При этом будут выполняться соотношения

$$y = Rz, \tag{9.5}$$

$$x = QRz. \tag{9.6}$$

Формула (9.3) выражает переменные $x_1, x_2, \ldots, x_n$ через переменные $y_1, y_2, \ldots, y_n$, формула (9.5) — переменные $y_1, y_2, \ldots, y_n$ через переменные $z_1, z_2, \ldots, z_n$, а формула (9.6) — переменные $x_1, x_2, \ldots, x_n$ через переменные $z_1, z_2, \ldots, z_n$. Матрицей результирующего преобразования переменных является матрица QR, т.е. произведение матриц последовательных преобразований (9.3) и (9.5).

Линейное преобразование переменных $x = Qy$, *являясь преобразованием координат векторов при изменении базиса, имеет невырожденную матрицу* Q. Такое преобразование переменных называют ***невырожденным***.

9.3. Преобразование квадратичной формы при линейном преобразовании переменных

Пусть дана квадратичная форма

$$f(x_1, x_2, \ldots, x_n) = x^{\mathrm{T}} A x \tag{9.7}$$

с матрицей A. Применим к ее переменным $x_1, x_2, \ldots, x_n$ невырожденное линейное преобразование переменных

$$\begin{cases} x_1 = q_{11}y_1 + q_{12}y_2 + \ldots + q_{1n}y_n, \\ x_2 = q_{21}y_1 + q_{22}y_2 + \ldots + q_{2n}y_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n = q_{n1}y_1 + q_{n2}y_2 + \ldots + q_{nn}y_n. \end{cases}, \tag{9.8}$$

или, что то же самое,

$$x = Qy. \tag{9.9}$$

Подставив в форму (9.7) вместо каждой переменной $x_1, x_2, \ldots, x_n$ выражение через переменные $y_1, y_2, \ldots, y_n$ по формулам (9.8), получим квадратичную форму

$$g(y_1, y_2, \ldots, y_n) = (Qy)^{\mathrm{T}} A (Qy) = y^{\mathrm{T}} Q^{\mathrm{T}} A Q y. \tag{9.10}$$

Вид квадратичной формы $g(y_1, y_2, \ldots, y_n)$ означает, что ее матрицей является матрица

$$B = Q^{\mathrm{T}} A Q. \tag{9.11}$$

Таким образом, при линейном преобразовании переменных с матрицей Q матрица A квадратичной формы f преобразуется в матрицу $Q^{\mathrm{T}}AQ$. При этом поскольку Q — невырожденная матрица, то из соотношения (9.11) следует, что ранг матрицы B равен рангу матрицы A, т.е. ***ранг матрицы квадратичной формы не меняется при невырожденном линейном преобразовании ее переменных***. Поэтому ранг матрицы квадратичной формы будем называть ***рангом квадратичной формы***. Отметим, что определитель матрицы квадратичной формы при линейной замене переменных изменяется, так как из равенства (9.11) вытекает, что

$$|B| = |A|\,|Q|^2. \tag{9.12}$$

Напомним, что переход к новым переменным мы рассматриваем как переход к новому базису в линейном пространстве. Поэтому квадратичную форму (9.10) в новых переменных y_1, y_2, ..., y_n следует рассматривать как квадратичную форму (9.7) в новом базисе.

Замечание. Не следует смешивать правило $Q^{\mathrm{T}}AQ$ изменения матрицы A квадратичной формы при переходе от одного базиса к другому с правилом $Q^{-1}AQ$ (см. разд. 5.4) изменения матрицы A линейного оператора при переходе от одного базиса к другому. Матрица $Q^{\mathrm{T}}AQ$ в общем случае не является матрицей того же оператора. Совпадение матриц $Q^{\mathrm{T}}AQ$ и $Q^{-1}AQ$ будет лишь в случае ортогональной матрицы Q, т.е. при $Q^{\mathrm{T}} = Q^{-1}$. Это происходит, например, при переходе от ортонормированного базиса к другому ортонормированному базису.

Пусть квадратичная форма $f(x_1, x_2, \ldots, x_n)$ некоторым невырожденным линейным преобразованием переменных приведена к виду

$$g(y_1, y_2, \ldots, y_n) = c_1 y_1^2 + c_2 y_2^2 + \ldots + c_n y_n^2, \tag{9.13}$$

содержащему только квадраты переменных. Такой вид квадратичной формы f называют ***каноническим видом***, а базис, в котором квадратичная форма имеет канонический вид, — ***каноническим базисом*** квадратичной формы. Матрица квадратичной формы (9.13) является диагональной матрицей

$$\begin{pmatrix} c_1 & & 0 \\ & \ddots & \\ 0 & & c_n \end{pmatrix}. \tag{9.14}$$

Приведение квадратичной формы к каноническому виду может осуществляться разными способами. Иными словами, квадратичная форма имеет различные канонические базисы, в которых вид квадратичной формы будет различаться. Однако поскольку при невырожденном линейном преобразовании переменных ранг матрицы квадратичной формы не меняется, то число ненулевых слагаемых в любом каноническом виде квадратичной формы одно и то же и равно рангу квадратичной формы.

9.4. Приведение квадратичной формы к каноническому виду

Если квадратичную форму подвергнуть линейному преобразованию, то получится квадратичная форма от новых переменных с другими коэффициентами. Естественно задаться целью с помощью надлежащего линейного преобразования переменных привести квадратичную форму к более простому виду. Но что понимать под более простым видом? Оказывается, что верно следующее утверждение.

Теорема 9.1. *Любую квадратичную форму с помощью невырожденного линейного преобразования переменных можно привести к каноническому виду.*

Привести квадратичную форму к каноническому виду можно ***методом Лагранжа***. Основная идея этого метода состоит в последовательном выделении в квадратичной форме полных квадратов по каждой переменной. Для выделения полного квадрата по переменной x_i необходимо, чтобы в квадратичной форме присутствовало слагаемое с квадратом этой переменной. Если в квадратичной форме нет членов с квадратами переменных, то выполняют специальное невырожденное линейное преобразование переменных так, чтобы в квадратичной форме появились члены с квадратами переменных. Так, если все $a_{ii} = 0$, но $a_{ij} \neq 0$ для некоторых номеров i и j, то, выполнив невырожденное преобразование переменных

$$x_i = y_i + y_j, \quad x_j = y_i - y_j, \quad x_k = y_k \quad \text{при} \quad k \neq i, j,$$

получим, что член $2a_{ij}x_ix_j$ квадратичной формы примет вид:

$$2a_{ij}(y_i + y_j)(y_i - y_j) = 2a_{ij}y_i^2 - 2a_{ij}y_j^2,$$

т.е. в квадратичной форме появятся члены с квадратами переменных y_i и y_j, причем они не могут сократиться с другими членами формы, так как в каждый другой ее член входит y_k при $k \neq i, j$. Таким образом, можно считать, что в квадратичной форме

$$f(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j$$

есть члены с квадратами переменных.

Пусть, например, в квадратичной форме f есть член с квадратом переменной x_1, т.е. $a_{11} \neq 0$. Соберем в f все члены, содержащие x_1, и дополним их сумму до полного квадрата. Тогда получим

$$\begin{aligned} f(x_1, x_2, \ldots, x_n) = \left(a_{11}x_1^2 + 2a_{12}x_1x_2 + \ldots + 2a_{1n}x_1x_n\right) + \\ + a_{22}x_2^2 + 2a_{23}x_2x_3 + \ldots + a_{nn}x_n^2 = \\ = \frac{1}{a_{11}}\left(a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n\right)^2 + g(x_2, x_3, \ldots, x_n), \end{aligned}$$

где $g(x_2, x_3, \ldots, x_n)$ — квадратичная форма уже только от переменных x_2, x_3, …, x_n.

Введем новые переменные

$$\begin{aligned} y_1 &= a_{11}x_1 + \ldots + a_{1n}x_n, \\ y_2 &= x_2, \\ &\cdots\cdots \\ y_n &= x_n. \end{aligned}$$

В новых переменных квадратичная форма f примет вид:

$$f(y_1, y_2, \ldots, y_n) = \frac{1}{a_{11}} y_1^2 + g(y_2, y_3, \ldots, y_n).$$

С квадратичной формой g можно поступать аналогично. Не более чем через $n-1$ шагов придем к каноническому виду квадратичной формы f. Результирующее преобразование переменных будет равно произведению последовательно выполненных преобразований переменных. Пусть Q — его матрица, A — матрица квадратичной формы, C — диагональная матрица полученного канонического вида. Тогда формула (9.11) принимает вид: $C = Q^{\mathrm{T}} A Q$.

Пример 9.1. Методом Лагранжа привести к каноническому виду квадратичную форму

$$f(x_1, x_2, x_n) = 2x_1x_2 + 2x_1x_3 - 6x_2x_3$$

и указать невырожденное преобразование переменных, осуществляющее такое преобразование квадратичной формы.

Решение. В данной квадратичной форме отсутствуют члены с квадратами переменных, но есть, например, член $2x_1x_2$. Поэтому совершим сначала невырожденное преобразование переменных

$$x_1 = y_1 + y_2, \quad x_2 = y_1 - y_2, \quad x_3 = y_3$$

с матрицей

$$Q_1 = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

В результате получим:

$$f = 2y_1^2 - 2y_2^2 - 4y_1y_3 + 8y_2y_3.$$

Здесь коэффициент при y_1^2 отличен от нуля. Поэтому можно выделить полный квадрат по y_1:

$$f = (2y_1^2 - 4y_1y_3) - 2y_2^2 + 8y_2y_3 = \frac{1}{2}(2y_1 - 2y_3)^2 - 2y_2^2 + 8y_2y_3 - 2y_3^2.$$

Введем новые переменные

$$z_1 = 2y_1 - 2y_3, \quad z_2 = y_2, \quad z_3 = y_3,$$

что равносильно линейному преобразованию переменных

$$y_1 = \frac{1}{2}z_1 + z_3, \quad y_1 = z_2, \quad y_3 = z_3$$

с матрицей

$$Q_2 = \begin{pmatrix} 1/2 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

В новых переменных квадратичная форма f принимает вид:

$$f = \frac{1}{2}z_1^2 - 2z_2^2 + 8z_2z_3 - 2z_3^2.$$

В квадратичной форме

$$g = -2z_2^2 + 8z_2z_3 - 2z_3^2$$

выделим полный квадрат по z_2:

$$f = \frac{1}{2}z_1^2 + \frac{1}{-2}(-2z_2 + 4z_3)^2 + 6z_3^2.$$

Введем новые переменные

$$t_1 = z_1, \quad t_2 = -2z_2 + 4z_3, \quad t_3 = z_3,$$

что равносильно линейному преобразованию переменных

$$z_1 = t_1, \quad z_2 = -\frac{1}{2}t_1 + 2t_2, \quad z_3 = t_3$$

с матрицей

$$Q_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1/2 & 2 \\ 0 & 0 & 1 \end{pmatrix}.$$

В новых переменных квадратичная форма f уже имеет канонический вид:

$$f = \frac{1}{2}t_1^2 - \frac{1}{2}t_2^2 + 6t_3^2.$$

Линейное преобразование переменных, сразу приводящее квадратичную форму f к полученному каноническому виду, имеет матрицу

$$Q = Q_1Q_2Q_3 = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1/2 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1/2 & 2 \\ 0 & 0 & 1 \end{pmatrix} =$$

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

т.е. определяется формулами

$$\begin{cases} x_1 = \frac{1}{2}t_1 - \frac{1}{2}t_2 + 3t_3, \\ x_2 = \frac{1}{2}t_1 + \frac{1}{2}t_2 - t_3, \\ x_3 = t_3. \end{cases}$$

Столбцы матрицы Q этого преобразования переменных, т.е. векторы

$$e_1' = \left(\frac{1}{2},\ \frac{1}{2},\ 0\right)^{\mathrm{T}}, \quad e_2' = \left(-\frac{1}{2},\ \frac{1}{2},\ 0\right)^{\mathrm{T}}, \quad e_3' = (3,\ -1,\ 1)^{\mathrm{T}},$$

составляют канонический базис квадратичной формы f.

Пусть квадратичная форма $f(x_1,x_2,\ldots,x_n)$ приведена к каноническому виду

$$f(y_1,y_2,\ldots,y_n)=c_1y_1^2+c_2y_2^2+c_ky_k^2-c_{k+1}y_{k+1}^2-\ldots c_ry_r^2,$$

$$c_i>0,\quad i=1,2,\ldots,r,$$

в котором положительные слагаемые предшествуют отрицательным (это условие может быть выполнено путем выбора соответствующего порядка переменных). Выполним дополнительное линейное преобразование переменных

$$z_1=\sqrt{c_1}y_1,\quad \ldots,\quad z_r=\sqrt{c_r}y_r,\quad z_{r+1}=y_{r+1},\quad \ldots,\quad z_n=y_n.$$

В результате квадратичная форма преобразуется к виду

$$f(z_1,z_2,\ldots,z_n)=z_1^2+z_2^2+\ldots+z_k^2-z_{k+1}^2-\ldots-z_n^2.$$

Такой вид квадратичной формы называют ее ***нормальным видом***.

В предыдущем примере квадратичная форма

$$f(x_1,x_2,x_3)=2x_1x_2+2x_1x_3-6x_2x_3$$

была приведена к каноническому виду

$$f(t_1,t_2,t_3)=\frac{1}{2}t_1^2-\frac{1}{2}t_2^2+6t_3^2.$$

Выполнив еще одно преобразование переменных

$$u_1=\frac{1}{\sqrt{2}}t_1,\quad u_2=\frac{1}{\sqrt{2}}t_2,\quad u_3=\sqrt{6}t_3,$$

получим нормальный вид квадратичной формы f:

$$f(u_1,u_2,u_3)=u_1^2-u_2^2+u_3^2.$$

9.5. Закон инерции квадратичных форм

Приводя разными способами квадратичную форму к каноническому виду, будем получать разные канонические виды. Квадраты переменных, входящие в канонический вид с положительными коэффициентами, для краткости будем называть ***положительными квадратами***, а квадраты, входящие в канонический вид с отрицательными коэффициентами, — ***отрицательными квадратами***.

Для квадратичных форм имеет место следующий ***закон инерции***.

Теорема 9.2. *Число положительных квадратов, как и число отрицательных квадратов, в любом каноническом виде данной квадратичной формы одно и то же и не зависит от того, каким невырожденным линейным преобразованием переменных получен канонический вид.*

▷ Пусть квадратичная форма f имеет канонический вид

$$f(x_1,x_2,\ldots,x_n)=c_1x_1^2+c_2x_2^2+\ldots+c_px_p^2-c_{p+1}x_{p+1}^2-\ldots-c_rx_r^2$$

в базисе e_1, e_2, ..., e_n (здесь $c_i>0$, $i=1,2,\ldots,r$) и канонический вид

$$f(y_1,y_2,\ldots,y_n)=d_1x_1^2+d_2x_2^2+\ldots+d_kx_k^2-d_{k+1}x_{k+1}^2-\ldots-d_rx_r^2$$

в базисе e'_1, e'_2, ..., e'_n (здесь $d_i>0$, $i=1,2,\ldots,r$). Общее количество квадратов в двух канонических видах одинаково и совпадает с рангом квадратичной формы.

Допустим, что число положительных квадратов в первом каноническом виде больше, чем во втором, т.е. $p>k$. Рассмотрим в линейном пространстве подпространства $L_1=$ $=\langle e_1,\,e_2,\,\ldots,\,e_p\rangle$ и $L_2=\langle e'_{k+1},\,e'_{k+2},\,\ldots,\,e'_n\rangle$. Так как сумма размерностей $p+(n-k)=n+(p-k)$ этих подпространств превышает размерность n рассматриваемого линейного пространства, то они имеют ненулевое пересечение, т.е. $L_1 \cap L_2 \neq$ $\neq \{0\}$. Следовательно, существует ненулевой вектор x, одновременно принадлежащий и L_1, и L_2. Из условия $x \in L_1$ вытекает представление

$$x = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_p e_p,$$

а из условия $x \in L_2$ — представление

$$x = \beta_{k+1} e'_{k+1} + \beta_{k+2} e'_{k+2} + \ldots + \beta_n e'_n.$$

Из первого канонического вида квадратичной формы вытекает, что

$$f(x) = c_1\alpha_1^2 + c_2\alpha_2^2 + \ldots + c_p\alpha_p^2 > 0,$$

а из второго — что

$$f(x) = -d_{k+1}y_{k+1}^2 - d_{k+2}y_{k+2}^2 - \ldots - d_n y_n^2 < 0.$$

Полученное противоречие показывает, что предположение $p > k$ неверно. Но тогда по тем же причинам неверно и предположение $k > p$. Следовательно, $p = k$ и количество квадратов одного знака в двух канонических видах данной квадратичной формы совпадает. ▶

Число положительных квадратов в каноническом виде квадратичной формы называют ее ***положительным индексом инерции*** и обозначают через i_+; число отрицательных квадратов в нормальном виде квадратичной формы называют ее ***отрицательным индексом инерции*** и обозначают через i_-. Очевидно, что сумма положительного и отрицательного индексов инерции совпадает с рангом квадратичной формы. Разность положительного и отрицательного индексов инерции $i_+ - i_-$ называют ***сигнатурой квадратичной формы*** и обозначают через s. По известным индексам инерции i_+ и i_- определяются и ранг квадратичной формы $r = i_+ + i_-$, и ее сигнатура $s = i_+ - i_-$. Наоборот, если известны ранг r и сигнатура s квадратичной формы, то из соотношений $r = i_+ + i_-$, $s = i_+ - i_-$ определяются индексы инерции i_+ и i_-.

9.6. Знакоопределенные квадратичные формы

Квадратичную форму $f(x_1, x_2, \ldots, x_n)$ называют ***положительно*** (***отрицательно***) ***определенной*** или ***положительно*** (***отрицательно***) ***знакопостоянной***, если она на любом ненулевом векторе принимает положительное (отрицательное) значение, т.е. если

$$x^{\mathrm{T}}Ax > 0 \quad (x^{\mathrm{T}}Ax < 0) \quad \text{при} \quad x \neq 0.$$

Если квадратичная форма на любом ненулевом векторе принимает неотрицательное (неположительное) значение, т.е.

$$x^{\mathrm{T}}Ax \geqslant 0 \quad (x^{\mathrm{T}}Ax \leqslant 0) \quad \text{при} \quad x \neq 0,$$

то ее называют ***положительно*** (***отрицательно***) ***полуопределенной*** или ***неотрицательной*** (***неположительной***).

Квадратичную форму, принимающую как положительные, так и отрицательные значения, называют ***неопределенной*** или ***знакопеременной***.

Если квадратичная форма $f(x_1, x_2, \ldots, x_n)$ положительно определенная, то квадратичная форма $-f(x_1, x_2, \ldots, x_n)$ отрицательно определенная.

У положительно определенной формы все коэффициенты при квадратах переменных, определитель ее матрицы и все характеристические числа матрицы положительны.

Приведем два признака положительной (отрицательной) определенности квадратичных форм.

1. ***Для того чтобы квадратичная форма от n переменных была положительно (отрицательно) определенной, необходимо и достаточно, чтобы она приводилась к каноническому виду с n положительными (отрицательными) квадратами.***

2. ***Критерий Сильвестра. Для того чтобы квадратичная форма от n переменных была положитель-***

но определенной, необходимо и достаточно, чтобы все угловые миноры ее матрицы A были положительными, т.е. чтобы

$$\Delta_1 = a_{11} > 0, \quad \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0,$$

$$\Delta_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} > 0, \quad \ldots, \quad \Delta_n = |A| > 0.$$

Для того чтобы квадратичная форма $f(x_1, x_2, \ldots, x_n)$ была отрицательно определенной, необходимо и достаточно, чтобы знаки угловых миноров матрицы A квадратичной формы чередовались, начиная со знака минус.

Доказательство первого признака. Пусть квадратичная форма f в каком-либо каноническом базисе e_1, e_2, ..., e_n имеет вид:

$$f(y_1, y_2, \ldots, y_n) = c_1 y_1^2 + c_2 y_2^2 + \ldots + c_n y_n^2.$$

Если квадратичная форма положительно определенная, то для любого вектора $x \neq 0$ имеем $f(x) > 0$. В частности, $f(e_i) = c_i > 0$, $i = 1, 2, \ldots, n$.

Если в указанном каноническом виде все коэффициенты c_i положительные, то для любого ненулевого вектора x с координатами $y = (y_1, y_2, \ldots, y_n)$ получаем

$$f(x) = c_1 y_1^2 + c_2 y_2^2 + \ldots + c_n y_n^2 > 0,$$

так как все слагаемые неотрицательны, но хотя бы одно слагаемое является ненулевым.

Доказательство критерия Сильвестра. Критерий Сильвестра вытекает из свойств квадратичной формы, рассматриваемой в различных подпространствах данного линейного пространства. Сначала остановимся на случае положительно определенной квадратичной формы.

Замечание. Определитель матрицы квадратичной формы в любом базисе имеет один и тот же знак. Иначе говоря, при невырожденном линейном преобразовании матрица квадратичной формы сохраняет знак своего определителя. Этот вывод можно сделать из равенства (9.12), в котором определитель $|Q|$ матрицы невырожденного линейного преобразования переменных отличен от нуля. В частности, у положительно определенной квадратичной формы в любом базисе определитель матрицы положителен, так как в соответствии с первым признаком это выполняется в каноническом базисе.

Предположим, что квадратичная форма f положительно определенная и пусть в некотором базисе e_1, e_2, ..., e_n она имеет следующий вид:

$$f(x) = x^{\mathrm{T}}Ax = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j. \tag{9.15}$$

Рассмотрим квадратичную форму f как квадратичную форму f_k на подпространстве $L_k = \langle e_1, \ldots, e_k \rangle$. Тогда

$$f_k(x) = \sum_{i=1}^{k}\sum_{j=1}^{k} a_{ij}x_ix_j,$$

матрица квадратичной формы f_k состоит из элементов матрицы A квадратичной формы F, расположенных в первых k строках и первых k столбцах, а определитель матрицы квадратичной формы f_k представляет собой k-й угловой минор матрицы A. Квадратичная форма f_k положительно определенная, так как в подпространстве L_k принимает те же значения, что и квадратичная форма f. Поэтому определитель квадратичной формы f_k, т.е. k-й угловой минор, положителен.

Обратное утверждение будем доказывать методом математической индукции по размерности линейного пространства. Для квадратичных форм в одномерном линейном пространстве утверждение критерия очевидно, так как квадратичная форма от одной переменной имеет вид $f(x_1) = a_{11}x_1^2$, а ее положительная определенность равносильна условию $a_{11} > 0$, т.е. положительности единственного углового минора.

Пусть обратная часть критерия Сильвестра верна для всех квадратичных форм в $(n-1)$-мерных линейных пространствах. Рассмотрим квадратичную форму f в n-мерном пространстве, и пусть в некотором базисе e_1, e_2, ..., e_n эта квадратичная форма имеет вид (9.15). Запишем ее в виде

$$f(x_1,x_2,\ldots,x_n)=\varphi(x_1,x_2,\ldots,x_{n-1})+ \\ +2\sum_{i=1}^{n-1}a_{in}x_ix_n+a_{nn}x_n^2, \quad (9.16)$$

где $\varphi(x_1,x_2,\ldots,x_{n-1})$ — квадратичная форма от $n-1$ переменных, составленная из всех слагаемых, не содержащих x_n. Квадратичная форма φ представляет собой сужение квадратичной формы f на подпространство $L_{n-1}=\langle e_1,\ \ldots,\ e_{n-1}\rangle$. Ее матрица получается из матрицы A квадратичной формы f вычеркиванием последней строки и последнего столбца. Следовательно, все угловые миноры матрицы квадратичной формы φ в то же время являются угловыми минорами матрицы A. Значит, они положительны. В соответствии с предположением математической индукции заключаем, что квадратичная форма φ положительно определена. Выберем в подпространстве L_{n-1} базис e'_1, e'_2, ..., e'_{n-1}, в котором квадратичная форма имеет нормальный вид. По первому признаку положительной определенности все квадраты в нормальном виде квадратичной формы положительны, т.е. в выбранном базисе

$$\varphi(y_1,y_2,\ldots,y_{n-1})=y_1^2+y_2^2+\ldots+y_{n-1}^2.$$

Добавив к базису e'_1, e'_2, ..., e'_{n-1} подпространства L_{n-1} вектор $e'_n=e_n$, получим базис всего линейного пространства, в котором квадратичная форма f имеет вид

$$f(y_1,y_2,\ldots,y_n)=y_1^2+y_2^2+\ldots+y_{n-1}^2+ \\ +2\sum_{i=1}^{n-1}b_{in}y_iy_n+b_{nn}y_n^2. \quad (9.17)$$

Выделим в этом выражении полные квадраты по переменным $y_1, y_2, \ldots, y_n$:

$$f(y_1, y_2, \ldots, y_n) = \sum_{i=1}^{n-1} \left[(y_i + b_{in} y_n)^2 - b_{in}^2 y_n^2 \right] + b_{nn} y_n^2 =$$

$$= \sum_{i=1}^{n-1} (y_i + b_{in} y_n)^2 + c_{nn} y_n^2,$$

где $c_{nn} = b_{nn} - b_{1n}^2 - \ldots - b_{n-1,n}^2$. Таким образом, в результате невырожденной замены переменных

$$\begin{cases} z_1 = y_1 + b_{1n} y_n, \\ z_2 = y_2 + b_{2n} y_n, \\ \ldots\ldots\ldots\ldots \\ z_{n-1} = y_{n-1} + b_{n-1,n} y_n, \\ z_n = y_n \end{cases}$$

квадратичная форма f приводится к каноническому виду

$$f(z_1, z_2, \ldots, z_n) = z_1^2 + z_2^2 + \ldots + z_{n-1}^2 + c_{nn} z_n^2. \tag{9.18}$$

Определитель матрицы квадратичной формы в каноническом виде равен произведению коэффициентов при квадратах, в данном случае c_{nn}. Но определитель матрицы квадратичной формы при переходе к новому базису сохраняет знак. Поэтому c_{nn} имеет тот же знак, что и определитель $\Delta_n = |A|$ матрицы квадратичной формы в исходном базисе, т.е. $c_{nn} > 0$. По первому признаку квадратичная форма с каноническим видом (9.18) является положительно определенной.

Итак, мы показали, что, во-первых, обратное утверждение критерия Сильвестра верно для одномерных пространств, во-вторых, если оно верно для $(n-1)$-мерных пространств, то оно верно и для n-мерных пространств. В соответствии с методом математической индукции заключаем, что критерий Сильвестра верен для любых конечномерных пространств.

Перейдем к случаю отрицательно определенной квадратичной формы. Если квадратичная форма f отрицательно определенная, т.е. принимает только отрицательные значения, то квадратичная форма $-f$ — положительно определенная, и наоборот. Поэтому критерием неотрицательной определенности квадратичной формы f с матрицей A является положительность всех угловых миноров матрицы

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

квадратичной формы $-f$, т.е.

$$-a_{11} > 0, \quad \begin{vmatrix} -a_{11} & -a_{12} \\ -a_{21} & -a_{22} \end{vmatrix} > 0, \quad \begin{vmatrix} -a_{11} & -a_{12} & -a_{13} \\ -a_{21} & -a_{22} & -a_{23} \\ -a_{31} & -a_{32} & -a_{33} \end{vmatrix} > 0, \quad \dots$$

Но это означает, что

$$a_{11} < 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} < 0, \quad \dots$$

или, другими словами, чередование знаков угловых миноров.

Пример 9.2. Квадратичная форма

$$f(x_1, x_2, \dots, x_n) = 2x_1^2 + x_2^2 + 11x_3^2 - 2x_1x_2 + 4x_1x_3 - 6x_2x_3$$

с матрицей

$$\begin{pmatrix} 2 & -1 & 2 \\ -1 & 1 & -3 \\ 2 & -3 & 11 \end{pmatrix}$$

положительно определенная, так как все угловые миноры матрицы A

$$\Delta_1 = 2, \quad \Delta_2 = \begin{vmatrix} 2 & -1 \\ -1 & 1 \end{vmatrix} = 1, \quad \Delta_3 = \begin{vmatrix} 2 & -1 & 2 \\ -1 & 1 & -3 \\ 2 & -3 & 11 \end{vmatrix} = 1$$

положительные.

На практике при выяснении вопроса о положительной (отрицательной) определенности квадратичной формы целесообразно пользоваться критерием Сильвестра, особенно в случае, когда квадратичная форма зависит от небольшого числа переменных. При большом числе переменных часто более предпочтительным является первый признак, связанный с приведением квадратичной формы к каноническому виду. Это объясняется тем, что приведение квадратичной формы к каноническому виду требует примерно столько же операций, сколько нужно для вычисления лишь одного определителя матрицы квадратичной формы. Кроме того, по каноническому виду квадратичной формы можно судить о ее неопределенности (если есть положительные и отрицательные квадраты) и полуопределенности (если все члены канонического вида либо положительные, либо отрицательные, но их меньше, чем число переменных в квадратичной форме), чего нельзя сделать по критерию Сильвестра. Например, у квадратичной формы

$$f(x_1,x_2,x_3) = 2x_1x_2 + 2x_1x_3 - 6x_2x_3$$

матрица

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

имеет угловые миноры

$$\Delta_1 = 0, \quad \Delta_2 = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = -1, \quad \Delta_3 = \begin{vmatrix} 0 & 1 & 1 \\ 1 & 0 & -3 \\ 1 & -3 & 0 \end{vmatrix} = -6.$$

По критерию Сильвестра можем лишь заключить, что данная квадратичная форма не является положительно (отрицательно) определенной. А по каноническому виду рассматриваемой квадратичной формы $f(t_1,t_2,t_3) = \frac{1}{2}t_1^2 - \frac{1}{2}t_2^2 + 6t_3^2$ (см. пример 9.1) нетрудно сделать вывод, что она знакопеременная, так как в ее каноническом виде есть положительные и отрицательные квадраты.

9.7. Распадающиеся квадратичные формы

Квадратичную форму называют ***распадающейся***, если она может быть представлена в виде произведения двух линейных множителей. Например, квадратичная форма $f(x_1,x_2) = x_1^2 - x_2^2$ распадающаяся, так как она представима в виде

$$f(x_1,x_2) = (x_1+x_2)(x_1-x_2).$$

Теорема 9.3. Действительная квадратичная форма от n переменных распадается тогда и только тогда, когда либо ее ранг равен единице, либо ранг равен двум, а сигнатура равна нулю.

$\triangleright$ Пусть квадратичная форма f распадающаяся, т.е. представима в виде

$$f(x_1,x_2,\ldots,x_n) = (a_1x_1+a_2x_2+\ldots+a_nx_n)\times \\ \times(b_1x_1+b_2x_2+\ldots+b_nx_n). \quad (9.19)$$

Если в линейных множителях этого разложения все коэффициенты пропорциональны, то

$$b_1x_1+b_2x_2+\ldots+b_nx_n = k(a_1x_1+a_2x_2+\ldots+a_nx_n)$$

и квадратичная форма f оказывается представленной в виде

$$f(x_1,x_2,\ldots,x_n) = k(a_1x_1+a_2x_2+\ldots+a_nx_n)^2.$$

Предположим для определенности, что $a_1 \neq 0$. Тогда, полагая

$$y_1 = a_1x_1+a_2x_2+\ldots+a_nx_n,$$
$$y_i = x_i, \quad i = 2, 3, \ldots, n,$$

получим невырожденное линейное преобразование переменных, в результате которого квадратичная форма приводится к виду $f(y_1,y_2,\ldots,y_n) = ky_1^2$. Это означает, что ранг квадратичной формы f равен единице.

