МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное агентство по образованию

Пензенский государственный университет
архитектуры и строительства

В.А. Смирнов
А.М. Данилов

ЭЛЕМЕНТЫ ЛИНЕЙНОЙ АЛГЕБРЫ И АНАЛИТИЧЕСКОЙ ГЕОМЕТРИИ

Часть I
Линейные пространства, матрицы, системы линейных алгебраических уравнений

Рекомендовано редсоветом университета в качестве учебного пособия для студентов, обучающихся по специальности 230201 «Информационные системы и технологии»

Пенза 2006

УДК 516.0 (512.8)
ББК 22.151
С 50

Рецензент – проректор по информатизации, доктор химических наук, профессор А.Н. Кошев (Пензенский государственный университет архитектуры и строительства).

Смирнов В.А.

С 50 Элементы линейной алгебры и аналитической геометрии. Часть I. Линейные пространства, матрицы, системы линейных алгебраических уравнений: Учебное пособие / В.А. Смирнов, А.М. Данилов. – Пенза: ПГУАС, 2006. – 113 с.

Излагается материал, составляющий основу теории конечномерных линейных пространств. Рассматриваются алгебра матриц, элементарные преобразования матриц, элементы теории определителей, системы линейных алгебраических уравнений. Многие разделы сопровождаются необходимыми примерами, кратко обозначающими содержание практических занятий.

Учебное пособие подготовлено на кафедре высшей математики университета и предназначено для студентов вузов, обучающихся по специальности 230201 «Информационные системы и технологии».

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ

Настоящее учебное пособие является первой частью объединенного курса линейной алгебры и аналитической геометрии (составная часть курса «Алгебра, геометрия, анализ»).

В первой части рассматриваются основы теории конечномерных линейных пространств, алгебра матриц и системы линейных алгебраических уравнений. К изданию готовятся вторая и третья части, в которых будут изложены основы теории пространств с квадратичной метрикой, аналитическая геометрия на плоскости и в пространстве, основы теории линейных преобразований.

Многие разделы сопровождаются кратким перечнем контрольных вопросов и упражнениями для самостоятельной работы. Если упражнение заимствовано из другого источника, то дается соответствующая ссылка.

Учебное пособие подготовлено на кафедре высшей математики ПГУАС на основании лекций и практических занятий по дисциплинам «Высшая математика» и «Алгебра, геометрия, анализ». Пособие предназначено для студентов вузов, обучающихся по специальности 230201 «Информационные системы и технологии», однако может оказаться полезным и для студентов других специальностей.

ВВЕДЕНИЕ

В программе специальности 230201 «Информационные системы и технологии» математика и смежные с ней дисциплины занимают одно из ведущих мест. Построение программы предполагает, что курсы линейной алгебры и аналитической геометрии предваряют изучение остальных разделов. Поэтому изложение именно этих курсов должно в наибольшей мере способствовать решению важной и непростой задачи – *сформировать такой стиль мышления, который в дальнейшем будет помогать успешному применению математических методов в прикладных областях*.

Курс линейной алгебры насыщен теоретическими положениями, которые в большинстве случаев весьма просто доказываются. По нашему мнению, в процессе изложения материала эти доказательства должны быть сохранены. Характер рассуждений при доказательстве различных положений теории помогает развить «алгоритмический» стиль мышления – стиль, опирающийся прежде всего на логику, а не на интуицию.

Структура первой части пособия следующая. Изложение материала первой главы предваряется необходимыми сведениями из теории множеств. Приводятся первоначальные сведения о комплексных числах. Как правило, указанные разделы излагаются в первых главах учебников математического анализа, однако по нашему мнению данные разделы должны предварять изложение всего курса «Алгебра, геометрия, анализ».

Понятие линейного пространства, как это и принято, вводится аксиоматически; за определением следуют многочисленные примеры линейных пространств. Особое внимание уделяется линейному пространству векторов-отрезков и координатному (арифметическому) линейному пространству. Подробно рассматриваются вопросы, связанные с линейной зависимостью и базисом, линейным изоморфизмом, понятиями подпространства, суммы и пересечения пространств.

Изложение основ теории матриц во второй главе начинается с основных определений и понятий, связанных с этими объектами. Понятия линейной зависимости строк и столбцов, вырожденности и ранга вводятся до (и независимо от) понятия детерминанта. В процессе исследования широко используются элементарные преобразования матриц.

Детерминанты от нулевого до второго порядков вводятся аксиоматически. Доказывается существование и единственность детерминанта произвольного порядка, устанавливаются его свойства. Приводится краткий сравнительный анализ прямых методов вычисления детерминантов.

Исследование систем линейных уравнений в третьей главе полностью основано на методах линейной алгебры. Изложение методов Гаусса полностью подготовлено материалом первой и второй глав, поэтому сопровождается минимальными пояснениями. Приводится вывод формул Крамера. Вводится понятие обратной матрицы. Кратко рассматривается задача собственных значений (подробное рассмотрение этой задачи будет выполнено при изложении теории линейных операторов).

Глава 1. Основания линейной алгебры

В математике изучаются абстрактные связи и отношения между элементами множеств произвольной природы. Эти отношения чрезвычайно просты и по этой причине их удается выразить со строгостью, недоступной для естественных и общественных дисциплин.

Необходимо отметить, что *окружающая действительность не является предметом исследования в естественных науках. Исследованию доступны только модели – умозрительные конструкции, которые должны отражать существенные свойства реального объекта*. Широкое распространение моделирования связано именно с высокой общностью моделей, для построения и последующего анализа которых привлекаются методы математики.

1.1. Множества

Суть *аксиоматического метода* состоит в перечислении основных *неопределяемых* понятий, отношения между которыми фиксируются системой *аксиом*. Все остальные понятия и отношения теории выводятся посредством логических рассуждений.

Наиболее общим понятием математики является понятие *множества*. Дать определение этому понятию не удается; можно лишь пояснить, что близкими по содержанию являются такие понятия, как *набор* и *совокупность* объектов (возможно, разнородных). Подобного описания вполне достаточно, так как при построении теории природа самих объектов игнорируется: предметом исследования являются лишь свойства *операций*, *отношений* между объектами определенных множеств (в частности, в линейной алгебре рассматриваются отношения между объектами *линейных пространств*).

Для обозначения того, что объект a является элементом[1] множества A, используют знак *принадлежности*:

$$a \in A .$$

Если объект a не является элементом множества A, то записывают:

$$a \notin A .$$

Определить множество можно двумя способами. Первый способ состоит в перечислении всех элементов; при этом список элементов заключают в фигурные скобки. Например, множества $\{1,2,3\}$, $\{3,1,2\}$,

[1] Отношение «являться элементом» здесь мы также считаем неопределяемым.

$\{1,2,2,3\}$ и $\{3,3,1,1,1,2\}$ состоят из трех чисел 1, 2 и 3; множество $\{0,2,-2,4,-4,6,-6,...\}$ состоит из всех четных чисел, и т.д.[1]

Второй способ состоит в указании *характеристического свойства*, которым обладают те и только те элементы, которые принадлежат данному множеству. Например, множество всех четных чисел можно определить так:

$$\{x \mid x = 2k,\ k \in Z\}$$

(читается «множество x, таких, что $x = 2k$»); здесь Z − множество целых чисел.

Два множества называют *равными*, если они состоят из одних и тех же элементов. Например:

$$\{3,3,1,1,1,2\} = \{1,2,3\}.$$

Равенство множеств обладает такими же свойствами, что и равенство чисел:

1. $A = A$ (равенство *рефлексивно*).
2. Если $A = B$, то $B = A$ (равенство *симметрично*).
3. Если $A = B$ и $B = C$, то $A = C$ (равенство *транзитивно*).

Отношение, обладающее тремя указанными свойствами, называют *отношением эквивалентности*; подобные отношения нам часто будут встречаться в дальнейшем.

Множество A называют *подмножеством* множества B, если любой элемент множества A является элементом множества B. Отношение «являться подмножеством» записывают при помощи знака включения:

$$A \subset B$$

(читается «A включается в B» или «B включает A»). Если при этом множества A и B не равны, то A называют *собственным* подмножеством B[2]. Используя последнее определение, можно иначе сформулировать понятие равенства: два множества A и B равны тогда и только тогда, когда A является подмножеством B и множество B является подмножеством A.

---

1 Приведенный пример свидетельствует о том, что в записи множества порядок перечисления элементов не имеет значения.

2 Часто знак «$\subset$» используют для обозначения только собственных подмножеств, в то время как для отношения «включается или равно» используют знак «$\subseteq$».

Множество $\varnothing = \{\ \}$, не содержащее ни одного элемента, называется *пустым*. Пустое множество является подмножеством любого другого. Множество $\{\varnothing\}$, состоящее из пустого множества, само уже не является пустым.

Объединением (или *суммой*) множеств A и B называют множество $A \cup B$, состоящее из тех и только тех элементов, которые принадлежат хотя бы одному из множеств A или B (некоторый элемент принадлежит объединению множеств тогда и только тогда, когда он принадлежит или множеству A, или множеству B).

Пересечением (или *произведением*) множеств A и B называют множество $A \cap B$, состоящее из тех и только тех элементов, которые принадлежат каждому из множеств A и B (некоторый элемент принадлежит пересечению множеств тогда и только тогда, когда он одновременно принадлежит и множеству A, и множеству B).

Разностью множеств A и B называют множество $A \setminus B$, состоящее из тех и только тех элементов, которые принадлежат множеству A, но не принадлежат множеству B (некоторый элемент принадлежит разности $A \setminus B$ множеств тогда и только тогда, когда он одновременно принадлежит множеству A и не принадлежит множеству B).

Определения операций объединения, пересечения и разности указывают на то, что эти операции можно читать как «или», «и» и «без». Например, запись $(A \cap B) \setminus C$ читается «A и B в скобках без C».

Если пересечение множеств A и B пустое:

$$A \cap B = \varnothing,$$

то говорят, что множества A и B *не пересекаются*. В этом случае $A \setminus B = A$, $B \setminus A = B$.

В процессе рассуждений (но не в процессе доказательств!), связанных с операциями над множествами, весьма полезными оказываются *диаграммы Венна*. Множества на этих диаграммах обозначают в виде геометрических фигур произвольной формы. Интуитивные представления объединения (рис. 1.1), пересечения (рис. 1.2) и разности (рис. 1.3) геометрических фигур помогают составить представление о результатах операций над множествами. Например, на рис. 1.4 изображен результат операции $(A \cup B) \setminus C$.

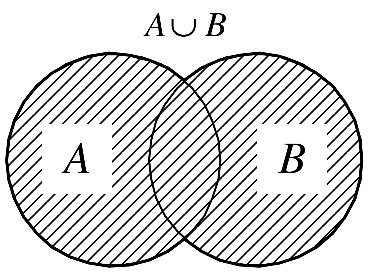

Рис. 1.1. Объединение множеств

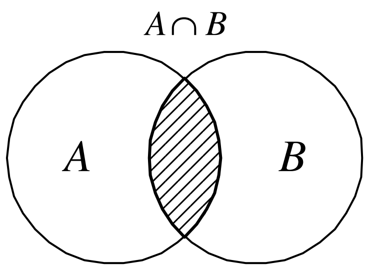

Рис. 1.2. Пересечение множеств

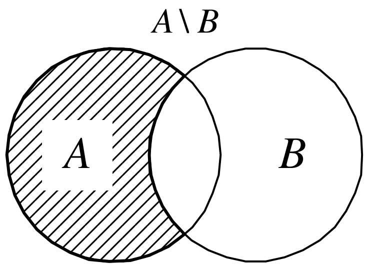

Рис. 1.3. Разность множеств

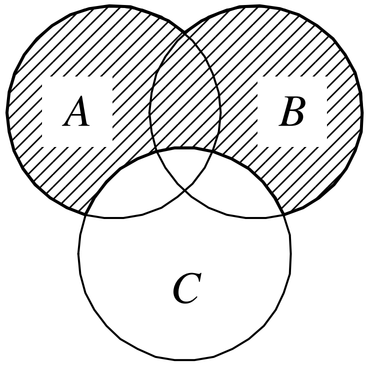

Рис. 1.4. Результат выполнения операции $(A \cup B) \setminus C$

Многие свойства операций объединения, пересечения и разности множеств аналогичны свойствам суммы, произведения и разности чисел, однако имеются и отличия.

1. $A \cup A = A$.
2. $A \cap A = A$.
3. $A \setminus A = \varnothing$.
4. $A \cup \varnothing = A$.
5. $A \cap \varnothing = \varnothing$.
6. $A \setminus \varnothing = A$.
7. $A \cup B = B \cup A$ (объединение *коммутативно*[1]).
8. $A \cap B = B \cap A$ (пересечение коммутативно).
9. $A \cup (B \cup C) = (A \cup B) \cup C$ (объединение *ассоциативно*[2]).
10. $A \cap (B \cap C) = (A \cap B) \cap C$ (пересечение ассоциативно).

---

[1] От лат. «изменяю».
[2] От лат. «соединяю».

11. $A\cap(B\cup C)=(A\cap B)\cup(A\cap C)$ (пересечение *дистрибутивно*[1] относительно объединения).

12. $(A\cup B)\setminus(A\cap B)=(A\setminus B)\cup(B\setminus A)$ (*симметрическая разность* множеств).

13. $A\cap(B\setminus C)=(A\cap B)\setminus(A\cap C)$ (пересечение дистрибутивно относительно разности).

Доказательство этих свойств основано на определениях соответствующих операций. Докажем, например, свойство 11. Нам требуется доказать справедливость включений:

$$A\cap(B\cup C)\subset(A\cap B)\cup(A\cap C),$$

$$(A\cap B)\cup(A\cap C)\subset A\cap(B\cup C).$$

Пусть некоторый элемент принадлежит множеству $A\cap(B\cup C)$. Тогда он принадлежит множеству A и, по меньшей мере, хотя бы одному из множеств B или C. Но тогда он либо одновременно принадлежит множествам A и B, либо одновременно принадлежит множествам A и C. Следовательно, он принадлежит объединению пересечений $(A\cap B)\cup(A\cap C)$. Обратно, пусть элемент принадлежит множеству $(A\cap B)\cup(A\cap C)$. Тогда он принадлежит или пересечению $A\cap B$, или пересечению $A\cap C$. Следовательно, он принадлежит множеству A и хотя бы одному из множеств B или C, т.е. пересечению $A\cap(B\cup C)$. Свойство доказано.

Контрольные вопросы и упражнения

1. Сформулируйте определения понятий равенства, объединения, пересечения и разности множеств.

2. В чем отличие понятий подмножества и собственного подмножества?

3 [5]. Старейший математик среди шахматистов и старейший шахматист среди математиков – это один и тот же человек или (возможно) разные?

4 [5]. Лучший математик среди шахматистов и лучший шахматист среди математиков – это один и тот же человек или (возможно) разные?

5. Проведите доказательство свойств 7, 9, 12 и 13.

---

[1] От лат. «распределяю».

6. Докажите, что $(A \cap B) \setminus C = (A \setminus C) \cap (B \setminus C)$.

7. Докажите, что $C \setminus (A \cap B) = (C \setminus A) \cup (C \setminus B)$.

8. Докажите, что $C \setminus (A \cup B) = (C \setminus A) \cap (C \setminus B)$.

9 [5]. Существуют ли такие множества *A*, *B* и *C*, для которых выполнено $A \cap B \neq \varnothing$, $A \cap C = \varnothing$, но $(A \cap B) \setminus C = \varnothing$?

1.2. Первоначальные сведения о комплексных числах[1]

Сумма, разность, произведение и частное (в предположении о том, что делитель отличен от нуля) двух действительных чисел сами также являются действительными числами. Говорят, что множество действительных чисел *замкнуто* по отношению к четырем арифметическим действиям.

Однако многие операции на множестве действительных чисел[2] не выполнимы. Например, среди действительных чисел нет корней уравнения $x^2 + 1 = 0$. Дальнейшее расширение числовых множеств (от натуральных чисел – к целым, рациональным и действительным) приводит к множеству *комплексных чисел C*.

Комплексным числом называется упорядоченная пара[3] $z = (a, b)$ действительных чисел, действия над которой производятся в соответствии с тремя аксиомами:

1. $(a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2)$.

2. $\lambda(a, b) = (\lambda a, \lambda b)$, $\lambda \in R$.

3. $(a_1, b_1)(a_2, b_2) = (a_1 a_2 - b_1 b_2, a_1 b_2 + a_2 b_1)$.

Первое из чисел пары (a, b) называется *действительной частью* комплексного числа и обозначается

$$a = \operatorname{Re} z.$$

Второе число называется *мнимой частью*:

$$b = \operatorname{Im} z.$$

Число называется *чисто мнимым*, если его действительная часть равна нулю. Два комплексных числа равны тогда и только тогда,

[1] Материал этого пункта требуется только при рассмотрении задачи собственных значений. При первом чтении пункт можно пропустить.

[2] Напомним, что это множество обозначается буквой *R*.

[3] Пара называется упорядоченной, если указано, какой из элементов является первым, какой – вторым. Аналогично определяется понятие *упорядоченного набора*.

когда равны их действительные и мнимые части. Операции сравнения для комплексных чисел не определены.

Два числа называют *комплексно-сопряженными*, если их действительные части одинаковы, а мнимые отличаются знаком:

$$(a,b)^* = (a,-b).$$

Для деления двух комплексных чисел числитель и знаменатель дроби следует умножить на число, комплексно-сопряженное знаменателю. Мнимая часть знаменателя полученной дроби оказывается равной нулю:

$$\frac{(a,b)}{(c,d)} = \frac{(a,b)(c,d)^*}{(c,d)(c,d)^*} = \frac{(a,b)(c,-d)}{(c,d)(c,-d)} = \frac{(ac+bd, bc-ad)}{(c^2+d^2,0)}.$$

Модулем $|z|$ комплексного числа $z=(a,b)$ называют арифметический квадратный корень из произведения числа на комплексно-сопряженное:

$$|z| = \sqrt{zz^*} = \sqrt{a^2+b^2}.$$

Действия над комплексными числами можно производить по правилам алгебры действительных чисел, если каждому действительному числу a поставить в соответствие число $(a,0)$, ввести обозначение

$$i = (0,1),$$

и записать комплексное число в *алгебраической форме*:

$$(a,b) = a + ib.$$

Чисто мнимое число $i=(0,1)$ называют *мнимой единицей*. Найдем его натуральные степени; очевидно, $i^1 = i$. Далее, в соответствии с аксиомой 3 имеем:

$$i^2 = (0,1)(0,1) = (0\cdot 0 - 1\cdot 1, 0\cdot 1 + 0\cdot 1) = (-1,0) = -1;$$

$$i^3 = i^2 i = (-1)i = -i;$$

$$i^4 = \left(i^2\right)^2 = (-1)^2 = 1;$$

$$i^5 = i^4 i = 1\cdot i = i;$$

$i^6 = i^5 i = i^2 = -1$,

и т.д.

Обратимся к целочисленным степеням:

$i^0 = \frac{i}{i} = 1$;

$i^{-1} = \frac{1}{i} = \frac{i}{i^2} = \frac{i}{(-1)} = -i$;

$i^{-2} = \frac{1}{i^2} = \frac{1}{(-1)} = -1$;

$i^{-3} = \frac{i}{i^4} = \frac{i}{1} = i$,

и т.д.

Окончательно:

$$i^n = \begin{cases} 1, & n = 4k \\ i, & n = 4k+1 \\ -1, & n = 4k+2 \\ -i, & n = 4k+3 \end{cases}, \quad k \in Z.$$

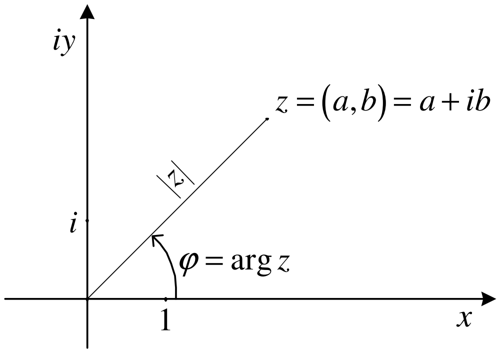

Рис. 1.5. Комплексная плоскость

Каждому комплексному числу (a,b) можно поставить в соответствие его *аффикс* – точку (a,b) на *комплексной плоскости* (рис. 1.5). Ось абсцисс этой плоскости называют *действительной осью*, ось ординат – *мнимой осью*.

Аргументом $\operatorname{Arg} z$ *комплексного числа* $z = (a,b)$ называют угол между положительным направлением действительной оси и отрезком, соединяющим начало координат и аффикс. Аргумент определен с точностью до слагаемого, кратного 2π. *Главным значением* $\arg z$ аргумента называют значение, принадлежащее интервалу $[0;2\pi)$.

Аффиксы действительных чисел расположены на действительной оси комплексной плоскости. Поэтому главные значения аргумента действительных чисел равны нулю (для положительных) или π (для отрицательных чисел).

Аффиксы число мнимых чисел расположены на мнимой оси комплексной плоскости. Поэтому главные значения аргумента этих чисел равны $\pm\dfrac{\pi}{2}$.

Так как для $z=(a,b)$ выполнено:

$$\begin{cases} a=|z|\cos(\arg z) \\ b=|z|\sin(\arg z) \end{cases}, \tag{1.1}$$

то число $z=(a,b)$ может быть записано в форме:

$$z=|z|\left(\cos(\arg z)+i\sin(\arg z)\right),$$

которую называют *тригонометрической формой* записи комплексного числа.

Произведение чисел z_1 и z_2, заданных в тригонометрической форме:

$$z_1 z_2=|z_1|\left(\cos(\arg z_1)+i\sin(\arg z_1)\right)|z_1|\left(\cos(\arg z_1)+i\sin(\arg z_1)\right)=$$
$$=|z_1||z_2|\left(\cos(\arg z_1+\arg z_2)+i\sin(\arg z_1+\arg z_2)\right),$$

т.е. *при умножении модули перемножаются,яаяаргументыяскладываются*.

В частности:

$$z^n=|z|^n\left(\cos(n\arg z)+i\sin(n\arg z)\right).$$

Последнее соотношение называется - *ормулойяМуавра*.

Корнем $w=\sqrt[n]{z}$ из комплексного числа z называется решение уравнения $w^n=z$. Из формулы Муавра следует, что корень n-ной степени имеет ровно n различных значений:

$$\sqrt[n]{z}=\sqrt[n]{|z|}\left(\cos\left(\frac{\arg z+2\pi k}{n}\right)+i\sin\left(\frac{\arg z+2\pi k}{n}\right)\right),$$

где $\sqrt[n]{|z|}$ – арифметический корень (действительное число), $k=0,1,\ldots,n-1$.

В алгебре доказывается, что на множестве комплексных чисел алгебраическое уравнение n-й степени имеет ровно n корней[1], считая с кратными.

---

[1] Следствие из т.н. *основной теоремы алгебры*.

Например, квадратное уравнение

$$x^2 - 3x + 2 = 0$$

имеет корни $x_1 = 1$ и $x_2 = 2$;

уравнение

$$x^2 - 2x + 5 = 0$$

имеет корни $x_1 = 1 + 2i$ и $x_2 = 1 - 2i$;

уравнение

$$x^2 - 4x + 4 = 0$$

имеет один *двукратный* корень $x = 2$.

Контрольные вопросы и упражнения

1. Что называют комплексным числом?

