

МИНОБРНАУКИ РОССИИ

Санкт-Петербургский государственный электротехнический
университет „ЛЭТИ“

ЛИНЕЙНАЯ АЛГЕБРА

Учебное электронное пособие

Санкт-Петербург
Издательство СПбГЭТУ „ЛЭТИ“
2012

УДК 512.8(07)
ББК В 143я7
Л 59

Авторы: А. Л. Белопольский, Н. А. Бодунов, А. Л. Меркулов,
А. П. Щеглова.
Л 59 Линейная алгебра: Учеб. эл. пособие. СПб.: Изд-во СПбГЭТУ
„ЛЭТИ“, 2012. 140 с.

ISBN 978–5–7629–1250–1

Излагаются элементы теории матриц, определителей, линейных пространств и их приложений к теории систем линейных уравнений и аналитической геометрии, входящие в программу курса алгебры и геометрии. Является первой частью дисциплины „Алгебра и геометрия“.

Предназначено для студентов факультетов электротехники и автоматизации, электроники, экономики и менеджмента и открытого факультета.

УДК 512.8(07)
ББК В 143я7

Рецензенты: кафедра высшей математики СПбГПУ;
д-р физ.-мат. наук Н. В. Смородина (СПбГУ).

Утверждено
редакционно-издательским советом университета
в качестве учебного электронного пособия

ISBN 978–5–7629–1250–1

© СПбГЭТУ „ЛЭТИ“, 2012

Введение

Данное издание предназначено для студентов первого курса всех технических направлений и специальностей ЭТУ и написано на основе курса, читаемого авторами. В нем излагается значительная часть материала, входящего обычно в раздел „Линейная алгебра и аналитическая геометрия“ курса „Алгебра и геометрия“, необходимого для различных разделов математического анализа и многих приложений. Это прежде всего элементы теории матриц и определителей, теория систем линейных уравнений, векторная алгебра. Рассмотрены фундаментальные понятия линейного пространства, а также определения и свойства унитарных и евклидовых пространств. Задачи аналитической геометрии, в особенности линейные, рассматриваются как приложение алгебры векторов, поэтому в пособии не разделяются линейные задачи на плоскости и в пространстве. Широкое использование векторной алгебры позволяет не вводить много различных форм уравнений прямых и плоскостей и не рассматривать многие частные задачи и формулы.

Теория систем линейных уравнений излагается на основе метода полного исключения неизвестных (метод Гаусса–Жордана). Это, а также широкое использование понятий линейной зависимости и линейной независимости позволило значительно упростить изложение вопроса о совместности систем и сделало теорему Кронекера–Капелли совершенно очевидной.

Теория определителей строится индуктивным методом. Это позволяет избежать изучения теории перестановок, которая в курсе нигде не используется. Все рассуждения, как правило, проводятся с полными доказательствами. Данное издание является переработанным вариантом пособия [1].

Отметим также, что при изложении общей теории линейных пространств авторы старались придерживаться идей, предложенных в замечательной книге И. М. Гельфанда [2].

В данном издании не рассматриваются вопросы, связанные с линейными операторами, собственными числами матриц и теорией квадратичных форм. Перечисленные темы являются основным содержанием следующего пособия [3].

Отметим некоторые стандартные обозначения, принятые в издании. Множество вещественных чисел обозначается символом $\mathbb{R}$; множество комплексных чисел – $\mathbb{C}$, при этом, если $z \in \mathbb{C}$, то $\Re(z)$ – вещественная часть комплексного числа z , а $\Im z$ – его мнимая часть. Знак ■ отмечает конец доказательства; знак ● – конец примера, а знак $\otimes$ – конец замечания.

личными буквами. Например, если в системе

$$\begin{cases} 2x + 3y = 5, \\ -y + 4x = 1 \end{cases}$$

принять $x = x_1$, $y = x_2$, то ее запись в виде (1.1) такова:

$$\begin{cases} 2x_1 + 3x_2 = 5, \\ 4x_1 - x_2 = 1, \end{cases} \quad (1.2)$$

где $a_{11} = 2$, $a_{12} = 3$, $a_{21} = 4$, $a_{22} = -1$, $b_1 = 5$, $b_2 = 1$. Если же принять $x = t_2$, $y = t_1$, то получим другую запись исходной системы:

$$\begin{cases} 3t_1 + 2t_2 = 5, \\ -t_1 + 4t_2 = 1, \end{cases}$$

где $a_{11} = 3$, $a_{12} = 2$, $a_{21} = -1$, $a_{22} = 4$, $b_1 = 5$, $b_2 = 1$. Как видно из этого примера, изменение порядка следования неизвестных, т. е. их перенумерация приводит к изменению вида записи этой системы — во всех уравнениях системы переставляются коэффициенты при соответствующих неизвестных.

Не меняется система, а меняется ее запись в виде (1.1) и при произвольной перестановке ее уравнений. Например, рассмотренная система помимо записи (1.2) допускает эквивалентную ей запись в виде

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

где $a_{11} = 4$, $a_{12} = -1$, $a_{21} = 2$, $a_{22} = 3$, $b_1 = 1$, $b_2 = 5$.

Замечание 1.1. Число неизвестных n системы, как ранее отмечалось, всегда предполагается заранее заданным. Например, рассмотренная система (1.2) является системой двух уравнений с двумя неизвестными. Внешне совпадающая с (1.2) система

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

является иной системой, а именно системой с тремя неизвестными. $\otimes$

Определение 1.1. Решением системы (1.1) называется такой набор чисел $(x_1^0, x_2^0, \dots, x_n^0)$, при подстановке которого в систему (1.1) (x_1^0 вместо x_1 , x_2^0 вместо x_2 и т. д.) все уравнения этой системы становятся верными числовыми равенствами.

Если система (1.1) не имеет решений (множество решений пусто), то она называется несовместной. Если система (1.1) имеет решение, то говорят, что она совместна.

Решить систему (1.1) — значит, найти все ее решения или доказать, что система является несовместной.

Рассмотрим простейший случай, а именно, одно уравнение с одним неизвестным (в этом частном случае все индексы в записи системы (1.1) можно опустить):

$$ax = b.$$

Возможны только следующие 3 случая:

- 1) если $a \neq 0$, то уравнение имеет единственное решение $x = b/a$;
- 2) если $a = 0$, $b \neq 0$, то уравнение решений не имеет;
- 3) если $a = b = 0$, то уравнение имеет бесконечно много решений (x — любое число).

Оказывается, что только эти 3 возможности существуют и для произвольной системы (1.1). Данный факт с необходимыми уточнениями будет установлен в дальнейшем.

Определение 1.2. Две системы вида (1.1) с одинаковым числом неизвестных называются равносильными, если множества решений этих систем совпадают.

В частности, равносильными являются любые две несовместные системы с одинаковым числом неизвестных.

Систему (1.1) можно задать следующей числовой таблицей:

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array} \right), \quad (1.3)$$

составленной из коэффициентов системы и ее правых частей. Часто таблицу (1.3) называют расширенной матрицей системы (1.1). При этом таблицу из коэффициентов системы (1.1)

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

называют матрицей этой системы. Если две системы равносильны, то соответствующие им расширенные матрицы будем связывать логическим знаком эквивалентности: “ $\Longleftrightarrow$ ”.

Иногда в таблице (1.3) над столбцами коэффициентов системы указывают порядковые номера соответствующих неизвестных. Во избежание ошибок это обязательно следует делать в тех случаях, когда будет производиться перенумерация неизвестных. Например, системе (1.2) соответствуют две таблицы

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

В основе большинства известных методов решения системы (1.1) лежит идея преобразования этой системы в некоторую другую систему, ей равносильную, решать которую оказывается проще, чем исходную систему. Такие преобразования обычно называют *допустимыми*¹ преобразованиями системы (1.1). Ясно, что любое допустимое преобразование системы (1.1) приводит к преобразованию соответствующей расширенной матрицы (1.3). Перечислим типы допустимых преобразований системы (1.1):

1) перестановка уравнений в системе (перестановка строк в расширенной матрице системы);

2) перестановка неизвестных во всех уравнениях системы (перестановка столбцов коэффициентов при этих неизвестных в расширенной матрице системы с обязательным указанием измененного порядка следования неизвестных);

3) умножение обеих частей произвольного уравнения системы на число, отличное от нуля (умножение на это число всех элементов соответствующей строки расширенной матрицы системы);

4) прибавление к одному из уравнений системы другого ее уравнения, умноженного на число (прибавление к элементам одной строки расширенной матрицы системы соответствующих элементов другой ее строки, умноженных на число).

Перечисленные преобразования сохраняют неизменным множество решений системы, т. е. справедливо следующее утверждение.

Утверждение 1.1. *Допустимые преобразования переводят систему в равносильную. Это означает, что системы, соответствующие парам*

¹ В некоторых учебниках данные преобразования называются *элементарными*.

расширенных матриц, равносильны¹:

$$\left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{i1} & \cdots & a_{in} & b_i \\ \vdots & \ddots & \vdots & \vdots \\ a_{k1} & \cdots & a_{kn} & b_k \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array} \right) \Leftrightarrow \left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{k1} & \cdots & a_{kn} & b_k \\ \vdots & \ddots & \vdots & \vdots \\ a_{i1} & \cdots & a_{in} & b_i \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array} \right) \quad (1.4)$$

(перестановка строк в расширенной матрице системы);

$$\begin{aligned} & \left(\begin{array}{cccc|cccc} 1 & \cdots & i & \cdots & k & \cdots & n & \\ \hline a_{11} & \cdots & a_{1i} & \cdots & a_{1k} & \cdots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mi} & \cdots & a_{mk} & \cdots & a_{mn} & b_m \end{array} \right) \Leftrightarrow \\ & \Leftrightarrow \left(\begin{array}{cccc|cccc} 1 & \cdots & k & \cdots & i & \cdots & n & \\ \hline a_{11} & \cdots & a_{1k} & \cdots & a_{1i} & \cdots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mk} & \cdots & a_{mi} & \cdots & a_{mn} & b_m \end{array} \right) \end{aligned} \quad (1.5)$$

(перестановка столбцов, кроме последнего, с обязательным указанием измененного порядка следования неизвестных).

Для любого числа $\lambda \in \mathbb{C}$ (или $\lambda \in \mathbb{R}$) и $\lambda \neq 0$ справедливо

$$\left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{k1} & \cdots & a_{kn} & b_k \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array} \right) \Leftrightarrow \left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ \lambda a_{k1} & \cdots & \lambda a_{kn} & \lambda b_k \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array} \right) \quad (1.6)$$

(умножение на число $\lambda \neq 0$ всех элементов соответствующей строки расширенной матрицы системы).

¹ Здесь считаем для справедливости записи, что $m \geq 4$, $n \geq 6$, $2 < i < k \leq m-1$ и $2 < i < k-1 \leq n-2$; изменения, необходимые в записи (1.4)–(1.7) при других m , n , i , k , очевидны.

Для любого числа $\lambda \in \mathbb{C}$ (или $\lambda \in \mathbb{R}$)

$$\left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{i1} & \cdots & a_{in} & b_i \\ \vdots & \ddots & \vdots & \vdots \\ a_{k1} & \cdots & a_{kn} & b_k \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array} \right) \Leftrightarrow \left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{i1} & \cdots & a_{in} & b_i \\ \vdots & \ddots & \vdots & \vdots \\ a_{k1} + \lambda a_{i1} & \cdots & a_{kn} + \lambda a_{in} & b_k + \lambda b_i \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array} \right) \quad (1.7)$$

(прибавление к элементам одной строки расширенной матрицы системы соответствующих элементов другой ее строки, умноженных на число).

Доказательство. Для преобразования первого типа (1.4) это совершенно очевидно.

Докажем равносильность преобразования второго типа при перестановке первого и второго неизвестных. Заметим, что в силу коммутативности сложения уравнение

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1$$

равносильно уравнению

$$a_{12}x_2 + a_{11}x_1 + \cdots + a_{1n}x_n = b_1.$$

Это же верно и для остальных уравнений системы (1.1), поэтому таблица

$$\left(\begin{array}{ccccc|c} 2 & 1 & 3 & \cdots & n & \\ \hline a_{12} & a_{11} & a_{13} & \cdots & a_{1n} & b_1 \\ a_{22} & a_{21} & a_{23} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m2} & a_{m1} & a_{m3} & \cdots & a_{mn} & b_m \end{array} \right)$$

соответствует системе, равносильной системе (1.1). Аналогично доказывается возможность перестановки произвольных неизвестных. Это означает, что таблица, полученная произвольной перестановкой местами столбцов (кроме последнего) исходной таблицы, соответствует системе, равносильной исходной.

Докажем, что преобразование третьего типа переводит систему (1.1) в равносильную ей.

Если $(x_1^0, x_2^0, \dots, x_n^0)$ – решение системы (1.1), то, согласно определению 1.1, все уравнения этой системы при замене x_k на x_k^0 , $1 \leq k \leq n$, становятся верными числовыми равенствами. В частности, если

$$a_{k1}x_1^0 + a_{k2}x_2^0 + \cdots + a_{kn}x_n^0 = b_k$$

$$\lambda a_{k1}x_1^0 + \lambda a_{k2}x_2^0 + \dots + \lambda a_{kn}x_n^0 = \lambda b_k$$

тоже верное числовое равенство при любом λ . Следовательно, набор чисел $(x_1^0, x_2^0, \dots, x_n^0)$ является решением системы

[illegible]

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \lambda a_{k1} & \lambda a_{k2} & \cdots & \lambda a_{kn} & \lambda b_k \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix}, \quad (1.8)$$
$$a_{k1}\tilde{x}_1^0 + a_{k2}\tilde{x}_2^0 + \cdots + a_{kn}\tilde{x}_n^0 = b_k.$$

Докажем, наконец, что преобразование четвертого типа также переводит систему (1.1) в равносильную ей.

Если $\lambda = 0$, то две рассматриваемые системы просто совпадают и доказательства не требуется. Пусть $\lambda \neq 0$. Если $(x_1^0, x_2^0, \dots, x_n^0)$ – решение системы (1.1), тогда все уравнения этой системы при замене x_k на x_k^0 становятся верными числовыми равенствами. Прибавляя к верному равенству $a_{k1}x_1^0 + a_{k2}x_2^0 + \dots + a_{kn}x_n^0 = b_k$ верное числовое равенство $\lambda a_{i1}x_1^0 + \lambda a_{i2}x_2^0 + \dots + \lambda a_{in}x_n^0 = \lambda b_i$, получим верное числовое равенство

$$(a_{k1} + \lambda a_{j1})x_1^0 + (a_{k2} + \lambda a_{j2})x_2^0 + \cdots + (a_{kn} + \lambda a_{jn})x_n^0 = b_k + \lambda b_j,$$

[illegible]
$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} & b_i \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{k1} + \lambda a_{i1} & a_{k2} + \lambda a_{i2} & \cdots & a_{kn} + \lambda a_{in} & b_k + \lambda b_i \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array} \right), \quad (1.9)$$

Отметим еще 2 важных свойства системы (1.1).

$$0x_1 + 0x_2 + \cdots + 0x_n = b, \quad (1.10)$$

Доказательство. Действительно, ни при каком наборе значений неизвестных $(x_1^0, x_2^0, \dots, x_n^0)$ уравнение (1.10) не может стать верным числовым равенством. ■

$$0x_1 + 0x_2 + \cdots + 0x_n = 0 \quad (1.11)$$

11

Доказательство. Ясно, что уравнение (1.11) является верным числовым равенством ($0 = 0$) при любом наборе значений неизвестных, и поэтому добавление такого уравнения к системе (1.1) не меняет множества ее решений. ■

Следствие 1.1. *Если в системе (1.1) содержится уравнение вида (1.11) или несколько таких уравнений, то эти уравнения можно исключить из системы (1.1), получив при этом систему с меньшим числом уравнений, равносильную исходной.*

Замечание 1.2. Однако если все уравнения системы (1.1) имеют вид (1.11) (такой случай возможен), то из этой системы исключаются все уравнения, кроме одного, так как в любой системе вида (1.1) должно быть хотя бы одно уравнение ($m \geq 1$). В этом случае система (1.1) равносильна одному уравнению (1.11), и, разумеется, решением такой системы служит любой набор $(x_1^0, x_2^0, \dots, x_n^0)$ значений неизвестных (сравните с рассмотренным ранее третьим случаем для уравнения $ax = b$). ⊗

1.2. Метод полного исключения неизвестных (метод Гаусса – Жордана)

Метод (алгоритм) Гаусса – Жордана является одним из методов решения системы (1.1) и сводится к последовательному применению допустимых преобразований системы.

Если в исходной системе (1.1) все коэффициенты a_{ij} равны нулю, то она либо несовместна (среди чисел b_i имеется хотя бы одно отличное от нуля), либо имеет бесконечно много решений (все правые части b_i равны нулю – в этом случае, как было отмечено в замечании 1.2, любой набор чисел $(x_1^0, x_2^0, \dots, x_n^0)$ служит ее решением).

Рассмотрим теперь случай, когда среди коэффициентов a_{ij} системы (1.1) имеются ненулевые, и опишем метод решения такой системы как последовательное (пошаговое) выполнение перечисленных допустимых преобразований расширенной матрицы системы.

Замечание 1.3. Как отмечалось ранее, перестановка столбцов в матрице системы приводит к изменению порядка следования неизвестных. Однако во избежание излишней громоздкости изложения не будем при описании метода указывать над столбцами матрицы номера соответствующих неизвестных (это будет сделано в рассматриваемых далее примерах). Это означает, что в применяемых обозначениях второй индекс j коэффициента a_{ij} соответствует номеру столбца и может, в общем случае, отличаться от номера соответствующей неизвестной. ⊗

Шаг 1.

1. Выберем какой-либо ненулевой элемент a_{ij} матрицы системы и назовем его ведущим элементом. Переставляя строки и столбцы расширенной матрицы (т.е. используя допустимые преобразования первого и второго типа), поместим ведущий элемент в первую строку и первый столбец (разумеется, если в качестве ведущего элемента был выбран элемент a_{11} , то никаких преобразований не делается). Полученную расширенную матрицу запишем в виде

$$\left(\begin{array}{cccc|c} a_{11}^1 & a_{12}^1 & \cdots & a_{1n}^1 & b_1^1 \\ a_{21}^1 & a_{22}^1 & \cdots & a_{2n}^1 & b_2^1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1}^1 & a_{m2}^1 & \cdots & a_{mn}^1 & b_m^1 \end{array} \right) \quad (1.12)$$

(верхний индекс указывает номер шага). В этой матрице $a_{11}^1 \neq 0$.

2. Умножим первую строку матрицы (1.12) на $1/a_{11}^1$ (это – преобразование третьего типа). Получим равносильную систему с расширенной матрицей

$$\left(\begin{array}{cccc|c} 1 & \tilde{a}_{12}^1 & \cdots & \tilde{a}_{1n}^1 & \tilde{b}_1^1 \\ a_{21}^1 & a_{22}^1 & \cdots & a_{2n}^1 & b_2^1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1}^1 & a_{m2}^1 & \cdots & a_{mn}^1 & b_m^1 \end{array} \right), \quad (1.13)$$

где $\tilde{a}_{1j}^1 = a_{1j}^1/a_{11}^1$, $j = 2, 3, \dots, n$ и $\tilde{b}_1^1 = b_1^1/a_{11}^1$.

3. Ко второй строке матрицы (1.13) прибавим 1-ю строку, умноженную на $(-a_{21}^1)$, к третьей строке прибавим 1-ю, умноженную на $(-a_{31}^1)$, и т.д. до последней строки, к которой прибавим 1-ю строку, умноженную на $(-a_{m1}^1)$ (это – преобразование четвертого типа). В результате получим равносильную систему с расширенной матрицей вида

$$\left(\begin{array}{cccc|c} 1 & \tilde{a}_{12}^1 & \cdots & \tilde{a}_{1n}^1 & \tilde{b}_1^1 \\ 0 & \tilde{a}_{22}^1 & \cdots & \tilde{a}_{2n}^1 & \tilde{b}_2^1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \tilde{a}_{m2}^1 & \cdots & \tilde{a}_{mn}^1 & \tilde{b}_m^1 \end{array} \right). \quad (1.14)$$

Если в матрице (1.14) содержатся нулевые строки, то исключим их (см. следствие 1.1); при этом число уравнений в системе уменьшится. Если же эта матрица содержит строку вида $[0 \ 0 \ \cdots \ 0 \ | \ b]$, где $b \neq 0$, то согласно утверждению 1.2, соответствующая система несовместна, а значит, несов-

местна и исходная система (1.1). В этом случае работа метода заканчивается.

Пусть в результате первого шага не обнаружена несовместность системы, а в получившейся матрице после отбрасывания возможных нулевых строк число оставшихся строк равно m_1 ($m_1 \leq m$), при этом $m_1 \geq 2$. Записывая эту матрицу в виде

$$\left(\begin{array}{cccc|c} 1 & a_{12}^2 & \cdots & a_{1n}^2 & b_1^2 \\ 0 & a_{22}^2 & \cdots & a_{2n}^2 & b_2^2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & a_{m_1 2}^2 & \cdots & a_{m_1 n}^2 & b_{m_1}^2 \end{array} \right), \quad (1.15)$$

переходим ко второму шагу. Отметим, что в каждой строке матрицы (1.15), начиная со второй, хотя бы один из коэффициентов a_{ij}^2 отличен от нуля.

Шаг 2.

1. Среди коэффициентов a_{ij}^2 , $2 \leq i \leq m_1$, $2 \leq j \leq n$, выберем какой-либо ненулевой и назовем его ведущим элементом второго шага. Переставляя, если нужно, строки и столбцы матрицы (1.15), поместим ведущий элемент во вторую строку и второй столбец. Для получившейся расширенной матрицы сохраним прежнюю запись в виде (1.15), считая теперь, что $a_{22}^2 \neq 0$.

2. Умножим вторую строку матрицы (1.15) на $\frac{1}{a_{22}^2}$. Получим равносильную систему с расширенной матрицей

$$\left(\begin{array}{ccccc|c} 1 & a_{12}^2 & a_{13}^2 & \cdots & a_{1n}^2 & b_1^2 \\ 0 & 1 & \tilde{a}_{23}^2 & \cdots & \tilde{a}_{2n}^2 & \tilde{b}_2^2 \\ 0 & a_{32}^2 & a_{33}^2 & \cdots & a_{3n}^2 & b_3^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & a_{m_1 2}^2 & a_{m_1 3}^2 & \cdots & a_{m_1 n}^2 & b_{m_1}^2 \end{array} \right), \quad (1.16)$$

где, очевидно, $\tilde{a}_{2j}^2 = \frac{a_{2j}^2}{a_{22}^2}$, $j = 3, 4, \dots, n$, и $\tilde{b}_2^2 = \frac{b_2^2}{a_{22}^2}$.

3. К первой строке матрицы (1.16) прибавим вторую строку, умноженную на $(-a_{12}^2)$, к третьей строке прибавим вторую, умноженную на $(-a_{32}^2)$, и т.д. до последней строки, к которой прибавим вторую строку, умноженную на $(-a_{m_1 2}^2)$. В результате получим равносильную систему с расширенной матрицей вида

$$\left(\begin{array}{ccccc|c} 1 & 0 & \tilde{a}_{13}^2 & \cdots & \tilde{a}_{1n}^2 & \tilde{b}_1^2 \\ 0 & 1 & \tilde{a}_{23}^2 & \cdots & \tilde{a}_{2n}^2 & \tilde{b}_2^2 \\ 0 & 0 & \tilde{a}_{33}^2 & \cdots & \tilde{a}_{3n}^2 & \tilde{b}_3^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \tilde{a}_{m_1 3}^2 & \cdots & \tilde{a}_{m_1 n}^2 & \tilde{b}_{m_1}^2 \end{array} \right). \quad (1.17)$$

Как и на первом шаге, если в матрице (1.17) содержатся нулевые строки, то они отбрасываются (число уравнений в системе уменьшается), если же в этой матрице содержится строка вида $[0 \ 0 \ \cdots \ 0 \mid b]$, $b \neq 0$, то соответствующая система вместе с исходной системой несовместна.

Итак, в результате выполнения второго шага метода Гаусса–Жордана либо устанавливается несовместность системы, либо она приводится к равносильной системе с расширенной матрицей вида

$$\left(\begin{array}{ccccc|c} 1 & 0 & a_{13}^3 & \cdots & a_{1n}^3 & b_1^3 \\ 0 & 1 & a_{23}^3 & \cdots & a_{2n}^3 & b_2^3 \\ 0 & 0 & a_{33}^3 & \cdots & a_{3n}^3 & b_3^3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & a_{m_2 3}^3 & \cdots & a_{m_2 n}^3 & b_{m_2}^3 \end{array} \right), \quad (1.18)$$

где $2 \leq m_2 \leq m_1$. Если $m_2 > 2$ (в этом случае в (1.18) в каждой строке, начиная с третьей, хотя бы один из коэффициентов a_{ij}^3 отличен от нуля), то переходим к третьему шагу и т. д. Задачей очередного i -го шага (если он выполняется) является получение в i -м столбце матрицы системы единицы на месте элемента a_{ii}^i и нулей – на месте остальных элементов этого столбца. При этом предыдущие столбцы остаются неизменными.

Если в процессе выполнения шагов метода (алгоритма) Гаусса–Жордана не устанавливается несовместность системы, то этот процесс всегда заканчивается (после некоторого k -го шага, $1 \leq k \leq n$) получением либо (когда $k = n$) системы с расширенной матрицей вида

$$\left(\begin{array}{ccccc|c} 1 & 0 & 0 & \cdots & 0 & b_1^n \\ 0 & 1 & 0 & \cdots & 0 & b_2^n \\ 0 & 0 & 1 & \cdots & 0 & b_3^n \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & b_n^n \end{array} \right) \quad (1.19)$$

(верхний индекс „ n “ указывает номер последнего шага), либо (когда $k < n$)

системы с расширенной матрицей вида

$$\left(\begin{array}{cccccc|c} 1 & 0 & 0 & \cdots & 0 & a_{1,k+1}^k & \cdots & a_{1n}^k & b_1^k \\ 0 & 1 & 0 & \cdots & 0 & a_{2,k+1}^k & \cdots & a_{2n}^k & b_2^k \\ 0 & 0 & 1 & \cdots & 0 & a_{3,k+1}^k & \cdots & a_{3n}^k & b_3^k \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & a_{k,k+1}^k & \cdots & a_{kn}^k & b_k^k \end{array} \right). \quad (1.20)$$

Каждая строка матрицы (1.19) соответствует уравнению с одним неизвестным, коэффициент при котором равен 1, а единственное значение неизвестного, удовлетворяющее этому уравнению, совпадает с правой частью этого уравнения. Таким образом, если алгоритм Гаусса – Жордана приводит систему (1.1) к равносильной системе с расширенной матрицей вида (1.19), то система (1.1) имеет единственное решение. Соответствующий единственный набор значений неизвестных $(x_1^0, x_2^0, \dots, x_n^0)$ легко находится по последнему столбцу матрицы (1.19). Если при выполнении метода Гаусса – Жордана столбцы матрицы системы не переставлялись, то

$$\begin{cases} x_1^0 = b_1^n, \\ x_2^0 = b_2^n, \\ \dots\dots\dots \\ x_n^0 = b_n^n. \end{cases}$$

Рассмотрим теперь систему с расширенной матрицей (1.20). Для простоты и наглядности изложения предположим, что в процессе приведения исходной системы (1.1) к системе с расширенной матрицей (1.20) порядок следования неизвестных в уравнениях системы не менялся, т. е. столбцы матрицы системы не переставлялись. По расширенной матрице (1.20) запишем соответствующие уравнения системы

$$\begin{cases} x_1 + a_{1,k+1}^k x_{k+1} + \cdots + a_{1n}^k x_n = b_1^k, \\ x_2 + a_{2,k+1}^k x_{k+1} + \cdots + a_{2n}^k x_n = b_2^k, \\ \dots\dots\dots \\ x_k + a_{k,k+1}^k x_{k+1} + \cdots + a_{kn}^k x_n = b_k^k. \end{cases} \quad (1.21)$$

Очевидно, что, задавая произвольно $x_{k+1}, \dots, x_n$ и вычисляя по формулам

$$\begin{cases} x_1 = b_1^k - a_{1,k+1}^k x_{k+1} - \cdots - a_{1n}^k x_n, \\ x_2 = b_2^k - a_{2,k+1}^k x_{k+1} - \cdots - a_{2n}^k x_n, \\ \dots\dots\dots \\ x_k = b_k^k - a_{k,k+1}^k x_{k+1} - \cdots - a_{kn}^k x_n \end{cases}$$

значения неизвестных $x_1, \dots, x_k$, получим все решения рассматриваемой системы.

Если в процессе приведения системы (1.1) к системе с расширенной матрицей (1.20) менялся порядок следования неизвестных, т. е. переставлялись столбцы в матрице системы, то уравнения системы, соответствующей матрице (1.20), будут отличаться от уравнений системы (1.21) только номерами отдельных, или даже всех, неизвестных. Однако структура всего множества решений системы не меняется: k неизвестных системы однозначно выражаются через остальные $(n - k)$ неизвестные, которые можно задавать произвольно. Таким образом, если алгоритм Гаусса – Жордана приводит систему (1.1) к равносильной системе с расширенной матрицей вида (1.20), то система имеет бесконечно много решений.

Замечание 1.4. Отметим, что при численной реализации алгоритма Гаусса–Жордана на ЭВМ может оказаться, что ведущий элемент на некотором шаге алгоритма очень мал по сравнению с другими элементами соответствующей строки. В этом случае в результате деления элементов строки на ведущий элемент может возникнуть переполнение разрядной сетки вычислительной машины и, как следствие, аварийный останов выполнения программы. Именно поэтому все машинные алгоритмы решения систем линейных уравнений предусматривают выбор в качестве ведущего элемента каждого шага наибольшего по модулю элемента из соответствующего множества элементов матрицы системы. Если таких элементов несколько, то в качестве ведущего элемента выбирается любой из них. $\otimes$

1.3. Примеры использования метода полного исключения неизвестных для решения систем линейных уравнений

Пример 1.1. Решить систему

$$\begin{cases} 1x_1 + 2x_2 + 3x_3 = 14, \\ 4x_1 + 5x_2 + 6x_3 = 32, \\ 7x_1 + 9x_2 + 8x_3 = 49, \\ 2x_1 + 2x_2 - 1x_3 = 3. \end{cases} \quad (1.22)$$

Системе (1.22) соответствует таблица

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 14 \\ 4 & 5 & 6 & 32 \\ 7 & 9 & 8 & 49 \\ 2 & 2 & -1 & 3 \end{array} \right).$$

Будем решать систему методом Гаусса–Жордана с выбором в качестве ведущего максимального по модулю элемента (см. замечание 1.4).

Шаг 1.

1. Ведущим элементом является элемент $a_{32} = 9$. Применяя преобразование первого типа к строкам с номерами 1 и 3 и преобразование второго типа к столбцам 1 и 2, получим новую таблицу с первым элементом первой строки $a_{11}^1 = a_{32} = 9$.

2. Первую строку умножим на $\frac{1}{a_{11}^1} = \frac{1}{9}$ – это преобразование третьего типа.

3. Применяем преобразование четвертого типа к первой и второй строкам с $\lambda = -a_{21}^1 = -5$; к первой и третьей строкам с $\lambda = -a_{31}^1 = -2$; к первой и четвертой строкам с $\lambda = -a_{41}^1 = -2$. В результате получим расширенную матрицу, в первом столбце которой все элементы, кроме первого, равны нулю (эта матрица соответствует системе, в которой из всех уравнений, кроме первого, исключено неизвестное x_2 , коэффициенты при котором образуют первый столбец).

Указанным преобразованиями первого шага соответствует следующая последовательность расширенных матриц, которые соответствуют системам, равносильным исходной системе (1.22):

$$\left(\begin{array}{ccc|c} 2 & 1 & 3 & 49 \\ 9 & 7 & 8 & 32 \\ 5 & 4 & 6 & 14 \\ 2 & 1 & 3 & 3 \end{array} \right) \Leftrightarrow \left(\begin{array}{ccc|c} 2 & 1 & 3 & 49 \\ 1 & \frac{7}{9} & \frac{8}{9} & \frac{49}{9} \\ 5 & 4 & 6 & 32 \\ 2 & 1 & 3 & 14 \\ 2 & 2 & -1 & 3 \end{array} \right) \Leftrightarrow \left(\begin{array}{ccc|c} 2 & 1 & 3 & 49 \\ 1 & \frac{7}{9} & \frac{8}{9} & \frac{49}{9} \\ 0 & \frac{1}{9} & \frac{14}{9} & \frac{43}{9} \\ 0 & -\frac{5}{9} & \frac{11}{9} & \frac{28}{9} \\ 0 & \frac{4}{9} & -\frac{25}{9} & \frac{71}{9} \end{array} \right).$$

Так как в последней таблице нет строк вида $[000|b]$, $b \neq 0$, и нулевых строк, значит, не установлена несовместность системы и количество уравнений не уменьшается. Переходим к следующему шагу.

Шаг 2.

1. Ведущий элемент $a_{43}^1 = \frac{-25}{9}$. Переставляем вторую и четвертую строки и второй и третий столбцы таблицы. После этого ведущий элемент будет на втором месте во второй строке, т. е. $a_{22}^2 = \frac{-25}{9}$.

2. Делим элементы второй строки на ведущий элемент (умножаем на $\frac{1}{a_{22}^2} = \frac{-9}{25}$).

3. Используя допустимые преобразования четвертого типа, получаем

во втором столбце нули на месте всех элементов, кроме второго.

Преобразованиям второго шага соответствует следующая последовательность расширенных матриц:

$$\left(\begin{array}{ccc|c} 2 & 3 & 1 & \\ \hline 1 & \frac{8}{9} & \frac{7}{9} & \frac{49}{9} \\ 0 & -\frac{25}{9} & \frac{4}{9} & -\frac{71}{9} \\ 0 & \frac{11}{9} & -\frac{5}{9} & \frac{28}{9} \\ 0 & \frac{14}{9} & \frac{1}{9} & \frac{43}{9} \end{array} \right) \Leftrightarrow \left(\begin{array}{ccc|c} 2 & 3 & 1 & \\ \hline 1 & \frac{8}{9} & \frac{7}{9} & \frac{49}{9} \\ 0 & 1 & -\frac{4}{25} & \frac{71}{25} \\ 0 & \frac{11}{9} & -\frac{5}{9} & \frac{28}{9} \\ 0 & \frac{14}{9} & \frac{1}{9} & \frac{43}{9} \end{array} \right) \Leftrightarrow \left(\begin{array}{ccc|c} 2 & 3 & 1 & \\ \hline 1 & 0 & \frac{23}{25} & \frac{73}{25} \\ 0 & 1 & -\frac{4}{25} & \frac{71}{25} \\ 0 & 0 & -\frac{9}{25} & -\frac{9}{25} \\ 0 & 0 & \frac{9}{25} & \frac{9}{25} \end{array} \right).$$

Так как в последней таблице нет строк вида $[000|b]$, $b \neq 0$, и нулевых строк, то и на этом шаге не установлена несовместность системы и количество уравнений не уменьшается.

Шаг 3.

1. Выберем ведущий элемент $a_{31}^2 = -9/25$ (так как $|a_{41}^2| = |a_{31}^2| = 9/25$, то в качестве ведущего может быть выбран любой из элементов a_{31}^2, a_{41}^2). Ведущий элемент уже стоит на третьем месте в третьей строке, и перестановок не требуется.

2. Делим элементы третьей строки на ведущий элемент.

3. Получаем нули на месте всех элементов третьего столбца, кроме третьего элемента.

В результате выполнения третьего шага получается следующая последовательность расширенных матриц:

$$\left(\begin{array}{ccc|c} 2 & 3 & 1 & \\ \hline 1 & 0 & \frac{23}{25} & \frac{73}{25} \\ 0 & 1 & -\frac{4}{25} & \frac{71}{25} \\ 0 & 0 & -\frac{9}{25} & -\frac{9}{25} \\ 0 & 0 & \frac{9}{25} & \frac{9}{25} \end{array} \right) \Leftrightarrow \left(\begin{array}{ccc|c} 2 & 3 & 1 & \\ \hline 1 & 0 & \frac{23}{25} & \frac{73}{25} \\ 0 & 1 & -\frac{4}{25} & \frac{71}{25} \\ 0 & 0 & 1 & 1 \\ 0 & 0 & \frac{9}{25} & \frac{9}{25} \end{array} \right) \Leftrightarrow \left(\begin{array}{ccc|c} 2 & 3 & 1 & \\ \hline 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array} \right).$$

Так как в последней расширенной матрице последняя строка состоит из нулей, то ее можно удалить из таблицы (или вычеркнуть последнее уравнение из системы), не теряя равносильности. На этом выполнение метода Гаусса–Жордана заканчивается. Окончательной расширенной мат-

рицей является матрица $\left(\begin{array}{ccc|c} 2 & 3 & 1 & 2 \\ 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 1 \end{array} \right)$, имеющая вид (1.19), которой

соответствует система $\begin{cases} x_2 = 2, \\ x_3 = 3, \\ x_1 = 1. \end{cases}$

Итак, система (1.22) имеет единственное решение: $\begin{cases} x_1 = 1, \\ x_2 = 2, \\ x_3 = 3. \end{cases} \bullet$

Пример 1.2. Решить систему

$$\begin{cases} 3x_1 - 2x_2 + 2x_3 + 1x_4 = 5, \\ 1x_1 + 3x_2 + 2x_3 - 2x_4 = 7, \\ 2x_1 + 2x_2 + 0x_3 + 0x_4 = 6, \\ 2x_1 + 1x_2 - 1x_3 + 1x_4 = 3. \end{cases} \quad (1.23)$$

Системе (1.23) соответствует таблица $\left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 5 \\ 3 & -2 & 2 & 1 & 7 \\ 1 & 3 & 2 & -2 & 6 \\ 2 & 2 & 0 & 0 & 3 \end{array} \right)$.

Будем решать систему методом Гаусса–Жордана.

Шаг 1.

1. Ведущий элемент $a_{11} = 3$. Ведущий элемент уже стоит на первом месте в первой строке, и перестановок не требуется.

2. Делим элементы первой строки на ведущий элемент.

3. Получаем нули на месте всех элементов первого столбца, кроме первого элемента.

В результате выполнения первого шага получается следующая последовательность расширенных матриц:

$$\left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 5 \\ 3 & -2 & 2 & 1 & 7 \\ 1 & 3 & 2 & -2 & 6 \\ 2 & 2 & 0 & 0 & 3 \end{array} \right) \Leftrightarrow \left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 5 \\ 1 & \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} & \frac{5}{3} \\ 0 & \frac{8}{3} & \frac{8}{3} & -\frac{8}{3} & \frac{16}{3} \\ 0 & \frac{4}{3} & \frac{4}{3} & -\frac{4}{3} & \frac{8}{3} \\ 0 & \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{array} \right).$$

Так как в последней расширенной матрице нет строк вида $[000|b]$, $b \neq 0$, и нулевых строк, то не установлена несовместность системы и количество уравнений не уменьшается.

Шаг 2.

1. Выберем ведущий элемент $a_{22}^1 = 8/3$ (так как $|a_{22}^1| = |a_{23}^1| = |a_{24}^1| = 8/3$, ведущим может быть выбран любой из элементов, a_{22}^1 , a_{23}^1 или a_{24}^1). Ведущий элемент уже стоит на втором месте во второй строке, и перестановок не требуется.

2. Делим элементы второй строки на ведущий элемент.

3. Получаем нули на месте всех элементов второго столбца, кроме второго элемента.

В результате выполнения второго шага получается следующая последовательность расширенных матриц:

$$\left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 5 \\ 1 & \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} & \frac{5}{3} \\ 0 & 1 & 1 & -1 & 2 \\ 0 & \frac{4}{3} & \frac{4}{3} & \frac{4}{3} & \frac{8}{3} \\ 0 & \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{array} \right) \Leftrightarrow \left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 5 \\ 1 & 0 & -1 & 1 & 1 \\ 0 & 1 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right).$$

Так как в последней расширенной матрице 2 последние строки состоят из нулей, их можно удалить из таблицы (или вычеркнуть 2 последних уравнения из системы), не теряя равносильности. На этом заканчивается выполнение метода Гаусса–Жордана. Окончательной расширенной матрицей является матрица вида (1.20)

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

Множеством решений системы (а значит, и исходной системы (1.23)) является множество наборов

$$(1 + x_3 - x_4, 2 - x_3 + x_4, x_3, x_4),$$

где числа x_3 и x_4 могут быть выбраны произвольно. •

Пример 1.3. Решить систему

$$\begin{cases} 1x_1 + 2x_2 + 3x_3 = 1, \\ 4x_1 + 5x_2 + 6x_3 = 2, \\ 7x_1 + 8x_2 + 9x_3 = 5. \end{cases} \quad (1.24)$$

Убедитесь самостоятельно, что после выполнения двух шагов метода Гаусса–Жордана получим систему, которой соответствует таблица

$$\left(\begin{array}{ccc|c} 3 & 1 & 2 & \\ \hline 1 & 0 & 1/2 & 1/6 \\ 0 & 1 & 1/2 & 1/2 \\ 0 & 0 & 0 & -1 \end{array} \right),$$

равносильную системе (1.24).