Если в линейных множителях разложения (9.19) коэффициенты не являются пропорциональными, например $\frac{a_1}{b_1} \neq \frac{a_2}{b_2}$, то уравнения

$$\begin{aligned}
y_1 &= a_1x_1 + a_2x_2 + \ldots + a_nx_n, \\
y_2 &= b_1x_1 + b_2x_2 + \ldots + b_nx_n, \\
y_i &= x_i, \quad i = 3, 4, \ldots, n,
\end{aligned}$$

определяют невырожденное линейное преобразование переменных, в результате которого квадратичная форма f приводится к виду $f(y_1, y_2, \ldots, y_n) = y_1y_2$. С помощью дополнительного невырожденного преобразования переменных

$$\begin{aligned}
y_1 &= z_1 + z_2, \\
y_2 &= z - 1 - z_2, \\
y_i &= z_i, \quad i = 3, 4, \ldots, n,
\end{aligned}$$

квадратичная форма приводится к каноническому виду

$$f(z_1, z_2, \ldots, z_n) = z_1^2 - z_2^2,$$

из которого заключаем, что квадратичная форма f имеет ранг 2 и сигнатуру 0.

Докажем обратное утверждение теоремы. Если ранг квадратичной формы равен единице, то невырожденным преобразованием переменных

$$\begin{cases}
y_1 = \alpha_{11}x_1 + \alpha_{12}x_2 + \ldots + \alpha_{1n}x_n, \\
y_2 = \alpha_{21}x_1 + \alpha_{22}x_2 + \ldots + \alpha_{2n}x_n, \\
\dots\dots\dots\dots\dots\dots \\
y_n = \alpha_{n1}x_1 + \alpha_{n2}x_2 + \ldots + \alpha_{nn}x_n
\end{cases} \tag{9.20}$$

она приводится к каноническому виду $f(y_1, y_2, \ldots, y_n) = a_{11}y_1^2$. Здесь легко выделить линейные множители $a_{11}y_1$ и y_1, которые в исходных переменных дают следующее представление:

$$f(x_1, x_2, \ldots, x_n) = a_{11}\left(\alpha_{11}x_1 + \alpha_{12}x_2 + \ldots + \alpha_{1n}x_n\right)^2.$$

Следовательно, квадратичная форма f распадающаяся.

Пусть квадратичная форма f имеет ранг 2 и сигнатуру 0. В этом случае невырожденным преобразованием переменных вида (9.20) она приводится к каноническому виду $f(y_1, y_2, \ldots, y_n) = y_1^2 - y_2^2$. Учитывая тождество

$$y_1^2 - y_2^2 = (y_1 + y_2)(y_1 - y_2),$$

получаем следующее представление квадратичной формы в исходных переменных:

$$\begin{aligned} f(x_1, x_2, \ldots, x_n) &= \\ &= \bigl[(\alpha_{11} + \alpha_{21})x_1 + (\alpha_{12} + \alpha_{22})x_2 + \ldots + (\alpha_{1n} + \alpha_{2n})x_n\bigr] \times \\ &\qquad \times \bigl[(\alpha_{11} - \alpha_{21})x_1 + (\alpha_{12} - \alpha_{22})x_2 + \ldots + (\alpha_{1n} - \alpha_{2n})x_n\bigr]. \end{aligned}$$

Отсюда следует, что квадратичная форма распадающаяся. ▶

Пример 9.3. Квадратичные формы

$$f_1(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 + 2x_1x_2 + 2x_1x_3 + 2x_2x_3,$$
$$f_2(x_1, x_2, x_3) = x_1^2 - x_2^2 - x_3^2 - 2x_2x_3$$

приводятся соответственно к виду

$$f_1(x_1, x_2, x_3) = (x_1 + x_2 + x_3)^2,$$
$$f_2(x_1, x_2, x_3) = x_1^2 - (x_2 + x_3)^2 = (x_1 + x_2 + x_3)(x_1 - x_2 - x_3).$$

Значит, эти формы распадающиеся, причем у первой из них ранг равен единице, а у второй ранг равен двум, сигнатура — нулю.

9.8. Квадратичные формы в евклидовом пространстве

Пусть в евклидовом пространстве E дана квадратичная форма, которая в некотором ортонормированном базисе имеет вид $f(x) = x^{\mathrm{T}}Ax$. Здесь A — симметрическая матрица, которая в ортонормированном базисе является матрицей симметрического оператора. Согласно теореме 8.20 симметрическая матрица A некоторой ортогональной матрицей Q приводится к диагональному виду

$$Q^{-1}AQ = \Lambda \tag{9.21}$$

с характеристическими числами матрицы A по диагонали. Матрица Q в этом соотношении состоит из столбцов координат ортонормированных собственных векторов матрицы A по соответствующим характеристическим числам.

Применим к квадратичной форме $f(x) = x^{\mathrm{T}}Ax$ ортогональное преобразование переменных $x = Qy$ с матрицей Q из соотношения (9.21). Тогда матрица A квадратичной формы $f(x) = x^{\mathrm{T}}Ax$ преобразуется в матрицу $Q^{\mathrm{T}}AQ$. Так как матрица Q ортогональная, то $Q^{\mathrm{T}} = Q^{-1}$. Поэтому выражение $Q^{\mathrm{T}}AQ$ принимает вид $Q^{-1}AQ$. Это означает, что матрица A квадратичной формы $f(x) = x^{\mathrm{T}}Ax$ при ортогональном преобразовании $x = Qy$ с матрицей Q преобразуется так же, как матрица A симметрического оператора, т.е. по закону (9.21). Таким образом, ортогональное преобразование переменных $x = Qy$ приводит квадратичную форму $f(x) = x^{\mathrm{T}}Ax$ к каноническому виду, в котором коэффициентами при квадратах переменных являются характеристические числа матрицы A, причем канонический базис пространства E для этой квадратичной формы состоит из ортонормированной системы собственных векторов матрицы A.

Приведение квадратичной формы к каноническому виду в ортонормированном базисе называют ***приведением квадратичной формы к главным осям.*** Из сказанного выше вытекает следующее правило построения канонического вида квадратичной формы.

1. Найти собственные значения матрицы квадратичной формы, считая каждое собственное значение столько раз, какова его алгебраическая кратность.

2. По найденным собственным значениям λ_1, λ_2, ..., λ_n записать канонический вид квадратичной формы в главных осях

$$f(y_1, y_2, \ldots, y_n) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2. \tag{9.22}$$

Для определения ортогонального преобразования, приводящего квадратичную форму к каноническому виду, или, другими словами, для построения канонического ортонормированного базиса данной квадратичной формы необходимо в дополнение к предыдущему выполнить следующее.

1. Для каждого собственного значения λ_i матрицы квадратичной формы построить фундаментальную систему решений однородной системы $(A-\lambda_i E)x=0$.

2. Каждую из найденных ФСР ортонормировать.

3. Ортонормированные ФСР, полученные для каждого собственного значения, объединить в одну систему. Эта система будет каноническим базисом квадратичной формы.

4. Столбцы координат ортонормированных ФСР записать в матрицу. Эта матрица будет матрицей Q ортогонального преобразования переменных, приводящего квадратичную форму к каноническому виду в главных осях.

Пример 9.4. Привести квадратичную форму

$$f(x_1,x_2,x_3)=x_1^2-2x_2^2+x_3^2+4x_1x_2-8x_1x_3-4x_2x_3$$

к главным осям и указать ортогональное преобразование переменных, осуществляющее такое приведение.

Решение. Данная квадратичная форма имеет матрицу

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

Корнями ее характеристического многочлена

$$|A-\lambda E|=\begin{vmatrix} 1-\lambda & 2 & -4 \\ 2 & -2-\lambda & -2 \\ -4 & -2 & 1-\lambda \end{vmatrix}=-\lambda^3+27\lambda+54$$

являются $\lambda_1=6$, $\lambda_2=\lambda_3=-3$. Поэтому рассматриваемая квадратичная форма в главных осях имеет канонический вид

$$f=6y_1^2-3y_2^2-3y_3^2.$$

Перейдем к определению матрицы Q ортогонального преобразования переменных. Чтобы сделать это, построим фундаментальные системы решений однородных систем уравнений $(A-\lambda_i E)x=0$ и ортонормируем их. При $\lambda=6$ система $(A-\lambda E)x=0$ имеет вид:

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

Ее общим решением является $x=(-x_3,\ -x_3/2,\ x_3)^{\mathrm{T}}$, фундаментальная система решений состоит из одного вектора, например $b_1=(2,\ 1,\ -2)^{\mathrm{T}}$. Нормируя его, получим вектор

$$e_1'=\left(\frac{2}{3},\ \frac{1}{3},\ -\frac{2}{3}\right)^{\mathrm{T}}.$$

При $\lambda=-3$ система $(A-\lambda E)x=0$ имеет вид:

$$\begin{cases} 4x_1+2x_2-4x_3=0, \\ 2x_1+x_2-2x_3=0, \\ -4x_1-2x_2+4x_3=0. \end{cases}$$

Ее общим решением является $x=(x_1,\ -2x_1+2x_3,\ x_3)^{\mathrm{T}}$, а фундаментальная система решений состоит из двух векторов, например $b_2=(1,\ 2,\ 2)^{\mathrm{T}}$ и $b_3=(2,\ -2,\ 1)^{\mathrm{T}}$. Эти векторы уже ортогональны, поэтому их не надо ортогонализировать, достаточно лишь нормировать. После нормирования получим векторы

$$e_2'=\left(\frac{1}{3},\ \frac{2}{3},\ \frac{2}{3}\right)^{\mathrm{T}},\quad e_3'=\left(\frac{2}{3},\ -\frac{2}{3},\ \frac{1}{3}\right)^{\mathrm{T}}.$$

Векторы e_1', e_2', e_3' составляют канонический ортонормированный базис данной квадратичной формы. Из столбцов их координат строим искомую ортогональную матрицу

$$Q=\begin{pmatrix} \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{pmatrix}.$$

По строкам этой матрицы записываем искомое ортогональное преобразование переменных:

$$x_1 = \frac{2}{3}y_1 + \frac{1}{3}y_2 + \frac{2}{3}y_3,$$
$$x_2 = \frac{1}{3}y_1 + \frac{2}{3}y_2 - \frac{2}{3}y_3,$$
$$x_3 = -\frac{2}{3}y_1 + \frac{2}{3}y_2 + \frac{1}{3}y_3.$$

Поскольку собственные числа матрицы квадратичной формы являются коэффициентами этой формы, приведенной к главным осям, заключаем, что по собственным числам можно делать заключения о положительной (отрицательной) определенности квадратичной формы. Сводка соответствующих результатов приведена в табл. 9.1.

Таблица 9.1

Тип квадратичной формы	Вид собственных чисел
Положительно определенная	Все собственные числа положительны
Отрицательно определенная	Все собственные числа отрицательны
Положительно полуопределенная	Все собственные числа неотрицательны
Отрицательно полуопределенная	Все собственные числа неположительны
Неопределенная	Есть и положительные, и отрицательные собственные числа

9.9. Пары квадратичных форм

Пусть даны две квадратичные формы от n переменных

$$f(x_1, x_2, \ldots, x_n) \quad \text{и} \quad g(x_1, x_2, \ldots, x_n).$$

Можно доказать (см. [16]), что в общем случае не существует невырожденного преобразования переменных, одновременно приводящего две формы к каноническому виду. Если же одна из данных форм, например первая, положительно определенная, то такое преобразование переменных существует. Действительно, совершим преобразование переменных $x = Ty$, приводящее первую форму к нормальному виду

$$f(y_1, y_2, \ldots, y_n) = y_1^2 + y_2^2 + \ldots + y_n^2.$$

Этим же преобразованием переменных вторая квадратичная форма преобразуется в квадратичную форму $g(y_1, y_2, \ldots, y_n)$. Совершим теперь ортогональное преобразование переменных $y = Qz$, приводящее форму $g(y_1, y_2, \ldots, y_n)$ к каноническому виду в главных осях:

$$g(z_1, z_2, \ldots, z_n) = \lambda_1 z_1^2 + \lambda_2 z_2^2 + \ldots + \lambda_n z_n^2.$$

Матрицей квадратичной формы в переменных y_1, y_2, $\ldots$, y_n является единичная матрица. В результате ортогонального преобразования, определяемого матрицей Q, единичная матрица снова перейдет в единичную матрицу. Это следует из равенства $Q^{\mathrm{T}} E Q = E$. Поэтому в переменных z_1, z_2, $\ldots$, z_n квадратичная форма f будет иметь вид:

$$f(z_1, z_2, \ldots, z_n) = z_1^2 + z_2^2 + \ldots + z_n^2.$$

Таким образом, преобразование переменных $x = TQz$ с матрицей TQ одновременно приводит первую форму к нормальному виду, а вторую — к каноническому. Поясним это на примере.

Пример 9.5. Даны квадратичные формы

$$f(x_1, x_2) = x_1^2 - 2x_1 x_2 + 4x_2^2 \quad \text{и} \quad g(x_1, x_2) = -4x_1 x_2.$$

Привести линейным невырожденным преобразованием переменных первую квадратичную форму к нормальному, а вторую — к каноническому виду и указать преобразование переменных, осуществляющее такое приведение квадратичных форм.

Решение. Матрица квадратичной формы f в переменных x_1, x_2 имеет вид:

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

Ее угловые миноры $\Delta_1 = 1$ и $\Delta_2 = |A| = 3$ положительны. Поэтому квадратичная форма f положительно определенная. Приведем ее к каноническому виду. Для этого выделим в ней полный квадрат по x_1:

$$f(x_1, x_2) = (x_1^2 - 2x_1x_2) + 4x_2^2 = (x_1 - x_2)^2 + 3x_2^2.$$

Введем новые переменные $y_1 = x_1 - x_2$, $y_2 = x_2$, которым соответствует невырожденное преобразование переменных $x_1 = y_1 + y_2$, $x_2 = y_2$ с матрицей

$$T_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

В новых переменных квадратичная форма f имеет канонический вид $f(y_1, y_2) = y_1^2 + 3y_2^2$. Выполним дополнительное преобразование переменных $z_1 = y_1$, $z_2 = \sqrt{3}y_2$, или $y_1 = z_1$, $y_2 = \frac{1}{\sqrt{3}}z_2$ с матрицей

$$T_2 = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{\sqrt{3}} \end{pmatrix}.$$

В переменных z_1, z_2 квадратичная форма f имеет нормальный вид $f(z_1, z_2) = z_1^2 + z_2^2$.

Невырожденное преобразование переменных, результирующее два выполненных преобразования, имеет матрицу

$$T = T_1T_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{\sqrt{3}} \end{pmatrix} = \begin{pmatrix} 1 & \frac{1}{\sqrt{3}} \\ 0 & \frac{1}{\sqrt{3}} \end{pmatrix}.$$

Поэтому оно записывается следующим образом:

$$x_1 = z_1 + \frac{1}{\sqrt{3}} z_2, \quad x_2 = \frac{1}{\sqrt{3}} z_2.$$

Запишем форму g в переменных z_1, z_2:

$$g(z_1, z_2) = -4\left(z_1 + \frac{1}{\sqrt{3}} z_2\right) \frac{1}{\sqrt{3}} z_2 = -\frac{4}{\sqrt{3}} z_1 z_2 - \frac{4}{3} z_2^2.$$

Эту форму будем приводить к каноническому виду в главных осях. Матрица

$$B = \begin{pmatrix} 0 & -\frac{2}{\sqrt{3}} \\ -\frac{2}{\sqrt{3}} & -\frac{4}{3} \end{pmatrix}$$

квадратичной формы g имеет характеристический многочлен

$$|B - \lambda E| = \lambda^2 + \frac{4}{3}\lambda - \frac{4}{3}.$$

Его корнями являются $\lambda_1 = -2$, $\lambda_2 = \frac{2}{3}$. Поэтому форма g в главных осях имеет канонический вид $g(t_1, t_2) = -2t_1^2 + \frac{2}{3} t_2^2$. Чтобы найти преобразование переменных, осуществляющее приведение формы g с матрицей B к главным осям, построим фундаментальные системы решений для систем $(B - \lambda_i E)z = 0$ и ортонормируем их.

При $\lambda = -2$ система $(B - \lambda E)z = 0$ имеет вид:

$$\begin{cases} 2z_1 - \frac{2}{\sqrt{3}} z_2 = 0, \\ -\frac{2}{\sqrt{3}} z_1 + \frac{2}{\sqrt{3}} z_2 = 0. \end{cases}$$

Ее общее решение $z = \left(z_2/\sqrt{3},\ z_2\right)^{\mathrm{T}}$ имеет одно свободное неизвестное, а ФСР состоит из одного решения, например, $b_1 = \left(1,\ \sqrt{3}\right)^{\mathrm{T}}$. Нормируя его, получим вектор $e_1' = \frac{1}{2}\left(1,\ \sqrt{3}\right)^{\mathrm{T}}$.

При $\lambda = 2/3$ аналогичным образом построим вектор $e_2' = \frac{1}{2}(\sqrt{3}, \ -1)^{\mathrm{T}}$. Из столбцов координат векторов e_1' и e_2' составим матрицу

$$Q = \frac{1}{2}\begin{pmatrix} 1 & \sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}.$$

Для этой матрицы запишем соответствующее преобразование переменных

$$z_1 = \frac{1}{2}t_1 + \frac{\sqrt{3}}{2}t_2, \qquad z_2 = \frac{\sqrt{3}}{2}t_1 - \frac{1}{2}t_2.$$

При этом преобразовании квадратичная форма $f(x_1, z_2) = z_1^2 + z_2^2$ преобразуется в форму $f(t_1, t_2) = t_1^2 + t_2^2$. Матрицей искомого линейного преобразования переменных будет матрица

$$TQ = \begin{pmatrix} 1 & \frac{1}{\sqrt{3}} \\ 0 & \frac{1}{\sqrt{3}} \end{pmatrix} \frac{1}{2}\begin{pmatrix} 1 & \sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix} = \begin{pmatrix} 1 & \frac{\sqrt{3}}{2} \\ \frac{1}{2} & -\frac{\sqrt{3}}{6} \end{pmatrix}.$$

Для этой матрицы записываем искомое преобразование переменных

$$x_1 = t_1 + \frac{\sqrt{3}}{3}t_2, \qquad x_2 = \frac{1}{2} - \frac{\sqrt{3}}{6}t_2.$$

Этим преобразованием переменных квадратичная форма f приводится к нормальному виду $f(t_1, t_2) = t_1^2 + t_2^2$, а квадратичная форма g — к каноническому виду $g(t_1, t_2) = -2t_1^2 + \frac{2}{3}t_2^2$.

9.10. Квадратичные формы в комплексном линейном пространстве

Квадратичную форму от комплексных переменных с комплексными коэффициентами иногда определяют так же, как была определена действительная квадратичная форма (см. разд. 9.1). Свойства такой квадратичной формы аналогичны свойствам действительной квадратичной формы (см. [16],

[28]). Однако чаще используют другой подход к понятию комплексной квадратичной формы.

Квадратичной формой от комплексных переменных $x_1, x_2, \ldots, x_n$ называют выражение

$$f(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i\overline{x_j}. \tag{9.23}$$

Переменные $x_1, x_2, \ldots, x_n$ можно рассматривать как координаты вектора x в некотором фиксированном базисе комплексного n-мерного линейного пространства X. Поэтому квадратичную форму (9.23) можно рассматривать как числовую функцию векторного аргумента $x \in X$. Матрицу

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

составленную из коэффициентов квадратичной формы (9.23), называют ***матрицей*** этой ***квадратичной формы***. Непосредственной проверкой легко убедиться в том, что квадратичную форму (9.23) кратко можно записать в виде

$$f(x) = x^{\mathrm{T}} A\overline{x}.$$

При линейном преобразовании переменных $x = Qy$ с матрицей Q матрица A квадратичной формы (9.23) преобразуется в матрицу

$$B = Q^{\mathrm{T}} A\overline{Q}. \tag{9.24}$$

Квадратичную форму (9.23) называют ***эрмитовой***, если ее коэффициенты удовлетворяют условиям

$$a_{ij} = \overline{a_{ji}}. \tag{9.25}$$

В силу этих условий выполняются соотношения

$$\overline{A} = A^{\mathrm{T}}, \quad A = \overline{A}^{\mathrm{T}} = A^{*}.$$

Следовательно, ***матрица эрмитовой квадратичной формы является эрмитовой.*** В ней элементы главной диагонали — вещественные числа. Эрмитова квадратичная форма принимает только вещественные значения, так как

$$\overline{x^{\mathrm{T}} A \overline{x}} = \overline{x}^{\mathrm{T}} \overline{A} x = \overline{x}^{\mathrm{T}} A^{\mathrm{T}} x = (x^{\mathrm{T}} A \overline{x})^{\mathrm{T}} = x^{\mathrm{T}} A \overline{x}.$$

Теория действительных квадратичных форм с незначительными изменениями переносится на комплексный случай. Приведем основные результаты для эрмитовых квадратичных форм.

Любая эрмитова квадратичная форма некоторым невырожденным линейным преобразованием переменных приводится к каноническому виду

$$\boldsymbol{f(y_1, y_2, \ldots, y_n) = c_1 y_1 \overline{y_1} + \ldots + c_n y_n \overline{y_n} =}$$
$$\boldsymbol{= c_1 |y_1|^2 + c_2 |y_2|^2 + \ldots + c_n |y_n|^2} \quad (9.26)$$

с действительными коэффициентами $\boldsymbol{c_1}$, $\boldsymbol{c_2}$, ..., $\boldsymbol{c_n}$.

Такое приведение эрмитовой формы можно осуществить, как и в случае действительных форм, методом Лагранжа. Сначала заметим, что если в эрмитовой квадратичной форме нет квадратов модулей переменных, то их можно получить, выполнив дополнительное преобразование переменных. Действительно, если в эрмитовой форме

$$f(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i \overline{x_j}$$

все коэффициенты $a_{ii} = 0$, то в ней есть два слагаемых $a_{ij} x_i \overline{x_j} + a_{ji} x_j \overline{x_i}$, для которых $a_{ij} = \overline{a_{ji}} \neq 0$. Выполним вспомогательное преобразование переменных

$$x_i = a_{ji}(y_i + y_j), \quad x_j = y_i - y_j, \quad x_k = y_k \quad \text{при} \quad k \neq i, j.$$

Тогда рассматриваемая сумма членов преобразуется следующим образом:

$$a_{ij} x_i \overline{x_j} + a_{ji} x_j \overline{x_i} = |a_{ij}|^2 (y_i + y_j)(\overline{y_i} - y_j) +$$
$$+ |a_{ij}|^2 (y_i - y_j)(\overline{y_i} + \overline{y_j}) = 2|a_{ij}|^2 |y_i|^2 - 2|a_{ij}|^2 |y_j|^2.$$

Следовательно, после такого преобразования переменных в квадратичной форме появляются члены с квадратами модулей переменных.

Замечание. Если в эрмитовой квадратичной форме все коэффициенты a_{ii} равны нулю и коэффициент $a_{ij}=\overline{a_{ij}}$, отличный от нуля, не является чисто мнимым числом, то для получения квадратов модулей переменных можно также пользоваться преобразованием переменных по формулам

$$x_i = y_i + y_j, \qquad x_j = y_i - y_j, \qquad x_k = y_k \quad \text{при} \quad k \neq i, j,$$

так как выражение $a_{ij}x_i\overline{x_j} + a_{ji}x_j\overline{x_i}$ в этом случае преобразуется в выражение

$$(a_{ij}+a_{ji})|y_i|^2 - (a_{ij}+a_{ji})|y_j|^2 + (a_{ji}-a_{ij})y_i\overline{y_j} + (a_{ij}-a_{ji})y_j\overline{y_i}.$$

В силу сделанного замечания можно предполагать, что в эрмитовой квадратичной форме есть отличные от нуля коэффициенты a_{ii}. Пусть, например, $a_{11} \neq 0$. Тогда в форме $f(x_1, x_2, \ldots, x_n)$ можно собрать все члены вида $a_{i1}x_i\overline{x_1}$ и форму преобразовать следующим образом:

$$\begin{aligned}
&f(x_1, x_2, \ldots, x_n) = \\
&\quad = (a_{11}x_1\overline{x_1} + a_{21}x_2\overline{x_1} + \ldots + a_{n1}x_n\overline{x_1}) + a_{12}x_1\overline{x_2} + \ldots \\
&\quad \ldots + a_{1n}x_1\overline{x_n} + a_{22}x_2\overline{x_2} + a_{23}x_2\overline{x_3} + a_{32}x_3\overline{x_2} + a_{nn}x_b\overline{x_n} = \\
&= \frac{1}{\overline{a_{11}}}(a_{11}x_1 + a_{21}x_2 + \ldots + a_{n1}x_n)\left(\overline{a_{11}x_1} + \overline{a_{21}x_2} + \ldots + \overline{a_{n1}x_n}\right) + \\
&\qquad\qquad + g(x_2, x_3, \ldots, x_n) = \\
&\qquad = \frac{1}{\overline{a_{11}}}\left|a_{11}x_1 + a_{21}x_2 + \ldots + a_{n1}x_n\right|^2 + g(x_2, x_3, \ldots, x_n).
\end{aligned}$$

Если теперь ввести новые переменные

$$y_1 = a_{11}x_1 + a_{21}x_2 + \ldots + a_{n1}x_n, \quad y_k = x_k \quad \text{при} \quad k \neq 1,$$

то форма f в новых переменных будет иметь вид:

$$f = \frac{1}{a_{11}}|y_1|^2 + g(y_2, y_3, \ldots, y_n).$$

Таким образом, в форме f выделен квадрат модуля переменной y_1. С формой $g(y_2, y_3, \ldots, y_n)$ можно поступать аналогично. Не более чем через $n-1$ шагов придем к каноническому виду формы f.

Пример 9.6. Методом Лагранжа привести к каноническому виду эрмитову квадратичную форму

$$f(x_1,x_2,x_3) = (1+i)x_1\overline{x_2} + (1-i)x_2\overline{x_1} + ix_1\overline{x_3} - \\ - ix_3\overline{x_1} + (2+2i)x_2\overline{x_3} + (2-2i)x_3\overline{x_2}$$

и указать невырожденное линейное преобразование переменных, осуществляющее такое преобразование квадратичной формы f.

Решение. В данной квадратичной форме нет членов с квадратами модулей переменных, но есть, например, член $(1+i)x_1\overline{x_2}$. Поэтому сначала выполним преобразование переменных по правилу

$$x_1 = (1-i)(y_1+y_2), \quad x_2 = y_1 - y_2, \quad x_3 = y_3$$

с матрицей

$$Q_1 = \begin{pmatrix} 1-i & 1-i & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

В новых переменных квадратичная форма f принимает вид:

$$f(y_1,y_2,y_3) = 2(y_1+y_2)(\overline{y_1}-\overline{y_2}) + 2(y_1-y_2)(\overline{y_1}+\overline{y_2}) + \\ + (1+i)(y_1+y_2)\overline{y_3} + (1-i)y_3(\overline{y_1}+\overline{y_3}) + (2+2i)(y_1-y_2)\overline{y_3} + \\ + (2-2i)y_3(\overline{y_1}-\overline{y_2}) = 4y_1\overline{y_1} - 4y_2\overline{y_2} + (3+3i)y_1\overline{y_3} + \\ + (3-3i)y_3\overline{y_1} - (1+i)y_2\overline{y_3} - (1-i)y_3\overline{y_2}.$$

В полученной квадратичной форме выделим квадрат модуля по переменной y_1:

$$f(y_1,y_2,y_3) = \left[4y_1\overline{y_1} + (3-3i)y_3\overline{y_1}\right] - 4y_2\overline{y_2} + (3+3i)y_1\overline{y_3} - \\ - (1+i)y_2\overline{y_3} - (1-i)y_3\overline{y_2} = \frac{1}{4}\left[4y_1 + (3-3i)y_3\right]\left[4\overline{y_1} + (3+3i)\overline{y_3}\right] - \\ - 4y_2\overline{y_2} - (1+i)y_2\overline{y_3} - (1-i)y_3\overline{y_2} - \frac{9}{2}y_3\overline{y_3}.$$

Введем новые переменные

$$z_1 = 4y_1 + (3-3i)y_3, \qquad z_2 = y_2, \qquad z_3 = y_3,$$

соответствующие преобразованию переменных

$$y_1 = \frac{1}{4}z_1 - \frac{3-3i}{4}z_3, \qquad y_2 = z_2, \qquad y_3 = z_3$$

с матрицей

$$Q_2 = \begin{pmatrix} \frac{1}{4} & 0 & -\frac{3-3i}{4} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

В новых переменных квадратичная форма f принимает вид:

$$f(z_1, z_2, z_3) = \frac{1}{4}z_1\overline{z_1} - 4z_2\overline{z_2} - (1+i)z_2\overline{z_3} - (1-i)z_3\overline{z_2} - \frac{9}{2}z_3\overline{z_3}.$$

Здесь выделим квадрат модуля по z_2:

$$\begin{aligned} f(z_1, z_2, z_3) &= \\ &= \frac{1}{4}z_1\overline{z_1} + \left[-4z_2\overline{z_2} - (1-i)z_3\overline{z_2}\right] - (1+i)z_2\overline{z_3} - \frac{9}{2}z_3\overline{z_3} = \\ &= \frac{1}{4}z_1\overline{z_1} - \frac{1}{4}\left[-4z_2 - (1-i)z_3\right] \times \\ &\qquad \times \left[-4\overline{z_2} - (1+i)\overline{z_3}\right] - 4z_3\overline{z_3}. \qquad (9.27) \end{aligned}$$

Введем новые переменные

$$t_1 = z_1, \quad t_2 = -4z_2 - (1-i)z_3, \quad t_3 = z_3,$$

соответствующие невырожденному преобразованию

$$z_1 = t_1, \quad z_2 = -\frac{1}{4}t_2 - \frac{1-i}{4}t_3, \quad z_3 = t_3$$

с матрицей

$$Q_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{4} & -\frac{1-i}{4} \\ 0 & 0 & 1 \end{pmatrix}.$$

В новых переменных квадратичная форма f уже принимает канонический вид:

$$f(t_1,t_2,t_3)=\frac{1}{4}t_1\overline{t_1}-\frac{1}{4}t_2\overline{t_2}-4t_3\overline{t_3}=\frac{1}{4}|t_1|^2-\frac{1}{4}|t_2|^2-4|t_3|^2.$$

Найдем матрицу результирующего преобразования переменных:

$$Q=Q_1Q_2Q_3=\begin{pmatrix}1-i & 1-i & 0\\ 1 & -1 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}\frac{1}{4} & 0 & -\frac{3-3i}{4}\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\times$$

$$\times\begin{pmatrix}1 & 0 & 0\\ 0 & -\frac{1}{4} & -\frac{1-i}{4}\\ 0 & 0 & 1\end{pmatrix}=\begin{pmatrix}\frac{1-i}{4} & -\frac{1-i}{4} & 2i\\ \frac{1}{4} & \frac{1}{4} & -\frac{1-i}{2}\\ 0 & 0 & 1\end{pmatrix}.$$

По строкам матрицы Q запишем само искомое преобразование переменных:

$$\begin{cases}x_1=\frac{1-i}{4}t_1-\frac{1-i}{4}t_2+2it_3,\\ x_2=\frac{1}{4}t_1+\frac{1}{4}t_2-\frac{1-i}{2}t_3,\\ x_3=t_3.\end{cases}$$

На эрмитовы квадратичные формы переносятся определение нормального вида квадратичной формы и закон инерции квадратичных форм.