2. Что называют действительной и мнимой частями комплексного числа? Какое число называют комплексно сопряженным данному?

3. Чему равны действительная и мнимая части произведения двух комплексных чисел?

4. Какое число называют мнимой единицей?

5. Что называют модулем и аргументом комплексного числа?

6. Используя аксиомы 1…3, выполните действия:

a) $2(2,-3)+(-4,7)$; b) $-(-1,5)+2(3,-1)$; c) $(2,-1)(3,4)$; d) $(5,-2)(1,3)$.

7. Выполните действия с комплексными числами, представленными в алгебраической форме:

a) $2(2+3i)+(1-i)$; b) $(5+i)(1-i)$; c) $(2+i)(3-2i)(1+3i)$; d) $(3+2i)^2$;

e) $(2-i)^4$; f) $(1+i)^8$; g) $\dfrac{4+2i}{1-i}$; h) $\dfrac{4-3i}{2+i}$.

8. Используя соотношение (1.1), представьте комплексные числа в тригонометрической форме (найдите модули и главные значения аргументов комплексных чисел):

a) $z=1$; b) $z=-3$; c) $z=-5i$; d) $z=1-i$; e) $z=1+i\sqrt{3}$; f) $z=\sqrt{3}-i$.

9. Докажите, что при любых $\{\alpha,\beta\} \subset R$ справедливо

$$\left|\frac{\alpha+i\beta}{\alpha-i\beta}\right| = 1.$$

10. Пользуясь формулой Муавра, вычислите:

$$\text{a) } (1+i)^{10}; \text{ b) } \left(\sqrt{3}+i\right)^5.$$

11. Найдите все значения корней:

$$\text{a) } \sqrt[3]{8}; \text{ b) } \sqrt[4]{81}; \text{ c) } \sqrt[6]{-1}.$$

12. Решите уравнения:

$$\text{a) } x^2-2x+2=0; \text{ b) } x^2-4x+5=0; \text{ c) } x^2+6x+25=0.$$

13*. Докажите, что если число $z=a+ib$ не принадлежит отрицательной действительной полуоси, то найдется $\beta \in R$, такое, что

$$a+ib=\sqrt{a^2+b^2}\,\frac{1+i\beta}{1-i\beta}.$$

1.3. Сокращенная запись суммирования

В дальнейшем нам будут часто встречаться суммы, содержащие произвольное (или же вполне определенное, но большое) число слагаемых. В подобных случаях для сокращения записи используют символ суммы, обозначая все слагаемые буквой с одним или несколькими индексами:

$$\sum_{i=1}^{n} a_i = a_1 + a_2 + \ldots + a_n$$

(читается «сумма a_i по i от 1 до n»).

Индекс i называют *индексом суммирования*. Числа 1 и n называют *нижним* и *верхним пределами суммирования*. Если пределы суммирования ясны из контекста, то указывают только индекс:

$$\sum_{i} a_i .$$

Все условия, определяющие множество значений индексов суммирования, можно записывать под нижним пределом. Например:

$$\sum_{1\le i\le n} a_i = \sum_{i=1}^{n} a_i .$$

Если, в частности, необходимо пропустить k-е слагаемое, то под нижним пределом следует записать неравенство:

$$\sum_{\substack{i=1\\ i\ne k}}^{n} a_i = a_1 + a_2 + \ldots a_{k-1} + a_{k+1} + \ldots + a_n .$$

Обозначение индекса суммирования может быть изменено (говорят, что индекс суммирования является *немым*, или *связанным*):

$$\sum_{i=1}^{n} a_i = \sum_{k=1}^{n} a_k .$$

Известные свойства арифметических операций естественно обобщаются для символа суммы.

1. Множитель, не зависящий от индекса суммирования, можно выносить за знак суммы:

$$\sum_{i=1}^{n} \alpha a_i = \alpha \sum_{i=1}^{n} a_i .$$

2. Если суммирование идет по нескольким различными индексам, каждый из которых меняется независимо от других, то порядок суммирования можно менять, не меняя пределов. Например, для двойной суммы

$$\sum_{i=1}^{n} \sum_{k=1}^{m} a_{ik} = \sum_{k=1}^{m} \sum_{i=1}^{n} a_{ik} .$$

Подразумевается, что в левой части слагаемые сначала суммируются по индексу k, а затем – по индексу i:

$$\sum_{i=1}^{n} \sum_{k=1}^{m} a_{ik} = \sum_{i=1}^{n} \left(a_{i1} + a_{i2} + ... + a_{im} \right) ,$$

в то время как в правой части суммирование первоначально выполняется по индексу i:

$$\sum_{k=1}^{m} \sum_{i=1}^{n} a_{ik} = \sum_{k=1}^{m} \left(a_{1k} + a_{2k} + ... + a_{nk} \right) .$$

Следствием двух последних правил является правило 3.

3. При суммировании по нескольким индексам за знак внутренней суммы можно выносить множитель, зависящий только от индекса внешней суммы. Например:

$$\sum_{i=1}^{n} \sum_{k=1}^{m} a_i b_k = \sum_{i=1}^{n} a_i \sum_{k=1}^{m} b_k = \sum_{k=1}^{m} b_k \sum_{i=1}^{n} a_i ;$$

$$\sum_{i=1}^{n} \sum_{j=1}^{m} \sum_{k=1}^{l} \lambda a_{ijk} b_{ij} c_i = \lambda \sum_{i=1}^{n} c_i \sum_{j=1}^{m} b_{ij} \sum_{k=1}^{l} a_{ijk} .$$

Если опущен символ суммы (и не указаны индексы суммирования), то суммирование подразумевается по тем индексам, которые встречаются дважды. Например:

$$\sum_i a_i b_i c_k = a_i b_i c_k$$

(говорят, что индекс i – *немой*, а индекс k – *свободный*),

$$\sum_k \sum_j a_{ik} b_{kj} c_{js} = a_{ik} b_{kj} c_{js},$$

и т.д.

Следует обратить внимание на то, что результат суммирования *зависит* от свободного индекса (по этому индексу суммирования нет). Так, для первого примера

$$a_i b_i c_k = a_1 b_1 c_k + a_2 b_2 c_k + \ldots$$

Подобная запись сумм распространена в прикладных дисциплинах, однако требует некоторого навыка; по возможности мы будем ее избегать.

Контрольные вопросы и упражнения

1. Раскройте суммы (представьте в развернутом виде):

$$\text{a) } \sum_{k=1}^{5} a_k\text{; b) } \sum_{k=1}^{3} \alpha_k \sum_{s=1}^{2} x_{ks}.$$

2. Выясните число слагаемых в суммах:

$$\text{a) } \sum_{k=1}^{10} a_k\text{; b) } \sum_{k=2}^{7}\sum_{s=3}^{5} a_{ks}\text{; c) } \sum_{i=1}^{2}\sum_{j=1}^{3}\sum_{k=1}^{4}\sum_{s=1}^{5} a_{ijks}\text{; d) } \sum_{i=1}^{10}\sum_{j=i}^{10} a_{ij}.$$

3. Вычислите:

$$\text{a) } \sum_{k=1}^{10} 2k\text{; b) } \sum_{k=1}^{4} 2^k\text{; c) } \sum_{k=1}^{5}\left(1+k+k^2\right)\text{; d) } \sum_{k=1}^{3}\sum_{s=1}^{k} sk^2.$$

1.4. Линейное пространство

Пусть задано множество $L = \{\mathbf{a}, \mathbf{b}, \mathbf{c}, \ldots\}$, содержащее элементы произвольной природы. Пусть также $U = \{\alpha, \beta, \gamma, \ldots\}$ – числовое множество (его элементы мы будем называть *скалярами*).

Пусть на множестве L определено отношение равенства, являющееся отношением эквивалентности (т.е. удовлетворяющее свойствам рефлексивности, симметричности и транзитивности).

Для элементов множества L определим операции *сложения* и *умножения на число*, такие что:

– сумма элементов из L есть вновь элемент из L (множество L *замкнуто по сложению*);

– произведение элемента из L на число из U есть вновь элемент из L (множество L *замкнуто по умножению на скаляр*).

Пусть выполнены аксиомы:

1. $\mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a}$ (сложение коммутативно).

2. $(\mathbf{a}+\mathbf{b})+\mathbf{c}=\mathbf{a}+(\mathbf{b}+\mathbf{c})$ (сложение ассоциативно).

3. Существует *нулевой* элемент (*ноль*) по сложению $\mathbf{0}\in L$, такой что $\mathbf{a}+\mathbf{0}=\mathbf{a}$.

4. Для любого элемента $\mathbf{a}\in L$ существует элемент $\mathbf{b}\in L$, *противоположный элементу* $\mathbf{a}$ *по сложению*: $\mathbf{a}+\mathbf{b}=\mathbf{0}$.

5. Существует *единица* 1 – число, для которого $1\cdot\mathbf{a}=\mathbf{a}$.

6. $\alpha(\beta\mathbf{a})=(\alpha\beta)\mathbf{a}$ (умножение на число ассоциативно).

7. $(\alpha+\beta)\mathbf{a}=\alpha\mathbf{a}+\beta\mathbf{a}$ (распределительный закон для элемента из множества L: умножение на элемент из L дистрибутивно относительно сложения чисел).

8. $\alpha(\mathbf{a}+\mathbf{b})=\alpha\mathbf{a}+\alpha\mathbf{b}$ (распределительный закон для скаляра: умножение на число дистрибутивно относительно сложения элементов из множества L).

Тогда операции сложения и умножения на число называют *линейными*, а множество L называют *линейным пространством*, или *линейным многообразием*. Если при этом множество U совпадает с множеством действительных чисел, то L называют *действительным линейным пространством*; если U совпадает с множеством комплексных чисел, то L называют *комплексным линейным пространством*.

Элементы линейного пространства называют *векторами*[1]; представления, связанные с понятием вектора, часто помогают находить геометрический смысл в полученных результатах.

---

[1] В дальнейшем как «геометрические» векторы (*векторы-отрезки*), так и векторы произвольного линейного пространства мы будем обозначать строчными буквами, выделяя их полужирным шрифтом.

Нулевой элемент по сложению называют *нулевым вектором*. Вектор **c** называют *разностью* векторов **a** и **b**, если $\mathbf{c}+\mathbf{b}=\mathbf{a}$; записывают: $\mathbf{c}=\mathbf{a}-\mathbf{b}$.

Линейное пространство называют *нулевым*, или *тривиальным*, если оно содержит единственный нулевой вектор.

Аксиомы 1...8 позволяют установить основные свойства линейного пространства.

Предложение 1.4.1. *Линейное пространство содержит только один нулевой вектор.*

Действительно, пусть **a** и **b** – нулевые векторы. Тогда в силу аксиом 1 и 3:

$$\mathbf{b}=\mathbf{b}+\mathbf{a}=\mathbf{a}+\mathbf{b}=\mathbf{a};$$

векторы **a** и **b** совпадают.

Предложение 1.4.2. *Для любого вектора существует единственный противоположный.*

Действительно, пусть **b** и **c** – противоположные векторы для **a**: $\mathbf{a}+\mathbf{b}=\mathbf{0}$, $\mathbf{a}+\mathbf{c}=\mathbf{0}$.

На основании аксиом 1...4 имеем

$$\mathbf{b}=\mathbf{b}+\mathbf{0}=\mathbf{b}+(\mathbf{a}+\mathbf{c})=(\mathbf{b}+\mathbf{a})+\mathbf{c}=(\mathbf{a}+\mathbf{b})+\mathbf{c}=\mathbf{0}+\mathbf{c}=\mathbf{c};$$

векторы **b** и **c** совпадают.

Предложение 1.4.3. *Произведение любого вектора на число* 0 *равно нулевому вектору.*

Для доказательства рассмотрим вектор **b**, противоположный вектору **a**. Тогда

$$0\cdot\mathbf{a}=0\cdot\mathbf{a}+\mathbf{0}=0\cdot\mathbf{a}+(\mathbf{a}+\mathbf{b})=(0+1)\mathbf{a}+\mathbf{b}=\mathbf{a}+\mathbf{b}=\mathbf{0}.$$

Предложение 1.4.4. *Произведение нулевого вектора* **0** *на любое число равно нулевому вектору.*

Действительно, пусть **a** – произвольный вектор; тогда

$$\alpha\cdot\mathbf{0}=\alpha(0\cdot\mathbf{a})=(\alpha\cdot 0)\mathbf{a}=0\cdot\mathbf{a}=\mathbf{0}.$$

Предложение 1.4.5. *Произведение ненулевого вектора на число, не равное нулю, всегда есть ненулевой вектор.*

Действительно, пусть $\alpha\neq 0$, $\mathbf{a}\neq\mathbf{0}$. Предположим $\alpha\mathbf{a}=\mathbf{0}$; тогда

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

что противоречит условию $\mathbf{a}\neq\mathbf{0}$. Предложение доказано.

Предложение 1.4.6. *Если* α *и* β *– различные числа, и вектор* $\mathbf{a}$ *– ненулевой, то векторы* $\alpha\mathbf{a}$ *и* $\beta\mathbf{a}$ *также различны.*

Предположим $\alpha\mathbf{a} = \beta\mathbf{a}$. Тогда $\alpha\mathbf{a} + (-\beta)\mathbf{a} = \mathbf{0}$, и должно быть

$$(\alpha - \beta)\mathbf{a} = \mathbf{0},$$

что в силу $\alpha \neq \beta$ и $\mathbf{a} \neq \mathbf{0}$ противоречит условию. Предложение доказано.

Предложение 1.4.7. *Если действительное линейное пространство не является нулевым, то оно содержит бесконечно много элементов.*

Для доказательства достаточно заметить, что если вектор $\mathbf{a}$ – ненулевой и действительные числа α и β не равны, то векторы $\alpha\mathbf{a}$ и $\beta\mathbf{a}$ будут различны. В силу бесконечности множества действительных чисел линейное пространство также будет содержать бесконечно много элементов.

Предложение 1.4.8. *Произведение вектора* $\mathbf{a}$ *на число* -1 *противоположно вектору* $\mathbf{a}$.

Действительно:

$$\mathbf{a} + (-1)\mathbf{a} = (1-1)\mathbf{a} = 0 \cdot \mathbf{a} = \mathbf{0}.$$

Предложение 1.4.9. *Для любых векторов* $\mathbf{a}$ *и* $\mathbf{b}$ *разность* $\mathbf{a} - \mathbf{b}$ *существует и единственна.*

Для доказательства существования рассмотрим вектор

$$\mathbf{c} = \mathbf{a} + (-1)\mathbf{b},$$

представляющий собой сумму вектора $\mathbf{a}$ и вектора, противоположного вектору $\mathbf{b}$. Имеем:

$$\mathbf{c} + \mathbf{b} = \mathbf{a} + (-1)\mathbf{b} + \mathbf{b} = \mathbf{a} + (-1+1)\mathbf{b} = \mathbf{a} + 0 \cdot \mathbf{b} = \mathbf{a}.$$

Докажем единственность. Пусть вектор $\mathbf{c}$ является разностью векторов $\mathbf{a}$ и $\mathbf{b}$: $\mathbf{c} + \mathbf{b} = \mathbf{a}$. Тогда

$$\mathbf{c} = \mathbf{c} + 0 \cdot \mathbf{b} = \mathbf{c} + (1-1)\mathbf{b} = \mathbf{c} + \mathbf{b} + (-1)\mathbf{b} = \mathbf{a} + (-1)\mathbf{b};$$

следовательно, разность всегда можно представить в виде суммы $\mathbf{c} = \mathbf{a} + (-1)\mathbf{b}$. Предложение доказано.

Замечание 1.1. Важно, что все приведенные выше предложения доказаны для *произвольного* линейного пространства (безотносительно к природе его элементов). Доказательства основаны исключительно на аксиомах 1...8.

Контрольные вопросы и упражнения

1. Перечислите аксиомы линейного пространства.
2. Какое линейное пространство называют тривиальным?

1.5. Примеры линейных пространств

1. Пространство векторов-отрезков. *Вектором-отрезком* $\mathbf{a} = \mathbf{AB}$ назовем упорядоченную пару точек A и B. Расстояние между начальной и конечной точками назовем *длиной* (или *модулем*) вектора:

$$|\mathbf{a}| = AB\,.$$

Модуль вектора мы часто будем обозначать той же буквой, что и сам вектор, но набранной обычным шрифтом:

$$a = |\mathbf{a}|\,.$$

Нулевым вектором назовем вектор $\mathbf{0}$, начало и конец которого совпадают. Модуль нулевого вектора равен нулю.

Векторы, параллельные одной и той же прямой, назовем *коллинеарными*. Нулевой вектор мы будем считать коллинеарным любому другому вектору.

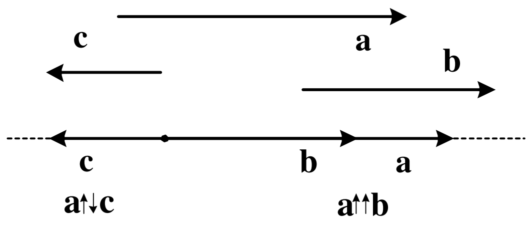

Рис. 1.6 Сонаправленные и противоположно направленные векторы

Пусть векторы коллинеарны. Совместим их начальные точки; если при этом конечные точки окажутся по одну сторону от (общей) начальной, то векторы называются *сонаправленными*. Если конечные точки оказываются по разные стороны от начальной, то векторы называются *противоположно направленными* (рис. 1.6).

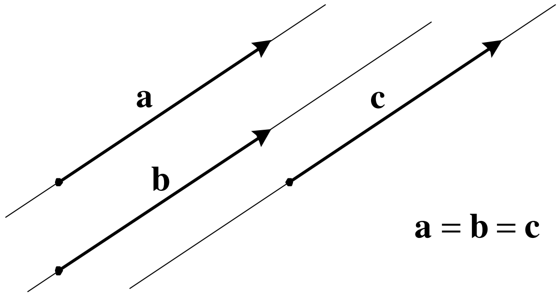

Рис. 1.7. Равные векторы

На множестве векторов-отрезков определим отношение равенства следующим образом: два вектора равны тогда и только тогда, когда они *сонаправлены и имеют одинаковые модули* (рис. 1.7). Нетрудно убедиться, что введенное таким образом отношение равенства рефлексивно, симметрично и транзитивно.

Замечание 1.2. В дальнейшем мы не будем делать никакого различия между равными векторами, считая их *одним и тем же* вектором. Однако следует помнить, что потребности практики могут сделать целесообразным иное определение равенства. Например, *связанные* векторы считаются равными тогда и только тогда, когда совпадают их начальные и конечные точки. *Скользящие* векторы считаются равными тогда и только тогда, когда они лежат на одной прямой, сонаправлены и имеют равные модули.

Суммой двух векторов **a** и **b** назовем вектор $\mathbf{c}=\mathbf{a}+\mathbf{b}$, замыкающий построенную на векторах **a** и **b** ломаную. При построении ломаной конечная точка вектора **a** совмещается с начальной точкой вектора **b**, начальная точка вектора **c** совмещается с начальной точкой вектора **a**, конечная точка вектора **c** совмещается с конечной точкой вектора **b** (рис. 1.8). Это определение суммы векторов называют *правилом треугольника*.

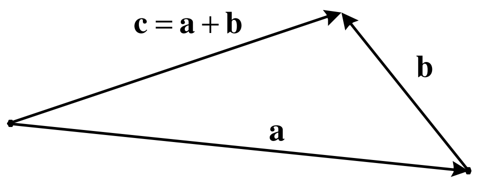

Рис. 1.8. Сумма векторов

Произведением вектора **a** на число λ назовем вектор $\mathbf{c}=\lambda\mathbf{a}$, модуль которого равен произведению модулей $c=|\lambda|a$ и который сонаправлен вектору **a** при $\lambda>0$ или противоположно направлен при $\lambda<0$ (рис. 1.9).

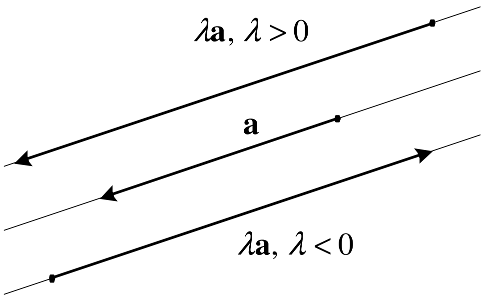

Рис. 1.9. Произведение вектора на число

Определенные подобным образом операции сложения и умножения на число замкнуты (и сумма, и произведение на число сами по определению являются *векторами-отрезками*).

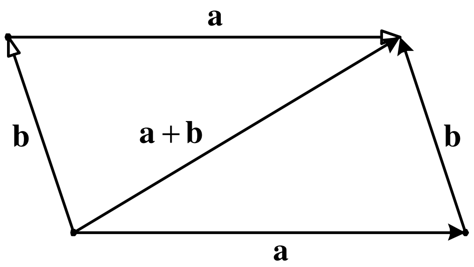

Рис. 1.10. Коммутативность сложения

Коммутативность сложения (аксиома 1) иллюстрируется *правилом параллелограмма* (рис. 1.10).

Для доказательства ассоциативности сложения достаточно построить векторы $(\mathbf{a}+\mathbf{b})+\mathbf{c}$ и $\mathbf{a}+(\mathbf{b}+\mathbf{c})$ (рис. 1.11).

Нулем по сложению в пространстве векторов-отрезков является нулевой вектор. Два противоположно направленных вектора являются элементами, противоположными по сложению.

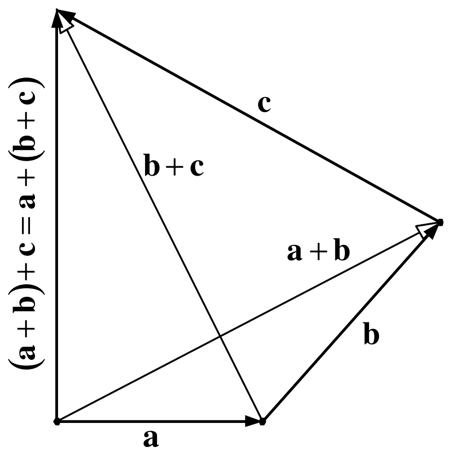

Рис. 1.11. Ассоциативность сложения

Единицей умножения является число 1.

Для доказательства ассоциативности умножения на скаляр заметим, что векторы $\alpha(\beta\mathbf{a})$ и $(\alpha\beta)\mathbf{a}$ имеют равные модули:

$$|\alpha(\beta\mathbf{a})| = |\alpha||\beta\mathbf{a}| = |\alpha||\beta|a =$$

$$= |\alpha\beta|a = |(\alpha\beta)\mathbf{a}|.$$

Пусть оба числа α и β положительны. Тогда

$$\alpha(\beta\mathbf{a}) \uparrow\uparrow (\beta\mathbf{a}) \uparrow\uparrow \mathbf{a}; \ (\alpha\beta)\mathbf{a} \uparrow\uparrow \mathbf{a}.$$

Следовательно:

$$\alpha(\beta\mathbf{a}) \uparrow\uparrow \alpha(\beta\mathbf{a}).$$

Сонаправленность векторов $\alpha(\beta\mathbf{a})$ и $\alpha(\beta\mathbf{a})$ для трех других сочетаний знаков чисел α и β доказывается аналогично.