Наличие строки $[0 \ 0 \ 0 \ | \ -1]$ в этой таблице доказывает несовместность полученной системы, а следовательно, и несовместность системы (1.24). •

1.4. Матрицы. Определения. Равенство матриц

Определение 1.3. Числовой матрицей A размера $m \times n$ называется совокупность mn чисел $a_{lj} \in \mathbb{C}$, расположенных в виде прямоугольной таблицы

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \quad (1.25)$$

где m – число строк, а n – число столбцов.

Матрицы будем обозначать прописными латинскими буквами, а числа, образующие матрицу (элементы матрицы), – соответствующими строчными буквами с двумя индексами. При этом первый индекс элемента матрицы обозначает номер строки, а второй – номер столбца, в которых расположен данный элемент. Кроме того, часто используется обозначение

$$A = \{a_{lj}\}, \quad 1 \leq l \leq m, \quad 1 \leq j \leq n.$$

Если число строк равно числу столбцов ($m = n$), то матрица называется *квадратной*, а число n – ее порядком. Например, матрица $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ – квадратная матрица второго порядка.

Среди квадратных матриц выделяют *единичную* матрицу

$$I_n = \underbrace{\left[\begin{array}{cccc} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{array} \right]}_n \Bigg\} n.$$

Когда порядок матрицы I_n понятен, то индекс n опускается.

Элементы квадратной матрицы $a_{11}, a_{22}, \dots, a_{nn}$ образуют так называемую *главную диагональ* (или просто *диагональ*) матрицы. Если на главной диагонали расположены числа $\lambda_1, \lambda_2, \dots, \lambda_n$, а остальные элементы равны нулю, то матрица называется *диагональной*. При этом часто используется обозначение

$$\begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} = \text{diag}[\lambda_1, \lambda_2, \dots, \lambda_n].$$

Таким образом, $I_n = \text{diag} \underbrace{[1, 1, \dots, 1]}_n$.

Замечание 1.5. Если все элементы матрицы A вещественны ($a_{lj} \in \mathbb{R}$), то матрицу A называют вещественной, а в случае $a_{lj} \in \mathbb{C}$ – комплексной. $\otimes$

Часто удобно рассматривать матрицу как совокупность ее столбцов или как совокупность ее строк. Пусть матрица A задана в виде (1.25). Рассмотрим n матриц размера $m \times 1$ (их называют столбцами высотой m):

$$A_{(1)} = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, A_{(2)} = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \dots, A_{(n)} = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix},$$

при этом будем писать $A = [A_{(1)} A_{(2)} \dots A_{(n)}]$. Если ввести m матриц размера $1 \times n$ (их называют строками длиной n):

$$A^{(1)} = [a_{11} \ a_{12} \ \dots \ a_{1n}], \ A^{(2)} = [a_{21} \ a_{22} \ \dots \ a_{2n}], \dots, A^{(m)} = [a_{m1} \ a_{m2} \ \dots \ a_{mn}],$$

то в этом случае будем писать

$$A = \begin{bmatrix} A^{(1)} \\ A^{(2)} \\ \vdots \\ A^{(m)} \end{bmatrix}.$$

Определение 1.4. Множество всех комплексных (вещественных) матриц размера $m \times n$ обозначается через $\mathcal{M}_{m \times n}(\mathbb{C})$ ($\mathcal{M}_{m \times n}(\mathbb{R})$).

Если понятно, какие числа используются (или если это безразлично), то допустима сокращенная запись $\mathcal{M}_{m \times n}$. Если матрицы квадратные, то пишут $\mathcal{M}_n(\mathbb{C}) = \mathcal{M}_{n \times n}(\mathbb{C})$, $\mathcal{M}_n(\mathbb{R}) = \mathcal{M}_{n \times n}(\mathbb{R})$ ($\mathcal{M}_n = \mathcal{M}_{n \times n}$).

Определение 1.5. Матрицы A и B равны, если: 1) они имеют одинаковые размеры ($A, B \in \mathcal{M}_{m \times n}$); 2) все элементы одной матрицы равны соответствующим элементам другой ($a_{lj} = b_{lj}$; $l = 1, 2, \dots, m$; $j = 1, 2, \dots, n$).

1.5. Линейные операции с матрицами

Определение 1.6. Пусть $A, B \in \mathcal{M}_{m \times n}(\mathbb{C})$. Суммой матриц A и B называется матрица $C = A + B \in \mathcal{M}_{m \times n}(\mathbb{C})$, элементы которой равны суммам соответствующих элементов матриц A и B :

$$C = A + B \iff c_{lj} = a_{lj} + b_{lj}; \quad l = 1, 2, \dots, m; \quad j = 1, 2, \dots, n. \quad (1.26)$$

Произведением матрицы $A \in \mathcal{M}_{m \times n}(\mathbb{C})$ на число $\lambda \in \mathbb{C}$ называется матрица $D = \lambda A \in \mathcal{M}_{m \times n}(\mathbb{C})$, элементы которой получаются умножением элементов матрицы A на число λ :

$$D = \lambda A \iff d_{lj} = \lambda a_{lj}; \quad l = 1, 2, \dots, m; \quad j = 1, 2, \dots, n. \quad (1.27)$$

Замечание 1.6. На множестве $\mathcal{M}_{m \times n}(\mathbb{R})$ аналогично определяются операции сложения матриц и умножения матрицы на вещественное число. Для этого достаточно в определении 1.6 заменить $\mathbb{C}$ на $\mathbb{R}$.

Матрицу из множества $\mathcal{M}_{m \times n}$, все элементы которой равны нулю, будем называть нулевой матрицей $\Theta_{m \times n}$ (или просто Θ). Из формулы (1.26) сразу следует, что

$$A + \Theta = A \quad (A, \Theta \in \mathcal{M}_{m \times n}).$$

Очевидно (см. (1.27)), что $1A = A$. Если через $(-A)$ обозначить матрицу $(-1)A$, то

$$A + (-A) = \Theta.$$

Разность двух матриц $A, B \in \mathcal{M}_{m \times n}(\mathbb{C})$ определяется формулой

$$A - B = A + (-B).$$

Ясно, что если $C = A - B$, то $c_{lj} = a_{lj} - b_{lj}$, где $l = 1, 2, \dots, m$; $j = 1, 2, \dots, n$.

Теорема 1.1. Введенные в определении 1.6 линейные операции обладают следующими свойствами:

1. $A + B = B + A$.
2. $A + (B + C) = (A + B) + C$.
3. $\lambda(A + B) = \lambda A + \lambda B$.
4. $(\lambda + \mu)A = \lambda A + \mu A$.
5. $(\lambda\mu)A = \lambda(\mu A)$,

где $A, B, C \in \mathcal{M}_{m \times n}(\mathbb{C})$ и $\lambda, \mu \in \mathbb{C}$.

Доказательство. Докажем первое из перечисленных свойств. Согласно (1.26),

$$C = A + B \iff c_{lj} = a_{lj} + b_{lj}, \quad l = 1, 2, \dots, m; \quad j = 1, 2, \dots, n,$$

и

$$D = B + A \iff d_{lj} = b_{lj} + a_{lj}, \quad l = 1, 2, \dots, m; \quad j = 1, 2, \dots, n.$$

Учитывая равенство $a_{lj} + b_{lj} = b_{lj} + a_{lj}$ и определение 1.5 получим $C = D$. Остальные свойства доказываются аналогично. ■

Для матрицы $C \in \mathcal{M}_{m \times n}(\mathbb{C})$ с элементами $c_{lj} = a_{lj} + ib_{lj}$, где $a_{lj}, b_{lj} \in \mathbb{R}$ и i – мнимая единица, очевидно, справедливо представление

$$C = A + iB, \quad A = \{a_{lj}\} \in \mathcal{M}_{m \times n}(\mathbb{R}), \quad B = \{b_{lj}\} \in \mathcal{M}_{m \times n}(\mathbb{R}).$$

Как и для комплексных чисел, матрицы A и B называются, соответственно, вещественной и мнимой частями матрицы C и обозначаются $A = \Re C$, $B = \Im C$. Равенство двух матриц $C_1 = C_2$ с комплексными элементами $c_{lj}^{(1)} = a_{lj}^{(1)} + ib_{lj}^{(1)}$, $c_{lj}^{(2)} = a_{lj}^{(2)} + ib_{lj}^{(2)}$ означает равенство двух пар вещественных матриц:

$$C_1 = C_2 \iff \begin{cases} A_1 = \Re C_1 = A_2 = \Re C_2, \\ B_1 = \Im C_1 = B_2 = \Im C_2. \end{cases}$$

1.6. Транспонированная и сопряженная матрицы

Определение 1.7. Пусть $A \in \mathcal{M}_{m \times n}$, матрица $B = A' \in \mathcal{M}_{n \times m}$ называется транспонированной¹ по отношению к матрице A , если $b_{lj} = a_{jl}$, где $l = 1, 2, \dots, n$, $j = 1, 2, \dots, m$.

Переход от матрицы A к матрице A' обычно называется операцией транспонирования.

В соответствии с этим определением элементы l -го столбца, где $l = 1, 2, \dots, n$ (или j -й строки, где $j = 1, 2, \dots, m$), матрицы A совпадают с соответствующими элементами l -й строки (j -го столбца) матрицы A' .

Например, если $A = \begin{bmatrix} -1 & 2 & 5 \\ -3 & 4 & -7 \end{bmatrix}$, то $A' = \begin{bmatrix} -1 & -3 \\ 2 & 4 \\ 5 & -7 \end{bmatrix}$.

Утверждение 1.4. Справедливы равенства:

- 1) $(\alpha A)' = \alpha A'$;
- 2) $(A + B)' = A' + B'$; 3) $(A')' = A$.

Доказательство. Докажем, например, последнее соотношение. Пусть $B = A' \in \mathcal{M}_{n \times m}$, тогда, согласно определению 1.7, для элементов b'_{jl} матрицы $B' = (A')'$ можно написать: $b'_{jl} = b_{lj} = a_{jl}$, где $l = 1, 2, \dots, n$, $j = 1, 2, \dots, m$, что и требовалось доказать. Аналогично доказываются первые 2 соотношения. ■

¹ Иногда транспонированную матрицу обозначают A^t .

Определение 1.8. Пусть $C \in \mathcal{M}_{m \times n}(\mathbb{C})$, матрица $D = \overline{C} \in \mathcal{M}_{m \times n}(\mathbb{C})$ называется комплексно-сопряженной по отношению к матрице C , если $d_{lj} = \overline{c_{lj}}$, где $l = 1, 2, \dots, m$; $j = 1, 2, \dots, n$. Если $c_{lj} = a_{lj} + ib_{lj}$, $a_{lj}, b_{lj} \in \mathbb{R}$, то $d_{lj} = a_{lj} - ib_{lj}$.

Утверждение 1.5. Справедливы равенства:

- 1) $\overline{(\alpha A)} = \overline{\alpha} \overline{A}$;
- 2) $\overline{(A + B)} = \overline{A} + \overline{B}$;
- 3) $\overline{(\overline{A})} = A$;
- 4) $\overline{(A')} = (\overline{A})'$.

Справедливость перечисленных свойств следует из свойств операции комплексного сопряжения.

Отметим, что если $\overline{A} = A$, то $A \in \mathcal{M}_{m \times n}(\mathbb{R})$, т. е. все элементы матрицы A являются вещественными числами.

Определение 1.9. Пусть $C \in \mathcal{M}_{m \times n}(\mathbb{C})$, матрица

$$D = C^* \in \mathcal{M}_{n \times m}(\mathbb{C})$$

называется сопряженной по отношению к матрице C , если $d_{lj} = \overline{c_{jl}}$, где $l = 1, 2, \dots, n$; $j = 1, 2, \dots, m$. Если $c_{jl} = a_{jl} + ib_{jl}$, где $a_{jl}, b_{jl} \in \mathbb{R}$, то $d_{lj} = a_{jl} - ib_{jl}$.

Утверждение 1.6. Справедливы равенства:

- 1) $(\alpha A)^* = \overline{\alpha} A^*$;
- 2) $(A + B)^* = A^* + B^*$;
- 3) $(A^*)^* = A$; 4) $A^* = \overline{(A')} = (\overline{A})'$.

Доказать перечисленные свойства предоставляется читателю.

Пример 1.4. Если матрица $C = \begin{bmatrix} 1+i & 2-i \\ 3 & -1-i \\ i & 1 \end{bmatrix}$, то

$$C^* = \begin{bmatrix} 1-i & 3 & -i \\ 2+i & -1+i & 1 \end{bmatrix}.$$

Если ввести обозначение $C = A + iB$, $A, B \in \mathcal{M}_{3 \times 2}(\mathbb{R})$, то

$$A = \begin{bmatrix} 1 & 2 \\ 3 & -1 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -1 \\ 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad A' = \begin{bmatrix} 1 & 3 & 0 \\ 2 & -1 & 1 \end{bmatrix}, \quad B' = \begin{bmatrix} 1 & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix}$$

и, следовательно, $A' - iB' = \begin{bmatrix} 1-i & 3 & -i \\ 2+i & -1+i & 1 \end{bmatrix} = C^*.$ •

Определение 1.10. Квадратная матрица A называется симметричной, если $A' = A$.

Определение 1.11. Квадратная матрица A называется самосопряженной, если $A^* = A$.

Замечание 1.7. Если матрица A симметрична, то $a_{lj} = a_{jl}$. Здесь $1 \leq l \leq n$, $1 \leq j \leq n$, где n – порядок матрицы A . Если A самосопряженная, то $a_{lj} = \overline{a_{jl}}$ и, следовательно, $a_{ll} \in \mathbb{R}$, т. е. на диагонали самосопряженной матрицы стоят вещественные числа. $\otimes$

1.7. Умножение матриц

Определение 1.12. Произведением матрицы-строки длиной n

$A = [a_1 \ a_2 \ \dots \ a_n]$ на матрицу-столбец высотой n $B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$ называется

одноэлементная матрица $C = [c] \in \mathcal{M}_{1 \times 1}$ (пишут $C = AB$), где

$$c = a_1 b_1 + a_2 b_2 + \dots + a_n b_n = \sum_{p=1}^n a_p b_p, \quad (1.28)$$

т. е. число c равно сумме произведений элементов a_i и b_i с одинаковыми номерами.

Произведением матрицы $A \in \mathcal{M}_{m \times n}$

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} A^{(1)} \\ A^{(2)} \\ \vdots \\ A^{(m)} \end{bmatrix} \quad (1.29)$$

на матрицу-столбец (высотой n) $B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$ называется матрица-столбец

(высотой m) $C = AB = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix}$, где $c_i = A^{(i)} B$. С учетом (1.28):

$$c_i = a_{i1} b_1 + a_{i2} b_2 + \dots + a_{in} b_n = \sum_{p=1}^n a_{ip} b_p. \quad (1.30)$$

Произведением матрицы $A \in \mathcal{M}_{m \times n}$ (1.29) на матрицу $B \in \mathcal{M}_{n \times k}$

$$B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1k} \\ b_{21} & b_{22} & \dots & b_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nk} \end{bmatrix} = [B_{(1)} \ B_{(2)} \ \dots \ B_{(k)}]$$

называется матрица $C = \{c_{lj}\} = AB = [C_{(1)} \ C_{(2)} \ \dots \ C_{(k)}] \in \mathcal{M}_{m \times k}$, где $C_{(j)} = AB_{(j)}$, $j = 1, 2, \dots, k$. С учетом (1.30) для нахождения элементов матрицы C получаем следующую формулу:

$$c_{lj} = A^{(l)} B_{(j)} = a_{l1}b_{1j} + a_{l2}b_{2j} + \dots + a_{ln}b_{nj} = \sum_{p=1}^n a_{lp}b_{pj}, \quad (1.31)$$

где $l = 1, 2, \dots, m$; $j = 1, 2, \dots, k$.

Замечание 1.8. Как следует из равенства (1.31), размеры перемножаемых матриц должны быть согласованы, а именно, число столбцов первой матрицы должно равняться числу строк второй матрицы. Если это условие не выполнено, то произведение матриц не определено. $\otimes$

Пример 1.5.

$$A = \begin{bmatrix} 1 & -3 & 2 \\ 2 & 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 5 & 2 & 4 \\ 1 & -1 & 1 & 3 \\ 2 & 1 & 1 & 1 \end{bmatrix}, \quad C = AB = \begin{bmatrix} 1 & 10 & 1 & -3 \\ 1 & 9 & 5 & 11 \end{bmatrix}.$$

Элементы c_{lj} матрицы C находят по формулам (1.31), например:

$$c_{12} = A^{(1)} B_{(2)} = \begin{bmatrix} 1 & -3 & 2 \end{bmatrix} \begin{bmatrix} 5 \\ -1 \\ 1 \end{bmatrix} = 1 \cdot 5 + (-3) \cdot (-1) + 2 \cdot 1 = 10. \quad \bullet$$

Умножение произвольных матриц не коммутативно, т.е. $AB \neq BA$. Например, для

$$A = [1 \ 2], \quad B = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad AB = [1], \quad BA = \begin{bmatrix} -1 & -2 \\ 1 & 2 \end{bmatrix}$$

и, конечно, $AB \neq BA$. Равенство

$$AB = BA \quad (1.32)$$

если и возможно, то только для квадратных матриц одинакового порядка. Действительно, пусть матрица $A \in \mathcal{M}_{m \times n}$, а $B \in \mathcal{M}_{p \times k}$, тогда, так как определено AB , то $n = p$, и так как определено BA , то $k = m$. Значит, $AB \in \mathcal{M}_m$, $BA \in \mathcal{M}_n$, а из равенства (1.32) следует, что $n = m$. Но

и этого недостаточно. Пусть, например, $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$. Тогда

$$AB = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \neq BA = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}.$$

Определение 1.13. Если для квадратных матриц A и B одинакового порядка выполняется равенство (1.32), то эти матрицы называются коммутирующими или перестановочными.

Легко проверить, что для любой квадратной матрицы $A \in \mathcal{M}_n$ и единичной матрицы $I = I_n \in \mathcal{M}_n$ того же порядка выполнено

$$AI = IA = A. \quad (1.33)$$

Это равенство оправдывает название „единичная матрица“. Отметим, что единичная матрица единственна. Действительно, если бы другая матрица обладала этим же свойством

$$AE = EA = A, \quad (1.34)$$

то из (1.34) $IE = I$ и из (1.33) $IE = E$, откуда следует: $I = E$.

Пусть $A \in \mathcal{M}_{m \times n}$ – произвольная матрица, тогда умножение ее на согласованную по размеру нулевую матрицу дает $A\Theta_{n \times k} = \Theta_{m \times k}$ и $\Theta_{k \times m}A = \Theta_{k \times n}$.

Установим теперь некоторые свойства введенной операции умножения матриц. Предполагается, что размеры всех матриц согласованы, т.е. все рассматриваемые произведения определены.

Теорема 1.2. Операция умножения матриц обладает следующими свойствами:

1) ассоциативностью:

$$(AB)C = A(BC); \quad (1.35)$$

2) дистрибутивностью:

$$A(B + C) = AB + AC, \quad (A + B)C = AC + BC; \quad (1.36)$$

3) если λ – число, то

$$\lambda(AB) = (\lambda A)B = A(\lambda B); \quad (1.37)$$

4)

$$(AB)' = B'A', \quad (AB)^* = B^*A^*. \quad (1.38)$$

Доказательство. Докажем сначала ассоциативность умножения матриц. Пусть $A \in \mathcal{M}_{m \times n}$, $B \in \mathcal{M}_{n \times k}$, $C \in \mathcal{M}_{k \times p}$. Определим $AB = U \in$

$\in \mathcal{M}_{m \times k}$, $BC = V \in \mathcal{M}_{m \times p}$, $(AB)C = S \in \mathcal{M}_{m \times p}$ и $A(BC) = T \in \mathcal{M}_{m \times p}$. Найдем выражение для элементов s_{gl} матрицы S . Согласно (1.31), где $s_{gl} =$

$$= U^{(g)}C_{(l)} = \sum_{j=1}^k u_{gj}c_{jl}, \quad u_{gj} = A^{(g)}B_{(j)} = \sum_{r=1}^n a_{gr}b_{rj}. \quad \text{Следовательно,}$$

$$s_{gl} = \sum_{j=1}^k \left(\sum_{r=1}^n a_{gr}b_{rj} \right) c_{jl} = \sum_{j=1}^k \sum_{r=1}^n a_{gr}b_{rj}c_{jl}. \quad (1.39)$$

Для элементов t_{gl} матрицы T , согласно (1.31), где

$$t_{gl} = A^{(g)}V_{(l)} = \sum_{r=1}^n a_{gr}v_{rl}, \quad v_{rl} = B^{(r)}C_{(l)} = \sum_{j=1}^k b_{rj}c_{jl}.$$

Следовательно, получаем

$$t_{gl} = \sum_{r=1}^n a_{gr} \left(\sum_{j=1}^k b_{rj}c_{jl} \right) = \sum_{r=1}^n \sum_{j=1}^k a_{gr}b_{rj}c_{jl}. \quad (1.40)$$

Сравнивая (1.39) и (1.40) видим, что соответствующие элементы матриц S и T совпадают, и, следовательно, равенство (1.35) доказано.

Докажем первую из формул (1.36). Пусть $A \in \mathcal{M}_{m \times n}$, а $B, C \in \mathcal{M}_{n \times k}$. Введем обозначения: $D = A(B + C)$, $U = AB$, $V = AC$. Запишем выражения для элементов матриц D , U и V :

$$d_{lj} = A^{(l)}(B_{(j)} + C_{(j)}) = \sum_{r=1}^n a_{lr}(b_{rj} + c_{rj}) = \sum_{r=1}^n (a_{lr}b_{rj} + a_{lr}c_{rj}), \quad (1.41)$$

$$u_{lj} = A^{(l)}B_{(j)} = \sum_{r=1}^n a_{lr}b_{rj}, \quad v_{lj} = A^{(l)}C_{(j)} = \sum_{r=1}^n a_{lr}c_{rj}. \quad (1.42)$$

Сравнивая (1.41) и (1.42) видим, что $d_{lj} = u_{lj} + v_{lj}$, значит, первое из равенств (1.36) доказано.

Докажем справедливость (1.37). Обозначим через $C = \lambda(AB)$, $D = (\lambda A)B$, $G = A(\lambda B)$, тогда, учитывая (1.31), можно написать:

$$c_{lj} = \lambda(A^{(l)}B_{(j)}) = \lambda \sum_{r=1}^n a_{lr}b_{rj},$$

$$d_{lj} = (\lambda A^{(l)})B_{(j)} = \sum_{r=1}^n (\lambda a_{lr})b_{rj},$$

$$g_{lj} = A^{(l)}(\lambda B_{(j)}) = \sum_{r=1}^n a_{lr}(\lambda b_{rj}).$$

Полученные выражения совпадают, что и доказывает (1.37).

Докажем теперь справедливость первого из равенств (1.38). Пусть $A \in \mathcal{M}_{m \times n}$ и $B \in \mathcal{M}_{n \times k}$, тогда, согласно определению 1.7, $B' \in \mathcal{M}_{k \times n}$ и $A' \in \mathcal{M}_{n \times m}$. Обозначим $AB = C \in \mathcal{M}_{m \times k}$, $G = B'A' = G \in \mathcal{M}_{k \times m}$, $(AB)' = D \in \mathcal{M}_{k \times m}$. Тогда

$$d_{lj} = c_{jl} = A^{(j)}B_{(l)} = \sum_{r=1}^n a_{jr}b_{rl}, \quad g_{lj} = (B')^{(l)}(A')_{(j)} = \sum_{r=1}^n b_{rl}a_{jr}.$$

Значит, $d_{lj} = g_{lj}$, т. е. $(AB)' = B'A'$. Аналогично доказывается (учитывая определение 1.9), что $(AB)^* = B^*A^*$. ■

Следствие 1.2. $(ABC)' = C'B'A'$, $(ABC)^* = C^*B^*A^*$.

Доказательство. Из равенства (1.35) и первого из равенств (1.38) следует $(ABC)' = C'(AB)' = C'B'A'$. Аналогично доказывается второе равенство. ■

Если A – квадратная матрица, то можно определить ее степени: $A^0 = I$ – единичная матрица того же порядка, далее по индукции:

$$A^1 = A, \quad A^2 = AA, \quad A^3 = AA^2 = AAA, \quad \dots, \quad A^n = AA^{n-1} = \underbrace{A \cdots A}_n.$$

Если матрица A является диагональной, $A = \text{diag}[\lambda_1, \lambda_2, \dots, \lambda_n]$, то

$$A^2 = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = \begin{bmatrix} \lambda_1^2 & 0 & \cdots & 0 \\ 0 & \lambda_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^2 \end{bmatrix} =$$

$$= \text{diag}[\lambda_1^2, \lambda_2^2, \dots, \lambda_n^2],$$

$$\dots \dots \dots$$

$$A^m = \text{diag}[\lambda_1^m, \lambda_2^m, \dots, \lambda_n^m].$$

Пусть $\mathcal{P}(x) = b_0 + b_1x + \cdots + b_mx^m$ – многочлен, где $b_i \in \mathbb{C}(\mathbb{R})$. Если A – квадратная матрица с элементами из $\mathbb{C}(\mathbb{R})$, то можно определить многочлен от матрицы

$$\mathcal{P}(A) = b_0A^0 + b_1A^1 + \cdots + b_mA^m = b_0I + b_1A + \cdots + b_mA^m,$$

значение которого является квадратной матрицей того же порядка.

1.8. Матричная запись системы линейных алгебраических уравнений

В качестве примера использования определения 1.12 приведем матричную запись системы линейных алгебраических уравнений (1.1) из m

уравнений с n неизвестными:

[illegible]

Для задания системы достаточно указать числа a_{ij} – коэффициенты системы (1.1) и числа b_j – правые части системы. Очевидно, что набору чисел a_{ij} соответствует матрица

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix},$$

которую, как указывалось ранее, называют матрицей системы (1.1), а правые части системы запишем в виде матрицы-столбца

$$B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

Введем также в рассмотрение матрицу-столбец из неизвестных системы

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Тогда, используя операцию умножения матриц (определение 1.12) и условие равенства матриц (определение 1.5), рассматриваемую систему (1.1) можно записать в виде матричного уравнения

$$AX = B. \quad (1.43)$$

Действительно, приравнивая i -е элементы матриц AX и B , получаем i -е уравнение исходной системы

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i, \quad i = 1, 2, \dots, m.$$

2. ОПРЕДЕЛИТЕЛИ. ТЕОРЕМА КРАМЕРА. МАТРИЧНЫЕ УРАВНЕНИЯ

2.1. Определители квадратных матриц

Здесь будет определена некоторая функция, ставящая в соответствие каждой квадратной матрице A число – определитель (или детерминант) матрицы A . Определитель матрицы A обозначается $\det(A)$. Следуя [4], дадим индукционное определение этой функции.

Определение 2.1. 1. *Определитель матрицы первого порядка $A = [a_{11}]$ принимается равным ее единственному элементу: $\det(A) = a_{11}$.*

2. *Предположим, что известно вычисление определителей матриц до порядка $(n-1)$ включительно. Рассмотрим квадратную матрицу n -го порядка:*

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1k} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{ik} & \cdots & a_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nk} & \cdots & a_{nn} \end{bmatrix} \quad i.$$

k

Если исключить i -ю строку и k -й столбец, то получим квадратную матрицу $(n-1)$ -го порядка $\tilde{A}_{ik}$, определитель которой назовем дополнительным минором (или просто минором) исходной матрицы A , отвечающим элементу a_{ik} (минор элемента a_{ik} матрицы A), и обозначим его $M_{ik}(A)$ ¹. Таким образом,

$$M_{ik}(A) = \det(\tilde{A}_{ik}) = \det \begin{bmatrix} a_{11} & \cdots & a_{1,k-1} & a_{1,k+1} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,k-1} & a_{i-1,k+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,k-1} & a_{i+1,k+1} & \cdots & a_{i+1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{n,k-1} & a_{n,k+1} & \cdots & a_{nn} \end{bmatrix}.$$

Число $A_{ik} = (-1)^{i+k} M_{ik}(A)$ назовем алгебраическим дополнением элемента a_{ik} матрицы A .

3. *Определитель квадратной матрицы n -го порядка определим как сумму произведений всех элементов первой строки на их алгебраические дополнения:*

$$\det(A) = a_{11}A_{11} + \cdots + a_{1n}A_{1n} = \sum_{k=1}^n a_{1k}A_{1k} = \sum_{k=1}^n (-1)^{1+k} a_{1k}M_{1k}(A). \quad (2.1)$$

¹ Если очевидно, минор какой матрицы рассматривается, то обозначаем его M_{ik} .

Формулу (2.1) обычно называют разложением определителя матрицы по элементам первой строки, или просто по первой строке.

Может показаться, что определение определителя неконструктивно: определитель матрицы n -го порядка выражается через определители матриц $(n - 1)$ -го порядка, которые сами не определены. Но для определения чисел $M_{1k}(A)$ можно воспользоваться той же формулой (2.1), тем самым выразив миноры $M_{1k}(A)$ через определители матриц $(n - 2)$ -го порядка. Продолжим этот процесс, пока не придем к матрицам первого порядка, а для них определитель определен непосредственно.

Пример 2.1. Вычислим определители матриц второго и третьего порядков. Для матрицы

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

имеем $M_{11} = a_{22}$, $M_{12} = a_{21}$, $A_{11} = (-1)^{1+1}a_{22} = a_{22}$, $A_{12} = (-1)^{2+1}a_{21} = -a_{21}$, и, следовательно, согласно (2.1),

$$\det(A) = \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

Для матрицы

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

очевидно, что

$$M_{11} = \det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}, \quad M_{12} = \det \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix}, \quad M_{13} = \det \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$$

и

$$\begin{aligned} \det(A) &= a_{11} \det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} - a_{12} \det \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} + a_{13} \det \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} = \bullet \\ &= a_{11}a_{22}a_{33} - a_{11}a_{32}a_{23} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{21}a_{32}a_{13} - a_{13}a_{22}a_{31}. \end{aligned}$$

Вычислим определитель единичной матрицы $I_n = \text{diag}[1, 1, \dots, 1]$. Для нее $M_{11} = \det(I_{n-1})$, где I_{n-1} — единичная матрица порядка $n - 1$, и $a_{1k} = 0$, если $k \neq 1$. Поэтому $\det(I_n) = \det(I_{n-1})$. Применяя это равенство $(n - 1)$ раз, получим $\det(I_n) = \det(I_1) = 1$.

Дадим оценку числа N_n арифметических операций, которое необходимо произвести, чтобы вычислить определитель матрицы n -го порядка по формуле (2.1). Очевидно, что $N_2 = 3$. Для того чтобы вычислить определитель матрицы n -го порядка, надо, прежде всего, вычислить n определителей матриц $(n - 1)$ -го порядка, т. е. произвести nN_{n-1} операций, а затем

произвести n умножений на элементы первой строки и $n - 1$ действий сложения. Следовательно, $N_n = nN_{n-1} + 2n - 1$ или $N_n + 2 = n(N_{n-1} + 2) + 1$. Значит,

$$N_n + 2 > n(N_{n-1} + 2).$$

Это неравенство справедливо для любого n , в частности,

$$\begin{aligned} N_{n-1} + 2 &> (n-1)(N_{n-2} + 2), \\ N_{n-2} + 2 &> (n-2)(N_{n-3} + 2), \\ &\dots\dots\dots \\ N_3 + 2 &> 3(N_2 + 2). \end{aligned}$$

Следовательно,

$$\begin{aligned} N_n + 2 &> n(N_{n-1} + 2) > n(n-1)(N_{n-2} + 2) > \\ &> n(n-1)(n-2)(N_{n-3} + 2) > \dots > n(n-1) \dots 3(N_2 + 2) = \frac{5}{2}n!. \end{aligned}$$

Таким образом, справедлива оценка

$$N_n > \frac{5}{2}n! - 2. \quad (2.2)$$

Из оценки (2.2) следует, в частности, что при вычислении определителя матрицы десятого порядка, исходя из его определения, т. е. по формуле (2.1), потребуется выполнить около $9 \cdot 10^6$ операций. Такое количество операций невозможно произвести без быстродействующих ЭВМ. В случае очень больших n ($n \gg 10$) применение формулы (2.1) для вычисления определителей становится бессмысленным даже с использованием ЭВМ.

Далее будет рассмотрен способ вычисления определителей, требующий значительно меньшего числа операций. Но прежде установим ряд свойств, которые окажутся полезными как при исследовании, так и при вычислении определителей матриц.

2.2. Основные свойства определителей

Доказательства приведенных здесь свойств проводятся методом полной математической индукции и представляют определенную трудность для студентов, поэтому при первом чтении их можно пропустить.

Утверждение 2.1. *Для каждой матрицы A порядка $n \geq 2$ справедливо равенство*

$$\det(A) = \sum_{i=1}^n a_{i1}A_{i1} = \sum_{i=1}^n (-1)^{i+1}a_{i1}M_{i1}(A). \quad (2.3)$$

Формула (2.3) называется разложением определителя матрицы по первому столбцу.

Доказательство. Очевидно, что для матриц второго порядка равенство (2.3) справедливо. Предполагая справедливость (2.3) для матриц порядка $n - 1$ (индукционное предположение), докажем это равенство для матриц порядка n . Для этого перепишем формулу (2.1), определяющую число $\det(A)$, выделив первый член суммы:

$$\det(A) = a_{11}M_{11}(A) + \sum_{k=2}^n (-1)^{1+k} a_{1k} M_{1k}(A). \quad (2.4)$$

При любом $k \geq 2$ в матрицу $\tilde{A}_{1k}$ (которая получается из матрицы исключением первой строки и k -го столбца) входит (без своего первого элемента) первый столбец матрицы A . Пользуясь предположением индукции, можно разложить $M_{1k}(A)$ по этому столбцу. При этом следует учесть, что i -я строка матрицы A ($i \geq 2$) в матрицу $\tilde{A}_{1k}$ входит под номером $i - 1$, поэтому при любом $k \geq 2$

$$M_{1k}(A) = \sum_{i=2}^n (-1)^i a_{i1} M_{i1}(\tilde{A}_{1k}).$$

Здесь $M_{i1}(\tilde{A}_{1k})$ – определитель матрицы порядка $n - 2$, получаемой из матрицы $\tilde{A}_{1k}$ исключением ее $(i - 1)$ -й строки и первого столбца или, что то же самое, получаемой из A исключением первой и i -й строк и первого и k -го столбцов. Отметим, что $M_{i1}(\tilde{A}_{1k}) = M_{1k}(\tilde{A}_{i1})$, так как это определители одной и той же матрицы, которая получается из матрицы A исключением первой и i -й строк и первого и k -го столбцов.

Подставляя в (2.4) полученное выражение для $M_{1k}(A)$, находим

$$\begin{aligned} \det(A) &= a_{11}M_{11}(A) + \sum_{k=2}^n \left((-1)^{1+k} a_{1k} \sum_{i=2}^n (-1)^i a_{i1} M_{i1}(\tilde{A}_{1k}) \right) = \\ &= a_{11}M_{11}(A) + \sum_{k=2}^n \sum_{i=2}^n (-1)^{1+k+i} a_{1k} a_{i1} M_{i1}(\tilde{A}_{1k}) = \\ &= a_{11}M_{11}(A) + \sum_{k=2}^n \sum_{i=2}^n (-1)^{1+k+i} a_{1k} a_{i1} M_{1k}(\tilde{A}_{i1}). \end{aligned}$$

Изменим порядок суммирования и вынесем множитель, не зависящий от k , за знак внутренней суммы:

$$\det(A) = a_{11}M_{11}(A) + \sum_{i=2}^n (-1)^{i+1} a_{i1} \sum_{k=2}^n (-1)^k a_{1k} M_{1k}(\tilde{A}_{i1}).$$

Нетрудно заметить, что внутренняя сумма представляет собой результат применения формулы (2.1) к минору $M_{i1}(A)$. Действительно,

$$M_{i1}(A) = \sum_{k=2}^n (-1)^k a_{1k} M_{1k}(\tilde{A}_{i1}),$$

так как в матрице $\tilde{A}_{i1}$, определителем которой является минор $M_{i1}(A)$, по сравнению с A , исключен первый столбец, i -я строка и все номера оставшихся столбцов уменьшены на 1. Теперь можно записать (2.4) в виде

$$\det(A) = a_{11}M_{11}(A) + \sum_{i=2}^n (-1)^{i+1} a_{i1}M_{i1}(A),$$

что совпадает с доказываемым равенством (2.3). ■

Рассмотрим частный случай квадратной матрицы, так называемую верхнюю треугольную матрицу

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_{nn} \end{bmatrix},$$

у которой все элементы, стоящие ниже главной диагонали, равны нулю ($a_{ik} = 0$ при $i > k$).

Следствие 2.1. *Если матрица A является верхней треугольной, то $\det(A) = a_{11}a_{22} \cdots a_{nn}$.*

Доказательство. Если использовать разложение $\det(A)$ по первому столбцу (формула (2.3)), то получим:

$$\begin{aligned} \det(A) &= a_{11} \det \begin{bmatrix} a_{22} & a_{23} & \dots & a_{2n} \\ 0 & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix} = a_{11}a_{22} \det \begin{bmatrix} a_{33} & \dots & a_{3n} \\ \vdots & \ddots & \vdots \\ 0 & \dots & a_{nn} \end{bmatrix} = \\ &= a_{11}a_{22} \cdots a_{nn}. \end{aligned}$$

Значит, определитель верхней треугольной матрицы равен произведению ее диагональных элементов и для его вычисления требуется $n - 1$ операция умножения. ■

Совершенно аналогично, используя (2.1), получаем, что определитель нижней треугольной матрицы

$$B = \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix},$$

у которой $a_{ik} = 0$ при $k > i$, также равен произведению диагональных элементов $\det(B) = a_{11}a_{22} \cdots a_{nn}$.

Утверждение 2.2. *Для любой квадратной матрицы*

$$\det(A') = \det(A).$$

Доказательство. Докажем это утверждение по индукции. Для матриц первого порядка оно очевидно. Предположим, что предложение верно для матриц порядка $n - 1$, и докажем его для матриц порядка n . Пусть $\tilde{A}_{1j}$ – матрица, получаемая из A исключением j -го столбца и первой строки, а $\tilde{B}_{j1}$ – матрица, получаемая из A' исключением j -й строки и первого столбца. Легко видеть, что $\tilde{B}_{j1} = (\tilde{A}_{1j})'$, поэтому из утверждения индукции следует, что $\det(\tilde{B}_{j1}) = \det(\tilde{A}_{1j})$, или, иными словами, минор $M_{1j}(A)$ элемента a_{1j} в матрице A равен минору $M_{j1}(B)$ элемента b_{j1} в матрице $B = A'$. Кроме того, $a_{1j} = b_{j1}$, и разложение $\det(A)$ по первой строке совпадает с разложением $\det(B)$ по первому столбцу, т. е.

$$\det(A) = \sum_{j=1}^n (-1)^{1+j} a_{1j} M_{1j}(A) = \sum_{j=1}^n (-1)^{j+1} a_{j1} M_{j1}(B) = \det(B)$$

(при получении последнего равенства использовано (2.3)). ■

Утверждение 2.3. *Для любой квадратной матрицы*

$$\det(\bar{A}) = \overline{\det(A)}.$$

Другими словами, операции комплексного сопряжения матрицы и вычисления ее определителя перестановочны. Доказательство этого утверждения проводится по индукции с учетом свойств операции комплексного сопряжения.

Следствие 2.2. *Для любой квадратной матрицы*

$$\det(A^*) = \overline{\det(A)}.$$

Для доказательства следствия достаточно воспользоваться утверждениями 2.2 и 2.3.

Замечание 2.1. Из утверждения 2.2 следует, что если справедливо утверждение об определителях, касающееся строк соответствующих матриц, то верно и аналогичное утверждение, касающееся столбцов, и наоборот. В связи с этим следующие далее утверждения достаточно доказать только для строк (или столбцов). $\otimes$

Утверждение 2.4. Пусть $k < l$, обозначим

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kn} \\ \vdots & \ddots & \vdots \\ a_{l1} & \cdots & a_{ln} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}, \quad B = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{l1} & \cdots & a_{ln} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kn} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix},$$

тогда $\det(B) = -\det(A)$.

Заметим, что матрица B получена из матрицы A перестановкой l -й и k -й строк. Аналогичное утверждение справедливо и при перестановке столбцов.