В унитарном пространстве эрмитовой квадратичной форме с матрицей A соответствует эрмитов оператор с матрицей $\overline{A}=A^{\mathrm{T}}$. Как и в действительном случае (см. разд. 9.8), это соответствие используется при обосновании следующего утверждения.

Теорема 9.4. *Любая эрмитова квадратичная форма f с матрицей A от n переменных x_1, x_2, ..., x_n некоторым* *унитарным преобразованием переменных приводится к каноническому виду*

$$f(y_1, y_2, \ldots, y_n) = \lambda_1 y_1 \overline{y_1} + \lambda_2 y_2 \overline{y_2} + \ldots + \lambda_n y_n \overline{y_n} = \\ = \lambda_1 |y_1|^2 + \lambda_2 |y_1|^2 + \ldots + \lambda_n |y_n|^2,$$

причем действительные коэффициенты λ_1, λ_2, $\ldots$, λ_n *этого канонического вида являются характеристическими числами матрицы* $\overline{A}$ *(а следовательно, матрицы* A*), а канонический ортонормированный базис квадратичной формы* f *состоит из ортонормированных собственных векторов матрицы* $\overline{A} = A^{\mathrm{T}}$.

Из столбцов координат векторов этого базиса строится унитарная матрица Q, удовлетворяющая условию $Q^{\mathrm{T}} A \overline{Q} = \Lambda$. По строкам матрицы Q записывается унитарное преобразование переменных, приводящее эрмитову квадратичную форму к главным осям. В качестве унитарной матрицы T, удовлетворяющей условию $T^{-1}AT = \Lambda$, естественно брать матрицу $T = \overline{Q}$.

Пример 9.7. Привести к главным осям эрмитову квадратичную форму

$$f(x_1, x_2) = 3x_1\overline{x_1} + (2+2i)x_1\overline{x_2} + (2-2i)x_2\overline{x_1} + x_2\overline{x_2}$$

и указать унитарное преобразование переменных, осуществляющее это преобразование.

Решение. Матрицей данной эрмитовой квадратичной формы является эрмитова матрица

$$A = \begin{pmatrix} 3 & 2+2i \\ 2-2i & 1 \end{pmatrix}.$$

Комплексно сопряженная матрица имеет вид:

$$\overline{A} = \begin{pmatrix} 3 & 2-2i \\ 2+2i & 1 \end{pmatrix}.$$

Характеристический многочлен

$$|\overline{A}-\lambda E| = \begin{vmatrix} 3-\lambda & 2-2i \\ 2+2i & 1-\lambda \end{vmatrix} = \lambda^2 - 4\lambda - 5$$

имеет корни $\lambda_1 = 5$, $\lambda_2 = -1$. Поэтому форма f в главных осях имеет вид:

$$f(y_1, y_2) = 5|y_1|^2 - |y_2|^2.$$

Перейдем к построению ортонормированного канонического базиса эрмитовой квадратичной формы f.

При $\lambda = 5$ система $(\overline{A} - \lambda E)X = 0$, или

$$\begin{cases} -2x_1 + (2-2i)x_2 = 0, \\ (2+2i)x_2 - 4x_2 = 0, \end{cases}$$

имеет фундаментальную систему решений из одного вектора, например, $b_1 = (1-i,\ 1)^{\mathrm{T}}$. Нормируя его, получим:

$$e_1' = \left(\frac{1-i}{\sqrt{3}},\ \frac{1}{\sqrt{3}}\right)^{\mathrm{T}}.$$

При $\lambda = -1$ аналогичным образом получаем вектор

$$e_2' = \left(\frac{1-i}{\sqrt{6}},\ \frac{2}{\sqrt{6}}\right)^{\mathrm{T}}.$$

Векторы e_1' и e_2' образуют канонический ортонормированный базис эрмитовой формы f. Из столбцов координат векторов e_1', e_2' составим матрицу

$$Q = \begin{pmatrix} \dfrac{1-i}{\sqrt{3}} & \dfrac{1-i}{\sqrt{6}} \\ \dfrac{1}{\sqrt{3}} & -\dfrac{2}{\sqrt{6}} \end{pmatrix}$$

и по ее строкам запишем искомое преобразование переменных

$$x_1 = \frac{1-i}{\sqrt{3}}y_1 + \frac{1-i}{\sqrt{6}}y_2, \qquad x_2 = \frac{1}{\sqrt{3}}y_1 - \frac{2}{\sqrt{6}}y_2.$$

Матрица Q удовлетворяет условию $Q^{\mathrm{T}}A\overline{Q} = \Lambda$, а матрица $T = \overline{Q}$ — условию $T^{-1}AT = \Lambda$.

В заключение отметим, что на эрмитовы квадратичные формы переносятся определение *положительно* (*отрицательно*) *определенной квадратичной формы* и *критерий Сильвестра*, а также результаты об одновременном приведении пары квадратичных форм к каноническому виду (см. разд. 9.9).

Упражнения

9.1. Записать в матричном виде квадратичную форму:

1) $f(x_1,x_2,x_3)=x_1^2-2x_2^2-2x_3^2-4x_1x_2+4x_1x_3+8x_2x_3$;

2) $f(x_1,x_2,x_3)=3x_1^2+6x_2^2+3x_3^2-4x_1x_2-8x_1x_3-4x_2x_3$;

3) $f(x_1,x_2,x_3)=2x_1^2+5x_2^2+5x_3^2-4x_1x_2-4x_1x_3-8x_2x_3$.

9.2. Записать квадратичную форму $f(x_1,x_2,x_3)$ по ее матрице:

1) $\begin{pmatrix} 2 & -2 & -1 \\ -2 & 5 & 2 \\ -1 & 2 & 2 \end{pmatrix}$; 2) $\begin{pmatrix} 18 & 6 & -6 \\ 6 & 21 & 0 \\ -6 & 0 & 16 \end{pmatrix}$; 3) $\begin{pmatrix} 8 & 4 & -1 \\ 4 & -7 & 4 \\ -1 & 4 & 8 \end{pmatrix}$;

4) $\begin{pmatrix} 11 & 8 & 5 \\ 8 & 5 & -10 \\ 2 & -10 & 2 \end{pmatrix}$; 5) $\begin{pmatrix} 2 & -2 & 2 \\ -2 & 3 & 0 \\ 2 & 0 & 1 \end{pmatrix}$; 6) $\begin{pmatrix} 3 & -3 & -1 \\ -3 & 3 & 1 \\ -1 & 1 & 7 \end{pmatrix}$.

9.3. Методом Лагранжа привести квадратичную форму к каноническому виду и указать невырожденное преобразование переменных, осуществляющее такое приведение квадратичной формы:

1) $f(x_1,x_2,x_3)=x_1^2+2x_1x_2+2x_2^2+4x_2x_3+5x_3^2$;

2) $f(x_1,x_2,x_3)=x_1^2-4x_1x_2+2x_1x_3+4x_2+x_3^2$;

3) $f(x_1,x_2,x_3)=x_1x_2+x_1x_3+2x_1x_4+2x_2x_3+4x_3x_4$;

4) $f(x_1,x_2,x_3)=x_1x_2+2x_1x_3+4x_2x_3$;

5) $f(x_1,x_2,x_3)=x_1x_2+4x_1x_3+6x_2x_3$;

6) $f(x_1,x_2,x_3)=2x_1x_2+4x_1x_3+8x_2x_3$;

7) $f(x_1,x_2,x_3)=x_1^2+x_2^2+3x_3^2+2x_1x_2+4x_1x_3+6x_2x_3$;

8) $f(x_1,x_2,x_3)=2x_1^2+x_2^2+2x_3^2-4x_1x_2-8x_1x_3+4x_2x_3$;

9) $f(x_1,x_2,x_3)=2x_1^2+5x_2^2-3x_3^2-4x_1x_2+4x_1x_3-8x_2x_3$;

10) $f(x_1,x_2,x_3)=2x_1^2+6x_2^2+5x_3^2-8x_1x_2-6x_2x_3$.

9.4. Квадратичные формы из упражнения 9.3 привести к нормальному виду и указать для каждой из них ранг, положительный и отрицательный индекс инерции и сигнатуру.

9.5. Выяснить вопрос о положительной (отрицательной) определенности квадратичных форм с матрицей:

1) $\begin{pmatrix} 2 & -2 & -1 \\ -2 & 5 & 2 \\ -1 & 2 & 2 \end{pmatrix}$; 2) $\begin{pmatrix} 1 & -2 & 0 \\ -2 & 2 & -2 \\ 0 & -2 & 3 \end{pmatrix}$; 3) $\begin{pmatrix} -8 & 6 & 4 \\ 6 & -9 & 2 \\ 4 & 2 & -4 \end{pmatrix}$;

4) $\begin{pmatrix} 14 & 2 & -4 \\ 2 & 17 & 2 \\ -4 & 2 & 14 \end{pmatrix}$; 5) $\begin{pmatrix} -5 & -2 & 4 \\ -2 & -1 & 2 \\ 4 & 2 & -5 \end{pmatrix}$; 6) $\begin{pmatrix} 3 & 2 & 0 \\ 2 & 4 & -2 \\ 0 & -2 & 5 \end{pmatrix}$;

7) $\begin{pmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{pmatrix}$; 8) $\begin{pmatrix} 17 & -8 & 4 \\ -8 & 17 & -4 \\ 4 & -4 & 11 \end{pmatrix}$; 9) $\begin{pmatrix} 5 & 2 & -1 \\ 2 & 1 & -1 \\ -1 & -1 & 5 \end{pmatrix}$.

9.6. Привести к главным осям квадратичные формы:

1) $f(x_1,x_2,x_3) = 2x_1^2 + 5x_2^2 + 2x_3^2 - 4x_1x_2 - 2x_1x_3 + 4x_2x_3$;
2) $f(x_1,x_2,x_3) = -3x_1^2 + 4x_1x_2 + 10x_1x_3 - 4x_2x_3$;
3) $f(x_1,x_2,x_3) = -x_1^2 + x_2^2 - 5x_3^2 + 6x_1x_2 + 4x_2x_3$.

9.7. Найти ортогональное преобразование переменных, приводящее квадратичную форму к каноническому виду, и записать этот канонический вид:

1) $f(x_1,x_2,x_3) = 2x_1^2 + x_2^2 - 4x_1x_2 - 4x_2x_3$;
2) $f(x_1,x_2,x_3) = x_1^2 + 2x_2^2 + 3x_3^2 + 4x_1x_2 - 4x_2x_3$;
3) $f(x_1,x_2,x_3) = 3x_1^2 + 4x_2^2 + 5x_3^2 + 4x_1x_2 - 4x_2x_3$;
4) $f(x_1,x_2,x_3) = 2x_1^2 + 5x_2^2 + 5x_3^2 + 4x_1x_2 - 4x_1x_3 - 8x_2x_3$;
5) $f(x_1,x_2,x_3) = 5x_1^2 + 6x_2^2 + 4x_3^2 - 4x_1x_2 - 4x_1x_3$;
6) $f(x_1,x_2,x_3) = 7x_1^2 + 5x_2^2 + 3x_3^2 - 8x_1x_2 + 8x_2x_3$.

9.8. По характеристическим числам матриц квадратичных форм из упражнений 9.6 и 9.7 сделать заключение о положительной (отрицательной) определенности, неопределенности или полуопределенности этих квадратичных форм.

9.9. Найти невырожденное линейное преобразование переменных, приводящее одну из данных квадратичных форм к нормальному виду, а другую — к каноническому виду:

1) $f(x_1,x_2,x_3)=x_1^2+26x_2^2+10x_1x_2,$
$g(x_1,x_2,x_3)=x_1^2+56x_2^2+16x_1x_2;$

2) $f(x_1,x_2,x_3)=x_1^2+4x_2^2+2x_3+2x_1x_3,$
$g(x_1,x_2,x_3)=8x_1^2-28x_2^2+14x_3^2+16x_1x_2+14x_1x_3+32x_2x_3;$

3) $f(x_1,x_2,x_3)=x_1^2+17x_2^2+3x_3^2+4x_1x_3-2x_1x_3-14x_2x_3,$
$g(x_1,x_2,x_3)=x_1^2-15x_2^2+4x_1x_2-2x_1x_3+6x_2x_3.$

9.10. Привести к каноническому виду эрмитову квадратичную форму и указать невырожденное преобразование переменных, осуществляющее такое приведение квадратичной формы:

1) $f=(2+i)x_1\overline{x_2}+(2-i)x_2\overline{x_1}+ix_1\overline{x_1}-ix_3\overline{x_1}+ix_2\overline{x_3}-ix_3\overline{x_2};$
2) $f=(2+i)x_1\overline{x_2}+(2-i)x_2\overline{x_1}+ix_1\overline{x_3}-ix_3\overline{x_1}+ix_2\overline{x_3}-ix_3\overline{x_2};$
3) $f=x_1\overline{x_1}+(1+i)x_1\overline{x_2}+(1-i)x_2\overline{x_1}+ix_1\overline{x_3}-ix_3\overline{x_1};$
4) $f=x_1\overline{x_1}+x_2\overline{x_2}+x_3\overline{x_3}+(1+i)x_1\overline{x_3}+(1-i)x_3\overline{x_1}+ix_2\overline{x_3}-$
$-ix_3\overline{x_2};$
5) $f=x_1\overline{x_1}+x_2\overline{x_2}+x_3\overline{x_3}+ix_1\overline{x_2}-ix_2\overline{x_1}+(1+i)x_2\overline{x_3}+$
$+(1-i)x_3\overline{x_2}.$

9.11. Убедиться в положительной определенности эрмитовой квадратичной формы:

1) $f=3x_1\overline{x_1}-ix_1\overline{x_2}+ix_2\overline{x_1}+3x_2\overline{x_2}+4x_3\overline{x_3};$
2) $f=3x_1\overline{x_1}+(2-i)x_1\overline{x_2}+(2+i)x_2\overline{x_1}+7x_2\overline{x_2}+4x_3\overline{x_3};$
3) $f=ix_1\overline{x_2}-ix_2\overline{x_1}+2x_2\overline{x_2}+4x_3\overline{x_3}.$

9.12. Найти унитарное преобразование переменных, приводящее эрмитову квадратичную форму f к каноническому виду, и записать этот канонический вид, если:

1) $f=3x_1\overline{x_1}-ix_1\overline{x_2}+ix_2\overline{x_1}+3x_2\overline{x_2}+9x_3\overline{x_3};$
2) $f=3x_1\overline{x_1}+(2-i)x_1\overline{x_2}+(2+i)x_2\overline{x_1}+7x_2\overline{x_2}+16x_3\overline{x_3};$
3) $f=3x_1\overline{x_1}+(2+2i)x_1\overline{x_2}+(2-2i)x_2\overline{x_1}+x_2\overline{x_2}+9x_3\overline{x_3};$
4) $f=ix_1\overline{x_2}-ix_2\overline{x_1}+2x_2\overline{x_2}+4x_2\overline{x_2};$
5) $f=ix_1\overline{x_2}-ix_2\overline{x_1}+9x_2\overline{x_2};$
6) $f=3x_1\overline{x_1}+(2+2i)x_1\overline{x_2}+(2-2i)x_2\overline{x_1}+5x_2\overline{x_2}+x_3\overline{x_3};$
7) $f=3x_1\overline{x_1}+(2-2i)x_1\overline{x_2}+(2+2i)x_2\overline{x_1}+10x_2\overline{x_2}+4x_3\overline{x_3}.$

Глава 10 ИТЕРАЦИОННЫЕ МЕТОДЫ РЕШЕНИЯ СИСТЕМ ЛИНЕЙНЫХ УРАВНЕНИЙ

10.1. Метод итераций

При большом числе неизвестных в системе линейных уравнений применение метода Гаусса и других методов, дающих точное решение, становится весьма затруднительным. В таких случаях для решения системы могут оказаться более удобными приближенные методы. Рассмотрим ***метод итераций*** (***метод последовательных приближений***). Решение системы при этом методе получается как предел последовательности приближенных решений, которые находятся с помощью некоторого единообразного процесса, называемого ***процессом итераций***.

Принцип построения итерационного процесса состоит в следующем. Пусть дана квадратная система

$$Ax = b \tag{10.1}$$

с невырожденной матрицей A n-го порядка. Представим эту систему в виде

$$x = Bx + c, \tag{10.2}$$

где $x = (x_1,\ x_2,\ \ldots,\ x_n)^{\text{т}}$, $c = (c_1,\ c_2,\ \ldots,\ c_n)^{\text{т}}$, B — квадратная матрица порядка n. К такому виду системы можно прийти, например, если в системе (10.1) первое уравнение разрешить относительно x_1, второе — относительно x_2 и т.д.

Систему (10.2) называют ***приведенной*** и решают ***методом последовательных приближений.*** За нулевое приближение принимают какой-либо вектор

$$x^{(0)} = \left(x_1^{(0)},\, x_2^{(0)},\, \ldots,\, x_n^{(0)}\right)^{\mathrm{T}}.$$

На практике часто за нулевое приближение берут столбец c свободных членов системы (10.2). Подставив $x^{(0)}$ вместо x в правую часть системы (10.2), получают первое приближение $x^{(1)} = \left(x_1^{(1)},\, x_2^{(1)},\, \ldots,\, x_n^{(1)}\right)^{\mathrm{T}}$. С ним поступают аналогично и т.д., т.е. действуют по правилу

$$x^{(k+1)} = Bx^{(k)} + c, \quad k = 0, 1, 2, \ldots \tag{10.3}$$

В результате получают последовательность приближенных решений

$$x^{(0)}, \quad x^{(1)}, \quad \ldots, \quad x^{(k)}, \quad \ldots$$

Если эта последовательность имеет предел

$$x = \lim_{k\to\infty} x^{(k)},$$

то этот предел является решением системы (10.2), а следовательно, и системы (10.1). Процесс итераций быстро сходится, если диагональные коэффициенты a_{11}, a_{22}, …, a_{nn} исходной системы значительно преобладают по абсолютной величине над остальными ее коэффициентами, или, что то же самое, коэффициенты a_{ij} системы (10.2) достаточно малы по абсолютной величине.

Более точно этот факт формулируется следующим образом.

Теорема 10.1. *Процесс итераций для системы линейных уравнений (10.2) сходится к единственному ее решению, если какая-либо норма матрицы этой системы меньше* *единицы, в частности, если выполняется хотя бы одно из следующих условий:*

$$\begin{cases} \|B\|_1 = \max\limits_j \sum\limits_i |b_{ij}| < 1, \\ \|B\|_\infty = \max\limits_i \sum\limits_j |a_{ij}| < 1, \\ \|B\|_E = \sqrt{\sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} |b_{ij}|^2} < 1, \\ \|B\| = \sqrt{\max \lambda_{B^{\mathrm{T}}B}} < 1, \end{cases} \tag{10.4}$$

где $\max \lambda_{B^{\mathrm{T}}B}$ *— наибольшее характеристическое число матрицы* $B^{\mathrm{T}}B$.

Погрешность приближений в общем случае можно оценить по формуле

$$\left\|x - x^{(k)}\right\| \leqslant \frac{\left\|x^{(k+1)} - x^{(k)}\right\|}{1 - \|B\|} \tag{10.5}$$

или по формуле

$$\left\|x - x^{(k)}\right\| \leqslant \frac{\|N\|^{k+1}}{1 - \|B\|} \|c\|, \tag{10.6}$$

если за нулевое приближение выбран вектор $x^{(0)} = c$.

Из формулы (10.6) можно найти номер k нужной итерации, который обеспечит необходимую точность приближенного решения. На практике процесс итераций обычно приостанавливают, когда во всех координатах приближенных решений стабилизируется нужное число десятичных знаков после запятой.

Пример 10.1. Методом итераций решить систему

$$\begin{cases} 10x_1 + 2x_2 + x_3 = 10, \\ x_1 + 10x_2 + 2x_3 = 12, \\ x_1 + x_2 + 10x_3 = 8. \end{cases}$$

Решение. Разрешив первое уравнение относительно x_1, второе — относительно x_2, третье — относительно x_3, придем к системе

$$\begin{cases} x_1 = 1 - 0{,}2x_2 - 0{,}1x_3, \\ x_2 = 1{,}2 - 0{,}1x_1 - 0{,}2x_3, \\ x_3 = 0{,}8 - 0{,}1x_1 - 0{,}1x_2. \end{cases} \tag{10.7}$$

Матрица этой системы имеет вид:

$$B = \begin{pmatrix} 0 & -0{,}2 & -0{,}1 \\ -0{,}1 & 0 & -0{,}2 \\ -0{,}1 & -0{,}1 & 0 \end{pmatrix}.$$

Ее норма

$$\|B\|_1 = \max\{0{,}1+0{,}1,\ 0{,}2+0+0{,}1,\ 0{,}1+0{,}2+0\} = 0{,}3 < 1.$$

Поэтому процесс итераций для системы (10.7) будет сходящимся. За нулевое приближение примем

$$x_1^{(0)} = 1, \qquad x_2^{(0)} = 1{,}2, \qquad x_3^{(0)} = 0{,}8.$$

Подставляя эти значения соответственно вместо x_1, x_2, x_3 в правые части уравнений системы (*10.7*), получим:

$$x_1^{(1)} = 0{,}68, \qquad x_2^{(1)} = 0{,}94, \qquad x_3^{(1)} = 0{,}58.$$

С полученным приближением поступим аналогично и т.д.

Результаты вычислений, округленные до трех десятичных знаков после запятой, приведены в табл. 10.1.

Таблица 10.1

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$
0	1	1,2	0,8
1	0,68	0,94	0,58
2	0,754	1,016	0,638
3	0,733	0,997	0,623
4	0,738	1,002	0,627
5	0,737	1,001	0,626
6	0,737	1,001	0,626

Процесс остановлен, так как в значениях каждой неизвестной x_1, x_2, x_3 стабилизировалось по три десятичных знака после запятой. Таким образом, можно принять

$$x \approx x_5 = (0{,}737,\ 1{,}001,\ 0{,}626)^{\mathrm{T}}.$$

Оценим погрешность пятого приближения. По формуле (10.6), вычислив $\|B\|_1 = 0{,}3$, $\|c\|_1 = 1 + 1{,}2 + 0{,}8 = 3$, получим

$$\left\|x - x^{(5)}\right\|_1 \leqslant \frac{(0{,}3)^6}{1 - 0{,}3} \cdot 3 = 0{,}001.$$

Это означает, что в пятом приближении каждая неизвестная имеет не менее чем по два верных десятичных знака после запятой.

10.2. Метод Зейделя

Другим итерационным методом решения систем линейных уравнений является метод Зейделя. Он представляет собой некоторые видоизменения метода итераций. В нем при вычислении $(k+1)$-го приближения неизвестной x_i используются уже вычисленные значения $(k+1)$-го приближения для неизвестных x_1, x_2, …, x_{i-1}. Если для приведенной системы

$$\begin{cases} x_1 = b_{11}x_1 + b_{12}x_2 + \ldots + b_{1n}x_n + c_1, \\ x_2 = b_{21}x_1 + b_{22}x_2 + \ldots + b_{2n}x_n + c_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n = b_{n1}x_1 + b_{n2}x_2 + \ldots + b_{nn}x_n + c_n \end{cases}$$

уже найдено k-е приближение $x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)}$, то ее $(k+1)$-е приближение определяется формулами

$$\begin{cases} x_1^{(k+1)} = b_{11}x_1^{(k)} + b_{12}x_2^{(k)} + \ldots + b_{1n}x_n^{(k)} + c_1, \\ x_2^{(k+1)} = b_{21}x_1^{(k+1)} + b_{22}x_2^{(k)} + \ldots + b_{2n}x_n^{(k)} + c_2, \\ x_3^{(k+1)} = b_{31}x_1^{(k+1)} + b_{32}x_2^{(k+1)} + \ldots + b_{3n}x_n^{(k)} + c_3, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n^{(k+1)} = b_{n1}x_1^{(k+1)} + b_{n2}x_2^{(k+1)} + \ldots \\ \qquad\qquad \ldots + b_{n-1,n}x_{n-1}^{(k+1)} + b_{nn}x_n^{(k)} + c_3, \end{cases} \tag{10.8}$$

$k = 0, 1, 2, \ldots$

Пример 10.2. Методом Зейделя решить систему

$$\begin{cases} 10x_1 - x_2 + x_3 = 10, \\ x_1 + 10x_2 - 3x_3 = 8, \\ x_1 - 2x_2 + 10x_3 = 9. \end{cases}$$

Решение. Разрешив первое уравнение относительно x_1, второе — относительно x_2, третье — относительно x_3, придем к системе

$$\begin{cases} x_1 = 0{,}1x_2 - 0{,}1x_3 + 1, \\ x_2 = -0{,}1x_1 + 0{,}3x_3 + 0{,}8, \\ x_3 = -0{,}1x_1 + 0{,}2x_2 + 0{,}9, \end{cases}$$

удобной для проведения метода Зейделя. За нулевое приближение примем $x_1^{(0)} = 1$, $x_2^{(0)} = 0{,}8$, $x_3^{(0)} = 0{,}9$ и по формулам (10.8) получим:

при $k = 0$

$$x_1^{(1)} = 1 + 0{,}1 \cdot 0{,}8 - 0{,}1 \cdot 0{,}9 = 0{,}99,$$
$$x_2^{(1)} = 0{,}8 - 0{,}1 \cdot 0{,}99 + 0{,}3 \cdot 0{,}9 = 0{,}971,$$
$$x_3^{(1)} = 0{,}9 - 0{,}1 \cdot 0{,}99 + 0{,}2 \cdot 0{,}971 = 0{,}9952;$$

при $k = 1$

$$x_1^{(2)} = 1 + 0{,}1 \cdot 0{,}971 - 0{,}1 \cdot 0{,}9952 = 0{,}9976,$$
$$x_2^{(2)} = 0{,}8 - 0{,}1 \cdot 0{,}9976 + 0{,}3 \cdot 0{,}9952 = 0{,}9989,$$
$$x_3^{(2)} = 0{,}9 - 0{,}1 \cdot 0{,}9976 + 0{,}2 \cdot 0{,}9989 = 1{,}0000$$

и т.д. Результаты вычислений, округленные до четырех десятичных знаков после запятой, приведены в табл. 10.2.

Таблица 10.2

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$
0	1	0,8	0,9
1	0,99	0,971	0,9952
2	0,9976	0,9989	1,0000
3	0,9998	1,0000	1,0000
4	1,0000	1,0000	1,0000
5	1,0000	1,0000	1,0000

Процесс остановлен, так как в значениях каждой неизвестной x_1, x_2, x_3 стабилизировались по четыре десятичных знака после запятой. За искомое решение можно принять $x = (1,\ 1,\ 1)^{\mathrm{T}}$.

10.3. Приведение линейной системы к виду, удобному для итераций

Если в исходной системе $AX = b$ хотя бы некоторые диагональные коэффициенты не преобладают по абсолютной величине над остальными коэффициентами, то, комбинируя уравнения этой системы, стараются заменить ее эквивалентной системой, в которой все диагональные коэффициенты уже будут обладать нужным свойством. Так, в системе

$$\begin{cases} 5x_1 + 9x_2 - 10x_3 = 10, \\ 5x_1 - 8x_2 + 9x_3 = 20, \\ 3x_1 + 2x_2 + 10x_3 = 30 \end{cases}$$

не все диагональные элементы преобладают по абсолютной величине над остальными. Чтобы исправить положение, сложим первое и второе уравнения и результат запишем первым уравнением; из второго уравнения вычтем третье и результат запишем вторым уравнением; третье уравнение оставим без изменения. В результате получим систему

$$\begin{cases} 10x_1 + x_2 - x_3 = 10, \\ 2x_1 - 10x_2 - x_3 = -10, \\ 3x_1 + 2x_2 + 10x_3 = 30, \end{cases}$$

в которой уже все диагональные коэффициенты доминируют по абсолютной величине над остальными. От этой системы обычным путем приходим к системе

$$\begin{cases} x_1 = 3 - 0{,}1x_2 + 0{,}1x_3, \\ x_2 = 1 + 0{,}2x_1 - 0{,}1x_3, \\ x_3 = 3 - 0{,}3x_1 - 0{,}2x_2, \end{cases}$$

уже удобной для итераций.

Такой подход при больших размерах системы не всегда легко приводит к цели. Поэтому на практике обычно идут другими путями. Например, в случае системы $Ax = b$ с симметрической положительно определенной матрицей A сначала от системы $Ax = b$ переходят к эквивалентной ей системе $\tau Ax = \tau b$, затем полагают $\tau A = E - E + \tau A$ и переходят к системе

$$x = (E - \tau A)x + \tau b. \tag{10.9}$$

Остается подобрать множитель τ так, чтобы какая-либо норма матрицы $B = E - \tau A$ была меньше единицы. Если положить

$$\tau = \frac{2}{\lambda_1 + \lambda_2}, \tag{10.10}$$

где λ_1 и λ_2 — наибольшее и наименьшее характеристические числа матрицы A, то спектральная норма матрицы $B = E - \tau A$ системы (10.9), т.е. $\|B\| = \sqrt{\max \lambda_{B^{\mathrm{T}}B}}$, равная

$$\|B\| = \frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2}, \tag{10.11}$$

будет меньше единицы. Этого достаточно для сходимости процесса итераций, применяемого к системе (10.9).

В силу трудоемкости вычисления характеристических чисел матрицы A вместо формулы (10.10) на практике обычно пользуются формулой

$$\tau = \frac{2}{\sigma}, \tag{10.12}$$

где за σ берут число, большее любой из легко вычисляемых норм, например, норм

$$\|A\|_1 = \max_j \sum_i |a_{ij}|, \qquad \|A\|_\infty = \max_i \sum_j |a_{ij}|.$$

В общем случае, когда невырожденная матрица A системы $Ax = b$ не является симметрической и в ней не все диагональные элементы преобладают по абсолютной величине над остальными элементами, сначала от системы $Ax = b$ переходят к системе $A^{\mathrm{T}}Ax = b$, имеющей симметрическую положительно определенную матрицу $A^{\mathrm{T}}A$. К этой системе можно применить описанный выше метод приведения системы к виду, удобному для применения метода итераций.

Пример 10.3. Привести систему

$$\begin{cases} 2x_1 - 2x_2 - x_3 = 4, \\ -2x_1 + 5x_2 + 2x_3 = 8, \\ -x_1 + 2x_2 + 2x_3 = 12 \end{cases}$$

к виду, удобному для решения методом итераций.