Докажем выполнение распределительного закона для вектора-отрезка (аксиома 7). Пусть, вновь, оба числа α и β положительны. Тогда три вектора $(\alpha+\beta)\mathbf{a}$, $\alpha\mathbf{a}$ и $\beta\mathbf{a}$ сонаправлены; следовательно, сонаправлеными будут векторы $(\alpha+\beta)\mathbf{a}$ и $\alpha\mathbf{a}+\beta\mathbf{a}$. Учитывая, что модуль суммы сонаправленных векторов равен сумме модулей слагаемых, получим:

$$|(\alpha+\beta)\mathbf{a}| = |\alpha+\beta|a = (|\alpha|+|\beta|)a = |\alpha|a + |\beta|a = |\alpha\mathbf{a}| + |\beta\mathbf{a}| = |\alpha\mathbf{a}+\beta\mathbf{a}|.$$

Аналогично доказывается сонаправленность и равенство модулей векторов $(\alpha+\beta)\mathbf{a}$ и $\alpha\mathbf{a}+\beta\mathbf{a}$ для трех других сочетаний знаков.

Докажем выполнение распределительного закона для скаляра (аксиома 8). Пусть, для определенности, $\alpha > 0$. Пользуясь правилом параллелограмма, построим векторы $\mathbf{a}+\mathbf{b}$ и $\alpha(\mathbf{a}+\mathbf{b})$. Выполнение аксиомы 8 следует из подобия параллелограммов (рис. 1.12): при умножении векторов $\mathbf{a}$ и $\mathbf{b}$ на число α диагональ $\mathbf{a}+\mathbf{b}$ также увеличивается в α раз.

Таким образом, для определенного отношения равенства, операций сложения и умножения векторов-отрезков на скаляр выполняются все аксиомы 1…8; следовательно, данное множество является линейным пространством.

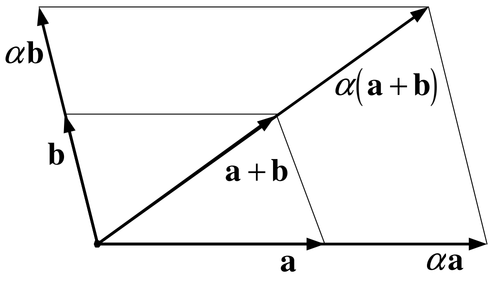

Рис. 1.12. При умножение слагаемых на одно и то же число их сумма также умножается на это число

2. Пространство многочленов. Рассмотрим множество, элементами которого являются многочлены n-й степени

$$P_n(x) = \sum_{k=0}^{n} a_k x^k .$$

Отношение равенства, операции сложения и умножения многочлена на число можно определить так же, как это делается в алгебре. Два многочлена будем считать равными, если равны все их коэффициенты (иначе: многочлены $P_n(x)$ и $Q_n(x)$ равны тогда и только тогда, когда равенство их значений имеет место при любом значении x).

Суммой многочленов будем считать многочлен, коэффициенты которого равны суммам коэффициентов многочленов-слагаемых:

$$\sum_{k=0}^{n} a_k x^k + \sum_{k=0}^{n} b_k x^k = \sum_{k=0}^{n} (a_k + b_k) x^k .$$

Произведением многочлена на число будем считать многочлен, коэффициенты которого равны произведениям коэффициентов исходного многочлена на это число:

$$\alpha \sum_{k=0}^{n} a_k x^k = \sum_{k=0}^{n} (\alpha a_k) x^k .$$

Тогда выполнение аксиом 1…8 следует из свойств арифметических операций на множестве действительных чисел. Например, для аксиомы 1 имеем:

$$\sum_{k=0}^{n} a_k x^k + \sum_{k=0}^{n} b_k x^k = \sum_{k=0}^{n} (a_k + b_k) x^k = \sum_{k=0}^{n} (b_k + a_k) x^k = \sum_{k=0}^{n} b_k x^k + \sum_{k=0}^{n} a_k x^k .$$

Для аксиомы 7:

$$(\alpha+\beta)\sum_{k=0}^{n} a_k x^k = \sum_{k=0}^{n}\left((\alpha+\beta)a_k\right)x^k = \sum_{k=0}^{n}(\alpha a_k + \beta a_k)x^k =$$

$$= \sum_{k=0}^{n}\left((\alpha a_k)x^k + (\beta a_k)x^k\right) = \sum_{k=0}^{n}(\alpha a_k)x^k + \sum_{k=0}^{n}(\beta a_k)x^k = \alpha\sum_{k=0}^{n} a_k x^k + \beta\sum_{k=0}^{n} a_k x^k ,$$

и т.д.

Множество многочленов n-й степени является линейным пространством. Нулем по сложению в этом пространстве является нулевой многочлен

$$\theta_n(x) = 0 .$$

Два многочлена с противоположными по знаку элементами противоположны по сложению. Единицей является число 1.

3. Пространство решений однородного линейного алгебраического уравнения. Рассмотрим множество, состоящее из всевозможных упорядоченных наборов чисел, обращающих алгебраическое уравнение $x + y - z = 0$ в тождество. Два набора (x_1, y_1, z_1) и (x_2, y_2, z_2) будем считать равными, если равны все их соответствующие элементы; сумму и произведение определим как

$$(x_1, y_1, z_1) + (x_2, y_2, z_2) = (x_1 + x_2, y_1 + y_2, z_1 + z_2);$$

$$\alpha(x, y, z) = (\alpha x, \alpha y, \alpha z) .$$

Легко убедиться, что множество решений замкнуто относительно введенных таким образом операций. Действительно, пусть (x, y, z), (x_1, y_1, z_1) и (x_2, y_2, z_2) являются решениями уравнения; тогда

$$(x_1 + x_2) + (y_1 + y_2) - (z_1 + z_2) = (x_1 + y_1 - z_1) + (x_2 + y_2 - z_2) = 0 + 0 = 0 ;$$

$$\alpha x + \alpha y - \alpha z = \alpha(x + y - z) = \alpha \cdot 0 = 0 ;$$

сумма двух решений и произведение решения на число α также являются решениями.

Нулем по сложению является набор $(0,0,0)$ (принадлежность его множеству решений проверяется непосредственно). Два набора с противоположными элементами противоположны друг другу по сложению. Единицей является число 1.

4. Пространство решений однородного дифференциального уравнения. Рассмотрим множество, элементами которого являются заданные на множестве действительных чисел функции[1] $f(x)$, удовлетворяющие *однородному дифференциальному уравнению второго порядка*

$$f''(x) + f(x) = 0.$$

На этом множестве обычным образом можно ввести операции сложения и умножения на число. Замкнутость относительно этих операций следует из линейности операции дифференцирования (производная суммы равна сумме производных, постоянный множитель можно выносить за знак производной).

Нетрудно проверить, что данное множество также является линейным пространством. Нулем по сложению является функция, тождественно равная нулю $f(x)=0$. Две функции, отличающиеся знаком, противоположны по сложению. Единицей является число 1.

5. Координатное пространство. Рассмотрим множество упорядоченных наборов, образованных из n действительных чисел[2]

$$\mathbf{x} = (x_1, x_2, ..., x_n).$$

Два набора $\mathbf{x} = (x_1, x_2, ..., x_n)$ и $\mathbf{y} = (y_1, y_2, ..., y_n)$ будем считать равными тогда и только тогда, когда равны числа, находящиеся на соответствующих местах: $x_i = y_i$, $i = \overline{1,n}$ (запись $i = \overline{1,n}$ означает, что соотношение $x_i = y_i$ должно выполняться для значений индекса, равных 1, 2, …, n; читается «i меняется от 1 до n»).

Суммой упорядоченных наборов назовем упорядоченный набор, на соответствующих местах которого находятся суммы чисел:

$$\mathbf{x} + \mathbf{y} = (x_1, x_2, ..., x_n) + (y_1, y_2, ..., y_n) = (x_1 + y_1, x_2 + y_2, ..., x_n + y_n).$$

Произведением упорядоченного набора на число назовем упорядоченный набор, на соответствующих местах которого находятся произведения:

$$\alpha\mathbf{x} = (\alpha x_1, \alpha x_2, ..., \alpha x_n).$$

Используя свойства арифметических операций над действительными числами, можно доказать, что для введенных таким образом отношения равенства, операций сложения и умножения на число выполняются все аксиомы 1…8.

---

[1] Говорят, что на множестве задана *числовая функция*, если каждому элементу этого множества по определенному правилу сопоставлено некоторое число.

[2] Подобные наборы часто называют *кортежами*.

Докажем, например, коммутативность сложения:

$$(x_1, x_2, ..., x_n) + (y_1, y_2, ..., y_n) = (x_1 + y_1, x_2 + y_2, ..., x_n + y_n) =$$
$$= (y_1 + x_1, y_2 + x_2, ..., y_n + x_n) = (y_1, y_2, ..., y_n) + (x_1, x_2, ..., x_n),$$

Докажем распределительный закон для скаляра:

$$\alpha((x_1, x_2, ..., x_n) + (y_1, y_2, ..., y_n)) = \alpha(x_1 + y_1, x_2 + y_2, ..., x_n + y_n) =$$
$$= (\alpha(x_1 + y_1), \alpha(x_2 + y_2), ..., \alpha(x_n + y_n)) =$$
$$= (\alpha x_1 + \alpha y_1, a x_2 + \alpha y_2, ..., \alpha x_n + \alpha y_n) =$$
$$= (\alpha x_1, \alpha x_2, ..., \alpha x_n) + (\alpha y_1, \alpha y_2, ..., \alpha y_n) =$$
$$= \alpha(x_1, x_2, ..., x_n) + \alpha(y_1, y_2, ..., y_n).$$

Выполнение остальных аксиом доказывается аналогично. Нулем по сложению является набор, состоящий из n нулей:

$$\mathbf{0} = (0, 0, ..., 0).$$

Два набора с противоположными по знаку элементами противоположны по сложению. Единицей является число 1.

Следовательно, множество упорядоченных наборов из n чисел является линейным пространством. Это пространство имеет в линейной алгебре особое значение; его называют *координатным* (или *арифметическим*) пространством.

1.6. Линейная зависимость

Любое линейное пространство определяется приведенной в п. 1.4 системой аксиом. Поэтому неудивительно, что следствия этих аксиом также имеют силу для любого линейного пространства.

Вместе с этим рассмотренные выше примеры показывают, что элементами линейного пространства могут быть объекты различной природы. Природа элементов, как и определения отношения равенства, операций сложения и умножения на число оказывают влияние и на характер рассуждений, необходимых для вывода различных положений, и на «наглядность» полученных результатов.

В координатном пространстве рассуждения удается проводить средствами алгебры, опираясь в основном на свойства действительных чисел. В пространстве векторов-отрезков полученные результаты допускают геометрическую интерпретацию. Поэтому важнейшей задачей является определение способа, позволяющего *взаимно одно-*

значно[1] сопоставить элементы координатного пространства элементам произвольного линейного пространства. Решению этой задачи посвящены пп. 1.6...0.

Пусть $\alpha_1, \alpha_2, \ldots, \alpha_n$ – элементы числового множества. *Линейной комбинацией* конечного числа векторов $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ линейного пространства называется сумма

$$\sum_{k=1}^{n} \alpha_k \mathbf{a}_k = \alpha_1 \mathbf{a}_1 + \alpha_2 \mathbf{a}_2 + \ldots + \alpha_n \mathbf{a}_n .$$

Если вектор $\mathbf{x}$ является линейной комбинацией векторов $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$:

$$\mathbf{x} = \sum_{k=1}^{n} \alpha_k \mathbf{a}_k ,$$

то говорят, что вектор $\mathbf{x}$ *линейно выражается* через векторы $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$.

Числа $\alpha_1, \alpha_2, \ldots, \alpha_n$ называют коэффициентами линейной комбинации. Если все коэффициенты равны нулю, то линейная комбинация называется *тривиальной* (*простейшей*). Если хотя бы один из коэффициентов отличен от нуля, то линейная комбинация называется *нетривиальной*.

Систему векторов $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ линейного пространства называют *линейно зависимой*, если существует нетривиальная линейная комбинация векторов этой системы, равная нулевому вектору:

$$\sum_{k=1}^{n} \alpha_k \mathbf{a}_k = \mathbf{0} .$$

Если такой линейной комбинации нет, то систему векторов $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ называют *линейно независимой*. Система, не содержащая ни одного вектора, по определению считается линейно независимой.

Линейное пространство L называется *n-мерным*, если в нем можно выделить n линейно независимых векторов, но любые $n+1$ векторов линейно зависимы. При этом число n называется *размерностью* пространства L и обозначается

$$n = \dim L .$$

---

[1] Соответствие между элементами множеств A и B называется взаимно однозначным, если:

1) каждому элементу $a \in A$ соответствует единственный элемент $b \in B$;
2) каждому элементу $b \in B$ соответствует единственный элемент $a \in A$.

Линейное пространство называют *бесконечномерным*, если в нем можно выделить любое число линейно независимых векторов. Далее (если не указано обратное) рассматриваются только конечномерные пространства.

Тривиальное линейное пространство имеет нулевую размерность (в таком пространстве система, содержащая хотя бы один – нулевой! – вектор, будет линейно зависимой).

Теорема 1.6.1. *Для того чтобы система векторов* $\mathbf{a}_1$, $\mathbf{a}_2$, …, $\mathbf{a}_n$ *была линейно зависимой, необходимо и достаточно, чтобы хотя бы один из векторов линейно выражался через остальные.*

Докажем необходимость. Пусть векторы $\mathbf{a}_1$, $\mathbf{a}_2$, …, $\mathbf{a}_n$ линейно зависимы:

$$\sum_{k=1}^{n} \alpha_k \mathbf{a}_k = \mathbf{0},$$

причем среди чисел α_1, α_2, …, α_n есть отличные от нуля.

Пусть $\alpha_m \neq 0$; тогда, перенося m-е слагаемое в другую часть и разделив обе части на $-\alpha_m$, получим:

$$\mathbf{a}_m = -\frac{1}{\alpha_m} \sum_{\substack{k=1 \\ k\neq m}}^{n} \alpha_k \mathbf{a}_k .$$

Докажем достаточность. Пусть вектор $\mathbf{a}_m$ линейно выражается через остальные:

$$\mathbf{a}_m = \sum_{\substack{k=1 \\ k\neq m}}^{n} \alpha_k \mathbf{a}_k .$$

Полагая $\alpha_m = -1 \neq 0$ и перенося $\mathbf{a}_m$ в другую часть, получим нетривиальную линейную комбинацию, равную нулевому вектору:

$$\sum_{k=1}^{n} \alpha_k \mathbf{a}_k = \mathbf{0}.$$

Два вектора $\mathbf{a} = (a_1, a_2, ..., a_n)$ и $\mathbf{b} = (b_1, b_2, ..., b_n)$ координатного пространства K_n линейно зависимы тогда и только тогда, когда их координаты пропорциональны:

$$\frac{a_1}{b_1} = \frac{a_2}{b_2} = ... = \frac{a_n}{b_n} = \lambda .$$

Действительно, в этом случае $a_1 = \lambda b_1$, $a_2 = \lambda b_2$, …, $a_n = \lambda b_n$ и вектор **a** линейно выражается через вектор **b**:

$$\mathbf{a} = \lambda \mathbf{b}.$$

Теорема 1.6.2. *Если среди векторов системы* $\mathbf{a}_1$, $\mathbf{a}_2$, …, $\mathbf{a}_n$ *имеется хотя бы один нулевой вектор, то система линейно зависима.*

Действительно, пусть $\mathbf{a}_m = \mathbf{0}$. Тогда этот вектор можно линейно выразить через остальные: $\mathbf{a}_m = \sum\limits_{\substack{k=1 \\ k \neq m}}^{n} 0 \cdot \mathbf{a}_k$.

Теорема 1.6.3. *Если часть системы* $\mathbf{a}_1$, $\mathbf{a}_2$, …, $\mathbf{a}_n$ *линейно зависима, то и вся система линейно зависима.*

Действительно, пусть для первых m векторов выполнено

$$\sum_{k=1}^{m} \alpha_k \mathbf{a}_k = \mathbf{0},$$

причем среди чисел α_1, α_2, …, α_m есть отличные от нуля. Но тогда существует нетривиальная и равная нулю линейная комбинация всех векторов системы:

$$\sum_{k=1}^{m} \alpha_k \mathbf{a}_k + \sum_{k=m+1}^{n} 0 \cdot \mathbf{a}_k = \mathbf{0}.$$

Теорема 1.6.4. *Если система линейно независима, то любая ее часть также линейно независима* (очевидно, что если часть системы линейно зависима, то в силу предыдущей теоремы вся система также должна быть линейно зависима).

В пространстве векторов-отрезков два вектора линейно зависимы тогда и только тогда, когда они коллинеарны. Действительно, пусть векторы **a** и **b** коллинеарны. Если хотя бы один из них – нулевой, то они линейно зависимы. Пусть векторы **a** и **b** ненулевые; тогда из определения операции умножения вектора-отрезка на число следует существование числа λ, такого, что $\mathbf{a} = \lambda \mathbf{b}$. Поэтому имеется нулевая нетривиальная линейная комбинация:

$$1 \cdot \mathbf{a} + (-\lambda)\mathbf{b} = \mathbf{0}.$$

Обратно, пусть существует нулевая нетривиальная линейная комбинация $\alpha\mathbf{a}+\beta\mathbf{b}=\mathbf{0}$. Положим для определенности $\alpha \neq 0$. Тогда

$$\mathbf{a}=-\frac{\beta}{\alpha}\mathbf{b}=\lambda\mathbf{b}.$$

Следовательно, вектор **a** коллинеарен вектору **b**.

Контрольные вопросы и упражнения

1. Что называют линейной комбинацией векторов?
2. Какую линейную комбинацию называют тривиальной (нетривиальной)?
3. Какую систему векторов называют линейно зависимой (линейно независимой)?
4. Что называют размерностью линейного пространства?
5. Сформулируйте необходимое и достаточное условие линейной зависимости векторов.
6. Будут ли в пространстве векторов-отрезков линейно зависимы три вектора, параллельные одной и той же плоскости?

1.7. Базис и координаты

Упорядоченный набор линейно независимых векторов $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ линейного пространства L называется *базисом*, если любой вектор $\mathbf{x}$ пространства L линейно выражается через векторы $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$:

$$\mathbf{x}=\sum_{k=1}^{n} x_k \mathbf{e}_k .$$

Последнее равенство называют *разложением* вектора $\mathbf{x}$ по базису $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$. Коэффициенты линейной комбинации $x_1, x_2, \dots, x_n$ называют *координатами* вектора $\mathbf{x}$ в базисе $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$.

Записывают:

$$\mathbf{x}=\left(x_1, x_2, \dots, x_n\right),$$

или:

$$\mathbf{x}=\left(x_k\right),\ \ k=\overline{1,n}.$$

Теорема 1.7.1. *Если размерность линейного пространства равна n, то любые n линейно независимых векторов этого пространства образуют базис.*

Действительно, пусть система векторов $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ линейно независима. Добавим к ней произвольный вектор $\mathbf{x}$. Тогда полученная система будет содержать $n+1$ вектор и окажется линейно зависимой, т.е. линейная комбинация

$$\alpha_0\mathbf{x}+\sum_{k=1}^{n}\alpha_k\mathbf{e}_k=\mathbf{0}$$

возможна с ненулевыми коэффициентами.

В этой линейной комбинации коэффициент α_0 должен быть отличным от нуля (ибо в противном случае окажется, что векторы $\mathbf{e}_1$, $\mathbf{e}_2$, …, $\mathbf{e}_n$ линейно зависимы). Поэтому

$$\mathbf{x}=-\frac{1}{\alpha_0}\sum_{k=1}^{n}x_k\mathbf{e}_k\ ;$$

вектор $\mathbf{x}$ линейно выражается через векторы $\mathbf{e}_1$, $\mathbf{e}_2$, …, $\mathbf{e}_n$.

Теорема 1.7.2. *Если линейное пространство имеет базис, состоящий из n векторов, то размерность пространства равна n.*

Доказательство. Пусть векторы $\mathbf{e}_1$, $\mathbf{e}_2$, …, $\mathbf{e}_n$ образуют базис. Тогда для произвольного вектора $\mathbf{x}$ справедливо

$$\mathbf{x}=\sum_{k=1}^{n}x_k\mathbf{e}_k\ ,$$

$$\left(\sum_{k=1}^{n}x_k\mathbf{e}_k\right)-\mathbf{x}=\mathbf{0}\,.$$

Следовательно, любая система из $n+1$ векторов линейно зависима. Теорема доказана.

Теорема 1.7.3. *Размерность координатного пространства, элементами которого являются упорядоченные наборы из n чисел, равна n.*

Для доказательства достаточно показать, что n векторов

$$\begin{aligned}
&\mathbf{e}_1=(1,0,\dots,0)\\
&\mathbf{e}_2=(0,1,\dots,0)\\
&\dots\\
&\mathbf{e}_n=(0,0,\dots,1)
\end{aligned}$$

образуют базис. Составим линейную комбинацию

$$\sum_{k=1}^{n}\alpha_k\mathbf{e}_k=(\alpha_1,\alpha_2,\dots,\alpha_n)\,.$$

Равенство этой линейной комбинации нулевому вектору $\mathbf{0}=(0,0,\ldots,0)$ возможно только тогда, когда все коэффициенты α_k равны нулю; следовательно, векторы $\mathbf{e}_1$, $\mathbf{e}_2$, …, $\mathbf{e}_n$ линейно независимы.

Далее, для произвольного вектора $\mathbf{x}=(x_1,x_2,\ldots,x_n)$ имеем разложение

$$(x_1,x_2,\ldots,x_n)=\sum_{k=1}^{n} x_k\mathbf{e}_k\ ;$$

следовательно, векторы $\mathbf{e}_1$, $\mathbf{e}_2$, …, $\mathbf{e}_n$ образуют базис.

Замечание 1.3. Базис, использованный при доказательстве последней теоремы, иногда называют *естественным* базисом n-мерного координатного пространства.

Теорема 1.7.4. *Любая линейно независимая система векторов конечномерного пространства может быть дополнена до базиса.*

Доказательство. Пусть система векторов $\mathbf{e}_1$, $\mathbf{e}_2$, …, $\mathbf{e}_m$ линейно независима. Если она является базисом, то теорема справедлива. Пусть она не является базисом; тогда существует вектор $\mathbf{e}_{m+1}$, который через векторы системы линейно не выражается. Добавляя этот вектор к системе, получим линейно независимую систему, содержащую $m+1$ вектор:

$$\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_m,\mathbf{e}_{m+1}.$$

Если полученная система является базисом, то теорема справедлива; в противном случае вновь добавим вектор, и т.д. В силу конечной размерности пространства указанный процесс не может повторяться бесконечно. Следовательно, на некотором шаге будет получен базис. Теорема доказана.

Теорема 1.7.5. *Разложение вектора* $\mathbf{x}$ *по базису единственно.*

Доказательство. Пусть наряду с разложением

$$\mathbf{x}=\sum_{k=1}^{n} x_k\mathbf{e}_k$$

существует разложение

$$\mathbf{x}=\sum_{k=1}^{n} x'_k\mathbf{e}_k\ .$$

Вычитая второе разложение из первого и меняя порядок суммирования, получим:

$$\sum_{k=1}^{n}\left(x_k - x_k'\right)\mathbf{e}_k = \mathbf{0}.$$

Базисные векторы $\mathbf{e}_1$, $\mathbf{e}_2$, …, $\mathbf{e}_n$ линейно независимы. Поэтому их нулевая линейная комбинация должна быть тривиальной:

$$x_k - x_k' = 0,\ \ k = \overline{1,n},$$

$$x_k = x_k',\ \ k = \overline{1,n}.$$

Теорема 1.7.6. *Все координаты нулевого вектора равны нулю при любом выборе базиса.*

Действительно, пусть в некотором базисе хотя бы одна из координат x_k нулевого вектора отлична от нуля. Тогда

$$x_k\mathbf{e}_k = \mathbf{0},$$

$$\mathbf{e}_k = \mathbf{0},$$

что противоречит определению базиса (наличие среди базисных векторов нулевого вектора должно сделать систему базисных векторов линейно зависимой).