Доказательство. Для матриц второго порядка это очевидно:

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{12}a_{21} = -(a_{12}a_{21} - a_{11}a_{22}) = -\det \begin{bmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{bmatrix}.$$

Пусть $n \geq 3$, предположим, что утверждение верно для матриц порядка $n-1$, и докажем его для матриц порядка n . Начнем доказательство для двух соседних строк.

Пусть номера переставляемых строк в матрице A n -го порядка k и $k+1$. Напишем разложение определителя матрицы A по первому столбцу, выделив в нем 2 слагаемых, соответствующих переставляемым строкам:

$$\begin{aligned} \det(A) &= (-1)^{k+1}a_{k1}M_{k1}(A) + (-1)^{k+2}a_{k+1,1}M_{k+1,1}(A) + \\ &+ \sum_{i \neq k, k+1} (-1)^{i+1}a_{i1}M_{i1}(A). \end{aligned}$$

Аналогично, для матрицы B , которая получается из A перестановкой k -й и $(k+1)$ -й строк,

$$\begin{aligned} \det(B) &= (-1)^{k+1}a_{k+1,1}M_{k1}(B) + (-1)^{k+2}a_{k1}M_{k+1,1}(B) + \\ &+ \sum_{i \neq k, k+1} (-1)^{i+1}a_{i1}M_{i1}(B). \end{aligned}$$

При $i \neq k, k+1$ $M_{i1}(A)$ и $M_{i1}(B)$ являются определителями матриц $(n-1)$ -го порядка, отличающихся друг от друга перестановкой двух

соседних строк. Следовательно, по предположению индукции, $M_{i1}(B) = -M_{i1}(A)$ при $i \neq k, k+1$.

Матрицы $\tilde{A}_{k1}$ и $\tilde{B}_{k+1,1}$ совпадают, так как они получаются соответственно вычеркиванием k -й строки и первого столбца из матрицы A и $(k+1)$ -й строки и первого столбца из матрицы B . Поэтому

$$M_{k1}(A) = \det(\tilde{A}_{k1}) = \det(\tilde{B}_{k+1,1}) = M_{k+1,1}(B).$$

Аналогично показывается, что $M_{k1}(B) = M_{k+1,1}(A)$. Значит,

$$\begin{aligned} \det(B) &= (-1)^{k+1} a_{k+1,1} M_{k1}(B) + (-1)^{k+2} a_{k1} M_{k+1,1}(B) + \\ &+ \sum_{i \neq k, k+1} (-1)^{i+1} a_{i1} M_{i1}(B) = -(-1)^{k+2} a_{k+1,1} M_{k+1,1}(A) - \\ &- (-1)^{k+1} a_{k1} M_{k1}(A) - \sum_{i \neq k, k+1} (-1)^{i+1} a_{i1} M_{i1}(A) = -\det(A). \end{aligned}$$

Пусть теперь в матрице A порядка n переставляются строки с номерами i и j и пусть для определенности $i < j$. Тогда между i -й и j -й строками находятся $j-i-1$ строк. Перестановку i -й и j -й строк можно осуществить, переставляя только соседние строки: сначала j -ю строку переставляем последовательно с $j-i$ строками, стоящими над ней (последней из них будет i -я), затем i -ю переставляем на место j -й, меняя местами с каждой из $j-i-1$ строк ниже нее. Всего будет проделано нечетное число $2(j-i)-1$ перестановок соседних строк. Поскольку при каждой из них определитель умножается на (-1) (или, как часто говорят, меняет знак), то при перестановке i -й и j -й строк он также должен умножиться на (-1) . Например, если в матрице пятого порядка требуется перестановка второй и пятой строк, то ее можно осуществить следующей серией перестановок соседних строк:

$$\begin{array}{cccccc} 1 & 1 & 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 5 & 5 & 5 \\ 3 & \longrightarrow 3 & \longrightarrow 5 & \longrightarrow 2 & \longrightarrow 3 & \longrightarrow 3 \\ 4 & 5 & 3 & 3 & 2 & 4 \\ 5 & 4 & 4 & 4 & 4 & 2 \end{array}$$

(здесь в столбец записаны номера строк в процессе перестановок). Всего требуется сделать 5 перестановок, т. е. нечетное число. ■

Следствие 2.3. *Если все элементы какого-нибудь столбца (или какой-нибудь строки) матрицы равны соответствующим элементам другого столбца (или строки), то определитель матрицы равен нулю.*

Другими словами, определитель матрицы с двумя равными столбцами (строками) равен нулю.

Доказательство. Если поменять местами одинаковые столбцы (или строки) в матрице, то она не изменится, но, согласно утверждению 2.4, получим, что $\det(A) = -\det(A)$, значит, $\det(A) = 0$. ■

Теорема 2.1. Для каждой матрицы A порядка $n \geq 2$ при произвольном i ($1 \leq i \leq n$) справедлива формула

$$\det(A) = \sum_{k=1}^n a_{ik} A_{ik} = \sum_{k=1}^n (-1)^{k+i} a_{ik} M_{ik}(A) \quad (2.5)$$

и при произвольном j ($1 \leq j \leq n$) – формула

$$\det(A) = \sum_{k=1}^n a_{kj} A_{kj} = \sum_{k=1}^n (-1)^{k+j} a_{kj} M_{kj}(A). \quad (2.6)$$

Равенство (2.5) называется разложением определителя по i -й строке, а (2.6) – разложением определителя по j -му столбцу.

Доказательство. Заметим, что при $i = 1$ формула (2.5) есть определение 2.1 определителя (формула (2.1)), а при $j = 1$ формула (2.6) совпадает с доказанной в утверждении 2.1 формулой (2.3) – разложением определителя по первому столбцу. Докажем формулу (2.5) при $i \geq 2$. Для этого переставим i -ю строку матрицы на первое место так, чтобы не нарушить порядок остальных строк. Нужная перестановка будет осуществлена, если переставить i -ю строку последовательно со всеми строками, расположенными выше нее. Выше i -й строки находится $i - 1$ строк, поэтому, если B – матрица, полученная после перестановки, то $\det(A) = (-1)^{i-1} \det(B)$. Разложив $\det(B)$ по первой строке (i -й строке матрицы A), получим:

$$\det(A) = (-1)^{i-1} \sum_{k=1}^n (-1)^{k+1} a_{ik} M_{1k}(B) = \sum_{k=1}^n (-1)^{k+i} a_{ik} M_{1k}(B).$$

Ясно, что матрица $\tilde{B}_{1k}$, полученная из матрицы B исключением первой строки и k -го столбца, совпадает с матрицей $\tilde{A}_{ik}$, полученной из матрицы A исключением i -й строки и k -го столбца. Отсюда следует, что

$$M_{1k}(B) = \det(\tilde{B}_{1k}) = \det(\tilde{A}_{ik}) = M_{ik}(A),$$

а это и доказывает нужное разложение (2.5).

Формулу (2.6) можно получить аналогично. ■

2.3. Дальнейшие свойства определителей

Все утверждения будем доказывать только для столбцов, несмотря на то, что в формулировках приведены утверждения и для строк. Это можно делать благодаря замечанию 2.1.

Утверждение 2.5. Пусть

$$A = \begin{bmatrix} a_{11} & \cdots & \lambda a_{1j} & \cdots & a_{1n} \\ a_{21} & \cdots & \lambda a_{2j} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & \lambda a_{nj} & \cdots & a_{nn} \end{bmatrix}, \quad B = \begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{bmatrix},$$

тогда $\det(A) = \lambda \det(B)$. Аналогичное утверждение справедливо и тогда, когда все элементы произвольной строки матрицы A содержат общий множитель.

Доказательство. Если разложить определитель матрицы A по j -му столбцу (формула (2.6)), то получим

$$\det(A) = \sum_{k=1}^n \lambda a_{kj} A_{kj} = \lambda \sum_{k=1}^n a_{kj} A_{kj} = \lambda \det(B).$$

При доказательстве утверждения для строки следует воспользоваться формулой (2.3). ■

Следствие 2.4. Если все элементы какого-нибудь столбца (или строки) матрицы A равны нулю, то определитель матрицы равен нулю.

Доказательство. Поскольку в этом случае все элементы столбца содержат общий множитель нуль, то из только что доказанного утверждения 2.5 следует, что $\det(A) = 0$. ■

Утверждение 2.6. Если все элементы какого-нибудь столбца (или строки) матрицы представлены в виде суммы двух слагаемых, то ее определитель равен сумме определителей двух матриц, у одной из которых соответствующий столбец (соответствующая строка) состоит из первых слагаемых, а у другой – из вторых слагаемых. При этом все остальные столбцы (или строки) не меняются.

Доказательство. Докажем, например, для j -го столбца. Разложив определитель по j -му столбцу (формула (2.6)), получим

$$\det \begin{bmatrix} a_{11} & \cdots & a'_{1j} + a''_{1j} & \cdots & a_{1n} \\ a_{21} & \cdots & a'_{2j} + a''_{2j} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a'_{nj} + a''_{nj} & \cdots & a_{nn} \end{bmatrix} = \sum_{k=1}^n (a'_{kj} + a''_{kj}) A_{kj} =$$

$$\begin{aligned}
&= \sum_{k=1}^n a'_{kj} A_{kj} + \sum_{k=1}^n a''_{kj} A_{kj} = \\
&= \det \begin{bmatrix} a_{11} & \cdots & a'_{1j} & \cdots & a_{1n} \\ a_{21} & \cdots & a'_{2j} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a'_{nj} & \cdots & a_{nn} \end{bmatrix} + \det \begin{bmatrix} a_{11} & \cdots & a''_{1j} & \cdots & a_{1n} \\ a_{21} & \cdots & a''_{2j} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a''_{nj} & \cdots & a_{nn} \end{bmatrix},
\end{aligned}$$

что и требовалось доказать. ■

Отметим, что при доказательстве утверждения 2.6 использовался тот факт, что алгебраические дополнения элементов произвольного столбца (строки) не зависят от самих элементов этого столбца (этой строки) и поэтому у всех трех рассмотренных матриц алгебраические дополнения соответствующих элементов j -го столбца совпадают.

Утверждение 2.7. *Если к элементам столбца (строки) матрицы прибавить соответствующие элементы другого столбца (строки), умноженные на один и тот же множитель, то определитель полученной матрицы будет равен определителю исходной.*

Доказательство. Действительно, прибавим, например, к первому столбцу элементы второго, умноженные на λ . Тогда, используя уже доказанные утверждения 2.6, 2.5 и следствие 2.3, получим:

$$\begin{aligned}
&\det \begin{bmatrix} a_{11} + \lambda a_{12} & a_{12} & \cdots & a_{1n} \\ a_{21} + \lambda a_{22} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} + \lambda a_{n2} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = \det \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} + \\
&+ \det \begin{bmatrix} \lambda a_{12} & a_{12} & \cdots & a_{1n} \\ \lambda a_{22} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda a_{n2} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = \det \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} + \\
&+ \lambda \det \begin{bmatrix} a_{12} & a_{12} & \cdots & a_{1n} \\ a_{22} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n2} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = \det \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix},
\end{aligned}$$

что и требовалось доказать. ■

Утверждение 2.8. *Сумма произведений элементов произвольного столбца (строки) матрицы на алгебраические дополнения соответствующих элементов другого столбца (строки) равна нулю.*

Доказательство. Заменим в матрице A столбец с номером k столбцом с номером l . Согласно следствию 2.3 определитель полученной матрицы будет равен нулю. Разложим ее определитель по элементам k -го столбца:

$$0 = \sum_{i=1}^n a_{il} A_{ik}.$$

Доказательство для строк совершенно аналогично. ■

Введем теперь обозначение

$$\delta_{lk} = \begin{cases} 1, & \text{если } l = k, \\ 0, & \text{если } l \neq k. \end{cases}$$

Знак δ_{lk} принято называть *символом Кронекера*. Из теоремы 2.1 и утверждения 2.8 следуют равенства

$$\sum_{i=1}^n a_{il} A_{ik} = \sum_{k=1}^n a_{li} A_{ki} = \delta_{lk} \det(A). \quad (2.7)$$

2.4. Определители ступенчатых матриц

Определение 2.2. *Квадратная матрица вида*

$$A = \left[\begin{array}{cccc|ccc} a_{11} & a_{12} & \cdots & a_{1m} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} & 0 & \cdots & 0 \\ \hline a_{m+1,1} & a_{m+1,2} & \cdots & a_{m+1,m} & a_{m+1,m+1} & \cdots & a_{m+1,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} & a_{n,m+1} & \cdots & a_{nn} \end{array} \right] =$$

$$= \left[\begin{array}{c|c} \overset{m}{A^{(1)}} & \overset{n-m}{0} \\ \hline \overset{m}{A^{(3)}} & \overset{n-m}{A^{(2)}} \end{array} \right]_{n-m}$$

называется *ступенчатой, или блочной*.

Утверждение 2.9. *Для ступенчатых матриц*

$$\det(A) = \det(A^{(1)}) \det(A^{(2)}),$$

т. е.

$$\det(A) = \det \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{bmatrix} \det \begin{bmatrix} a_{m+1,m+1} & \cdots & a_{m+1,n} \\ \vdots & \ddots & \vdots \\ a_{n,m+1} & \cdots & a_{nn} \end{bmatrix}. \quad (2.8)$$

Доказательство. Проведем индукцию по m , т. е. по порядку матрицы $A^{(1)}$. Пусть $m = 1$, раскладывая $\det(A)$ по 1-й строке, получим:

$$\begin{aligned} \det(A) &= \det \left[\begin{array}{c|ccc} a_{11} & 0 & \cdots & 0 \\ \hline a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{array} \right] = a_{11}A_{11} + a_{21}0 + \cdots + a_{m1}0 = \\ &= a_{11}M_{11}(A) = a_{11} \det \left[\begin{array}{ccc} a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n2} & \cdots & a_{nn} \end{array} \right], \end{aligned}$$

т. е. для $m = 1$ формула (2.8) справедлива.

Пусть формула (2.8) справедлива для матриц $A^{(1)}$ порядка $m - 1$. Докажем ее справедливость для матриц порядка m . Разложим определитель матрицы A по элементам первой строки:

$$\det(A) = \sum_{i=1}^n a_{1i}A_{1i} = \sum_{i=1}^m a_{1i}A_{1i} = \sum_{i=1}^m a_{1i}(-1)^{1+i}M_{1i}(A).$$

Заметим, что $M_{1i}(A) = \det(\tilde{A}_{1i})$, где матрица $\tilde{A}_{1i}$ является квадратной ступенчатой матрицей порядка $n - 1$, причем роль левой верхней матрицы в $\tilde{A}_{1i}$ играет матрица $\tilde{A}_{1i}^{(1)}$, а нижняя правая матрица есть $A^{(2)}$. По индукционному предположению

$$M_{1i}(A) = \det(\tilde{A}_{1i}^{(1)}) \det(A^{(2)}) = M_{1i}(A^{(1)}) \det(A^{(2)}).$$

Следовательно,

$$\begin{aligned} \det(A) &= \sum_{i=1}^m a_{1i}(-1)^{1+i}M_{1i}(A^{(1)}) \det(A^{(2)}) = \\ &= \det(A^{(2)}) \sum_{i=1}^m a_{1i}(-1)^{1+i}M_{1i}(A^{(1)}) = \det(A^{(2)}) \det(A^{(1)}), \end{aligned}$$

что и требовалось доказать. ■

Пример 2.2. $\det \left[\begin{array}{cc|ccc|cc} 1 & 2 & 0 & 0 & 0 & 0 & 0 \\ 3 & 4 & 0 & 0 & 0 & 0 & 0 \\ \hline 5 & 6 & 7 & 8 & 9 & 0 & 0 \\ 10 & 11 & 12 & 13 & 14 & 0 & 0 \\ \hline 1 & 2 & 3 & 4 & 5 & 0 & 0 \\ 8 & 9 & 10 & 11 & 12 & 13 & 14 \\ \hline 5 & 6 & 7 & 8 & 9 & 10 & 11 \end{array} \right] =$

$$\begin{aligned}
&= \det \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \det \left[\begin{array}{ccc|cc} 7 & 8 & 9 & 0 & 0 \\ 12 & 13 & 14 & 0 & 0 \\ 3 & 4 & 5 & 0 & 0 \\ \hline 10 & 11 & 12 & 13 & 14 \\ 7 & 8 & 9 & 10 & 11 \end{array} \right] = \\
&= \det \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \det \begin{bmatrix} 7 & 8 & 9 \\ 12 & 13 & 14 \\ 3 & 4 & 5 \end{bmatrix} \det \begin{bmatrix} 13 & 14 \\ 10 & 11 \end{bmatrix} = 0,
\end{aligned}$$

так как

$$\det \begin{bmatrix} 7 & 8 & 9 \\ 12 & 13 & 14 \\ 3 & 4 & 5 \end{bmatrix} = 7 \cdot 13 \cdot 5 - 7 \cdot 4 \cdot 14 - 8 \cdot 12 \cdot 5 + 8 \cdot 14 \cdot 3 + 9 \cdot 12 \cdot 4 - 9 \cdot 3 \cdot 13 = 0$$

(см. полученную ранее формулу для вычисления определителя матрицы третьего порядка). •

Утверждение 2.10. *Определитель произведения двух квадратных матриц (естественно, одинакового порядка) равен произведению определителей матриц-сомножителей:*

$$\det(AB) = \det(A) \det(B). \quad (2.9)$$

Доказательство. Пусть

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nn} \end{bmatrix}.$$

Из утверждения 2.9 следует:

$$\det(S) = \det \left[\begin{array}{cccc|ccc} a_{11} & a_{12} & \cdots & a_{1n} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 & \cdots & 0 \\ \hline -1 & 0 & \cdots & 0 & b_{11} & \cdots & b_{1n} \\ 0 & -1 & \cdots & 0 & b_{21} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -1 & b_{n1} & \cdots & b_{nn} \end{array} \right] = \det(A) \det(B).$$

Определитель матрицы S не изменится (утверждение 2.7), если к $(n+1)$ -му столбцу прибавим первый, умноженный на b_{11} , второй, умноженный на b_{21} , и т. д. до n -го, умноженного на b_{n1} ; а к $(n+2)$ -му столбцу прибавим первый, умноженный на b_{12} , второй, умноженный на b_{22} , и т. д. до n -го, умноженного на b_{n2} , и т. д.; наконец, к $(2n)$ -му столбцу прибавим первый, умноженный на b_{1n} , второй, умноженный на b_{2n} , и т. д. до n -го,

умноженного на b_{nn} . Если полученную при этом матрицу обозначить через T , то $\det(S) = \det(T)$ и при этом матрица T будет иметь вид

$$T = \left[\begin{array}{cccc|ccc} a_{11} & a_{12} & \cdots & a_{1n} & c_{11} & \cdots & c_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} & c_{21} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & c_{n1} & \cdots & c_{nn} \\ \hline -1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & -1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -1 & 0 & \cdots & 0 \end{array} \right],$$

где матрица C имеет вид

$$C = \left[\begin{array}{ccc|ccc} a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1n}b_{n1} & \cdots & a_{11}b_{1n} + a_{12}b_{2n} + \cdots + a_{1n}b_{nn} \\ a_{21}b_{11} + a_{22}b_{21} + \cdots + a_{2n}b_{n1} & \cdots & a_{21}b_{1n} + a_{22}b_{2n} + \cdots + a_{2n}b_{nn} \\ \vdots & \ddots & \vdots \\ a_{n1}b_{11} + a_{n2}b_{21} + \cdots + a_{n1}b_{n1} & \cdots & a_{n1}b_{1n} + a_{n2}b_{2n} + \cdots + a_{nn}b_{nn} \end{array} \right]$$

и при этом очевидно, что $C = AB$. Поменяем местами первую строку с $(n+1)$ -й, вторую – с $(n+2)$ -й, ... и n -ю – с $(2n)$ -й, т. е. сделаем n перестановок. Тогда, благодаря утверждению 2.4, получим

$$\begin{aligned} \det(T) &= (-1)^n \det \left[\begin{array}{cccc|ccc} -1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -1 & 0 & \cdots & 0 \\ \hline a_{11} & a_{12} & \cdots & a_{1n} & c_{11} & \cdots & c_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} & c_{21} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & c_{n1} & \cdots & c_{nn} \end{array} \right] = \\ &= (-1)^n \det \left[\begin{array}{cccc} -1 & 0 & \cdots & 0 \\ 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -1 \end{array} \right] \det(C) = (-1)^n (-1)^n \det(C) = \\ &= (-1)^{2n} \det(C) = \det(C) = \det(AB) \end{aligned}$$

(здесь вновь использовалось утверждение 2.9). ■

Другое доказательство (2.9) можно найти в учебнике В. И. Смирнова [4, с. 29].

2.5. Вычисление определителя матрицы методом Гаусса.

Оценка числа операций

Рассмотрим способ вычисления определителя, состоящий в приведении матрицы к верхней треугольной (метод Гаусса).

Шаг 1. Пусть дана матрица

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

Считаем, что $a_{11} \neq 0$. Если $a_{11} = 0$, то, переставляя строки, помещаем в левый верхний угол ненулевой элемент первого столбца (отметим, что при этом определитель умножается на (-1) , согласно утверждению 2.4); если все элементы первого столбца равны нулю, то $\det(A) = 0$. (Заметим, что в вычислительных алгоритмах на место элемента a_{11} помещаем наибольший по модулю элемент первого столбца – его называют ведущим элементом первого столбца.) Делим первую строку на a_{11} , т.е. выносим a_{11} за знак определителя (см. утверждение 2.5). Для этого потребуется $n-1$ операция:

$$\det(A) = a_{11} \det \begin{bmatrix} 1 & a_{12}^{(1)} & a_{13}^{(1)} & \dots & a_{1n}^{(1)} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{bmatrix}.$$

Прибавляем ко второй строке первую, умноженную на $-a_{21}$, – требуется $2(n-1)$ операция; к третьей строке – первую, умноженную на $-a_{31}$ – требуется $2(n-1)$ операция и т.д.; к n -й строке – первую, умноженную на $-a_{n1}$, – требуется $2(n-1)$ операция. Получим:

$$\det(A) = a_{11} \det \begin{bmatrix} 1 & a_{12}^{(1)} & a_{13}^{(1)} & \dots & a_{1n}^{(1)} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} & \dots & a_{2n}^{(1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}^{(1)} & a_{n3}^{(1)} & \dots & a_{nn}^{(1)} \end{bmatrix} = a_{11} \det \begin{bmatrix} a_{22}^{(1)} & a_{23}^{(1)} & \dots & a_{2n}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n2}^{(1)} & a_{n3}^{(1)} & \dots & a_{nn}^{(1)} \end{bmatrix}.$$

Всего на первом шаге выполнено

$$(n-1) + 2(n-1)(n-1) = (n-1) + 2(n-1)^2$$

арифметических операций.

Если все элементы первого столбца полученной матрицы равны нулю, то ее определитель, а следовательно, и $\det(A)$ равны нулю. Если это не так, то считаем, что $a_{22}^{(1)} \neq 0$ (см. 1-й шаг).

Шаг 2. Поступаем аналогично с полученной на первом шаге матрицей $(n - 1)$ -го порядка – вычисление ее определителя сведем к вычислению определителя порядка $n - 2$:

$$\det(A) = a_{11}a_{22}^{(1)} \det \begin{bmatrix} a_{33}^{(2)} & \dots & a_{3n}^{(2)} \\ \vdots & \ddots & \vdots \\ a_{n3}^{(2)} & \dots & a_{nn}^{(2)} \end{bmatrix}.$$

На это потребуется $(n - 2) + 2(n - 2)^2$ арифметических операций и т. д. Если не установлено равенство нулю $\det(A)$, то на последнем, $(n - 1)$ -м, шаге (переход от матрицы второго порядка к матрице первого порядка) потребуется $1 + 2 \cdot 1^2$ арифметических операций. После этого нужно еще перемножить все вынесенные за знак определителя множители и определитель полученной матрицы первого порядка – результат $n - 1$ действий умножения. Таким образом, общее число арифметических операций составляет:

$$\begin{aligned} M_n &= (n-1) + 2(n-1)^2 + (n-2) + 2(n-2)^2 + \dots + 1 + 2 \cdot 1^2 + (n-1) = \\ &= (n-1) + (n-2) + \dots + 1 + 2[(n-1)^2 + (n-2)^2 + \dots + 1^2] + (n-1) = \\ &= \frac{n(n-1)}{2} + \frac{(n-1)n(2n-1)}{3} + (n-1) = \frac{(n-1)(4n^2 + n + 6)}{6}. \end{aligned}$$

Отсюда видно, что при большом n потребуется примерно $\frac{2}{3}n^3$ операций. Для $n = 10$ вместо $9 \cdot 10^6$ (вспомните полученную ранее оценку числа операций (2.2)) при непосредственном вычислении определителя, исходя из его определения, придется проделать всего 624 операции.

Пример 2.3. Вычислить определитель матрицы

$$A = \begin{bmatrix} 1 & -2 & 3 & 5 \\ 2 & 3 & 4 & -1 \\ 5 & -2 & 1 & 2 \\ -3 & 4 & 2 & 3 \end{bmatrix}.$$

Для вычисления этого определителя нужно выполнить 3 шага метода Гаусса:

$$\det(A) = \det \begin{bmatrix} 1 & -2 & 3 & 5 \\ 2 & 3 & 4 & -1 \\ 5 & -2 & 1 & 2 \\ -3 & 4 & 2 & 3 \end{bmatrix} = \det \begin{bmatrix} 1 & -2 & 3 & 5 \\ 0 & 7 & -2 & -11 \\ 0 & 8 & -14 & -23 \\ 0 & -2 & 11 & 18 \end{bmatrix} =$$

$$\begin{aligned}
&= \det \begin{bmatrix} 7 & -2 & -11 \\ 8 & -14 & -23 \\ -2 & 11 & 18 \end{bmatrix} = 7 \det \begin{bmatrix} 1 & -\frac{2}{7} & -\frac{11}{7} \\ 8 & -14 & -23 \\ -2 & 11 & 18 \end{bmatrix} = \\
&= 7 \det \begin{bmatrix} 1 & -\frac{2}{7} & -\frac{11}{7} \\ 0 & -\frac{82}{7} & -\frac{73}{7} \\ 0 & \frac{73}{7} & \frac{104}{7} \end{bmatrix} = 7 \det \begin{bmatrix} -\frac{82}{7} & -\frac{73}{7} \\ \frac{73}{7} & \frac{104}{7} \end{bmatrix} = \\
&= 7 \left(-\frac{82}{7} \right) \det \begin{bmatrix} 1 & \frac{73}{82} \\ \frac{73}{7} & \frac{104}{7} \end{bmatrix} = -82 \det \begin{bmatrix} 1 & \frac{73}{82} \\ 0 & \frac{457}{82} \end{bmatrix} = -82 \left(\frac{457}{82} \right) = -457. \bullet
\end{aligned}$$

Рассмотрим матрицу порядка n

$$W = \begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{bmatrix}, \quad (2.10)$$

где $x_1, x_2, \dots, x_n$ — произвольные числа. Такая матрица возникает при решении прикладных задач и называется матрицей Вандермонда.

Пример 2.4. Вычислить определитель матрицы Вандермонда

$$W = \begin{bmatrix} 1 & x_1 & x_1^2 & x_1^3 \\ 1 & x_2 & x_2^2 & x_2^3 \\ 1 & x_3 & x_3^2 & x_3^3 \\ 1 & x_4 & x_4^2 & x_4^3 \end{bmatrix}.$$

Для вычисления определителя необходимо выполнить 3 шага метода Гаусса:

$$\begin{aligned}
\det(W) &= \det \begin{bmatrix} 1 & x_1 & x_1^2 & x_1^3 \\ 1 & x_2 & x_2^2 & x_2^3 \\ 1 & x_3 & x_3^2 & x_3^3 \\ 1 & x_4 & x_4^2 & x_4^3 \end{bmatrix} = \det \begin{bmatrix} 1 & x_1 & x_1^2 & x_1^3 \\ 0 & x_2 - x_1 & x_2^2 - x_1^2 & x_2^3 - x_1^3 \\ 0 & x_3 - x_1 & x_3^2 - x_1^2 & x_3^3 - x_1^3 \\ 0 & x_4 - x_1 & x_4^2 - x_1^2 & x_4^3 - x_1^3 \end{bmatrix} = \\
&= \det \begin{bmatrix} x_2 - x_1 & (x_2 - x_1)(x_2 + x_1) & (x_2 - x_1)(x_2^2 + x_2x_1 + x_1^2) \\ x_3 - x_1 & (x_3 - x_1)(x_3 + x_1) & (x_3 - x_1)(x_3^2 + x_3x_1 + x_1^2) \\ x_4 - x_1 & (x_4 - x_1)(x_4 + x_1) & (x_4 - x_1)(x_4^2 + x_4x_1 + x_1^2) \end{bmatrix} =
\end{aligned}$$

$$\begin{aligned}
&= (x_2 - x_1)(x_3 - x_1)(x_4 - x_1) \det \begin{bmatrix} 1 & x_2 + x_1 & x_2^2 + x_2x_1 + x_1^2 \\ 1 & x_3 + x_1 & x_3^2 + x_3x_1 + x_1^2 \\ 1 & x_4 + x_1 & x_4^2 + x_4x_1 + x_1^2 \end{bmatrix} = \\
&= (x_2 - x_1)(x_3 - x_1)(x_4 - x_1) \det \begin{bmatrix} 1 & x_2 + x_1 & x_2^2 + x_2x_1 + x_1^2 \\ 0 & x_3 - x_2 & x_3^2 - x_2^2 + x_3x_1 - x_2x_1 \\ 0 & x_4 + x_1 & x_4^2 - x_2^2 + x_4x_1 - x_2x_1 \end{bmatrix} = \\
&= (x_2 - x_1)(x_3 - x_1)(x_4 - x_1) \det \begin{bmatrix} x_3 - x_2 & (x_3 - x_2)(x_3 + x_2 + x_1) \\ x_4 + x_1 & (x_4 + x_1)(x_4 + x_2 + x_1) \end{bmatrix} = \\
&= (x_2 - x_1)(x_3 - x_1)(x_4 - x_1)(x_3 - x_2)(x_4 - x_1) \det \begin{bmatrix} 1 & x_3 + x_2 + x_1 \\ 1 & x_4 + x_2 + x_1 \end{bmatrix} = \\
&= (x_2 - x_1)(x_3 - x_1)(x_4 - x_1)(x_3 - x_2)(x_4 - x_1) \det \begin{bmatrix} 1 & x_3 + x_2 + x_1 \\ 0 & x_4 - x_3 \end{bmatrix} = \\
&= (x_2 - x_1)(x_3 - x_1)(x_4 - x_1)(x_3 - x_2)(x_4 - x_1)(x_4 - x_3). \bullet
\end{aligned}$$

Для определителя матрицы Вандермонда (2.10) порядка n можно доказать следующую общую формулу:

$$\det(W) = \prod_{i=2}^n \left(\prod_{k=1, k < i} (x_i - x_k) \right). \quad (2.11)$$

Из выражения (2.11), очевидно, вытекает следующее.

Утверждение 2.11. *Чтобы определитель матрицы Вандермонда был отличен от нуля, необходимо и достаточно, чтобы $x_i \neq x_k$ при $i \neq k$ ($i, k = 1, \dots, n$).*

2.6. Теорема Крамера

Рассмотрим систему линейных алгебраических уравнений (1.1) для случая, когда число уравнений равно числу неизвестных:

[illegible]

В этом случае матрица системы A будет квадратной.

Операции, входящие в метод полного исключения, изменяют матрицу системы и, следовательно, изменяют ее определитель. Однако справедлива следующая лемма.

Лемма 2.1. Любое допустимое преобразование метода Гаусса–Жордана изменяет значение определителя матрицы системы на множитель, отличный от нуля.

Доказательство. При выполнении преобразований первого и второго типов (перестановка строк и перестановка столбцов), согласно утверждению 2.4, определитель полученной матрицы равен определителю исходной, умноженному на (-1) . Преобразование третьего типа (умножение строки на число, не равное нулю) приводит к матрице, определитель которой отличается от определителя исходной матрицы (см. утверждение 2.5) на множитель, равный этому числу. Наконец, преобразование четвертого типа, согласно утверждению 2.7, не меняет значения определителя. ■

Определение 2.3. Если $\det(A) = 0$, то матрица A называется вырожденной, в противном случае — невырожденной.

Лемма 2.1 означает, что после любых допустимых преобразований метода Гаусса–Жордана невырожденная матрица системы переходит в невырожденную, а вырожденная матрица системы — в вырожденную.

Теорема 2.2 (Крамера). Чтобы система (2.12) имела единственное решение, необходимо и достаточно, чтобы матрица этой системы была невырожденна:

$$\det(A) = \det \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \neq 0.$$

Доказательство. $\Rightarrow$ Если система (2.12) имеет единственное решение, то, решая ее методом полного исключения, приходим к равносильной системе с расширенной матрицей (1.19). Определитель матрицы этой системы равен единице. Следовательно, согласно лемме 2.1, определитель матрицы исходной системы (2.12) отличен от нуля.

$\Leftarrow$ Если определитель матрицы системы (2.12) отличен от нуля, то в процессе выполнения алгоритма полного исключения в матрице системы нулевые строки не образуются, так как определитель матрицы с нулевой строкой равен нулю (см. следствие 2.4), что противоречило бы лемме 2.1. Значит, алгоритм закончится получением системы с единичной матрицей (системы, которой соответствует расширенная матрица (1.19)), имеющей единственное решение. ■

Получим теперь формулы Крамера, выражающие в явном виде решение системы (2.12). Пусть $(x_1^0, x_2^0, \dots, x_n^0)$ — числовые значения неизвестных, представляющие единственное решение системы (2.12), т. е. имеется

n верных числовых равенств:

[illegible]

Умножим первое равенство из (2.13) на число A_{1k} (алгебраическое дополнение элемента a_{1k}), второе — на A_{2k} и т. д., n -е — на A_{nk} . В результате получим n числовых равенств:

$$\begin{aligned} & a_{11}A_{1k}x_1^0 + \cdots + a_{1k}A_{1k}x_k^0 + \cdots + a_{1n}A_{1k}x_n^0 = b_1A_{1k}, \\ & \dots\dots\dots \\ & a_{k1}A_{kk}x_1^0 + \cdots + a_{kk}A_{kk}x_k^0 + \cdots + a_{kn}A_{kk}x_n^0 = b_kA_{kk}, \\ & \dots\dots\dots \\ & a_{n1}A_{nk}x_1^0 + \cdots + a_{nk}A_{nk}x_k^0 + \cdots + a_{nn}A_{nk}x_n^0 = b_nA_{nk}. \end{aligned}$$

Сложив полученные равенства и перегруппировав слагаемые найдем:

$$\begin{aligned} & (a_{11}A_{1k} + \cdots + a_{k1}A_{kk} + \cdots + a_{n1}A_{nk})x_1^0 + \cdots \\ & + \boxed{(a_{1k}A_{1k} + \cdots + a_{kk}A_{kk} + \cdots + a_{nk}A_{nk})} + \cdots \\ & + (a_{1n}A_{1k} + \cdots + a_{kn}A_{kk} + \cdots + a_{nn}A_{nk})x_n^0 = \\ & = b_1A_{1k} + \cdots + b_kA_{kk} + \cdots + b_nA_{nk}. \end{aligned}$$

Все выражения, стоящие в скобках, кроме выделенного, равны нулю (2.7). Выделенное выражение, согласно свойству определителя (2.5), равно $\det(A)$. Таким образом, получили:

$$\det(A)x_k^0 = b_1A_{1k} + \cdots + b_kA_{kk} + \cdots + b_nA_{nk}. \quad (2.14)$$

Заменим теперь в матрице A k -й столбец столбцом свободных членов системы (2.12). Полученную матрицу обозначим

$$B_k = \begin{bmatrix} a_{11} & \cdots & a_{1,k-1} & b_1 & a_{1,k+1} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{k,k-1} & b_k & a_{k,k+1} & \cdots & a_{kn} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{n,k-1} & b_n & a_{n,k+1} & \cdots & a_{nn} \end{bmatrix}.$$

Заметим, что алгебраические дополнения элементов b_i k -го столбца матрицы B_k такие же, как у элементов a_{ik} k -го столбца матрицы A , и поэтому они равны A_{ik} . По свойству определителя (2.5)

$$\det(B_k) = b_1 A_{1k} + \cdots + b_k A_{kk} + \cdots + b_n A_{nk}. \quad (2.15)$$

Из равенств (2.14) и (2.15) следует, что

$$x_k^0 \det(A) = \det(B_k), \quad k = 1, 2, \dots, n.$$

Суммируя изложенное, получаем следующее утверждение.

Утверждение 2.12 (формулы Крамера). Если $\det(A) \neq 0$, то решение системы (2.12) вычисляется по формулам:

$$x_1 = \frac{\det(B_1)}{\det(A)}, \quad x_2 = \frac{\det(B_2)}{\det(A)}, \quad \dots, \quad x_n = \frac{\det(B_n)}{\det(A)}, \quad (2.16)$$

где матрица B_k получается из матрицы системы A заменой k -го столбца столбцом свободных членов системы (2.12).

Пример 2.5. Решить систему

$$\begin{cases} 1x_1 + 2x_2 + 3x_3 = 14, \\ 4x_1 + 5x_2 + 6x_3 = 32, \\ 7x_1 + 9x_2 + 8x_3 = 49. \end{cases}$$

Так как $\det(A) = \det \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 9 & 8 \end{bmatrix} = 9$, то, согласно теореме Крамера, эта система имеет единственное решение и его можно найти по формулам Крамера. Вычислим определители:

$$\det(B_1) = \det \begin{bmatrix} 14 & 2 & 3 \\ 32 & 5 & 6 \\ 49 & 9 & 8 \end{bmatrix} = 9, \quad \det(B_2) = \det \begin{bmatrix} 1 & 14 & 3 \\ 4 & 32 & 6 \\ 7 & 49 & 8 \end{bmatrix} = 18,$$

$$\det(B_3) = \det \begin{bmatrix} 1 & 2 & 14 \\ 4 & 5 & 32 \\ 7 & 9 & 49 \end{bmatrix} = 27.$$

Тогда из формул (2.16) получим: $x_1 = \frac{9}{9} = 1$, $x_2 = \frac{18}{9} = 2$, $x_3 = \frac{27}{9} = 3$. •

Замечание 2.2. Как уже отмечалось, формулы Крамера имеют определенную теоретическую ценность. Их практическое использование (для поиска решения) невыгодно, так как при больших порядках системы ($n > 3$) объем вычислений чрезвычайно велик и метод полного исключения оказывается предпочтительным. ⊗

2.7. Однородные системы линейных уравнений

Сначала дадим определение:

Определение 2.4. Система линейных алгебраических уравнений с нулевым столбцом свободных членов называется однородной:

[illegible]

Всякая однородная система имеет нулевое решение $x_1 = x_2 = \dots = x_n = 0$ (тривиальное решение), т.е. она не может быть несовместна. В связи с этим для однородной системы возможны лишь 2 случая: единственное решение, т.е. равносильность ее системе, которой соответствует таблица (формула (1.19))

$$\left(\begin{array}{cccc|c} j_1 & j_2 & \cdots & j_n & \\ \hline 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{array} \right), \quad (2.18)$$

или бесконечное множество решений, т. е. равносильность ее системе, которой соответствует таблица (формула (1.20))

$$\left(\begin{array}{ccccccc|c} j_1 & j_2 & \cdots & j_s & j_{s+1} & \cdots & j_n & \\ \hline 1 & 0 & \cdots & 0 & a_{1j_{s+1}}^s & \cdots & a_{1j_n}^s & 0 \\ 0 & 1 & \cdots & 0 & a_{2j_{s+1}}^s & \cdots & a_{2j_n}^s & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & a_{sj_{s+1}}^s & \cdots & a_{sj_n}^s & 0 \end{array} \right). \quad (2.19)$$

Замечание 2.3. То, что однородная система (2.17) не может быть несовместной, следует также из того, что в процессе ее решения методом Гаусса—Жордана строки вида $[0 \ \cdots \ 0 \mid b]$, $b \neq 0$, не могут образоваться, так как при любых преобразованиях в правых частях уравнений всегда будут нули. $\otimes$

Теорема 2.3. Если в однородной системе число неизвестных n больше числа уравнений m , то система имеет ненулевые решения (имеет бесконечно много решений).

Доказательство. Применим метод полного исключения. В процессе решения количество неизвестных не меняется, а количество уравнений может лишь уменьшаться, т. е. всегда меньше числа неизвестных. Значит, получить систему вида (2.18) при выполнении алгоритма невозможно, поэтому получим систему вида (2.19), которая имеет бесконечное множество решений — части неизвестных придаются произвольные значения, а остальные неизвестные вычисляются. ■

Теорема 2.4. *Однородная система n уравнений с n неизвестными имеет единственное (тривиальное) решение тогда и только тогда, когда определитель ее матрицы не равен нулю.*

Доказательство. 1. Если $\det(A) \neq 0$, то, согласно теореме 2.2 (Крамера), система имеет единственное решение. Но однородная система всегда имеет нулевое (тривиальное) решение. Следовательно, в этом случае именно это нулевое решение и является единственным решением однородной системы.