Решение. В данной системе не все диагональные коэффициенты преобладают по абсолютной величине над остальными, но матрица системы симметрическая положительно определенная. Поэтому преобразуем систему, используя формулу (10.9). Для матрицы системы

$$A = \begin{pmatrix} 2 & -2 & -1 \\ -2 & 5 & 2 \\ -1 & 2 & 2 \end{pmatrix}$$

найдем ее характеристические числа $\lambda_1 = 7$, $\lambda_2 = \lambda_3 = 1$. По формуле (10.10) определяем:

$$\tau = \frac{2}{7+1} = \frac{1}{4},$$

а по формуле (10.9) составляем систему $x = (E - \frac{1}{4}A)x + \frac{1}{4}b$, эквивалентную данной. В подробной записи эта система имеет вид

$$\begin{cases} x_1 = \frac{1}{2}x_1 + \frac{1}{2}x_2 - \frac{1}{4}x_3 + 1, \\ x_2 = \frac{1}{2}x_1 - \frac{1}{4}x_2 - \frac{1}{2}x_3 + 2, \\ x_3 = \frac{1}{4}x_1 - \frac{1}{2}x_2 + \frac{1}{2}x_3 + 3. \end{cases}$$

В силу (10.11) спектральная норма матрицы B этой системы будет равна

$$\|B\| = \frac{7-1}{7+1} = \frac{3}{4} < 1.$$

Поэтому процесс итераций, примененный к этой системе, будет сходящимся. Если бы не были известны характеристические числа матрицы A, то можно было бы легко подсчитать нормы

$$\|A\|_1 = \|A\|_\infty = \max\{2+2+1,\ 2+5+2,\ 1+2+2\} = 9$$

и для определения τ воспользоваться формулой (10.12), полагая в ней σ, например, равным 10. Тогда получили бы $\tau = 1/5$ и пришли бы к системе

$$x = \left(E - \frac{1}{5}A\right)x + \frac{1}{5}b,$$

т.е.

$$\begin{cases} x_1 = \frac{3}{5}x_1 + \frac{2}{5}x_2 + \frac{1}{5}x_3 + \frac{4}{5}, \\ x_2 = \frac{2}{5}x_1 - \frac{2}{5}x_3 + \frac{8}{5}, \\ x_3 = \frac{1}{5}x_1 - \frac{2}{5}x_2 + \frac{3}{5}x_3 + \frac{12}{5}, \end{cases}$$

также удобной для решения методом итераций, так как

$$\left\|E - \frac{1}{5}A\right\| = \max_S\left|1 - \frac{1}{5}\lambda_S\right| = \max\left\{\left|1 - \frac{1}{5}\cdot 1\right|,\ \left|1 - \frac{1}{5}\cdot 7\right|\right\} = \frac{4}{5} < 1$$

(при подсчете спектральной нормы мы опирались на то, что спектральная норма симметрической матрицы равна наибольшему из модулей ее собственных значений).

Пример 10.4. Привести систему

$$\begin{cases} 6x_1 - 4x_2 - 3x_3 = -11, \\ 5x_1 + x_2 + 5x_3 = 22, \\ 4x_1 + 3x_2 - 4x_3 = -2 \end{cases}$$

к виду, удобному для решения методом итераций.

Решение. Матрица

$$A = \begin{pmatrix} 6 & -4 & -3 \\ 5 & 1 & 5 \\ 4 & 3 & -4 \end{pmatrix}$$

системы не является симметрической и в ней не все диагональные элементы преобладают по абсолютной величине. Поэтому перейдем к эквивалентной системе $A^{\mathrm{T}}Ax = A^{\mathrm{T}}b$, т.е. к системе

$$\begin{cases} 77x_1 - 7x_2 - 9x_3 = 36, \\ -7x_1 + 26x_2 + 5x_3 = 60, \\ -9x_1 + 5x_2 + 50x_3 = 151. \end{cases}$$

В симметрической матрице B модифицированной системы все диагональные элементы преобладают по абсолютной величине. Поэтому можно не прибегать к общему приему подбора параметра τ. Первое уравнение системы разрешим относительно x_1, второе — относительно x_2, третье — относительно x_3. В результате придем к системе, удобной для решения методом итераций:

$$\begin{cases} x_1 = \dfrac{1}{11}x_2 + \dfrac{9}{77}x_3 + \dfrac{36}{77}, \\ x_2 = \dfrac{7}{26}x_1 - \dfrac{5}{26}x_3 + \dfrac{30}{13}, \\ x_3 = \dfrac{9}{50}x_1 - \dfrac{1}{10}x_2 + \dfrac{151}{50}. \end{cases}$$

Для использования общего приема следует вычислить какую-либо норму матрицы $B = A^{\mathrm{T}}A$, например

$$\|B\|_1 = \|B\|_\infty = \max\{77+7+9,\ \ 7+26+5,\ \ 9+5+50\} = 93.$$

Далее, положив $\sigma = 100 > \|B\|_1$, найти

$$\tau = \frac{2}{\sigma} = \frac{1}{50}$$

и, используя найденный параметр, перейти к системе $x = (E - \tau B)x + \tau b$, в координатной форме имеющей вид:

$$\begin{cases} x_1 = -\dfrac{27}{50}x_1 + \dfrac{7}{50}x_2 + \dfrac{9}{50}x_3 + \dfrac{18}{25}, \\ x_2 = \dfrac{7}{50}x_1 + \dfrac{12}{25}x_2 - \dfrac{1}{10}x_3 + \dfrac{6}{5}, \\ x_3 = \dfrac{9}{50}x_1 - \dfrac{1}{10}x_2 + \dfrac{151}{50}. \end{cases}$$

Эту систему можно решать методом итераций.

Обобщением только что описанного подхода является следующий прием (см. [14], [24]). Сначала от системы $Ax = b$ с произвольной невырожденной матрицей, как и выше, переходят к системе $\tau Ax = \tau b$, а затем полагают $\tau A = B - B + \tau A$ и переходят к системе

$$Bx = (B - \tau A)x + \tau b. \tag{10.13}$$

Полученную систему решают методом последовательных приближений в соответствии с формулой

$$B\frac{x^{(k+1)} - x^{(k)}}{\tau} + Ax^{(k)} = b, \quad k = 0, 1, 2, \ldots \tag{10.14}$$

При этом τ выбирают так, чтобы процесс (10.14) итераций был сходящимся.

В случае $B = E$ получается метод простой итерации (см. разд. 10.1). Выбирая $B \neq E$ и подходящий параметр τ, приходят к различным модификациям метода итераций. Выбрав в качестве B какую-либо легко обратимую матрицу, получают общий неявный метод простой итерации. При

$$B = D = \begin{pmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix}, \quad \tau = 1,$$

получают модифицированный метод простой итерации. При $B = D + \tau L$ и таком выборе τ, что наибольшее по модулю собственное значение матрицы $E - \tau(D + \tau L)^{-1}A$ становится минимальным, приходим к ***методу верхней релаксации*** (см. [10], [14], [24], [32]).

Упражнения

10.1. Методом итерации решить системы линейных уравнений:

1) $\begin{cases} 10x_1 + 2x_2 + 3x_3 = 23, \\ 2x_1 + 10x_2 + 3x_3 = 31, \\ 2x_1 + 3x_2 + 10x_3 = 38; \end{cases}$ 2) $\begin{cases} 10x_1 + 3x_2 + x_3 = 35, \\ x_1 + 10x_2 + 2x_3 = 17, \\ x_1 + 2x_2 + 10x_3 = 25; \end{cases}$

3) $\begin{cases} 10x_1 + 2x_2 + x_3 = 25, \\ x_1 + 10x_2 + x_3 = 15, \\ x_1 + 3x_2 + 10x_3 = 35; \end{cases}$ 4) $\begin{cases} 10x_1 + 2x_2 + x_3 = 17, \\ x_1 + 10x_2 + x_3 = 32, \\ x_1 + 3x_2 + 10x_3 = 20; \end{cases}$

5) $\begin{cases} 10x_1 + 3x_2 - x_3 = 35, \\ 2x_1 + 10x_2 - x_3 = 25, \\ x_1 + x_2 + 10x_3 = 15; \end{cases}$ 6) $\begin{cases} 10x_1 + x_2 + 2x_3 = 32, \\ 2x_1 + 10x_2 + x_3 = 18, \\ 2x_1 + x_2 + 10x_3 = 36; \end{cases}$

7) $\begin{cases} 10x_1 + 2x_2 + x_3 = 31, \\ 2x_1 + 10x_2 + x_3 = 39, \\ 3x_1 + 2x_2 - 10x_3 = 38; \end{cases}$ 8) $\begin{cases} 10x_1 + 2x_2 + x_3 = 27, \\ x_1 - 10x_2 + x_3 = 24, \\ x_1 - x_2 + 10x_3 = 5; \end{cases}$

9) $\begin{cases} 10x_1 + 2x_2 + x_3 = 39, \\ x_1 + 10x_2 + x_3 = 28, \\ 3x_1 + 3x_2 - 10x_3 = 65; \end{cases}$ 10) $\begin{cases} 10x_1 - 2x_2 - x_3 = 43, \\ x_1 + 10x_2 + 3x_3 = 34, \\ x_1 + 2x_2 + 10x_3 = 39; \end{cases}$

11) $\begin{cases} 10x_1 + 2x_2 + x_3 = 27, \\ x_1 + 10x_2 + 3x_3 = 35, \\ 2x_1 + 3x_2 + 20x_3 = 33; \end{cases}$ 12) $\begin{cases} 20x_1 + 3x_2 + x_3 = 68, \\ 2x_1 + 20x_2 + 3x_3 = 41, \\ 3x_1 + x_2 + 10x_3 = 60. \end{cases}$

10.2. Методом Зейделя решить системы линейных уравнений из предыдущего упражнения.

10.3. Привести системы линейных уравнений к виду, удобному для проведения метода итераций:

1) $\begin{cases} 4x_1 - 3x_2 - 4x_3 = -1, \\ 7x_1 + x_2 + 4x_3 = 39, \\ 3x_1 + 4x_2 - 5x_3 = 14; \end{cases}$ 2) $\begin{cases} 6x_1 - 4x_2 - 3x_3 = -15, \\ 5x_1 + x_2 + 5x_3 = 38, \\ 4x_1 + 3x_2 - 4x_3 = -3; \end{cases}$

3) $\begin{cases} 3x_1 - 3x_2 + 4x_3 = 14, \\ 4x_1 - x_2 - 4x_3 = 15, \\ 4x_1 + 4x_2 + 3x_3 = 54; \end{cases}$ 4) $\begin{cases} 3x_1 - 4x_2 + 4x_3 = 2, \\ 4x_1 - x_2 - 5x_3 = -17, \\ 5x_1 + 5x_2 + 2x_3 = 43; \end{cases}$

5) $\begin{cases} 2x_1 - 3x_2 + 5x_3 = 4, \\ 3x_1 - 3x_2 - 4x_3 = -1, \\ 4x_1 + 5x_2 + x_3 = 32; \end{cases}$ 6) $\begin{cases} 2x_1 - 4x_2 + 5x_3 = 3, \\ 3x_1 - 3x_2 - 5x_3 = 4, \\ 4x_1 + 5x_2 + 2x_3 = 50. \end{cases}$

Глава 11 О ПРИБЛИЖЕННЫХ МЕТОДАХ ВЫЧИСЛЕНИЯ СОБСТВЕННЫХ ЗНАЧЕНИЙ И СОБСТВЕННЫХ ВЕКТОРОВ

11.1. Метод вращений (метод Якоби)

Для симметрической матрицы A при отыскании собственных значений и собственных векторов в настоящее время наиболее употребительным является ***метод вращений*** (***метод Якоби***). При его обосновании исходят из того, что определение собственных значений и собственных векторов симметрической матрицы A равносильно построению диагональной матрицы Λ и ортогональной матрицы T, связанных соотношением (см. разд. 8.9)

$$\Lambda = T^{-1}AT = T^{\mathrm{T}}AT. \tag{11.1}$$

Диагональные элементы матрицы Λ будут искомыми собственными значениями, а столбцы матрицы T — столбцами координат собственных векторов, соответствующих этим собственным значениям. При приближенном вычислении матриц Λ и T строят последовательность матриц

$$A_0 = A, \quad A_1, \quad A_2, \quad \ldots, \quad A_k, \quad \ldots \to \Lambda,$$

$$T_0 = E, \quad T_1, \quad T_2, \quad \ldots, \quad T_k, \quad \ldots \to T$$

по формулам

$$A_{k+1} = T_{ij}^{\mathrm{T}} A_k T_{ij}, \quad T_{k+1} = T_k T_{ij}, \tag{11.2}$$

где T_{ij} — матрицы простых вращений. Матрица T_k равна произведению всех матриц T_{ij}, примененных при построении матриц A_0, A_1, A_2, $\ldots$, A_k, причем матрицы T_{ij} в этом произведении перемножаются слева направо в том порядке, в каком они применялись. На k-м шаге принимают $\Lambda \approx A_k$, $T \approx T_k$.

Матрицу T_{ij} в формуле (11.2) строят следующим образом. В матрице A_k выбирают наибольший по модулю недиагональный элемент $a_{ij}^{(k)}$ и строят матрицу простого вращения

$$T_{ij} = \begin{array}{cl} \begin{pmatrix} 1 & & & \vdots & & & \vdots & & & \\ & \ddots & & \vdots & & & \vdots & & & \\ & & 1 & \vdots & & & \vdots & & & \\ \ldots & & & \cos\varphi & \ldots & & -\sin\varphi & \ldots & & \\ & & & \vdots & 1 & & \vdots & & & \\ & & & \vdots & & \ddots & \vdots & & & \\ & & & \vdots & & 1 & \vdots & & & \\ \ldots & & & \sin\varphi & \ldots & & \cos\varphi & \ldots & & \\ & & & \vdots & & & \vdots & 1 & & \\ & & & \vdots & & & \vdots & & \ddots & \\ & & & \vdots & & & \vdots & & & 1 \end{pmatrix} & \begin{array}{l} \\ \\ \\ i\text{-я строка} \\ \\ \\ \\ j\text{-я строка} \\ \\ \\ \\ \end{array} \end{array}, \tag{11.3}$$

выбирая угол φ так, чтобы в матрице A_{k+1} обратился в нуль элемент $a_{ij}^{(k+1)}$. Заметим, что если в правой части первого соотношения из (11.2) провести умножение матриц, то элемент $a_{ij}^{(k+1)}$ матрицы A_{k+1} будет иметь вид:

$$a_{ij}^{(k+1)} = (\cos^2\varphi - \sin^2\varphi) a_{ij}^{(k)} + \cos\varphi \sin\varphi (a_{jj}^{(k)} - a_{ii}^{(k)}).$$

Из равенства нулю этого выражения получаем

$$\operatorname{tg} 2\varphi = \frac{2a_{ij}^{(k)}}{a_{ii}^{(k)} - a_{jj}^{(k)}}, \quad |\varphi| \leqslant \frac{\pi}{4}. \tag{11.4}$$

Чтобы записать матрицу T_{ij}, нужно знать $\cos\varphi$ и $\sin\varphi$. Их легко найти по формулам

$$\cos 2\varphi = \frac{1}{\sqrt{1+\operatorname{tg}^2 2\varphi}},$$
$$\cos\varphi = \sqrt{\frac{1+\cos 2\varphi}{2}}, \quad \sin\varphi = \pm\sqrt{\frac{1-\cos 2\varphi}{2}}. \tag{11.5}$$

Чтобы выполнялось условие $|\varphi| \leqslant \frac{\pi}{4}$, знак у корня в формуле для $\sin\varphi$ выбирают тот же, что и у выражения

$$a_{ij}^{(k)}\left(a_{ii}^{(k)} - a_{jj}^{(k)}\right). \tag{11.6}$$

При вычислениях элементов матрицы A_{k+1} удобно пользоваться формулами (см. [14, с. 182])

$$a_{ml}^{(k+1)} = \begin{cases} a_{ml}^{(k)}, & m, l \neq i, j; \\ a_{mi}^{(k)}\cos\varphi + a_{mj}^{(k)}\sin\varphi, & m = i,\ l \neq i, j \text{ или } \\ & l = i,\ m \neq i, j; \\ -a_{mi}^{(k)}\sin\varphi + a_{mj}^{(k)}\cos\varphi, & m = j,\ l \neq i, j \text{ или } \\ & l = j,\ m \neq i, j; \\ a_{ii}^{(k)}\cos^2\varphi + 2a_{ij}^{(k)}\cos\varphi\sin\varphi + & m = l = i; \\ \qquad + a_{jj}^{(k)}\sin^2\varphi, & \\ a_{ii}^{(k)}\sin^2\varphi - 2a_{ij}^{(k)}\cos\varphi\sin\varphi + & m = l = j; \\ \qquad + a_{jj}^{(k)}\cos^2\varphi, & \\ (\cos^2\varphi - \sin^2\varphi)a_{ij}^{(k)} + & m = i,\ l = j \text{ или } \\ \qquad + \cos\varphi\sin\varphi(a_{jj}^{(k)} - a_{ii}^{(k)}), & m = j,\ l = i. \end{cases}$$

которые получаются при сравнении элементов матриц левой и правой частей первого равенства из (11.2) (элементы $a_{ij}^{(k+1)}$, $a_{ji}^{(k+1)}$ при выбранном φ по формулам (11.4) и (11.5) равны нулю).

Метод вращений обладает высокой скоростью сходимости (см. [6]).

Пример 11.1. Методом вращений найти собственные значения и собственные векторы матрицы

$$A=\begin{pmatrix}17&-2&-2\\-2&14&-4\\-2&-4&14\end{pmatrix}.$$

Решение. Положим $A_0=A$ и будем строить при $k=0$ по формулам (11.2) матрицы A_1 и T_1. Так как $\max|a_{ij}^{(0)}|=|a_{ij}^{(0)}|=$ $=4$, то $i=2$, $j=3$ и

$$T_{ij}=T_{23}=\begin{pmatrix}1&0&0\\0&\cos\varphi&-\sin\varphi\\0&\sin\varphi&\cos\varphi\end{pmatrix}.$$

Угол φ определяем по формуле (11.4):

$$\operatorname{tg}2\varphi=\frac{2a_{23}^{(0)}}{a_{22}^{(0)}-a_{33}^{(0)}}=\frac{-8}{14-14}=-\infty.$$

Бесконечное значение тангенса указывает на то, что угол равен $-\frac{\pi}{2}$. Следовательно, $\varphi=-\frac{\pi}{4}$. Отсюда находим $\cos\varphi=$ $=1/\sqrt{2}$, $-\sin\varphi=-1/\sqrt{2}$ и

$$T_{23}=\begin{pmatrix}1&0&0\\0&\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\\0&-\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\end{pmatrix}.$$

По первой формуле (11.2) вычисляем матрицу A_1:

$$A_1=T_{23}^{\mathrm{T}}A_0T_{23}=\begin{pmatrix}1&0&0\\0&\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{2}}\\0&\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\end{pmatrix}\begin{pmatrix}1&0&0\\0&\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\\0&-\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\end{pmatrix}=$$

$$=\begin{pmatrix}17&0&-\frac{4}{\sqrt{2}}\\0&18&0\\-\frac{4}{\sqrt{2}}&0&10\end{pmatrix}.$$

Следовательно, $T_1=T_0T_{23}=T_{23}$.

По тем же формулам (11.2) при $k = 1$ строим матрицы A_2 и T_2. В матрице A_1 имеем: $\max|a_{ij}^{(1)}| = |a_{13}^{(1)}| = 4/\sqrt{2}$. Поэтому $i = 1$, $j = 3$ и

$$T_{ij} = T_{13} = \begin{pmatrix} \cos\varphi & 0 & -\sin\varphi \\ 0 & 1 & 0 \\ \sin\varphi & 0 & \cos\varphi \end{pmatrix},$$

где угол φ определяется соотношением

$$\operatorname{tg} 2\varphi = \frac{2a_{13}^{(1)}}{a_{11}^{(1)} - a_{33}^{(1)}} = -\frac{8}{7\sqrt{2}},$$

откуда

$$\cos 2\varphi = \frac{1}{\sqrt{1 + \operatorname{tg}^2 2\varphi}} = \frac{7}{9}.$$

Вычислив

$$\cos\varphi = \sqrt{\frac{1 + \cos 2\varphi}{2}} = \frac{2\sqrt{2}}{3}, \qquad \sin\varphi = -\sqrt{\frac{1 - \cos 2\varphi}{2}} = -\frac{1}{3}$$

(поскольку $a_{13}^{(1)}(a_{11}^{(1)} - a_{33}^{(1)}) = -\frac{1}{\sqrt{2}}(17 - 10) < 0$, то в формуле для $\sin\varphi$ выбран знак „минус“), записываем матрицу T_{13}:

$$T_{13} = \begin{pmatrix} \frac{2\sqrt{2}}{3} & 0 & \frac{1}{3} \\ 0 & 1 & 0 \\ -\frac{1}{3} & 0 & \frac{2\sqrt{2}}{3} \end{pmatrix}.$$

Теперь по первой формуле (11.2) вычисляем матрицу A_2:

$$A_2 = T_{13}^{\mathrm{T}} A_1 T_{13} = \begin{pmatrix} \frac{2\sqrt{2}}{3} & 0 & \frac{1}{3} \\ 0 & 1 & 0 \\ -\frac{1}{3} & 0 & \frac{2\sqrt{2}}{3} \end{pmatrix} \begin{pmatrix} 17 & 0 & -\frac{4}{\sqrt{2}} \\ 0 & 18 & 0 \\ -\frac{4}{\sqrt{2}} & 0 & 10 \end{pmatrix} \times$$

$$\times \begin{pmatrix} \frac{2\sqrt{2}}{3} & 0 & \frac{1}{3} \\ 0 & 1 & 0 \\ -\frac{1}{3} & 0 & \frac{2\sqrt{2}}{3} \end{pmatrix} = \begin{pmatrix} 18 & 0 & 0 \\ 0 & 18 & 0 \\ 0 & 0 & 9 \end{pmatrix} = \Lambda,$$

а затем по второй формуле (11.2) — матрицу T_2:

$$T_2 = T = T_{23}T_{13} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{2\sqrt{2}}{3} & 0 & \frac{1}{3} \\ 0 & 1 & 0 \\ -\frac{1}{3} & 0 & \frac{2\sqrt{2}}{3} \end{pmatrix} =$$

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

По найденным матрицам Λ и T выписываем собственные значения $\lambda_1 = \lambda_2 = 18$, $\lambda_3 = 9$ и соответствующие им собственные векторы

$$e_1' = \frac{1}{3\sqrt{2}} \begin{pmatrix} 4 \\ -1 \\ -1 \end{pmatrix}, \quad e_2' = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \quad e_3' = \frac{1}{3} \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}.$$

Зная матрицы T и Λ, можно легко записать для матрицы A каноническое разложение $A = T\Lambda T^{-1} = T\Lambda T^{\text{т}}$, которое определяется ортогональной трансформирующей матрицей.

Метод Якоби применяют в случае произвольной действительной, а также в случае комплексной матрицы (см. [14], [29]). Если A — комплексная матрица, то вместо матрицы (11.3) простого вращения применяют унитарную матрицу

$$T_{ij} = \begin{array}{c} \begin{pmatrix} 1 & & & \vdots & & & \vdots & & & \\ & \ddots & & \vdots & & & \vdots & & & \\ & & 1 & \vdots & & & \vdots & & & \\ \cdots & & & \cos\varphi & \cdots & & -e^{i\psi}\sin\varphi & \cdots & & \\ & & & \vdots & 1 & & \vdots & & & \\ & & & \vdots & & \ddots & \vdots & & & \\ & & & \vdots & & 1 & \vdots & & & \\ \cdots & & & e^{-i\psi}\sin\varphi & \cdots & & \cos\varphi & \cdots & & \\ & & & \vdots & & & \vdots & 1 & & \\ & & & \vdots & & & \vdots & & \ddots & \\ & & & \vdots & & & \vdots & & & 1 \end{pmatrix} \end{array} \begin{array}{l} \\ \\ \\ i\text{-я строка} \\ \\ \\ \\ j\text{-я строка} \\ \\ \\ \\ \end{array}$$

Напомним, что по формуле Эйлера

$$e^{i\varphi} = \cos\varphi + i\sin\varphi.$$

Для эрмитовой матрицы A метод Якоби остается таким же, как и в случае действительной симметрической матрицы, а именно, для максимального уменьшения на k-м шаге (см. [14]) суммы квадратов модулей недиагональных элементов матрицы A_k номера i и j выбирают так, чтобы элемент $a_{ij}^{(k)}$ был наибольшим по модулю недиагональным элементом матрицы A_k, а углы ψ и φ находят из соотношений

$$\psi = \arg a_{ij}^{(k)}, \quad \operatorname{tg} 2\varphi = \frac{2|a_{ij}^{(k)}|}{a_{ii}^{(k)} - a_{jj}^{(k)}}, \quad |\varphi| \leqslant \frac{\pi}{4}.$$

Возможны и различные модификации этого метода. Например, применяют метод Якоби с циклическим перебором недиагональных элементов (см. [29]).

Пример 11.2. Методом Якоби найти собственные значения и собственные векторы эрмитовой матрицы

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

Решение. Положим $A_0 = A$ и построим матрицы T_{ij} и A_1. Для этого замечаем, что $\max|a_{ij}| = |a_{12}|$. Поэтому принимаем $i = 1$, $j = 2$. Поскольку

$$a_{ij} = a_{12} = -i = \cos\left(-\frac{\pi}{2}\right) + i\sin\left(-\frac{\pi}{2}\right),$$

то $\psi = \arg a_{12} = -\pi/2$. Далее находим

$$\operatorname{tg} 2\varphi = \frac{2|a_{12}|}{a_{11} - a_{22}} = \frac{2 \cdot 1}{3 - 3} = \infty.$$

Следовательно, $2\varphi=\pi/2$ и $\varphi=\pi/4$. Можно записать матрицу

$$T_{12}=\begin{pmatrix}\cos\frac{\pi}{4} & -e^{-i\pi/2}\sin\frac{\pi}{4} & 0\\ e^{i\pi/2}\sin\frac{\pi}{4} & \cos\frac{\pi}{4} & 0\\ 0 & 0 & 1\end{pmatrix}=\begin{pmatrix}\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2}i & 0\\ \frac{\sqrt{2}}{2}i & \frac{\sqrt{2}}{2} & 0\\ 0 & 0 & 1\end{pmatrix}.$$

Далее вычисляем матрицу

$$A_1=T_{12}^{-1}AT_{12}=T_{12}^{*}AT_{12}=$$

$$=\begin{pmatrix}\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2}i & 0\\ -\frac{\sqrt{2}}{2}i & \frac{\sqrt{2}}{2} & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}3 & -i & 0\\ i & 3 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2}i & 0\\ \frac{\sqrt{2}}{2}i & \frac{\sqrt{2}}{2} & 0\\ 0 & 0 & 1\end{pmatrix}=$$

$$=\begin{pmatrix}4 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 1\end{pmatrix}=\Lambda.$$

Матрица $A_1=\Lambda$ уже диагональная, а матрица T оказалась равной матрице T_{12}. По найденным матрицам Λ и T запишем собственные значения $\lambda_1=4$, $\lambda_2=2$, $\lambda_3=1$ матрицы A и соответствующие им собственные векторы

$$e_1=\frac{\sqrt{2}}{2}\begin{pmatrix}1\\ i\\ 0\end{pmatrix},\quad e_2=\frac{\sqrt{2}}{2}\begin{pmatrix}i\\ 1\\ 0\end{pmatrix},\quad e_3=\begin{pmatrix}0\\ 0\\ 1\end{pmatrix}.$$

В рассмотренных примерах процесс закончился соответственно на втором и первом шагах. В общем же случае такой процесс бесконечен, так как на каждом шаге вместо нулевых недиагональных элементов могут возникать ненулевые. Однако при выполнении каждого шага уменьшается с большой скоростью сумма квадратов недиагональных элементов. Процесс прекращают, когда все недиагональные элементы становятся пренебрежимо малыми по абсолютной величине.

Метод вращений хорошо реализуется на ЭВМ, причем существует большое количество вычислительных схем для его реализации. При любой из таких схем метод вращений представляет собой исключительно компактную процедуру, которая обеспечивает высокую точность вычислений. Точность нахождения собственных значений и собственных векторов этим методом оказывается сравнимой с длиной машинного слова. Стандартные программы к методу вращений содержатся во многих пакетах математического обеспечения.

Замечание. Другим надежным и быстро сходящимся методом вычисления собственных значений и собственных векторов положительно определенной симметрической (эрмитовой) матрицы может быть (см. разд. 8.18) итерационный процесс получения сингулярного разложения $A = Q\Sigma P^*$ такой матрицы косвенным путем с помощью отражений и вращений с последующим выписыванием из этого разложения собственных значений — диагональных элементов матрицы Σ и соответствующих им собственных векторов — столбцов матрицы P (см. [18], [29], [37]). В случае произвольной симметрической (эрмитовой) матрицы A этот процесс дает модули собственных значений матрицы A.

11.2. QR-алгоритм

QR-алгоритм — это наиболее употребительный в настоящее время итерационный метод отыскания характеристических чисел действительных и комплексных матриц. Стандартные программы к этому методу содержатся во многих пакетах математического обеспечения. QR-алгоритм состоит в следующем.

Для данной матрицы $A = A_0$ строят QR-разложение $A = Q_0R_0$ (см. разд. 8.17). В полученном разложении меняют местами множители. Это приводит к матрице $A_1 = R_0Q_0$, с которой поступают аналогично, и т.д. В результате получают последовательность матриц A_0, A_1, A_2, ..., A_k, ... Эти матрицы подобны, так как $A_1 = Q_0^{-1}A_0Q_0$, $A_2 = Q_1^{-1}A_1Q_1$ и т.д.

При довольно общих предположениях построенная последовательность матриц сходится к треугольной матрице. Диагональные элементы треугольной матрицы являются ее характеристическими числами и характеристическими числами всех подобных матриц, в частности матрицы $A_0 = A$.

Для ускорения QR-алгоритма желательно, чтобы на каждом его шаге быстро строилось QR-разложение очередной матрицы A_k. Для этого следует исходную матрицу с помощью ортогональных (унитарных) преобразований привести к подобной матрице $A_0 = U^{-1}AU$ более простого вида, например, почти треугольного вида или трехдиагональной формы. О приведении матрицы к подобной матрице почти треугольного вида см. разд. 8.10 и [3, с. 182–184]. QR-алгоритм практически всегда применяется к почти треугольным матрицам, а в симметрическом случае — к трехдиагональным.

Другим важным приемом ускорения QR-алгоритма является использование сдвигов σ_k для повышения скорости убывания по абсолютной величине элементов, лежащих ниже главной диагонали. С этой целью на $(k+1)$-м шаге QR-разложение строят не для матрицы A_k, а для матрицы $A_k - \sigma_k E$. Если $A_k - \sigma_k E = Q_k R_k$ — построенное QR-разложение, то полагают $A_{k+1} = R_k Q_k + \sigma_k E$. Такой переход от A_k к A_{k+1} не меняет характеристических чисел, поскольку матрицы A_k и A_{k+1} подобны. Действительно,

$$Q_k^{-1} A_k Q_k = Q_k^{-1}(Q_k R_k + \sigma_k E) Q_k = R_k Q_k + \sigma_k E = A_{k+1}.$$

На практике обычно в качестве сдвига σ_k принимают элемент матрицы A_k, стоящий в ее нижнем правом углу. При таком выборе сдвига обеспечивается высокая скорость сходимости к характеристическому числу λ_n.

Пусть в результате некоторого числа шагов QR-алгоритма пришли к матрице вида

$$\begin{pmatrix} * & \ldots & * & * \\ \cdot & \cdot & \cdot & \cdot \\ * & \ldots & * & * \\ 0 & \ldots & \varepsilon & \lambda_n' \end{pmatrix} \tag{11.7}$$

с пренебрежимо малым по абсолютной величине числом ε. Тогда принимают $\lambda_n = \lambda'_n$ и переходят к вычислению характеристического числа λ_{n-1}. При этом фактически работают уже с матрицей $(n-1)$-го порядка, расположенной в матрице (11.7) в ее левом верхнем углу. При этом за сдвиг σ_k теперь каждый раз принимают элемент, стоящий в правом нижнем углу матрицы $(n-1)$-го порядка, с которой работают. После того как вычислено λ_{n-1}, аналогично поступают при вычислении λ_{n-2} и т.д. Если матрица действительная, то QR-алгоритм можно реализовать в вещественной арифметике (см. [34, с. 142]).