Теорема 1.7.7. *При сложении двух любых векторов линейного пространства их координаты складываются. При умножении любого вектора на число* λ *все координаты этого вектора умножаются на* λ.

Доказательство. Пусть векторы $\mathbf{e}_1$, $\mathbf{e}_2$, …, $\mathbf{e}_n$ составляют базис и пусть в этом базисе координаты векторов $\mathbf{x}$ и $\mathbf{y}$ равны $\left(x_k\right)$ и $\left(y_i\right)$, соответственно.

Из аксиом линейного пространства следует:

$$\mathbf{x} + \mathbf{y} = \sum_{k=1}^{n} x_k\mathbf{e}_k + \sum_{k=1}^{n} y_k\mathbf{e}_k = \sum_{k=1}^{n}\left(x_k\mathbf{e}_k + y_k\mathbf{e}_k\right) = \sum_{k=1}^{n}\left(x_k + y_k\right)\mathbf{e}_k;$$

$$\lambda\mathbf{x} = \lambda\sum_{k=1}^{n} x_k\mathbf{e}_k = \sum_{k=1}^{n}\left(\lambda x_k\right)\mathbf{e}_k.$$

Контрольные вопросы и упражнения

1. Что называют базисом линейного пространства?

2. При каком условии n векторов будут являться базисом n-мерного линейного пространства?

3. Какова взаимосвязь между теоремой 1.7.7 и определением координатного пространства?

1.8. Изоморфизм линейных пространств

Как показано в пп. 1.6 и 1.7, определение базиса позволяет каждому элементу произвольного линейного пространства взаимно однозначно сопоставить элемент координатного пространства. При этом (*независимо* от конкретного определения равенства, суммы и произведения на число в исходном пространстве) координаты суммы будут равны суммам координат; координаты произведения вектора на число будут равны произведениям координат вектора на это число.

Линейные пространства X и Y называются *изоморфными* (точнее, *линейно изоморфными*), если между элементами $\mathbf{x} \in X$ и $\mathbf{y} \in Y$ можно установить взаимно-однозначное соответствие так, что:

1. Если вектор $\mathbf{x}_1 \in X$ соответствует вектору $\mathbf{y}_1 \in Y$ и вектор $\mathbf{x}_2 \in X$ соответствует вектору $\mathbf{y}_2 \in Y$, то сумме $\mathbf{x}_1 + \mathbf{x}_2$ векторов пространства X соответствует сумма $\mathbf{y}_1 + \mathbf{y}_2$ векторов пространства Y.

2. Если вектор $\mathbf{x} \in X$ соответствует вектору $\mathbf{y} \in Y$, то произведению $\alpha\mathbf{x}$ вектора пространства X на число соответствует произведение $\alpha\mathbf{y}$ вектора пространства Y на это же число.

Можно сказать, что линейный изоморфизм пространств − это *взаимно-однозначное соответствие, сохраняющее линейные операции*.

На множестве $\{L_1, L_2, ...\}$, элементами которого являются линейные пространства, отношение изоморфизма является отношением эквивалентности. Рефлексивность очевидна (если вектор $\mathbf{x} \in X$ сопоставить этому же вектору, то пространство X окажется изоморфным самому себе). Симметричность и транзитивность следуют из взаимной однозначности соответствия.

Из определения следует, что линейной комбинации

$$\sum_{k=1}^{n} \alpha_k \mathbf{x}_k$$

векторов пространства X соответствует линейная комбинация

$$\sum_{k=1}^{n}\alpha_k\mathbf{y}_k$$

векторов пространства Y. Поэтому изоморфизм сохраняет линейную зависимость (линейно зависимой системе пространства X соответствует линейно зависимая система пространства Y, и наоборот).

Из сохранения операций следует также, что произведению $0\cdot\mathbf{x}$ пространства X соответствует произведение $0\cdot\mathbf{y}$ пространства Y; следовательно, линейный изоморфизм сохраняет нулевой вектор (нулю по сложению пространства X соответствует ноль пространства Y).

Теорема 1.8.1. *Любое n-мерное линейное пространство изоморфно n-мерному координатному пространству K_n.*

Для доказательства выберем в n-мерном пространстве L базис $\mathbf{e}_1$, $\mathbf{e}_2$, …, $\mathbf{e}_n$ (это всегда можно сделать в силу теорем п. 1.7) и поставим в соответствие вектору

$$\mathbf{x}=\sum_{k=1}^{n}x_k\mathbf{e}_k$$

пространства L его координаты $(x_1,x_2,\ldots,x_n)\in K_n$.

В силу единственности разложения вектора $\mathbf{x}$ по базису такое соответствие будет взаимно однозначным. Далее, координаты суммы векторов равны суммам координат слагаемых, и координаты произведения вектора на число равны произведениям координат на это число:

$$\mathbf{x}+\mathbf{y}=\sum_{k=1}^{n}(x_k+y_k)\mathbf{e}_k\ ,\ \ \alpha\mathbf{x}=\sum_{k=1}^{n}(\alpha x_k)\mathbf{e}_k\ .$$

Поэтому такое соответствие сохраняет линейные операции. Следовательно, оно является линейным изоморфизмом.

Теорема 1.8.2. *Два пространства изоморфны тогда и только тогда, когда их размерность одинакова.*

Докажем необходимость. Пусть пространства изоморфны. Так как изоморфизм сохраняет отношение линейной зависимости, то максимальное число линейно независимых векторов в изоморфных пространствах одно и то же. Следовательно, размерность пространств одинакова.

Докажем достаточность. Пусть размерность пространств одинакова и равна n. Тогда каждое из этих пространств изоморфно n-мерному координатному пространству. Поэтому в силу транзитивности отношения линейного изоморфизма пространства изоморфны друг другу.

Как следствие, *два пространства не изоморфны тогда и только тогда, когда их размерность различна.*

Из доказанной теоремы и транзитивности отношения изоморфизма следует также, что *все пространства равной размерности изоморфны друг другу*.

Любое отношение эквивалентности, заданное на каком-либо множестве M, разбивает это множество на *классы эквивалентности* – множества M_1, M_2, ..., содержащие равные (в смысле определенного отношения эквивалентности) друг другу элементы. Например, обычное отношение равенства на множестве рациональных чисел (чисел вида m/n, где m и n – целые) разбивает это множество на классы эквивалентности, каждый из которых вместе с числом m/n включает все числа вида km/kn, $k \in Z$.

Отношение линейного изоморфизма разбивает все множество линейных пространств на классы, содержащие пространства равной размерности.

Контрольные вопросы и упражнения

1. Какие линейные пространства называют изоморфными?

2. Сформулируйте необходимое и достаточное условие изоморфности двух линейных пространств.

1.9. Подпространства и линейные оболочки

Если подмножество L' линейного пространства L замкнуто относительно введенных в L операций сложения и умножения на число, то L' называют *подпространством* (точнее, *линейным подпространством*) пространства L.

Тривиальное подпространство $\{\mathbf{0}\}$ и подпространство, совпадающее с исходным, называют *несобственными*; остальные подпространства называют *собственными*.

Теорема 1.9.1. *Подпространство L' пространства L само является линейным пространством.*

По определению, L' замкнуто относительно сложения и умножения на число. Так как $L' \subset L$, то для всех элементов множества L' выполнены аксиомы 1, 2, 5, ..., 8. Для доказательства достаточно показать выполнение аксиом 3 (наличие нуля) и 4 (наличие элемента, противоположного по сложению).

Пусть $\mathbf{x} \in L'$. В силу замкнутости L' относительно умножения на число имеем

$$-\mathbf{x} = (-1)\mathbf{x} \in L';$$

аксиома 4 выполнена. Но тогда в силу замкнутости L' относительно сложения

$$\mathbf{0} = \mathbf{x} + (-\mathbf{x}) \in L';$$

аксиома 3 также выполнена. Теорема доказана.

Теорема 1.9.2. *Любая линейно независимая система векторов подпространства L' линейно независима в пространстве L.*

Действительно, пусть система векторов $\mathbf{a}_1$, $\mathbf{a}_2$, …, $\mathbf{a}_m$ подпространства L' линейно независима. Так как L' является подмножеством L и определения линейных операций в L' и L совпадают, то равенство

$$\sum_{k=1}^{m} \alpha_k \mathbf{a}_k = \mathbf{0}$$

в пространстве L имеет смысл и возможно только для тривиальной линейной комбинации.

Теорема 1.9.3. *Размерность подпространства L' не выше размерности исходного пространства L.*

Предположим противное: пусть $n = \dim L$, $m = \dim L'$, причем $m > n$. Тогда в пространстве L' имеются m линейно независимых векторов, которые в пространстве L будут линейно зависимы, что невозможно в силу предыдущей теоремы. Теорема доказана.

Пусть даны пространство L и два его подпространства L_1 и L_2. Тогда для подпространств как для подмножеств множества L имеют смысл операции пересечения и объединения. Следующие теоремы устанавливают свойства множеств, полученных при выполнении этих операций.

Теорема 1.9.4. *Пересечение двух подпространств L_1 и L_2 пространства L также является подпространством.*

Пусть $L_3 = L_1 \cap L_2$, $\{\mathbf{x}, \mathbf{y}\} \subset L_3$. Тогда $\{\mathbf{x}, \mathbf{y}\} \subset L_1$ и $\{\mathbf{x}, \mathbf{y}\} \subset L_2$. Рассматривая векторы $\mathbf{x}$ и $\mathbf{y}$ как элементы подпространства L_1, приходим к заключению что сумма $\mathbf{x} + \mathbf{y}$ и произведение $\alpha\mathbf{x}$ принадлежат L_1. Рассматривая эти векторы как элементы подпространства L_2, находим, что сумма $\mathbf{x} + \mathbf{y}$ и произведение $\alpha\mathbf{x}$ принадлежат также и подпространству L_2. Но тогда сумма и произведение принадлежат пересечению $L_3 = L_1 \cap L_2$; следовательно, L_3 является подпространством.

Теорема 1.9.5. *Размерность пересечения двух пространств* $L_1 \cap L_2$ *не выше наибольшей из размерностей* $\dim L_1$, $\dim L_2$.

Достаточно заметить, что пересечение $L_1 \cap L_2$ является подпространством каждого из пространств L_1 и L_2 (в силу предыдущей теоремы и определения операции пересечения множеств); тогда его размерность не может превышать ни размерности L_1, ни размерности L_2. Теорема доказана.

Суммой (объединением) подпространств L_1 и L_2 пространства L называется множество $L_1 + L_2$ всех векторов, представимых в виде

$$\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2,$$

где $\mathbf{x}_1 \in L_1$, $\mathbf{x}_2 \in L_2$.

Теорема 1.9.6. *Сумма двух подпространств L_1 и L_2 пространства L также является подпространством.*

Пусть $L_3 = L_1 + L_2$, $\{\mathbf{x}, \mathbf{y}\} \subset L_3$; тогда $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$ и $\mathbf{y} = \mathbf{y}_1 + \mathbf{y}_2$, где $\{\mathbf{x}_1, \mathbf{y}_1\} \subset L_1$ и $\{\mathbf{x}_2, \mathbf{y}_2\} \subset L_2$.

Прежде всего заметим, что множество L_3 является подмножеством множества L. Действительно, если $\mathbf{x}_1 \in L_1$ и $\mathbf{x}_2 \in L_2$, то оба этих вектора принадлежат и множеству L; поэтому в силу замкнутости L относительно сложения сумма $\mathbf{x}_1 + \mathbf{x}_2$ также принадлежит L.

Далее,

$$\mathbf{x} + \mathbf{y} = (\mathbf{x}_1 + \mathbf{x}_2) + (\mathbf{y}_1 + \mathbf{y}_2) = (\mathbf{x}_1 + \mathbf{y}_1) + (\mathbf{x}_2 + \mathbf{y}_2) \in L_3,$$

$$\alpha\mathbf{x} = \alpha(\mathbf{x}_1 + \mathbf{x}_2) = \alpha\mathbf{x}_1 + \alpha\mathbf{x}_2 \in L_3;$$

множество L_3 замкнуто относительно сложения и умножения на число. Следовательно, L_3 является подпространством.

Теорема 1.9.7. *Размерность суммы двух пространств* $L_1 + L_2$ *не ниже наибольшей из размерностей* $\dim L_1$, $\dim L_2$.

Пусть $\mathbf{x}_1 \in L_1$, $\mathbf{x}_2 \in L_2$. Так как $\mathbf{0} \in L_1$ и $\mathbf{0} \in L_2$, то

$$\mathbf{x}_1 = \mathbf{x}_1 + \mathbf{0} \in L_1 + L_2,$$

$$\mathbf{x}_2 = \mathbf{0} + \mathbf{x}_2 \in L_1 + L_2;$$

следовательно, каждое из пространств L_1 и L_2 является подпространством суммы $L_1 + L_2$. Но тогда ни размерность L_1, ни размерность L_2 не может превышать размерность $L_1 + L_2$.

Если для любого $\mathbf{x} \in L_1 + L_2$ представление в виде

$$\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$$

единственно, то сумма пространств L_1 и L_2 называется *прямой суммой* и обозначается $L_1 \oplus L_2$.

Теорема 1.9.8. *Сумма $L_1 + L_2$ является прямой суммой тогда и только тогда, когда пересечение $L_1 \cap L_2$ равно тривиальному пространству* (т.е. когда L_1 и L_2 не имеют общих элементов, за исключением нуля).

Докажем необходимость. Пусть для любого $\mathbf{x} \in L_1 + L_2$ представление в виде

$$\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$$

единственно.

Предположим, что существует вектор $\mathbf{y} \in L_1 \cap L_2$. Очевидно, он также принадлежит и сумме $L_1 + L_2$ (см. предыдущую теорему), поэтому его представление имеет вид

$$\mathbf{y} = \mathbf{y} + \mathbf{y}$$

(здесь учтено, что $\mathbf{y} \in L_1$ и $\mathbf{y} \in L_2$); прибавляя к левой и правой части вектор, противоположный $\mathbf{y}$ по сложению, получим:

$$\mathbf{y} = \mathbf{0}.$$

Докажем достаточность. Пусть пространства L_1 и L_2 не имеют общих элементов, за исключением нуля. Предположим, что для вектора $\mathbf{x} \in L_1 + L_2$ существуют два представления $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$ и $\mathbf{x} = \mathbf{x}_1' + \mathbf{x}_2'$. Вычитая второе разложение из первого, получим:

$$(\mathbf{x}_1 - \mathbf{x}_1') + (\mathbf{x}_2 - \mathbf{x}_2') = \mathbf{0},$$

$$\mathbf{x}_1 - \mathbf{x}_1' = \mathbf{x}_2 - \mathbf{x}_2'.$$

В левой части последнего равенства находится вектор, принадлежащий пространству L_1, а в правой – вектор, принадлежащий L_2. Так как единственным общим вектором этих пространств является нулевой, то

$$\mathbf{x}_1 - \mathbf{x}_1' = \mathbf{x}_2 - \mathbf{x}_2' = \mathbf{0},$$

откуда $\mathbf{x}_1 = \mathbf{x}'$ и $\mathbf{x}_2 = \mathbf{x}_2'$; представления $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$ и $\mathbf{x} = \mathbf{x}_1' + \mathbf{x}_2'$ совпадают. Теорема доказана.

Для дальнейшего изложения вопросов, связанных с подпространствами, удобно ввести понятие *линейной оболочки*.

Пусть в линейном пространстве выбрана система векторов $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$. Множество всевозможных линейных комбинаций

$$\mathbf{x} = \sum_{k=1}^{m} \alpha_k \mathbf{a}_k$$

называется *линейной оболочкой* данной системы и обозначается[1]

$$L(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m).$$

Линейная оболочка пустого множества векторов по определению совпадает с линейной оболочкой $L(\mathbf{0})$ и равна тривиальному пространству $\{\mathbf{0}\}$.

Любая линейно независимая система векторов может быть дополнена до базиса. Следующая теорема показывает, что линейно зависимая система может быть *сокращена* до линейно независимой так, что ее линейная оболочка останется неизменной.

Теорема 1.9.9. *Из любой конечной системы векторов может быть выделена линейно независимая подсистема, линейная оболочка которой совпадает с линейной оболочкой исходной системы.*

Если система векторов $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ линейно независима, то теорема справедлива. Пусть система $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ линейно зависима. Тогда существует вектор $\mathbf{a}_i$, являющийся линейной комбинацией остальных векторов системы:

$$\mathbf{a}_i = \sum_{\substack{k=1 \\ k \neq i}}^{m} \beta_k \mathbf{a}_k .$$

При удалении этого вектора из системы ее линейная оболочка не изменится, так как любой вектор

$$\mathbf{x} = \sum_{k=1}^{m} \alpha_k \mathbf{a}_k$$

[1] Иногда используют обозначение $[\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m]$ или $\mathrm{span}(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m)$.

можно представить в виде линейной комбинации векторов системы $\mathbf{a}_1$, $\mathbf{a}_2$, …, $\mathbf{a}_{i-1}$, $\mathbf{a}_{i+1}$, …, $\mathbf{a}_m$, не содержащей вектора $\mathbf{a}_i$:

$$\mathbf{x} = \sum_{k=1}^{m} \alpha_k \mathbf{a}_k = \sum_{\substack{k=1 \\ k \neq i}}^{m} \alpha_k \mathbf{a}_k + \alpha_i \sum_{\substack{k=1 \\ k \neq i}}^{m} \beta_k \mathbf{a}_k = \sum_{\substack{k=1 \\ k \neq i}}^{m} \left(\alpha_k + \alpha_i \beta_k\right) \mathbf{a}_k .$$

Если система $\mathbf{a}_1$, $\mathbf{a}_2$, …, $\mathbf{a}_{i-1}$, $\mathbf{a}_{i+1}$, …, $\mathbf{a}_m$ линейно независима, то теорема справедлива. В противном случае следует вновь удалить вектор, являющийся линейной комбинацией остальных, и т.д. В силу конечности системы $\mathbf{a}_1$, $\mathbf{a}_2$, …, $\mathbf{a}_m$ процесс последовательного удаления на очередном этапе приведет к линейно независимой подсистеме (возможно, не содержащей ни одного вектора).

Введенное определение позволяет сказать, что *линейное пространство есть линейная оболочка любого из своих базисов*. Две следующие теоремы устанавливают строгое соответствие между понятиями *подпространства* и *линейной оболочки*.

Теорема 1.9.10. *Линейная оболочка любой системы векторов пространства L является подпространством пространства L.*

Прежде всего заметим, что линейная оболочка $L(\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_m)$ является подмножеством пространства L (в силу замкнутости L по сложению и умножению на число любая линейная комбинация векторов пространства L также является вектором этого пространства).

Пусть векторы $\mathbf{x}$ и $\mathbf{y}$ принадлежат линейной оболочке $L(\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_m)$:

$$\mathbf{x} = \sum_{k=1}^{m} \alpha_k \mathbf{a}_k ,$$

$$\mathbf{y} = \sum_{k=1}^{m} \beta_k \mathbf{a}_k$$

(такие векторы всегда можно найти, так как любая линейная оболочка содержит, по крайней мере, нулевой вектор). Тогда:

$$\mathbf{x} + \mathbf{y} = \sum_{k=1}^{m} \alpha_k \mathbf{a}_k + \sum_{k=1}^{m} \beta_k \mathbf{a}_k = \sum_{k=1}^{m} \left(\alpha_k + \beta_k\right) \mathbf{a}_k \in L(\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_m) ,$$

$$\lambda \mathbf{x} = \lambda \sum_{k=1}^{m} \alpha_k \mathbf{a}_k = \sum_{k=1}^{m} \left(\lambda \alpha_k\right) \mathbf{a}_k \in L(\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_m) ;$$

линейная оболочка замкнута относительно сложения и умножения на число. Следовательно, она является подпространством.

Как следствие двух предыдущих теорем, отметим, что *размерность линейной оболочки* $L(\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_m)$ *равна максимальному числу линейно независимых векторов, входящих в систему* $\mathbf{a}_1$, $\mathbf{a}_2$, …, $\mathbf{a}_m$.

Теорема 1.9.11. *В конечномерном пространстве любое подпространство есть линейная оболочка некоторой системы векторов.*

Действительно, в конечномерном пространстве любое подпространство также конечномерно (имеет размерность, не превышающую размерность исходного пространства). Если подпространство тривиально, то оно является линейной оболочкой $L(\mathbf{0})$ своего нулевого вектора и теорема справедлива.

Пусть подпространство нетривиально. Тогда оно имеет базис

$$\mathbf{a}_1,\ \mathbf{a}_2,\ \ldots,\ \mathbf{a}_m$$

и включает такие и только такие векторы, для которых

$$\mathbf{x} = \sum_{k=1}^{m} \alpha_k \mathbf{a}_k\ ,$$

т.е. является линейной оболочкой своего базиса.

Теорема 1.9.12. *Пусть в пространстве L имеются подпространства* L_k *и* L_l *размерностей* $\dim L_k = k$ *и* $\dim L_l = l$. *Тогда если размерность их пересечения равна m, то их сумма имеет размерность* $k + l - m$.

Доказательство. Выберем в пространстве $L_m = L_k \cap L_l$ базис

$$\mathbf{e}_1,\ \mathbf{e}_2,\ \ldots,\ \mathbf{e}_m \text{ (здесь } m \le k\ ,\ m \le l\text{)}.$$

Любая линейно независимая система векторов может быть дополнена до базиса (см. п. 1.7). Дополним систему $\mathbf{e}_1$, $\mathbf{e}_2$, …, $\mathbf{e}_m$ до базисов в пространствах L_k и L_l; пусть

$$\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_m, \mathbf{e}_{m+1}, ..., \mathbf{e}_k \ \text{– базис в } L_k;$$

$$\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_m, \mathbf{e}'_{m+1}, ..., \mathbf{e}'_l \text{– базис в } L_l.$$

Достаточно показать, что система из $k + l - m$ векторов

$$\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_m, \mathbf{e}_{m+1}, ..., \mathbf{e}_k, \mathbf{e}'_{m+1}, ..., \mathbf{e}'_l \qquad \text{(a)}$$

является базисом суммы $L_1 + L_2$. Пусть $\mathbf{x}_1 \in L_k$, $\mathbf{x}_2 \in L_l$. Имеем разложения:

$$\mathbf{x}_1 = \sum_{i=1}^{k} \alpha_i \mathbf{e}_i ,$$

$$\mathbf{x}_2 = \sum_{i=1}^{m} \beta_i \mathbf{e}_i + \sum_{i=m+1}^{l} \beta_i \mathbf{e}'_i .$$

Тогда вектор $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2 \in L_1 + L_2$ может быть разложен по векторам системы (a):

$$\mathbf{x} = \sum_{i=1}^{k} \alpha_i \mathbf{e}_i + \sum_{i=1}^{l} \beta_i \mathbf{e}_i = \sum_{i=1}^{m} (\alpha_i + \beta_i) \mathbf{e}_i + \sum_{i=m+1}^{k} \alpha_i \mathbf{e}_i + \sum_{i=m+1}^{l} \beta_i \mathbf{e}'_i .$$

Докажем, что векторы системы (a) линейно независимы. Первые k ее векторов линейно независимы, так как они образуют базис пространства L_k. Остальные $l - m$ векторов линейно независимы, так как входят в базис пространства L_l. Составим из векторов системы (a) линейную комбинацию

$$\sum_{i=1}^{k} \lambda_i \mathbf{e}_i + \sum_{j=m+1}^{l} \xi_j \mathbf{e}'_j = \mathbf{0}. \qquad \text{(b)}$$

Из последнего соотношения:

$$\sum_{i=1}^{k} \lambda_i \mathbf{e}_i = - \sum_{j=m+1}^{l} \xi_j \mathbf{e}'_j ,$$

поэтому в левой части (b) обе суммы или равны нулю по отдельности, или обе отличны от нуля. В первом случае из линейной независимости векторов $\mathbf{e}_1, \mathbf{e}_2, , \mathbf{e}_m, \mathbf{e}_{m+1}, ..., \mathbf{e}_k$ и $\mathbf{e}'_{m+1}, ..., \mathbf{e}'_l$ следует: $\lambda_i = 0$, $\xi_j = 0$. Линейная комбинация (b) тривиальна.