2. Пусть однородная система n уравнений с n неизвестными имеет единственное решение. Предположим, что $\det(A) = 0$. Но тогда на каждом шаге метода полного исключения будем получать систему с вырожденной матрицей. Это означает, что алгоритм не может закончиться получением системы вида (2.18), а обязательно завершится получением системы (2.19), которая, как известно, имеет бесконечно много решений. Это противоречит исходным условиям, следовательно, исходное предположение неверно, т. е. $\det(A) \neq 0$. ■

Следствие 2.5. *Для однородной системы n уравнений с n неизвестными равносильны следующие условия:*

- 1) $\det(A) = 0$,
- 2) *система имеет бесконечно много решений.*

2.8. Матричные уравнения

Ранее (см. 1.8) было установлено, что систему линейных алгебраических уравнений (1.1) можно записать в виде матричного уравнения (1.43) $AX = B$, где введены обозначения: A — матрица системы (матрица коэффициентов при неизвестных), B — матрица-столбец свободных членов, X — матрица-столбец неизвестных. Пусть теперь A и B — заданные числовые матрицы размером $m \times n$ и $m \times k$ соответственно. Запишем уравнение

$$AX = B. \quad (2.20)$$

Здесь искомый объект — матрица X размером $n \times k$.

Пример 2.6. Если

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 5 & 7 & 9 \\ 6 & 8 & 10 \end{bmatrix}, \quad X = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \end{bmatrix},$$

то матричное уравнение (2.20) представляет собой систему из шести (число элементов матрицы B) линейных уравнений с шестью (число элементов матрицы X) неизвестными. Если выполнить умножение в левой части (2.20), то получим матричное равенство

$$\begin{bmatrix} 1x_{11} + 2x_{21} & 1x_{12} + 2x_{22} & 1x_{13} + 2x_{23} \\ 3x_{11} + 4x_{21} & 3x_{12} + 4x_{22} & 3x_{13} + 4x_{23} \end{bmatrix} = \begin{bmatrix} 5 & 7 & 9 \\ 6 & 8 & 10 \end{bmatrix}.$$

Запишем теперь равенство первых столбцов матриц AX и B :

$$\begin{cases} 1x_{11} + 2x_{21} = 5, \\ 3x_{11} + 4x_{21} = 6. \end{cases} \quad (2.21)$$

Получилась система линейных уравнений с матрицей A . Эту систему можно записать в матричной форме:

$$AX_{(1)} = B_{(1)},$$

где $X_{(1)} = \begin{bmatrix} x_{11} \\ x_{21} \end{bmatrix}$ и $B_{(1)} = \begin{bmatrix} 5 \\ 6 \end{bmatrix}$ — первые столбцы матриц X и B соответственно.

Так же запишем равенство вторых, а затем третьих столбцов матриц AX и B :

$$\begin{cases} 1x_{12} + 2x_{22} = 7, \\ 3x_{12} + 4x_{22} = 8, \end{cases} \quad \begin{cases} 1x_{13} + 2x_{23} = 9, \\ 3x_{13} + 4x_{23} = 10 \end{cases} \quad (2.22)$$

или в матричной форме:

$$AX_{(2)} = B_{(2)}, \quad AX_{(3)} = B_{(3)},$$

где $X_{(2)}$, $B_{(2)}$ — вторые, а $X_{(3)}$, $B_{(3)}$ — третьи столбцы матриц X и B соответственно.

Видно, что одно матричное уравнение распалось на 3 матричных уравнения (по числу столбцов матрицы B), представляющих собой 3 *независимые* системы линейных уравнений с *общей матрицей* A . •

Вернемся теперь к рассмотрению общего случая уравнения (2.20). Обозначая

$$X = [X_{(1)} \ X_{(2)} \ \cdots \ X_{(k)}], \quad B = [B_{(1)} \ B_{(2)} \ \cdots \ B_{(k)}],$$

где, как и прежде, $X_{(i)}$, $B_{(i)}$ — i -е столбцы матриц X и B соответственно, легко видеть, что равенство столбцов матричного уравнения

$$A [X_{(1)} \ X_{(2)} \ \cdots \ X_{(k)}] = [B_{(1)} \ B_{(2)} \ \cdots \ B_{(k)}]$$

приводит к матричным уравнениям

$$AX_{(1)} = B_{(1)}, AX_{(2)} = B_{(2)}, \dots, AX_{(k)} = B_{(k)},$$

каждое из которых представляет собой систему из t уравнений с n неизвестными. Все эти системы имеют *общую матрицу* A . Решая эти системы поочередно (например, методом Гаусса–Жордана), можно последовательно найти все столбцы $X_{(i)}$ искомой матрицы X .

Отметим, что при решении системы методом полного исключения ход алгоритма определяется матрицей системы A . Например, для всех трех систем примера 2.6 (система (2.21) и две системы из (2.22)) все шаги метода Гаусса–Жордана будут одинаковыми – различия только в столбцах свободных членов. В связи с этим системы уравнений, получающиеся при “расщеплении” матричного уравнения, целесообразно решать все сразу, записывая расширенную матрицу, соответствующую матричному уравнению (2.20), в виде

$$\left(\begin{array}{ccc|ccc} a_{11} & \cdots & a_{1n} & b_{11} & \cdots & b_{1k} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_{m1} & \cdots & b_{mk} \end{array} \right) \quad (2.23)$$

и выполняя, как обычно, все шаги алгоритма полного исключения.

Проиллюстрируем это примером 2.7.

Пример 2.7. Рассмотрим матричное уравнение из примера 2.6, запишем его расширенную матрицу вида (2.23) и применим алгоритм Гаусса–Жордана:

$$\begin{aligned} & \left(\begin{array}{cc|ccc} 1 & 2 & 5 & 7 & 9 \\ 3 & 4 & 6 & 8 & 10 \end{array} \right) \Leftrightarrow \left(\begin{array}{cc|ccc} 1 & 2 & 5 & 7 & 9 \\ 0 & -2 & -9 & -13 & -17 \end{array} \right) \Leftrightarrow \\ & \Leftrightarrow \left(\begin{array}{cc|ccc} 1 & 2 & 5 & 7 & 9 \\ 0 & 1 & \frac{9}{2} & \frac{13}{2} & \frac{17}{2} \end{array} \right) \Leftrightarrow \left(\begin{array}{cc|ccc} 1 & 0 & -4 & -6 & -8 \\ 0 & 1 & \frac{9}{2} & \frac{13}{2} & \frac{17}{2} \end{array} \right). \end{aligned}$$

Справа от черты получили искомую матрицу $X = \begin{bmatrix} -4 & -6 & -8 \\ \frac{9}{2} & \frac{13}{2} & \frac{17}{2} \end{bmatrix}$. •

Рассмотрим теперь матричное уравнение вида

$$XA = B, \quad (2.24)$$

где матрицы $A \in \mathcal{M}_{m \times n}$ и $B \in \mathcal{M}_{k \times n}$, следовательно, $X \in \mathcal{M}_{k \times m}$.

Транспонируя обе части уравнения (2.24), получаем:

$$A'X' = B'. \quad (2.25)$$

Матричное уравнение (2.25) есть матричное уравнение рассмотренного ранее вида. Решая его, находим матрицу X' и, транспонируя результат, получаем искомую матрицу X .

Пример 2.8. Решить матричное уравнение

$$X \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 5 & 7 \\ 6 & 8 \end{bmatrix}.$$

Транспонируем обе части уравнения:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} X' = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}.$$

Решаем полученное матричное уравнение (без перестановки строк и столбцов):

$$\begin{aligned} \left(\begin{array}{cc|cc} 1 & 2 & 5 & 6 \\ 3 & 4 & 7 & 8 \end{array} \right) &\Leftrightarrow \left(\begin{array}{cc|cc} 1 & 2 & 5 & 6 \\ 0 & -2 & -8 & -10 \end{array} \right) \Leftrightarrow \\ &\Leftrightarrow \left(\begin{array}{cc|cc} 1 & 2 & 5 & 6 \\ 0 & 1 & 4 & 5 \end{array} \right) \Leftrightarrow \left(\begin{array}{cc|cc} 1 & 0 & -3 & -4 \\ 0 & 1 & 4 & 5 \end{array} \right). \end{aligned}$$

Получили $X' = \begin{bmatrix} -3 & -4 \\ 4 & 5 \end{bmatrix}$, следовательно, $X = \begin{bmatrix} -3 & 4 \\ -4 & 5 \end{bmatrix}$. •

2.9. Обратная матрица

Определение 2.5. Пусть A – квадратная матрица. Если

$$AX = I, \tag{2.26}$$

где I – единичная матрица того же порядка, что и матрица A , то матрицу X называют обратной матрицей для A . Обратную матрицу обозначают $X = A^{-1}$.

Утверждение 2.13. Вырожденная матрица не имеет обратной матрицы.

Доказательство. Пусть $\det(A) = 0$ и допустим противное: существует обратная матрица A^{-1} . Тогда

$$0 = \det(A) \det(A^{-1}) = \det(AA^{-1}) = \det(I) = 1 \neq 0.$$

Полученное противоречие доказывает утверждение. ■

Теорема 2.5. Произвольная невырожденная квадратная матрица имеет единственную обратную.

Доказательство. Пусть A – квадратная матрица порядка n и $\det(A) \neq 0$. Матричное уравнение (2.26) сводится к n системам линейных уравнений с общей матрицей коэффициентов A . Поскольку матрица A невырождена, то, согласно теореме 2.2 (Крамера), каждая система имеет единственное решение. Следовательно, и матричное уравнение (2.26) имеет единственное решение, т. е. существует единственная обратная матрица A^{-1} . ■

Утверждение 2.14. Если матрица A имеет обратную матрицу A^{-1} (т. е. $\det(A) \neq 0$), то $A^{-1}A = I$.

Доказательство. Рассмотрим матричное уравнение вида (2.24):

$$YA = I, \quad (2.27)$$

равносильное уравнению вида (2.25):

$$A'Y' = I' = I. \quad (2.28)$$

Поскольку $\det(A') = \det(A) \neq 0$, матричное уравнение (2.28) имеет единственное решение (см. доказательство теоремы 2.5), а значит, и уравнение (2.27) имеет единственное решение.

Остается доказать, что $Y = A^{-1}$. Умножим уравнение (2.27) справа на матрицу A^{-1} :

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

В силу ассоциативности умножения матриц и учитывая (2.26), получим:

$$(YA)A^{-1} = Y(AA^{-1}) = YI = Y.$$

Следовательно, $Y = IA^{-1} = A^{-1}$. ■

Пример 2.9. Найти для матрицы $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ обратную матрицу A^{-1} .

Решаем матричное уравнение $AX = I$:

$$\begin{aligned} \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array} \right) &\Leftrightarrow \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -2 & -3 & 1 \end{array} \right) \Leftrightarrow \\ &\Leftrightarrow \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & \frac{3}{2} & -\frac{1}{2} \end{array} \right) \Leftrightarrow \left(\begin{array}{cc|cc} 1 & 0 & -2 & 1 \\ 0 & 1 & \frac{3}{2} & -\frac{1}{2} \end{array} \right). \end{aligned}$$

Следовательно, $A^{-1} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}$. •

Иногда необходимо иметь явные выражения для элементов обратной матрицы. Получим их. Пусть

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}, \quad A^{-1} = X = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nn} \end{bmatrix}.$$

Система уравнений, из которой находят элементы k -го столбца обратной матрицы, имеет вид

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_{1k} \\ \vdots \\ x_{nk} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} k,$$

где справа – k -й столбец единичной матрицы I .

Согласно формулам Крамера (2.16) $x_{ik} = \frac{\det(B_{ki})}{\det(A)}$, где

$$B_{ki} = \begin{bmatrix} a_{11} & \cdots & a_{1i-1} & 0 & a_{1i+1} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{ki-1} & 1 & a_{ki+1} & \cdots & a_{kn} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{ni-1} & 0 & a_{ni+1} & \cdots & a_{nn} \end{bmatrix}.$$

Разлагая $\det(B_{ki})$ по элементам i -го столбца, получим $\det(B_{ki}) = A_{ki}$. Следовательно, элемент обратной матрицы, стоящий в i -й строке и k -м столбце:

$$x_{ik} = \frac{A_{ki}}{\det(A)}. \quad (2.29)$$

Формула (2.29) позволяет сформулировать алгоритм построения матрицы A^{-1} , обратной матрице A :

1) заменяем каждый элемент матрицы A его алгебраическим дополнением;

2) полученную матрицу транспонируем, в результате получаем матрицу A^{Π} , которая называется присоединенной по отношению к матрице A ;

3) каждый элемент присоединенной матрицы делим на $\det(A)$, т.е. умножаем матрицу A^{Π} на $\frac{1}{\det(A)}$.

Свойства обратной матрицы:

1) существует матрица, обратная к A^{-1} , причем $(A^{-1})^{-1} = A$;

2) если A и B – квадратные невырожденные матрицы, то

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

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

Рассмотрим матричное уравнение с квадратной невырожденной матрицей A

$$AX = B.$$

Если известна матрица A^{-1} , то решением этого уравнения будет:

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

3. КОНЕЧНОМЕРНЫЕ ЛИНЕЙНЫЕ ПРОСТРАНСТВА

3.1. Определение линейного пространства

В математике часто рассматриваются множества, для элементов которых определены операции сложения и умножения на число.

Пример 3.1. В школьном курсе геометрии вводятся векторы как направленные отрезки. Для этих векторов определены операции сложения (по правилу треугольника (параллелограмма)) и умножения на вещественные числа. При этом предполагается, что если 2 направленных отрезка можно совместить параллельным переносом, то они определяют один и тот же вектор. Это предположение позволяет все векторы откладывать от какой-либо одной точки, которую принимают обычно за начало некоторой координатной системы. •

Пример 3.2. В 1.5 на множестве $\mathcal{M}_{m \times n}(\mathbb{C})$ (матриц размером $m \times n$) определены операции сложения и умножения на комплексные числа. На множестве вещественных матриц ($\mathcal{M}_{m \times n}(\mathbb{R})$) определены операции сложения и умножения на вещественные числа. Выделим множества $\mathcal{M}_{n \times 1}(\mathbb{C})$ (или $\mathcal{M}_{n \times 1}(\mathbb{R})$) – матриц, состоящих из одного столбца высотой n комплексных (или вещественных) чисел. В дальнейшем будем пользоваться для $\mathcal{M}_{n \times 1}(\mathbb{C})$ и $\mathcal{M}_{n \times 1}(\mathbb{R})$ обозначениями $\mathbb{C}^n$ и $\mathbb{R}^n$ соответственно. •

Пример 3.3. В курсе математического анализа определяются операции сложения функций и умножения их на число. Эти операции определены, например, на множестве $C[a, b]$ – непрерывных на сегменте $[a, b]$ вещественных функций, и их результатом являются (как известно) также непрерывные функции. •

В приведенных примерах операции сложения и умножения на число определены для совершенно разных объектов. Для анализа свойств таких объектов с общей точки зрения введем понятие *линейного пространства*.

Определение 3.1. Множество $\mathbb{L}$ называется линейным пространством, если:

а) каждому двум элементам $x, y \in \mathbb{L}$ поставлен в соответствие элемент $z = x + y \in \mathbb{L}$, называемый суммой элементов x и y ;

б) каждому элементу $x \in \mathbb{L}$ и каждому числу $\lambda \in \mathbb{C}$ ($\lambda \in \mathbb{R}$) поставлен в соответствие элемент $u = \lambda x \in \mathbb{L}$, называемый произведением числа λ на элемент x ;

в) указанные операции удовлетворяют следующим требованиям (аксиомам):

1) $x + y = y + x$ (коммутативность сложения);

2) $(x + y) + z = x + (y + z)$ (ассоциативность сложения);

3) существует элемент $\Theta \in \mathbb{L}$, такой, что для любого $x \in \mathbb{L}$ выполнено равенство $x + \Theta = x$ (элемент Θ называется нулевым элементом);

4) для любого элемента $x \in \mathbb{L}$ существует противоположный элемент, обозначаемый через $(-x) \in \mathbb{L}$, такой, что $x + (-x) = \Theta$;

5) $1 \cdot x = x$;

6) $\alpha(\beta x) = (\alpha\beta)x$;

7) $(\alpha + \beta)x = \alpha x + \beta x$;

8) $\alpha(x + y) = \alpha x + \alpha y$.

Здесь x, y и z — элементы¹ линейного пространства, а α и β — числа.

Отметим, что один и тот же знак $+$ используется, к сожалению, для обозначения и суммы элементов линейного пространства, и суммы чисел.

Теорема 3.1. Для любого линейного пространства справедливы следующие утверждения:

1) нулевой элемент Θ в линейном пространстве единственен;

2) противоположный любому элементу $x \in \mathbb{L}$ элемент $(-x)$ единственен;

3) для любого элемента $x \in \mathbb{L}$ справедливо равенство

$$0 \cdot x = \Theta;$$

4) единственным противоположным элементом любому элементу $x \in \mathbb{L}$ служит

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

Доказательство. Приведем доказательство первого утверждения. Предположим, что в пространстве $\mathbb{L}$ существует 2 нулевых элемента Θ_1 и Θ_2 . Если принять $x = \Theta_1$, то из третьей аксиомы получим:

$$\Theta_1 + \Theta_2 = \Theta_1. \quad (3.1)$$

¹ В абстрактном пространстве $\mathbb{L}$ часто вместо „элемент“ используется термин „вектор“.

Если в той же аксиоме принять $x = \Theta_2$, то

$$\Theta_2 + \Theta_1 = \Theta_2. \quad (3.2)$$

Сравнивая формулы (3.1) и (3.2) и учитывая первую аксиому, находим $\Theta_1 = \Theta_2$, что и требовалось доказать.

Доказать оставшиеся утверждения предлагаем читателю. (Полное доказательство всех утверждений теоремы можно найти в учебном пособии [5, 2.1]). ■

Для проверки того, что некоторое множество с введенными в нем операциями сложения элементов и умножения элементов на число является линейным пространством, необходимо проверять выполнение аксиом 1–8 определения 3.1. Вернемся к приведенным ранее примерам. Для примера 3.2 выполнение аксиом было проверено в 1.5. Предлагаем читателю самостоятельно проверить выполнение аксиом для примеров 3.1 и 3.3. Множества, описанные в примерах 3.1–3.3, являются линейными пространствами.

Рассмотрим еще 2 примера.

Пример 3.4. Совокупность всех многочленов степени, не превышающей натурального числа n , с обычными операциями сложения многочленов и умножения их на число является линейным пространством, так как выполнение аксиом линейного пространства очевидно. При этом нулевым элементом является многочлен, тождественно равный нулю. ●

Пример 3.5. Множество многочленов степени n не является линейным пространством, так как сумма многочленов степени n может иметь меньшую степень, например: $t^n + t + (-t^n + 1) = t + 1$. ●

Замечание 3.1. Будем называть элементы линейного векторного пространства векторами. Не должен смущать тот факт, что этот термин употребляется и в более узком смысле (например, геометрические векторы примера 3.1), так как геометрические представления, связанные с ним, помогают прояснить, а часто и предвидеть ряд результатов.

Если числа, участвующие в определении линейного пространства, вещественны, то пространство называют вещественным линейным пространством, а его элементы – вещественными векторами. Если же эти числа комплексные, то $\mathbb{L}$ называют комплексным пространством. ⊗

3.2. Линейная зависимость векторов

Рассмотрим систему векторов $x_1, x_2, \dots, x_n \in \mathbb{L}$ в линейном пространстве.

Определение 3.2. Вектор

$$y = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n = \sum_{k=1}^n \alpha_k x_k \in \mathbb{L}, \quad n \geq 1,$$

где $\alpha_1, \alpha_2, \dots, \alpha_n$ – числа, называется линейной комбинацией векторов $x_1, x_2, \dots, x_n \in \mathbb{L}$ с коэффициентами $\alpha_1, \alpha_2, \dots, \alpha_n$.

Очевидно, что если $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$, то $y = \Theta$. Важную роль в дальнейшем играет вопрос: может ли линейная комбинация равняться нулевому вектору, если хотя бы один из коэффициентов $\alpha_k \neq 0$?

Определение 3.3. Система векторов $x_1, \dots, x_n$ называется линейно зависимой, если существуют числа $\alpha_1, \dots, \alpha_n$, из которых хотя бы одно отлично от нуля, такие, что

$$\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n = \Theta. \quad (3.3)$$

Система векторов $x_1, x_2, \dots, x_n$ называется линейно независимой, если равенство (3.3) выполнено только при

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

Приведем несколько простых свойств систем векторов, связанных с линейной зависимостью.

Утверждение 3.1. Если среди векторов $x_1, x_2, \dots, x_n$ содержится нулевой вектор Θ , то система векторов $x_1, x_2, \dots, x_n$ линейно зависима.

Доказательство. Предположим, что $x_k = \Theta$, тогда равенство

$$\sum_{k=1}^n \alpha_k x_k = \Theta$$

будет выполнено, если принять, например,

$$\alpha_1 = \alpha_2 = \cdots = \alpha_{k-1} = \alpha_{k+1} = \cdots = \alpha_n = 0 \text{ и } \alpha_k = 1,$$

следовательно, векторы $x_1, x_2, \dots, x_n$ линейно зависимы. ■

Утверждение 3.2. Если система векторов $x_1, \dots, x_n$ линейно зависима, то и система $x_1, \dots, x_n, y$, где y – произвольный вектор, линейно зависима.

Доказательство. По условию система векторов $x_1, x_2, \dots, x_n$ линейно зависима. Это означает, что выполнено равенство $\sum_{k=1}^n \alpha_k x_k = \Theta$, где среди чисел $\alpha_1, \alpha_2, \dots, \alpha_n$ хотя бы одно отлично от нуля. Тогда справедливо равенство

$$\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n + 0y = \Theta,$$

показывающее, что система векторов $x_1, x_2, \dots, x_n, y$ линейно зависима, поскольку среди чисел $\alpha_1, \dots, \alpha_n, 0$ имеются отличные от нуля. ■

Следствие 3.1. Если система векторов $x_1, \dots, x_n$ содержит линейно зависимую подсистему¹ k ($k \leq n$) векторов $x_{l_1}, x_{l_2}, \dots, x_{l_k}$, то исходная система векторов $x_1, \dots, x_n$ линейно зависима.

Доказательство. Если к подсистеме векторов $x_{l_1}, x_{l_2}, \dots, x_{l_k}$ последовательно добавлять оставшиеся в системе $x_1, \dots, x_n$ векторы, то, согласно утверждению 3.2, все получающиеся системы векторов линейно зависимы. ■

Утверждение 3.3. Если $x_1, x_2, \dots, x_n$ – система линейно независимых векторов, то любая подсистема k ($k \leq n$) векторов

$$x_{l_1}, x_{l_2}, \dots, x_{l_k}$$

линейно независима.

Доказательство. Предположим противное – система векторов

$$x_1, x_2, \dots, x_n$$

имеет линейно зависимую подсистему. Но тогда, согласно следствию 3.1, система векторов

$$x_1, x_2, \dots, x_n$$

должна быть линейно зависима, что противоречит условию. ■

Теорема 3.2. Система векторов $x_1, x_2, \dots, x_n$ линейно зависима тогда и только тогда, когда хотя бы один из этих векторов является линейной комбинацией остальных.

Доказательство. $\Rightarrow$ Пусть система векторов $x_1, \dots, x_n$ линейно зависима. Тогда найдутся такие числа $\alpha_1, \dots, \alpha_n$, что $\sum_{k=1}^n \alpha_k x_k = \Theta$, причем хотя бы одно из них не равно нулю, например $\alpha_l \neq 0$. Перенеся в этом равенстве все слагаемые, кроме l -го, в правую часть и умножив полученное равенство на $1/\alpha_l$, получим:

$$x_l = -\frac{\alpha_1}{\alpha_l} x_1 - \dots - \frac{\alpha_{l-1}}{\alpha_l} x_{l-1} - \frac{\alpha_{l+1}}{\alpha_l} x_{l+1} - \dots - \frac{\alpha_n}{\alpha_l} x_n,$$

т. е. вектор x_l представлен в виде линейной комбинации остальных векторов системы.

¹ Это означает, что система $x_1, \dots, x_n$ может быть получена присоединением к системе $x_{l_1}, x_{l_2}, \dots, x_{l_k}$ $n - k$ недостающих векторов.

⇐ Пусть вектор x_m есть линейная комбинация остальных векторов:

$$x_m = \beta_1 x_1 + \cdots + \beta_{m-1} x_{m-1} + \beta_{m+1} x_{m+1} + \cdots + \beta_n x_n.$$

Тогда выполнено равенство

$$\beta_1 x_1 + \cdots + \beta_{m-1} x_{m-1} - 1 x_m + \beta_{m+1} x_{m+1} + \cdots + \beta_n x_n = \Theta.$$

Приняв $\alpha_l = \begin{cases} \beta_l, & l \neq m, \\ -1, & l = m, \end{cases}$ получим $\sum_{k=1}^n \alpha_k x_k = \Theta$, где $\alpha_m = -1 \neq 0$, значит, система векторов $x_1, x_2, \dots, x_n$ линейно зависима. ■

3.3. Размерность линейного пространства

Определение 3.4. *Линейное пространство $\mathbb{L}$ называется N -мерным, если в нем: 1) существует линейно независимая система из N векторов и 2) любая система из $N + 1$ векторов линейно зависима. Число N называют размерностью пространства и записывают этот факт $N = \dim(\mathbb{L})$.*

Данное определение означает: размерность линейного пространства равна максимально возможному числу линейно независимых векторов этого пространства.

Заметим, что если в $\mathbb{L}$ для любого натурального числа N можно найти систему из N линейно независимых векторов, то $\mathbb{L}$ называется *бесконечномерным*. Подробное изучение бесконечномерных пространств выходит за рамки данного издания, поэтому будем рассматривать пространства конечного числа измерений.

Теорема 3.3. Пусть в линейном пространстве задан набор векторов $f_1, f_2, \dots, f_k$. Пусть каждый из векторов $g_1, g_2, \dots, g_l$ есть линейная комбинация векторов $f_1, f_2, \dots, f_k$:

$$\begin{aligned} g_1 &= \alpha_{11}f_1 + \alpha_{12}f_2 + \cdots + \alpha_{1k}f_k, \\ g_2 &= \alpha_{21}f_1 + \alpha_{22}f_2 + \cdots + \alpha_{2k}f_k, \\ &\dots\dots\dots \\ g_l &= \alpha_{l1}f_1 + \alpha_{l2}f_2 + \cdots + \alpha_{lk}f_k. \end{aligned} \tag{3.4}$$

Тогда если система векторов $g_1, g_2, \dots, g_l$ линейно независима, то $l \leq k$. Другими словами, среди линейных комбинаций k векторов

$$f_1, f_2, \dots, f_k$$

не может быть более k линейно независимых векторов.

Доказательство. Предположим, что $l > k$. Составим из векторов $g_1, g_2, \dots, g_l$ линейную комбинацию

$$L = \beta_1 g_1 + \beta_2 g_2 + \cdots + \beta_l g_l,$$

подставим в нее (3.4) и перегруппируем слагаемые:

$$\begin{aligned} L &= \beta_1(\alpha_{11}f_1 + \alpha_{12}f_2 + \cdots + \alpha_{1k}f_k) + \\ &+ \beta_2(\alpha_{21}f_1 + \alpha_{22}f_2 + \cdots + \alpha_{2k}f_k) + \\ &\dots\dots\dots \\ &+ \beta_l(\alpha_{l1}f_1 + \alpha_{l2}f_2 + \cdots + \alpha_{lk}f_k) = \\ &= (\alpha_{11}\beta_1 + \alpha_{21}\beta_2 + \cdots + \alpha_{l1}\beta_l)f_1 + \\ &+ (\alpha_{12}\beta_1 + \alpha_{22}\beta_2 + \cdots + \alpha_{l2}\beta_l)f_2 + \\ &\dots\dots\dots \\ &+ (\alpha_{1l}\beta_1 + \alpha_{2l}\beta_2 + \cdots + \alpha_{kl}\beta_l)f_k. \end{aligned}$$

Если в полученном выражении все коэффициенты при

$$f_1, f_2, \dots, f_k$$

равны нулю, то получим нулевой элемент Θ . Следовательно, если

$$\beta_1, \beta_2, \dots, \beta_l$$

удовлетворяют системе уравнений

[illegible]

TO

$$\beta_1 g_1 + \beta_2 g_2 + \cdots + \beta_l g_l = \Theta. \quad (3.6)$$

Так как в однородной системе (3.5) число неизвестных l больше числа уравнений k , то согласно теореме 2.3 система (3.5) имеет бесконечно много решений. Таким образом, существуют числа $\beta_1, \beta_2, \dots, \beta_l$, среди которых есть ненулевые, такие, что выполняется равенство (3.6). Следовательно, система векторов $q_1, q_2, \dots, q_l$ линейно зависима. ■

Из доказанной теоремы вытекает следующий важный результат.

Следствие 3.2. Если в пространстве $\mathbb{L}$ существует набор k линейно независимых векторов $f_1, \dots, f_k$ и при этом любой вектор из $\mathbb{L}$ представим в виде их линейной комбинации, то пространство $\mathbb{L}$ k -мерно.

Доказательство. Так как любой вектор из $\mathbb{L}$ можно представить в виде линейной комбинации $f_1, \dots, f_k$, то любой набор линейно независимых векторов $g_1, \dots, g_l$ удовлетворяет условиям теоремы 3.3 и, следовательно, $l \leq k$. Значит, в пространстве $\mathbb{L}$ максимальное число линейно независимых векторов равно k , т. е. $\dim(\mathbb{L}) = k$. ■

3.4. Базис и координаты в n -мерном пространстве

Определение 3.5. В n -мерном линейном пространстве $\mathbb{L}$ упорядоченный набор n линейно независимых векторов $e_1, e_2, \dots, e_n$ называется базисом этого пространства.

Определение 3.6. Пусть $e_1, e_2, \dots, e_n$ – базис в линейном пространстве $\mathbb{L}$ и вектор $x \in \mathbb{L}$ представлен в виде

$$x = \xi_1 e_1 + \xi_2 e_2 + \dots + \xi_n e_n. \quad (3.7)$$

Это представление называется разложением вектора x по базису

$$e_1, e_2, \dots, e_n,$$

а числа $\xi_1, \xi_2, \dots, \xi_n$ – координатами вектора x в этом базисе.

Теорема 3.4. Если $e_1, e_2, \dots, e_n$ – базис в n -мерном пространстве $\mathbb{L}$, то любой вектор $x \in \mathbb{L}$ можно единственным образом представить в виде линейной комбинации векторов этого базиса.

Доказательство. Присоединим к векторам базиса $e_1, e_2, \dots, e_n$ вектор $x \in \mathbb{L}$. Получим набор $n + 1$ векторов, которые, по определению 3.4, линейно зависимы, т. е.

$$\alpha_0 x + \alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_n e_n = \Theta, \quad (3.8)$$

где не все α_i равны нулю. При этом $\alpha_0 \neq 0$, так как иначе из (3.8) следует линейная зависимость $e_1, e_2, \dots, e_n$. Значит,

$$x = -\frac{\alpha_1}{\alpha_0} e_1 - \frac{\alpha_2}{\alpha_0} e_2 - \dots - \frac{\alpha_n}{\alpha_0} e_n.$$

Итак, осталось доказать единственность полученного разложения.

Предположим, что существуют 2 разложения вида (3.7):

$$x = \xi_1 e_1 + \xi_2 e_2 + \dots + \xi_n e_n$$

и

$$x = \xi'_1 e_1 + \xi'_2 e_2 + \dots + \xi'_n e_n.$$

Вычитая, получим

$$\Theta = (\xi_1 - \xi'_1)e_1 + (\xi_2 - \xi'_2)e_2 + \cdots + (\xi_n - \xi'_n)e_n.$$

В силу линейной независимости $e_1, e_2, \dots, e_n$ это возможно только в том случае, если

$$\xi_1 - \xi'_1 = \xi_2 - \xi'_2 = \cdots = \xi_n - \xi'_n = 0$$

или

$$\xi_1 = \xi'_1, \xi_2 = \xi'_2, \dots, \xi_n = \xi'_n.$$

Единственность доказана. ■

Выберем в линейном пространстве $\mathbb{L}$ базис $e_1, e_2, \dots, e_n$. Тогда каждому вектору $\mathbb{L}$ соответствует единственный набор n чисел – его координат в этом базисе.

Теорема 3.5. *При сложении двух векторов пространства $\mathbb{L}$ их координаты складываются. При умножении вектора на число все его координаты умножаются на это число.*

Доказательство. Пусть

$$x = \xi_1 e_1 + \xi_2 e_2 + \cdots + \xi_n e_n, \quad y = \eta_1 e_1 + \eta_2 e_2 + \cdots + \eta_n e_n.$$

Тогда

$$x + y = (\xi_1 + \eta_1)e_1 + (\xi_2 + \eta_2)e_2 + \cdots + (\xi_n + \eta_n)e_n,$$

$$\lambda x = \lambda \xi_1 e_1 + \lambda \xi_2 e_2 + \cdots + \lambda \xi_n e_n,$$

что и требовалось доказать. ■

Ясно также, что нулевой вектор Θ , и только он, имеет все координаты, равные нулю.

Укажем размерности и базисы в некоторых ранее рассмотренных пространствах.

Обратимся к примеру 3.2 – множеству $\mathcal{M}_{m \times n}(\mathbb{C})$ матриц размером $m \times n$. Пусть G_{ik} – матрица, в которой на пересечении i -й строки и k -го столбца стоит 1, а остальные элементы равны нулю:

$$G_{ik} = \begin{bmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{bmatrix} \begin{matrix} \\ \\ i. \\ \\ k \end{matrix} \quad (3.9)$$

Рассмотрим набор из mn матриц G_{ik} , $1 \leq i \leq m$, $1 \leq k \leq n$. Очевидно, что любая матрица $A \in \mathcal{M}_{m \times n}(\mathbb{C})$ с элементами a_{ik} представляется в виде

линейной комбинации матриц G_{ik} (3.9):

$$A = \sum_{i=1}^m \sum_{k=1}^n a_{ik} G_{ik}. \quad (3.10)$$

Докажите самостоятельно, что набор G_{ik} , $i = 1, \dots, m$, $k = 1, \dots, n$, линейно независим.

Значит, согласно следствию 3.2, пространство $\mathcal{M}_{m \times n}(\mathbb{C})$ имеет размерность mn , а $\{G_{ik}\}$ образуют базис, который будем называть *стандартным* базисом. Согласно формуле (3.10) числа a_{ik} являются координатами в стандартном базисе.

Рассмотрим пример 3.4. Векторами в этом пространстве являются многочлены степени не более n . Простейшим базисом является, например, совокупность линейно независимых векторов $e_1 = 1$, $e_2 = t$, ..., $e_{n+1} = t^n$. Действительно, если многочлен степени n равен нулю для любых t , то все его коэффициенты равны нулю.

Координатами многочлена

$$P(t) = a_n t^n + a_{n-1} t^{n-1} + \dots + a_1 t + a_0$$

в этом базисе являются его коэффициенты $a_0, a_1, \dots, a_{n-1}, a_n$. Размерность данного пространства $n + 1$.

Обратимся, наконец, к примеру 3.3. Пусть N – произвольное натуральное число. Непрерывные функции $f_1(t) = 1$, $f_2(t) = t$, ..., $f_N(t) = t^{N-1}$ образуют совокупность N линейно независимых векторов, так как если многочлен степени $N - 1$ равен нулю для любых $t \in [a, b]$, то все его коэффициенты равны нулю. Значит, в данном пространстве $C[a, b]$ для любого натурального числа N имеется набор из N линейно независимых функций, и, следовательно, пространство $C[a, b]$ бесконечномерно.

3.5. Изоморфность конечномерных пространств

Так как единственными операциями, которые введены в линейных пространствах, являются операции сложения векторов и умножения вектора на число, то естественно ввести следующее определение.

Определение 3.7. *Линейные пространства $\mathbb{L}$ и $\mathbb{L}'$ (оба вещественные или оба комплексные) называются изоморфными, если между элементами этих пространств можно установить взаимно-однозначное соответствие так, что если вектору $x \in \mathbb{L}$ соответствует вектор $x' \in \mathbb{L}'$ ($x \longleftrightarrow x'$), а вектору $y \in \mathbb{L}$ соответствует вектор $y' \in \mathbb{L}'$ ($y \longleftrightarrow y'$), то:*

1) вектору $(x + y) \in \mathbb{L}$ соответствует вектор $(x' + y') \in \mathbb{L}'$ ($(x + y) \longleftrightarrow (x' + y')$);

2) если λ – число, то вектору $(\lambda x) \in \mathbb{L}$ соответствует вектор $(\lambda x') \in \mathbb{L}'$ $((\lambda x) \longleftrightarrow (\lambda x'))$.

Из этого определения следует, что, если в изоморфных пространствах $\mathbb{L}$ и $\mathbb{L}'$ векторам $x_1, x_2, \dots, x_n \in \mathbb{L}$ соответствуют векторы $x'_1, x'_2, \dots, x'_n \in \mathbb{L}'$ $(x_i \longleftrightarrow x'_i)$, то линейной комбинации векторов $x_1, \dots, x_n$ с какими-то коэффициентами соответствует линейная комбинация векторов $x'_1, \dots, x'_n$ с теми же коэффициентами:

$$\lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_n x_n \longleftrightarrow \lambda_1 x'_1 + \lambda_2 x'_2 + \dots + \lambda_n x'_n.$$

Утверждение 3.4. Если пространства $\mathbb{L}$ и $\mathbb{L}'$ изоморфны, то набору линейно независимых векторов из $\mathbb{L}$ соответствует набор линейно независимых векторов из $\mathbb{L}'$, и наоборот.

Доказательство. Покажем сначала, что если пространства $\mathbb{L}$ и $\mathbb{L}'$ изоморфны, то нулевому вектору $\Theta \in \mathbb{L}$ соответствует нулевой вектор $\Theta' \in \mathbb{L}'$. Действительно, пусть $x \longleftrightarrow x'$, но тогда (согласно свойству 2 из определения 3.7) $\Theta = 0x \longleftrightarrow 0x' = \Theta'$.

Пусть набору векторов $x_1, x_2, \dots, x_n \in \mathbb{L}$ соответствует набор векторов $x'_1, x'_2, \dots, x'_n \in \mathbb{L}'$. Условия $\Theta \longleftrightarrow \Theta'$ и

$$\lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_n x_n \longleftrightarrow \lambda_1 x'_1 + \lambda_2 x'_2 + \dots + \lambda_n x'_n$$

означают, что из справедливости равенства

$$\lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_n x_n = \Theta$$

следует справедливость равенства

$$\lambda_1 x'_1 + \lambda_2 x'_2 + \dots + \lambda_n x'_n = \Theta'$$

с теми же коэффициентами $\lambda_1, \lambda_2, \dots, \lambda_n$, и наоборот. Отсюда следует, что рассматриваемые наборы векторов могут быть линейно зависимы или линейно независимы только одновременно. ■

Утверждение 3.5. Если 2 линейных пространства $\mathbb{L}$ и $\mathbb{L}'$ изоморфны, то их размерности равны $(\dim(\mathbb{L}) = \dim(\mathbb{L}'))$.

Доказательство. Пусть $\mathbb{L}$ и $\mathbb{L}'$ изоморфны. Из утверждения 3.4 следует, что максимальное число линейно независимых векторов в $\mathbb{L}$ и $\mathbb{L}'$ одно и то же, т.е. размерности $\mathbb{L}$ и $\mathbb{L}'$ равны. ■

Справедливо и обратное утверждение.

Теорема 3.6. Все конечномерные линейные пространства, имеющие одну и ту же размерность n , изоморфны друг другу.

Доказательство. Пусть $\mathbb{L}$ и $\mathbb{L}'$ – 2 линейных пространства размерности n . Выберем базис $e_1, e_2, \dots, e_n$ в пространстве $\mathbb{L}$ и базис $e'_1, e'_2, \dots, e'_n$ в $\mathbb{L}'$. Произвольному вектору $x \in \mathbb{L}$:

$$x = \xi_1 e_1 + \xi_2 e_2 + \dots + \xi_n e_n = \sum_{i=1}^n \xi_i e_i, \quad (3.11)$$

поставим в соответствие вектор $x' \in \mathbb{L}'$ по правилу

$$x' = \xi_1 e'_1 + \xi_2 e'_2 + \dots + \xi_n e'_n = \sum_{i=1}^n \xi_i e'_i$$

т. е. линейную комбинацию векторов e'_i с теми же коэффициентами, что и в (3.11).