Пример 11.3. С помощью QR-алгоритма найти характеристические числа матрицы

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

Решение. Чтобы показать, как проводятся циклы QR-алгоритма без сдвига и со сдвигом, начнем с того, что первый цикл проведем без сдвига. Поэтому будем строить QR-разложение матрицы $A_0 = A$. Сначала эту матрицу вращениями приведем к треугольному виду. Начнем с обнуления элемента $a_{32} = 2$. Для этого составим матрицу вращения

$$T_{32} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix}$$

и найдем $\cos\varphi$ и $\sin\varphi$ исходя из равенства нулю элемента матрицы $T_{32}(A_1 - E)$ в третьей строке втором столбце, т.е. из равенства

$$a_{22}\sin\varphi + a_{32}\cos\varphi = 2\sin\varphi + 2\cos\varphi = 0.$$

Отсюда получаем: $\operatorname{tg}\varphi = -1$. Поэтому

$$\cos\varphi = \frac{1}{\sqrt{1+\operatorname{tg}^2\varphi}} = \frac{1}{\sqrt{2}}, \quad \sin\varphi = \frac{\operatorname{tg}\varphi}{\sqrt{1+\operatorname{tg}^2\varphi}} = -\frac{1}{\sqrt{2}}.$$

Следовательно,

$$T_{32} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

и

$$T_{32}A_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{4}{\sqrt{2}} & \frac{4}{\sqrt{2}} \\ 0 & 0 & \frac{2}{\sqrt{2}} \end{pmatrix} = R_0.$$

Из этого равенства получается QR-разложение

$$A_0 = T_{32}^{-1}R_0 = T_{32}^{\mathrm{T}}R_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{4}{\sqrt{2}} & \frac{4}{\sqrt{2}} \\ 0 & 0 & \frac{2}{\sqrt{2}} \end{pmatrix} = Q_0R_0.$$

В заключение цикла строим матрицу

$$A_1 = R_0Q_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{4}{\sqrt{2}} & \frac{4}{\sqrt{2}} \\ 0 & 0 & \frac{2}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 1 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & 4 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

Следующий цикл проведем со сдвигом $\sigma = a_{33}^{(1)} = 1$. Поэтому будем строить QR-разложение матрицы

$$A_1 - \sigma E = A_1 - E = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & 3 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Сначала эту матрицу вращениями приведем к треугольному виду. Здесь следует обнулить лишь элемент $A_{32}^{(1)} = 1$. Для этого составим матрицу вращения

$$T_{32} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

и найдем $\cos\varphi$ и $\sin\varphi$ из равенства нулю элемента матрицы $T_{32}(A_1 - E)$ в третьей строке втором столбце, т.е. из равенства

$$a_{22}\sin\varphi + a_{32}\cos\varphi = 3\sin\varphi + \cos\varphi = 0.$$

Отсюда получаем: $\operatorname{tg}\varphi = -1/3$ и

$$\cos\varphi = \frac{1}{\sqrt{1+\operatorname{tg}^2\varphi}} = \frac{3}{\sqrt{10}}, \qquad \sin\varphi = \frac{\operatorname{tg}\varphi}{\sqrt{1+\operatorname{tg}^2\varphi}} = -\frac{1}{\sqrt{10}}.$$

Следовательно,

$$T_{32} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ 0 & -\frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \end{pmatrix}.$$

Из равенства

$$T_{32}(A_1 - E) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ 0 & -\frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \end{pmatrix} \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & 3 & 0 \\ 0 & 1 & 0 \end{pmatrix} =$$

$$= \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \sqrt{10} & 0 \\ 0 & 0 & 0 \end{pmatrix} = R_1$$

получаем QR-разложение:

$$A_1 - E = T_{32}^{-1}R_1 = T_{32}^{\mathrm{T}}R_1 = Q_1R_1 =$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{3}{\sqrt{10}} & -\frac{1}{\sqrt{10}} \\ 0 & \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \end{pmatrix} \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \sqrt{10} & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

В заключение цикла строим матрицу

$$A_2 = R_1Q_1 + E = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \sqrt{10} & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{3}{\sqrt{10}} & -\frac{1}{\sqrt{10}} \\ 0 & \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \end{pmatrix} + E =$$

$$= \begin{pmatrix} 0 & \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ 0 & 3 & -1 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ 0 & 4 & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Эта матрица уже треугольная с характеристическими числами $\lambda_1 = 1$, $\lambda_2 = 4$, $\lambda_3 = 1$. Эти же числа являются характеристическими числами матрицы A.

11.3. Степенной метод

При отыскании наибольшего по абсолютной величине собственного значения λ_1 матрицы A и соответствующего ему собственного вектора $X_1 = (x_1, x_2, \ldots, x_n)^{\text{т}}$ можно пользоваться следующим итерационным методом, называемым ***степенным***.

В качестве нулевого приближения к искомому вектору x_λ берут произвольный вектор $X_1^{(0)} = \left(x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)}\right)^{\text{т}}$ и последовательно строят следующие приближения:

$$X_1^{(1)} = AX_1^{(0)}, \quad X_1^{(2)} = A^2X_1^{(0)}, \quad \ldots, \quad X_1^{(k)} = A^kX_1^{(0)}, \quad \ldots \qquad (11.8)$$

Если эта последовательность векторов сходится при $k \to \infty$ к некоторому вектору, то этот вектор будет искомым собственным вектором, а для соответствующих координат векторов последовательности (11.8) будут выполняться соотношения

$$\lim_{k\to\infty} \frac{x_i^{(k)}}{x_i^{(k-1)}} = \lambda_1 \qquad (11.9)$$

для любого индекса $i = 1, 2, \ldots, n$, при котором отношение, стоящее под знаком предела, имеет смысл.

За вектор X_1 приближенно можно принять вектор $X_1^{(k)} = A^k X_1^{(0)}$ при достаточно большом значении k или ему пропорциональный вектор, а за собственное значение λ_1 — любое из имеющих смысл отношений

$$\lambda_1^{(k)} = \frac{x_i^{(k)}}{x_i^{(k-1)}}, \quad i = 1, 2, \ldots, n. \tag{11.10}$$

Итерационный процесс останавливают, если в $\lambda_1^{(k)}$ стабилизируется достаточное число десятичных знаков после запятой.

Если процесс окажется несходящимся (последнее легко заметить по сильно „прыгающим“ значениям соотношений (11.10) при изменении k), то следует изменить начальный вектор $X_1^{(0)}$ и процесс повторить.

Чтобы избежать чрезмерного роста по абсолютной величине координат векторов $A^k X_1^{(0)}$, $k = 1, 2, \ldots$, иногда целесообразно все их координаты умножать на какие-либо числа α_k. Например, удобным является деление всех координат векторов $A^k X_1^{(0)}$ на модуль их первых координат или нормирование этих векторов. При этом вместо последовательности векторов (11.8) получают последовательность векторов

$$Y_1^{(1)} = \alpha_1 A X_1^{(0)}, \quad Y_1^{(2)} = \alpha_2 A^2 X_1^{(0)}, \quad \ldots, \quad Y_1^{(k)} = \alpha_k A^k X_1^{(0)}, \quad \ldots$$

Для получения λ_1 в этом случае следует брать отношения координат векторов $AY_1^{(k)}$ и $Y_1^{(k)}$, а за вектор X_1 — вектор $Y_1^{(k)}$. Более подробные рекомендации об этом можно найти в [3], [32]. Там же можно познакомиться с различными модификациями степенного метода и с обратным степенным методом.

П р и м е р 11.4. Для матрицы

$$A = \begin{pmatrix} 1 & -3 & 4 \\ 4 & -7 & 8 \\ 6 & -7 & 7 \end{pmatrix}$$

степенным методом найти наибольшее по абсолютной величине собственное значение λ_1 и соответствующий ему собственный вектор X_1.

Решение. За нулевое приближение вектора X_1 примем вектор $X_1^{(0)} = (1,\ 1,\ 1)^{\text{т}}$. Далее построим векторы $X_1^{(k)} = A^k X^{(0)}$ для $k = 1, \ldots, 10$. Результаты приведены в табл. 11.1.

Таблица 11.1

$X_1^{(0)}$	$A^1X_1^{(0)}$	$A^2X_1^{(0)}$	$A^3X_1^{(0)}$	$A^4X_1^{(0)}$	$A^5X_1^{(0)}$
1	2	11	24	85	238
1	5	21	49	169	477
1	6	19	52	165	482

$A^6X_1^{(0)}$	$A^7X_1^{(0)}$	$A^8X_1^{(0)}$	$A^9X_1^{(0)}$	$A^{10}X_1^{(0)}$
735	2180	6569	19674	59059
1469	4361	13137	39349	118117
1463	4368	13139	39358	118107

Теперь вычислим λ_1 по формуле (11.10) при $k = 9,\ 10$ и $i = 3$:

$$\lambda_1^{(9)} = 2{,}997791, \qquad \lambda_1^{(10)} = 3{,}000838.$$

Мы видим, что $\lambda_1^{(9)}$ и $\lambda_1^{(10)}$ мало отличаются одно от другого. Поэтому, округляя до второго десятичного знака, можно принять $\lambda_1 = 3{,}00$, а в качестве собственного вектора взять

$$X_1 \approx X_1^{(10)} = A^{10} X_1^{(0)} = (59059,\ 118117,\ 118107)^{\text{т}}.$$

Естественно вектор X_1 заменить ему пропорциональным, например вектором

$$Y_1 = \frac{1}{59059} A^{10} X_1^{(0)} = (1,\ 2,\ 2)^{\text{т}}$$

или ортом

$$\frac{X_1}{|X_1|} = (0{,}333333,\ 0{,}666666,\ 0{,}666666)^{\text{т}}.$$

При переходе, например, от $A^9 X_1^{(0)}$ к $A^{10} X_1^{(0)}$, чтобы избежать больших по абсолютной величине координат, вместо вектора $A^9X_1^{(0)}$ можно было взять вектор

$$Y_1^{(9)} = \frac{1}{19674}A^9X_1^{(0)} = (1,\ 2{,}0000508,\ 2{,}0005082)^{\text{т}}.$$

Тогда вместо $A^{10}X_1^{(0)}$ мы имели бы:

$$AY_1^{(9)} = (3{,}0018804,\ 6{,}0037100,\ 6{,}0032018)^{\text{т}}$$

и

$$\lambda_1^{(10)} = \frac{3{,}0018804}{1} = 3{,}0018804.$$

Степенной метод можно приспособить к задаче вычисления других собственных значений и собственных векторов матрицы. Так, для вычисления второго по величине собственного значения λ_2 и соответствующего ему собственного вектора X_2 можно комбинировать степенной метод с ***методом исчерпывания*** (см. [3], [10]), т.е. поступать следующим образом: найти степенным методом или из системы $(A^{\text{т}} - \lambda_1 E)U = 0$ собственный вектор U_1 матрицы $A^{\text{т}}$, принадлежащий собственному значению λ_1, и положить

$$X_1' = \frac{U_1}{(X_1, U_1)}.$$

Так построенный собственный вектор X_1' матрицы $A^{\text{т}}$ удовлетворяет условию $(X_1, X_1') = X_1^{\text{т}}X_1' = 1$. Если матрица A симметрическая, то за вектор X_1' принимают нормированный по длине вектор X_1. Далее следует составить матрицу

$$A_1 = A - \lambda_1 X_1 (X_1')^{\text{т}}.$$

Очевидно, что процесс конструирования матрицы A_1 легко программируется и что полученную при этом программу можно включить в качестве подпрограммы в общую программу по отысканию собственных значений и собственных векторов матрицы A.

Все собственные векторы $X_1, X_2, \ldots, X_n$ матрицы A являются также собственными векторами матрицы A_1 и наоборот. При этом соответствующие собственные значения сохраняются, за исключением λ_1, которому в матрице A_1 соответствует нулевое собственное значение.

Наибольшее значение μ_1 и ему принадлежащий собственный вектор V_1 матрицы A_1 (их опять можно найти степенным методом) являются соответственно собственным значением λ_2 и ему принадлежащим собственным вектором X_2 матрицы A. С матрицей A_1 можно поступать аналогично и т.д. В итоге можно найти все собственные значения и собственные векторы матрицы A. Поясним такой подход на примере.

Пример 11.5. Для матрицы

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

найти собственные значения и собственные векторы.

Решение. Найдем степенным методом для матрицы A собственное значение λ_1 и ему принадлежащий собственный вектор X_1. Для этого за нулевое приближение вектора X_1 примем, например, вектор $X_1^{(0)} = (1,\ 1,\ 1)^{\mathrm{T}}$ и построим векторы $A^k X_1^{(0)}$ для $k = 1, 2, 3, 4$. Результаты вычислений приведены в табл. 11.2.

Таблица 11.2

$X_1^{(0)}$	$A^1X_1^{(0)}$	$A^2X_1^{(0)}$	$A^3X_1^{(0)}$	$A^4X_1^{(0)}$
1	5	25	125	625
1	5	25	125	625
1	5	25	125	625

Теперь по формуле (11.10) при $k = 4$ находим

$$\lambda_1 \approx \lambda_1^{(4)} = \frac{625}{125} = 5,$$

за вектор X_1 примем вектор

$$X_1 = \frac{1}{625} A^4 X_1^{(0)} = (1,\ 1,\ 1)^{\mathrm{T}}.$$

Чтобы найти собственное значение λ_2 и собственный вектор X_2 матрицы A, сначала из системы $(A^{\mathrm{T}} - \lambda_1 E)U = 0$ или степенным методом найдем собственный вектор $U_1 = (1,\ -1,\ 1)^{\mathrm{T}}$ матрицы A^{T}, принадлежащий собственному значению $\lambda_1 = 5$, и положим

$$X_1' = \frac{U_1}{(X_1, U_1)} = (1,\ -1,\ 1)^{\mathrm{T}}.$$

Затем сконструируем матрицу

$$A_1 = A - \lambda_1 X_1 (X_1')^{\mathrm{T}} = A - 5\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}(1,\ -1,\ 1) = \begin{pmatrix} 0 & 1 & -1 \\ -3 & 6 & -3 \\ -3 & 5 & -2 \end{pmatrix}.$$

Для матрицы A_1 степенным методом найдем собственное значение μ_1 и ему принадлежащий собственный вектор V_1. Для этого за нулевое приближение этого вектора примем, например, вектор $V_1^{(0)} = (0,\ 1,\ 1)^{\mathrm{T}}$ (брать $V_1^{(0)} = (1,\ 1,\ 1)^{\mathrm{T}}$ здесь нельзя, так как иначе получилось бы $A_1 V_1^{(0)} = (0,\ 0,\ 0)^{\mathrm{T}}$, что невозможно для дальнейшего проведения степенного метода) и построим векторы $A_1^k V_1^{(0)}$, $k = 1, 2, 3, 4$. Результаты вычислений приведены в табл. 11.3.

Таблица 11.3

$V_1^{(0)}$	$A_1^1 V_1^{(0)}$	$A_1^2 V_1^{(0)}$	$A_1^3 V_1^{(0)}$	$A_1^4 V_1^{(0)}$
0	0	0	0	0
1	3	9	27	81
1	3	9	27	81

По формуле (11.10) при $k = 4$ и $i = 2, 3$ находим

$$\mu_1 = \frac{81}{27} = 3.$$

За вектор V_1 примем $V_1^{(4)} = A_1^4 V_1^{(0)} = (0,\ 81,\ 81)^{\text{т}}$. Поэтому для матрицы A имеем $\lambda_2 = 3$. За вектор X_2 можно принять вектор

$$X_2 = \frac{1}{81} V_1 = (0,\ 1,\ 1)^{\text{т}}.$$

Собственное значение $\lambda_3 = 1$ матрицы A найдем из соотношения

$$\lambda_1 + \lambda_2 + \lambda_3 = a_{11} + a_{22} + a_{33},$$

а собственный вектор $X_3 = (1,\ 0,\ -1)^{\text{т}}$ определим, решив систему $(A - \lambda_3 E)X = 0$.

11.4. Метод скалярных произведений

При отыскании наибольшего по абсолютной величине собственного значения λ_1 матрицы A можно воспользоваться также ***методом скалярных произведений***, который состоит в следующем. Берут произвольный ненулевой вектор $X_1^{(0)}$, строят две последовательности векторов

$$AX_1^{(0)},\quad A^2X_1^{(0)},\quad \ldots,\quad A^kX_1^{(0)},\quad \ldots,$$

$$A^{\text{т}}X_1^{(0)},\quad (A^{\text{т}})^2X_1^{(0)},\quad \ldots,\quad (A^{\text{т}})^kX_1^{(0)},\quad \ldots$$

и вычисляют соответствующие приближения к λ_1 по формуле

$$\lambda_1^{(k)} = \frac{\left(A^kX_1^{(0)},\ (A^{\text{т}})^kX_1^{(0)}\right)}{\left(A^{k-1}X_1^{(0)},\ (A^{\text{т}})^kX_1^{(0)}\right)}. \tag{11.11}$$

Процесс останавливают, когда в $\lambda_1^{(k)}$ стабилизируется достаточное число знаков после запятой, и принимают $\lambda_1 \approx \lambda_1^{(k)}$.

Как и в случае степенного метода, для избежания чрезмерного роста по абсолютной величине координат векторов $A^kX_1^{(0)}$ целесообразно все их координаты умножать на какие-либо числа α_k, например на

$$\alpha_k = \frac{1}{|X_1^{(k)}|} \quad \text{или} \quad \alpha_k = \frac{1}{|A^kX^{(0)}|}.$$

Аналогично следует поступать и с векторами $(A^{\text{т}})^k X_1^{(0)}$. В результате получатся две последовательности векторов

$$Y^{(1)}=\alpha_1 A x^{(0)},\ Y^{(2)}=\alpha_2 A^2 x^{(0)},\ \ldots,\ Y^{(k)}=\alpha_k A^k x^{(0)},\ \ldots,$$

$$Z^{(1)}=\beta_1 (A^{\text{т}}) x^{(0)},\ Z^{(2)}=\beta_2 (A^{\text{т}})^2 x^{(0)},\ \ldots,\ Z^{(k)}=\beta_k (A^{\text{т}})^k x^{(0)},\ \ldots,$$

и $\lambda_1^{(k)}$ находят по формуле

$$\lambda_1^{(k)} = \frac{\left(AY^{(k)}, A^{\text{т}} Z^{(k)}\right)}{\left(Y^{(k)}, A^{\text{т}} Z^{(k)}\right)}.$$

Собственный вектор X_1 находят, как и в предыдущем случае, т.е. принимают

$$X_1 \approx A^k X_1^{(0)} \quad \text{или} \quad X_1 \approx \alpha_k A^k X_1^{(0)} = Y^{(k)}.$$

Заметим, что этот метод значительно упрощается, если матрица A симметрическая, т.е. если $A = A^{\text{т}}$.

Пример 11.6. Методом скалярных произведений найти наибольшее по абсолютной величине собственное значение λ_1 и соответствующий ему собственный вектор матрицы

$$A = \begin{pmatrix} 1 & -3 & 4 \\ 4 & -7 & 8 \\ 6 & -7 & 7 \end{pmatrix}.$$

Решение. Примем за начальный вектор $X_1^{(0)} = (1,\ 1,\ 1)^{\text{т}}$. Соответствующая последовательность векторов $X_1^{(k)} = A^k X_1^{(0)}$, $k = 1, \ldots, 10$, уже вычислена в предыдущем примере. Вычислим векторы $(A^{\text{т}})^k X_1^{(0)}$, $k = 1, 2, 3, 4$. Результаты приведены в табл. 11.4.

Таблица 11.4

$X_1^{(0)}$	$(A^{\text{т}})^1 X_1^{(0)}$	$(A^{\text{т}})^2 X_1^{(0)}$	$(A^{\text{т}})^3 X_1^{(0)}$	$(A^{\text{т}})^4 X_1^{(0)}$
1	11	57	115	433
1	−17	−47	−129	−415
1	19	41	139	401

Вычислим λ_1 по формуле (11.11) при $k=4$:

$$\lambda_1 \approx \lambda_1^{(4)} = \frac{\left(A^4 X_1^{(0)},\, (A^{\text{т}})^4 X_1^{(0)}\right)}{\left(A^3 X_1^{(0)},\, (A^{\text{т}})^4 X_1^{(0)}\right)} = \frac{(85, 169, 165)\begin{pmatrix} 433 \\ -415 \\ 401 \end{pmatrix}}{(24, 49, 52)\begin{pmatrix} 433 \\ -415 \\ 401 \end{pmatrix}} = 3{,}0002,$$

X_1 — по формуле $Xx_1 \approx A^k X_1^{(0)}$ при $k=4$:

$$X_1 \approx A^4 X_1^{(0)} = (85,\ 169,\ 165)^{\text{т}}.$$

Мы видим, что формула (11.11) уже при $k=4$ дала результат, близкий к тому, какой в предыдущем примере был получен при $k=10$. Поэтому метод скалярных произведений рассматривают как ускорение степенного метода (см. [32]).

Метод скалярных произведений, как и степенной метод, можно применять в сочетании с методом исчерпывания для отыскания других собственных значений и собственных векторов матрицы.

Упражнения

11.1. Методом вращений найти собственные значения и собственные векторы следующих симметрических матриц:

1) $\begin{pmatrix} 2 & -2 & 0 \\ -2 & 1 & -2 \\ 0 & -2 & 0 \end{pmatrix}$; 2) $\begin{pmatrix} 1 & -2 & 0 \\ -2 & 2 & -2 \\ 0 & -2 & 3 \end{pmatrix}$; 3) $\begin{pmatrix} 3 & 2 & 0 \\ 2 & 4 & -2 \\ 0 & -2 & 5 \end{pmatrix}$;

4) $\begin{pmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{pmatrix}$; 5) $\begin{pmatrix} 1 & -2 & 2 \\ -2 & -2 & 4 \\ 2 & 4 & -2 \end{pmatrix}$; 6) $\begin{pmatrix} 7 & -4 & 0 \\ -4 & 5 & 4 \\ 0 & 4 & 3 \end{pmatrix}$;

7) $\begin{pmatrix} 3 & -2 & 4 \\ -2 & 6 & -2 \\ -4 & -2 & 3 \end{pmatrix}$; 8) $\begin{pmatrix} 5 & -2 & -2 \\ -2 & 6 & 0 \\ -2 & 0 & 4 \end{pmatrix}$; 9) $\begin{pmatrix} 6 & -2 & 2 \\ -2 & 5 & 0 \\ 2 & 0 & 7 \end{pmatrix}$;

$$10)\begin{pmatrix}11 & 2 & -8\\ 2 & 2 & 10\\ -8 & 10 & 5\end{pmatrix};\quad 11)\begin{pmatrix}17 & -8 & 4\\ -8 & 17 & -4\\ 4 & -4 & 11\end{pmatrix};\quad 12)\begin{pmatrix}1 & 2 & 1\\ 2 & -5 & 2\\ 1 & 2 & 5\end{pmatrix}.$$

11.2. Методом Якоби найти собственные значения и собственные векторы следующих матриц:

$$1)\begin{pmatrix}3 & -i & 0\\ i & 3 & 0\\ 0 & 0 & 3\end{pmatrix};\qquad 2)\begin{pmatrix}3 & 2+2i & 0\\ 2-2i & 10 & 0\\ 0 & 0 & 4\end{pmatrix};$$

$$3)\begin{pmatrix}3 & 2-i & 0\\ 2+i & 7 & 0\\ 0 & 0 & 3\end{pmatrix};\qquad 4)\begin{pmatrix}0 & i & 0\\ -i & 0 & 0\\ 0 & 0 & 1\end{pmatrix};$$

$$5)\begin{pmatrix}3 & 2+2i & 0\\ 2-2i & 1 & 0\\ 0 & 0 & 4\end{pmatrix};\qquad 6)\begin{pmatrix}1 & 1+i & 0\\ 1-i & 0 & 0\\ 0 & 0 & 1\end{pmatrix};$$

$$7)\begin{pmatrix}1 & 0 & 1-i\\ 0 & 1 & 0\\ 1+i & 0 & 2\end{pmatrix};\qquad 8)\begin{pmatrix}3 & 2+2i & 0\\ 2-2i & 5 & 0\\ 0 & 0 & 1\end{pmatrix};$$

$$9)\begin{pmatrix}3 & 2-2i & 0\\ 2+2i & 1 & 0\\ 0 & 0 & 1\end{pmatrix}.$$

11.3. С помощью QR-алгоритма найти характеристические числа матриц из упражнения 11.1 и следующих матриц:

$$1)\begin{pmatrix}4 & -1 & -2\\ 2 & 1 & -2\\ 1 & -1 & 1\end{pmatrix};\quad 2)\begin{pmatrix}1 & -3 & 4\\ 4 & -7 & 8\\ 6 & -7 & 7\end{pmatrix};\quad 3)\begin{pmatrix}4 & -2 & 2\\ -5 & 7 & -5\\ -6 & 6 & -4\end{pmatrix};$$

$$4)\begin{pmatrix}7 & -12 & -2\\ 3 & -4 & 0\\ -2 & 0 & -2\end{pmatrix};\quad 5)\begin{pmatrix}2 & -1 & 2\\ 5 & -3 & 3\\ -1 & 0 & -2\end{pmatrix};\quad 6)\begin{pmatrix}-1 & 3 & -1\\ -3 & 5 & -1\\ -3 & 3 & 1\end{pmatrix}.$$

11.4. Степенным методом найти наибольшие по абсолютной величине собственные значения и соответствующие им собственные векторы матриц из упражнения 11.1.

11.5. Степенным методом в комбинации с методом исчерпывания найти собственные значения и собственные векторы матриц из упражнения 11.1.

11.6. Методом скалярных произведений найти наибольшие по абсолютной величине собственные значения и соответствующие им собственные векторы матриц из упражнения 11.1.

11.7. Методом скалярных произведений в комбинации с методом исчерпывания найти собственные значения и собственные векторы матриц из упражнения 11.1.

11.8. Найти собственные значения и собственные векторы следующих положительно определенных симметрических матриц:

$$1)\begin{pmatrix} 2 & -2 & -1 \\ -2 & 5 & 2 \\ -1 & 2 & 2 \end{pmatrix};\quad 2)\begin{pmatrix} 18 & 6 & -6 \\ 6 & 21 & 0 \\ -6 & 0 & 16 \end{pmatrix};\quad 3)\begin{pmatrix} 14 & 2 & -4 \\ 2 & 17 & 2 \\ -4 & 2 & 14 \end{pmatrix};$$

$$4)\begin{pmatrix} 6 & 2 & -2 \\ 2 & 7 & 0 \\ -2 & 0 & 5 \end{pmatrix};\quad 5)\begin{pmatrix} 5 & 2 & -4 \\ 2 & 1 & -2 \\ -4 & -2 & 5 \end{pmatrix};\quad 6)\begin{pmatrix} 3 & 2 & 0 \\ 2 & 4 & -2 \\ 0 & -2 & 5 \end{pmatrix};$$

$$7)\begin{pmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{pmatrix};\quad 8)\begin{pmatrix} 17 & -8 & 4 \\ -8 & 17 & -4 \\ 4 & -4 & 11 \end{pmatrix};\quad 9)\begin{pmatrix} 44 & -22 & 26 \\ -22 & 29 & -4 \\ 26 & -4 & 53 \end{pmatrix};$$

$$10)\begin{pmatrix} 9 & -16 & -8 \\ -16 & 33 & 16 \\ -8 & 16 & 9 \end{pmatrix};\qquad 11)\begin{pmatrix} 13 & 14 & -4 \\ 14 & 24 & -18 \\ -4 & -18 & 29 \end{pmatrix};$$

$$12)\begin{pmatrix} 12 & 22 & -22 \\ 22 & 45 & -44 \\ -22 & -44 & 45 \end{pmatrix};\qquad 13)\begin{pmatrix} 8 & -6 & -4 \\ -6 & 9 & -2 \\ -4 & -2 & 4 \end{pmatrix};$$

$$14)\begin{pmatrix} 11 & -7 & -9 \\ -7 & 11 & 9 \\ -9 & 9 & 27 \end{pmatrix};\qquad 15)\begin{pmatrix} 29 & -16 & -4 \\ -16 & 17 & 16 \\ -4 & 16 & 29 \end{pmatrix},$$

строя для них сингулярное разложение (см. разд. 8.18 и 11.1).

Глава 12 ЭЛЕМЕНТЫ n-МЕРНОЙ АНАЛИТИЧЕСКОЙ ГЕОМЕТРИИ

12.1. Аффинные пространства

Пусть дано линейное пространство X над полем P и непустое точечное множество $\mathcal{A}$. Множество $\mathcal{A}$ называют ***аффинным***, или ***точечно-векторным***, ***пространством***, связанным с линейным пространством X, а его элементы — точками, если выполняются следующие аксиомы:

1) каждой упорядоченной паре точек A и B множества $\mathcal{A}$ поставлен во взаимно однозначное соответствие вектор x из линейного пространства X, обозначаемый $x = \overline{AB}$;

2) для каждой точки $A \in \mathcal{A}$ и каждого вектора $x \in X$ существует единственная точка $B \in \mathcal{A}$, такая, что $x = \overline{AB}$;

3) для любых точек $A, B, C \in \mathcal{A}$ выполняется „правило треугольника“

$$\overline{AB} + \overline{BC} = \overline{AC}.$$

Векторы из линейного пространства X называют ***свободными векторами аффинного пространства*** $\mathcal{A}$. Если линейное пространство X n-мерное, то и аффинное пространство $\mathcal{A}$ называют ***n-мерным аффинным пространством***.

Примерами конечномерных аффинных пространств служат одномерное, двумерное и трехмерное пространства, изучаемые в аналитической геометрии. Приведем ещё один пример аффинного пространства. Будем называть векторами строки $(\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_n,\ 0)$, $\alpha_i \in P$, а точками — строки

$(\alpha_1, \alpha_2, \ldots, \alpha_n, 1)$, $\alpha_i \in P$. Операции над векторами определим равенствами

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

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

Точкам $(\alpha_1, \ldots, \alpha_n, 1)$ и $(\beta_1, \ldots, \beta_n, 1)$ поставим в соответствие вектор $(\beta_1 - \alpha_1, \ldots, \beta_n - \alpha_n, 0)$. Непосредственной проверкой легко убедиться в том, что множество X всех векторов с указанными операциями образует линейное пространство над полем P, а множество всех точек — n-мерное аффинное пространство, связанное с линейным пространством X_n.

В дальнейшем нас будут интересовать лишь конечномерные аффинные пространства, связанные с действительными линейными пространствами. Такие аффинные пространства соответственно будем называть ***конечномерными действительными аффинными пространствами***.

12.2. Координаты в аффинном пространстве

Пусть дано n-мерное аффинное пространство $\mathcal{A}$, связанное с линейным пространством X.

Системой координат, или ***репером***, в аффинном пространстве $\mathcal{A}$ называют пару (O, e) из некоторой точки $O \in \mathcal{A}$, называемой ***началом координат***, и некоторого базиса $e = (e_1, e_2, \ldots, e_n)$ в линейном пространстве X.

Координатами точки $A \in \mathcal{A}$ в системе координат (O, e) называют координаты вектора $\overline{OA} \in X$ (***радиус-вектора точки***) в базисе e, т.е. коэффициенты $\alpha_1, \alpha_2, \ldots, \alpha_n$ в представлении

$$\overline{OA} = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n = (e_1, e_2, \ldots, e_n)\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = e[OA]_e,$$

и записывают $A(\alpha_1, \alpha_2, \ldots, \alpha_n)_e$. Индекс e обычно опускают, если это не вызывает путаницы.