Предположим, что обе суммы в левой части (b) отличны от нуля. Обозначим

$$\mathbf{y} = \sum_{i=1}^{k} \lambda_i \mathbf{e}_i ;$$

вектор $\mathbf{y}$ может быть разложен по базису пространства L_k, следовательно, $\mathbf{y} \in L_k$.

С другой стороны,

$$\mathbf{y} = -\sum_{j=m+1}^{l} \xi_j \mathbf{e}'_j = \sum_{j=m+1}^{l} \left(-\xi_j\right) \mathbf{e}'_j ,$$

откуда $\mathbf{y} \in L_l$. Поэтому $\mathbf{y} \in L_k \cap L_l$ и может быть разложен базису этого пространства:

$$\mathbf{y} = \sum_{j=1}^{m} \xi_j \mathbf{e}_j .$$

Таким образом

$$\mathbf{y} = \sum_{j=1}^{m} \xi_j \mathbf{e}_j = \sum_{j=m+1}^{l} \left(-\xi_j\right) \mathbf{e}'_j ,$$

$$\sum_{j=1}^{m} \xi_j \mathbf{e}_j + \sum_{j=m+1}^{l} \left(-\xi_j\right) \mathbf{e}'_j = \mathbf{0} .$$

Но векторы $\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_m, \mathbf{e}'_{m+1}, ..., \mathbf{e}'_l$ линейно независимы, так как составляют базис пространства L_l. Поэтому все $\xi_j = 0$ и вторая сумма в левой части (b) оказывается равной нулю. Полученное противоречие свидетельствует, что векторы системы (a) линейно независимы. Теорема доказана.

Утверждение доказанной теоремы можно сформулировать иначе: *размерность суммы равна сумме размерностей за вычетом размерности пересечения*.

Как следствие, *размерность прямой суммы двух подпространств равна сумме размерностей слагаемых*.

Пусть система векторов $\mathbf{b}_1$, $\mathbf{b}_2$, …, $\mathbf{b}_n$ является базисом пространства L, а система $\mathbf{a}_1$, $\mathbf{a}_2$, …, $\mathbf{a}_m$ является базисом подпространства $L' \subset L$ (изложенное выше показывает, что m не должно превышать n). Если вектор $\mathbf{x} \in L'$, то этот вектор может быть разложен по векторам $\mathbf{a}_1$, $\mathbf{a}_2$, …, $\mathbf{a}_m$:

$$\mathbf{x} = \sum_{k=1}^{m} \alpha_k \mathbf{a}_k .$$

Пусть векторы $\mathbf{a}_k$, $k=\overline{1,m}$ в базисе $\mathbf{b}_1$, $\mathbf{b}_2$, …, $\mathbf{b}_n$ имеют координаты

$$\mathbf{a}_k = (\beta_{1k}, \beta_{2k}, ..., \beta_{nk}),\ k=\overline{1,m},$$

т.е.

$$\mathbf{a}_k = \sum_{i=1}^{n} \beta_{ik} \mathbf{b}_i,\ k=\overline{1,m}.$$

Тогда

$$\mathbf{x} = \sum_{k=1}^{m} \alpha_k \mathbf{a}_k = \sum_{k=1}^{m} \alpha_k \sum_{i=1}^{n} \beta_{ik} \mathbf{b}_i = \sum_{i=1}^{n} \left(\sum_{k=1}^{m} \beta_{ik} \alpha_k \right) \mathbf{b}_i .$$

С другой стороны, $\mathbf{x} \in L$, поэтому его можно разложить по векторам базиса $\mathbf{b}_1$, $\mathbf{b}_2$, …, $\mathbf{b}_n$:

$$\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{b}_i .$$

Сравнивая два последних разложения, приходим к заключению

$$x_i = \sum_{k=1}^{m} \beta_{ik} \alpha_k,\ i=\overline{1,n}.$$

В развернутой форме:

$$\begin{cases} x_1 = \alpha_1\beta_{11} + \alpha_2\beta_{12} + ... + \alpha_m\beta_{1m} \\ x_2 = \alpha_1\beta_{21} + \alpha_2\beta_{22} + ... + \alpha_m\beta_{2m} \\ ... \\ x_n = \alpha_1\beta_{n1} + \alpha_2\beta_{n2} + ... + \alpha_m\beta_{nm} \end{cases}.$$

Последняя система называется *параметрическим уравнением подпространства*.

Обратимся к n-мерному координатному пространству

$$K_n = \{(x_1, x_2, ..., x_n)\}.$$

Его подпространством является, в частности, множество тех наборов $(x_1, x_2, ..., x_n)$, на элементы которых наложены ограничения в виде однородных линейных алгебраических уравнений. Например, требуя $x_n = 0$, получим подмножество:

$$K_{n-1} = \{(x_1, x_2, ..., x_{n-1}, 0)\}.$$

Это подмножество является подпространством, так как

$$(x_1, x_2, ..., x_{n-1}, 0) + (y_1, y_2, ..., y_{n-1}, 0) =$$

$$= (x_1 + y_1, x_2 + y_2, ..., x_{n-1} + y_{n-1}, 0) \in K_{n-1};$$

$$\alpha(x_1, x_2, ..., x_{n-1}, 0) = (\alpha x_1, \alpha x_2, ..., \alpha x_{n-1}, 0) \in K_{n-1}.$$

Если на элементы набора $(x_1, x_2, ..., x_n)$ наложено ограничение

$$A_1x_1 + A_2x_2 + ... + A_nx_n = 0$$

(среди чисел A_1, A_2, …, A_n должны быть отличные от нуля), то полученное подпространство

$$K_{n-1} = \left\{\left(x_1, x_2, ..., x_{n-1}, -\frac{1}{A_n}(A_1x_1 + A_2x_2 + ... + A_{n-1}x_{n-1})\right)\right\}$$

называется $(n-1)$-*мерной гиперплоскостью* (здесь предполагается, что $A_n \neq 0$).

При этом само ограничение

$$A_1x_1 + A_2x_2 + ... + A_nx_n = 0$$

называется *уравнением гиперплоскости*.

Рассмотрим одномерное подпространство K_1 пространства K_n. Пусть вектор $\mathbf{d}$ является базисом K_1. Тогда любой другой вектор $\mathbf{x} \in K_1$ может быть представлен в виде $\mathbf{x} = t\mathbf{d}$.

Если в пространстве K_n вектор $\mathbf{d}$ имеет координаты d_1, d_2, …, d_n, то координаты вектора $\mathbf{x}$ в пространстве K_n равны

$$\begin{cases} x_1 = td_1 \\ x_2 = td_2 \\ ... \\ x_n = td_n \end{cases}.$$

Последняя система называется *параметрическим уравнением прямой*. Число t (единственную координату вектора $\mathbf{x}$ в пространстве K_1) называют *параметром*. В пространстве K_n вектор $\mathbf{d}$ называют *направляющим вектором прямой*.

Контрольные вопросы и упражнения

1. Что называют подпространством линейного пространства?

2. Будет ли двумерное координатное пространство – множество упорядоченных троек $(\alpha, \beta, 0)$ действительных чисел – подпростран-

ством трехмерного (множества упорядоченных троек (α, β, γ) действительных чисел)?

3. Что называют линейной оболочкой системы векторов?

4. Даны два вектора $(\alpha, 0)$ и $(0, \beta)$, принадлежащие одномерным координатным пространствам L_x и L_y, соответственно (здесь α и β – произвольные и не связанные между собой действительные числа). Чем будет являться линейная оболочка этих векторов? Какова размерность линейной оболочки? Что будет представлять собой пересечение пространств L_x и L_y?

Глава 2. Матричная алгебра

Выше было отмечено, что предметом линейной алгебры являются отношения между элементами линейных пространств. Естественная запись линейных операций требует привлечения индексов и символа суммы; использование подобной записи существенно сокращает громоздкость выкладок.

Поэтому целесообразно определить объект, правила действий над которым таковы, что ни индексация, ни символ суммы при записи операций не требуются совсем.

2.1. Матрица

Матрицей размера $m{\times}n$ называется прямоугольная таблица объектов произвольной природы, содержащая m строк и n столбцов. Если указанные объекты являются числами (действительными или комплексными), то матрица называется *числовой*; далее будут рассматриватся только числовые матрицы.

Матрицы мы будем обозначать полужирными прописными буквами:

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

или, в сокращенной записи:

$$\mathbf{A} = \left(a_{ij}\right),\ i = \overline{1,m},\ \ j = \overline{1,n}.$$

Для обозначения матриц можно использовать не только круглые, но двойные прямые скобки:

$$\mathbf{A} = \left\| a_{ij} \right\|.$$

Числа a_{ij} называются *элементами* матрицы. Индексы i, j означают, что элемент расположен на пересечении строки с номером i и столбца с номером j. Часто номер строки удобно записывать как верхний индекс[1] (не следует путать его с показателем степени); в этих обозначениях

$$\mathbf{A} = \left(a^i_j\right).$$

Если число строк m совпадает с числом столбцов n, то матрица (размера $n \times n$) называется *квадратной*, а число n называется *порядкомя*матрицы. Если $m \neq n$, то матрица называется *прямоугольной* (в этом случае числа m и n можно называть ее *порядками*).

Главной диагональю (или просто *диагональю*) квадратной (или прямоугольной) матрицы называется диагональ, на которой находятся элементы a_{ii}, индексы строки и столбца которых одинаковы. Другая диагональ квадратной матрицы называется *побочной* (для прямоугольной матрицы понятие побочной диагонали не определено). Элементы главной диагонали называются *диагональными*, все остальные элементы называются *внедиагональными*.

Матрица называется *д, агональной*, если все ее внедиагональные элементы равны нулю. Если диагональные элементы квадратной диагональной матрицы равны между собой, то матрица называется *скалярной*.

Если все диагональные элементы скалярной матрицы n-го порядка равны единице, то такая матрица

$$\mathbf{E}_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & & & \\ 0 & 0 & \dots & 1 \end{pmatrix},$$

называется *единичной матрицей* порядка n.

Квадратная матрица называется *верхнейя*(или*яправой*) треугольной, если в ней равны нулю все элементы, расположенные ниже главной диагонали. Квадратная матрица называется *н, жнейя*(илия *левой*) треугольной, если равны нулю все элементы, расположенные выше главной диагонали. Диагональная матрица является одновременно и верхней, и нижней треугольной. Если диагональные эле-

[1] Мы по возможности будем избегать подобной записи, применяя ее только в тех случаях, когда она облегчает правильную интерпретацию выкладок.

менты верхней треугольной матрицы равны нулю, то матрица называется *наддиагональной*; если равны нулю диагональные элементы нижней треугольной матрицы, то матрица называется *поддиагональной*.

Если в квадратной матрице n-го порядка отличны от нуля только те элементы, которые расположены «рядом» с главной диагональю (т.е. $a_{ij} = 0$ при $|i-j| > k$; $0 \le k \le n-1$), то матрицу называют *ленточной*. Примером такой матрицы является диагональная (для ее элементов при $|i-j| > 0$ выполнено условие $a_{ij} = 0$).

В частности, если элементы a_{ij} матрицы $\mathbf{A}$ равны нулю при $|i-j| > 1$, то матрицу называют *трехдиагональной*. Такая матрица имеет вид:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & 0 & 0 & \dots & 0 \\ a_{21} & a_{22} & a_{23} & 0 & \dots & 0 \\ 0 & a_{32} & a_{33} & a_{34} & \dots & 0 \\ \dots & & & & & \\ 0 & \dots & 0 & a_{n-1\,n-2} & a_{n-1\,n-1} & a_{n-1\,n} \\ 0 & \dots & 0 & 0 & a_{n\,n-1} & a_{nn} \end{pmatrix}.$$

Аналогично определяется *пятидиагональная* матрица.

Пусть матрица $\mathbf{A} = (a_{ij})$ имеет размер $m \times n$. Матрица $\mathbf{A}^T$ размера $n \times m$ называется *транспонированной* по отношению к матрице $\mathbf{A}$, если элементы i-го столбца матрицы $\mathbf{A}^T$ равны соответствующим элементам i-й строки матрицы $\mathbf{A}$:

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

Такая замена строк столбцами называется *транспонированиемя* матрицы. Очевидно, $\left(\mathbf{A}^T\right)^T = \mathbf{A}$.

Квадратная матрица порядка n называется *симметрической* (или *симметричной*), если она совпадает со своей транспонированной: $\mathbf{A}^T = \mathbf{A}$. Для такой матрицы выполнено условие $a_{ij} = a_{ji}$.

Квадратная матрица называется *кососимметрической* (или *антисимметричной*), если для ее элементов выполнено условие $a_{ij} = -a_{ji}$. В частности, $a_{ii} = -a_{ii}$; диагональные элементы кососимметрической матрицы по необходимости равны нулю.

Матрица размера $m\times 1$, состоящая из одного столбца, называется *вектор-столбцом* (или просто *столбцом*) длины m. Матрица размера $1\times n$, состоящая из одной строки, называется *вектор-строкой* (или просто *строкой*) длины n. Строки и столбцы можно обозначать так же, как и векторы (полужирными строчными буквами).

Матрицу $\mathbf{A} = \left(a_j^i\right)$ размера $m\times n$ можно понимать как столбец

$$\mathbf{A} = \begin{pmatrix} \mathbf{a}^1 \\ \mathbf{a}^2 \\ \ldots \\ \mathbf{a}^m \end{pmatrix}$$

из m строк

$$\mathbf{a}^i = \begin{pmatrix} a_1^i & a_2^i & \ldots & a_n^i \end{pmatrix},\ i = \overline{1,m}.$$

Эту же матрицу можно понимать как строку

$$\mathbf{A} = \begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \ldots & \mathbf{a}_n \end{pmatrix}$$

из n столбцов

$$\mathbf{a}_j = \begin{pmatrix} a_j^1 \\ a_j^2 \\ \ldots \\ a_j^m \end{pmatrix},\ \ j = \overline{1,n}.$$

Столбцы и строки называют *рядами* матрицы. Для сокращения записи столбцы часто представляют в виде $\mathbf{a}_j = \begin{pmatrix} a_j^1 & a_j^2 & \ldots & a_j^m \end{pmatrix}^T$, или $\mathbf{a}_j = \mathrm{Col}\begin{pmatrix} a_j^1 & a_j^2 & \ldots & a_j^m \end{pmatrix}$.

Пусть дана матрица $\mathbf{A} = \left(a_{ij}\right)$. Выберем в ней r строк с номерами $i_1, i_2, \ldots, i_r$ и s столбцов с номерами $j_1, j_2, \ldots, j_s$ (предполагается, что номера строк и столбцов выбраны в порядке возрастания). Матрица

$$\mathbf{A}' = \begin{pmatrix} a_{i_1 j_1} & a_{i_1 j_2} & \ldots & a_{i_1 j_s} \\ a_{i_2 j_1} & a_{i_2 j_2} & \ldots & a_{i_2 j_s} \\ \ldots & & & \\ a_{i_r j_1} & a_{i_r j_2} & \ldots & a_{i_r j_s} \end{pmatrix}$$

размера $r \times s$, образованная элементами, расположенными на пересечении выбранных строк и столбцов, называется *подматрицей* матрицы $\mathbf{A}$.

Если матрица разделена на несколько подматриц, занимающих смежные строки и столбцы, то говорят, что матрица является *блочной*. Например, матрицу

$$\mathbf{A} = \begin{pmatrix} 1 & 3 & -1 \\ 2 & 4 & -2 \\ 7 & 8 & 9 \end{pmatrix}$$

можно разбить на четыре блока

$$\mathbf{A}_1 = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}, \ \mathbf{A}_2 = \begin{pmatrix} -1 \\ -2 \end{pmatrix}, \ \mathbf{A}_3 = \begin{pmatrix} 7 & 8 \end{pmatrix}, \ \mathbf{A}_4 = \begin{pmatrix} 9 \end{pmatrix},$$

и записать в виде

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_2 \\ \mathbf{A}_3 & \mathbf{A}_4 \end{pmatrix}.$$

При транспонировании блочной матрицы в ней транспонируется каждый блок и затем транспонируются сами блоки. Например:

$$\begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_2 \\ \mathbf{A}_3 & \mathbf{A}_4 \end{pmatrix}^T = \begin{pmatrix} \mathbf{A}_1^T & \mathbf{A}_3^T \\ \mathbf{A}_2^T & \mathbf{A}_4^T \end{pmatrix}.$$

Контрольные вопросы и упражнения

1. Что называют матрицей?

2. На пересечении каких рядов матрицы $\mathbf{A}$ расположен элемент a_{37}?

3. Что называют главной и побочной диагоналями матрицы?

4. Можно ли сказать, что элементы a_{21} и a_{12} матрицы **A** второго порядка являются ее диагональными элементами?

5. Является ли матрица

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

диагональной? Является ли она скалярной?

6. Какая матрица называется верхней треугольной? Нижней треугольной?

7. В чем различие между верхней треугольной и наддиагональной матрицами?

8. Является ли матрица

$$\begin{pmatrix} 1 & -2 & 4 \\ 2 & 6 & -1 \\ 0 & 3 & -4 \end{pmatrix}$$

трехдиагональной? Является ли она пятидиагональной?

9. Что называют транспонированием матрицы?

10. Является ли матрица

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

симметрической? Является ли она кососимметрической?

11. Можно ли в матрице

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

выделить симметрическую подматрицу? Можно ли в ней выделить кососимметрическую подматрицу?

12. Можно ли верхнюю треугольную матрицу

$$\begin{pmatrix} 1 & 2 & 4 & 0 \\ 0 & 1 & 5 & 6 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

разбить на четыре блока так, чтобы каждый из них сам был треугольным? Единственно ли такое разбиение? При каком числе блоков указанное разбиение заведомо возможно и единственно?

2.2. Линейное пространство матриц

Отношение равенства, операции сложения и умножения на число для матриц определяются точно так же, как и для векторов координатного пространства.

Две матрицы $\mathbf{A} = (a_{ij})$ и $\mathbf{B} = (b_{ij})$ одинаковых размеров $m \times n$ равны тогда и только тогда, когда равны все их соответствующие элементы:

$$a_{ij} = b_{ij}, \ i = \overline{1, m}, \ j = \overline{1, n}.$$

Как и в случае векторов координатного пространства, следствием подобного «поэлементного» определения является выполнение для отношения равенства матриц всех свойств отношения равенства их элементов. Поэтому отношение равенства числовых матриц есть отношение эквивалентности.

Суммой двух матриц **A** и **B** одинаковых размеров $m \times n$ называется матрица **C** того же размера, элементы которой равны суммам соответствующих элементов матриц **A** и **B**:

$$c_{ij} = a_{ij} + b_{ij}, \ i = \overline{1, m}, \ j = \overline{1, n}.$$

Произведением матрицы **A** на число λ называется матрица **C**, элементы которой равны произведениям элементов матрицы **A** на это число:

$$c_{ij} = \lambda a_{ij}, \ i = \overline{1, m}, \ j = \overline{1, n}.$$

Сложение блочных матриц и умножение блочной матрицы на число можно выполнять поблочно (очевидно, что слагаемые должны быть разбиты на блоки равных размеров).

Например:

$$\begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_2 \\ \mathbf{A}_3 & \mathbf{A}_4 \end{pmatrix} + \begin{pmatrix} \mathbf{B}_1 & \mathbf{B}_2 \\ \mathbf{B}_3 & \mathbf{B}_4 \end{pmatrix} = \begin{pmatrix} \mathbf{A}_1 + \mathbf{B}_1 & \mathbf{A}_2 + \mathbf{B}_2 \\ \mathbf{A}_3 + \mathbf{B}_3 & \mathbf{A}_4 + \mathbf{B}_4 \end{pmatrix};$$

$$\lambda \begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_2 \\ \mathbf{A}_3 & \mathbf{A}_4 \end{pmatrix} = \begin{pmatrix} \lambda\mathbf{A}_1 & \lambda\mathbf{A}_2 \\ \lambda\mathbf{A}_3 & \lambda\mathbf{A}_4 \end{pmatrix}.$$

Для введенных операций сложения и умножения на число выполняются все восемь аксиом линейного пространства. Таким образом, *множество всех матриц размера m×n является линейным пространством*[1].

Нулем по сложению в этом пространстве является *нулевая матрица*

$$\mathbf{O} = \begin{pmatrix} 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \ldots & & & \\ 0 & 0 & \ldots & 0 \end{pmatrix}$$

− матрица, все элементы которой равны нулю.

Матрицей, противоположной матрице $\mathbf{A}$ по сложению, является матрица

$$-\mathbf{A} = -1 \cdot \mathbf{A} = \left(-a_{ij}\right);$$

разностью $\mathbf{A} - \mathbf{B}$ матриц $\mathbf{A}$ и $\mathbf{B}$ называется сумма $\mathbf{A} + \left(-\mathbf{B}\right)$.

Используя определения сложения матриц и умножения матрицы на число, можно иначе сформулировать понятие кососимметрической матрицы. А именно, *квадратная матрица называется кососимметрической, если транспонирование переводит ее в противоположную*: $\mathbf{A}^T = -\mathbf{A}$.

Контрольные вопросы и упражнения

1. Что называется суммой матриц? Что называется произведением матрицы на число?

2. Даны матрицы

[1] Это пространство линейно изоморфно $m{\cdot}n$ - мерному координатному пространству.

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

Вычислите:

$$2\mathbf{A} + \mathbf{B}^T; \ \mathbf{A}^T - \mathbf{B}; \ \left(-3\mathbf{A} - 2\mathbf{B}^T\right)^T.$$

3. Определена ли для матриц предыдущего примера сумма $\mathbf{A} + \mathbf{B}$? Сумма $\mathbf{A}^T + \mathbf{B}^T$?