Это соответствие взаимно-однозначно. Действительно, каждый вектор $x \in \mathbb{L}$ однозначно представляется в виде (3.11), поэтому числа ξ_i , а следовательно, и вектор $x' \in \mathbb{L}'$ определяются по x однозначно. Пространства $\mathbb{L}$ и $\mathbb{L}'$ в данном построении равноправны, значит, каждому $x' \in \mathbb{L}'$ отвечает вектор из $\mathbb{L}$, и притом один.

Проверим свойства 1 и 2 из определения 3.7. Пусть $x = \sum_{i=1}^n \xi_i e_i \in \mathbb{L}$

и $y = \sum_{i=1}^n \eta_i e_i \in \mathbb{L}$, тогда $x + y = \sum_{i=1}^n (\xi_i + \eta_i) e_i$ и по построению вектору $x + y$ соответствует вектор

$$\sum_{i=1}^n (\xi_i + \eta_i) e'_i = \sum_{i=1}^n \xi_i e'_i + \sum_{i=1}^n \eta_i e'_i = x' + y'.$$

Аналогично,

$$\lambda x = \sum_{i=1}^n (\lambda \xi_i) e_i \longleftrightarrow \sum_{i=1}^n (\lambda \xi_i) e'_i = \lambda \sum_{i=1}^n \xi_i e'_i = \lambda x'.$$

Итак, изоморфизм $\mathbb{L}$ и $\mathbb{L}'$ доказан. ■

Из теоремы 3.6 следует, что самой существенной характеристикой конечномерного линейного пространства является его размерность.

3.6. n -мерные векторные пространства $\mathbb{C}^n$ и $\mathbb{R}^n$

Изоморфность конечномерных пространств одинаковой размерности позволяет при изучении линейных пространств фиксированной размерности ограничиться рассмотрением какого-то одного (произвольного) пространства.

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad x_1, x_2, \dots, x_n \in \mathbb{C}(\mathbb{R}).$$

Рассмотрим множество, состоящее из t векторов:

$$\vec{a}_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix}, \quad \vec{a}_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix}, \dots, \quad \vec{a}_m = \begin{bmatrix} a_{1m} \\ a_{2m} \\ \vdots \\ a_{nm} \end{bmatrix}. \quad (3.12)$$

$$\lambda_1 \vec{a}_1 + \lambda_2 \vec{a}_2 + \cdots + \lambda_m \vec{a}_m = \vec{0}$$
[illegible]
$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix},$$

¹ Напомним, что в абстрактном пространстве $\mathbb{L}$ нулевой вектор обозначался Θ .

Отметим, что всякие m векторов (3.12) n -мерного пространства $\mathbb{C}^n$ (или $\mathbb{R}^n$) составляют при $m > n$ линейно зависящую систему векторов.

Действительно, в однородной системе (3.13) в этом случае число неизвестных m больше числа уравнений n , поэтому данная система имеет бесконечно много (отличных от тривиального) решений (см. теорему 2.3). А это и означает, что система векторов линейно зависима.

Пусть теперь дана произвольная система n векторов:

$$\vec{a}_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix}, \quad \vec{a}_2 = \begin{bmatrix} a_{21} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix}, \dots, \quad \vec{a}_n = \begin{bmatrix} a_{n1} \\ a_{n2} \\ \vdots \\ a_{nn} \end{bmatrix}. \quad (3.14)$$

Теорема 3.7. *Для того чтобы система векторов была линейно независимой, необходимо и достаточно, чтобы*

$$\det(A) = \det \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \neq 0.$$

Доказательство. В этом случае векторному уравнению

$$\lambda_1 \vec{a}_1 + \lambda_2 \vec{a}_2 + \cdots + \lambda_n \vec{a}_n = \vec{0}$$

соответствует однородная система линейных уравнений с квадратной матрицей коэффициентов A :

[illegible]

которая, согласно теореме 2.4, имеет только тривиальное решение тогда и только тогда, когда $\det(A) \neq 0$. ■

Следствие 3.3. Если $\det(A) = 0$, то набор векторов, составленный из столбцов матрицы A , линейно зависим.

Таким образом, если взять произвольную невырожденную квадратную матрицу порядка n , то ее столбцы (а также транспонированные строки) образуют множество из n линейно независимых векторов. Согласно определению 3.5, такой набор векторов является базисом. Например, взяв $A = I$ (I – единичная матрица порядка n), получаем стандартный базис

пространств $\mathbb{C}^n$ и $\mathbb{R}^n$:

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \dots, \vec{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

Приведем другие примеры базисов:

1. Векторы $\vec{a}_1 = \begin{bmatrix} 1 \\ i \end{bmatrix}$, $\vec{a}_2 = \begin{bmatrix} i \\ -1 + i \end{bmatrix}$ образуют базис в $\mathbb{C}^2$, так как

$$\det \begin{bmatrix} 1 & i \\ i & -1+i \end{bmatrix} = -1 + i - i^2 = i \neq 0.$$

2. Векторы $\vec{a}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\vec{a}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $\vec{a}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ образуют базис в $\mathbb{R}^3$ (а также в $\mathbb{C}^3$), так как

$$\det \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} = 1 \neq 0.$$

Заметим, что векторы $\vec{a}_2, \vec{a}_1, \vec{a}_3$, согласно определению 3.5, образуют уже другой базис того же пространства (важен порядок векторов).

Из изложенного следует, что существует сколько угодно различных базисов в n -мерном пространстве $\mathbb{C}^n(\mathbb{R}^n)$. Действительно, достаточно взять столбцы любой невырожденной матрицы соответствующего размера.

Рассмотрим более подробно задачу нахождения координат вектора при его разложении по базису в пространстве $\mathbb{C}^n(\mathbb{R}^n)$. Пусть векторы $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$ образуют базис в $\mathbb{C}^n$, а $\vec{b} \in \mathbb{C}^n$ – произвольный вектор этого пространства. Рассмотрим уравнение

$$\vec{b} = \beta_1 \vec{a}_1 + \beta_2 \vec{a}_2 + \cdots + \beta_n \vec{a}_n.$$

Оно равносильно системе линейных уравнений

[illegible]

Система (3.15) имеет единственное решение по теореме Крамера (теорема 2.2). Для нахождения координат $\beta_1, \beta_2, \dots, \beta_n$ вектора $\vec{b}$ в базисе

$\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$ достаточно решить систему (3.15), например методом Гаусса–Жордана.

Из изложенного следует равносильность следующих утверждений:

- 1) $\det(A) \neq 0$;
- 2) столбцы матрицы A образуют базис в $\mathbb{C}^n$;
- 3) система линейных уравнений (3.15) имеет (при любом столбце свободных членов $\vec{b}$) единственное решение.

3.7. Подпространства линейного пространства

Определение 3.8. *Подпространством $\mathbb{L}'$ пространства $\mathbb{L}$ называется совокупность таких элементов из $\mathbb{L}$, что они сами образуют линейное пространство с введенными в $\mathbb{L}$ операциями сложения и умножения на число.*

Другими словами, совокупность $\mathbb{L}'$ элементов из $\mathbb{L}$ образует линейное подпространство пространства $\mathbb{L}$, если из $x \in \mathbb{L}'$, $y \in \mathbb{L}'$ следует $x + y \in \mathbb{L}'$ и $\alpha x \in \mathbb{L}'$ для любого числа α . Выполнение аксиом линейного пространства при этом гарантируется их выполнением в пространстве $\mathbb{L}$.

Очевидно, что любое подпространство $\mathbb{L}'$ пространства $\mathbb{L}$ содержит нулевой элемент $\Theta \in \mathbb{L}$.

Так как подпространство само есть линейное пространство, то все конструкции линейного пространства (базис, размерность и т. д.) применимы и к подпространству. Очевидно, что в подпространстве не может быть больше линейно независимых векторов, чем во всем пространстве, т. е. *размерность любого подпространства не превосходит размерности всего пространства.*

Приведем примеры подпространств.

Пример 3.6. Нулевое подпространство, т. е. подпространство, состоящее из единственного элемента, – Θ . Естественно, что $\dim(\Theta) = 0$. •

Пример 3.7. Все пространство $\mathbb{L}$. •

Нулевое подпространство и все пространство называются обычно *несобственными* подпространствами.

Пример 3.8. В пространстве $\mathbb{C}^n$ совокупность всех тех векторов $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, для которых $x_1 = 0$, образует подпространство. •

Пример 3.9. В пространстве $\mathbb{C}[a, b]$ совокупность многочленов степени не выше n является подпространством. •

Определение 3.9. Возьмем в линейном пространстве $\mathbb{L}$ произвольное конечное множество векторов $f_1, f_2, \dots, f_m$; совокупность всевозможных линейных комбинаций выбранных векторов называется линейной оболочкой векторов $f_1, f_2, \dots, f_m$, которую будем обозначать $\mathcal{L}(f_1, f_2, \dots, f_m)$.

Утверждение 3.6. Линейная оболочка произвольного множества векторов $f_1, f_2, \dots, f_m$ линейного пространства $\mathbb{L}$ является подпространством этого пространства.

Доказательство. Складывая между собой и умножая на числа линейные комбинации векторов $f_1, f_2, \dots, f_m$, снова получаем линейные комбинации этих же векторов, т.е. элементы линейной оболочки. Согласно определению 3.8, это означает, что линейная оболочка является подпространством. ■

Линейная оболочка векторов $f_1, f_2, \dots, f_m$ называется *подпространством, порожденным векторами* $f_1, f_2, \dots, f_m$.

Утверждение 3.7. Подпространство $\mathbb{L}'$, порожденное линейно независимыми векторами $e_1, e_2, \dots, e_k$, является k -мерным, и векторы

$$e_1, e_2, \dots, e_k$$

образуют в нем базис.

Доказательство. В $\mathbb{L}'$ имеется система k линейно независимых векторов – векторы $e_1, e_2, \dots, e_k$. Если $x_1, x_2, \dots, x_l$ – произвольные линейно независимые векторы из $\mathbb{L}'$, то они являются линейными комбинациями векторов

$$e_1, e_2, \dots, e_k,$$

так как $\mathbb{L}'$ есть линейная оболочка векторов $e_1, e_2, \dots, e_k$. Значит, согласно теореме 3.3, $l \leq k$ и, следовательно, $\mathbb{L}'$ k -мерно. Набор векторов $e_1, e_2, \dots, e_k$ есть один из возможных базисов в $\mathbb{L}'$. ■

Самыми простыми (если исключить нулевое подпространство) являются одномерные подпространства. Базис всякого такого подпространства состоит из одного ненулевого вектора $e_1 \in \mathbb{L}$. Таким образом, одномерное подпространство состоит из векторов вида αe_1 , где α – произвольное число.

Аналогично, векторы вида $\alpha e_1 + \beta e_2$, где e_1 и e_2 – фиксированные линейно независимые векторы, а α и β – произвольные числа, образуют двухмерное подпространство.

Утверждение 3.8. В пространстве $\mathbb{C}^n(\mathbb{R}^n)$ множество X векто-

ров $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, удовлетворяющих соотношению

$$\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n = 0, \quad (3.16)$$

где $\alpha_1, \alpha_2, \dots, \alpha_n$ – фиксированные числа, не все равные нулю, образует подпространство размерности $n - 1$.

Доказательство. Предположим, что в (3.16) $\alpha_1 \neq 0$, и рассмотрим набор векторов

$$\vec{f}_1 = \begin{bmatrix} -\frac{\alpha_2}{\alpha_1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \vec{f}_2 = \begin{bmatrix} -\frac{\alpha_3}{\alpha_1} \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \dots, \quad \vec{f}_{n-1} = \begin{bmatrix} -\frac{\alpha_n}{\alpha_1} \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix},$$

очевидно линейно независимый. Непосредственно проверяется, что любой вектор $\vec{x} \in X$ можно записать в виде

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = x_2 \vec{f}_1 + x_3 \vec{f}_2 + \dots + x_n \vec{f}_{n-1},$$

т. е. представить в виде линейной комбинации $n - 1$ векторов, причем набор векторов $\vec{f}_1, \vec{f}_2, \dots, \vec{f}_{n-1}$ линейно независим. ■

3.8. Разложение линейного пространства $\mathbb{L}$ в прямую сумму подпространств

Пусть заданы 2 подпространства $\mathbb{L}_1$ и $\mathbb{L}_2$ n -мерного линейного пространства $\mathbb{L}$.

Определение 3.10. Если любой вектор x пространства $\mathbb{L}$ можно, и притом единственным образом, представить как сумму двух векторов

$$x = x_1 + x_2,$$

где $x_1 \in \mathbb{L}_1$, а $x_2 \in \mathbb{L}_2$, то говорят, что пространство $\mathbb{L}$ разложено в прямую сумму подпространств $\mathbb{L}_1$ и $\mathbb{L}_2$.

Это обычно записывают так:

$$\mathbb{L} = \mathbb{L}_1 \dot{+} \mathbb{L}_2.$$

Теорема 3.8. Пусть для подпространств $\mathbb{L}_1$ и $\mathbb{L}_2$ пространства $\mathbb{L}$ выполнены условия:

1. Подпространства $\mathbb{L}_1$ и $\mathbb{L}_2$ имеют только один общий вектор $x = \Theta$ (нулевой вектор).
2. Сумма размерностей этих подпространств равна размерности пространства $\mathbb{L}$ ($\dim \mathbb{L}_1 + \dim \mathbb{L}_2 = \dim \mathbb{L}$).

Тогда линейное пространство $\mathbb{L}$ разлагается в прямую сумму подпространств $\mathbb{L}_1$ и $\mathbb{L}_2$.

Доказательство. Выберем некоторый базис $e_1, \dots, e_k$ в подпространстве $\mathbb{L}_1$ и базис $f_1, \dots, f_l$ в подпространстве $\mathbb{L}_2$. Поскольку сумма размерностей $\mathbb{L}_1$ и $\mathbb{L}_2$ есть n , то $k + l = n$.

Покажем, что векторы $e_1, \dots, e_k, f_1, \dots, f_l$ линейно независимы, т. е. образуют базис пространства $\mathbb{L}$. Действительно, пусть

$$\lambda_1 e_1 + \dots + \lambda_k e_k + \mu_1 f_1 + \dots + \mu_l f_l = \Theta;$$

отсюда

$$\lambda_1 e_1 + \dots + \lambda_k e_k = -\mu_1 f_1 - \dots - \mu_l f_l.$$

Левая часть этого соотношения есть вектор из $\mathbb{L}_1$, а правая из $-\mathbb{L}_2$. Так как, по условию, единственный общий вектор $\mathbb{L}_1$ и $\mathbb{L}_2$ есть нулевой вектор, то

$$\lambda_1 e_1 + \dots + \lambda_k e_k = \Theta, \tag{3.17}$$

$$\mu_1 f_1 + \dots + \mu_l f_l = \Theta. \tag{3.18}$$

Но каждый из наборов $e_1, \dots, e_k$ и $f_1, \dots, f_l$ состоит из линейно независимых векторов, так как это базисы в $\mathbb{L}_1$ и $\mathbb{L}_2$. Поэтому из (3.17) следует, что

$$\lambda_1 = \dots = \lambda_k = 0,$$

а из (3.18) следует, что

$$\mu_1 = \dots = \mu_l = 0.$$

Следовательно, система $e_1, \dots, e_k, f_1, \dots, f_l$ состоит из n линейно независимых векторов, т. е. это базис в пространстве $\mathbb{L}$.

Произвольный вектор x из $\mathbb{L}$ можно разложить по векторам этого базиса:

$$x = \alpha_1 e_1 + \dots + \alpha_k e_k + \beta_1 f_1 + \dots + \beta_l f_l.$$

При этом

$$x_1 = \alpha_1 e_1 + \dots + \alpha_k e_k \in \mathbb{L}_1$$

и

$$x_2 = \beta_1 f_1 + \cdots + \beta_l f_l \in \mathbb{L}_2.$$

Таким образом,

$$x = x_1 + x_2,$$

где $x_1 \in \mathbb{L}_1$ и $x_2 \in \mathbb{L}_2$. Покажем, что это разложение единственно.

Предположим, что существуют 2 разложения:

$$x = x_1 + x_2,$$

где $x_1 \in \mathbb{L}_1$, $x_2 \in \mathbb{L}_2$, и

$$x = x'_1 + x'_2,$$

где $x'_1 \in \mathbb{L}_1$, $x'_2 \in \mathbb{L}_2$. Вычитая второе равенство из первого, получаем:

$$x_1 - x'_1 = x'_2 - x_2.$$

Так как вектор, стоящий в левой части равенства, принадлежит $\mathbb{L}_1$, а вектор, стоящий в правой части, принадлежит $\mathbb{L}_2$, то каждый из этих векторов равен нулю, т. е. $x'_1 = x_1$, $x'_2 = x_2$. Единственность разложения доказана. ■

3.9. Сумма и пересечение подпространств

Допустим, что задано 2 произвольных подпространства $\mathbb{L}_1$ и $\mathbb{L}_2$ линейного пространства $\mathbb{L}$. Легко проверить, что совокупность векторов, принадлежащих обоим этим подпространствам, также есть подпространство $\mathbb{L}_0$ пространства $\mathbb{L}$. Это подпространство называется *пересечением* $\mathbb{L}_1$ и $\mathbb{L}_2$ и обозначается

$$\mathbb{L}_0 = \mathbb{L}_1 \cap \mathbb{L}_2.$$

По двум подпространствам $\mathbb{L}_1$ и $\mathbb{L}_2$ можно построить еще одно подпространство, которое называется их *суммой*. Векторами этого подпространства являются всевозможные суммы вида

$$x = x_1 + x_2, \tag{3.19}$$

где $x_1 \in \mathbb{L}_1$, $x_2 \in \mathbb{L}_2$.

Легко проверить, что элементы вида (3.19) образуют подпространство. Это подпространство $\widetilde{\mathbb{L}}$ называется суммой подпространств $\mathbb{L}_1$ и $\mathbb{L}_2$ и обозначается

$$\widetilde{\mathbb{L}} = \mathbb{L}_1 + \mathbb{L}_2.$$

Заметим, что, в отличие от прямой суммы двух подпространств, запись элемента из $\widetilde{\mathbb{L}}$ в виде (3.19) может быть неоднозначной. Докажем сначала вспомогательный результат.

Лемма 3.1. Пусть в подпространстве $\mathbb{L}_1$ с $\dim(\mathbb{L}_1) = k$ имеется l ($l < k$) линейно независимых векторов $f_1, f_2, \dots, f_l$, тогда набор векторов $f_1, f_2, \dots, f_l$ можно дополнить до базиса в $\mathbb{L}_1$.

Доказательство. Так как $l < k$, в $\mathbb{L}_1$ найдется ненулевой вектор (обозначим его f_{l+1}), который не является линейной комбинацией векторов $f_1, f_2, \dots, f_l$, т.е. набор векторов $f_1, f_2, \dots, f_l, f_{l+1}$ – линейно независим (почему?). Если бы такого вектора не нашлось, то это означало бы, что размерность $\mathbb{L}_1$ равна l (любой вектор можно представить в виде линейной комбинации l линейно независимых векторов $f_1, f_2, \dots, f_l$). Если $l + 1 = k$, то лемма доказана, если $l + 1 < k$, то в $\mathbb{L}_1$ найдется вектор f_{l+2} , не являющийся линейной комбинацией векторов $f_1, f_2, \dots, f_l, f_{l+1}$, т.е. получаем набор $f_1, f_2, \dots, f_{l+2}$ линейно независимых векторов. Если $l+2 = k$, то лемма доказана, если $l + 2 < k$, то этот процесс можно продолжить до получения набора из k линейно независимых векторов $f_1, f_2, \dots, f_l, f_{l+1}, \dots, f_k$. ■

Теорема 3.9. Пусть заданы 2 подпространства $\mathbb{L}_1$ и $\mathbb{L}_2$ пространства $\mathbb{L}$. Тогда сумма размерностей $\mathbb{L}_1$ и $\mathbb{L}_2$ равна размерности их суммы плюс размерность их пересечения, т.е. выполнено:

$$\dim(\mathbb{L}_1) + \dim(\mathbb{L}_2) = \dim(\mathbb{L}_1 + \mathbb{L}_2) + \dim(\mathbb{L}_1 \cap \mathbb{L}_2).$$

Доказательство. Выберем в пересечении $\mathbb{L}_0 = \mathbb{L}_1 \cap \mathbb{L}_2$ базис

$$e_1, \dots, e_k. \quad (3.20)$$

Дополним набор векторов (3.20) до базиса в $\mathbb{L}_1$ (см. лемму 3.1):

$$e_1, \dots, e_k, f_1, \dots, f_l \quad (3.21)$$

и дополним (3.20) до базиса в $\mathbb{L}_2$:

$$e_1, \dots, e_k, g_1, \dots, g_m. \quad (3.22)$$

Отметим частные случаи: если $\mathbb{L}_0 = \mathbb{L}_1 \cap \mathbb{L}_2 = \Theta$, то отсутствуют векторы $e_1, \dots, e_k$; если $\mathbb{L}_2 = \mathbb{L}_1 \cap \mathbb{L}_2$, то отсутствуют векторы $f_1, \dots, f_l$; если $\mathbb{L}_1 = \mathbb{L}_1 \cap \mathbb{L}_2$, то отсутствуют векторы $g_1, \dots, g_m$.

Рассмотрим наиболее общий случай (предположим, что множества векторов $e_1, \dots, e_k$, $f_1, \dots, f_l$ и $g_1, \dots, g_m$ не пусты), доказательства в частных случаях можно провести с незначительными изменениями. Покажем, что набор векторов

$$f_1, \dots, f_l, e_1, \dots, e_k, g_1, \dots, g_m \quad (3.23)$$

образует базис в сумме $\tilde{\mathbb{L}} = \mathbb{L}_1 + \mathbb{L}_2$. Сначала покажем, что эти векторы линейно независимы. Действительно, пусть

$$\lambda_1 f_1 + \dots + \lambda_l f_l + \mu_1 e_1 + \dots + \mu_k e_k + \nu_1 g_1 + \dots + \nu_m g_m = \Theta.$$

Тогда

$$\lambda_1 f_1 + \dots + \lambda_l f_l + \mu_1 e_1 + \dots + \mu_k e_k = -\nu_1 g_1 - \dots - \nu_m g_m.$$

Левая часть этого равенства есть вектор из $\mathbb{L}_1$, правая – из $\mathbb{L}_2$. Таким образом, эта правая часть есть одновременно вектор из $\mathbb{L}_1$ и из $\mathbb{L}_2$, т. е. принадлежит $\mathbb{L}_0$, и значит, выражается как линейная комбинация векторов базиса $e_1, \dots, e_k$ подпространства $\mathbb{L}_0$:

$$-\nu_1 g_1 - \dots - \nu_m g_m = c_1 e_1 + \dots + c_k e_k.$$

В силу линейной независимости векторов (3.22) это возможно только когда все коэффициенты – нули. В частности, $\nu_1 = \dots = \nu_m = 0$, т. е. выполнено

$$\lambda_1 f_1 + \dots + \lambda_l f_l + \mu_1 e_1 + \dots + \mu_k e_k = \Theta.$$

Из линейной независимости векторов (3.21) получаем, что и все коэффициенты $\lambda_1, \dots, \lambda_l, \mu_1, \dots, \mu_k$ равны нулю. Таким образом, линейная независимость системы (3.23) доказана.

Покажем теперь, что всякий вектор $x \in \widetilde{\mathbb{L}}$ выражается как линейная комбинация векторов этой системы. По определению $\widetilde{\mathbb{L}}$, вектор x можно представить в виде

$$x = x_1 + x_2,$$

где $x_1 \in \mathbb{L}_1$, $x_2 \in \mathbb{L}_2$. Так как $x_1 \in \mathbb{L}_1$, то его можно представить как линейную комбинацию векторов (3.21). Аналогично $x_2 \in \mathbb{L}_2$ и x_2 можно представить как линейную комбинацию векторов (3.22). Складывая, получим, что вектор x представим как линейная комбинация системы (3.23).

Итак, видим, что векторы

$$f_1, \dots, f_l, e_1, \dots, e_k, g_1, \dots, g_m,$$

с одной стороны, линейно независимы, и, с другой стороны, всякий вектор из $\widetilde{\mathbb{L}}$ есть их линейная комбинация. В силу следствия 3.2 отсюда следует, что эти векторы образуют базис в $\widetilde{\mathbb{L}}$. Итак, имеем k векторов (3.20), образующих базис в $\mathbb{L}_0$, $k + l$ векторов (3.21), образующих базис в $\mathbb{L}_1$, $k + m$ векторов (3.22), образующих базис в $\mathbb{L}_2$, и $k + l + m$ векторов (3.23), образующих базис в $\widetilde{\mathbb{L}} = \mathbb{L}_1 + \mathbb{L}_2$. Таким образом, утверждение теоремы, означающее, что

$$(k + l) + (k + m) = (k + l + m) + k,$$

доказано. ■

Из доказанной теоремы следует, что любым двум 15-мерным подпространствам „тесно“ в 28-мерном пространстве – они пересекаются по крайней мере по двухмерному подпространству. Действительно, сумма их размерностей равна 30, а размерность суммы не может, конечно, превосходить размерности всего пространства, т. е. 28.

3.10. Ранг матрицы

Рассмотрим произвольную матрицу A размера $m \times n$. Пусть $k \leq m$ и $k \leq n$. Выделим в этой матрице какие-нибудь k строк и k столбцов. Элементы, стоящие на пересечении этих строк и столбцов, образуют квадратную матрицу k -го порядка. Ее определитель называется минором k -го порядка матрицы A .

Пример 3.10. Пусть дана матрица

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 0 & -1 & 1 \\ 0 & 4 & 5 & 0 \end{bmatrix}.$$

Для этой матрицы существует 12 миноров первого порядка, 18 миноров второго порядка и 4 минора третьего порядка. Например, если выделить первую и третью строки и первый и третий столбцы матрицы, то получим $\det \begin{bmatrix} 1 & 3 \\ 0 & 5 \end{bmatrix} = 5$ – один из миноров второго порядка для матрицы A . •

Определение 3.11. Рангом ненулевой матрицы A называется наивысший порядок отличных от нуля миноров этой матрицы. Обозначается $r(A)$ или $\text{rank}(A)$. Ранг нулевой матрицы полагаем равным нулю ($\text{rank}(\Theta) = 0$).

Таким образом, если ранг матрицы равен r , то среди миноров этой матрицы есть по крайней мере один минор r -го порядка, отличный от нуля, в то время как все ее миноры порядка $r + 1$ и выше равны нулю.

Определение 3.12. Если $\text{rank}(A) = r$, то любой отличный от нуля минор порядка r называется базисным минором матрицы A .

Пример 3.11. Рассмотрим матрицу

$$A = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 0 & -1 \\ 0 & 4 & 5 \end{bmatrix}.$$

Определитель этой матрицы

$$\det(A) = 3 \det \begin{bmatrix} 0 & -1 \\ 4 & 5 \end{bmatrix} - 2 \det \begin{bmatrix} 2 & 1 \\ 4 & 5 \end{bmatrix} = 3 \cdot 4 - 2 \cdot 6 = 0,$$

но так как $\det \begin{bmatrix} 3 & 2 \\ 2 & 0 \end{bmatrix} = -4 \neq 0$, то $\text{rank}(A) = 2$. •

Теорема 3.10. Ранг транспонированной¹ матрицы равен рангу исходной, т. е. $\text{rank}(A) = \text{rank}(A')$.

Доказательство. Согласно утверждению 2.2 для определителей, каждый минор транспонированной матрицы равен некоторому минору исходной матрицы, и наоборот. Следовательно, $r(A) = r(A')$. ■

Будем рассматривать векторы, которые соответствуют столбцам и строкам матрицы (при этом произвольную строку $(m \times n)$ -матрицы рассматриваем как n -мерный вектор, равный транспонированной строке), и говорить, что столбцы или строки линейно независимы, если линейно независима соответствующая система векторов.

Теорема 3.11. Если ранг матрицы равен r , то в этой матрице можно найти r линейно независимых строк (столбцов), причем все остальные ее строки (столбцы), если они есть, выражаются в виде линейных комбинаций этих r линейно независимых строк (столбцов).

Доказательство. Пусть дана матрица A размера $m \times n$, причем $\text{rank}(A) = r$. Предположим для определенности, что базисному минору r -го порядка этой матрицы соответствует подматрица, расположенная в левом верхнем углу, т. е. что

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \text{ и } \Delta = \det \begin{bmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & \ddots & \vdots \\ a_{r1} & \cdots & a_{rr} \end{bmatrix} \neq 0.$$

Докажем, что в таком случае первые r строк матрицы A будут линейно независимы. (Если отличен от нуля не этот, а какой-нибудь другой минор r -го порядка матрицы A , то линейно независимыми будут именно те строки, которые образуют этот базисный минор.) Предположим, что, наоборот, эти строки¹ линейно зависимы. Тогда одна из них, пусть, для определенности, $A^{(r)}$, линейно выражается через остальные:

$$A^{(r)} = \alpha_1 A^{(1)} + \alpha_2 A^{(2)} + \cdots + \alpha_{r-1} A^{(r-1)}.$$

Вычтем из r -й строки матрицы A первую строку, умноженную на α_1 , вторую, умноженную на α_2 , и т. д., наконец, $(r-1)$ -ю строку, умноженную на α_{r-1} . После таких преобразований r -я строка полученной матрицы $\tilde{A}$ окажется состоящей из одних нулей. При этом определитель Δ не должен

¹ Напомним, что для матрицы A транспонированная матрица обозначается A' (см. определение 1.7).

¹ Заметим, что для матрицы A ее i -я строка обозначается $A^{(i)}$, а ее j -й столбец — $A_{(j)}$ (см. с. 23).

Докажем теперь вторую часть теоремы – о том, что все остальные строки матрицы A линейно выражаются через ее первые r строк. Пусть $r < k \leq m$ и $1 \leq l \leq n$. Рассмотрим определитель матрицы $(r+1)$ -го порядка:

Он равен нулю при всех k и l : если $l \leq r$, то у него 2 одинаковых столбца, если же $l > r$, то это минор $(r+1)$ -го порядка матрицы ранга r , и поэтому $D = 0$. Разложим определитель D по элементам последнего столбца:

где через A_i обозначены соответствующие алгебраические дополнения. Заметим, что алгебраические дополнения $A_1, \dots, A_{r+1}$ элементов последнего столбца зависят от k , но не зависят от l . Кроме того, $A_{r+1} = (-1)^{2r+2} \Delta \neq 0$, и, следовательно, разделив на A_{r+1} , получим:

где $c_i = -\frac{A_i}{A_{r+1}}$ не зависят от l . Подставляя $l = 1, 2, \dots, n$, получим:

Но эти равенства означают, что k -я строка матрицы A линейно выражается через ее первые r строк:

Доказательство для столбцов проводится совершенно аналогично, с использованием матрицы A' . Ее ранг также равен r , согласно теореме 3.10, а строками служат столбцы матрицы A . ■

Следствие 3.4. Если $\text{rank}(A) = r$, то максимальное число линейно независимых столбцов матрицы A равно максимальному числу линейно независимых строк и равно r .

Следствие 3.5. *Размерности линейных оболочек столбцов произвольной прямоугольной матрицы A*

$$A_{(1)}, A_{(2)}, \dots, A_{(n)}$$

и линейных оболочек транспонированных строк матрицы A

$$\left(A^{(1)}\right)', \left(A^{(2)}\right)', \dots, \left(A^{(m)}\right)'$$

совпадают и равны рангу матрицы A ($\text{rank}(A)$).

Данное следствие вытекает из следствий 3.4 и 3.2.

Для вычисления ранга матрицы ее сначала приводят к наиболее простому виду с помощью так называемых допустимых преобразований. Допустимыми называются (аналогично допустимым преобразованиям, рассмотренным в гл. 1) следующие преобразования матрицы:

- 1) перестановка строк или столбцов;
- 2) умножение всех элементов строки или столбца на любое число, отличное от нуля;
- 3) прибавление ко всем элементам строки или столбца соответствующих элементов другой строки или столбца, умноженных на одно и то же число.

Теорема 3.12. *При элементарных преобразованиях матрицы ее ранг не меняется.*

Доказательство. С учетом теоремы 3.10 достаточно доказать эту теорему только для столбцов.

Пусть в матрице A произвольно переставлены ее столбцы. Покажем, что эта операция не меняет ранг матрицы. Действительно, размерность линейной оболочки векторов

$$A_{(1)}, A_{(2)}, \dots, A_{(n)}$$

не зависит от того, в каком порядке они записаны; следовательно, согласно следствию 3.5, и ранг матрицы не зависит от порядка ее столбцов.

При умножении всех элементов j -го столбца матрицы на число $c \neq 0$ из ее миноров, которые не содержат элементов j -го столбца, не изменятся, а остальные миноры умножаются на $c \neq 0$ (утверждение 2.2), т. е. если минор равнялся нулю, то он останется равным нулю, если же он был отличен от нуля, то также будет не равным нулю. Это означает, что наивысший порядок отличного от нуля минора матрицы не изменится, следовательно, второе допустимое преобразование не изменяет ранг матрицы.

Пусть к j -му столбцу матрицы A прибавлен m -й столбец этой же матрицы, умноженный на число λ . Это означает, что система векторов

$$A_{(1)}, \dots, A_{(j)}, \dots, A_{(m)}, \dots, A_{(n)}$$

заменена на систему

$$A_{(1)}, \dots, A_{(j)} + \lambda A_{(m)}, \dots, A_{(m)}, \dots, A_{(n)}.$$

Покажем, что линейные оболочки $\mathbb{L}_1$ и $\mathbb{L}_2$ обеих этих систем совпадают. Действительно, все векторы второй системы входят в линейную оболочку первой системы, поэтому $\mathbb{L}_2 \subset \mathbb{L}_1$. С другой стороны, равенство

$$a_{(j)} = (a_{(j)} + \lambda a_{(m)}) - \lambda a_{(m)}$$

показывает, что вектор $a_{(j)}$ входит в линейную оболочку векторов второй системы; так как все остальные векторы первой системы, очевидно, также входят в эту линейную оболочку, то $\mathbb{L}_1 \subset \mathbb{L}_2$. Отсюда вытекает, что $\mathbb{L}_1 = \mathbb{L}_2$, поэтому ранг матрицы A не меняется в результате третьего допустимого преобразования. ■

Утверждение 3.9. *Исключение из матрицы столбца, состоящего из нулей, дает матрицу того же ранга (но другого размера).*

Аналогичное утверждение справедливо и для строки, состоящей из нулей.

Доказательство. Пусть в матрице имеется столбец, состоящий из нулей, тогда ему отвечает нулевой вектор Θ . Очевидно, что исключение нулевого вектора из системы векторов $a_{(1)}, a_{(2)}, \dots, a_{(n)}$ не меняет линейной оболочки векторов $a_{(1)}, a_{(2)}, \dots, a_{(n)}$, следовательно, не изменяются ее размерность и ранг матрицы. Как обычно, утверждение для строк следует из теоремы 3.10. ■

Для нахождения ранга матрицы ее при помощи допустимых преобразований приводят либо к единичной, либо к матрице вида

$$\underbrace{\left[\begin{array}{ccccc} 1 & 0 & \cdots & 0 & \cdots \\ 0 & 1 & \cdots & 0 & \cdots \\ \vdots & \vdots & \ddots & \vdots & \ddots \\ 0 & 0 & \cdots & 1 & \cdots \end{array} \right]}_{r+s},$$

ранг которой, очевидно, равен r . Это всегда возможно при помощи элементарных преобразований алгоритма Гаусса–Жордана (см. гл. 1) с учетом теоремы 3.12 и утверждения 3.9.

Пример 3.12. Найдём ранг матрицы

$$A = \begin{bmatrix} 3 & 2 & 1 & 2 \\ 2 & 0 & -1 & 1 \\ 0 & 4 & 5 & 1 \end{bmatrix}.$$

$\boxed{\Leftarrow}$ Пусть $\text{rank}(A) = \text{rank}(B) = k$. Докажем, что (3.24) совместна. Предположим для определенности, что отличному от нуля минору порядка k матрицы A соответствует ее подматрица, расположенная в ее левом верхнем углу (это можно добиться перестановкой строк – уравнений системы и столбцов):

Тогда первые k строк матрицы B линейно независимы, а так как ранг ее равен k , то остальные строки матрицы B линейно выражаются через первые ее k строк. Но это означает, что последние $m - k$ уравнений системы (3.24) получаются как суммы первых k уравнений с некоторыми коэффициентами (говорят, что последние $m - k$ уравнений являются следствиями первых k уравнений системы). Следовательно, достаточно решить первые k уравнений системы; их решения автоматически будут удовлетворять и остальным $m - k$ уравнениям. Далее возможны 2 случая:

[illegible]

2) $k < n$. Возьмем первые k уравнений системы и, оставив в левых частях системы первые k неизвестных, остальные перенесем в правые части:

Назовем неизвестные $x_{k+1}, \dots, x_n$ свободными. Им можно придавать какие угодно значения, получая при этом соответствующие значения неизвестных $x_1, \dots, x_k$ из системы (3.25), в которой определитель матрицы $D \neq 0$

(по теореме Крамера она имеет единственное решение при любых правых частях). Система (3.25) совместна и имеет бесконечно много решений. ■

Замечание 3.2. Отметим, что метод Гауса–Жордана решения системы линейных алгебраических уравнений в общем случае требует перестановок как строк, так и столбцов в исходной матрице. При этом такие перестановки определены, в общем случае, неоднозначно. Несмотря на эту неоднозначность из теоремы Кронекера–Капелли следует, что если система имеет решение, то в методе Гауса–Жордана при любых допустимых перестановках строк и столбцов конечная матрица всегда имеет одно и то же число ненулевых строк, равное рангу исходной матрицы. ⊗

Пример 3.13. Рассмотрим систему

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

Так как $\det \begin{bmatrix} 1 & 2 & 3 \\ 1 & -1 & 1 \\ 1 & 3 & -1 \end{bmatrix} = 14 \neq 0$, то $\text{rank}(A) = 3$. Так как

$$\det(B) = \det \begin{bmatrix} 1 & 2 & 3 & 2 \\ 1 & -1 & 1 & 0 \\ 1 & 3 & -1 & -2 \\ 3 & 4 & 3 & 0 \end{bmatrix} = 0,$$

то $\text{rank}(B) < 4$, но $\text{rank}(B) \geq r(A)$. Следовательно, $\text{rank}(B) = \text{rank}(A) = 3$, и система совместна.

Действительно, в данной системе четвертое уравнение является следствием первых трех (четвертое уравнение равно сумме первых трех уравнений), и его можно исключить из системы. Из первых трех уравнений системы находим: $x_1 = -1$, $x_2 = 0$, $x_3 = 1$ (например, по формулам Крамера). •

Пример 3.14. Пусть

$$\begin{cases} x_1 + 2x_2 + 3x_3 - x_4 = 0, \\ x_1 - x_2 + x_3 + 2x_4 = 4, \\ x_1 + 5x_2 + 5x_3 - 4x_4 = -4, \\ x_1 + 8x_2 + 7x_3 - 7x_4 = -8. \end{cases}$$

Для вычисления рангов матриц A и B можно применить алгоритм Гаусса–Жордана к расширенной матрице системы. Заметим, что в этом

случае удастся решить задачу без перестановок строк и столбцов:

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

$$\Leftrightarrow \begin{pmatrix} 1 & 2 & 3 & -1 & | & 0 \\ 0 & 1 & \frac{2}{3} & -1 & | & -\frac{4}{3} \\ 0 & 3 & 2 & -3 & | & -4 \\ 0 & 6 & 4 & -6 & | & -8 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1 & 0 & \frac{5}{3} & 1 & | & \frac{8}{3} \\ 0 & 1 & \frac{2}{3} & -1 & | & -\frac{4}{3} \\ 0 & 0 & 0 & 0 & | & 0 \\ 0 & 0 & 0 & 0 & | & 0 \end{pmatrix}.$$

Отсюда следует, что $\text{rank}(A) = \text{rank}(B) = 2$ и, следовательно, данная система совместна.

Отметим, что нахождение рангов матриц A и B для конкретной системы методом Гаусса–Жордана приводит одновременно к решению системы. Действительно, из последней таблицы получается:

$$\begin{cases} x_1 = 8/3 - 5/3x_3 - x_4, \\ x_2 = -4/3 - 2/3x_3 + x_4. \end{cases} \bullet$$

Пример 3.15. Пусть

$$\begin{cases} x_1 + 2x_2 + 3x_3 - x_4 = 0, \\ x_1 - x_2 + x_3 + 2x_4 = 4, \\ x_1 + 5x_2 + 5x_3 - 4x_4 = -4, \\ x_1 + 8x_2 + 7x_3 - 7x_4 = 6. \end{cases}$$

Так как $\text{rank}(A) = 2$, $\text{rank}(B) = 3$ (проверьте), то данная система несовместна. $\bullet$

3.12. Однородные системы уравнений

Рассмотрим подробно случай однородных систем. Ранее было показано, что такие системы всегда совместны. Выясним, когда однородная система имеет единственное решение и когда она имеет бесконечно много решений. Пусть n – число неизвестных и A – матрица системы.