Пусть в аффинном пространстве $\mathcal{A}$ две точки $A(\alpha_1, \ldots, \alpha_n)$ и $B(\beta_1, \ldots, \beta_n)$ заданы координатами в системе координат (O, e). Тогда

$$\overline{OA} = \alpha_1 e_1 + \ldots + \alpha_n e_n = e\alpha, \quad \overline{OB} = \beta_1 e_1 + \ldots + \beta_n e_n = e\beta$$

и из правила треугольника $\overline{OA} + \overline{AB} = \overline{OB}$ получаем соотношение

$$\overline{AB} = \overline{OB} - \overline{OA} = e(\beta - \alpha).$$

Таким образом, координаты $x_1, \ldots, x_n$ вектора $\overline{AB}$ в базисе e вычисляются по формулам

$$x_1 = \beta_1 - \alpha_1, \quad \ldots, \quad x_n = \beta_n - \alpha_n,$$

т.е. равны разностям соответствующих координат конца и начала вектора.

Пусть в аффинном пространстве $\mathcal{A}$ даны две системы координат (O, e), где $O \in \mathcal{A}$, $e = (e_1, \ldots, e_n)$ — базис в X, и (O, e'), где $O' \in \mathcal{A}$, $e' = (e_1', \ldots, e_n')$ — базис в X. Пусть известны координаты точки O' в системе координат (O, e), т.е. $O'(\gamma_1, \gamma_2, \ldots, \gamma_n)_e$, и матрица перехода $T = (t_{ij})$ из базиса e в базис e'. Выясним, как связаны координаты $A(\alpha_1, \alpha_2, \ldots, \alpha_n)_e$ произвольной точки $A \in \mathcal{A}$ в системе координат (O, e) с ее координатами $A(\alpha_1', \alpha_2', \ldots, \alpha_n')_{e'}$ в системе координат (O', e'). Исходя из известных координат точки O' в системе координат (O, e) заключаем, что

$$\overline{OO'} = \gamma_1 e_1 + \ldots + \gamma_n e_n = e\gamma,$$

где $\gamma = (\gamma_1, \ldots, \gamma_n)^{\mathrm{T}}$. Исходя из координат точки A в обеих системах координат получаем еще два равенства

$$\overline{OA} = \alpha_1 e_1 + \cdots + \alpha_n e_n = e\alpha, \quad \overline{O'A} = \alpha_1' e_1 + \cdots + \alpha_n' e_n = e'\alpha',$$

где $\alpha = (\alpha_1, \ldots, \alpha_n)^{\mathrm{T}}$, $\alpha' = (\alpha'_1, \ldots, \alpha'_n)^{\mathrm{T}}$. Наконец, базисы e и e' с помощью матрицы перехода связаны соотношением $e' = eT$. Учитывая все это, находим

$$e\alpha = \overline{OA} = \overline{OO'} + \overline{O'A} = e\gamma + e'\alpha' = e\gamma + eT\alpha' = e(\gamma + T\alpha').$$

Последнее равенство — это равенство двух разложений в базисе e. Из него вытекает равенство коэффициентов, т.е.

$$\alpha = \gamma + T\alpha', \tag{12.1}$$

или в подробной записи:

$$\begin{cases} \alpha_1 = \gamma_1 + t_{11}\alpha'_1 + t_{12}\alpha'_2 + \ldots + t_{1n}\alpha'_n, \\ \alpha_2 = \gamma_2 + t_{21}\alpha'_1 + t_{22}\alpha'_2 + \ldots + t_{2n}\alpha'_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \alpha_n = \gamma_n + t_{n1}\alpha'_1 + t_{n2}\alpha'_2 + \ldots + t_{nn}\alpha'_n. \end{cases}$$

Учитывая невырожденность матрицы T, из (12.1) можно получить другое соотношение, разрешенное относительно α':

$$\alpha' = T^{-1}\alpha - T^{-1}\gamma. \tag{12.2}$$

Формула (12.1) дает выражение старых координат точки A через ее новые координаты, а формула (12.2) — выражение новых координат точки через ее старые координаты.

12.3. Плоскости в аффинном пространстве

Пусть в n-мерном аффинном пространстве $\mathcal{A}$, связанном с линейным пространством X, дана точка M_0, а в линейном пространстве X — k-мерное подпространство L. Множество точек в $\mathcal{A}$, для которых вектор $\overline{M_0M}$ принадлежит L, т.е. множество, описываемое уравнением

$$r = L + r_0, \tag{12.3}$$

где $r = \overline{OM}$, $r_0 = \overline{OM_0}$, называют ***k-мерной плоскостью***, проходящей через точку M_0 в направлении подпространства L. Точку M_0 называют ***начальной точкой плоскости***, точку M — ***текущей точкой плоскости***, подпространство L — ***направляющим подпространством плоскости***.

Всякая k-мерная плоскость в аффинном пространстве $\mathcal{A}$ сама является аффинным пространством размерности k.

Уравнение (12.3) показывает, что рассматриваемая плоскость получается *параллельным переносом* подпространства L на вектор r_0. Очевидно, что нульмерная плоскость состоит лишь из одной точки M_0. Поэтому каждую точку аффинного пространства можно рассматривать как нульмерную плоскость. Само аффинное пространство является n-мерной плоскостью. Одномерную плоскость обычно называют ***прямой***, $(n-1)$-мерную плоскость — ***гиперплоскостью***.

В определение плоскости введена ее начальная точка. В действительности все точки плоскости равноправны, так как в качестве начальной точки плоскости можно взять любую точку этой плоскости. Из этого следует, что ***две плоскости, имеющие одну общую точку и одно и то же направляющее подпространство, совпадают***.

Пусть в аффинном пространстве $\mathcal{A}$ выбрана система координат (O, e), где $O \in \mathcal{A}$, $e = (e_1, e_2, \ldots, e_n)$ — базис линейного пространства свободных векторов, и дана k-мерная плоскость $r = L + r_0$, проходящая через точку $M_0\left(x_1^0, x_2^0, \ldots, x_n^0\right)_e$ параллельно подпространству $L = \langle a_1, a_2, \ldots, a_k\rangle$ (векторы a_1, a_2, ..., a_k линейно независимые, так как подпространство L k-мерное). Тогда рассматриваемую плоскость можно записать в виде

$$r = t_1 a_1 + \ldots + t_k a_k + r_0, \tag{12.4}$$

где t_1, ..., t_k — параметры, принимающие произвольные числовые значения из поля P независимо друг от друга.

Уравнение (12.4) называют ***параметрическим уравнением*** рассматриваемой плоскости ***в векторной форме***.

Пусть в выбранной системе координат

$$[a_1]_e = (a_{11},\ a_{21},\ \ldots,\ a_{n1})^{\mathrm{T}}, \quad \ldots, \quad [a_k]_e = (a_{1k},\ a_{2k},\ \ldots,\ a_{nk})^{\mathrm{T}},$$
$$[r]_e = (x_1,\ x_2,\ \ldots,\ x_n)^{\mathrm{T}}, \qquad [r_0]_e = \left(x_1^0,\ x_2^0,\ \ldots,\ x_n^0\right)^{\mathrm{T}}.$$

Тогда, переходя от векторного равенства (12.4) к покоординатным равенствам, получим систему уравнений

$$\begin{cases} x_1 = a_{11}t_1 + a_{12}t_2 + \ldots + a_{1k}t_k + x_1^0, \\ x_2 = a_{21}t_1 + a_{22}t_2 + \ldots + a_{2k}t_k + x_2^0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n = a_{n1}t_1 + a_{n2}t_2 + \ldots + a_{nk}t_k + x_n^0. \end{cases} \tag{12.5}$$

Систему (12.5) называют ***параметрическими уравнениями k-мерной плоскости***, проходящей через точку M_0 в направлении подпространства L.

В частном случае $k = 1$ k-мерная плоскость есть прямая. При этом подпространство L порождается одним вектором a с координатами $(a_1,\ \ldots,\ a_n)^{\mathrm{T}}$, а параметрические уравнения прямой имеют вид:

$$\begin{cases} x_1 = a_1 t + x_1^0, \\ \ldots\ldots\ldots \\ x_n = a_n t + x_n^0. \end{cases} \tag{12.6}$$

При $t \geqslant t_0$ на прямой (12.6) выделяется луч, а при $t_1 \leqslant t \leqslant$ $\leqslant t_2$ — отрезок.

От параметрических уравнений прямой (12.6) легко перейти к ее ***каноническим уравнениям***

$$\frac{x_1 - x_1^0}{a_1} = \frac{x_2 - x_2^0}{a_1} = \ldots = \frac{x_n - x_n^0}{a_n}. \tag{12.7}$$

Исключив из параметрических уравнений (12.5) k-мерной плоскости все параметры, получим ее ***общие уравнения***

$$\begin{cases} A_{11}x_1 + \ldots + A_{1n}x_n + D_1 = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ A_{n-k,1}x_1 + \ldots + A_{n-k,n}x_n + D_{n-k} = 0. \end{cases} \tag{12.8}$$

В частности, гиперплоскость, соответствующая случаю $k = n - 1$, задается одним уравнением

$$A_1x_1 + \ldots + A_nx_n + D = 0. \qquad (12.9)$$

В системе уравнений (12.8) каждое отдельное уравнение можно рассматривать как уравнение гиперплоскости, а всю систему уравнений — как определение k-мерной плоскости пересечением k гиперплоскостей. Это наглядно иллюстрируется на примере прямых в трехмерном пространстве, так как любая такая прямая может рассматриваться как пересечение двух плоскостей.

Получив уравнения прямых и плоскостей в аффинном пространстве, можно решать все вопросы аналитической геометрии относительно прямых и плоскостей в этом пространстве, т.е. вопросы относительно прямых и плоскостей, не связанные с измерением длин и углов. В частности, здесь можно развить теорию выпуклых множеств и выпуклых многогранников, нужную для линейного программирования.

12.4. Гиперповерхности второго порядка

Множество точек M n-мерного аффинного пространства $\mathcal{A}$, координаты которых удовлетворяют в выбранной системе координат уравнению второй степени

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j + 2\sum_{i=1}^{n} a_ix_i + a = 0, \qquad (12.10)$$

где $a_{ij} = a_{ji}$, $i, j = 1, 2, \ldots, n$, называют ***гиперповерхностью второго порядка***, или ***квадрикой***.

В двумерном аффинном пространстве квадрики называют ***линиями второго порядка***. Как и в аналитической геометрии для линий и поверхностей второго порядка, здесь рассматривают (см. [2], [28]) взаимное расположение прямой и квадрики, асимптотические направления, центр, диаметральные плоскости и диаметры квадрик, а также вопросы, связанные с упрощением уравнений квадрик и их аффинной классификацией. Мы остановимся лишь на упрощении уравнений квадрик и на их аффинной классификации.

Известно, что существует (см. разд. 9.4) невырожденное линейное преобразование переменных

$$x_i = \sum_{k=1}^{n} q_{ik} y_k, \quad i = 1, 2, \ldots, n, \tag{12.11}$$

приводящее квадратичную форму $\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} a_{ij}x_i x_j$ к каноническому виду $\sum\limits_{i=1}^{r} \lambda_i y_i^2$, где r — натуральное число, не превосходящее n.

Преобразование переменных (12.11) будем рассматривать как формулы преобразования координат. Подставив выражения для x_i из (12.11) в левую часть уравнения (12.10), приведем это уравнение к виду

$$\sum_{i=1}^{r} \lambda_i y_i^2 + 2\sum_{i=1}^{n} c_i y_i + a = 0. \tag{12.12}$$

Если $\lambda_i \neq 0$, то выделением полного квадрата по y_i и переносом начала координат в уравнении (12.12) можно уничтожить член $2c_i y_i$. Действительно, в этом случае сумму $\lambda_i y_i^2 + 2c_i y_i$ можно представить в виде

$$\lambda_i y_i^2 + 2c_i y_i = \lambda_i\Big(y_i + \frac{c_i}{\lambda_i}\Big) - \frac{c_i^2}{\lambda_i}.$$

Полагая

$$\widetilde{x}_i = y_i + \frac{c_i}{\lambda_i}, \quad \widetilde{x}_j = y_j \text{ при } j \neq i,$$

получим, что в уравнении (12.12) коэффициент при $\widetilde{x}_i^2$ останется равным λ_i, член с первой степенью $\widetilde{x}_i$ исчезнет и изменится свободный член. Проделав так со всеми переменными в уравнении (12.12), приведем это уравнение к виду

$$\sum_{i=1}^{r}\lambda_i\widetilde{x}_i^2+2\sum_{j=r+1}^{n}a'_j\widetilde{x}_j+b=0. \qquad (12.13)$$

Рассмотрим возможные случаи относительно коэффициентов a'_j и b в уравнении (12.13).

Если все коэффициенты a'_j нулевые, а $b\neq 0$, то уравнение (12.13) можно записать в виде

$$\sum_{i=1}^{r}\varepsilon_i\widetilde{x}_i^2=1,$$

где $\varepsilon_i=-\dfrac{\lambda_i}{b}$. Преобразуя далее координаты по формулам

$$X_i=\begin{cases}\sqrt{|\varepsilon_i|}\,\widetilde{x}_i, & \varepsilon_i\neq 0;\\ \widetilde{x}_i, & \varepsilon_i=0,\end{cases}$$

последнее уравнение приведем к виду

$$\sum_{i=1}^{r}\mu_iX_i^2=1, \qquad (12.14)$$

где $\mu_i=\pm1$ и $0<r\leqslant n$.

Если в уравнении (12.13) все $a'_j=0$ и $b=0$, то, полагая

$$X_i=\begin{cases}\sqrt{|\lambda_i|}\,\widetilde{x}_i, & \lambda_i\neq 0;\\ \widetilde{x}_i, & \lambda_i=0,\end{cases}$$

уравнение (12.13) приведем к виду

$$\sum_{i=1}^{r}\mu_iX_i^2=0, \qquad (12.15)$$

где $\mu_i=\pm1$, $0<r\leqslant n$. Это уравнение определяет конус с $(n-r)$-*вершиной*.

Если в уравнении (12.13) среди коэффициентов a'_j есть отличные от нуля, то, полагая

$$X_i = \begin{cases} \widetilde{x}_i, & i \neq r+1; \\ -\sum\limits_{j=r+1}^{n} a'_j \widetilde{x}_j - \frac{b}{2}, & i = r+1, \end{cases} \tag{12.16}$$

уравнение (12.13) приведем к виду

$$\sum_{i=1}^{r} \lambda_i X_i^2 - 2X_{r+1} = 0, \tag{12.17}$$

где r — натуральное число, меньшее n.

Уравнения (12.14), (12.15), (12.17) определяют следующие аффинные классы квадрик в $\mathcal{A}_n$:

1) если в уравнении (12.14) $r = n$ и все $\mu_i = 1$, то оно определяет ***эллипсоид***

$$X_1^2 + \ldots + X_n^2 = 1;$$

2) если в уравнении (12.14) $r = n$ и все $\mu_i = -1$, то оно определяет ***мнимый эллипсоид***

$$X_1^2 + \ldots + X_n^2 = -1;$$

3) если в уравнении (12.14) $r = n$ и μ_i разных знаков, например, $\mu_1 = \ldots = \mu_k = 1$, $\mu_{k+1} = \ldots = \mu_n = -1$, то оно определяет ***гиперболоид***

$$X_1^2 + \ldots + X_k^2 - X_{k+1}^2 - \ldots - X_n^2 = 1;$$

4) уравнение (12.15) определяет ***конусы***

$$\mu_1 X_1^2 + \ldots + \mu_n X_n^2 = 0,$$

$\mu_i = \pm 1$, с точечной вершиной при $r = n$ и с $(n-r)$-вершиной при $r < n$, а именно, если $\mu_1, \ldots, \mu_n$ имеют одинаковые знаки, то будет ***мнимый конус***; если $\mu_1, \ldots, \mu_n$ имеют разные знаки, то будет ***действительный конус***;

5) при $r = n - 1$ уравнение (12.17) определяет ***параболоиды***;

6) при $r < n$ уравнение (12.14) определяет ***эллиптические*** и ***гиперболические цилиндры*** в зависимости от того, одинаковые или разные знаки у коэффициентов $\mu_1, \ldots, \mu_r$;

7) при $r < n - 1$ уравнение (12.17) определяет ***параболические цилиндры.***

Замечание. Первые пять аффинных классов поверхностей являются основными; оставшиеся два класса повторяют основные в подпространствах меньшей размерности. Более подробная классификация квадрик в n-мерном пространстве приведена в [22]. Из полученных результатов непосредственно вытекает аффинная классификация линий второго порядка на плоскости и поверхностей второго порядка в пространстве.

Пример 12.1. В трехмерном аффинном пространстве $\mathcal{A}$ задана квадрика уравнением

$$x_1^2 + 2x_2^2 + 5x_3^2 - 2x_1x_2 + 4x_2x_3 + 2x_1 - 4x_2 + 4x_3 - 17 = 0.$$

Упростить уравнение квадрики, установить ее аффинный класс, указать формулы перехода к новой системе координат, координаты нового начала и новых координатных векторов.

Решение. Начнем с приведения квадратичной формы

$$f(x_1, x_2, x_3) = x_1^2 + 2x_2^2 + 5x_3^2 - 2x_1x_2 + 4x_2x_3$$

методом Лагранжа к каноническому виду. Для этого сначала в ней выделим полные квадраты по переменным x_1, x_2, x_3. Тогда получим

$$f(x_1, x_2, x_3) = (x_1^2 - 2x_1x_2) + 2x_2^2 + 5x_3^2 + 4x_2x_3 =$$
$$= (x_1 - x_2)^2 + (x_2 + 4x_2x_3) + 5x_3^2 = (x_1 - x_2)^2 + (x_2 + 2x_3)^2 + x_3^2.$$

Введем новые переменные, полагая $y_1 = x_1 - x_2$, $y_2 = x_2 + 2x_3$, $y_3 = x_3$, что соответствует замене переменных $x_1 = y_1 + y_2 - 2y_3$, $x_2 = y_2 - 2y_3$, $x_3 = y_3$. В новых переменных рассматриваемая квадратичная форма принимает канонический вид $y_1^2+y_2^2+y_3^2$. Этим же преобразованием переменных уравнение квадрики преобразуется в уравнение

$$y_1^2+y_2^2+y_3^2+2(y_1+y_2-2y_3)-4(y_2-2y_3)+4y_3+17=0,$$

эквивалентное уравнению

$$y_1^2+y_2^2+y_3^2+2y_1-2y_2+8y_3+17=0.$$

Выделив в левой части этого уравнения полные квадраты по переменным y_1, y_2, y_3, запишем его в виде

$$(y_1+1)^2+(y_2-1)^2+(y_3+4)^2-1=0.$$

Введем новые переменные $z_1=y_1+1$, $z_2=y_2-1$, $z_3=y_3+4$, соответствующие замене переменных $y_1=z_1-1$, $y_2=z_2+1$, $y_3=z_3-4$. В новых переменных уравнение квадрики имеет вид:

$$z_1^2+z_2^2+z_3^2=1.$$

Результирующее преобразование переменных имеет вид

$$\begin{cases} x_1=y_1+y_2-2y_3=z_1+z_2-2z_3+8, \\ x_2=y_2-2y_3=z_2-2z_3+9, \\ x_3=y_3=z_3+9. \end{cases}$$

Таким образом, новая система координат имеет начало в точке $O'(8,\ 9,\ 9)$ и координатные векторы $e_1'=(1,\ 0,\ 0)^{\mathrm{T}}$, $e_2'=(1,\ 1,\ 0)^{\mathrm{T}}$, $e_3'=(-2,\ -2,\ 1)^{\mathrm{T}}$. В этой системе координат рассматриваемая квадрика имеет уравнение

$$z_1^2+z_2^2+z_3^2=1.$$

Поэтому она является эллипсоидом.

Пример 12.2. В трехмерном аффинном пространстве квадрика задана уравнением

$$2x_1^2+5x_2^2+2x_3^2-2x_1x_2-4x_1x_2+2x_2x_3-2x_1+10x_2+2x_3+3=0.$$

Упростить уравнение этой квадрики и установить ее аффинный класс.

Решение. Сначала методом Лагранжа приведем квадратичную форму, входящую в уравнение квадрики, к каноническому виду. Для этого выделим в ней полный квадрат по x_1. Тогда получим

$$f(x_1,x_2,x_3) = \frac{1}{2}(2x_1 - x_2 - 2x_3)^2 + 5x_2 + 2x_3^2 + 2x_2x_3 -$$

$$- \frac{1}{2}x_2^2 - 2x_3^2 - 2x_2x_3 = \frac{1}{2}(2x_1 - x_2 - 2x_3)^2 + \frac{9}{2}x_2^2.$$

Введя новые переменные $y_1 = 2x_1 - x_2 - 2x_3$, $y_2 = x_2$, $y_3 = x_3$, соответствующие замене переменных

$$x_1 = \frac{1}{2}\big(x_1 - y_2 + 2y_3\big), \quad x_2 = y_2, \quad x_3 = y_3,$$

квадратичную форму f приведем к виду

$$f(y_1,y_2,y_3) = \frac{1}{2}y_1^2 + \frac{9}{2}y_2^2,$$

а уравнение квадрики — к виду

$$\frac{1}{2}y_1^2 + \frac{9}{2}y_2^2 - y_1 + 9y_2 + 3 = 0.$$

В полученном уравнении выделим полные квадраты по y_1 и y_2. Тогда получим

$$\frac{1}{2}(y_1 - 1)^2 + \frac{9}{2}(y_2 + 1)^2 - 2 = 0.$$

Введем новые переменные $z_1 = y_1 - 1$, $z_2 = y_2 + 1$, $z_3 = y_3$, соответствующие замене переменных $y_1 = z_1 + 1$, $y_2 = z_2 - 1$, $y_3 = z_3$. Тогда уравнение квадрики запишется в виде

$$\frac{1}{2}z_1^2 + \frac{9}{2}z_2^2 - 2 = 0,$$

или
$$\frac{z_1^2}{4} + \frac{z_2^2}{4/9} = 1.$$

Выполним дополнительное преобразование переменных $z_1 = 2u_1$, $z_2 = \frac{2}{3}u_2$, $z_3 = u_3$. Тогда уравнение квадрики примет канонический вид
$$u_1^2 + u_2^2 = 1.$$

Из канонического вида следует, что рассматриваемая квадрика является эллиптическим цилиндром.

Пример 12.3. В трехмерном аффинном пространстве квадрика задана уравнением

$$x_1^2 + 4x_2^2 + 9x_3^2 - 4x_1x_2 - 6x_1x_3 + 12x_2x_3 + 2x_1 - 2x_2 + 4x_3 + 3 = 0.$$

Упростить уравнение этой квадрики и установить ее аффинный класс.

Решение. Выделяя в квадратичной форме $f(x_1, x_2, x_3) = x_1^2 + 4x_2^2 + 9x_3^2 - 4x_1x_2 - 6x_1x_3 + 12x_2x_3$ полный квадрат по переменной x_1, получим

$$f(x_1, x_2, x_3) = (x_1 - 2x_2 - 3x_3)^2.$$

Вводя новые переменные $y_1 = x_1 - 2x_2 - 3x_3$, $y_2 = x_2$, $y_3 = x_3$, соответствующие замене переменных $x_1 = y_1 + 2y_2 + 3y_3$, $x_2 = y_2$, $x_3 = y_3$, приведем квадратичную форму к виду $f(y_1, y_2, y_3) = y_1^2$, а уравнение квадрики – к виду

$$y_1^2 + 2y_1 + 2y_2 + 10y_3 + 3 = 0.$$

В этом уравнении выделим полный квадрат по y_1. Тогда получим
$$(y_1 + 1)^2 + 2y_2 + 10y_3 + 2 = 0.$$

Введя новые переменные $z_1 = y_1 + 1$, $z_2 = y_2$, $z_3 = y_3$, соответствующие замене переменных $y_1 = z_1 - 1$, $y_2 = z_2$, $y_3 = z_3$, приведем уравнение квадрики к виду

$$z_1^2 + 2z_2 + 10z_3 + 2 = 0.$$

Положим далее $u_1 = z_1$, $u_2 = -z_2 - 5z_3 - 1$, $u_3 = z_3$, что соответствует замене переменных $z_1 = u_1$, $z_2 = u_2 - 5u_3 - 1$, $z_3 = u_3$. Тогда уравнение квадрики примет канонический вид

$$u_1^2 = 2u_2.$$

Исходя из канонического вида квадрики заключаем, что рассматриваемая квадрика является параболическим цилиндром.

12.5. Точечно-векторное евклидово пространство

Аффинное пространство называют ***точечно-векторным евклидовым пространством***, или ***евклидовым пространством***, если связанное с ним линейное пространство является евклидовым векторным пространством. Мы будем рассматривать лишь конечномерные евклидовы пространства, т.е. такие точечно-векторные пространства, с которыми связаны n-мерные евклидовы векторные пространства. Трехмерное точечно-векторное евклидово пространство представляет собой трехмерное пространство, изучаемое в аналитической геометрии.

Любое конечномерное аффинное пространство можно превратить в точечно-векторное евклидово пространство, так как линейное пространство, связанное с аффинным пространством, всегда можно превратить в евклидово векторное пространство, задав в нем скалярное умножение векторов.

Система координат, координаты точек и векторов в конечномерном точечно-векторном евклидовом пространстве вводятся так же, как в аффинном пространстве. При этом систему координат (O, e) называют ***прямоугольной***, если базис $e = (e_1, \ldots, e_n)$ векторного пространства, связанного с аффинным, ортонормированный.

Формулы (12.1) преобразования координат также сохраняются. При этом если совершается переход от одной прямоугольной системы координат к другой прямоугольной системе координат, то матрица T в этих формулах будет ортогональной. Например, на плоскости (двумерное евклидово пространство) при повороте осей координат на угол α формулы преобразования прямоугольных координат будут иметь вид:

$$\begin{cases} x = x'\cos\alpha - y'\sin\alpha + x_0, \\ y = x'\sin\alpha + y'\cos\alpha + y_0. \end{cases}$$

Поскольку точечно-векторное евклидово пространство является в то же время и аффинным пространством, то в нем сохраняется геометрия прямых, плоскостей и квадрик. При этом появляется возможность рассматривать еще вопросы, связанные с измерением длин и углов. В связи с этим геометрия в n-мерном точечно-векторном евклидовом пространстве в точности напоминает изученную ранее аналитическую геометрию, а в случае двух и трех измерений совпадает с ней. Так, например, умножив скалярно векторное уравнение гиперплоскости

$$r = L + r_0, \tag{12.18}$$

заданное в некоторой прямоугольной системе координат, на вектор $n = (a_1, \ldots, a_n)$, ортогональный $(n-1)$-мерному подпространству L, придадим этому уравнению вид

$$(n, r) = (n, r_0), \tag{12.19}$$

что в координатной форме дает уравнение гиперплоскости

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n + d = 0, \tag{12.20}$$

где $d = -(n, r_0)$. Таким образом, как и для плоскости в аналитической геометрии, коэффициентами при переменных в общем уравнении гиперплоскости в прямоугольной системе координат служат координаты ортогонального к этой гиперплоскости вектора.

Уравнение гиперплоскости (12.20), как и уравнение плоскости в аналитической геометрии, можно привести к нормальному виду

$$\frac{a_1x_1 + \ldots + a_nx_n + d}{\pm\sqrt{a_1^2 + \ldots + a_n^2}} = 0, \tag{12.21}$$

где знак у корня, как и в аналитической геометрии, выбирается противоположным знаку d.

Нормальное уравнение позволяет получить формулу расстояния ρ от точки $M_0(x_1^0, \ldots, x_n^0)$ до гиперплоскости (12.21):

$$\rho = \frac{|a_1x_1^0 + \ldots + a_nx_n^0 + d|}{\sqrt{a_1^2 + \ldots + a_n^2}}. \tag{12.22}$$

Как видим, в точечно-векторном аффинном пространстве можно воспроизвести все результаты, которые известны в рамках аналитической геометрии.

12.6. Гиперповерхности второго порядка в евклидовом пространстве

Пусть в прямоугольной системе координат евклидова пространства E гиперповерхность второго порядка (квадрика) задана уравнением

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j + 2\sum_{i=1}^{n} a_ix_i + a = 0, \tag{12.23}$$

где коэффициенты a_{ij} квадратичной формы квадрики удовлетворяют условию $a_{ij} = a_{ji}$ для всех пар индексов.

Как и в аналитической геометрии, имеющей дело с линиями и поверхностями второго порядка, в точечно-векторном евклидовом пространстве обычно изучают вопросы упрощения уравнений квадрик, главные направления, метрические инварианты и метрическую классификацию квадрик (см. [2], [4], [28]). Мы остановимся лишь на упрощении уравнений квадрик и на их метрической классификации. При этом все будем делать так же, как это делалось для квадрик в аффинном пространстве, с той лишь разницей, что каждый раз будем пользоваться прямоугольными системами координат.

Известно (см. разд. 9.8), что существует ортогональное преобразование переменных

$$x_i = \sum_{k=1}^{n} q_{ik} y_k, \quad i = 1, 2, \ldots, n, \tag{12.24}$$

приводящее квадратичную форму $\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} a_{ij}x_i x_j$ квадрики к каноническому виду $\sum\limits_{i=1}^{r} \lambda_i y_i^2$. Преобразования переменных (12.23) будем рассматривать как преобразования прямоугольных координат в евклидовом пространстве. Подставляя выражения для x_i из (12.24) в левую часть уравнения (12.23), приведем это уравнение к виду

$$\sum_{i=1}^{r} \lambda_i y_i^2 + 2\sum_{i=1}^{n} c_i y_i + a = 0. \tag{12.25}$$

Если $\lambda_i \neq 0$, то выделением полного квадрата по переменной y_i и переносом начала координат можно уничтожить в уравнении (12.25) член с первой степенью этой переменной. Действительно, в этом случае сумму $\lambda_i y_i^2 + 2c_i y_i$ можно представить в виде

$$\lambda_i y_i^2 + 2c_i y_i = \lambda_i\left(y_i + \frac{c_i}{\lambda_i}\right)^2 - \frac{c_i^2}{\lambda_i}.$$

Полагая $X_i = y_i + c_i/\lambda_i$, $X_j = y_j$ при $j \neq i$, получим, что в уравнении (12.25) коэффициент при X_i^2 останется равным λ_i, член с первой степенью X_i исчезнет и изменится свободный член. Проделав так со всеми переменными в уравнении (12.25), приведем это уравнение к виду

$$\sum_{i=1}^{r} \lambda_i X_i^2 + 2\sum_{j=r+1}^{n} a_j' X_j + b = 0, \tag{12.26}$$

где r — натуральное число, не превышающее n.

Рассмотрим возможные случаи относительно коэффициентов a'_j в уравнении (12.26).

Если все $a'_j = 0$, то имеем приведенное уравнение квадрики первого рода

$$\sum_{i=1}^{r} \lambda_i X_i^2 + b = 0. \tag{12.27}$$

Если среди коэффициентов a'_j в уравнении (12.26) есть ненулевые, то, полагая

$$\begin{cases} z_1 = X_1, \\ \dots\dots\dots\dots\dots\dots \\ z_r = X_r, \\ z_{r+1} = -\dfrac{\sum\limits_{j=r+1}^{n} a'_j X_j + \dfrac{b}{2}}{\sqrt{{a'_{r+1}}^2 + \ldots + {a'_n}^2}}, \\ z_{r+2} = \gamma_{r+2,1} X_1 + \ldots + \gamma_{r+2,n} X_n, \\ \dots\dots\dots\dots\dots\dots \\ z_n = \gamma_{n,1} X_1 + \ldots + \gamma_{n,n} X_n, \end{cases} \tag{12.28}$$

уравнение (12.26) приведем к виду

$$\sum_{i=1}^{r} \lambda_i z_i^2 - 2p z_{r+1} = 0, \tag{12.29}$$

где $p = \sqrt{{a'_{r+1}}^2 + \ldots + {a'_n}^2}$.