2.3. Умножение матриц

Пусть m чисел $z_1, z_2, \ldots, z_m$ *линейно выражаются* через p чисел $y_1, y_2, \ldots, y_p$:

$$z_i = \sum_{k=1}^{p} a_{ik} y_k, \ i = \overline{1, m}, \tag{a}$$

а числа $y_1, y_2, \ldots, y_p$ линейно выражаются через числа $x_1, x_2, \ldots, x_n$:

$$y_k = \sum_{j=1}^{n} b_{kj} x_j, \ k = \overline{1, p}.$$

Тогда числа z_i можно линейно выразить через x_j:

$$z_i = \sum_{k=1}^{p} a_{ik} \sum_{j=1}^{n} b_{kj} x_j = \sum_{j=1}^{n} \left(\sum_{k=1}^{p} a_{ik} b_{kj} \right) x_j, \ i = \overline{1, m}. \tag{b}$$

Коэффициенты

$$\sum_{k=1}^{p} a_{ik} b_{kj}$$

в этих m выражениях устанавливают правило *умножения матриц*.

Именно, пусть даны матрица $\mathbf{A} = (a_{ik})$ размера $m \times p$ и матрица $\mathbf{B} = (b_{kj})$ размера $p \times n$. *Произведением матриц* $\mathbf{A}$ и $\mathbf{B}$ называется матрица $\mathbf{C} = (c_{ij})$ размера $m \times n$, элементы c_{ij} которой равны суммам произведений соответствующих элементов i-й строки матрицы $\mathbf{A}$ и j-го столбца матрицы $\mathbf{B}$:

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}, \ i = \overline{1, m}, \ j = \overline{1, n}.$$

Например, пусть требуется найти произведение матриц

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

В первом сомножителе две строки, поэтому произведение также имеет две строки. Во втором сомножителе два столбца, поэтому произведение также имеет два столбца. Каждый элемент произведения есть сумма из трех слагаемых (число столбцов в первом сомножителе, равное числу строк во втором):

$$\begin{pmatrix} 1 & 3 & 6 \\ 2 & 4 & 7 \end{pmatrix}\begin{pmatrix} -2 & -1 \\ 5 & -3 \\ -4 & -5 \end{pmatrix} = \begin{pmatrix} 1\cdot(-2)+3\cdot 5+6\cdot(-4) & 1\cdot(-1)+3\cdot(-3)+6\cdot(-5) \\ 2\cdot(-2)+4\cdot 5+7\cdot(-4) & 2\cdot(-1)+4\cdot(-3)+7\cdot(-5) \end{pmatrix} =$$

$$= \begin{pmatrix} -2+15-24 & -1-9-30 \\ -4+20-28 & -2-12-35 \end{pmatrix} = \begin{pmatrix} -11 & -40 \\ -12 & -49 \end{pmatrix}.$$

Подобное определение операции умножения матриц позволяет записать соотношение (a) в виде

$$\mathbf{z} = \mathbf{A}\mathbf{y},$$

а соотношение (b) – в виде

$$\mathbf{z} = (\mathbf{A}\mathbf{B})\mathbf{x},$$

где $\mathbf{z}$, $\mathbf{y}$ и $\mathbf{x}$ – вектор-столбцы высоты m, p и n, соответственно.

Из определения следует, что умножать можно не любые матрицы. Умножение возможно только в том случае, если число столбцов в первом множителе совпадает с числом строк во втором[1]. В частности, можно умножать две квадратные матрицы одного порядка; при этом результат является матрицей того же порядка.

[1] Такие матрицы иногда называют *согласованными*.

Произведение двух квадратных диагональных матриц $\mathbf{A} = (a_{ij})$ и $\mathbf{B} = (b_{ij})$ равных порядков является диагональной матрицей того же порядка. Действительно, если при $i \neq j$ выполнено $a_{ij} = 0$ и $b_{ij} = 0$, то элементы произведения $\mathbf{C} = \mathbf{AB} = (c_{ij})$ оказываются равными

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} = a_{ii} b_{ij} = \begin{cases} a_{ii} b_{ij}, & i = j \\ 0, & i \neq j \end{cases}.$$

Пусть квадратные матрицы $\mathbf{A} = (a_{ij})$ и $\mathbf{B} = (b_{ij})$ равных порядков являются верхними треугольными: $a_{ij} = b_{ij} = 0$ при $i > j$. Тогда элементы их произведения $\mathbf{C} = \mathbf{AB}$ равны

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

Следовательно, $c_{ij} = 0$ при $i > j$ и произведение $\mathbf{C} = \mathbf{AB}$ также является верхней треугольной матрицей.

Аналогично показывается, что произведение двух нижних треугольных матриц есть нижняя треугольная матрица.

Произведение матрицы **A** размера $m \times n$ и столбца **B** высоты n есть столбец высоты m:

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & & & \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \ldots \\ b_n \end{pmatrix} = \begin{pmatrix} a_{11} b_1 + a_{12} b_2 + \ldots + a_{1n} b_n \\ a_{21} b_1 + a_{22} b_2 + \ldots + a_{2n} b_n \\ \ldots \\ a_{m1} b_1 + a_{m2} b_2 + \ldots + a_{mn} b_n \end{pmatrix}.$$

Поэтому можно говорить, что j-й столбец произведения двух (прямоугольных) матриц есть линейная комбинация столбцов первого сомножителя с коэффициентами, равными элементам j-го столбца второго сомножителя.

Произведение строки **B** длины n на матрицу **A** размера $m \times n$ есть строка длины n:

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

поэтому i-я строка произведения есть линейная комбинация строк второго сомножителя с коэффициентами, равными элементам i-й строки первого сомножителя.

Для операции умножения матриц выполняются свойства:

1. Умножение матриц ассоциативно:

$$\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}.$$

Действительно, пусть $\mathbf{A} = (a_{ik})$, $\mathbf{B} = (b_{ks})$, $\mathbf{C} = (c_{sj})$, $k = \overline{1,p}$, $s = \overline{1,l}$. Тогда

$$\mathbf{A}(\mathbf{BC}) = \left(\sum_{k=1}^{p} a_{ik} \left(\sum_{s=1}^{l} b_{ks} c_{sj} \right) \right) = \left(\sum_{k=1}^{p} \sum_{s=1}^{l} a_{ik} b_{ks} c_{sj} \right) = \left(\sum_{s=1}^{l} \left(\sum_{k=1}^{p} a_{ik} b_{ks} \right) c_{sj} \right) = (\mathbf{AB})\mathbf{C}.$$

2. Умножение дистрибутивно относительно сложения:

$$(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}; \ \ \mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}.$$

Действительно, если $\mathbf{A} = (a_{ik})$, $\mathbf{B} = (b_{ik})$, $\mathbf{C} = (c_{kj})$, $k = \overline{1,p}$, то

$$(\mathbf{A} + \mathbf{B})\mathbf{C} = \left(\sum_{k=1}^{p} (a_{ik} + b_{ik}) c_{kj} \right) = \left(\sum_{k=1}^{p} a_{ik} c_{kj} + \sum_{k=1}^{p} b_{ik} c_{kj} \right) =$$

$$= \left(\sum_{k=1}^{p} a_{ik} c_{kj} \right) + \left(\sum_{k=1}^{p} b_{ik} c_{kj} \right) = \mathbf{AC} + \mathbf{AB}.$$

3. $\alpha(\mathbf{AB}) = (\alpha\mathbf{A})\mathbf{B} = \mathbf{A}(\alpha\mathbf{B})$.

Доказательство. Пусть $\mathbf{A} = (a_{ik})$, $\mathbf{B} = (b_{kj})$, $k = \overline{1,p}$. Тогда

$$\alpha(\mathbf{AB}) = \alpha \left(\sum_{k=1}^{p} a_{ik} b_{kj} \right) = \left(\sum_{k=1}^{p} (\alpha a_{ik}) b_{kj} \right) = (\alpha\mathbf{A})\mathbf{B};$$

$$\alpha(\mathbf{AB}) = \alpha \left(\sum_{k=1}^{p} a_{ik} b_{kj} \right) = \left(\sum_{k=1}^{p} a_{ik} (\alpha b_{kj}) \right) = \mathbf{A}(\alpha\mathbf{B}).$$

4. $\left(\mathbf{AB}\right)^T = \mathbf{B}^T\mathbf{A}^T$.

Действительно, пусть $\mathbf{A} = \left(a_{ik}\right)$, $\mathbf{B} = \left(b_{kj}\right)$, $\mathbf{A}^T = \left(a_{ki}\right)$, $\mathbf{B}^T = \left(b_{jk}\right)$, $k = \overline{1, p}$. Элемент, находящийся на пересечении i-й строки и j-го столбца произведения $\mathbf{C} = \mathbf{AB}$ равен

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj} .$$

В то же время, элемент, находящийся на пересечении j-й строки и i-го столбца произведения $\mathbf{D} = \mathbf{B}^T\mathbf{A}^T$ равен

$$d_{ji} = \sum_{k=1}^{p} b_{jk} a_{ki} ,$$

поэтому при любых i и j выполнено условие $c_{ij} = d_{ji}$.

Данное свойство легко распространяется на произведение, содержащее произвольное число сомножителей:

$$\left(\mathbf{AB}\ldots\mathbf{CD}\right)^T = \mathbf{D}^T\mathbf{C}^T\ldots\mathbf{B}^T\mathbf{A}^T .$$

Из свойства 4 следует, что если $\mathbf{A}$ – произвольная прямоугольная матрица, то матрицы $\mathbf{A}^T\mathbf{A}$ и $\mathbf{AA}^T$ являются симметрическими[1]. Действительно,

$$\left(\mathbf{A}^T\mathbf{A}\right)^T = \mathbf{A}^T\left(\mathbf{A}^T\right)^T = \mathbf{A}^T\mathbf{A} ;$$

$$\left(\mathbf{AA}^T\right)^T = \left(\mathbf{A}^T\right)^T\mathbf{A}^T = \mathbf{AA}^T .$$

5. Произведение любой матрицы (квадратной или прямоугольной) на нулевую есть нулевая матрица (возможно, другого размера).

6. Если $\mathbf{A}$ – квадратная матрица порядка n, и $\mathbf{E}$ – единичная матрица того же порядка, то

$$\mathbf{AE} = \mathbf{EA} = \mathbf{A} .$$

Нетрудно убедиться, что умножение матриц не коммутативно: перестановка сомножителей в общем случае меняет произведение.

---

[1] Подобные произведения часто встречаются в анализе при использовании т.н. *метода наименьших квадратов*.

Например:

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

в то же время

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

Если для квадратных матриц $\mathbf{A}$ и $\mathbf{B}$ выполнено условие $\mathbf{AB}=\mathbf{BA}$, то матрицы называют *коммутирующими*. В частности, свойство 6 утверждает что любая квадратная матрица порядка n коммутирует с единичной матрицей. Из этого свойства следует также, что если для произвольной квадратной матрицы $\mathbf{A}$ и матрицы $\mathbf{E}'$ выполнено условие

$$\mathbf{E}'\mathbf{A}=\mathbf{AE}'=\mathbf{A},$$

то матрица $\mathbf{E}'$ является единичной. Действительно, полагая $\mathbf{A}=\mathbf{E}$ получим:

$$\mathbf{E}'\mathbf{E}=\mathbf{EE}'=\mathbf{E}.$$

Но, с другой стороны,

$$\mathbf{E}'\mathbf{E}=\mathbf{EE}'=\mathbf{E}',$$

откуда

$$\mathbf{E}'=\mathbf{E}.$$

Контрольные вопросы и упражнения

1. Запишите формулу, выражающую элементы произведения матриц через элементы сомножителей.
2. Выполните умножение

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

3. Выполнено ли для матрицы

$$\mathbf{C}=\begin{pmatrix}1 & a & c\\ 1 & b & b\\ 1 & c & a\end{pmatrix}\begin{pmatrix}1 & a & c\\ 1 & b & b\\ 1 & c & a\end{pmatrix}^T$$

свойство $a_{ij}=a_{4-j,4-i}$?

4. Пусть матрица $\mathbf{A}$ – кососимметрическая. Будет ли произведение $\mathbf{A}(-\mathbf{A})$ также являться кососимметрической матрицей? Симметрической матрицей?

5. Определено ли для матриц

$$\mathbf{A}=\begin{pmatrix}5 & -4\\ -3 & 2\\ 4 & 8\end{pmatrix} \text{ и } \mathbf{B}=\mathbf{A}^T$$

произведение $\mathbf{AB}$? Определено ли произведение $\mathbf{A}^T\mathbf{B}$? Произведение $\mathbf{A}^T\mathbf{B}^T$?

6. Вычислите

$$\begin{pmatrix}15 & 31 & 23\\ 17 & -11 & 8\\ 26 & -18 & 5\end{pmatrix}\begin{pmatrix}-2 & 1\\ 3 & -5\\ -7 & 9\end{pmatrix}+\begin{pmatrix}15 & 31 & 23\\ 17 & -11 & 8\\ 26 & -18 & 5\end{pmatrix}\begin{pmatrix}2 & -1\\ -3 & 5\\ 7 & -9\end{pmatrix}.$$

7. Говорят, что алгебраическая операция *антикоммутативна*, если смена порядка операндов переводит результат операции в противоположный ему по сложению: $ab=-ba$. Является умножение матриц (в общем случае) антикоммутативным? Является ли оно антикоммутативным для матриц

$$\begin{pmatrix}0 & -a\\ a & 0\end{pmatrix} \text{ и } \begin{pmatrix}0 & b\\ b & 0\end{pmatrix}?$$

2.4. Элементарные преобразования

Элементарными преобразованиями строк матрицы называют:

1. Умножение любой строки на отличное от нуля число.
2. Замену любой строки на сумму ее с другой строкой.

Аналогично определяются элементарные преобразования столбцов.

В результате последовательного применения нескольких элементарных преобразований можно выполнить:

– замену строки разностью ее с другой строкой;

– перестановку строк.

Замена строки $\mathbf{a}^k$ разностью $\mathbf{a}^k - \mathbf{a}^s$ осуществляется как последовательность трех преобразований:

– умножение s-й строки на -1;

– замена k-й строки суммой ее с s-й строкой;

– умножение s-й строки на -1.

Перестановка строк $\mathbf{a}^k$ и $\mathbf{a}^s$ может быть осуществлена как последовательность четырех преобразований:

– замена k-й строки разностью ее с s-й;

– замена s-й строки суммой с k-й;

– замена k-й строки разностью с s-й;

– умножение k-й строки на -1.

Если матрица $\mathbf{B}$ получена из матрицы $\mathbf{A}$ при помощи элементарных преобразований, то говорят, что $\mathbf{A}$ *редуцируема* (*сводима*) к матрице $\mathbf{B}$. Так как любая матрица редуцируема сама к себе (соответствующая последовательность не содержит ни одного элементарного преобразования), то отношение редуцируемости рефлексивно.

Элементарные преобразования *обратимы*: если матрица $\mathbf{B}$ может быть получена из матрицы $\mathbf{A}$ при помощи элементарных преобразований, то существует последовательность преобразований, переводящая $\mathbf{B}$ в $\mathbf{A}$. Действительно, умножение строки на число $\alpha \neq 0$ может быть обращено умножением на число $1/\alpha$; замена строки $\mathbf{a}^k$ суммой $\mathbf{a}^k + \mathbf{a}^s$ может быть обращена заменой разностью $\mathbf{a}^k - \mathbf{a}^s$. Поэтому отношение редуцируемости симметрично.

Наконец, если матрица $\mathbf{B}$ может быть получена из матрицы $\mathbf{A}$ и матрица $\mathbf{C}$ может быть получена из матрицы $\mathbf{B}$ при помощи элементарных преобразований, то существует последовательность элементарных преобразований, переводящая матрицу $\mathbf{A}$ в матрицу $\mathbf{C}$.

Таким образом, для отношения редуцируемости выполнены все свойства отношения эквивалентности. Поэтому данное отношение разбивает все множество матриц размеров $m \times n$ на классы, представители которых редуцируемы друг к другу.

Например, любая матрица второго порядка элементарными преобразованиями строк редуцируется к одной из следующих матриц:

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

где a – произвольное число (возможно, равное нулю).

Отношение редуцируемости часто обозначают знаком эквивалентности:

$$\mathbf{A} \sim \mathbf{B}$$

(читается «**A** эквивалентна **B**»), однако следует помнить, что *из редуцируемости матриц друг к другу не следует их равенство* (хотя обратное, очевидно, выполнено).

Элементарные преобразования строк матрицы **A** можно представить как умножение матрицы справа на некоторую квадратную матрицу, вид которой полностью определяется преобразованием и не зависит от **A**. *Тождественному преобразованию* (преобразованию, которое не меняет матрицу) соответствует умножение на единичную матрицу.

Умножение k-й строки матрицы **A** (размера $m \times n$) на число λ можно представить как умножение матрицы **A** на матрицу

$$\mathbf{S}_{k,\lambda} = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & \dots & 0 \\ \dots & & & & & \\ 0 & 0 & \dots & \lambda & \dots & 0 \\ \dots & & & & & \\ 0 & 0 & 0 & 0 & \dots & 1 \end{pmatrix},$$

полученную из единичной $\mathbf{E}_n$ заменой k-й единицы на диагонали числом λ. Действительно, элементы произведения $\mathbf{S}_{k,\lambda}\mathbf{A} = \mathbf{A}' = \left(a'_{ij}\right)$ равны

$$a'_{ij} = \sum_{l=1}^{m} s_{il} a_{lj} = s_{ii} a_{ij} = \begin{cases} \lambda a_{ij}, & i = k \\ a_{ij}, & i \neq k \end{cases}.$$

Пусть $\mathbf{C}_{kl}$ – матрица, полученная из $\mathbf{E}_n$ заменой на единицу нулевого элемента на пересечении k-й строки и l-го столбца. Тогда умножение на $\mathbf{C}_{kl}$ соответствует замене k-й строки на сумму ее l-й строкой.

$$\mathbf{C}_{kl}\mathbf{A} = \mathbf{A}'' = \left(a''_{ij}\right) = \sum_{s=1}^{m} c_{is}a_{sj} = \begin{cases} a_{lj} + a_{ij}, & i = k \\ a_{ij}, & i \neq k \end{cases}.$$

Матрицы $\mathbf{S}_{k,\lambda}$ и $\mathbf{C}_{kl}$, умножение на которые соответствует элементарным преобразованиям, называют *элементарными матрицами*.

Последовательное выполнение элементарных преобразований строк матрицы **A** сводится к последовательному умножению на элементарные матрицы. При этом матрица, соответствующая первому преобразованию, является в произведении ближайшей к матрице **A**.

Например, замена второй строки матрицы **A** разностью ее с первой строкой осуществляется последовательным умножением матрицы **A** справа на матрицы $\mathbf{S}_{1,-1}$, $\mathbf{C}_{21}$ и $\mathbf{S}_{1,-1}$. В частности, для матрицы размера $2 \times n$:

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

Указанное преобразование соответствует умножению на матрицу, полученную из единичной заменой на -1 нулевого элемента на пересечении второй строки и первого столбца.

Матрицу преобразования, состоящего в замене k-й строки разностью ее с l-й строкой, можно обозначить $\mathbf{C}_{kl}^{-1}$. Целесообразность подобного обозначения связана с обратимостью элементарных преобразований: сложение k-й строки с l-й строкой и последующее вычитание l-й строки из k-й строки (выполненные в произвольном порядке) не меняют k-ю строку.

Иначе говоря, при любых k и l выполнено условие

$$\mathbf{C}_{kl}^{-1}\mathbf{C}_{kl}\mathbf{A} = \mathbf{C}_{kl}\mathbf{C}_{kl}^{-1}\mathbf{A} = \mathbf{A},$$

следовательно,

$$\mathbf{C}_{kl}^{-1}\mathbf{C}_{kl} = \mathbf{C}_{kl}\mathbf{C}_{kl}^{-1} = \mathbf{E},$$

где **E** – единичная матрица того же порядка, что и $\mathbf{C}_{kl}$.

Так, для приведенного выше примера:

$$\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1\cdot1+0\cdot1 & 1\cdot0+0\cdot1 \\ -1\cdot1+1\cdot1 & -1\cdot0+1\cdot1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix};$$

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1\cdot1+0\cdot(-1) & 1\cdot0+0\cdot1 \\ 1\cdot1+1\cdot(-1) & 1\cdot0+1\cdot1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Для матрицы **A** размера $2{\times}n$ перестановка первой и второй строк может быть представлена умножением

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

В общем случае перестановка строк соответствует умножению на матрицу, полученную из единичной перестановкой соответствующих строк. Обращение указанного преобразования можно представить как повторное умножение на ту же самую матрицу (перестановка строк с номерами k и l, выполненная дважды, не меняет матрицу). В частности, для приведенного примера:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0\cdot0+1\cdot1 & 0\cdot1+1\cdot0 \\ 1\cdot0+0\cdot1 & 1\cdot1+0\cdot0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Подводя итог, можно отметить, что *любое элементарное преобразование осуществляется умножением на матрицу, полученную аналогичным преобразованием из единичной матрицы.*

Контрольные вопросы и упражнения

1. Какие действия называют элементарными преобразованиями строк? Элементарными преобразованиями столбцов?

2. Укажите последовательность элементарных преобразований строк, переводящую матрицу $\begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}$ в матрицу $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

3. Редуцируемы ли матрицы предыдущего упражнения к единичной матрице?

4. Запишите матрицу, представляющую редукцию (последовательность элементарных преобразований строк) матриц упражнения 2.

Рассматривая матрицу $\mathbf{A} = \left(a_j^i\right)$ размера $m \times n$ как строку из столбцов или столбец из строк, ее ряды – столбцы $\mathbf{a}_j$ или строки $\mathbf{a}^i$ – можно считать элементами координатных линейных пространств (размерностей m или n, соответственно). Отношение линейной зависимости рядов матрицы определяется совершенно аналогично отношению линейной зависимости элементов произвольного линейного пространства: ряды матрицы называются линейно зависимыми, если существует нетривиальная и равная нулю (точнее, нулевому вектору соответствующего координатного пространства) их линейная комбинация.

Пусть матрица $\mathbf{A}$ имеет порядок n. Выберем в координатном пространстве K_n естественный базис (см. п. 1.7)

$$\mathbf{e}_1 = (1, 0, \ldots, 0),\ \mathbf{e}_2 = (0, 1, \ldots, 0),\ \ldots,\ \mathbf{e}_n = (0, 0, \ldots, 1).$$

Матрица, рядами которой являются векторы естественного базиса, является единичной матрицей. Из доказанного выше следует, что ряды единичной матрицы линейно независимы и любой ряд (вектор-строка или вектор-столбец) с тем же числом элементов может быть по ним разложен.

Если строки матрицы линейно зависимы, то матрицу мы будем называть *вырожденной* (иначе – *сингулярной*, или *особенной*). Из изложенного в п. 1.6 следует, что матрица вырождена тогда и только тогда, когда одна из ее строк является линейной комбинацией остальных. В частности, вырожденной будет матрица, имеющая хотя бы одну нулевую строку. Примером невырожденной матрицы является единичная.

Теорема 2.5.1. *Элементарные преобразования строк сохраняют линейную зависимость между ними.*

Доказательство. Пусть строки матрицы $\mathbf{A}$ линейно независимы. Тогда в линейной комбинации

$$\sum_{i=1}^{m} \alpha_i \mathbf{a}^i = \mathbf{0}$$

все коэффициенты α_i нулевые.