Теорема 3.14. Если $k = \text{rank}(A)$, то при $k = n$ система имеет единственное (нулевое) решение, при $k < n$ система имеет бесконечно много решений.

Доказательство. Если $k = n$, то имеем квадратную матрицу, определитель которой отличен от нуля, и, следовательно, по теореме Крамера, система имеет единственное (нулевое) решение.

Если $k < n$, то, решая систему методом Гаусса–Жордана, приходим к равносильной системе с расширенной матрицей вида (2.19), которая имеет бесконечно много решений (см. 1.2). ■

Рассмотрим подробнее случай $k < n$. Система имеет вид

[illegible]

или в матричной записи $A\vec{x} = \vec{0}$, где

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}; \quad \vec{0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Пусть

$$x_1 = \alpha_1, \ x_2 = \alpha_2, \ \dots, \ x_n = \alpha_n; \quad \vec{e}_1 = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}$$

– какое-нибудь ненулевое решение системы (3.26), т. е. $A\vec{e}_1 = \vec{0}$. Вектор $c\vec{e}_1$, где c – произвольное число, тоже является решением системы (3.26),

так как $A(c\vec{e}_1) = cA\vec{e}_1 = c\vec{0} = \vec{0}$. Если $\vec{e}_2 = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}$ – какое-нибудь другое

решение системы (3.26), то при любых c_1, c_2 линейная комбинация $c_1\vec{e}_1 + c_2\vec{e}_2$ этих решений тоже будет решением системы (3.26):

$$A(c_1\vec{e}_1 + c_2\vec{e}_2) = Ac_1\vec{e}_1 + Ac_2\vec{e}_2 = c_1A\vec{e}_1 + c_2A\vec{e}_2 = c_1\vec{0} + c_2\vec{0} = \vec{0}.$$

Отсюда следует, что любая линейная комбинация решений однородной системы (3.26) тоже будет ее решением. Нетрудно доказать, что множество решений однородной системы является линейным пространством (пространством решений). Определим размерность этого пространства и найдем его базис.

Определение 3.13. *Линейно независимая система решений*

$$\vec{e}_1, \vec{e}_2, \dots, \vec{e}_p$$

системы (3.26) называется фундаментальной, если каждое решение $\vec{x}$ системы (3.26) является линейной комбинацией решений $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_r$:

$$\vec{x} = c_1 \vec{e}_1 + c_2 \vec{e}_2 + \cdots + c_p \vec{e}_p. \quad (3.27)$$

Это означает, что любая фундаментальная система решений

$$\vec{e}_1, \vec{e}_2, \dots, \vec{e}_p$$

является базисом пространства решений и размерность этого пространства равна p . Формулу (3.27) обычно называют общим решением однородной системы (3.26).

Теорема 3.15. Если ранг k матрицы однородной системы уравнений (3.26) меньше n , то размерность пространства решений этой системы $p = n - k$.

Доказательство. Пусть $\text{rank}(A) = k$, $k < n$, и пусть, для определенности, ненулевой минор k -го порядка соответствует подматрице, стоящей в левом верхнем углу матрицы A :

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}, \quad D = \det \begin{bmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{bmatrix} \neq 0.$$

В этом случае исходная система (3.26) равносильна следующей:

[illegible]

Действительно, так как $D \neq 0$, то последние $n - k$ строк являются линейными комбинациями первых n строк и, следовательно, элементарными операциями можно последние $n - k$ строк превратить в нулевые и исключить из системы. Переносим свободные неизвестные $x_{k+1}, \dots, x_n$ первых k уравнений системы (3.28) в правые части, получим систему относительно неизвестных $x_1, \dots, x_k$ с ненулевым определителем:

[illegible]

Пусть $x_{k+1} = 1, x_{k+2} = 0, \dots, x_n = 0$. Тогда система (3.29) имеет единственное решение $x_1 = \alpha_1, x_2 = \alpha_2, \dots, x_k = \alpha_k$. Это дает нам вектор-решение системы (3.28), а значит, и системы (3.26):

$$\vec{x}_1 = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Аналогично, придавая свободным неизвестным значения $x_{k+1} = 0, x_{k+2} = 1, \dots, x_n = 0$ и вычисляя соответствующие значения неизвестных $x_1 = \beta_1, x_2 = \beta_2, \dots, x_k = \beta_k$, получаем вектор-решение

$$\vec{x}_2 = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

и т. д. Так найдем $p = n - k$ решений системы (3.26):

$$\vec{x}_1 = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \vec{x}_2 = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \dots, \quad \vec{x}_p = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_k \\ 0 \\ \vdots \\ 1 \end{bmatrix}. \quad (3.30)$$

Эти p векторов линейно независимы, так как ранг составленной из них матрицы, очевидно, равен p (нижняя подматрица есть единичная матрица порядка p , и ее определитель не равен 0). Покажем теперь, что решения (3.30) образуют фундаментальную систему решений системы (3.26). Для этого остается показать, что каждое решение системы (3.26) является ли-

нейной комбинацией векторов $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p$. Итак, пусть

$$\vec{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_k \\ a_{k+1} \\ \vdots \\ a_n \end{bmatrix}$$

– произвольное решение системы (3.26). Рассмотрим вектор

$$\vec{b} = \vec{a} - a_{k+1}\vec{x}_1 - a_{k+2}\vec{x}_2 - \dots - a_n\vec{x}_p.$$

Легко видеть, что все элементы, стоящие на последних p местах этого вектора (последние p координат), равны нулю:

$$\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Будучи линейной комбинацией решений, вектор $\vec{b}$ является решением системы (3.26), следовательно, вектор $[b_1 \dots b_k]'$ является решением системы (3.29) с нулевыми правыми частями. Поскольку определитель этой системы отличен от нуля, то по теореме Крамера получаем, что $b_1 = \dots = b_k = 0$. Таким образом, $\vec{b} = \vec{0}$, следовательно,

$$\vec{a} = a_{k+1}\vec{x}_1 + a_{k+2}\vec{x}_2 + \dots + a_n\vec{x}_p,$$

что и требовалось доказать. ■

Заметим, что для получения фундаментальной системы решений можно было бы придавать свободным неизвестным и какие угодно другие значения, лишь бы определитель соответствующей матрицы p -го ранга был отличен от нуля. Так, можно найти сколько угодно фундаментальных систем решений, каждая из которых состоит из p векторов-решений.

Рассмотрим теперь неоднородную систему

$$A\vec{x} = \vec{b} \tag{3.31}$$

и соответствующую ей однородную систему

$$A\vec{x} = \vec{0}. \tag{3.32}$$

Теорема 3.16. Любое решение $\vec{x}$ неоднородной системы (3.31) представимо в виде

$$\vec{x} = \vec{x}_0 + c_1 \vec{e}_1 + c_2 \vec{e}_2 + \cdots + c_p \vec{e}_p, \quad (3.33)$$

где $\vec{x}_0$ – некоторое фиксированное решение системы ((3.31), $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_p$ – фундаментальная система решений однородной системы ((3.32), а $c_1, c_2, \dots, c_p$ – произвольные числа.

Доказательство. Покажем, что вектор $\vec{x}$, определенный формулой (3.33), при любых $c_1, c_2, \dots, c_r$ является решением системы (3.31). Действительно,

$$\begin{aligned} A(\vec{x}_0 + c_1\vec{e}_1 + c_2\vec{e}_2 + \cdots + c_p\vec{e}_p) &= A\vec{x}_0 + c_1A\vec{e}_1 + c_2A\vec{e}_2 + \cdots + c_pA\vec{e}_p = \\ &= \vec{b} + \vec{0} + \vec{0} + \cdots + \vec{0} = \vec{b}. \end{aligned}$$

Докажем теперь, что любое решение неоднородной системы (3.31) может быть представлено в виде (3.33). Пусть $\vec{x}_0$ – какое-то фиксированное решение системы (3.31), а $\vec{x}$ – любое другое ее решение. Тогда $\vec{x}_1 = \vec{x} - \vec{x}_0$ будет решением системы (3.32). Действительно,

$$A\vec{x}_1 = A(\vec{x} - \vec{x}_0) = A\vec{x} - A\vec{x}_0 = \vec{b} - \vec{b} = \vec{0}.$$

С учетом (3.27) $\vec{x}_1 = c_1 \vec{e}_1 + c_2 \vec{e}_2 + \dots + c_p \vec{e}_p$, следовательно,

$$\vec{x} = \vec{x}_0 + c_1\vec{e}_1 + c_2\vec{e}_2 + \cdots + c_p\vec{e}_p,$$

где $c_1, c_2, \dots, c_p$ – произвольные числа. ■

Формулу (3.33) обычно называют общим решением неоднородной системы (3.31).

3.13. Преобразование координат вектора при изменении базиса

Как уже отмечалось, выбор базиса в линейном пространстве неоднозначен. Рассмотрим задачу о нахождении координат вектора в одном базисе по его координатам в другом базисе. Пусть $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ и $\vec{e}'_1, \vec{e}'_2, \dots, \vec{e}'_n$ – 2 базиса n -мерного пространства. Каждый вектор $\vec{e}'_i$ разложим по векторам первого базиса:

$$\begin{aligned} \vec{e}'_1 &= a_{11}\vec{e}_1 + a_{21}\vec{e}_2 + \cdots + a_{n1}\vec{e}_n, \\ \vec{e}'_2 &= a_{12}\vec{e}_1 + a_{22}\vec{e}_2 + \cdots + a_{n2}\vec{e}_n, \\ &\dots\dots\dots \\ \vec{e}'_n &= a_{1n}\vec{e}_1 + a_{2n}\vec{e}_2 + \cdots + a_{nn}\vec{e}_n. \end{aligned} \tag{3.34}$$

Рассмотрим матрицу

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

Столбцы этой матрицы составлены из координат векторов второго базиса в первом базисе. Так как базисные векторы линейно независимы, матрица A – невырождена ($\det(A) \neq 0$).

Возьмем произвольный вектор $\vec{x}$ и обозначим через ξ_i его координаты в первом базисе, а через ξ'_i его координаты во втором базисе:

$$\vec{x} = \xi_1 \vec{e}_1 + \xi_2 \vec{e}_2 + \cdots + \xi_n \vec{e}_n; \quad (3.35)$$

$$\vec{x} = \xi'_1 \vec{e}_1 + \xi'_2 \vec{e}_2 + \cdots + \xi'_n \vec{e}_n. \quad (3.36)$$

Подставляя (3.34) в (3.36), получим:

[illegible]

Если перегруппировать слагаемые, то

[illegible]

Так как разложение по базису единственно, то, сравнивая последнее выражение с (3.35), получим:

$$\begin{aligned} \xi_1 &= a_{11}\xi'_1 + a_{12}\xi'_2 + \cdots + a_{1n}\xi'_n, \\ \xi_2 &= a_{21}\xi'_1 + a_{22}\xi'_2 + \cdots + a_{2n}\xi'_n, \\ . &. \\ \xi_n &= a_{n1}\xi'_1 + a_{n2}\xi'_2 + \cdots + a_{nn}\xi'_n. \end{aligned} \tag{3.37}$$

Если координаты вектора $\vec{x}$ в заданных базисах записать в виде матриц-столбцов:

$$\vec{\xi} = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix}, \quad \vec{\xi}' = \begin{bmatrix} \xi'_1 \\ \xi'_2 \\ \vdots \\ \xi'_n \end{bmatrix},$$

то равенства (3.37) можно переписать в матричной форме:

$$\vec{\xi} = A\vec{\xi}'. \quad (3.38)$$

Так как A – невырожденная матрица (существует A^{-1}), из (3.38) следует, что

$$\vec{\xi}' = A^{-1}\vec{\xi}.$$

Отметим, что столбцами матрицы A^{-1} являются координаты векторов первого базиса $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ во втором базисе $\vec{e}'_1, \vec{e}'_2, \dots, \vec{e}'_n$.

4. ЕВКЛИДОВЫ И УНИТАРНЫЕ ПРОСТРАНСТВА

4.1. Скалярное произведение в линейном пространстве

Рассмотрим комплексное линейное пространство $\mathbb{H}$.

Определение 4.1. Если каждой паре векторов $\vec{x}, \vec{y} \in \mathbb{H}$ поставлено в соответствие комплексное число $(\vec{x}, \vec{y})$, причем:

- 1) $(\vec{x}, \vec{x}) \in \mathbb{R}$, при этом $(\vec{x}, \vec{x}) \geq 0$ и $(\vec{x}, \vec{x}) = 0 \iff \vec{x} = \vec{0}$;
- 2) $(\vec{x}, \vec{y}) = \overline{(\vec{y}, \vec{x})}$;
- 3) $(\alpha\vec{x}, \vec{y}) = \alpha(\vec{x}, \vec{y})$ для любого числа $\alpha \in \mathbb{C}$;
- 4) $((\vec{x} + \vec{y}), \vec{z}) = (\vec{x}, \vec{z}) + (\vec{y}, \vec{z})$ для любых $\vec{x}, \vec{y}, \vec{z} \in \mathbb{H}$,

то число $(\vec{x}, \vec{y})$ называется скалярным произведением векторов $\vec{x}, \vec{y}$.

Свойства 1–4 принято называть аксиомами скалярного произведения. Вещественное число $(\vec{x}, \vec{x})$ обычно называют скалярным квадратом вектора $\vec{x}$.

Определение 4.2. Комплексное линейное пространство $\mathbb{H}$ со скалярным произведением называется унитарным.

Введем теперь скалярное произведение векторов в вещественном линейном пространстве $\mathbb{E}$.

Определение 4.3. Если каждой паре векторов $\vec{x}, \vec{y} \in \mathbb{E}$ поставлено в соответствие вещественное число $(\vec{x}, \vec{y})$, причем:

- 1) $(\vec{x}, \vec{x}) \geq 0$, при этом $(\vec{x}, \vec{x}) = 0 \iff \vec{x} = \vec{0}$;
- 2) $(\vec{x}, \vec{y}) = (\vec{y}, \vec{x})$;
- 3) $(\alpha\vec{x}, \vec{y}) = \alpha(\vec{x}, \vec{y})$ для любого числа $\alpha \in \mathbb{R}$;
- 4) $((\vec{x} + \vec{y}), \vec{z}) = (\vec{x}, \vec{z}) + (\vec{y}, \vec{z})$ для любых $\vec{x}, \vec{y}, \vec{z} \in \mathbb{E}$,

то число $(\vec{x}, \vec{y})$ называется скалярным произведением векторов $\vec{x}, \vec{y}$.

Подчеркнем, что вторая аксиома в определении 4.3 означает коммутативность скалярного произведения в вещественном линейном пространстве.

Определение 4.4. *Вещественное линейное пространство $\mathbb{E}$ со скалярным произведением называется евклидовым.*

Приведем простейшие следствия из аксиом скалярного произведения:

1. Для любых $\vec{x}, \vec{y} \in \mathbb{H}$ и любого числа $\alpha \in \mathbb{C}$

$$(\vec{x}, \alpha \vec{y}) = \overline{\alpha} (\vec{x}, \vec{y}). \quad (4.1)$$

Действительно,

$$(\vec{x}, \alpha \vec{y}) = \overline{(\alpha \vec{y}, \vec{x})} = \overline{\alpha (\vec{y}, \vec{x})} = \overline{\alpha} \overline{(\vec{y}, \vec{x})} = \overline{\alpha} (\vec{x}, \vec{y}).$$

2. Для любых $\vec{x}, \vec{y} \in \mathbb{H}$

$$(\vec{x}, (\vec{y} + \vec{z})) = (\vec{x}, \vec{y}) + (\vec{x}, \vec{z}). \quad (4.2)$$

Действительно,

$$(\vec{x}, (\vec{y} + \vec{z})) = \overline{((\vec{y} + \vec{z}), \vec{x})} = \overline{(\vec{y}, \vec{x}) + (\vec{z}, \vec{x})} = \overline{(\vec{y}, \vec{x})} + \overline{(\vec{z}, \vec{x})} = (\vec{x}, \vec{y}) + (\vec{x}, \vec{z}).$$

3. Каков бы ни был вектор $\vec{x} \in \mathbb{H}$, имеем

$$(\vec{x}, \vec{0}) = 0. \quad (4.3)$$

Действительно, $\vec{0} = 0\vec{x}$. Тогда $(\vec{x}, \vec{0}) = 0(\vec{x}, \vec{x}) = 0$.

В евклидовом пространстве свойство (4.1) превращается в

$$(\vec{x}, \alpha \vec{y}) = \alpha (\vec{x}, \vec{y}),$$

а свойства (4.2) и (4.3) сохраняются.

Из свойств (4.1)–(4.3) следует, что со скалярным произведением можно производить те же формальные преобразования, что и с обычным умножением. Единственное различие состоит в том, что в случае унитарного пространства коэффициент из второго множителя скалярного произведения выносится с комплексным сопряжением. Например,

$$\begin{aligned} (\alpha_1 \vec{x}_1 + \alpha_2 \vec{x}_2 + \alpha_3 \vec{x}_3, \beta_1 \vec{y}_1 + \beta_2 \vec{y}_2) &= \alpha_1 \overline{\beta_1} (\vec{x}_1, \vec{y}_1) + \alpha_1 \overline{\beta_2} (\vec{x}_1, \vec{y}_2) + \\ &+ \alpha_2 \overline{\beta_1} (\vec{x}_2, \vec{y}_1) + \alpha_2 \overline{\beta_2} (\vec{x}_2, \vec{y}_2) + \alpha_3 \overline{\beta_1} (\vec{x}_3, \vec{y}_1) + \alpha_3 \overline{\beta_2} (\vec{x}_3, \vec{y}_2). \end{aligned}$$

Теорема 4.1 (неравенство Коши—Буняковского). *Для любых двух векторов $\vec{x}, \vec{y}$ линейного пространства со скалярным произведением справедливо неравенство*

$$|(\vec{x}, \vec{y})| \leq \sqrt{(\vec{x}, \vec{x})} \sqrt{(\vec{y}, \vec{y})}.$$

Докажем теорему для случая унитарного пространства. Для евклидова пространства доказывается аналогично (см., например, [2] и [6]).

Доказательство. Если $\vec{y} = \vec{0}$, то неравенство очевидно. Предположим, что $\vec{y} \neq \vec{0}$, и обозначим $\lambda = \frac{(\vec{x}, \vec{y})}{(\vec{y}, \vec{y})}$. Рассмотрим скалярный квадрат вектора $\vec{x} - \lambda\vec{y}$:

$$\begin{aligned} ((\vec{x} - \lambda\vec{y}), (\vec{x} - \lambda\vec{y})) &= (\vec{x}, \vec{x}) - (\lambda\vec{y}, \vec{x}) - (\vec{x}, \lambda\vec{y}) + (\lambda\vec{y}, \lambda\vec{y}) = \\ &= (\vec{x}, \vec{x}) - \lambda\overline{(\vec{x}, \vec{y})} - \bar{\lambda}(\vec{x}, \vec{y}) + \lambda\bar{\lambda}(\vec{y}, \vec{y}) = \\ &= (\vec{x}, \vec{x}) - \frac{(\vec{x}, \vec{y})\overline{(\vec{x}, \vec{y})}}{(\vec{y}, \vec{y})} - \frac{\overline{(\vec{x}, \vec{y})}(\vec{x}, \vec{y})}{(\vec{y}, \vec{y})} + \frac{(\vec{x}, \vec{y})\overline{(\vec{x}, \vec{y})}(\vec{y}, \vec{y})}{(\vec{y}, \vec{y})^2} = \\ &= (\vec{x}, \vec{x}) - \frac{|(\vec{x}, \vec{y})|^2}{(\vec{y}, \vec{y})}. \end{aligned}$$

По первой аксиоме скалярный квадрат вектора неотрицателен, следовательно,

$$(\vec{x}, \vec{x}) - \frac{|(\vec{x}, \vec{y})|^2}{(\vec{y}, \vec{y})} \geq 0.$$

Отсюда $(\vec{x}, \vec{x})(\vec{y}, \vec{y}) - |(\vec{x}, \vec{y})|^2 \geq 0$ и $|(x, y)| \leq \sqrt{(x, x)}\sqrt{(y, y)}$. ■

В пространстве $\mathbb{C}^n$ скалярное произведение может быть введено сле-

дующим образом: если $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, $\vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$, то

$$(\vec{x}, \vec{y}) = x_1\overline{y_1} + x_2\overline{y_2} + \cdots + x_n\overline{y_n} = \sum_{i=1}^n x_i\overline{y_i}. \quad (4.4)$$

Замечание 4.1. Если векторы $\vec{x}$ и $\vec{y}$ рассматривать как матрицы-столбцы, то введенное скалярное произведение можно записать в виде про-

изведения матриц¹: $(\vec{x}, \vec{y}) = [\overline{y_1} \ \overline{y_2} \ \cdots \ \overline{y_n}] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \vec{y}^* \vec{x} \cdot \otimes$

Легко проверить выполнение всех аксиом скалярного произведения, например, первой аксиомы:

$$(\vec{x}, \vec{x}) = x_1\overline{x_1} + \cdots + x_n\overline{x_n} = |x_1|^2 + \cdots + |x_n|^2 \geq 0,$$

так как сумма модулей комплексных чисел всегда неотрицательна и обращается в нуль, только если все $x_i = 0$, $i = 1, \dots, n$, т. е. $\vec{x} = \vec{0}$.

¹ Напомним, что для матрицы A сопряженная матрица обозначается A^* (см. определение 1.7).

Аналогично, в вещественном пространстве $\mathbb{R}^n$ скалярное произведение можно ввести по формуле

$$(\vec{x}, \vec{y}) = x_1y_1 + x_2y_2 + \cdots + x_ny_n = \sum_{i=1}^n x_iy_i. \quad (4.5)$$

Введенные по формулам (4.4) и (4.5) скалярные произведения в $\mathbb{C}^n$ и $\mathbb{R}^n$ обычно называют стандартными.

Определение 4.5. Два унитарных (евклидовых) пространства $\mathbb{H}$ и $\mathbb{H}'$ называются изоморфными, если они изоморфны как линейные пространства и этот изоморфизм таков, что он сохраняет скалярное произведение соответствующих векторов: если $\vec{x} \longleftrightarrow \vec{x}'$ и $\vec{y} \longleftrightarrow \vec{y}'$, то $(\vec{x}, \vec{y}) = (\vec{x}', \vec{y}')$.

По аналогии с теоремой 3.6 можно доказать следующую теорему.

Теорема 4.2. Все унитарные (евклидовы) пространства одинаковой размерности изоморфны друг другу.

4.2. Норма в линейном пространстве

Рассмотрим линейное пространство $\mathbb{U}$.

Определение 4.6. Если каждому вектору $\vec{x} \in \mathbb{U}$ поставлено в соответствие вещественное число $\|\vec{x}\|$, причем:

- 1) $\|\vec{x}\| \geq 0$, при этом $\|\vec{x}\| = 0 \Leftrightarrow \vec{x} = \vec{0}$;
 - 2) $\|\alpha\vec{x}\| = |\alpha| \cdot \|\vec{x}\|$ для любого числа α ;
 - 3) $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$ для любых $\vec{x}, \vec{y} \in \mathbb{U}$,
- то это число $\|\vec{x}\|$ называется нормой вектора $\vec{x}$.

Свойства 1–3 принято называть аксиомами нормы, при этом третью аксиому часто называют неравенством треугольника.

Определение 4.7. Линейное пространство, в котором введена норма, называется нормированным пространством.

Если в линейном пространстве задано скалярное произведение, то в качестве нормы вектора $\|\vec{x}\|$ можно взять, например,

$$\|\vec{x}\| = \sqrt{(\vec{x}, \vec{x})}. \quad (4.6)$$

Данная норма называется нормой, порожденной скалярным произведением. Проверим выполнение аксиом для нормы, заданной формулой (4.6).

По первой аксиоме скалярного произведения $(\vec{x}, \vec{x}) \geq 0$ и $(\vec{x}, \vec{x}) = 0$, только если $\vec{x} = \vec{0}$. Следовательно, и $\|\vec{x}\| \geq 0$, причем равенство достигается только для нулевого вектора. Значит, первая аксиома нормы выполнена.

Вторая аксиома нормы выполнена в силу второй аксиомы скалярного произведения и формулы (4.1). Действительно,

$$\|\alpha\vec{x}\| = \sqrt{(\alpha\vec{x}, \alpha\vec{x})} = \sqrt{\alpha\bar{\alpha}(\vec{x}, \vec{x})} = \sqrt{|\alpha|^2(\vec{x}, \vec{x})} = |\alpha|\sqrt{(\vec{x}, \vec{x})} = |\alpha| \cdot \|\vec{x}\|.$$

Отметим, что, используя (4.6), неравенство Коши–Буняковского можно переписать в виде

$$|(\vec{x}, \vec{y})| \leq \|\vec{x}\| \|\vec{y}\|. \quad (4.7)$$

Для проверки справедливости третьей аксиомы найдем квадрат нормы суммы двух векторов:

$$\|\vec{x} + \vec{y}\|^2 = ((\vec{x} + \vec{y}), (\vec{x} + \vec{y})) = \|\vec{x}\|^2 + (\vec{x}, \vec{y}) + (\vec{y}, \vec{x}) + \|\vec{y}\|^2.$$

Так как модуль суммы комплексных чисел не превосходит суммы модулей слагаемых, а норма – неотрицательное число, то можно записать:

$$\begin{aligned} \|\vec{x} + \vec{y}\|^2 &= |(\vec{x} + \vec{y})|^2 = |\|\vec{x}\|^2 + (\vec{x}, \vec{y}) + (\vec{y}, \vec{x}) + \|\vec{y}\|^2| \leq \\ &\leq |\|\vec{x}\|^2| + |(\vec{x}, \vec{y})| + |(\vec{y}, \vec{x})| + |\|\vec{y}\|^2| = \|\vec{x}\|^2 + |(\vec{x}, \vec{y})| + |(\vec{y}, \vec{x})| + \|\vec{y}\|^2. \end{aligned}$$

Тогда, учитывая (4.7), получим:

$$\|\vec{x} + \vec{y}\|^2 \leq \|\vec{x}\|^2 + 2\|\vec{x}\|\|\vec{y}\| + \|\vec{y}\|^2 = (\|\vec{x}\| + \|\vec{y}\|)^2$$

и, извлекая квадратный корень из обеих частей неравенства, найдем:

$$\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|.$$

Таким образом, соотношение (4.6) действительно задает норму.

Если в пространстве $\mathbb{C}^n$ ввести стандартное скалярное произведение (формула (4.4)), то порожденная им норма имеет вид

$$\|\vec{x}\| = \sqrt{x_1\bar{x}_1 + x_2\bar{x}_2 + \dots + x_n\bar{x}_n} = \sqrt{|x_1|^2 + |x_2|^2 + \dots + |x_n|^2}.$$

Для стандартного скалярного произведения в $\mathbb{R}^n$

$$\|\vec{x}\| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}. \quad (4.8)$$

Следует отметить, что скалярное произведение в $\mathbb{C}^n$ можно задавать различным образом, и порождаемые различными скалярными произведениями нормы, в общем случае, будут различными. Кроме того, норму можно задавать вообще независимо от скалярного произведения. Например, в пространстве $\mathbb{C}^n$ часто используются следующие нормы:

$$\|\vec{x}\|_1 = |x_1| + |x_2| + \dots + |x_n|$$

и

$$\|\vec{x}\|_{\infty} = \max_{k=1, \dots, n} |x_k|,$$

которые не порождены никаким скалярным произведением.

Читателю предлагается самостоятельно проверить, что $\|\vec{x}\|_1$ и $\|\vec{x}\|_{\infty}$ действительно являются нормами в пространстве $\mathbb{C}^n$ (проверить выполнение трех аксиом нормы).

Несмотря на то, что нормы в конечномерном линейном пространстве $\mathbb{U}$ можно вводить различным образом, между ними имеется связь.

Определение 4.8. Нормы $\|\vec{x}\|_I$ и $\|\vec{x}\|_{II}$ в $\mathbb{U}$ называются эквивалентными, если существуют такие положительные числа α и β , что для всех векторов $\vec{x} \in \mathbb{U}$ выполнены неравенства

$$\alpha \|\vec{x}\|_I \leq \|\vec{x}\|_{II} \leq \beta \|\vec{x}\|_I.$$

Теорема 4.3. В конечномерном линейном пространстве любые две нормы эквивалентны.

Доказательство данной теоремы выходит за рамки настоящего издания, но его можно найти, например, в пособии [7, с. 44].

4.3. Матрица Грама

Определение 4.9. Если в унитарном пространстве задан набор векторов $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_m$, то квадратная матрица

$$G = G(\vec{x}_1, \dots, \vec{x}_m) = \begin{bmatrix} (\vec{x}_1, \vec{x}_1) & (\vec{x}_2, \vec{x}_1) & \cdots & (\vec{x}_m, \vec{x}_1) \\ (\vec{x}_1, \vec{x}_2) & (\vec{x}_2, \vec{x}_2) & \cdots & (\vec{x}_m, \vec{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ (\vec{x}_1, \vec{x}_m) & (\vec{x}_2, \vec{x}_m) & \cdots & (\vec{x}_m, \vec{x}_m) \end{bmatrix}$$

называется матрицей Грама.

Отметим, что матрица Грама G является самосопряженной (см. определение 1.11). Действительно, для любых i, j

$$g_{ij} = (\vec{x}_j, \vec{x}_i) = \overline{(\vec{x}_i, \vec{x}_j)} = \overline{g_{ji}}.$$

Теорема 4.4. Система векторов $\vec{x}_1, \dots, \vec{x}_m$ в унитарном пространстве линейно зависима тогда и только тогда, когда матрица Грама вырожденная, т. е. $\det(G(\vec{x}_1, \dots, \vec{x}_m)) = 0$.

Доказательство. $\Rightarrow$ Пусть $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_m$ – набор линейно зависимых векторов, т. е., согласно определению 3.3, существует набор чисел

$\alpha_1, \alpha_2, \dots, \alpha_m$ ($|\alpha_1| + |\alpha_2| + \dots + |\alpha_m| \neq 0$), удовлетворяющий векторному уравнению

$$\alpha_1 \vec{x}_1 + \alpha_2 \vec{x}_2 + \cdots + \alpha_m \vec{x}_m = \vec{0}.$$

Умножая это уравнение скалярно на $\vec{x}_i$, где $i = 1, \dots, m$, получим систему уравнений

[illegible]

Рассматривая эту систему уравнений относительно чисел $\alpha_1, \dots, \alpha_m$, видим, что однородная система имеет нетривиальные решения, следовательно, согласно следствию 2.5, матрица системы – а это и есть матрица Грама – вырожденная.

 Пусть $\det(G) = 0$. Рассмотрим однородную систему линейных уравнений

[illegible]

Согласно следствию 2.5 существует нетривиальное решение

$$\alpha_1^0, \alpha_2^0, \dots, \alpha_m^0 \quad (|\alpha_1^0| + |\alpha_2^0| + \dots + |\alpha_m^0| \neq 0)$$

этой системы. Подставим его в систему и перепишем полученные равенства в виде

[illegible]

Умножим первое равенство на $\overline{\alpha_1^0}$, второе – на $\overline{\alpha_2^0}$ и т. д. и затем сложим все равенства:

$$\begin{aligned} & \overline{\alpha_1^0}(\alpha_1^0\vec{x}_1 + \alpha_2^0\vec{x}_2 + \cdots + \alpha_m^0\vec{x}_m, \vec{x}_1) + \\ & + \overline{\alpha_2^0}(\alpha_1^0\vec{x}_1 + \alpha_2^0\vec{x}_2 + \cdots + \alpha_m^0\vec{x}_m, \vec{x}_2) + \\ & \dots\dots\dots \\ & + \overline{\alpha_m^0}(\alpha_1^0\vec{x}_1 + \alpha_2^0\vec{x}_2 + \cdots + \alpha_m^0\vec{x}_m, \vec{x}_m) = \\ & = (\alpha_1^0\vec{x}_1 + \alpha_2^0\vec{x}_2 + \cdots + \alpha_m^0\vec{x}_m, \alpha_1^0\vec{x}_1) + \\ & + (\alpha_1^0\vec{x}_1 + \alpha_2^0\vec{x}_2 + \cdots + \alpha_m^0\vec{x}_m, \alpha_2^0\vec{x}_2) + \\ & \dots\dots\dots \end{aligned}$$

$$\begin{aligned} &+ (\alpha_1^0 \vec{x}_1 + \alpha_2^0 \vec{x}_2 + \cdots + \alpha_m^0 \vec{x}_m, \alpha_m^0 \vec{x}_m) = \\ &= (\alpha_1^0 \vec{x}_1 + \alpha_2^0 \vec{x}_2 + \cdots + \alpha_m^0 \vec{x}_m, \alpha_1^0 \vec{x}_1 + \alpha_2^0 \vec{x}_2 + \cdots + \alpha_m^0 \vec{x}_m) = 0. \end{aligned}$$

По первой аксиоме скалярного произведения

$$\alpha_1^0 \vec{x}_1 + \alpha_2^0 \vec{x}_2 + \cdots + \alpha_m^0 \vec{x}_m = \vec{0},$$

следовательно, система векторов $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_m$ линейно зависима. ■

Следствие 4.1. Система векторов $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_m$ линейно независима тогда и только тогда, когда $\det(G) \neq 0$.

Пусть в унитарном пространстве $\mathbb{H}$ размерностью n задан базис

$$\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n.$$

Разложим произвольный вектор $\vec{x} \in \mathbb{H}$ по этому базису:

$$\alpha_1 \vec{e}_1 + \alpha_2 \vec{e}_2 + \cdots + \alpha_n \vec{e}_n = \vec{x}.$$

Так как $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ является базисом, то существует единственное решение этого уравнения $\alpha_1, \alpha_2, \dots, \alpha_n$. Для его нахождения последовательно скалярно умножим это уравнение на векторы $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ и получим систему линейных уравнений

[illegible]

Так как система векторов $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ линейно независима, то матрица Грама невырожденная и, следовательно, полученная система уравнений имеет единственное решение. Решив систему (4.9) (например, методом Гаусса—Жордана), найдем координаты вектора в базисе $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$.

Рассмотрим отдельно пространство $\mathbb{C}^n$. Пусть заданы векторы

$$\vec{x}_1, \vec{x}_2, \dots, \vec{x}_m.$$

Составим из них матрицу X размером $n \times m$:

$$X = [\vec{x}_1 \ \vec{x}_2 \ \cdots \ \vec{x}_m].$$

Тогда, учитывая замечание 4.1, матрицу Грама можно записать в виде

$$G = X^* X. \quad (4.10)$$

4.4. Ортогональный и ортонормированный базис

Рассмотрим унитарное пространство $\mathbb{H}$. Считаем, что в нем введена норма, порожденная скалярным произведением.

Определение 4.10. Векторы $\vec{x}, \vec{y} \in \mathbb{H}$ называются ортогональными, если $(\vec{x}, \vec{y}) = 0$.

Заметим, что нулевой вектор $\vec{0}$ ортогонален любому вектору, поэтому представляет интерес ортогональность ненулевых векторов. Заметим также, что если векторы $\vec{x}$ и $\vec{y}$ ортогональны, то для любых чисел λ и μ векторы $\lambda\vec{x}$ и $\mu\vec{y}$ также ортогональны:

$$(\lambda\vec{x}, \mu\vec{y}) = \lambda\bar{\mu}(\vec{x}, \vec{y}) = 0.$$

Определение 4.11. Ортогональной системой векторов называется набор попарно ортогональных ненулевых векторов, т. е. такой набор векторов $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_m$, что $(\vec{x}_k, \vec{x}_i) = 0$ при $k \neq i$ и $(\vec{x}_k, \vec{x}_k) \neq 0$.

Теорема 4.5. Ортогональная система векторов линейно независима.

Доказательство. Пусть $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_m$ — ортогональная система. Составим матрицу Грама для этой системы. Все диагональные элементы матрицы $(\vec{x}_k, \vec{x}_k)$ отличны от нуля, а остальные элементы — нули. Значит, определитель матрицы Грама (равный в этом случае произведению ее диагональных элементов) не равен нулю. Следовательно, согласно следствию 4.1, набор векторов $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_m$ линейно независим. ■

Определение 4.12. Система векторов называется ортонормированной, если она ортогональна и норма каждого вектора системы равна единице.

Отметим, что если задан ненулевой вектор $\vec{x}$, то вектор $\vec{e} = \frac{1}{\|\vec{x}\|}\vec{x}$ будет иметь норму, равную единице; действительно,

$$\|\vec{e}\| = \left\| \frac{1}{\|\vec{x}\|}\vec{x} \right\| = \frac{1}{\|\vec{x}\|}\|\vec{x}\| = 1.$$

Такая операция называется нормированием вектора, поэтому, если задана ортогональная система $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_m$, то система

$$\vec{e}_1 = \frac{1}{\|\vec{x}_1\|}\vec{x}_1, \vec{e}_2 = \frac{1}{\|\vec{x}_2\|}\vec{x}_2, \dots, \vec{e}_m = \frac{1}{\|\vec{x}_m\|}\vec{x}_m$$

будет ортонормированной.

$$\alpha_k = \frac{(\vec{x}, \vec{e}_k)}{(\vec{e}_k, \vec{e}_k)}, \quad k = 1, 2, \dots, n. \quad (4.11)$$

Если в качестве базиса $\vec{e}_1, \dots, \vec{e}_n$ выбрана ортонормированная система (ортонормированный базис), то коэффициенты Фурье (координаты вектора) приобретают вид

$$\alpha_k = (\vec{x}, \vec{e}_k), \quad k = 1, 2, \dots, n.$$

Отметим еще одно существенное свойство ортонормированных базисов. Пусть в унитарном пространстве $\mathbb{H}$ задан ортонормированный базис $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$. Возьмем 2 вектора $\vec{x}$ и $\vec{y}$ и разложим их по базису:

$$\vec{x} = \alpha_1 \vec{e}_1 + \alpha_2 \vec{e}_2 + \cdots + \alpha_n \vec{e}_n,$$

$$\vec{y} = \beta_1 \vec{e}_1 + \beta_2 \vec{e}_2 + \cdots + \beta_n \vec{e}_n.$$

[illegible]
$$(\vec{x}, \vec{y}) = \alpha_1 \overline{\beta_1} + \alpha_2 \overline{\beta_2} + \cdots + \alpha_n \overline{\beta_n}.$$

В любом ортонормированном базисе скалярное произведение векторов выражается через их координаты по той же формуле, что и стандартное скалярное произведение в пространстве $\mathbb{C}^n$.

Отметим, что стандартный базис пространства $\mathbb{C}^n$

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \dots, \quad \vec{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

ортонормированный.

Теорема 4.6 (Пифагора). Пусть задана ортогональная система векторов $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_k$, тогда

$$\|\vec{e}_1 + \vec{e}_2 + \dots + \vec{e}_k\|^2 = \|\vec{e}_1\|^2 + \|\vec{e}_2\|^2 + \dots + \|\vec{e}_k\|^2. \quad (4.12)$$

Доказательство. Используем метод математической индукции. Докажем (4.12) сначала для двух векторов: если $(\vec{e}_1, \vec{e}_2) = 0$, то

$$\|\vec{e}_1 + \vec{e}_2\|^2 = \|\vec{e}_1\|^2 + \|\vec{e}_2\|^2. \quad (4.13)$$

Действительно,

$$\|\vec{e}_1 + \vec{e}_2\|^2 = (\vec{e}_1 + \vec{e}_2, \vec{e}_1 + \vec{e}_2) = \|\vec{e}_1\|^2 + (\vec{e}_1, \vec{e}_2) + (\vec{e}_2, \vec{e}_1) + \|\vec{e}_2\|^2 = \|\vec{e}_1\|^2 + \|\vec{e}_2\|^2,$$

так как $(\vec{e}_1, \vec{e}_2) = (\vec{e}_2, \vec{e}_1) = 0$.