Уравнение (12.29) называют ***приведенным уравнением квадрики второго рода***.

В формулах (12.28) преобразования прямоугольных координат коэффициенты γ_{ij}, $i = r+2, r+3, \ldots, n$, $j = 1, 2, \ldots, n$, выбирают так, чтобы матрица этого преобразования координат была ортогональной. Это равносильно дополнению ортонормированной системы n-мерных векторов-строк

$$\begin{gathered}(1,\ 0,\ 0,\ \ldots,\ 0),\\ (0,\ 1,\ 0,\ \ldots,\ 0),\\ \ldots\ldots\ldots\ldots\ldots\\ \left(0,\ \ldots,\ 0,\ \frac{a'_{r+1}}{p},\ \ldots,\ \frac{a'_n}{p}\right),\end{gathered}$$

где $p=\sqrt{{a'_{r+1}}^2+\ldots+{a'_n}^2}$, до ортонормированного базиса в пространстве n-мерных векторов-строк.

Уравнения (12.27) и (12.29) определяют следующие метрические классы квадрик в n-мерном евклидовом пространстве:

1) если в уравнении (12.27) $r=n$ и все коэффициенты $\lambda_1,\ \ldots,\ \lambda_n$ одного знака, а коэффициент b противоположного им знака, то уравнение определяет эллипсоид;

2) если в уравнении (12.27) $r=n$ и все коэффициенты $\lambda_1,\ \ldots,\ \lambda_n,\ b$ одного знака, то уравнение определяет мнимый эллипсоид;

3) если в уравнении (12.27) $r=n$, $b\neq 0$, а коэффициенты $\lambda_1,\ \ldots,\ \lambda_n$ имеют разные знаки, то уравнение определяет гиперболоид;

4) уравнение (12.27) при $b=0$ определяет конус с точечной вершиной при $r=n$ и с $(n-r)$-вершиной при $r<n$ (мнимый конус, если $\lambda_1,\ \ldots,\ \lambda_n$ имеют одинаковые знаки, и действительный конус в противном случае);

5) при $r=n-1$ уравнение (12.29) определяет параболоид;

6) при $r<n$ уравнение (12.27) определяет эллиптический или гиперболический цилиндр в зависимости от знаков коэффициентов $\lambda_1,\ \ldots,\ \lambda_r$;

7) при $r<n-1$ уравнение (12.29) определяет параболический цилиндр.

Более подробная классификация квадрик в евклидовом пространстве приведена в [23]. Из полученных результатов непосредственно вытекает метрическая классификация линий второго порядка на плоскости и поверхностей второго порядка в пространстве. Приведем эти классификации.

Метрическая классификация линий второго порядка на плоскости

1. $\dfrac{x_1^2}{a^2}+\dfrac{x_2^2}{b^2}=1$ — эллипс;
2. $\dfrac{x_1^2}{a^2}+\dfrac{x_2^2}{b^2}=-1$ — мнимый эллипс;
3. $\dfrac{x_1^2}{a^2}-\dfrac{x_2^2}{b^2}=1$ — гипербола;
4. $\dfrac{x_1^2}{a^2}-\dfrac{x_2^2}{b^2}=0$ — пара действительных пересекающихся прямых;
5. $\dfrac{x_1^2}{a^2}+\dfrac{x_2^2}{b^2}=0$ — пара мнимых пересекающихся прямых;
6. $x_1^2=2px_2$ — парабола;
7. $x_1^2-a^2=0$ — пара действительных параллельных прямых;
8. $x_1^2+a^2=0$ — пара мнимых параллельных прямых;
9. $x_1^2=0$ — пара совпавших прямых.

Метрическая классификация поверхностей второго порядка в трехмерном пространстве

1. $\dfrac{x_1^2}{a^2}+\dfrac{x_2^2}{b^2}+\dfrac{x_3^2}{c^2}=1$ — эллипсоид;
2. $\dfrac{x_1^2}{a^2}+\dfrac{x_2^2}{b^2}+\dfrac{x_3^2}{c^2}=-1$ — мнимый эллипсоид;
3. $\dfrac{x_1^2}{a^2}+\dfrac{x_2^2}{b^2}-\dfrac{x_3^2}{c^2}=1$ — однополостный гиперболоид;
4. $\dfrac{x_1^2}{a^2}+\dfrac{x_2^2}{b^2}-\dfrac{x_3^2}{c^2}=-1$ — двуполостный гиперболоид;
5. $\dfrac{x_1^2}{a^2}+\dfrac{x_2^2}{b^2}-\dfrac{x_3^2}{c^2}=0$ — конус;
6. $\dfrac{x_1^2}{a^2}+\dfrac{x_2^2}{b^2}+\dfrac{x_3^2}{c^2}=0$ — мнимый конус;
7. $\dfrac{x_1^2}{a^2}+\dfrac{x_2^2}{b^2}=2x_3$ — эллиптический параболоид;
8. $\dfrac{x_1^2}{a^2}-\dfrac{x_2^2}{b^2}=2x_3$ — гиперболический параболоид;
9. $\dfrac{x_1^2}{a^2}+\dfrac{x_2^2}{b^2}=1$ — эллиптический цилиндр;

35-4350

10. $\frac{x_1^2}{a^2}+\frac{x_2^2}{b^2}=-1$ — мнимый эллиптический цилиндр;
11. $\frac{x_1^2}{a^2}-\frac{x_2^2}{b^2}=1$ — гиперболический цилиндр;
12. $\frac{x_1^2}{a^2}-\frac{x_2^2}{b^2}=0$ — пара действительных пересекающихся плоскостей;
13. $\frac{x_1^2}{a^2}+\frac{x_2^2}{b^2}=0$ — пара мнимых пересекающихся плоскостей;
14. $x_1^2=2px_2$ — параболический цилиндр;
15. $x_1^2-a^2=0$ — пара действительных параллельных плоскостей;
16. $x_1^2+a^2=0$ — пара мнимых параллельных плоскостей;
17. $x_1^2=0$ — пара совпавших плоскостей.

Пример 12.4. В четырехмерном евклидовом пространстве в прямоугольной системе координат квадрика задана уравнением

$$9x_1^2+5x_2^2+5x_3^2+8x_4^2+8x_2x_3-4x_2x_4+4x_3x_4+\\+18x_1-54x_2+54x_3-54x_4+207=0.$$

Преобразованием прямоугольных координат упростить уравнение квадрики, указать преобразование координат, осуществляющее такое упрощение уравнения квадрики, и установить ее метрический класс.

Решение. Сначала приведем квадратичную форму

$$9x_1^2+5x_2^2+5x_3^2+8x_4^2+8x_2x_3-4x_2x_4+4x_3x_4$$

к каноническому виду в главных осях. Матрицей этой квадратичной формы является

$$A=\begin{pmatrix}9&0&0&0\\0&5&4&-2\\0&4&5&2\\0&-2&2&8\end{pmatrix}.$$

Ее характеристический многочлен $|A - \lambda E| = \lambda(\lambda - 9)^3$ имеет корни $\lambda_1 = \lambda_2 = \lambda_3 = 9$, $\lambda_4 = 0$. Для каждого собственного значения λ_i построим фундаментальную систему решений однородной системы уравнений $(A - \lambda_i E)x = 0$ и ортонормируем их.

При $\lambda_i = 9$ система $(A - \lambda_i E)x = 0$ имеет вид:

$$\begin{cases} -4x_2 + 4x_3 - 2x_3 = 0, \\ 4x_2 - 4x_3 + 2x_3 = 0, \\ -2x_2 + 2x_3 - x_3 = 0. \end{cases}$$

Ее общее решение $x = (x_1,\ x_2,\ x_3,\ -2x_2 + 2x_3)^{\mathrm{T}}$ имеет три свободных неизвестных, а ФСР состоит из трех векторов, например, $b_1 = (1,\ 0,\ 0,\ 0)^{\mathrm{T}}$, $b_2 = (0,\ 1,\ 2,\ 2)^{\mathrm{T}}$, $b_3 = (0,\ 2,\ 1,\ -2)^{\mathrm{T}}$. Так как векторы b_1, b_2, b_3 оказались попарно ортогональными, то остается лишь их нормировать. После нормирования получим векторы

$$e_1' = (1,\ 0,\ 0,\ 0)^{\mathrm{T}}, \quad e_2' = \frac{1}{3}(0,\ 1,\ 2,\ 2)^{\mathrm{T}}, \quad e_3' = \frac{1}{3}(0,\ 2,\ 1,\ -2)^{\mathrm{T}}.$$

При $\lambda_i = 0$ система $(A - \lambda_i E)x = 0$ имеет вид:

$$\begin{cases} 9x_1 = 0, \\ 5x_2 + 4x_3 - 2x_4 = 0, \\ 4x_2 + 5x_3 + 2x_4 = 0, \\ -2x_2 + 2x_3 + 8x_4 = 0. \end{cases}$$

Ее общее решение $x = (0,\ 2x_4,\ -2x_4,\ x_4)^{\mathrm{T}}$ имеет одно свободное неизвестное, а ФСР состоит из одного решения, например, $b_4 = (0,\ 2,\ -2,\ 1)^{\mathrm{T}}$. Нормируя его, получаем вектор $e_4' = \frac{1}{3}(0,\ 2,\ -2,\ 1)^{\mathrm{T}}$.

Из столбцов координат векторов e_1, e_2, e_3, e_4 составим матрицу

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

и по ней запишем преобразование координат:

$$\begin{cases} x_1 = y_1, \\ x_2 = \frac{1}{3}y_2 + \frac{2}{3}y_3 + \frac{2}{3}y_4, \\ x_3 = \frac{2}{3}2y_2 + \frac{1}{3}y_3 - \frac{2}{3}y_4, \\ x_4 = \frac{2}{3}y_2 - \frac{2}{3}y_3 + \frac{1}{3}y_4. \end{cases}$$

При этом преобразовании координат рассматриваемая квадратичная форма приводится к каноническому виду $9y_1^2 + 9y_2^2 + 9y_3^2$, а уравнение квадрики — к виду

$$9y_1^2 + 9y_2^2 + 9y_3^2 + 18y_1 - 18(y_2 + 2y_3 + 2y_4) + \\ + 18(2y_2 + y_3 - 2y_4) - 18(2y_2 - 2y_3 + y_4) + 207 = 0,$$

или

$$9y_1^2 + 9y_2^2 + 9y_3^2 + 18y_1 - 18y_2 + 18y_3 - 90y_4 + 207 = 0.$$

В полученном уравнении квадрики выделим полные квадраты по переменным y_1, y_2, y_3. Уравнение квадрики при этом примет вид:

$$9(y_1 + 1)^2 + 9(y_2 - 1)^2 + 9(y_3 + 1)^2 - 90y_4 + 180 = 0.$$

Теперь введем новые переменные $z_1 = y_1 + 1$, $z_2 = y_2 - 1$, $z_3 = y_3 + 1$, $z_4 = y_4$, соответствующие преобразованию переменных $y_1 = z_1 - 1$, $y_2 = z_2 + 1$, $y_3 = z_3 - 1$, $y_4 - z_4$. Тогда уравнение квадрики преобразуется в уравнение

$$9z_1^2 + 9z_2^2 + 9z_3^2 - 2(45z_4 - 90) = 0.$$

Далее введем новые переменные $X_1 = z_1$, $X_2 = z_2$, $X_3 = z_3$, $X_4 = \frac{45z_4 - 90}{9} = z_4 - 2$, соответствующие преобразованию переменных $z_1 = X_1$, $z_2 = X_2$, $z_3 = X_3$, $z_4 = X_4 + 2$. В новых координатах рассматриваемая квадрика имеет уравнение

$$9X_1^2 + 9X_2^2 + 9X_3^2 - 90X_4 = 0,$$

или

$$\frac{X_1^2}{5}+\frac{X_2^2}{5}+\frac{X_3^2}{5}-2X_4=0.$$

Из канонического вида квадрики следует, что она является эллиптическим параболоидом.

Результирующее преобразование прямоугольных координат определяется формулами

$$x_1=y_1=z_1+1=X_1+1,$$

$$x_2=\frac{1}{3}y_2+\frac{2}{3}y_3+\frac{2}{3}y_4=\frac{1}{3}z_2+\frac{2}{3}z_3+\frac{2}{3}z_4-1=$$
$$=\frac{1}{3}X_2+\frac{2}{3}X_3+\frac{2}{3}X_4+1,$$

$$x_3=\frac{2}{3}y_2+\frac{1}{3}y_3-\frac{2}{3}y_4=\frac{2}{3}z_2+\frac{1}{3}z_3-\frac{2}{3}z_4+1=$$
$$=\frac{2}{3}X_2+\frac{1}{3}X_3-\frac{2}{3}X_4-1,$$

$$x_4=\frac{2}{3}y_2-\frac{2}{3}y_3+\frac{1}{3}y_4=\frac{2}{3}z_2-\frac{2}{3}z_3+\frac{1}{3}z_4+4=$$
$$=\frac{2}{3}X_2-\frac{2}{3}X_3+\frac{1}{3}X_4+2.$$

Новая система координат $(O',\ e')$ **имеет начало** $O'(1,\ 1,\ -1,\ 2)$ **и базис из векторов**

$$e_1'=\begin{pmatrix}1\\0\\0\\0\end{pmatrix},\quad e_2'=\frac{1}{3}\begin{pmatrix}0\\1\\2\\2\end{pmatrix},\quad e_3'=\frac{1}{3}\begin{pmatrix}0\\2\\1\\-2\end{pmatrix},\quad e_4'=\frac{1}{3}\begin{pmatrix}0\\2\\-2\\1\end{pmatrix}.$$

Пример 12.5. В двумерном евклидовом пространстве преобразованием прямоугольных координат упростить уравнение квадрики

$$x_1^2+x_1x_2+x_2^2-3x_1-6x_2+3=0,$$

указать преобразование прямоугольных координат, осуществляющее такое упрощение уравнения квадрики, и установить ее метрический класс.

Решение. Квадратичная форма $x_1^2 + x_1x_2 + x_2^2$ имеет матрицу

$$A = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1 \end{pmatrix}.$$

Ее характеристический многочлен

$$|A - \lambda E| = \begin{vmatrix} 1-\lambda & 1/2 \\ 1/2 & 1-\lambda \end{vmatrix} = \lambda^2 - 2\lambda + \frac{3}{4}$$

имеет корни $\lambda_1 = 3/2$ и $\lambda_2 = 1/2$. Поэтому рассматриваемая квадратичная форма в главных осях имеет канонический вид

$$\frac{3}{2}y_1^2 + \frac{1}{2}y_2^2.$$

Перейдем к построению матрицы ортогонального преобразования переменных, приводящего рассматриваемую квадратичную форму к каноническому виду в главных осях. Для этого необходимо найти фундаментальные системы решений однородных систем уравнений $(A - \lambda_i E)x = 0$ и ортонормировать их.

При $\lambda_i = 3/2$ система $(A - \lambda_i E)x = 0$ имеет вид:

$$\begin{cases} -x_1 + x_2 = 0, \\ x_1 - x_2 = 0. \end{cases}$$

Ее общим решением является $x = (x_2,\ x_2)^{\mathrm{T}}$, а фундаментальная система решений состоит из одного вектора, например, $b_1 = (1,\ 1)^{\mathrm{T}}$. Нормируя его, получим вектор $e_1' = \frac{1}{\sqrt{2}}(1,\ 1)^{\mathrm{T}}$.

При $\lambda_i = 1/2$ аналогичным образом построим вектор $e_2' = \frac{1}{\sqrt{2}}(-1,\ 1)^{\mathrm{T}}$. Векторы e_1' и e_2' уже ортогональны, так как принадлежат различным собственным значениям. Они составляют канонический ортонормированный базис данной квадратичной формы. Из столбцов их координат строим ортогональную матрицу

$$Q = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

По строкам этой матрицы записываем искомое ортогональное преобразование переменных:

$$x_1 = \frac{1}{\sqrt{2}}y_1 - \frac{1}{\sqrt{2}}y_2, \quad x_2 = \frac{1}{\sqrt{2}}y_1 + \frac{1}{\sqrt{2}}y_2.$$

Используя это преобразование переменных, преобразуем уравнение квадрики:

$$\frac{3}{2}y_1^2 + \frac{1}{2}y_2^2 - \frac{9\sqrt{2}}{2}y_1 - \frac{3\sqrt{2}}{2}y_2 + 3 = 0.$$

Для удобства умножим это уравнение на 2:

$$3y_1^2 + y_2^2 - 9\sqrt{2}y_1 - 3\sqrt{2}y_2 + 6 = 0.$$

Выделим в нем полные квадраты по y_1 и y_2. Тогда получим

$$3\left(y_1 - \frac{3\sqrt{2}}{2}\right)^2 + \left(y_2 - \frac{3\sqrt{2}}{2}\right)^2 - 12 = 0.$$

Вводя переменные

$$X_1 = y_1 - \frac{3\sqrt{2}}{2}, \qquad X_2 = y_2 - \frac{3\sqrt{2}}{2},$$

соответствующие преобразованию переменных

$$y_1 = X_1 + \frac{3\sqrt{2}}{2}, \qquad y_2 = X_2 + \frac{3\sqrt{2}}{2},$$

уравнение квадрики приведем к виду $3X_1^2 + X_2^2 - 12 = 0$, откуда получаем канонический вид

$$\frac{X_1^2}{4} + \frac{X_2^2}{12} = 1.$$

Следовательно, рассматриваемая квадрика на плоскости является эллипсом.

Результирующее преобразование прямоугольных координат определяется формулами

$$x_1 = \frac{1}{\sqrt{2}}y_1 - \frac{1}{\sqrt{2}}y_n = \frac{1}{\sqrt{2}}\left(X_1 + \frac{3\sqrt{2}}{2}\right) - \\ - \frac{1}{\sqrt{2}}\left(X_2 + \frac{3\sqrt{2}}{2}\right) = \frac{1}{\sqrt{2}}X_1 - \frac{1}{\sqrt{2}}X_2,$$

$$x_2 = \frac{1}{\sqrt{2}}y_1 + \frac{1}{\sqrt{2}}y_n = \frac{1}{\sqrt{2}}\left(X_1 + \frac{3\sqrt{2}}{2}\right) + \\ + \frac{1}{\sqrt{2}}\left(X_2 + \frac{3\sqrt{2}}{2}\right) = \frac{1}{\sqrt{2}}X_1 + \frac{1}{\sqrt{2}}X_2 + 3.$$

Новая прямоугольная система координат (O', e_1', e_2') имеет начало $O'(0, 3)$ и координатные векторы

$$e_1' = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad e_2' = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Пример 12.6. В трехмерном пространстве преобразованием прямоугольных координат упростить уравнение квадрики

$$x_1^2 + x_2^2 + 4x_3^2 - 2x_1x_2 + 4x_1x_3 - 4x_2x_3 - 72x_1 + 12 = 0.$$

Решение. Сначала приведем к главным осям квадратичную форму

$$x_1^2 + x_2^2 + 4x_3^2 - 2x_1x_2 + 4x_1x_3 - 4x_2x_3.$$

Матрицей этой квадратичной формы является

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

Ее характеристический многочлен

$$|A-\lambda E| = \begin{vmatrix} 1-\lambda & -1 & 2 \\ -1 & 1-\lambda & -2 \\ 2 & -2 & 4-\lambda \end{vmatrix} = -\lambda^2(\lambda-6)$$

имеет корни $\lambda_1 = 6$, $\lambda_2 = \lambda_3 = 0$. Для каждого собственного значения λ_i построим фундаментальную систему решений системы уравнений $(A-\lambda_i E)x = 0$, а затем все эти ФСР ортонормируем.

Для $\lambda_1 = 6$ система $(A-\lambda_1 E)x = 0$ имеет вид:

$$\begin{cases} -5x_1 - x_2 + 2x_3 = 0, \\ -x_1 - 5x_2 - 2x_3 = 0, \\ 2x_1 - 2x_2 - 2x_3 = 0. \end{cases}$$

Ее общее решение $x = (x_1,\ -x_1,\ 2x_1)^{\mathrm{T}}$ имеет одно свободное неизвестное, а ФСР состоит из одного вектора, например $b_1 = (1,\ -1,\ 2)^{\mathrm{T}}$. Нормируя его, получим: $e_1' = \frac{1}{\sqrt{6}}(1,\ -1,\ 2)^{\mathrm{T}}$.

При $\lambda_2 = 0$ система $(A-\lambda_2 E)x = 0$ имеет вид:

$$\begin{cases} x_1 - x_2 + 2x_3 = 0, \\ -x_1 + x_2 - 2x_3 = 0, \\ 2x_1 - 2x_2 + 4x_3 = 0. \end{cases}$$

Ее общее решение $x = (x_2 - 2x_3,\ x_2,\ x_3)^{\mathrm{T}}$ имеет два свободных неизвестных, а фундаментальная система решений состоит из двух векторов, например $b_2 = (1,\ 1,\ 0)^{\mathrm{T}}$ и $b_3 = (-1,\ 1,\ 1)^{\mathrm{T}}$. Так как b_2 и b_3 уже ортогональны, то остается их лишь нормировать. После нормирования получим

$$e_2' = \frac{1}{\sqrt{2}}(1,\ 1,\ 0)^{\mathrm{T}}, \qquad e_3' = \frac{1}{\sqrt{3}}(-1,\ 1,\ 1)^{\mathrm{T}}.$$

Из столбцов координат векторов e_1', e_2', e_3' составим матрицу

$$Q = \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \end{pmatrix}$$

и по ней запишем преобразование прямоугольных координат

$$\begin{cases} x_1 = \dfrac{1}{\sqrt{6}}y_1 + \dfrac{1}{\sqrt{2}}y_2 - \dfrac{1}{\sqrt{3}}y_3, \\ x_2 = -\dfrac{1}{\sqrt{6}}y_1 + \dfrac{1}{\sqrt{2}}y_2 + \dfrac{1}{\sqrt{3}}y_3, \\ x_3 = \dfrac{2}{\sqrt{6}}y_1 + \dfrac{1}{\sqrt{3}}y_3. \end{cases}$$

При этом преобразовании координат рассматриваемая квадратичная форма приводится к каноническому виду $6y_1^2$, а все уравнение квадрики приводится к виду

$$6y_1^2 - 72\left(\frac{1}{\sqrt{6}}y_1 + \frac{1}{\sqrt{2}}y_2 - \frac{1}{\sqrt{3}}y_3\right) + 12 = 0,$$

что эквивалентно уравнению

$$y_1^2 - 2\sqrt{6}y_1 - 6\sqrt{2}y_2 + 4\sqrt{3}y_3 + 2 = 0.$$

В этом уравнении квадрики выделим полный квадрат по y_1. Уравнение квадрики при этом примет вид:

$$\left(y_1 - \sqrt{6}\right)^2 - 6\sqrt{2}y_2 + 4\sqrt{3}y_3 - 4 = 0.$$

Далее введем переменные $z_1 = y_1 - \sqrt{6}$, $z_2 = y_2$, $z_3 = y_3$, соответствующие преобразованию переменных $y_1 = z_1 + \sqrt{6}$, $z_2 = y_2$, $z_3 = y_3$. Тогда уравнение квадрики преобразуется к виду $z_1^2 - 6\sqrt{2}z_2 + 4\sqrt{3}z_3 - 4 = 0$, или

$$z_1^2 - 2(3\sqrt{2}z_2 - 2\sqrt{3}z_3 + 2) = 0.$$

Введем переменные в соответствии с формулами

$$\begin{aligned} X_1 &= z_1, \\ X_2 &= \frac{3\sqrt{2}z_2 - 2\sqrt{3}z_3 + 2}{\sqrt{(3\sqrt{2})^2 + (2\sqrt{3})^2}} = \frac{3}{\sqrt{15}}z_2 - \frac{2}{\sqrt{10}}z_3 + \frac{2}{\sqrt{30}}, \\ X_3 &= \gamma_1 z_1 + \gamma_2 z_2 + \gamma_3 z_3. \end{aligned}$$

Коэффициенты γ_1, γ_2, γ_3 выберем так, чтобы матрица рассматриваемого преобразования переменных была ортогональной, т.е. чтобы векторы-строки

$$a_1 = (1,\ 0,\ 0), \quad a_2 = \left(0,\ \frac{3}{\sqrt{15}},\ -\frac{2}{\sqrt{10}}\right), \quad a_3 = (\gamma_1,\ \gamma_2,\ \gamma_3)$$

составляли ортонормированную систему векторов. Так как система векторов a_1, a_2 ортонормированная, то координаты вектора a_3 следует искать из условий

$$\begin{cases} (a_1, a_3) = \gamma_1 = 0, \\ (a_2, a_3) = \dfrac{3}{\sqrt{15}}\gamma_2 - \dfrac{2}{\sqrt{10}}\gamma_3 = 0. \end{cases}$$

Затем найденный вектор a_3 нужно еще нормировать.

Проделав это, получим: $\gamma_1 = 0$, $\gamma_2 = \frac{2}{\sqrt{10}}$, $\gamma_3 = \frac{3}{\sqrt{15}}$.

Итак, новые переменные выражаются через старые формулами

$$\begin{aligned} X_1 &= z_1, \\ X_2 &= \frac{3}{\sqrt{15}}z_2 - \frac{2}{\sqrt{10}}z_3 + \frac{2}{\sqrt{30}}, \\ X_3 &= \frac{2}{\sqrt{10}}z_2 + \frac{3}{\sqrt{15}}z_3, \end{aligned}$$

а преобразование переменных имеет вид:

$$\begin{aligned} z_1 &= X_1, \\ z_2 &= \frac{3}{\sqrt{15}}X_2 + \frac{2}{\sqrt{10}}X_3 - \frac{2}{\sqrt{5}}, \\ z_3 &= -\frac{2}{\sqrt{10}}z_2 + \frac{3}{\sqrt{15}}z_3 + \frac{2\sqrt{3}}{15}. \end{aligned}$$

В новых координатах рассматриваемая квадрика имеет каноническое уравнение $X_1^2 = 2\sqrt{30}X_2$. Оно показывает, что эта квадрика представляет собой параболический цилиндр.

Упражнения

12.1. Привести к нормальному виду уравнение квадрики, установить ее вид, выписать формулы преобразования координат, координаты нового начала и новых координатных векторов относительно старой системы координат:

1) $x_1^2 - 2x_1x_2 + x_2^2 + 2x_1 - 5 = 0$;
2) $3x_1^2 + 10x_1x_2 + 3x_2^2 + 2x_1 + 14x_2 - 53 = 0$;
3) $x_1^2 + 4x_1x_2 + 5x_2^2 - 2x_1 - 2x_2 - 2 = 0$;
4) $x_1^2 + x_2^2 + 8x_3^2 - 2x_1x_2 - 6x_1x_3 + 4x_2x_3 + 2x_1 + 2x_2 + 4x_3 - {} - 14 = 0$;
5) $x_1^2 + 2x_2^2 + 2x_3^2 - 2x_1x_2 - 2x_1x_2 - 2x_1 + 8x_3 - 6 = 0$;
6) $4x_1^2 + 6x_2^2 + 5x_3^2 + 4x_1x_3 - 8x_2 - 4x_3 + 3 = 0$;
7) $x_1x_2 + x_2x_3 + x_3x_4 + x_4x_1 + 2x_1 - 4x_2 + 10x_3 + 6x_4 + 8 = 0$;
8) $x_1^2 + 2x_2^2 + 2x_3^2 + x_4^2 + 2x_1x_2 + 2x_1x_3 + 4x_2x_3 + 2x_3x_4 + 2x_1 + {} + 2x_2 + 4x_3 - 8x_4 = 0$;
9) $x_1^2 + 2x_1x_2 + x_2^2 - 2x_3^2 - 4x_3x_4 - 2x_4^2 + 4x_1 - 8x_2 + 6x_3 - {} - 10x_4 + 10 = 0$.

12.2. Преобразованием прямоугольных кординат привести к каноническому виду уравнение квадрики, записать преобразование координат, координаты нового начала и новых координатных векторов относительно старой системы координат:

1) $5x_1^2 + 8x_1x_2 + 5x_2^2 - 18x_1 - 18x_2 + 9 = 0$;
2) $x_1^2 - 2x_1x_2 + x_2^2 + 2x_1 - 2x_2 - 3 = 0$;
3) $3x_1^2 + 4x_1x_2 + 2x_2^2 - 4x_2 - 9 = 0$;
4) $2x_1^2 + 3x_2^2 + x_3^2 - 4x_1x_2 + 4x_1x_3 - 12x_1 - 6x_3 + 4 = 0$;
5) $x_1^2 - 2x_2^2 + x_3^2 + 4x_1x_2 - 8x_1x_3 - 4x_2x_3 + 2x_1 - 4x_3 - 1 = 0$;
6) $8x_1^2 + 4x_2^2 + 5x_3^2 - 4x_1x_3 - 44x_1 - 2x_3 + 29 = 0$;
7) $2x_1x_2 - 6x_1x_3 - 6x_2x_4 + 2x_3x_4 + 4x_1 - 8x_2 + 8x_3 - 4x_4 - 9 = {} = 0$;
8) $5x_1^2 + 5x_2^2 + 5x_3^2 + 5x_4^2 - 10x_1x_2 + 2x_1x_3 + 6x_1x_4 + 6x_2x_3 + {} + 2x_2x_4 - 10x_3x_4 + 24x_3 - 24x_4 - 18 = 0$;
9) $3x_1^2 + 8x_1x_2 - 3x_2^2 + 4x_3^2 - 4x_3x_4 + x_4^2 + 10\sqrt{5}x_3 - 20\sqrt{5}x_4 - {} - 14 = 0$.

Приложение. ВЫЧИСЛЕНИЕ ХАРАКТЕРИСТИЧЕСКОГО МНОГОЧЛЕНА

Характеристическим многочленом матрицы A является определитель $|A - \lambda E|$ характеристической матрицы $A - \lambda E$. Непосредственное вычисление этого определителя для сколь-нибудь высокого порядка матрицы громоздко и весьма трудоемко. Поэтому разработаны специальнын методы развертывания определителя характеристической матрицы, не требующие его непосредственного вычисления (методы А.Н. Крылова, А.М. Данилевского, Леверье, Д.К. Фаддеева, интерполяционный метод и др.). О таких методах см. [10], [15], [31], [32]. Кратко изложим некоторые из этих методов.

1. Метод А.Н. Крылова. Пусть

$$|A-\lambda E|=(-1)^n(\lambda^n+p_1\lambda^{n-1}+p_2\lambda^{n-2}+\ldots+p_{n-1}\lambda+p_n) \quad (12.1)$$

— характеристический многочлен матрицы A n-го порядка. В силу теоремы 5.9 (Гамильтона — Кэли) будем иметь:

$$A^n+p_1A^{n-1}+p_2A^{n-2}+\ldots+p_{n-1}A+p_nE=0.$$

Умножив обе части этого равенства справа на произвольный вектор $x^{(0)}$, получим:

$$A^nx^{(0)}+p_1A^{n-1}x^{(0)}+p_2A^{n-2}x^{(0)}+\ldots$$
$$\ldots+p_{n-1}Ax^{(0)}+p_nx^{(0)}=0. \quad (12.2)$$

Полагая $x^{(k)}=A^kx^{(0)}$, $k=1,2,\ldots,n$, запишем равенство (12.2) в виде:

$$x^{(n)}+p_1x^{(n-1)}+\ldots+p_{n-1}x^{(1)}+p_nx^{(0)}=0. \quad (12.3)$$

В силу соотношений $x^{(k)}=A^kx^{(0)}$, $k=1,2,\ldots,n$ выполняются равенства

$$x^{(k)}=Ax^{(k-1)}, \quad k=1,2,\ldots,n, \quad (12.4)$$

которые удобно использовать для последовательного вычисления векторов $x^{(1)}$, $x^{(2)}$, ..., $x^{(n)}$.