Пусть матрица **B** получена из матрицы **A** умножением s-й строки на отличное от нуля число λ. Рассмотрим линейную комбинацию строк $\mathbf{b}^i$ полученной матрицы

$$\sum_{i=1}^{m} \beta_i \mathbf{b}^i = \beta_s \lambda \mathbf{a}^s + \sum_{\substack{i=1 \\ i \neq s}}^{m} \beta_i \mathbf{a}^i .$$

Так как строки $\mathbf{a}^i$ исходной матрицы линейно независимы, то их линейная комбинация

$$\beta_s \lambda \mathbf{a}^s + \sum_{\substack{i=1 \\ i \neq s}}^{m} \beta_i \mathbf{a}^i = \mathbf{0}$$

должна быть тривиальной (равенство возможно только с нулевыми коэффициентами); следовательно, $\beta_s \lambda = 0$. Отсюда в силу $\lambda \neq 0$ получим $\beta_s = 0$; линейная комбинация

$$\sum_{i=1}^{m} \beta_i \mathbf{b}^i = \mathbf{0}$$

также тривиальна.

Пусть матрица **B** получена из матрицы **A** заменой s-й строки на сумму ее с k-й строкой. Рассмотрим линейную комбинацию

$$\sum_{i=1}^{m} \beta_i \mathbf{b}^i = \beta_s \left(\mathbf{a}^s + \mathbf{a}^k \right) + \sum_{\substack{i=1 \\ i \neq s}}^{m} \beta_i \mathbf{a}^i = \beta_s \mathbf{a}^s + (\beta_k + \beta_s) \mathbf{a}^k + \sum_{\substack{i=1 \\ i \notin \{k,s\}}}^{m} \beta_i \mathbf{a}^i .$$

Так как строки $\mathbf{a}^i$ исходной матрицы линейно независимы, то равенство

$$\beta_s \mathbf{a}^s + (\beta_k + \beta_s) \mathbf{a}^k + \sum_{\substack{i=1 \\ i \notin \{k,s\}}}^{m} \beta_i \mathbf{a}^i = \mathbf{0}$$

возможно только с нулевыми коэффициентами. Следовательно, $\beta_k + \beta_s = 0$. Поэтому (в силу $\beta_s = 0$) коэффициент β_k также равен нулю. Линейная комбинация

$$\sum_{i=1}^{m} \beta_i \mathbf{b}^i = \mathbf{0}$$

тривиальна.

Пусть теперь строки матрицы **A** линейно зависимы. Предположим, что линейно независимыми будут строки матрицы $\mathbf{B} = \mathbf{TA}$, полученной в результате элементарного преобразования матрицы **A**. Так как элементарные преобразования обратимы, то существует преобразование

$$\mathbf{T}_1\mathbf{B} = \mathbf{A},$$

переводящее матрицу **B** в матрицу **A**. Но тогда из предыдущего следует, что строки матрицы **A** также линейно независимы. Полученное противоречие завершает доказательство теоремы.

Как следствие, *элементарные преобразования строк переводят невырожденную матрицу в невырожденную, а вырожденную матрицу – в вырожденную*.

Теорема 2.5.2. *Элементарные преобразования столбцов сохраняют линейную зависимость между ними*.

Доказательство выполняется аналогично.

Теорема 2.5.3. *Элементарные преобразования строк сохраняют линейную зависимость между столбцами*.

Доказательство. Предположим, что столбцы матрицы **A** линейно зависимы.

Тогда некоторый k-й столбец линейно выражается через остальные:

$$\mathbf{a}_k = \sum_{\substack{j=1 \\ j \neq k}}^{n} \alpha_j \mathbf{a}_j,$$

или, в развернутой форме

$$a_{ik} = \sum_{\substack{j=1 \\ j \neq k}}^{n} \alpha_j a_{ij}, \; i = \overline{1, m}, \tag{a}$$

где m и n − число строк и столбцов матрицы **A**.

Пусть матрица **B** получена из матрицы **A** умножением s-й строки на число λ:

$$b_{ij} = \begin{cases} a_{ij}, & i \neq s \\ \lambda a_{ij}, & i = s \end{cases}.$$

Тогда для любого $i \neq s$ зависимость (а) выполнена. При $i = s$ имеем

$$b_{sk} = \lambda \sum_{\substack{j=1 \\ j \neq k}}^{n} \alpha_j a_{sj} = \sum_{\substack{j=1 \\ j \neq k}}^{n} \alpha_j \left(\lambda a_{sj} \right) = \sum_{\substack{j=1 \\ j \neq k}}^{n} \alpha_j b_{sj} ;$$

следовательно, k-й столбец матрицы **B** линейно выражается через остальные.

Пусть матрица **B** получена из матрицы **A** заменой s-й строки на сумму ее с l-й строкой:

$$b_{ij} = \begin{cases} a_{ij}, i \neq s \\ a_{ij} + a_{lj}, i = s \end{cases}.$$

Для любого $i \neq s$ зависимость (а) выполнена. При $i = s$ получим

$$b_{sk} = \sum_{\substack{j=1 \\ j \neq k}}^{n} \alpha_j a_{sj} + \sum_{\substack{j=1 \\ j \neq k}}^{n} \alpha_j a_{lj} = \sum_{\substack{j=1 \\ j \neq k}}^{n} \alpha_j \left(a_{sj} + a_{lj} \right) = \sum_{\substack{j=1 \\ j \neq k}}^{n} \alpha_j b_{sj} ;$$

следовательно, k-й столбец линейно выражается через остальные.

Пусть теперь столбцы матрицы **A** линейно независимы. Предположим, что линейно зависимыми будут столбцы матрицы $\mathbf{B} = \mathbf{TA}$, полученной в результате элементарного преобразования матрицы **A**. Так как элементарные преобразования обратимы, то существует преобразование

$$\mathbf{T}_1 \mathbf{B} = \mathbf{A} ,$$

переводящее матрицу **B** в матрицу **A**. Но тогда из предыдущего следует, что столбцы матрицы **A** также линейно зависимы. Полученное противоречие доказывает теорему.

Теорема 2.5.4. *Элементарные преобразования столбцов сохраняют линейную зависимость между строками.*

Доказательство выполняется аналогично.

Как следствие, *элементарные преобразования* (*строк или столбцов*) *переводят невырожденную матрицу в невырожденную, а вырожденную матрицу – в вырожденную.*

Теорема 2.5.5. *Любая невырожденная квадратная матрица элементарными преобразованиями строк может быть редуцирована к единичной матрице того же порядка.*

Доказательство этой теоремы можно выполнить *конструктивно*, фактически установив способ редукции. Рассмотрим невырожденную матрицу

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

В первой строке матрицы **A** обязательно найдется элемент, отличный от нуля. Пусть этот элемент a_{1s_1} расположен в столбце с номером s_1. Разделим первую строку на a_{1s_1} и затем для всех $i = \overline{2,n}$ вычтем из i-й строки первую строку, умноженную на a_{is_1}.

Получим матрицу

$$\mathbf{B} = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1s_1-1} & 1 & b_{1s_1+1} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2s_1-1} & 0 & b_{2s_1+1} & \dots & b_{2n} \\ \dots & & & & & & & \\ b_{n1} & b_{n2} & \dots & b_{ns_1-1} & 0 & b_{ns_1+1} & \dots & b_{nn} \end{pmatrix},$$

в которой s_1-й столбец является первым столбцом единичной матрицы.

Матрица **B** невырождена, так как получена из матрицы **A** элементарными преобразованиями. Следовательно, в ее второй строке есть ненулевой элемент b_{2s_2}, причем $s_2 \neq s_1$ (так как в столбце s_1 все элементы кроме b_{1s_1} уже обращены в ноль). Разделим вторую строку на элемент b_{2s_2} и заменим первую строку и строки с номерами 3, 4, …, n на разности со второй строкой, умноженной на b_{is_2}. Получим матрицу, у которой столбец с номером s_2 является вторым столбцом единичной матрицы.

Выполняя преобразования далее, на n-м шаге получим матрицу, все столбцы которой совпадают со столбцами единичной матрицы (и находятся в позициях s_1, s_2, …, s_n). Но тогда строки этой матрицы также совпадают со строками единичной матрицы. Выполняя перестановку строк, можно расположить их в естественном порядке. Теорема доказана.

Метод преобразования, использованный при доказательстве теоремы, называется *методом Гаусса* (точнее – *методом Жордана–Гаусса с выбором ведущего элемента по строке*). Различные варианты этого метода широко используются в вычислительной практике.

Теорема 2.5.6. *Квадратная матрица невырождена тогда и только тогда, когда она может быть представлена в виде произведения элементарных матриц*.

Докажем необходимость. В силу предыдущей теоремы для любой невырожденной матрицы **A** найдется последовательность элементарных преобразований, с помощью которых эта матрица может быть превращена в единичную. Так как элементарные преобразования обратимы, то существуют элементарные матрицы $\mathbf{T}_1$, $\mathbf{T}_2$, …, $\mathbf{T}_M$, последовательное умножение на которые превращает единичную матрицу в **A**:

$$\mathbf{T}_1\mathbf{T}_2\ldots\mathbf{T}_M\mathbf{E} = \mathbf{A},$$

откуда $\mathbf{A} = \mathbf{T}_1\mathbf{T}_2\ldots\mathbf{T}_M$.

Докажем достаточность. Пусть матрица **A** представлена в виде произведения

$$\mathbf{A} = \mathbf{T}_1\mathbf{T}_2\ldots\mathbf{T}_M\mathbf{E}$$

элементарных матриц. Тогда произведение $\mathbf{T}_1\mathbf{T}_2\ldots\mathbf{T}_M$ является результатом элементарных преобразований строк единичной матрицы, которая невырождена. Поэтому произведение $\mathbf{T}_1\mathbf{T}_2\ldots\mathbf{T}_M$ также невырождено. Теорема доказана.

Теорема 2.5.7. *Квадратная матрица невырождена тогда и только тогда, когда ее столбцы линейно независимы*.

Докажем необходимость. Если матрица невырождена, то элементарными преобразованиями ее можно превратить в единичную, столбцы которой линейно независимы. Следовательно, столбцы исходной матрицы также линейно независимы.

Докажем достаточность. Пусть столбцы матрицы **A** линейно независимы. Тогда линейно независимыми будут строки транспонированной матрицы $\mathbf{A}^T$, и матрица $\mathbf{A}^T$ оказывается невырожденной. Поэтому ее столбцы линейно независимы. Но тогда линейно независимыми будут строки матрицы $\left(\mathbf{A}^T\right)^T = \mathbf{A}$, следовательно, матрица $\mathbf{A}$ невырождена. Теорема доказана.

Утверждение последней теоремы можно сформулировать иначе – *матрицы* $\mathbf{A}$ *и* $\mathbf{A}^T$ *или обе вырождены, или обе невырождены*.

Контрольные вопросы и упражнения

1. Какую матрицу называют вырожденной?

2. Можно ли элементарными преобразованиями строк и столбцов превратить сингулярную матрицу в несингулярную? Превратить несингулярную матрицу в сингулярную?

3. Выясните, являются ли вырожденными матрицы:

$$\begin{pmatrix} 1 & 2 \\ 3 & 3 \end{pmatrix};\ \begin{pmatrix} 1 & 2 & -2 \\ 3 & -5 & 1 \\ -1 & -3 & -1 \end{pmatrix};\ \begin{pmatrix} 1 & 2 & -2 \\ 3 & -5 & 1 \\ -1 & 9 & -5 \end{pmatrix};\ \begin{pmatrix} 1 & 3 & -1 \\ 2 & -5 & 9 \\ -2 & 1 & -5 \end{pmatrix}.$$

2.6. Ранг матрицы

Пусть в матрице размеров $m{\times}n$ имеется невырожденная квадратная подматрица порядка r, но все подматрицы большего порядка вырождены. Тогда указанная подматрица называется *базисной*, а ее порядок r называется *рангом матрицы* и обозначается $\operatorname{rang} A$. Строки и столбцы, входящие в базисную подматрицу, называются *базисными* строками и столбцами.

Нулевая матрица не содержит никакой невырожденной подматрицы, поэтому ее ранг равен нулю. Ранг единичной матрицы совпадает с ее порядком.

Предложение 2.6.1. *Ранг матрицы не меняется при транспонировании*. Действительно, при транспонировании матрицы все ее подматрицы также транспонируются. При этом (см. п. 2.5) невырожденные подматрицы остаются невырожденными, а вырожденные – вырожденными.

Предложение 2.6.2. *Ранг любой подматрицы* $\mathbf{A}'$ *матрицы* $\mathbf{A}$ *не превосходит ранга матрицы* $\mathbf{A}$. Действительно, любая невырожденная подматрица, входящая в $\mathbf{A}'$, входит и в матрицу $\mathbf{A}$.

Предложение 2.6.3. *Ранг матрицы не меняется при элементарных преобразованиях*. Действительно, элементарные преобразования (строк или столбцов) переводят любую невырожденную подматрицу в невырожденную, а любую вырожденную – в вырожденную.

Теорема 2.6.1. *Система из r строк линейно независима тогда и только тогда, когда в этих строках найдется невырожденная квадратная подматрица порядка r.*

Ограничимся доказательством достаточности. Предположим, что система из r строк линейно зависима. Рассмотрим произвольную квадратную подматрицу порядка r, расположенную в этих строках. Если строки линейно зависимы, то линейно зависимыми (с теми же коэффициентами) будут и отрезки этих строк, составляющие подматрицу. Поэтому подматрица вырождена. Следовательно (в силу произвольности подматрицы), если в r строках найдется невырожденная подматрица порядка r, то строки линейно независимы.

Теорема 2.6.2 (*теорема о ранге*). *Максимальное число линейно независимых строк совпадает с максимальным числом линейно независимых столбцов и равно рангу матрицы.*

Действительно, пусть число линейно независимых строк матрицы $\mathbf{A}$ равно r, но любые $p > r$ строк уже линейно зависимы. Тогда в r линейно независимых строках найдется невырожденная подматрица порядка r, но любая подматрица, расположенная в $p > r$ строках, будет вырожденной. Следовательно, ранг матрицы равен r. Далее, максимальное число линейно независимых столбцов матрицы $\mathbf{A}$ совпадает с максимальным числом линейно независимых строк матрицы $\mathbf{A}^T$, а значит – совпадает с рангом матрицы $\mathbf{A}^T$. Так как ранг не меняется при транспонировании, то максимальное число линейно независимых столбцов матрицы $\mathbf{A}$ равно рангу $\mathbf{A}$. Теорема доказана.

Предложение 2.6.4. *Ранг матрицы не меняется при умножении ее на невырожденную матрицу.*

Действительно, если матрица $\mathbf{A}$ невырождена, то она может быть представлена в виде произведения

$$\mathbf{A} = \mathbf{T}_1\mathbf{T}_2\ldots\mathbf{T}_M$$

элементарных матриц. Следовательно, если для матрицы $\mathbf{B}$ определено произведение $\mathbf{AB}$, то

$$\mathbf{AB} = \mathbf{T}_1\mathbf{T}_2\ldots\mathbf{T}_M\mathbf{B}.$$

Так как элементарные преобразования строк не меняют ранга, то

$$\operatorname{rang}\mathbf{AB} = \operatorname{rang}\mathbf{T}_1\mathbf{T}_2\ldots\mathbf{T}_M\mathbf{B} = \operatorname{rang}\mathbf{T}_2\ldots\mathbf{T}_M\mathbf{B} = \ldots = \operatorname{rang}\mathbf{T}_M\mathbf{B} = \operatorname{rang}\mathbf{B}.$$

Пусть для матрицы $\mathbf{C}$ определено произведение $\mathbf{CA}$. Так как элементарные преобразования столбцов не меняют ранга, то

$$\operatorname{rang}\mathbf{CA} = \operatorname{rang}\mathbf{CT}_1\mathbf{T}_2\ldots\mathbf{T}_M = \operatorname{rang}\mathbf{CT}_1\mathbf{T}_2\ldots\mathbf{T}_{M-1} = \ldots = \operatorname{rang}\mathbf{CT}_1 = \operatorname{rang}\mathbf{C}.$$

Предложение 2.6.5. *Ранг произведения двух матриц не превосходит рангов сомножителей.*

Действительно, пусть определено произведение $\mathbf{AB}$ (число столбцов в этом произведении совпадает с числом столбцов матрицы $\mathbf{B}$). Рассмотрим матрицу $\mathbf{D}$, составленную из строк матриц $\mathbf{B}$ и $\mathbf{AB}$. Так как $\mathbf{AB}$ является подматрицей матрицы $\mathbf{D}$, то ее ранг не превосходит ранга $\mathbf{D}$.

Далее, строки $\mathbf{AB}$ являются линейными комбинациями строк матрицы $\mathbf{B}$ (коэффициенты в i-й линейной комбинации равны элементам i-й строки матрицы $\mathbf{A}$). Поэтому, не меняя ранга матрицы $\mathbf{D}$, элементарными преобразованиями строк можно превратить все строки подматрицы $\mathbf{AB}$ в нулевые. Добавление нулевой строки не создает новых невырожденных подматриц, поэтому

$$\operatorname{rang}\mathbf{D} = \operatorname{rang}\mathbf{B},$$

$$\operatorname{rang}\mathbf{AB} \le \operatorname{rang}\mathbf{B}.$$

Аналогично доказывается что $\operatorname{rang}\mathbf{AB} \le \operatorname{rang}\mathbf{A}$.

Для вычисления ранга матрицы она элементарными преобразованиями приводится к такому виду, для которого базисная подматрица легко определяется.

Говорят, что матрица размеров $m\times n$ имеет *приведенно-ступенчатую форму*, если первые r ее столбцов являются первыми столбцами единичной матрицы порядка m, и, в случае $m > r$, нижние $m-r$ строк являются нулевыми.

Теорема 2.6.3. *Любой матрице элементарными преобразованиями строк и перестановкой столбцов можно придать приведенно-ступенчатую форму.*

Доказательство выполним конструктивно. Если матрица нулевая, то она уже имеет приведенно-ступенчатую форму. Иначе применим метод Гаусса. Если в очередной строке матрицы имеется ненулевой элемент, то его можно использовать для исключения всех остальных элементов, расположенных в этом же столбце. Если строка нулевая, то она переставляется с очередной ненулевой строкой и преобразования продолжаются.

Пусть после очередного r-го шага ненулевые строки исчерпаны. Тогда r столбцов матрицы превращены в некоторые столбцы единичной матрицы. Поэтому перестановкой строк и столбцов можно придать матрице приведенно-ступенчатую форму.

Предложение 2.6.6. Если все диагональные элементы треугольной матрицы отличны от нуля, то ранг этой матрицы совпадает с наименьшим из ее порядков.

В силу предложения 2.6.1 доказательство достаточно провести для верхней треугольной матрицы. В каждой i-й строке этой матрицы элемент a_{ii} отличен от нуля, поэтому используем его для исключения элементов, расположенных выше. Затем выполним деление i-й строки на элемент a_{ii}. Если наименьший из размеров матрицы равен r, то после r шагов ненулевые строки будут исчерпаны, а подматрица в первых r строках и r столбцах превращена в единичную.

Последнее предложение облегчает нахождение ранга. В процессе преобразований нет необходимости доводить матрицу до приведенно-ступенчатой формы – достаточно довести ее до треугольной.

Пусть, например, требуется найти ранг матрицы

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

Заменим вторую строку на разность ее с первой строкой, умноженной на 2, третью строку – на разность с первой, умноженной на 3; четвертую – на разность с первой:

$$\mathbf{A} \sim \begin{pmatrix} 1 & -1 & 2 & 3 \\ 0 & 3 & -5 & -2 \\ 0 & 5 & -8 & -8 \\ 0 & 3 & -5 & -2 \end{pmatrix}.$$

Заменим четвертую строку на разность со второй. Третью строку умножим на 3 и заменим на разность со второй, умноженной на 5.

Получим

$$\mathbf{A} \sim \mathbf{B} = \begin{pmatrix} 1 & -1 & 2 & 3 \\ 0 & 3 & -5 & -2 \\ 0 & 0 & 1 & -14 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Осталось исключить элементы a_{12}, a_{13} и a_{23}. Полученная матрица

$$\begin{pmatrix} 1 & -1 & 0 & 31 \\ 0 & 3 & 0 & -72 \\ 0 & 0 & 1 & -14 \\ 0 & 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 0 & 31 \\ 0 & 1 & 0 & -24 \\ 0 & 0 & 1 & -14 \\ 0 & 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 0 & 31 \\ 0 & 1 & 0 & -24 \\ 0 & 0 & 1 & -14 \\ 0 & 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 7 \\ 0 & 1 & 0 & -24 \\ 0 & 0 & 1 & -14 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

имеет приведенную ступенчатую форму. Она содержит строку из нулей и поэтому вырождена, однако ее подматрица, расположенная в первых трех строках и трех столбцах, является единичной (и поэтому невырождена). Следовательно,

$$\operatorname{rang} \mathbf{A} = 3.$$

Вновь отметим, что вывод о ранге матрицы можно было сделать уже после преобразования ее к треугольному виду: число ненулевых строк матрицы **B** равно трем (в этой матрице можно выделить треугольную подматрицу третьего порядка, диагональные элементы которой отличны от нуля).

Контрольные вопросы и упражнения

1. Какую подматрицу называют базисной? Что называют рангом матрицы?

2. Можно ли элементарными преобразованиями строк и столбцов изменить ранг матрицы?

3. Найти ранг матриц

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

4. Каковы порядок и ранг произведения

$$(1 \quad 2 \quad 3 \quad 4)^T(1 \quad 2 \quad 3 \quad 4)?$$

(проверить непосредственно и найти исходя из предложения 2.6.4).

2.7. Возведение матрицы в целочисленную степень

Пусть задана квадратная матрица $\mathbf{A} = (a_{ij})$ порядка n. Тогда для этой матрицы определено произведение $\mathbf{AA}$; его естественно обозначить $\mathbf{A}^2$. Аналогично определяется любая натуральная степень квадратной матрицы.

Например,

$$\begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}^3 = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 4 \\ -2 & -1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} -5 & 2 \\ -1 & -5 \end{pmatrix}.$$

В качестве нулевой степени квадратной матрицы по определению принимается единичная матрица того же порядка.

Естественно возникает вопрос, что следует принять в качестве отрицательной степени матрицы $\mathbf{A}$.

Для ненулевого действительного числа a по определению полагают

$$aa^{-1} = a^{-1}a = 1;$$

аналогично определяется отрицательная степень матрицы. Именно:

Матрица $\mathbf{A}^{-1}$ *называется обратной для матрицы* $\mathbf{A}$, *если выполнено условие*

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

где $\mathbf{E}$ − единичная матрица того же порядка, что и $\mathbf{A}$. Нахождение матрицы $\mathbf{A}^{-1}$ называют *обращением матрицы* $\mathbf{A}$.

Из определения следует ряд свойств.

1. $\left(\mathbf{A}^{-1}\right)^{-1} = \mathbf{A}$. Для доказательства достаточно умножить обе части слева на матрицу $\mathbf{A}^{-1}$:

$$\mathbf{A}^{-1}\left(\mathbf{A}^{-1}\right)^{-1} = \mathbf{A}^{-1}\mathbf{A}, \ \mathbf{A}^{-1}\left(\mathbf{A}^{-1}\right)^{-1} = \mathbf{E}.$$

2. $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$. Действительно, в силу ассоциативности умножения матриц:

$$(\mathbf{AB})\left(\mathbf{B}^{-1}\mathbf{A}^{-1}\right) = \mathbf{A}\left(\mathbf{BB}^{-1}\right)\mathbf{A}^{-1} = \mathbf{AEA}^{-1} = \mathbf{AA}^{-1} = \mathbf{E};$$

аналогично показывается, что $\left(\mathbf{B}^{-1}\mathbf{A}^{-1}\right)(\mathbf{AB}) = \mathbf{E}$.