Предположим, что (4.12) выполнено для векторов $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_{k-1}$, и докажем справедливость (4.12) для векторов $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_k$. Введем обозначение: $\vec{a} = \vec{e}_1 + \vec{e}_2 + \dots + \vec{e}_{k-1}$. В силу индукционного предположения $\|\vec{a}\|^2 = \|\vec{e}_1\|^2 + \|\vec{e}_2\|^2 + \dots + \|\vec{e}_{k-1}\|^2$, а благодаря попарной ортогональности векторов $(\vec{a}, \vec{e}_k) = 0$. Используя доказанное равенство (4.13), получим:

$$\|\vec{e}_1 + \vec{e}_2 + \dots + \vec{e}_k\|^2 = \|\vec{a} + \vec{e}_k\|^2 = \|\vec{a}\|^2 + \|\vec{e}_k\|^2 = \|\vec{e}_1\|^2 + \|\vec{e}_2\|^2 + \dots + \|\vec{e}_k\|^2.$$

Итак, доказательство методом математической индукции закончено. ■

4.5. Алгоритм Грама–Шмидта

Рассмотрим процесс ортогонализации системы векторов, который называется алгоритмом Грама–Шмидта. Пусть дана линейно независимая система m векторов $\vec{f}_1, \dots, \vec{f}_m$. Построим m попарно ортогональных векторов $\vec{e}_k$, $k = 1, \dots, m$, таких, что $\vec{e}_k \in \mathcal{L}(\vec{f}_1, \dots, \vec{f}_k)$ (см. определение 3.9).

Примем $\vec{e}_1 = \vec{f}_1$. Вектор $\vec{e}_2$ будем искать в виде $\vec{e}_2 = \vec{f}_2 + \lambda \vec{e}_1$, при этом α подберем так, чтобы $(\vec{e}_1, \vec{e}_2) = 0$, т. е. $(\vec{f}_2 + \lambda \vec{e}_1, \vec{e}_1) = 0$. Тогда

$$\lambda = -\frac{(\vec{f}_2, \vec{e}_1)}{(\vec{e}_1, \vec{e}_1)}.$$

Вектор $\vec{e}_2$ отличен от нулевого, так как в противном случае система векторов $\vec{f}_2$ и $\vec{f}_1$ была бы линейно зависима.

Предположим, что попарно ортогональные и отличные от нуля векторы $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_{k-1}$ уже построены. Вектор $\vec{e}_k$ ищем в виде

$$\vec{e}_k = \vec{f}_k + \lambda_1 \vec{e}_1 + \dots + \lambda_{k-1} \vec{e}_{k-1}, \quad (4.14)$$

т.е. вектор $\vec{e}_k$ получается из вектора $\vec{f}_k$ „исправлением“ его с помощью линейной комбинации уже построенных векторов $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_{k-1}$.

Коэффициенты $\lambda_1, \lambda_2, \dots, \lambda_{k-1}$ в (4.14) находим из условия ортогональности вектора

$$\vec{e}_k = \vec{f}_k + \lambda_1 \vec{e}_1 + \dots + \lambda_{k-1} \vec{e}_{k-1}$$

к векторам $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_{k-1}$:

$$(\vec{f}_k + \lambda_1 \vec{e}_1 + \dots + \lambda_{k-1} \vec{e}_{k-1}, \vec{e}_1) = 0,$$

$$(\vec{f}_k + \lambda_1 \vec{e}_1 + \dots + \lambda_{k-1} \vec{e}_{k-1}, \vec{e}_2) = 0,$$

.....

$$(\vec{f}_k + \lambda_1 \vec{e}_1 + \dots + \lambda_{k-1} \vec{e}_{k-1}, \vec{e}_{k-1}) = 0.$$

Так как векторы $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_{k-1}$ попарно ортогональны, то эти равенства записываются в виде

$$(\vec{f}_k, \vec{e}_1) + \lambda_1 (\vec{e}_1, \vec{e}_1) = 0,$$

$$(\vec{f}_k, \vec{e}_2) + \lambda_2 (\vec{e}_2, \vec{e}_2) = 0,$$

.....

$$(\vec{f}_k, \vec{e}_{k-1}) + \lambda_{k-1} (\vec{e}_{k-1}, \vec{e}_{k-1}) = 0.$$

Отсюда

$$\lambda_1 = -\frac{(\vec{f}_k, \vec{e}_1)}{(\vec{e}_1, \vec{e}_1)}, \quad \lambda_2 = -\frac{(\vec{f}_k, \vec{e}_2)}{(\vec{e}_2, \vec{e}_2)}, \quad \dots, \quad \lambda_{k-1} = -\frac{(\vec{f}_k, \vec{e}_{k-1})}{(\vec{e}_{k-1}, \vec{e}_{k-1})}.$$

Для доказательства того, что построенный вектор $\vec{e}_k$ отличен от нулевого вектора, используем линейную независимость системы векторов

$$\vec{f}_1, \dots, \vec{f}_m.$$

Предварительно отметим, что вектор $\vec{e}_k$ есть линейная комбинация векторов $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_{k-1}, \vec{f}_k$. Но вектор $\vec{e}_{k-1}$ можно заменить линейной комбинацией вектора $\vec{f}_{k-1}$ и векторов $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_{k-2}$ и т.д. Окончательно получим, что вектор $\vec{e}_k$ записывается в виде

$$\vec{e}_k = a_1 \vec{f}_1 + a_2 \vec{f}_2 + \dots + a_{k-1} \vec{f}_{k-1} + \vec{f}_k, \quad (4.15)$$

т. е. $\vec{e}_k \in \mathcal{L}(\vec{f}_1, \dots, \vec{f}_k)$. Теперь ясно, что $\vec{e}_k \neq \vec{0}$. Действительно, в противном случае правая часть равенства (4.15) была бы нулевым вектором, что противоречит линейной независимости векторов $\vec{f}_1, \dots, \vec{f}_k$, так как коэффициент при $\vec{f}_k$ равен 1. Итак, доказано, что $\vec{e}_k \neq \vec{0}$.

Таким образом, по векторам $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_{k-1}$ и $\vec{f}_k$ построен вектор $\vec{e}_k$. Аналогично по $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_k$ и $\vec{f}_{k+1}$ построим $\vec{e}_{k+1}$ и т. д. Продолжая этот процесс до тех пор, пока не будут исчерпаны заданные векторы $\vec{f}_1, \dots, \vec{f}_m$, получим m отличных от нуля и попарно ортогональных векторов

$$\vec{e}_1, \vec{e}_2, \dots, \vec{e}_m.$$

Теорема 4.7. *Во всяком n -мерном пространстве существуют ортогональные базисы.*

Доказательство. По определению n -мерного пространства (см. определение 3.4), в нем существует линейно независимая система векторов (базис) $\vec{f}_1, \dots, \vec{f}_n$. С помощью процесса ортогонализации из него можно построить ортогональный базис $\vec{e}_1, \dots, \vec{e}_n$. ■

Если векторы $\vec{e}_k$ нормировать:

$$\vec{e}'_k = \frac{1}{\|\vec{e}_k\|} \vec{e}_k,$$

то базис $\vec{e}'_1, \vec{e}'_2, \dots, \vec{e}'_m$ будет, как нетрудно видеть, ортонормированным базисом. Отметим, что из данного базиса $\vec{f}_1, \dots, \vec{f}_n$ можно построить различные ортогональные базисы, если начинать процесс ортогонализации с разных векторов $\vec{f}_k$.

Пример 4.1. Рассмотрим в пространстве $\mathbb{R}^3$ базис

$$\vec{f}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{f}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{f}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

Построим ортогональный базис методом Грама–Шмидта. В качестве первого вектора возьмем $\vec{e}_1 = \vec{f}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, а $\vec{e}_2 = \vec{f}_2 + \lambda \vec{e}_1 = \begin{bmatrix} 1 + \lambda \\ 1 + \lambda \\ 1 \end{bmatrix}$, из условия ортогональности $0 = (\vec{e}_2, \vec{e}_1) = 2 + 3\lambda$ найдем $\lambda = -2/3$, следовательно, $\vec{e}_2 = \begin{bmatrix} 1/3 \\ 1/3 \\ -2/3 \end{bmatrix}$. Найдем третий вектор в виде

$$\vec{e}_3 = \vec{f}_3 + \lambda_1 \vec{e}_1 + \lambda_2 \vec{e}_2 = \begin{bmatrix} 1 + \lambda_1 + \lambda_2/3 \\ \lambda_1 + \lambda_2/3 \\ \lambda_1 - 2\lambda_2/3 \end{bmatrix},$$

запишем условия ортогональности в виде системы

$$\begin{cases} 0 = (\vec{e}_3, \vec{e}_1) = 1 + 3\lambda_1, \\ 0 = (\vec{e}_3, \vec{e}_2) = 1/3 + 2\lambda_2/3, \end{cases}$$

откуда $\lambda_1 = -1/3$ и $\lambda_2 = -1/2$, следовательно, $\vec{e}_3 = \begin{bmatrix} 1/2 \\ -1/2 \\ 0 \end{bmatrix}$. Итак, в результате получили ортогональный базис

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{e}_2 = \begin{bmatrix} 1/3 \\ 1/3 \\ -2/3 \end{bmatrix}, \quad \vec{e}_3 = \begin{bmatrix} 1/2 \\ -1/2 \\ 0 \end{bmatrix}.$$

Если эти векторы нормировать, то получится следующий ортонормированный базис:

$$\vec{e}_1 = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}, \quad \vec{e}_2 = \begin{bmatrix} 1/\sqrt{6} \\ 1/\sqrt{6} \\ -2/\sqrt{6} \end{bmatrix}, \quad \vec{e}_3 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{bmatrix}.$$

Предоставляем читателю самостоятельно убедиться в том, что если в качестве исходного базиса выбрать

$$\vec{h}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{h}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{h}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

(который отличается от рассмотренного только порядком векторов), то в результате процесса ортогонализации получится базис

$$\vec{g}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{g}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{g}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \bullet$$

4.6. Расстояние от точки до подпространства

Рассмотрим унитарное пространство $\mathbb{H}$ (по-прежнему считаем, что норма векторов в нем порождена скалярным произведением).

Определение 4.14. Если $\mathbb{H}_1$ – подпространство унитарного пространства $\mathbb{H}$, то будем говорить, что вектор $\vec{h}$ ортогонален подпространству $\mathbb{H}_1$, если он ортогонален любому вектору $\vec{x}$ из $\mathbb{H}_1$ (т. е. $(\vec{h}, \vec{x}) = 0$ для любых $\vec{x} \in \mathbb{H}_1$).

Если вектор $\vec{h}$ ортогонален векторам $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_m$, то он ортогонален любой их линейной комбинации. Действительно, из равенства $(\vec{h}, \vec{e}_i) = 0$ следует, что для любых чисел $\lambda_1, \lambda_2, \dots, \lambda_m$

$$(h, \lambda_1 \vec{e}_1 + \lambda_2 \vec{e}_2 + \dots + \lambda_m \vec{e}_m) = 0.$$

Следовательно, для того чтобы вектор $\vec{h}$ был ортогонален m -мерному подпространству $\mathbb{H}_1$, необходимо и достаточно, чтобы он был ортогонален m линейно независимым векторам из $\mathbb{H}_1$ (базису в $\mathbb{H}_1$).

Пусть в пространстве $\mathbb{H}$ заданы вектор $\vec{f} \in \mathbb{H}$ и m -мерное подпространство $\mathbb{H}_1$.

Определение 4.15. Вектор $\vec{f}_0 \in \mathbb{H}_1$ называется ортогональной проекцией вектора $\vec{f}$ на подпространство $\mathbb{H}_1$, если вектор $\vec{h} = \vec{f} - \vec{f}_0$ ортогонален $\mathbb{H}_1$.

Далее будет показано, что для любого вектора $\vec{f} \in \mathbb{H}$ его ортогональная проекция на произвольное подпространство $\mathbb{H}_1$ существует и единственна.

Теорема 4.8. Если $\vec{f}_0$ – ортогональная проекция вектора $\vec{f} \in \mathbb{H}$ на подпространство $\mathbb{H}_1$, то для любого вектора $\vec{f}_1 \in \mathbb{H}_1$, такого, что $\vec{f}_1 \neq \vec{f}_0$, выполнено неравенство

$$\|\vec{f} - \vec{f}_0\| < \|\vec{f} - \vec{f}_1\|.$$

Доказательство. Действительно, вектор $\vec{f}_0 - \vec{f}_1$ как разность двух векторов из $\mathbb{H}_1$ принадлежит $\mathbb{H}_1$ и, следовательно, ортогонален вектору $\vec{h} = \vec{f} - \vec{f}_0$. По теореме Пифагора (см. теорему 4.6) имеем:

$$\|\vec{f} - \vec{f}_0\|^2 + \|\vec{f}_0 - \vec{f}_1\|^2 = \|\vec{f} - \vec{f}_0 + \vec{f}_0 - \vec{f}_1\|^2 = \|\vec{f} - \vec{f}_1\|^2,$$

и поскольку $\|\vec{f}_0 - \vec{f}_1\| > 0$, то $\|\vec{f} - \vec{f}_0\| < \|\vec{f} - \vec{f}_1\|$. ■

Покажем теперь, как вычислить для $\vec{f}$ его ортогональную проекцию $\vec{f}_0$ на подпространство $\mathbb{H}_1$. Пусть базис подпространства $\mathbb{H}_1$ состоит из векторов $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_m$. Будем искать вектор $\vec{f}_0$ в виде

$$\vec{f}_0 = c_1 \vec{e}_1 + c_2 \vec{e}_2 + \dots + c_m \vec{e}_m, \quad (4.16)$$

где коэффициенты c_k найдем из условия ортогональности $\vec{f} - \vec{f}_0$ к $\mathbb{H}_1$. Для справедливости этого условия необходимо и достаточно, чтобы выполнялось m равенств $(\vec{f} - \vec{f}_0, \vec{e}_k) = 0$ ($k = 1, 2, \dots, m$), т. е.

$$(\vec{f}_0, \vec{e}_k) = (\vec{f}, \vec{e}_k).$$

Подставляя сюда вместо $\vec{f}_0$ его выражение (4.16), получаем систему m уравнений

$$c_1(\vec{e}_1, \vec{e}_k) + c_2(\vec{e}_2, \vec{e}_k) + \cdots + c_m(\vec{e}_m, \vec{e}_k) = (\vec{f}, \vec{e}_k), \quad k = 1, 2, \dots, m, \quad (4.17)$$

относительно чисел c_i , которая имеет единственное решение, так как матрица этой системы невырождена.

Рассмотрим отдельно часто встречающийся случай, когда базис $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_m$ – ортонормированный. В этом случае задача решается особенно просто. В таком базисе система (4.17) превращается в систему равенств

$$c_i = (\vec{f}, \vec{e}_i),$$

сразу определяющих нужные коэффициенты.

4.7. Псевдорешение системы линейных уравнений

Рассмотрим систему линейных уравнений

[illegible]

Введя следующие обозначения:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{C}^n, \quad \vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \in \mathbb{C}^m,$$

запишем систему (4.18) в матричной форме:

$$A\vec{x} = \vec{b}. \quad (4.19)$$

Возьмем произвольный элемент $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{C}^n$ и построим вектор

$$\vec{d}(\vec{x}) = \vec{b} - A\vec{x}. \quad (4.20)$$

Определение 4.16. Вектор $\vec{d}(\vec{x})$ вида (4.20) называется вектором невязки системы (4.19) (или просто невязкой системы (4.19)).

Определение 4.17. Псевдорешением системы (4.19) называется вектор $\vec{x}_0 \in \mathbb{C}^n$, который минимизирует норму невязки, т. е. для любых $\vec{x} \in \mathbb{C}^n$ выполняется неравенство

$$\|\vec{d}(\vec{x})\| \geq \|\vec{d}(\vec{x}_0)\|. \quad (4.21)$$

Заметим, что если система (4.19) совместна и $\vec{x}_0$ – ее решение, то это решение является также и псевдорешением, так как в этом случае $\vec{d}(\vec{x}_0) = \vec{0}$ и (4.21) очевидно выполнено.

Теорема 4.9. Любая система линейных уравнений имеет псевдорешение.

Доказательство. Рассмотрим столбцы матрицы A :

$$\vec{a}_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, \quad \vec{a}_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \quad \dots \quad \vec{a}_n = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

как элементы пространства $\mathbb{C}^m$. Линейная оболочка векторов $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$ – $\mathcal{L}(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n) = \mathbb{L}_1$ – представляет собой подпространство пространства $\mathbb{C}^m$ (см. утверждение 3.6).

В 4.6 было показано, что для любого вектора $\vec{b} \in \mathbb{C}^m$ существует единственный вектор $\vec{b}_0$ – ортогональная проекция вектора $\vec{b}$ на подпространство $\mathbb{L}_1$. Так как $\vec{b}_0$ – элемент линейной оболочки $\mathcal{L}(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n)$, то существует такой набор чисел $x_1^0, x_2^0, \dots, x_n^0$, что

$$\vec{b}_0 = x_1^0 \vec{a}_1 + x_2^0 \vec{a}_2 + \dots + x_n^0 \vec{a}_n. \quad (4.22)$$

Заметим, что если набор векторов $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$ линейно зависим, то представление $\vec{b}_0$ в виде линейной комбинации (4.22) не единственно. Выражение (4.22) можно переписать в матричном виде:

$$\vec{b}_0 = A\vec{x}_0, \quad \vec{x}_0 = \begin{bmatrix} x_1^0 \\ x_2^0 \\ \vdots \\ x_n^0 \end{bmatrix} \in \mathbb{C}^n.$$

Возьмем произвольный вектор $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{C}^n$, тогда вектор

$$\vec{y} = A\vec{x} = x_1 \vec{a}_1 + x_2 \vec{a}_2 + \dots + x_n \vec{a}_n \in \mathcal{L}(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n).$$

Пример 4.2 является частным случаем системы m уравнений с одним неизвестным

$$\begin{aligned}a_1x &= b_1, \\a_2x &= b_2, \\&\dots\dots\dots \\a_mx &= b_m,\end{aligned}$$

псевдорешение которой имеет вид

$$x = \frac{(\vec{b}, \vec{a}_1)}{(\vec{a}_1, \vec{a}_1)} = \frac{\sum_{k=1}^m a_k b_k}{\sum_{k=1}^m a_k^2}.$$

5. ЭЛЕМЕНТЫ АНАЛИТИЧЕСКОЙ ГЕОМЕТРИИ

Рассмотрим решение ряда геометрических задач средствами линейной алгебры. Переход от геометрических задач к алгебраическим осуществляется прежде всего введением какой-либо координатной системы. Ограничимся рассмотрением декартовой прямоугольной системы координат.

5.1. Пространство направленных отрезков (геометрических векторов)

Вначале в достаточно сжатой форме повторим хорошо известные из школьного курса геометрии факты, касающиеся направленных отрезков (геометрических векторов). Направленные отрезки будем рассматривать в окружающем нас пространстве, причем договоримся не путать это пространство с введенным ранее линейным пространством.

Обозначать направленные отрезки будем одной или двумя буквами со стрелкой наверху: $\vec{a}$ или $\overrightarrow{AB}$. Во втором случае предполагается, что точка A – начало, а точка B – конец направленного отрезка $\overrightarrow{AB}$. Будем считать, что выбрана некоторая единица измерения (масштаб) и, следовательно, каждый направленный отрезок $\vec{a}$ имеет определенную неотрицательную длину $|\vec{a}|$. Кроме того, предполагаем, что направленные отрезки не меняются при параллельном переносе, т. е. что все сонаправленные и одинаковые по длине направленные отрезки отождествляются. Направленный отрезок нулевой длины (его начало и конец совпадают) назовем нулевым направленным отрезком и будем обозначать $\vec{0}$. Его длина равна нулю, а направление для него не имеет смысла.

Определение 5.1. *Направленные отрезки $\vec{a}$ и $\vec{b}$, расположенные на одной прямой или на параллельных прямых, называются коллинеарными*

и обозначаются $\vec{a} \parallel \vec{b}$.

Определение 5.2. Направленные отрезки $\vec{a}$, $\vec{b}$ и $\vec{c}$, расположенные в одной плоскости или в параллельных плоскостях, называются *компланарными*, в противном случае – *некомпланарными* (рис. 5.1).

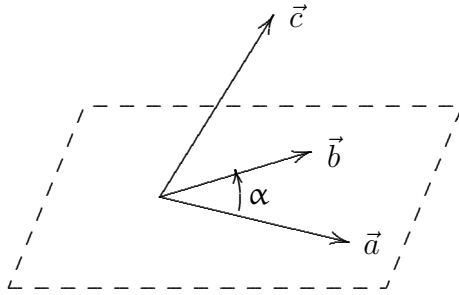


Рис. 5.1

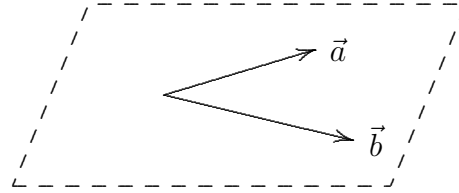


Рис. 5.2

Замечание 5.1. В силу предположения о неизменности направленных отрезков при параллельном переносе можно считать, что все направленные отрезки имеют общее начало в некоторой произвольно выбранной точке пространства. Тогда очевидно, что любые 2 направленных отрезка $\vec{a}$ и $\vec{b}$ всегда лежат в одной плоскости (рис. 5.2). В случае, когда направленные отрезки $\vec{a}$, $\vec{b}$ и $\vec{c}$ некомпланарны, любой из них не лежит в той плоскости, в которой расположены другие 2 направленных отрезка (рис. 5.1). $\otimes$

Ясно, что если направленные отрезки некомпланарны, то никакие 2 из них не могут быть коллинеарными и, конечно же, ни один из них не является нулевым.

Определение 5.3. Упорядоченная тройка некомпланарных направленных отрезков $\vec{a}$, $\vec{b}$, $\vec{c}$ называется *правой*, если из конца третьего направленного отрезка $\vec{c}$ кратчайший поворот (на какой-то угол α , $0 < \alpha < \pi$) от первого направленного отрезка $\vec{a}$ ко второму $\vec{b}$ виден как движение против часовой стрелки (рис. 5.1). В противном случае тройка называется *левой*.

Замечание 5.2. Нетрудно убедиться в том, что если направленные отрезки $\vec{a}$, $\vec{b}$, $\vec{c}$ образуют правую тройку, то $\vec{b}$, $\vec{c}$, $\vec{a}$ и $\vec{c}$, $\vec{a}$, $\vec{b}$ также образуют правые тройки, а векторы $\vec{a}$, $\vec{c}$, $\vec{b}$, векторы $\vec{b}$, $\vec{a}$, $\vec{c}$ и векторы $\vec{c}$, $\vec{b}$, $\vec{a}$ образуют левые тройки. $\otimes$

Определение 5.4. Если совместить начало направленного отрезка $\vec{b}$ с концом направленного отрезка $\vec{a}$, то направленный отрезок $\vec{c}$, идущий из начала $\vec{a}$ в конец $\vec{b}$, называют *суммой направленных отрезков $\vec{a}$ и $\vec{b}$* и

обозначают $\vec{c} = \vec{a} + \vec{b}$ (рис. 5.3).

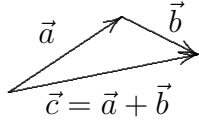


Рис. 5.3

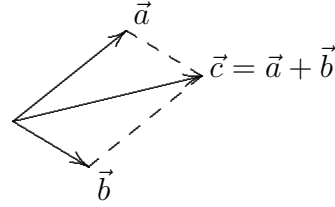


Рис. 5.4

Это правило сложения направленных отрезков часто называют правилом треугольника.

Если один из направленных отрезков нулевой (пусть для определенности $\vec{b} = \vec{0}$), то, конечно же, $\vec{a} + \vec{0} = \vec{a}$. Если $\vec{a} \neq 0$, $\vec{b} \neq 0$ и, кроме того, $\vec{a}$ и $\vec{b}$ неколлинеарны, то сумма $\vec{a} + \vec{b}$ может быть найдена по правилу параллелограмма: совмещаем начала направленных отрезков $\vec{a}$ и $\vec{b}$, строим на них как на смежных сторонах параллелограмм и находим сумму $\vec{a} + \vec{b}$ как диагональ этого параллелограмма, начало которой совпадает с началом $\vec{a}$ и $\vec{b}$ (рис. 5.4).

Определение 5.5. Произведением числа λ на $\vec{a}$ называется направленный отрезок $\vec{b}$ (пишут $\vec{b} = \lambda\vec{a}$), длина которого $|\vec{b}| = |\lambda| |\vec{a}|$, а направление при $\vec{a} \neq 0$ совпадает с направлением $\vec{a}$ при $\lambda > 0$ и противоположно $\vec{a}$ при $\lambda < 0$. При $\lambda = 0$ длина $\vec{b}$ равна нулю и вектор $\vec{b}$ является, таким образом, нулевым вектором.

Можно проверить, что введенные операции сложения направленных отрезков и умножения их на число обладают следующими свойствами:

- 1) $\vec{a} + \vec{b} = \vec{b} + \vec{a}$;
- 2) $\vec{a} + (\vec{b} + \vec{c}) = (\vec{a} + \vec{b}) + \vec{c}$;
- 3) $\lambda(\mu\vec{a}) = (\lambda\mu)\vec{a}$;
- 4) $(\lambda + \mu)\vec{a} = \lambda\vec{a} + \mu\vec{a}$;
- 5) $\lambda(\vec{a} + \vec{b}) = \lambda\vec{a} + \lambda\vec{b}$,

где λ и μ – вещественные числа.

Назовем направленный отрезок $(-1)\vec{a}$ противоположным направленному отрезку $\vec{a}$ и будем обозначать его $-\vec{a}$.

Определение 5.6. Разностью направленных отрезков $\vec{a}$ и $\vec{b}$ называют направленный отрезок $\vec{c} = \vec{a} + (-\vec{b})$. Обозначают его $\vec{c} = \vec{a} - \vec{b}$.

Используя правило сложения направленных отрезков и неизменность

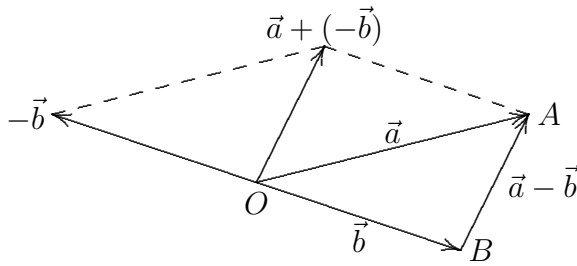


Рис. 5.5

направленных отрезков при параллельном переносе, легко понять, что разностью $\vec{a} - \vec{b}$ направленных отрезков $\vec{a}$ и $\vec{b}$ является направленный отрезок, соединяющий конец направленного отрезка $\vec{b}$ с концом направленного отрезка $\vec{a}$ (рис. 5.5).

Проверьте самостоятельно, что введенные операции сложения направленных отрезков и умножения направленного отрезка на число удовлетворяют всем аксиомам линейного пространства (см. определение 3.1), а значит, множество всех направленных отрезков является вещественным линейным пространством.

5.2. Изоморфизм пространства направленных отрезков и пространства $\mathbb{R}^3$

Векторам стандартного базиса в $\mathbb{R}^3$

$$\vec{i}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{i}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{i}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

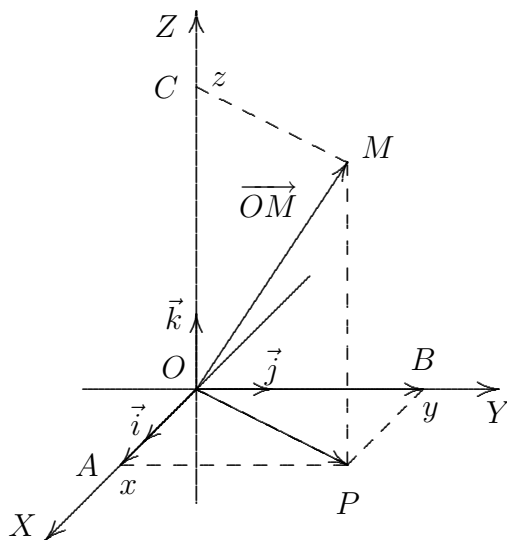


Рис. 5.6

поставим в соответствие какую-нибудь правую тройку попарно перпендикулярных направленных отрезков единичной длины $\vec{i}, \vec{j}, \vec{k}$ с общим началом в некоторой точке O :

$$\vec{i}_1 \longleftrightarrow \vec{i}, \quad \vec{i}_2 \longleftrightarrow \vec{j}, \quad \vec{i}_3 \longleftrightarrow \vec{k}.$$

Выберем декартову прямоугольную систему координат $OXYZ$ с началом координат в точке O (рис. 5.6), координатные оси OX , OY и OZ которой направлены так же, как направленные отрезки $\vec{i}, \vec{j}$ и $\vec{k}$ соответственно.

Считаем, что масштаб на координатных осях таков же, как выбранный ранее для измерения длин направленных отрезков. Кроме того, в силу

основного предположения о неизменности направленных отрезков при параллельном переносе, можно считать, что все направленные отрезки имеют начало в точке O и поэтому любой направленный отрезок $\overrightarrow{OM}$ однозначно определяется своей конечной точкой $M(x; y; z)$, где x, y, z – координаты точки M в выбранной системе координат. При этом, если A, B и C – проекции точки M на координатные оси OX, OY и OZ (рис. 5.6), то x, y и z – координаты точек A, B и C на соответствующих осях. Заметим, что точкам A, B и C соответствуют направленные отрезки $\overrightarrow{OA}, \overrightarrow{OB}$ и $\overrightarrow{OC}$, причем очевидны равенства

$$\overrightarrow{OA} = x\vec{i}, \quad \overrightarrow{OB} = y\vec{j}, \quad \overrightarrow{OC} = z\vec{k}. \quad (5.1)$$

Из правила сложения направленных отрезков следует, что

$$\overrightarrow{OM} = \overrightarrow{OP} + \overrightarrow{OC} = \overrightarrow{OA} + \overrightarrow{OB} + \overrightarrow{OC} \quad (\overrightarrow{OP} = \overrightarrow{OA} + \overrightarrow{OB}),$$

и с учетом (5.1) получаем:

$$\overrightarrow{OM} = x\vec{i} + y\vec{j} + z\vec{k}. \quad (5.2)$$

Итак, любой направленный отрезок $\overrightarrow{OM}$ представим в виде (5.2). Числа x, y и z называют координатами (проекциями) направленного отрезка $\overrightarrow{OM}$. С одной стороны, эти числа служат коэффициентами разложения (5.2) направленного отрезка $\overrightarrow{OM}$ по направленным отрезкам $\vec{i}, \vec{j}$ и $\vec{k}$, с другой стороны – это координаты точки M в декартовой прямоугольной системе координат, порожденной направленными отрезками $\vec{i}, \vec{j}$ и $\vec{k}$ (их называют координатными ортами).

Рассмотрим в $\mathbb{R}^3$ вектор

$$\vec{a} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = x\vec{i}_1 + y\vec{i}_2 + z\vec{i}_3. \quad (5.3)$$

Сравнение (5.2) и (5.3) позволяет сделать следующее утверждение.

Утверждение 5.1. Любому направленному отрезку $\overrightarrow{OM}$ соответствует вектор $\vec{a} \in \mathbb{R}^3$, координатами которого в стандартном базисе $\vec{i}_1, \vec{i}_2, \vec{i}_3$ служат проекции (координаты) направленного отрезка $\overrightarrow{OM}$ в некоторой фиксированной декартовой прямоугольной системе координат.

И наоборот, любому вектору $\vec{a} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3$ соответствует направленный отрезок $\overrightarrow{OM}$, определяемый формулой (5.2), где $\vec{i}, \vec{j}$ и $\vec{k}$ –

единичные направленные отрезки (координатные орты), лежащие на координатных осях выбранной системы координат.

Установленное соответствие является взаимно-однозначным, при этом:

$$1) \vec{i}_1 \longleftrightarrow \vec{i}, \vec{i}_2 \longleftrightarrow \vec{j}, \vec{i}_3 \longleftrightarrow \vec{k};$$

$$2) \vec{0} - \text{нулевой вектор в } \mathbb{R}^3;$$

$$3) \text{ если } \overrightarrow{OM} \longleftrightarrow \vec{a}, \text{ то } \lambda \overrightarrow{OM} \longleftrightarrow \lambda \vec{a} \text{ для вещественных чисел } \lambda;$$

$$4) \text{ если } \overrightarrow{OM_1} \longleftrightarrow \vec{a}_1 \text{ и } \overrightarrow{OM_2} \longleftrightarrow \vec{a}_2, \text{ то } \overrightarrow{OM_1} + \overrightarrow{OM_2} \longleftrightarrow \vec{a}_1 + \vec{a}_2.$$

Действительно, 1) справедливо по построению, свойство 2) – очевидно.

Докажем 3): если

$$\overrightarrow{OM} = x\vec{i} + y\vec{j} + z\vec{k} \longleftrightarrow \vec{a} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3,$$

то

$$\lambda \overrightarrow{OM} = \lambda x\vec{i} + \lambda y\vec{j} + \lambda z\vec{k} \longleftrightarrow \begin{bmatrix} \lambda x \\ \lambda y \\ \lambda z \end{bmatrix} = \lambda \vec{a}.$$

Докажем 4): если

$$\overrightarrow{OM_1} = x_1\vec{i} + y_1\vec{j} + z_1\vec{k} \longleftrightarrow \vec{a}_1 = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} \in \mathbb{R}^3 \text{ и}$$

$$\overrightarrow{OM_2} = x_2\vec{i} + y_2\vec{j} + z_2\vec{k} \longleftrightarrow \vec{a}_2 = \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} \in \mathbb{R}^3,$$

то

$$\begin{aligned} \overrightarrow{OM_1} + \overrightarrow{OM_2} &= (x_1 + x_2)\vec{i} + (y_1 + y_2)\vec{j} + (z_1 + z_2)\vec{k} \longleftrightarrow \\ &\longleftrightarrow \begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \\ z_1 + z_2 \end{bmatrix} = \vec{a}_1 + \vec{a}_2. \end{aligned}$$

Перечисленные свойства взаимно-однозначного соответствия между пространством направленных отрезков и пространством $\mathbb{R}^3$ показывают, что эти пространства изоморфны. Этот изоморфизм зависит, разумеется, от выбираемой системы координат (от координатных ортов $\vec{i}$, $\vec{j}$ и $\vec{k}$).

В дальнейшем будем считать, что декартова система координат выбрана и фиксирована, а потому можно отождествить любой направленный

отрезок $\overrightarrow{OM} = x\vec{i} + y\vec{j} + z\vec{k}$ с соответствующим ему вектором $\vec{a} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3$

и писать $\overrightarrow{OM} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$. В связи с этим впредь направленные отрезки будем

называть векторами и считать их элементами векторного пространства $\mathbb{R}^3$. При таком соглашении правомерно все предыдущие определения, построения и т. д., проводимые для векторов из $\mathbb{R}^3$, перенести на направленные отрезки. С другой стороны, некоторые геометрические факты, касающиеся направленных отрезков, можно перенести на векторы из $\mathbb{R}^3$.

В заключение отметим, что вектор $\overrightarrow{OM}$ часто называют радиусом-вектором точки M . В дальнейшем, если $\overrightarrow{OM}_1 = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}$ и $\overrightarrow{OM}_2 = \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix}$ – радиусы-векторы точек M_1 и M_2 , то $\overrightarrow{OM}_2 = \overrightarrow{OM}_1 + \overrightarrow{M_1M_2}$, т. е. $\overrightarrow{M_1M_2} = \overrightarrow{OM}_2 - \overrightarrow{OM}_1$ и, следовательно,

$$\overrightarrow{M_1M_2} = \begin{bmatrix} x_1 - x_2 \\ y_1 - y_2 \\ z_1 - z_2 \end{bmatrix}. \quad (5.4)$$

Отметим также, что длина направленного отрезка $\overrightarrow{OM} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ очевидно равна

$$|\overrightarrow{OM}| = \sqrt{x^2 + y^2 + z^2}, \quad (5.5)$$

и ей, таким образом, соответствует норма (4.8) вектора в $\mathbb{R}^3$. Поэтому иногда длину направленного отрезка обозначают так же, как и норму в $\mathbb{R}^3$: $\|\overrightarrow{OM}\|$. Наконец, из (5.4) и (5.5) следует, что расстояние между точками $M_1(x_1; y_1; z_1)$ и $M_2(x_2; y_2; z_2)$, очевидно равное $|\overrightarrow{M_1M_2}|$, вычисляется по формуле

$$|\overrightarrow{M_1M_2}| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}.$$

5.3. Скалярное произведение векторов в $\mathbb{R}^3$

В 4.1 было введено скалярное произведение для векторов из $\mathbb{R}^n$ (формула (4.5)). Для $\mathbb{R}^3$ эта формула имеет вид

$$(\vec{a}, \vec{b}) = a_1b_1 + a_2b_2 + a_3b_3, \quad (5.6)$$

где $\vec{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$ и $\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$.

Утверждение 5.2. Для любых двух ненулевых векторов $\vec{a}, \vec{b} \in \mathbb{R}^3$ определенное по формуле (5.6) скалярное произведение удовлетворяет соотношению

$$(\vec{a}, \vec{b}) = \|\vec{a}\| \|\vec{b}\| \cos \widehat{(\vec{a}, \vec{b})}, \quad (5.7)$$

где $\widehat{(\vec{a}, \vec{b})}$ – угол между векторами $\vec{a}$ и $\vec{b}$ ($0 \leq \widehat{(\vec{a}, \vec{b})} \leq \pi$).

Доказательство. Сначала докажем случай $0 < \widehat{(\vec{a}, \vec{b})} < \pi$. Рассмотрим треугольник OAB (см. рис. 5.5), стороны которого образуют векторы $\vec{a} = \overrightarrow{OA}$, $\vec{b} = \overrightarrow{OB}$ и $\vec{a} - \vec{b} = \overrightarrow{BA}$. По теореме косинусов

$$\|\vec{a} - \vec{b}\|^2 = \|\vec{a}\|^2 + \|\vec{b}\|^2 - 2\|\vec{a}\| \|\vec{b}\| \cos \widehat{(\vec{a}, \vec{b})}. \quad (5.8)$$

Квадрат длины вектора $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ составляет

$$\|\vec{x}\|^2 = (\vec{x}, \vec{x}) = x_1^2 + x_2^2 + x_3^2. \quad (5.9)$$

Пусть $\vec{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$, тогда $\vec{a} - \vec{b} = \begin{bmatrix} a_1 - b_1 \\ a_2 - b_2 \\ a_3 - b_3 \end{bmatrix}$. Из (5.8), учитывая (5.9), находим:

$$\begin{aligned} \|\vec{a}\| \|\vec{b}\| \cos \widehat{(\vec{a}, \vec{b})} &= \frac{1}{2}(\|\vec{a}\|^2 + \|\vec{b}\|^2 - \|\vec{a} - \vec{b}\|^2) = \\ &= \frac{1}{2}(a_1^2 + a_2^2 + a_3^2 + b_1^2 + b_2^2 + b_3^2 - [(a_1 - b_1)^2 + (a_2 - b_2)^2 + (a_3 - b_3)^2]) = \\ &= \frac{1}{2}(2a_1b_1 + 2a_2b_2 + 2a_3b_3) = a_1b_1 + a_2b_2 + a_3b_3 = (\vec{a}, \vec{b}). \end{aligned}$$

Рассмотрим теперь случаи $\widehat{(\vec{a}, \vec{b})} = 0$ и $\widehat{(\vec{a}, \vec{b})} = \pi$. Очевидно, что в этих случаях выполнено $\vec{b} = \lambda \vec{a}$, причем $\text{sign}(\lambda) = \cos(\widehat{(\vec{a}, \vec{b})})$. Если $\vec{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$, то

$$\vec{b} = \begin{bmatrix} \lambda a_1 \\ \lambda a_2 \\ \lambda a_3 \end{bmatrix}. \text{ Следовательно,}$$

$$(\vec{a}, \vec{b}) = \lambda a_1^2 + \lambda a_2^2 + \lambda a_3^2 = \lambda(a_1^2 + a_2^2 + a_3^2) = \lambda \|\vec{a}\|^2.$$

С другой стороны,

$$\begin{aligned} \|\vec{a}\| \|\vec{b}\| \cos(\widehat{(\vec{a}, \vec{b})}) &= \sqrt{a_1^2 + a_2^2 + a_3^2} \sqrt{\lambda^2 a_1^2 + \lambda^2 a_2^2 + \lambda^2 a_3^2} (\text{sign}(\lambda)) = \\ &= \|\vec{a}\| |\lambda| \|\vec{a}\| \text{sign}(\lambda) = \lambda \|\vec{a}\|^2. \blacksquare \end{aligned}$$

Замечание 5.3. Можно считать, что формула (5.7) справедлива и в случае, когда хотя бы один из векторов, $\vec{a}$ или $\vec{b}$, – нулевой вектор. В этом случае угол между векторами не определен, но обе части в (5.7) совпадают, так как равны нулю. $\otimes$

Из (5.7) следует, что если векторы $\vec{a}$ и $\vec{b}$ перпендикулярны (иногда говорят „ортогональны“, по аналогии с общим определением 4.10), то их скалярное произведение равно нулю. Для перпендикулярных векторов будем использовать обозначение $\vec{a} \perp \vec{b}$. Будем считать, что нулевой вектор $\vec{0}$ перпендикулярен любому вектору $\vec{a}$. Таким образом, условия $\vec{a} \perp \vec{b}$ и $(\vec{a}, \vec{b}) = 0$ являются равносильными.