Задав вектор $x^{(0)}$, мы тем самым определим все векторы $x^{(1)}$, $x^{(2)}$, ..., $x^{(n)}$, что позволяет рассматривать равенство (12.3) как систему линейных уравнений с неизвестными p_1, p_2, ..., p_n, записанную в векторной форме (см. разд. 3.3). Если эта система имеет единственное решение, то это решение представляет собой набор коэффициентов p_1, p_2, ..., p_n характеристического многочлена матрицы A. Если указанная система уравнений имеет бесконечно много решений, то следует изменить начальный вектор $x^{(0)}$ и повторить процесс вычислений.

Пример. Методом А.Н. Крылова найдем характеристический многочлен матрицы

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

Решение. Записав характеристический многочлен в виде

$$|A - \lambda E| = \lambda^4 + p_1\lambda^3 + p_2\lambda^2 + p_3\lambda + p_4,$$

для определения коэффициентов p_1, p_2, p_3, p_4 выберем в качестве начального вектор $x^{(0)} = (1,\ 2,\ 3,\ 0)^{\mathrm{T}}$ и по формулам (12.4) найдем векторы

$$x^{(1)} = \begin{pmatrix} -5 \\ -12 \\ 9 \\ 6 \end{pmatrix}, \quad x^{(2)} = \begin{pmatrix} 45 \\ 46 \\ 15 \\ 12 \end{pmatrix}, \quad x^{(3)} = \begin{pmatrix} -25 \\ -32 \\ 21 \\ 18 \end{pmatrix}, \quad x^{(4)} = \begin{pmatrix} 89 \\ 90 \\ 27 \\ 24 \end{pmatrix}.$$

Далее, используя найденные векторы $x^{(i)}$, $i = 0, 1, 2, 3, 4$, запишем векторное уравнение (12.3):

$$\begin{pmatrix} 89 \\ 90 \\ 27 \\ 24 \end{pmatrix} + p_1 \begin{pmatrix} -25 \\ -32 \\ 21 \\ 18 \end{pmatrix} + p_2 \begin{pmatrix} 45 \\ 46 \\ 15 \\ 12 \end{pmatrix} + p_3 \begin{pmatrix} -5 \\ -12 \\ 9 \\ 6 \end{pmatrix} + p_4 \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Это равенство является векторной формой записи системы линейных уравнений, имеющей единственное решение $p_1 = 0$, $p_2 = -2$, $p_3 = 0$, $p_4 = 1$. *Поэтому характеристический многочлен* матрицы A имеет вид:

$$|A - \lambda E| = \lambda^4 - 2\lambda^2 + 1.$$

2. Интерполяционный метод. Интерполяционный метод — это общий метод вычисления определителей, элементами которых являются многочлены одного переменного, в том числе и определителя характеристической матрицы (12.1).

Для вычисления характеристического многочлена (12.1) выберем n различных значений $\lambda_1, \lambda_2, \ldots, \lambda_n$ переменного λ и поочередно подставим в обе части равенства (12.1). В результате получим систему линейных уравнений

$$\begin{cases} (-1)^n\left(\lambda_1^n + p_1\lambda_1^{n-1} + \ldots + p_{n-1}\lambda_1 + p_n\right) = |A - \lambda_1 E|, \\ (-1)^n\left(\lambda_2^n + p_1\lambda_2^{n-1} + \ldots + p_{n-1}\lambda_2 + p_n\right) = |A - \lambda_2 E|, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ (-1)^n\left(\lambda_n^n + p_1\lambda_n^{n-1} + \ldots + p_{n-1}\lambda_n + p_n\right) = |A - \lambda_n E| \end{cases} \quad (12.5)$$

относительно неизвестных $p_1, p_2, \ldots, p_n$. Определителем этой системы является определитель Вандермонда, отличный от нуля, поскольку все значения $\lambda_1, \lambda_2, \ldots, \lambda_n$ различны. Следовательно, система (12.5) имеет единственное решение, определяющее коэффициенты характеристического многочлена.

Пример. Интерполяционным методом найдем характеристический многочлен матрицы

$$A = \begin{pmatrix} 1 & -3 & 0 & 3 \\ -2 & -6 & 0 & 13 \\ 0 & -3 & 1 & 3 \\ -1 & -4 & 0 & 8 \end{pmatrix}.$$

Решение. Для определения коэффициентов p_1, p_2, p_3, p_4 характеристического многочлена $|A-\lambda E| = \lambda^4 + p_1\lambda^3 + p_2\lambda^2 + p_3\lambda + p_4$ выберем четыре значения $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = 3$, $\lambda_4 = 4$ и для каждого из них вычислим соответствующий определитель:

$$|A-\lambda_1 E| = |A-E| = 0, \quad |A-\lambda_2 E| = |A-2E| = 1,$$

$$|A-\lambda_3 E| = |A-3E| = 16, \quad |A-\lambda_4 E| = |A-4E| = 81.$$

Подставляя выбранные значения λ_1, λ_2, λ_3, λ_4 и вычисленные определители в (12.5), получим систему линейных уравнений

$$\begin{cases} 1+p_1+p_2+p_3+p_4 = 0, \\ 16+8p_1+4p_2+2p_3+p_4 = 1, \\ 81+27p_1+9p_2+3p_3+p_4 = 16, \\ 256+64p_1+16p_2+4p_3+p_4 = 81, \end{cases}$$

имеющую единственное решение $p_1 = -4$, $p_2 = 6$, $p_3 = -4$, $p_4 = 1$. Следовательно, характеристический многочлен матрицы A имеет вид:

$$|A-\lambda E| = \lambda^4 - 4\lambda^3 + 6\lambda^2 - 4\lambda + 1.$$

3. Метод Д.К. Фаддеева. Этот метод для вычисления характеристического многочлена (12.1) матрицы A, записанного в виде

$$|A-\lambda E| = (-1)^n(\lambda^n - p_1\lambda^{n-1} - p_2\lambda^{n-2} - \ldots - p_{n-1}\lambda - p_n), \quad (12.6)$$

состоит в проведении следующих вычислений:

$$A_i = B_{i-1}A, \quad p_i = \frac{\mathrm{Sp}\,A_i}{i}, \quad B_i = A_i - p_i E, \quad i = 1, 2, \ldots, n, \quad (12.7)$$

где $B_0 = E$ (т.е. $A_1 = A$), а $\mathrm{Sp}\,A_i$ — след матрицы A_i, по определению равный сумме ее диагональных элементов.

Для обоснования приведенных формул запишем характеристический многочлен матрицы A в виде

$$|A-\lambda E| = (-1)^n(\lambda^n - c_1\lambda^{n-1} - c_2\lambda^{n-2} - \ldots - c_{n-1}\lambda - c_n),$$

и покажем, что коэффициенты c_k этого многочлена совпадают со значениями p_k, полученными по формулам (12.7). Доказательство проведем методом математической индукции по индексу чисел c_k и p_k. Известно, что коэффициент характеристического многочлена матрицы A при степени $n-1$ равен $(-1)^n \operatorname{Sp} A$. Следовательно, $c_1 = \operatorname{Sp} A = p_1$. Пусть уже доказано, что $p_i = c_i$ при $i = 1, \ldots, k-1$. Покажем, что $p_k = c_k$.

В силу построения матриц A_k в соответствии с формулами (12.7) и индуктивного предположения имеем:

$$A_k = A^k - p_1A^{k-1} - p_2A^{k-2} - \ldots - p_{k-1}A = \\ = A^k - c_1A^{k-1} - c_2A^{k-2} - \ldots - c_{k-1}A.$$

Поэтому

$$kp_k = \operatorname{Sp} A_k = \operatorname{Sp} A^{k-1} - c_1 \operatorname{Sp} A^{k-1} - \ldots - c_{k-1} \operatorname{Sp} A = \\ = S_k - c_1S_{k-1} - \ldots - c_{k-1}S_1,$$

где S_1, S_2, …, S_k — степенные суммы от характеристических чисел матрицы A. Но по формулам Ньютона $S_k - c_1S_{k-1} - \ldots - c_{k-1}S_1 = kc_k$. Следовательно, $kp_k = kc_k$ и $c_k = p_k$.

В заключение заметим, что в силу теоремы 5.9 (Гамильтона — Кэли)

$$B_n = A_n - p_nE = A^n - c_1A^{n-1} - \ldots - c_{n-1}A - c_nE = 0.$$

Равенство $B_n = 0$ может служить для контроля правильности вычислений.

Пример. Методом Д.К. Фаддеева вычислим характеристический многочлен матрицы

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

Решение. На первом шаге по формулам (12.7) находим:

$$A_1 = A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 0 \\ -2 & -2 & -1 \end{pmatrix}, \qquad p_1 = \operatorname{Sp} A_1 = 2,$$

$$B_1 = A_1 - p_1 E = \begin{pmatrix} -1 & 2 & 0 \\ 0 & 0 & 0 \\ -2 & -2 & -3 \end{pmatrix}.$$

Используя найденную матрицу B_1, на втором шаге получаем:

$$A_2 = AB_1 = \begin{pmatrix} -1 & 2 & 0 \\ 0 & 0 & 0 \\ 4 & -2 & 3 \end{pmatrix}, \qquad p_2 = \frac{\operatorname{Sp} A_2}{2} = 1,$$

$$B_2 = A_2 - p_2 E = \begin{pmatrix} -2 & 2 & 0 \\ 0 & -1 & 0 \\ 4 & -2 & 2 \end{pmatrix}.$$

Наконец, используя матрицу B_2, на третьем шаге вычислений определяем:

$$A_3 = AB_2 = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix}, \qquad p_3 = \frac{\operatorname{Sp} A_3}{3} = -2,$$

$$B_3 = A_3 - p_3 E = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Итак,

$$|A - \lambda E| = (-1)^3(\lambda^3 - 2\lambda^2 - \lambda + 2) = -\lambda^3 + 2\lambda^2 + \lambda - 2.$$

ЛИТЕРАТУРА

1. *Александров П.С.* Курс аналитической геометрии и линейной алгебры. – М.: Наука, 1979.
2. *Беклемишев Д.В.* Курс аналитической геометрии и линейной алгебры. – М.: Наука, 1984.
3. *Беклемишев Д.В.* Дополнительные главы линейной алгебры. – М.: Наука, 1983.
4. *Воеводин В.В.* Линейная алгебра. – М.: Наука, 1980.
5. *Воеводин В.В.* Вычислительные основы линейной алгебры. – М.: Наука, 1977.
6. *Воеводин В.В., Кузнецов В.А.* Матрицы и вычисления. – М.: Наука, 1984.
7. *Гантмахер Ф.Д.* Теория матриц. – М.: Наука, 1967.
8. *Гельфанд И.М.* Лекции по линейной алгебре. – М.: Наука, 1971.
9. *Головина Л.И.* Линейная алгебра и некоторые ее приложения. – М.: Наука, 1975.
10. *Демидович Б.И., Марон И.А.* Основы вычислительной математики. – М.: Наука, 1970.
11. *Ефимов Н.В., Розендорн Э.Р.* Линейная алгебра и многомерная геометрия. – М.: Наука, 1974.
12. *Загускин В.Л.* Справочник по численным методам решения уравнений. – М.: Физматгиз, 1960.
13. *Икрамов Х.Д.* Задачник по линейной алгебре. – М.: Наука, 1975.
14. *Ильин В.А., Позняк Э.Г.* Линейная алгебра. – М.: Наука, 1974.
15. *Крылов В.И. и др.* Вычислительные методы / В.И. Крылов, В.В. Бобков, П.И. Монастырский. – М.: Наука, 1976. Т.1.
16. *Курош А.Г.* Курс высшей алгебры. – М.: Физматгиз, 1962.
17. *Ланкастер П.* Теория матриц. – М.: Наука, 1978.
18. *Лоусон Ч., Хенсон Р.* Численное решение задач метода наименьших квадратов. – М.: Наука, 1986.
19. *Мальцев А.И.* Основы линейной алгебры. – М.: Наука, 1970.

20. Математический практикум / Г.Н. Положий, Н.А. Пахарева, И.З. Степаненко и др. – М.: Физматгиз, 1960.

21. *Проскуряков И.В.* Сборник задач по линейной алгебре. – М.: Наука, 1978.

22. *Розенфельд Б.А.* Многомерные пространства. – М.: Физматгиз, 1966.

23. *Рублев А.Н.* Курс линейной алгебры и аналитической геометрии. – М.: Высшая школа, 1972.

24. *Самарский А.А.*, *Гулин А.В.* Численные методы. – М.: Наука, 1989.

25. *Сесекин Н.Ф.* Основы линейной алгебры. – Свердловск, 1987.

26. *Сесекин Н.Ф.* Линейные алгебраические системы. – Свердловск, 1987.

27. *Стренг Г.* Линейная алгебра и ее приложения. – М.: Мир, 1980.

28. *Тышкевич Р.И.*, *Феденко А.С.* Линейная алгебра и аналитическая геометрия. – Минск: Вышэйшая школа, 1968.

29. *Уилкинсон Дж.*, *Райнш К.* Справочник алгоритмов на языке АЛГОЛ. Линейная алгебра. – М.: Машиностроение, 1976.

30. *Фаддеев Д.К.* Лекции по линейной алгебре. – М.: Наука, 1984.

31. *Фаддеев Д.К.*, *Соминский И.С.* Сборник задач по высшей алгебре. – М.: Наука, 1982.

32. *Фаддеев Д.К.*, *Фаддеева В.Н.* Вычислительные методы линейной алгебры. – М.–Л.: Физматгиз, 1963.

33. *Форсайт Дж.*, *Малькольм М.*, *Моулер К.* Машинные методы математических вычислений. – М.: Мир, 1980.

34. *Хорн Р.*, *Джонсон Ч.* Матричный анализ. – М.: Мир, 1989.

35. *Шевцов Г.С.* Линейная алгебра. – М.: Гардарики, 1999.

36. *Шилов Г.Е.* Математический анализ. Конечномерные линейные пространства. – М.: Наука, 1969.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

Аксиомы линейного пространства 81
– нормы 413
– скалярного произведения 266, 293
Алгебраическая кратность 181, 191
– операция 5
– – ассоциативная 6
– – коммутативная 6
Алгебраическое дополнение 32
Арифметический корень из неотрицательного оператора 313, 346
– – – неотрицательной матрицы 315, 346
Арифметическое пространство 81
Аффинное пространство 525
– – действительное конечномерное 526
– – n-мерное 525

Базис арифметического пространства естественный 106
– жорданов 217, 222
– корневой 217
– линейного пространства 105
– ортогональный 281
– ортонормированный 281
– системы векторов 90
Базисная подсистема 125
Базисные столбцы 94
– строки 95
Базисы сингулярные 364

Ведущий элемент 18
Вектор 80
– единичный 272
– пропорциональный 83
Векторы ортогональные 274
Вращение 327
Вторая форма сингулярного разложения 365
Высота вектора 216
Вычитание векторов 82
– матриц 52

Геометрическая кратность 196
Гиперболоид 534
Гиперплоскость 529
Гиперповерхность второго порядка 531, 541
Главная диагональ 49
Главное число 363
Группа 7
– абелева 7
– аддитивная 7
– конечная 7
– мультипликативная 7

Детерминант 24
Дефект линейного оператора 163
Дискриминант квадратичной формы 447
Длина вектора 272

Евклидово пространство 268, 539
Евклидовы пространства изоморфные 291

Жорданова клетка 221
– форма 215, 222
– цепочка 216
Жорданов базис *см. Базис жорданов*

Закон дистрибутивности 8
– инерции 457

Инвариантное подпространство оператора 193
Инвариантный множитель матрицы 241
Инверсия 13
Интерполяционный многочлен Лагранжа — Сильвестра 249

Канонические уравнения прямой 530
Канонический базис 451
– вид квадратичной формы 451
– – матрицы 241
Каноническое разложение *см. Разложение матрицы каноническое*
Квадратичная форма действительная 446
– – знакопеременная 459
– – неопределенная 459
– – неотрицательная 459
– – неположительная 459
– – от комплексных переменных 477
– – отрицательно знакопостоянная 459
– – – определенная 459
– – – полуопределенная 459
– – положительно знакопостоянная 459
Квадратичная форма положительно определенная 459
– – – полуопределенная 459
– – распадающаяся 466
– – эрмитова 477
Квадрика 531
Кольцо 8
– коммутативное 8
Компонента строки (столбца) 48
Компоненты матрицы 254
Конус 533, 534
– действительный 534
– мнимый 534
Координатное пространство 81
Координатный столбец 109
Координаты вектора 108
– точки 526
Корень многочлена матричный 183
Корневое подпространство оператора 215
Коэффициенты линейной комбинации 54
Критерий Сильвестра 459

Линейная зависимость векторов 83
– комбинация системы векторов 83
– – столбцов 53
– оболочка 139
Линейное подпространство 138
– преобразование переменных 448
– – – невырожденное 449
– пространство 80
– – бесконечномерное 104
– – действительное 81

Линейное пространство
  комплексное 81
--- конечномерное 104
-- n-мерное 106
Линейные пространства
  изоморфные 119
Линия второго порядка 531

Матрица 48
- вращения 328
- вырожденная 60
- Гивенса 328
- Грама 269
- действительная
  (вещественная) 48
- дефектная 198
- диагональная 49
- единичная 49
- квадратичной формы 447, 477
- квадратная 49
- квазидиагональная 222
- комплексная 48
- комплексно сопряженная 50
- кососимметрическая 50
- левая нижняя 49
-- (нижняя) треугольная 49
- линейного оператора 165
- невырожденная 60
- неособенная 60
- неотрицательная 315
- нормальная 351
- нулевая 49
- обратная 64
- ортогональная 283
- особенная 60
- отражения 330
- перестановки 62
- перехода 110
- положительно определенная
  269, 315

Матрица правая верхняя 49
-- (верхняя) треугольная 49
- присоединенная 65
- простая 198
- прямоугольная 48
- псевдообратная 381
- симметрическая 50
- системы линейных уравнений
  16
---- расширенная 17
- скалярная 49
- сопряженная 294
- союзная 65
- транспонированная 50
- трансформирующая 172, 223
- треугольная 49
- унитарная 298
-- элементарная 349
- характеристическая 180
- элементарная 62
- эрмитова 50, 294
- эрмитово-сопряженная 50
Матрицы перестановочные 57
- подобные 172
- равные 48
- эквивалентные 61, 64
Метод верхней релаксации 499
- вращений 501
- Гаусса 18
- Зейделя 492
- исчерпывания 517
- итераций 488
- Лагранжа 452
- наименьших квадратов 394
- окаймления 95
- последовательного
  исключения неизвестных 18
- последовательных
  приближений 488
- регуляризации 409

Метод скалярных произведений 520
– Якоби 501
Минор базисный 94
– главный 31
– дополнительный 32
– матрицы 91
– определителя 31
Многочлен, аннулируемый матрицей 183
– минимальный 185
Множество 5
– пустое 5
Модуль оператора 318

Направляющее подпространство плоскости 529
Начало координат 526
Начальная точка плоскости 529
Неизвестные главные 126
– свободные 126
Неравенство Коши — Буняковского 273
– треугольника 274
Норма 412
– евклидова 413
– кубическая 413
– октаэдрическая 413
– спектральная 415
– сферическая 413
Нормальный вид квадратичной формы 456
Норма матрицы индуцированная 415
– – обобщенная 414
– – согласованная 415
Нормирование вектора 272
Нормированное пространство 413

Область значений оператора 162
Образ вектора 161
Обратная операция 6
Обратный элемент 6
Общие уравнения плоскости 530
– – подпространства 142
Оператор 161
– вращения 327
– линейный 161
– – невырожденный 163, 163
– неотрицательный 312, 345
– *нормальный 350*
– нулевой 163
– обратный 179
– ортогональный 318
– отражения 327
– подобия 163
– положительно определенный 312, 345
– простого спектра 202
– простой структуры 199
– самосопряженный 305, 343
– симметрический 305
– сопряженный 299, 340
– тождественный 163
– унитарный 347
Операторы взаимно сопряженные 300
– линейные равные 163
– равные 175
Оператор эрмитов 343
Операция вычитания 81
Определитель 24
Определяющие векторы 330
Определяющий многочлен 249
Орт 272
Ортогонализация 276
Ортогональная составляющая 288

Ортогональное дополнение 285
Основная теорема о линейной зависимости векторов 88
Отражение 330
Отрицательные квадраты 457
Отрицательный индекс инерции 458

Параболоид 535
Параметрические уравнения подпространства 145
Параметрическое уравнение плоскости 529
Пересечение подпространств 146
Перестановка 12
– нечетная 13
– четная 13
Плоскость k-мерная 529
Подгруппа 8
Подстановка 15
– четная 15
Поле 8
– основное 81
Положительные квадраты 457
Положительный индекс инерции 458
Полярное разложение *см. Разложение матрицы полярное*
Порядок группы 7
Правило Крамера 43
Преобразование Хаусхолдера 331
Приведение к главным осям 469
Приведенное уравнение квадрики второго рода 543
Присоединенный вектор 216
Проекция на подпространство 146
– – – ортогональная 288
Произведение 5
– матриц 56
– матрицы на число 51
– операторов 177
– столбца на число 53
– элемента на число 80
Прообраз вектора 161
Противоположный элемент 7
Процесс итераций 488
– ортогонализации 276
Прямая 529
– сумма 146
Прямое дополнение 146
Псевдорешение 394
– нормальное 395

Радиус-вектор точки 526
Разложение вектора по базису 108
– матрицы каноническое 203
– – полярное 374
– – скелетное 101
– – спектральное 203
– – QR 354
– спектральное функции от матрицы 254
Размерность линейного пространства 106
Разность матриц 52
– элементов 81
Ранг квадратичной формы 447, 450
– линейного оператора 162
– матрицы 91
– системы векторов 90
– – линейных уравнений 125
Репер 526

Решение обобщенное 394
– системы линейных уравнений 124
– – общее 17
– – частное 17

Свободный вектор 525
– член уравнения 16
Сигнатура квадратичной формы 458
Симметрический вид квадратичной формы 447
Сингулярное разложение 365
– число 363
Сингулярные векторы 364
Система векторов линейно зависимая 84
– – – независимая 84
– – – – максимальная 87
– – ортогональная 274
– – ортонормированная 274
– координат 526
– – прямоугольная 539
– крамеровская 41
– линейных уравнений несовместная 124
– – – приведенная 489
– – – совместная 124
– неоднородная 17
– неопределенная 17
– несовместная 17
– нормальных уравнений **396**
– однородная 16
– – приведенная 136
– определенная 17
– совместная 17
Системы векторов эквивалентные 88
– линейных уравнений эквивалентные 17, 124

Скалярное произведение 266, 293
– умножение 266, 293
Скелетное разложение *см.* *Разложение матрицы скелетное*
Сложение 5
– матриц 52
Собственное значение 190
– подпространство оператора 193
Собственный вектор 190
– – левый 199
– – правый 199
Спектральное разложение *см.* *Разложение матрицы спектральное*
Спектр матрицы 181
– – простой 181
Степенной метод 514
Столбец 48
Строка 48
Сумма 5
– матриц 51
– операторов 175
– подпространств 146
– – прямая 146
– столбцов 53
– элементов 80

Текущая точка плоскости 529
Теорема Гамильтона — Кэли 184
– Кронекера — Капелли 124
– Лапласа 36
– о базисном миноре 94
– – рангах двух систем векторов 90
– – ранге матрицы 92

Точечно-векторное пространство 525
– – евклидово 539
Транспозиция 13
Транспонирование матрицы 50
– определителя 26
Трансформирование 173

Угол между векторами 272
Умножение 5
– матриц 56
Унитарное пространство 293
Уравнения базисные 125

Формулы Крамера 42
– преобразования координат 116
Фундаментальная система решений (ФСР) 130
Функция, определенная на спектре 251

Характеристический корень 181
– – линейного оператора 189
Характеристический многочлен 180
– – линейного оператора 189

Циклическое подпространство 217
Цилиндр гиперболический 535
– параболический 535
– эллиптический 535

Число обусловленности 417

Элементарное преобразование 17, 61
Элемент матрицы 48
– – диагональный 49
– множества 5
– нулевой 80
Эллипсоид 534
– мнимый 534

Ядро линейного оператора 162

λ-матрица 184
LU-разложение 74
$(m \times n)$-матрица 48
QR-алгоритм 509
QR-разложение *см.* *Разложение матрицы* QR

ОГЛАВЛЕНИЕ

Предисловие 3

Глава 1. Первоначальные сведения 5

1.1. Множества, алгебраические операции, группы, кольца, поля 5

1.2. Простые и двойные суммы 10

1.3. Перестановки и подстановки 12

Глава 2. Системы линейных уравнений. Определители 16

2.1. Метод последовательного исключения неизвестных (метод Гаусса) 16

2.2. Определители 24

2.3. Свойства определителей 26

2.4. Миноры и алгебраические дополнения 31

2.5. Разложение определителя по строке или столбцу. Теорема Лапласа 35

2.6. Вычисление определителей 38

2.7. Крамеровские системы 41

Упражнения 45

Глава 3. Матрицы и действия над ними 48

3.1. Первоначальные сведения о матрицах 48

3.2. Сложение матриц и умножение матрицы на число . . 51

3.3. Линейные комбинации столбцов (строк) 52

3.4. Умножение матриц 56

3.5. Элементарные преобразования над матрицами и элементарные матрицы 61

3.6. Обратная матрица 64

3.7. Простейшие матричные уравнения 72

3.8. Разложение квадратной матрицы на треугольные множители 74

Упражнения 77

Глава 4. Линейные пространства 80

4.1. Определение линейного пространства 80

4.2. Линейная зависимость векторов 83

4.3. Ранг матрицы. Скелетное разложение матрицы . . . 91

4.4. Базис и размерность пространства 104

4.5. Связь между базисами линейного пространства . . . 110

4.6. Преобразование координат вектора при переходе от базиса к базису 115

4.7. Изоморфизм линейных пространств 119

4.8. Системы линейных уравнений (общая теория) 123

4.9. Однородные системы линейных уравнений 129

4.10. Связь между решениями однородной и неоднородной систем 136

4.11. Линейные подпространства 138

Упражнения 154

Глава 5. Линейные операторы в линейных пространствах 161

5.1. Определение и примеры линейных операторов 161

5.2. Линейные операторы и матрицы 165

5.3. Выражение координат вектора-образа через координаты вектора-прообраза 169

5.4. Связь между матрицами линейного оператора в разных базисах 171

5.5. Действия с линейными операторами 175

5.6. Характеристический и минимальный многочлены . . 180

5.7. Собственные векторы и собственные значения линейного оператора 190

5.8. Линейные операторы простой структуры 199

Упражнения 209

Глава 6. Каноническая жорданова форма матрицы . . 215

6.1. Жорданов базис 215

6.2. Построение жорданова базиса и жордановой матрицы 223

6.3. Второй способ построения жордановой и трансформирующей матриц 236

6.4. Третий способ построения жордановой и трансформирующей матриц 240

6.5. К построению минимального многочлена 246

Упражнения 248

Глава 7. Функции от матриц 249

7.1. Интерполяционный многочлен Лагранжа — Сильвестра 249

7.2. Функции от матриц 251

7.3. Спектральное разложение матрицы $f(A)$ 254

7.4. Представление функций от матриц рядами 258

7.5. Некоторые приложения функций от матриц 259

Упражнения 264

Глава 8. Евклидовы и унитарные пространства 266
8.1. Определение евклидова пространства. Матрица Грама . . . 266
8.2. Длины и углы. Ортогональность. Процесс ортогонализации . . . 272
8.3. Ортонормированные базисы . . . 281
8.4. Ортогональные матрицы . . . 283
8.5. Ортогональное дополнение. Ортогональная проекция вектора на подпространство . . . 285
8.6. Изоморфизм евклидовых пространств . . . 291
8.7. Понятие об унитарном пространстве . . . 293
8.8. Сопряженные операторы в евклидовом пространстве 298
8.9. Симметрические (самосопряженные) операторы . . . 305
8.10. Ортогональные операторы . . . 318
8.11. Произвольные линейные операторы в евклидовом пространстве . . . 338
8.12. Сопряженные операторы в унитарном пространстве 340
8.13. Эрмитовы операторы . . . 343
8.14. Унитарные операторы . . . 347
8.15. Нормальные операторы . . . 350
8.16. Произвольные линейные операторы в унитарном пространстве . . . 354
8.17. QR-разложение матрицы . . . 354
8.18. Сингулярное разложение матрицы . . . 362
8.19. Полярное разложение матрицы . . . 374
8.20. Псевдообратная матрица . . . 381
8.21. Решение систем линейных уравнений методом наименьших квадратов . . . 393
8.22. Метод регуляризации для систем линейных уравнений 409
8.23. Нормы векторов и матриц . . . 412
8.24. Оценка погрешности решения системы линейных уравнений . . . 416
8.25. Отыскание устойчивого решения системы линейных уравнений . . . 420
8.26. Рекомендации к решению систем линейных уравнений на ЭВМ . . . 426
Упражнения . . . 433
Глава 9. Квадратичные формы . . . 446
9.1. Определение квадратичной формы . . . 446
9.2. Линейное преобразование переменных . . . 448
9.3. Преобразование квадратичной формы при линейном преобразовании переменных . . . 449

9.4. Приведение квадратичной формы к каноническому виду 451
9.5. Закон инерции квадратичных форм 457
9.6. Знакоопределенные квадратичные формы 459
9.7. Распадающиеся квадратичные формы 466
9.8. Квадратичные формы в евклидовом пространстве . 468
9.9. Пары квадратичных форм 472
9.10. Квадратичные формы в комплексном линейном пространстве 476
Упражнения 485

Глава 10. Итерационные методы решения систем линейных уравнений 488
10.1. Метод итераций 488
10.2. Метод Зейделя 492
10.3. Приведение линейной системы к виду, удобному для итераций 494
Упражнения 499

Глава 11. О приближенных методах вычисления собственных значений и собственных векторов . . 501
11.1. Метод вращений (метод Якоби) 501
11.2. QR-алгоритм 509
11.3. Степенной метод 514
11.4. Метод скалярных произведений 520
Упражнения 522

Глава 12. Элементы n-мерной аналитической геометрии 525
12.1. Аффинные пространства 525
12.2. Координаты в аффинном пространстве 526
12.3. Плоскости в аффинном пространстве 528
12.4. Гиперповерхности второго порядка 531
12.5. Точечно-векторное евклидово пространство 539
12.6. Гиперповерхности второго порядка в евклидовом пространстве 541
Упражнения 556

Приложение. Вычисление характеристического многочлена 557

Литература 563

Предметный указатель 565

Цель учебного пособия

«Линейная алгебра: теория и прикладные аспекты»

- помочь глубже усвоить идеи и методы предмета;
- показать, как использовать эти идеи и методы для решения практических задач;
- оказать помощь студентам при выполнении лабораторных и практических заданий;
- служить справочником для специалистов, применяющих методы линейной алгебры в практической деятельности.

ISBN 5-279-02557-7

9 785279 025572