Определение 5.7. Пусть даны 2 вектора $\vec{a}, \vec{b} \in \mathbb{R}^3$. Число, равное $\|\vec{b}\| \cos(\widehat{(\vec{a}, \vec{b})})$, называется проекцией вектора $\vec{b}$ на вектор $\vec{a}$ и обозначается $\text{Пр}_{\vec{a}} \vec{b}$.

Из этого определения следует, что в случае $\vec{a} \perp \vec{b}$ $\text{Пр}_{\vec{a}} \vec{b} = 0$. Если же векторы не перпендикулярны (и, следовательно, оба они ненулевые), то $\text{Пр}_{\vec{a}} \vec{b} > 0$, если угол между $\vec{a}$ и $\vec{b}$ острый, и $\text{Пр}_{\vec{a}} \vec{b} < 0$, если угол между $\vec{a}$ и $\vec{b}$ тупой (рис. 5.7).



Рис. 5.7

В первом случае $\text{Пр}_{\vec{a}} \vec{b} > 0$ и равна длине проекции отрезка $\vec{b}$ на прямую l , на которой лежит вектор $\vec{a}$, во втором случае $\text{Пр}_{\vec{a}} \vec{b} < 0$ и равна длине проекции отрезка $\vec{b}$ на прямую l , взятой со знаком “минус”.

Из определения 5.7 и формулы (5.7) следует, что

$$(\vec{a}, \vec{b}) = \|\vec{a}\| \text{Пр}_{\vec{a}} \vec{b} = \|\vec{b}\| \text{Пр}_{\vec{b}} \vec{a}.$$

5.4. Векторное произведение векторов в $\mathbb{R}^3$

Определение 5.8. Векторным произведением $\vec{a} \times \vec{b}$ векторов $\vec{a}$ и $\vec{b}$ из $\mathbb{R}^3$ называется вектор $\vec{c}$, такой, что:

- 1) вектор $\vec{c}$ ортогонален векторам $\vec{a}$ и $\vec{b}$;
- 2) $\|\vec{c}\| = \|\vec{a}\| \|\vec{b}\| \sin(\widehat{(\vec{a}, \vec{b})})$;
- 3) векторы $\vec{a}, \vec{b}, \vec{c}$ (если $\vec{a}$ и $\vec{b}$ неколлинеарны) образуют правую тройку.

Из п. 2 определения следует, что $\vec{a} \times \vec{b} = \vec{0} \Leftrightarrow \vec{a}$ и $\vec{b}$ коллинеарны, в частности, если хотя бы один из векторов, $\vec{a}$ или $\vec{b}$, нулевой. Заметим, что норма векторного произведения численно равна площади параллелограмма, построенного на векторах-сомножителях как на сторонах.

Утверждение 5.3. *Векторное произведение обладает следующими свойствами:*

- 1) *антикоммутативностью:* $\vec{a} \times \vec{b} = -(\vec{b} \times \vec{a})$;
- 2) *если λ – число, то $(\lambda \vec{a}) \times \vec{b} = \vec{a} \times (\lambda \vec{b}) = \lambda(\vec{a} \times \vec{b})$;*
- 3) *дистрибутивностью:*

$$(\vec{a} + \vec{c}) \times \vec{b} = \vec{a} \times \vec{b} + \vec{c} \times \vec{b} \quad (5.10)$$

и

$$\vec{a} \times (\vec{b} + \vec{c}) = \vec{a} \times \vec{b} + \vec{a} \times \vec{c}; \quad (5.11)$$

- 4) *для координатных ортов $\vec{i}, \vec{j}, \vec{k}$ справедливы соотношения:*

$$\begin{aligned} \vec{i} \times \vec{i} &= \vec{0}, & \vec{i} \times \vec{j} &= \vec{k}, & \vec{i} \times \vec{k} &= -\vec{j}, \\ \vec{j} \times \vec{i} &= -\vec{k}, & \vec{j} \times \vec{j} &= \vec{0}, & \vec{j} \times \vec{k} &= \vec{i}, \\ \vec{k} \times \vec{i} &= \vec{j}, & \vec{k} \times \vec{j} &= -\vec{i}, & \vec{k} \times \vec{k} &= \vec{0}. \end{aligned} \quad (5.12)$$

Доказательство. Докажем первое свойство. Согласно пп. 1 и 2 из определения 5.8, $\vec{a} \times \vec{b} \parallel \vec{b} \times \vec{a}$ и $\|\vec{a} \times \vec{b}\| = \|\vec{b} \times \vec{a}\|$, т. е.

$$\vec{a} \times \vec{b} = \pm(\vec{b} \times \vec{a}), \quad \vec{b} \times \vec{a} = \pm(\vec{a} \times \vec{b}) \quad (5.13)$$

(либо „+“, либо „–“). Но если $\vec{a}, \vec{b}, \vec{a} \times \vec{b}$ – правая тройка векторов (по определению 5.8), то $\vec{b}, \vec{a}, \vec{a} \times \vec{b}$ – левая, следовательно, $\vec{a} \times \vec{b} \neq \vec{b} \times \vec{a}$ и, учитывая (5.13), получаем: $\vec{a} \times \vec{b} = -(\vec{b} \times \vec{a})$.

Второе свойство при $\lambda = 0$ очевидно. Если $\lambda > 0$, то умножение $\vec{a}$ на λ увеличивает площадь параллелограмма в λ раз, а направление векторного произведения не меняется, следовательно,

$$(\lambda \vec{a}) \times \vec{b} = \lambda(\vec{a} \times \vec{b}).$$

Пусть $\lambda < 0$. Тогда, используя тот факт, что если $\vec{a}, \vec{b}, \vec{a} \times \vec{b}$ – правая тройка векторов, то $(-\vec{a}), \vec{b}, \vec{a} \times \vec{b}$ – левая тройка, получим:

$$(\lambda \vec{a}) \times \vec{b} = (-|\lambda| \vec{a}) \times \vec{b} = (|\lambda|(-\vec{a})) \times \vec{b} = |\lambda|((- \vec{a}) \times \vec{b}) = -|\lambda|(\vec{a} \times \vec{b}) = \lambda(\vec{a} \times \vec{b}).$$

Итак, для любых λ доказано

$$(\lambda \vec{a}) \times \vec{b} = \lambda(\vec{a} \times \vec{b}).$$

Если воспользоваться уже доказанным свойством антикоммутативности, то получим:

$$\vec{a} \times (\lambda \vec{b}) = -(\lambda \vec{b}) \times \vec{a} = -\lambda(\vec{b} \times \vec{a}) = \lambda(\vec{a} \times \vec{b}),$$

что завершает доказательство второго свойства.

Третье свойство будет доказано в 5.5.

Последнее, четвертое, свойство очевидно, если заметить, что площадь параллелограмма, построенного на единичных ортогональных векторах, равна единице, а векторы $\vec{i}$, $\vec{j}$, $\vec{k}$ образуют правую тройку. ■

Следствие 5.1. Для векторов $\vec{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$ и $\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$ их векторное

произведение можно записать в виде

$$\begin{aligned} \vec{a} \times \vec{b} &= \vec{i}(a_2b_3 - a_3b_2) - \vec{j}(a_1b_3 - a_3b_1) + \vec{k}(a_1b_2 - a_2b_1) = \\ &= \vec{i} \det \begin{bmatrix} a_2 & b_2 \\ a_3 & b_3 \end{bmatrix} - \vec{j} \det \begin{bmatrix} a_1 & b_1 \\ a_3 & b_3 \end{bmatrix} + \vec{k} \det \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}. \end{aligned} \quad (5.14)$$

Доказательство. Векторы $\vec{a}$ и $\vec{b}$ запишем в виде линейной комбинации векторов $\vec{i}$, $\vec{j}$, $\vec{k}$, тогда

$$\vec{a} \times \vec{b} = (a_1\vec{i} + a_2\vec{j} + a_3\vec{k}) \times (b_1\vec{i} + b_2\vec{j} + b_3\vec{k}).$$

Воспользовавшись свойствами 2 и 3 утверждения 5.3, получим:

$$\begin{aligned} \vec{a} \times \vec{b} &= a_1b_1(\vec{i} \times \vec{i}) + a_1b_2(\vec{i} \times \vec{j}) + a_1b_3(\vec{i} \times \vec{k}) + \\ &+ a_2b_1(\vec{j} \times \vec{i}) + a_2b_2(\vec{j} \times \vec{j}) + a_2b_3(\vec{j} \times \vec{k}) + \\ &+ a_3b_1(\vec{k} \times \vec{i}) + a_3b_2(\vec{k} \times \vec{j}) + a_3b_3(\vec{k} \times \vec{k}). \end{aligned}$$

Воспользовавшись формулами (5.12) получим равенство (5.14). ■

Отметим, что формулу (5.14) часто записывают в виде „символического“ определителя:

$$\vec{a} \times \vec{b} = \det \begin{bmatrix} \vec{i} & a_1 & b_1 \\ \vec{j} & a_2 & b_2 \\ \vec{k} & a_3 & b_3 \end{bmatrix},$$

формальное разложение которого по первому столбцу совпадает с (5.14).

5.5. Смешанное произведение векторов в $\mathbb{R}^3$

Определение 5.9. Скалярное произведение вектора $\vec{a}$ на векторное произведение векторов $\vec{b}$ и $\vec{c}$ $\lambda = (\vec{a}, \vec{b} \times \vec{c})$ называется смешанным произведением векторов $\vec{a}$, $\vec{b}$ и $\vec{c}$.

Очевидно, смешанное произведение трех векторов $\lambda = (\vec{a}, \vec{b} \times \vec{c})$ есть число.

Выясним геометрический смысл смешанного произведения. Заметим сначала, что если $\vec{a}$, $\vec{b}$ и $\vec{c}$ компланарны (лежат в одной плоскости), то $\vec{b} \times \vec{c}$ перпендикулярны плоскости, в которой лежат эти векторы, и, следовательно, $\vec{b} \times \vec{c}$ ортогонален вектору $\vec{a}$, откуда $(\vec{a}, \vec{b} \times \vec{c}) = 0$.

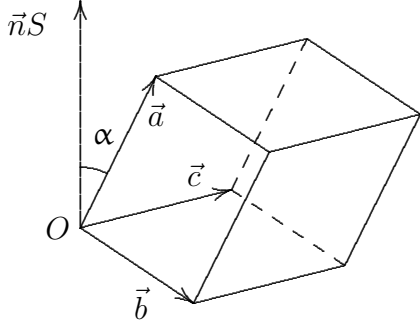


Рис. 5.8

Пусть $\vec{a}$, $\vec{b}$ и $\vec{c}$ – некопланарны. Построим на этих векторах как на ребрах (исходящих из одной вершины O) параллелепипед (рис. 5.8). Согласно определению 5.9 векторного произведения векторов $\vec{b}$ и $\vec{c}$, его длина $\|\vec{b} \times \vec{c}\| = S$ численно равна площади параллелограмма, построенного на векторах $\vec{b}$ и $\vec{c}$, т. е. площади основания параллелепипеда, следовательно, $\vec{b} \times \vec{c} = S\vec{n}$, где $\vec{n}$ – единичный вектор нормали к плоскости векторов $\vec{b}$ и $\vec{c}$, причем векторы $\vec{b}$, $\vec{c}$, $\vec{n}$ образуют правую тройку. Отсюда

$$(\vec{a}, \vec{b} \times \vec{c}) = \|\vec{b} \times \vec{c}\| \text{Пр}_{\vec{b} \times \vec{c}} \vec{a} = \|\vec{b} \times \vec{c}\| \text{Пр}_{\vec{n}} \vec{a} = S \text{Пр}_{\vec{n}} \vec{a} = S \|\vec{a}\| \cos(\widehat{\vec{a}, \vec{n}}).$$

Заметим, что $|\text{Пр}_{\vec{n}} \vec{a}| = h$ – высота параллелепипеда. Если $\vec{b}$, $\vec{c}$, $\vec{a}$ образуют правую тройку, то $\text{Пр}_{\vec{n}} \vec{a} = h$, и $\text{Пр}_{-\vec{n}} \vec{a} = -h$, если $\vec{b}$, $\vec{c}$, $\vec{a}$ образуют левую тройку. Если $\vec{b}$, $\vec{c}$, $\vec{a}$ образуют правую тройку, то угол $\alpha < \pi/2$, а если левую, то угол $\alpha > \pi/2$, следовательно,

$$|(\vec{a}, \vec{b} \times \vec{c})| = |S \text{Пр}_{\vec{n}} \vec{a}| = Sh = V,$$

где V – объем параллелепипеда.

Итак, смешанное произведение трех векторов равно объему параллелепипеда, построенного на этих векторах, если они образуют правую тройку, и равно объему параллелепипеда, взятому со знаком минус, если векторы образуют левую тройку.

Закончим теперь доказательство утверждения 5.3, т. е. дистрибутивности векторного произведения. Если $\vec{a}$, $\vec{b}$, $\vec{c}$ образуют правую тройку, то $V = (\vec{a}, \vec{b} \times \vec{c})$, где V – объем параллелепипеда. Очевидно, что векторы $\vec{c}$, $\vec{a}$, $\vec{b}$ тоже образуют правую тройку и построенный на них параллелепипед совпадает с параллелепипедом, построенным на векторах $\vec{a}$, $\vec{b}$, $\vec{c}$, следовательно, $V = (\vec{c}, \vec{a} \times \vec{b})$, значит,

$$(\vec{a}, \vec{b} \times \vec{c}) = (\vec{c}, \vec{a} \times \vec{b}). \quad (5.15)$$

Формула (5.15) также верна, если $\vec{a}$, $\vec{b}$, $\vec{c}$ образуют левую тройку. Пусть $\vec{x}$ – произвольный вектор. Тогда в силу справедливости (5.15) и линейности

скалярного произведения

$$\begin{aligned} (\vec{x}, \vec{a} \times (\vec{b} + \vec{c})) &= (\vec{x} \times \vec{a}, (\vec{b} + \vec{c})) = (\vec{x} \times \vec{a}, \vec{b}) + (\vec{x} \times \vec{a}, \vec{c}) = \\ &= (\vec{x}, \vec{a} \times \vec{b}) + (\vec{x}, \vec{a} \times \vec{c}) = (\vec{x}, (\vec{a} \times \vec{b} + \vec{a} \times \vec{c})). \end{aligned} \quad (5.16)$$

Из равенства $(\vec{x}, \vec{u}) = (\vec{x}, \vec{v})$, справедливого при любом $\vec{x}$, следует

$$(\vec{x}, \vec{u} - \vec{v}) = 0.$$

Если принять $\vec{x} = \vec{u} - \vec{v}$, то

$$(\vec{u} - \vec{v}, \vec{u} - \vec{v}) = \|\vec{u} - \vec{v}\|^2 = 0$$

и, следовательно, $\vec{u} = \vec{v}$. Значит, из (5.16) получим:

$$\vec{a} \times (\vec{b} + \vec{c}) = \vec{a} \times \vec{b} + \vec{a} \times \vec{c},$$

и формула дистрибутивности (5.11) доказана. Формула (5.10) получается из (5.11) двойным применением свойства антикоммутативности. ■

Найдем теперь формулу для вычисления смешанного произведения.

Пусть

$$\vec{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}, \quad \vec{c} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix},$$

тогда, согласно формуле (5.14),

$$\vec{b} \times \vec{c} = \begin{bmatrix} b_2 c_3 - b_3 c_2 \\ b_3 c_1 - b_1 c_3 \\ b_1 c_2 - b_2 c_1 \end{bmatrix}.$$

Отсюда для смешанного произведения получаем формулу

$$\begin{aligned} (\vec{a}, \vec{b} \times \vec{c}) &= a_1(b_2 c_3 - b_3 c_2) + a_2(b_3 c_1 - b_1 c_3) + a_3(b_1 c_2 - b_2 c_1) = \\ &= \det \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}. \end{aligned} \quad (5.17)$$

Поскольку из свойств определителя следует

$$\det \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} = \det \begin{bmatrix} b_1 & c_1 & a_1 \\ b_2 & c_2 & a_2 \\ b_3 & c_3 & a_3 \end{bmatrix} = \det \begin{bmatrix} c_1 & a_1 & b_1 \\ c_2 & a_2 & b_2 \\ c_3 & a_3 & b_3 \end{bmatrix},$$

то, учитывая (5.17) и свойство коммутативности скалярного произведения в евклидовом пространстве $\mathbb{R}^3$, получим:

$$(\vec{a}, \vec{b} \times \vec{c}) = (\vec{b}, \vec{c} \times \vec{a}) = (\vec{c}, \vec{a} \times \vec{b}) = (\vec{a} \times \vec{b}, \vec{c}).$$

Отсюда следует, что смешанное произведение не меняется при циклической перестановке векторов (рис. 5.9) и при перестановке знаков векторного

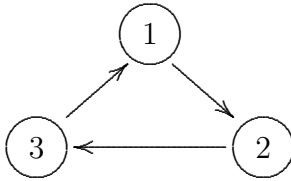


Рис. 5.9

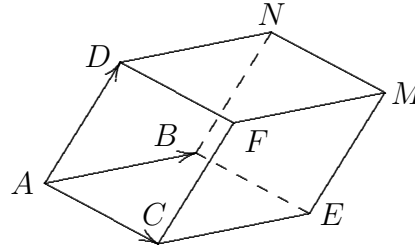


Рис. 5.10

и скалярного произведений, поэтому смешанное произведение трех векторов часто обозначают $\vec{a} \vec{b} \vec{c}$.

Пример 5.1. Найти объем пирамиды с вершинами, расположенными в точках $A(-1, 0, 2)$, $B(2, 1, 1)$, $C(3, 0, -1)$, $D(3, 2, 2)$. Построим векторы $\vec{AB} = \begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix}$, $\vec{AC} = \begin{bmatrix} 4 \\ 0 \\ -3 \end{bmatrix}$, $\vec{AD} = \begin{bmatrix} 4 \\ 2 \\ 0 \end{bmatrix}$ и на них построим параллелепипед $ACEBDFMN$ (рис. 5.10). Из школьного курса геометрии известно, что $V_{\text{пир}} = 1/6 V_{\text{пар}}$, следовательно, используя (5.17) получим:

$$\begin{aligned} V_{\text{пир}} &= \frac{1}{6} |\vec{AB} \vec{AC} \vec{AD}| = \frac{1}{6} \left| \det \begin{bmatrix} 3 & 4 & 4 \\ 1 & 0 & 2 \\ -1 & -3 & 0 \end{bmatrix} \right| = \\ &= \frac{1}{6} \left| \det \begin{bmatrix} 3 & 4 & -2 \\ 1 & 0 & 0 \\ -1 & -3 & 2 \end{bmatrix} \right| = \frac{1}{6} |(-1)(8 - 6)| = \frac{2}{6} = \frac{1}{3}. \bullet \end{aligned}$$

5.6. Уравнение плоскости

Будем считать, что уравнение плоскости задает некоторое множество точек, если координаты каждой точки этого множества удовлетворяют этому уравнению, а координаты любой точки, не принадлежащей этому множеству, уравнению не удовлетворяют. При этом, естественно, предполагается, что фиксирована некоторая система координат.

Рассмотрим в пространстве произвольную плоскость Π . Ее положение в пространстве (считаем, что фиксирована некоторая декартова прямоугольная система координат) однозначно определяется какой-либо точкой $M_0(x_0; y_0; z_0)$, принадлежащей этой плоскости, и нормальным векто-

ром $\vec{n} = \begin{bmatrix} A \\ B \\ C \end{bmatrix}$ (ненулевым вектором, перпендикулярным плоскости). Возь-

лем произвольную точку $M(x; y; z)$, лежащую в плоскости Π . Тогда вектор $\overrightarrow{M_0M}$ лежит в этой плоскости, и потому $\overrightarrow{M_0M} \perp \vec{n}$, т. е.

$$(\vec{n}, \overrightarrow{M_0M}) = 0. \quad (5.18)$$

Если точка M не лежит в плоскости Π , то вектор $\overrightarrow{M_0M}$ не перпендикулярен вектору $\vec{n}$, и тогда $(\vec{n}, \overrightarrow{M_0M}) \neq 0$. Таким образом, (5.18) является векторным уравнением рассматриваемой плоскости.

Запишем это уравнение в координатной форме. Поскольку

$$\overrightarrow{M_0M} = \begin{bmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{bmatrix},$$

уравнение (5.18) с учетом (5.6) принимает вид

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0. \quad (5.19)$$

Раскрывая в (5.19) скобки и обозначая $-Ax_0 - By_0 - Cz_0 = D$, можем записать это уравнение в виде

$$Ax + By + Cz + D = 0 \quad (A^2 + B^2 + C^2 > 0). \quad (5.20)$$

Уравнение вида (5.20) называют общим уравнением плоскости.

Таким образом, любая плоскость всегда может быть задана уравнением вида (5.20). Покажем теперь, что, в свою очередь, уравнение (5.20) всегда задает некоторую плоскость.

Возьмем произвольное уравнение (5.20). Найдём какое-нибудь решение $(x_0; y_0; z_0)$ этого уравнения (это уравнение имеет бесконечно много решений, что следует из общей теории систем линейных уравнений) – получим справедливое числовое равенство

$$Ax_0 + By_0 + Cz_0 + D = 0,$$

вычитая которое из уравнения (5.20), получим равносильное ему уравнение

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0.$$

Получили уравнение вида (5.19), которое, как было показано, задает плоскость, проходящую через точку $(x_0; y_0; z_0)$ с нормальным вектором $\begin{bmatrix} A \\ B \\ C \end{bmatrix}$.

Пример 5.2. Выведем уравнение плоскости, проходящей через 3 заданные точки $M_1(x_1; y_1; z_1)$, $M_2(x_2; y_2; z_2)$ и $M_3(x_3; y_3; z_3)$, не лежащие на одной прямой. Пусть $M(x; y; z)$ – произвольная точка этой плоскости. Тогда 3 вектора $\overrightarrow{M_1M}$, $\overrightarrow{M_1M_2}$ и $\overrightarrow{M_1M_3}$ лежат в данной плоскости, т. е. являют-

ся компланарными (рис. 5.11), и поэтому их смешанное произведение равно нулю:

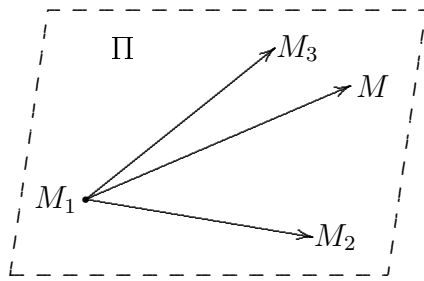


Рис. 5.11

$$\overrightarrow{M_1M} \overrightarrow{M_1M_2} \overrightarrow{M_1M_3} = 0. \quad (5.21)$$

Если точка M не лежит в рассматриваемой плоскости, то, разумеется, векторы $\overrightarrow{M_1M}$, $\overrightarrow{M_1M_2}$, $\overrightarrow{M_1M_3}$ не являются компланарными и условие (5.21) не выполняется. Таким образом, (5.21) является векторным уравнением плоскости, проходящей

через 3 точки. Запишем это уравнение в координатной форме. Имеем:

$$\overrightarrow{M_1M} = \begin{bmatrix} x - x_1 \\ y - y_1 \\ z - z_1 \end{bmatrix}, \quad \overrightarrow{M_1M_2} = \begin{bmatrix} x_2 - x_1 \\ y_2 - y_1 \\ z_2 - z_1 \end{bmatrix}, \quad \overrightarrow{M_1M_3} = \begin{bmatrix} x_3 - x_1 \\ y_3 - y_1 \\ z_3 - z_1 \end{bmatrix}$$

и с учетом (5.18) получаем уравнение вида

$$\det \begin{bmatrix} x - x_1 & x_2 - x_1 & x_3 - x_1 \\ y - y_1 & y_2 - y_1 & y_3 - y_1 \\ z - z_1 & z_2 - z_1 & z_3 - z_1 \end{bmatrix} = 0. \quad (5.22)$$

Заметим, что можно в уравнении (5.22) найти определитель матрицы, разложив его, например, по первой строке, и тогда получим общее уравнение плоскости вида (5.20). Однако чаще уравнение плоскости, проходящей через 3 заданные точки, записывают в форме (5.22), удобной для запоминания. •

Пример 5.3. Найти уравнение плоскости, проходящей через 3 точки,

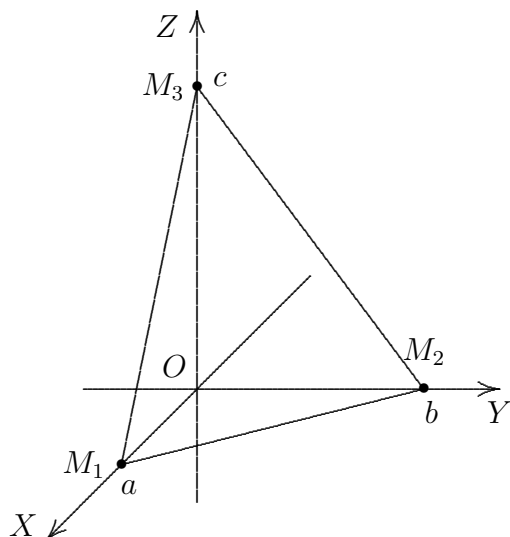


Рис. 5.12

лежащие на координатных осях OX , OY и OZ и имеющие на этих осях координаты a , b и c соответственно (предполагается, что $abc \neq 0$, т.е. что плоскость не проходит через начало координат (рис. 5.12)). Рассматриваемая плоскость проходит, таким образом, через 3 точки: $M_1(a; 0; 0)$, $M_2(0; b; 0)$ и $M_3(0; 0; c)$. Подставив координаты этих точек в уравнение (5.22), получим:

$$\det \begin{bmatrix} x - a & 0 - a & 0 - a \\ y - 0 & b - 0 & 0 - 0 \\ z - 0 & 0 - 0 & c - 0 \end{bmatrix} = 0.$$

Разложим определитель по первому столбцу:

$$(x - a)bc + yac + zab = 0.$$

Отсюда

$$bcx + acy + abz = abc.$$

Деля это уравнение на $abc \neq 0$, получим:

$$\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1. \quad (5.23)$$

Уравнение (5.23) называется уравнением плоскости в отрезках на осях. •

Пусть заданы две плоскости Π_1 и Π_2 :

$$\begin{aligned} A_1x + B_1y + C_1z + D_1 &= 0, \\ A_2x + B_2y + C_2z + D_2 &= 0. \end{aligned}$$

Обозначим через $\vec{n}_1 = \begin{bmatrix} A_1 \\ B_1 \\ C_1 \end{bmatrix}$, $\vec{n}_2 = \begin{bmatrix} A_2 \\ B_2 \\ C_2 \end{bmatrix}$ нормальные векторы плоскостей Π_1 и Π_2 соответственно, а через α – угол между этими векторами: $\alpha = (\widehat{\vec{n}_1, \vec{n}_2})$.

Определение 5.10. Углом $(\widehat{\Pi_1, \Pi_2})$ между плоскостями Π_1 и Π_2 называется угол α , если $0 \leq \alpha \leq \frac{\pi}{2}$, либо угол $(\pi - \alpha)$, если $\frac{\pi}{2} < \alpha \leq \pi$.

Согласно этому определению угол $(\widehat{\Pi_1, \Pi_2})$ всегда принадлежит отрезку $\left[0; \frac{\pi}{2}\right]$, при этом

$$\begin{aligned} \cos(\widehat{\Pi_1, \Pi_2}) &= \left| \cos(\widehat{\vec{n}_1, \vec{n}_2}) \right| = \frac{|(\vec{n}_1, \vec{n}_2)|}{\|\vec{n}_1\| \|\vec{n}_2\|} = \\ &= \frac{|A_1A_2 + B_1B_2 + C_1C_2|}{\sqrt{A_1^2 + B_1^2 + C_1^2} \sqrt{A_2^2 + B_2^2 + C_2^2}}. \end{aligned}$$

Отметим частные случаи взаимного расположения двух плоскостей:

1. Плоскости Π_1 и Π_2 перпендикулярны друг другу, если

$$\begin{aligned} (\widehat{\Pi_1, \Pi_2}) &= \frac{\pi}{2} \iff (\vec{n}_1, \vec{n}_2) = 0 \iff \\ &\iff A_1A_2 + B_1B_2 + C_1C_2 = 0. \end{aligned}$$

2. Плоскости Π_1 и Π_2 параллельны, если их нормальные векторы $\vec{n}_1$ и $\vec{n}_2$ коллинеарны, т. е. когда

$$\frac{A_1}{A_2} = \frac{B_1}{B_2} = \frac{C_1}{C_2}$$

(если в этом равенстве какой-то из знаменателей равен нулю, то равен нулю и соответствующий числитель).

3. В случае

$$\frac{A_1}{A_2} = \frac{B_1}{B_2} = \frac{C_1}{C_2} = \frac{D_1}{D_2}$$

плоскости Π_1 и Π_2 очевидно совпадают.

5.7. Уравнение прямой

Из школьного курса геометрии известно, что пересечением двух не совпадающих и не параллельных плоскостей является прямая. Верно и обратное утверждение: любая прямая в пространстве принадлежит по крайней мере двум различным плоскостям. Следовательно, прямая в пространстве может быть задана системой двух линейных уравнений с тремя неизвестными:

$$\begin{cases} A_1x + B_1y + C_1z + D_1 = 0, \\ A_2x + B_2y + C_2z + D_2 = 0. \end{cases} \quad (5.24)$$

Например, система $\begin{cases} x = 0, \\ y = 0 \end{cases}$ задает прямую – ось аппликат OZ .

Прямую L в пространстве можно задать и другими способами. Напри-

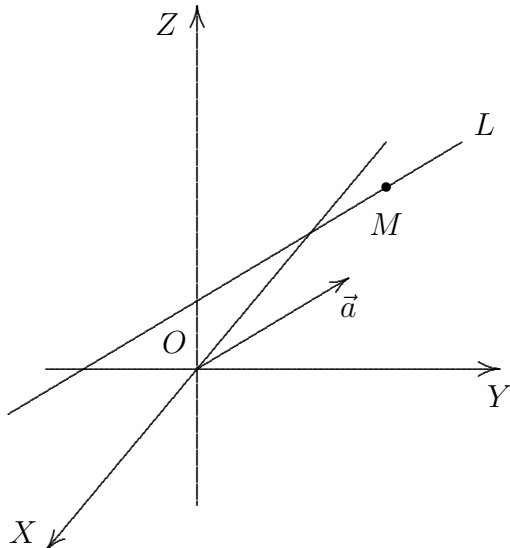


Рис. 5.13

мер, можно задать одну точку $M_0(x_0, y_0, z_0)$ на данной прямой и на-

правляющий вектор $\vec{a} = \begin{bmatrix} m \\ n \\ p \end{bmatrix}$ – лю-

бой вектор, параллельный данной прямой (рис. 5.13). Пусть $M(x, y, z)$ – произвольная точка на прямой L , тогда ясно, что направленный отрезок $\overrightarrow{M_0M}$ коллинеарен $\vec{a}$, т.е. найдется такое вещественное число t , что $\overrightarrow{M_0M} = t\vec{a}$. Верно и обратное: если $\overrightarrow{M_0M} = t\vec{a}$, то $M(x, y, z)$ лежит на данной прямой. Таким образом,

$$\overrightarrow{M_0M} = t\vec{a}, \quad t \in \mathbb{R} \quad (5.25)$$

– уравнение данной прямой. Запишем (5.25) в координатах:

$$\begin{cases} x - x_0 = tm, \\ y - y_0 = tn, \\ z - z_0 = tp, \end{cases} \quad \text{или} \quad \begin{cases} x = x_0 + tm, \\ y = y_0 + tn, \\ z = z_0 + tp, \end{cases} \quad t \in \mathbb{R}. \quad (5.26)$$

Уравнения (5.26) называют параметрическими уравнениями прямой в пространстве. Исключая из уравнений (5.26) параметр t , получаем:

$$\frac{x - x_0}{m} = \frac{y - y_0}{n} = \frac{z - z_0}{p}. \quad (5.27)$$

Уравнения (5.27) называют каноническими уравнениями прямой в пространстве. Заметим, что если одна из координат вектора $\vec{a}$ равна нулю, например, если $m = 0$, то $x - x_0 = 0$ или $x = x_0$, и нуль в знаменателе дроби в (5.27) означает, таким образом, что и числитель равен нулю.

Если прямая задана системой уравнений (5.24), то параметрическое уравнение (5.25) легко получить следующим образом. В качестве направ-

ляющего вектора прямой можно взять $\vec{a} = \vec{n}_1 \times \vec{n}_2$, где $\vec{n}_1 = \begin{bmatrix} A_1 \\ B_1 \\ C_1 \end{bmatrix}$ и

$\vec{n}_2 = \begin{bmatrix} A_2 \\ B_2 \\ C_2 \end{bmatrix}$, а в качестве координат точки $M_0(x_0, y_0, z_0)$ – любое решение

системы (5.24).

Пример 5.4. Найти уравнение прямой, проходящей через две точки $M_1(x_1, y_1, z_1)$ и $M_2(x_2, y_2, z_2)$. Ясно, что в данном случае в качестве M_0 можно взять любую из этих точек, например M_1 , а в качестве направляющего вектора прямой можно взять $\vec{a} = \overrightarrow{M_1M_2}$, т. е.

$$\vec{a} = \overrightarrow{M_1M_2} = \begin{bmatrix} x_2 - x_1 \\ y_2 - y_1 \\ z_2 - z_1 \end{bmatrix} = \begin{bmatrix} m \\ n \\ p \end{bmatrix}. \quad (5.28)$$

Подставляя (5.28) в (5.26), получаем:

$$\begin{cases} x = x_1 + t(x_2 - x_1), \\ y = y_1 + t(y_2 - y_1), \\ z = z_1 + t(z_2 - z_1), \end{cases} \quad t \in \mathbb{R}, \quad (5.29)$$

а уравнение (5.27) принимает вид

$$\frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1} = \frac{z - z_1}{z_2 - z_1}. \quad (5.30)$$

Итак, получили (5.29) – параметрическое уравнение прямой и (5.30) – каноническое уравнение прямой, проходящей через две точки. •

Пример 5.5. Выведем уравнение отрезка. Пусть заданы две точки $M_1(x_1, y_1, z_1)$ и $M_2(x_2, y_2, z_2)$ (рис. 5.14). Возьмем произвольную точку

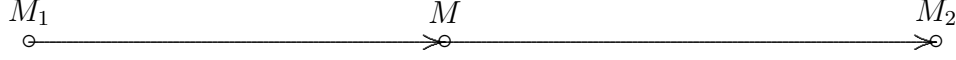


Рис. 5.14

$M(x, y, z)$ на отрезке M_1M_2 . Векторы $\overrightarrow{M_1M_2}$ и $\overrightarrow{M_1M}$ коллинеарны и сонаправлены, причем $\|\overrightarrow{M_1M}\| \leq \|\overrightarrow{M_1M_2}\|$, следовательно, существует $t \in [0, 1]$, такое, что

$$\overrightarrow{M_1M} = t\overrightarrow{M_1M_2}, \quad t \in [0, 1].$$

Это и есть векторное уравнение отрезка M_1M_2 . Записывая его в координатной форме, получим:

$$\begin{cases} x = x_1 + t(x_2 - x_1), \\ y = y_1 + t(y_2 - y_1), \\ z = z_1 + t(z_2 - z_1), \end{cases} \quad t \in [0, 1]. \quad (5.31)$$

Нетрудно видеть, что уравнения (5.31) полностью совпадают с уравнениями (5.29). Это естественно, так как отрезок M_1M_2 является частью прямой, проходящей через его конечные точки M_1 и M_2 . При этом, если на прямой параметр t принадлежит $\mathbb{R}$, то для отрезка этот параметр принимает значения из $[0, 1]$. •

Систему уравнений (5.31) часто записывают в виде

$$\begin{cases} x = (1 - t)x_1 + tx_2, \\ y = (1 - t)y_1 + ty_2, \\ z = (1 - t)z_1 + tz_2, \end{cases} \quad t \in [0, 1]. \quad (5.32)$$

При $t = 1/2$ формулы (5.32) дают координаты середины отрезка M_1M_2 :

$$x = \frac{x_1 + x_2}{2}, \quad y = \frac{y_1 + y_2}{2}, \quad z = \frac{z_1 + z_2}{2}.$$

Список литературы

1. Белопольский А. Л., Бодунов Н. А., Меркулов А. И. Линейная алгебра: Учеб. пособие. – СПб.: Изд-во СПбГЭТУ „ЛЭТИ“, 2000. – 152 с.
2. Гельфанд И. М. Лекции по линейной алгебре. – М.: Добросвет, 2009. – 320 с.
3. Колбина С. А., Пилюгин С. Ю. Линейная алгебра (дополнительные главы): Учеб. пособие. – СПб.: Изд-во СПбГЭТУ „ЛЭТИ“, 2009. – 60 с.
4. Беклемишев Д. В. Курс аналитической геометрии и линейной алгебры. – М.: Физматлит, 2009. – 312 с.
5. Смирнов В. И. Курс высшей математики: В 5 т. Т. 3, Ч. 1. – М.: Наука, 1974. – 324 с.
6. Шилов Г. Е. Математический анализ (конечномерные линейные пространства). – М.: Наука, 1969. – 432 с.
7. Беклемишев Д. В. Дополнительные главы линейной алгебры. – СПб.: Лань, 2008. – 496 с.

Оглавление

Введение	3
1. Системы линейных алгебраических уравнений. Матрицы	4
1.1. Системы линейных алгебраических уравнений. Равносильность	4
1.2. Метод полного исключения неизвестных (метод Гаусса – Жордана)	12
1.3. Примеры использования метода полного исключения неизвестных для решения систем линейных уравнений	17
1.4. Матрицы. Определения. Равенство матриц	22
1.5. Линейные операции с матрицами	24
1.6. Транспонированная и сопряженная матрицы	25
1.7. Умножение матриц	27
1.8. Матричная запись системы линейных алгебраических уравнений	31
2. Определители. Теорема Крамера. Матричные уравнения	33
2.1. Определители квадратных матриц	33
2.2. Основные свойства определителей	35
2.3. Дальнейшие свойства определителей	41
2.4. Определители ступенчатых матриц	44
2.5. Вычисление определителя матрицы методом Гаусса. Оценка числа операций	48
2.6. Теорема Крамера	51
2.7. Однородные системы линейных уравнений	55
2.8. Матричные уравнения	56
2.9. Обратная матрица	59
3. Конечномерные линейные пространства	62
3.1. Определение линейного пространства	62
3.2. Линейная зависимость векторов	64
3.3. Размерность линейного пространства	67
3.4. Базис и координаты в n -мерном пространстве	69
3.5. Изоморфность конечномерных пространств	71
3.6. n -мерные векторные пространства $\mathbb{C}^n$ и $\mathbb{R}^n$	73
3.7. Подпространства линейного пространства	77
3.8. Разложение линейного пространства $\mathbb{L}$ в прямую сумму подпространств	79
3.9. Сумма и пересечение подпространств	81
3.10. Ранг матрицы	83
3.11. Теорема Кронекера – Капелли	89
3.12. Однородные системы уравнений	92

3.13. Преобразование координат вектора при изменении базиса	97
4. Евклидовы и унитарные пространства	99
4.1. Скалярное произведение в линейном пространстве	99
4.2. Норма в линейном пространстве	102
4.3. Матрица Грама	104
4.4. Ортогональный и ортонормированный базис	107
4.5. Алгоритм Грама – Шмидта	109
4.6. Расстояние от точки до подпространства	112
4.7. Псевдорешение системы линейных уравнений	114
5. Элементы аналитической геометрии	117
5.1. Пространство направленных отрезков (геометрических векторов)	117
5.2. Изоморфизм пространства направленных отрезков и пространства $\mathbb{R}^3$	120
5.3. Скалярное произведение векторов в $\mathbb{R}^3$	123
5.4. Векторное произведение векторов в $\mathbb{R}^3$	125
5.5. Смешанное произведение векторов в $\mathbb{R}^3$	127
5.6. Уравнение плоскости	130
5.7. Уравнение прямой	134
Список литературы	137

Белопольский Андрей Львович
Бодунов Николай Александрович
Меркулов Александр Львович
Щеглова Александра Павловна

ЛИНЕЙНАЯ АЛГЕБРА

Учебное электронное пособие

Редактор Э. К. Долгатов

Подписано в печать . Формат 60×84 1/16. Бумага офсетная.
Печать офсетная. Гарнитура „Times New Roman“. Печ. л. 8,75.
Тираж 10 экз. Заказ

Издательство СПбГЭТУ „ЛЭТИ“
197376, С.-Петербург, ул. Проф. Попова, 5