Это свойство можно обобщить на произвольное число сомножителей:

$$(\mathbf{A}_1\mathbf{A}_2...\mathbf{A}_k)^{-1} = \mathbf{A}_k^{-1}\mathbf{A}_{k-1}^{-1}...\mathbf{A}_1^{-1}.$$

3. Матрица, обратная к транспонированной, может быть получена транспонированием обратной (операции обращения и транспонирования перестановочны):

$$\left(\mathbf{A}^T\right)^{-1} = \left(\mathbf{A}^{-1}\right)^T.$$

Доказательство основано на известном свойстве операции транспонирования: $(\mathbf{XY})^T = \mathbf{Y}^T\mathbf{X}^T$.

Транспонируя обе части равенства $\mathbf{AA}^{-1} = \mathbf{E}$, получим

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

Следовательно, матрица $\left(\mathbf{A}^{-1}\right)^T$ является обратной к матрице $\mathbf{A}^T$.

Предложение 2.7.1. *Если обратная матрица существует, то она единственна.*

Действительно, предположим что $\mathbf{A}_1$ и $\mathbf{A}_2$ являются для матрицы $\mathbf{A}$ обратными:

$$\mathbf{A}_1\mathbf{A} = \mathbf{AA}_1 = \mathbf{E}, \ \mathbf{A}_2\mathbf{A} = \mathbf{AA}_2 = \mathbf{E}.$$

Тогда

$$\mathbf{A}_2 = \mathbf{A}_2\mathbf{E} = \mathbf{A}_2(\mathbf{AA}_1) = (\mathbf{A}_2\mathbf{A})\mathbf{A}_1 = \mathbf{EA}_1 = \mathbf{A}_1.$$

Предложение 2.7.2. *Матрица имеет обратную тогда и только тогда, когда она невырождена.*

Докажем достаточность. Пусть матрица $\mathbf{A}$ невырождена. Тогда элементарными преобразованиями строк ее можно превратить в единичную:

$$\mathbf{T}_N\mathbf{T}_{N-1}...\mathbf{T}_1\mathbf{A} = \mathbf{E}.$$

Обозначая $\mathbf{A}_1 = \mathbf{T}_N \mathbf{T}_{N-1} \ldots \mathbf{T}_1$, получим: $\mathbf{A}_1 \mathbf{A} = \mathbf{E}$.

Далее, матрица $\mathbf{A}_1$ является произведением элементарных матриц и поэтому невырождена. Следовательно, существует матрица $\mathbf{A}_2$, такая, что

$$\mathbf{A}_2 \mathbf{A}_1 = \mathbf{E}.$$

Умножая последнее равенство справа на матрицу **A**, получим:

$$(\mathbf{A}_2 \mathbf{A}_1)\mathbf{A} = \mathbf{E}\mathbf{A}, \ \mathbf{A}_2 \mathbf{E} = \mathbf{A}, \ \mathbf{A}_2 = \mathbf{A},$$

и равенство $\mathbf{A}_2 \mathbf{A}_1 = \mathbf{E}$ принимает вид $\mathbf{A}\mathbf{A}_1 = \mathbf{E}$; таким образом, матрица $\mathbf{A}_1$ является обратной для матрицы **A**.

Докажем необходимость. Пусть матрица **A** имеет обратную. Покажем, что равная нулевому вектору линейная комбинация

$$\sum_{i=1}^{n} \beta_i \mathbf{a}^i = \mathbf{0}$$

строк матрицы **A** является тривиальной.

Действительно, такую линейную комбинацию можно записать в виде

$$\mathbf{b}\mathbf{A} = \mathbf{0},$$

где $\mathbf{b} = (\beta_1, \beta_2, \ldots, \beta_n)$. Умножая обе части справа на матрицу $\mathbf{A}^{-1}$, обратную к матрице **A**, получим:

$$\mathbf{b}\mathbf{A}\mathbf{A}^{-1} = \mathbf{0}\mathbf{A}^{-1}, \ \mathbf{b}\mathbf{E} = \mathbf{0}, \ \mathbf{b} = \mathbf{0}.$$

Следовательно, все коэффициенты β_i равны нулю. Поэтому строки матрицы линейно независимы и матрица **A** невырождена. Предложение доказано.

Предложение 2.7.3. *Матрица, обратная к диагональной, также является диагональной.*

Действительно, невырожденную диагональную матрицу можно представить в виде произведения

$$\mathbf{D} = \mathbf{S}_{1,a_{11}} \mathbf{S}_{2,a_{22}} \ldots \mathbf{S}_{n,a_{nn}}$$

элементарных матриц, каждая из которых получена из единичной матрицы $\mathbf{E}_n$ заменой k-й единицы на диагонали числом a_{kk}.

Поэтому

$$\mathbf{D}^{-1} = \left(\mathbf{S}_{1,a_{11}}\mathbf{S}_{2,a_{22}}...\mathbf{S}_{n,a_{nn}}\right)^{-1} = \mathbf{S}_{n,a_{nn}}^{-1}\mathbf{S}_{n-1,a_{n-1\,n-1}}^{-1}...\mathbf{S}_{1,a_{11}}^{-1} = \mathbf{S}_{n,\,1/a_{nn}}\mathbf{S}_{n-1,\,1/a_{n-1\,n-1}}...\mathbf{S}_{1,\,1/a_{11}}.$$

Так как произведение диагональных матриц в правой части само является диагональной матрицей, то $\mathbf{D}^{-1}$ – диагональная.

Заметим, что последнее предложение можно доказать иначе. Пусть $\mathbf{D} = \left(d_{ij}\right)$ – диагональная матрица: $d_{ij} = 0$ при $i \neq j$. Пусть $\mathbf{F} = \left(f_{ij}\right)$ – матрица, обратная к $\mathbf{D}$. Тогда должно быть выполнено условие $\mathbf{DF} = \mathbf{E}$, поэтому

$$\sum_{k=1}^{n} d_{ik} f_{kj} = \begin{cases} 1, i = j \\ 0, j \neq j \end{cases}, \quad d_{ii} f_{ij} = \begin{cases} 1, i = j \\ 0, j \neq j \end{cases}, \quad f_{ij} = \begin{cases} 1/d_{ii}, i = j \\ 0, j \neq j \end{cases}.$$

Предложение 2.7.4. *Матрица, обратная верхней треугольной, также является верхней треугольной.*

Доказательство выполняется аналогично (невырожденную верхнюю треугольную матрицу можно представить в виде произведения элементарных матриц, каждая из которых, вместе со своей обратной, является верхней треугольной).

Предложение 2.7.5. *Матрица, обратная к нижней треугольной, также является нижней треугольной.*

Для доказательства можно заметить, что если матрица $\mathbf{L}$ является нижней треугольной, то матрица $\mathbf{L}^T$ является верхней треугольной. Следовательно, матрица $\left(\mathbf{L}^T\right)^{-1} = \left(\mathbf{L}^{-1}\right)^T$ также является верхней треугольной. Поэтому $\mathbf{L}^{-1}$ – нижняя треугольная.

Теорема 2.7.1. *Пусть* $\mathbf{A}$ *– невырожденная матрица n-го порядка. Тогда любой столбец* $\mathbf{b}$ *высоты n можно представить в виде линейной комбинации столбцов матрицы* $\mathbf{A}$:

$$\mathbf{b} = \mathbf{Ax},$$

причем такое представление единственно.

Доказательство. Если матрица $\mathbf{A}$ невырождена, то она имеет обратную. В силу единственности обратной матрицы вектор-столбец $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ также определен однозначно. Умножая обе части равенства $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ слева на матрицу $\mathbf{A}$, получим:

$$\mathbf{Ax} = \mathbf{AA}^{-1}\mathbf{b}, \ \mathbf{Ax} = \mathbf{b}.$$

Теорема доказана.

Замечание 2.1. Утверждение последней теоремы можно сформулировать иначе: если матрица n-го порядка невырождена, то ее столбцы образуют базис координатного пространства K_n. Аналогичное утверждение справедливо для строк.

Способ обращения матрицы устанавливает предложение 2.7.2. Умножая обе части равенства

$$\mathbf{T}_N\mathbf{T}_{N-1}...\mathbf{T}_1\mathbf{A}=\mathbf{E}$$

справа на матрицу $\mathbf{A}^{-1}$, получим:

$$\mathbf{T}_N\mathbf{T}_{N-1}...\mathbf{T}_1\mathbf{A}\mathbf{A}^{-1}=\mathbf{E}\mathbf{A}^{-1},$$

$$\mathbf{T}_N\mathbf{T}_{N-1}...\mathbf{T}_1\mathbf{E}=\mathbf{A}^{-1}.$$

Следовательно, если элементарными преобразованиями строк матрица $\mathbf{A}$ приводится к единичной, то та же последовательность элементарных преобразований приводит единичную матрицу к матрице $\mathbf{A}^{-1}$.

Аналогично, умножая обе части выражения

$$\mathbf{A}\mathbf{T}_1\mathbf{T}_2...\mathbf{T}_M=\mathbf{E}$$

слева на матрицу $\mathbf{A}^{-1}$, найдем

$$\mathbf{A}^{-1}\mathbf{A}\mathbf{T}_1\mathbf{T}_2...\mathbf{T}_M=\mathbf{A}^{-1}\mathbf{E},$$

$$\mathbf{E}\mathbf{T}_1\mathbf{T}_2...\mathbf{T}_M=\mathbf{A}^{-1}.$$

Поэтому если элементарными преобразованиями столбцов матрица $\mathbf{A}$ приводится к единичной, то та же последовательность элементарных преобразований приводит единичную матрицу к матрице $\mathbf{A}^{-1}$.

Пусть, например, следует обратить матрицу

$$\mathbf{A}=\begin{pmatrix}1 & 0 & 1\\ 1 & 1 & 0\\ 3 & 1 & 1\end{pmatrix}.$$

Дополним ее справа единичной матрицей и, выполняя элементарные преобразования строк, приведем левую часть к единичной матрице. Для этого достаточно выполнить следующие преобразования:

– замену второй строки на разность с первой строкой, замену третьей строки на разность с утроенной первой строкой;

– замену третьей стоки на разность со второй строкой, умножение третьей строки на -1;

– замену второй строки на сумму с третьей строкой, замену первой строки на разность с третьей строкой.

Имеем:

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

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

Обратная матрица:

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

Обращение можно выполнить иначе. Дополним матрицу $\mathbf{A}$ снизу единичной матрицей и элементарными преобразованиями столбцов приведем к единичной матрице верхнюю часть:

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

Необходимость определения *дробной* степени квадратной матрицы обычно в приложениях не возникает, поэтому подробно соответствующий вопрос здесь мы не затрагиваем. Приведем лишь несколько примеров.

Пусть требуется *извлечь квадратный корень* из матрицы $\mathbf{A}$ второго порядка, т.е. найти матрицу $\mathbf{B}$ для которой $\mathbf{B}^2 = \mathbf{A}$. Обозначим

$$\mathbf{A} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix},\ \mathbf{B} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Тогда

$$\mathbf{B}^2 = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2+bc & ab+bd \\ ca+dc & cb+d^2 \end{pmatrix}.$$

Приравнивая элементы матриц **A** и $\mathbf{B}^2$, получим систему *нелинейных* уравнений относительно неизвестных a, b, c и d:

$$\begin{cases} a^2+bc=\alpha \\ ab+bd=\beta \\ ca+dc=\gamma \\ cb+d^2=\delta \end{cases}. \qquad (2.1)$$

Множество аналитических решений этой системы, записанное в конечном виде, отличается чрезвычайной громоздкостью. Сравнительно компактная запись решения возможна только в некоторых специальных случаях.

Пусть матрица **A** является единичной. Тогда система (2.1), из которой определяются элементы матрицы **B**, будет иметь вид:

$$\begin{cases} a^2+bc=1 \\ ab+bd=0 \\ ca+dc=0 \\ cb+d^2=1 \end{cases}.$$

Решениями этой системы будут:

$$\mathbf{B}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},\ \mathbf{B}_2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},\ \mathbf{B}_3 = \begin{pmatrix} \lambda & \mu \\ \dfrac{1-\lambda^2}{\mu} & -\lambda \end{pmatrix},\ \mathbf{B}_4 = \begin{pmatrix} \lambda & \dfrac{1-\lambda^2}{\mu} \\ \mu & -\lambda \end{pmatrix},$$

где λ и μ – произвольные числа.

Уже только приведенный пример свидетельствует, что корень квадратный из матрицы имеет бесконечно много значений.

Контрольные вопросы и упражнения

1. Какая матрица называется обратной к данной?

2. Сформулируйте необходимое и достаточное условие обратимости матрицы.

3. Вычислите

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}^2; \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}^4; \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}^6; \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{10}.$$

4*. Вычислите

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}^n.$$

5 [10]. Вычислите

$$\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}^n.$$

6 [9]. Вычислите

$\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}^n; \begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix}^n$.7. Обратите матрицы или докажите их вырожденность:

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

Выполните проверку полученных результатов.

8. Вычислите

$$16\begin{pmatrix} 3 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 3 \end{pmatrix}^{-2}.$$

9. Пусть дана прямоугольная матрица $\mathbf{A}$ размера $m\times n$, причем $m \neq n$. Говорят, что матрица $\mathbf{A}_1$ является *левой обратной*, если $\mathbf{A}_1\mathbf{A} = \mathbf{E}$; говорят, что $\mathbf{A}_2$ является *правой обратной*, если $\mathbf{A}\mathbf{A}_2 = \mathbf{E}$. Могут ли левая и правая обратные матрицы совпадать?

10*. Докажите, что при любом $z \in C$ квадрат симметрической матрицы

$$\begin{pmatrix} z & z^* \\ z^* & z \end{pmatrix}$$

будет образован только действительными числами (здесь z^* – число, комплексно сопряженное z).

11* [9]. Найдите все матрицы нулевого порядка, квадраты которых равны нулевой матрице.

12*. Укажите хотя бы одну матрицу **A**, для которой

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

2.8. Детерминант

Отношение линейной зависимости разбивает все системы векторов на два класса – линейно зависимые и линейно независимые. Однако может возникнуть необходимость выполнить сравнение двух систем векторов, принадлежащих одному классу.

Вопрос о том, в какой мере одна система векторов «более линейно независима» нежели другая, не лишен смысла. При поиске ответа на этот вопрос оказывается полезной такая числовая функция, которая равна нулю для вырожденных матриц и принимает некоторые ненулевые значения для невырожденных.

Матрица нулевого порядка (не содержащая ни одного элемента) по определению невырождена. *Детерминантом* (или *определителем*) *нулевого порядка* мы назовем число 1.

Матрица первого порядка, состоящая из одного элемента, вырождена тогда и только тогда, когда этот элемент нулевой. *Детерминантом первого порядка* матрицы из одного элемента мы назовем этот элемент:

$$\det(a_{11}) = a_{11}.$$

Два вектора координатного пространства линейно зависимы тогда и только тогда, когда их координаты пропорциональны. Поэтому квадратная матрица второго порядка

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

вырождена тогда и только тогда, когда элементы ее строк пропорциональны:

$$\begin{cases} a_{21} = \lambda a_{11} \\ a_{22} = \lambda a_{12} \end{cases},$$

где λ – любое число (возможно, ноль).

Условие пропорциональности элементов строк можно записать в виде:

$$\lambda a_{11} a_{22} = a_{21} \lambda a_{12},$$

или

$$a_{11} a_{22} - a_{21} a_{12} = 0.$$

Детерминантом второго порядка мы назовем разность элементов главной и побочной диагоналей соответствующей матрицы:

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

Данные определения позволяют установить три следующих важнейших свойства детерминанта.

Свойство 1. *Детерминант является линейным однородным многочленом от элементов любой строки*:

$$\det \mathbf{A} = \sum_{j=1}^{n} a_{ij} h_{ij},$$

где ни один из коэффициентов h_{ij} не зависит ни от одного из элементов i-й строки.

Действительно, для матрицы второго порядка

$$\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11} a_{22} - a_{12} a_{21} = a_{11} h_{11} + a_{12} h_{12},$$

где коэффициенты $h_{11} = a_{22}$, $h_{12} = -a_{21}$ не зависят ни от одного из элементов второй строки.

С другой стороны:

$$\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11} a_{22} - a_{12} a_{21} = a_{21} h_{21} + a_{22} h_{22},$$

где коэффициенты $h_{21} = -a_{12}$, $h_{22} = a_{11}$ не зависят ни от одного из элементов первой строки.

Свойство 2. *Детерминант вырожденной матрицы равен нулю*.

Это свойство, по сути, положено в основу определений. Например, для матрицы второго порядка:

$$\det\begin{pmatrix} a & b \\ \lambda a & \lambda b \end{pmatrix} = a\lambda b - b\lambda a = 0 .$$

Свойство 3. *Детерминант единичной матрицы равен единице*.

Действительно, детерминант нулевого порядка равен единице по определению. Единичная матрица первого порядка содержит в качестве единственного элемента единицу, поэтому

$$\det(1) = 1 .$$

Для матрицы второго порядка:

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

Теперь для квадратных матриц произвольного порядка можно следующим образом определить числовую функцию, которая в некотором роде является «мерой линейной независимости» рядов:

Детерминантом матрицы называется любая числовая функция, обладающая свойствами 1...3 (эти свойства теперь понимаются как *аксиомы*).

Существование и единственность подобной функции для матрицы порядка $n > 2$ *не очевидны и требуют доказательства*. Прежде всего следует изучить те свойства, которые следуют непосредственно из определения. До окончания доказательства термин *детерминант* будет использован именно в смысле *«любая функция, обладающая свойствами 1...3»*.

Предложение 2.8.1. *Если i-я строка матрицы* $\mathbf{A}$ *является линейной комбинацией* $\mathbf{a}^i = \alpha\mathbf{p} + \beta\mathbf{q}$ *двух строк* $\mathbf{p}$ *и* $\mathbf{q}$, *то детерминант равен*

$$\det \mathbf{A} = \alpha \det \mathbf{A}_p + \beta \det \mathbf{A}_q ,$$

где матрицы $\mathbf{A}_p$ и $\mathbf{A}_q$ получены из матрицы $\mathbf{A}$ заменой i-й строки на строки $\mathbf{p}$ и $\mathbf{q}$, соответственно.

Доказательство. Пусть i-я строка $\mathbf{A}$ является линейной комбинацией $\alpha\mathbf{p}+\beta\mathbf{q}$ строк $\mathbf{p}=(p_j)$ и $\mathbf{q}=(q_j)$. Тогда элементы этой строки равны

$$a_{ij}=\alpha p_j+\beta q_j .$$

Так как детерминант обладает свойством линейности по строкам, то

$$\det\mathbf{A}=\sum_{j=1}^{n}a_{ij}h_{ij}=\sum_{j=1}^{n}(\alpha p_j+\beta q_j)h_{ij}=\alpha\sum_{j=1}^{n}p_jh_{ij}+\beta\sum_{j=1}^{n}q_jh_{ij} .$$

Здесь коэффициенты h_{ij} не зависят от элементов i-й строки. Поэтому

$$\sum_{j=1}^{n}p_jh_{ij}=\det\mathbf{A}_p \text{ и } \sum_{j=1}^{n}q_jh_{ij}=\det\mathbf{A}_q .$$

Следовательно, $\det\mathbf{A}=\alpha\det\mathbf{A}_p+\beta\det\mathbf{A}_q$.

Последнее свойство называют свойством *линейности по строке* и часто формулируют в виде двух отдельных утверждений:

– *множитель элементов любой строки можно вынести за знак детерминанта*;

– *если какая-либо из строк матрицы* $\mathbf{A}$ *является суммой двух строк* $\mathbf{p}$ *и* $\mathbf{q}$, *то детерминант матрицы* $\mathbf{A}$ *равен сумме детерминантов матриц, получаемых из* $\mathbf{A}$ *заменой этой строки на строки* $\mathbf{p}$ *и* $\mathbf{q}$.

Предложение 2.8.2. *Если к некоторой строке матрицы прибавить другую ее строку, умноженную на любое число, то детерминант матрицы не изменится.*

Действительно, пусть матрица $\mathbf{B}$ получена заменой i-й строки матрицы $\mathbf{A}$ на сумму ее с другой k-й строкой, умноженной на число λ; пусть матрица $\mathbf{C}$ получена заменой i-й строки матрицы $\mathbf{A}$ на k-ю строку. Тогда по свойству линейности

$$\det\mathbf{B}=\det\mathbf{A}+\lambda\det\mathbf{C}.$$

Но матрица $\mathbf{C}$ содержит две одинаковые строки (в нее дважды входит k-я строка матрицы $\mathbf{A}$) и поэтому является вырожденной. Отсюда $\det\mathbf{C}=0$, и

$$\det\mathbf{B}=\det\mathbf{A} .$$

Предложение 2.8.3. *При перестановке двух строк детерминант меняет знак.*

Для доказательства выполним известную из п. 0 последовательность преобразований:

– замену k-й строки разностью ее с s-й;

– замену s-й строки суммой с k-й;

– замену k-й строки разностью с s-й;

– умножение k-й строки на -1.

Первые три преобразования не меняют детерминанта; при выполнении последнего преобразования детерминант меняет знак. Предложение доказано.

Предложение 2.8.4. *Если заданная на множестве квадратных матриц порядка $n>1$ числовая функция $f(\mathbf{A})$ линейна по строкам и для матриц с двумя одинаковыми строками равна нулю, то она равна нулю для всех вырожденных матриц.*

Доказательство. Пусть матрица $\mathbf{A}$ вырождена. Так как число строк $n>1$ и они линейно зависимы, то одна из строк является линейной комбинацией остальных:

$$\mathbf{a}^k = \sum_{\substack{i=1 \\ i\neq k}}^{n} \alpha_i \mathbf{a}^i .$$

Пусть матрицы $\mathbf{A}_i$ получены из матрицы $\mathbf{A}$ заменой i-й строки на k-ю строку (здесь $i=\overline{1,n}$, $i\neq k$). Тогда для этих матриц $f(\mathbf{A}_i)=0$.

Последовательно применяя свойство линейности по i-й строке, получим:

$$f(\mathbf{A}) = \alpha_i \sum_{\substack{i=1 \\ i\neq k}}^{n} f(\mathbf{A}_i) = 0 .$$

Предложение 2.8.5. *Детерминант элементарной матрицы $\mathbf{S}_{k,\lambda}$, полученной единичной заменой k-й единицы на диагонали числом λ, равен λ. Детерминант элементарной матрицы $\mathbf{C}_{kl}$, полученной из единичной заменой на единицу нулевого элемента на пересечении k-й строки и l-го столбца, равен* 1.

Доказательство. Согласно определению детерминанта и свойству линейности по строке имеем:

$$\det \mathbf{S}_{k,\lambda} = \lambda \det \mathbf{E} = \lambda ;$$

$$\det \mathbf{C}_{kl} = \det \mathbf{E} = 1.$$

Как следствие:

Предложение 2.8.6. *Если существуют две функции* $f_1(\mathbf{A})$ *и* $f_2(\mathbf{A})$*, удовлетворяющие определению детерминанта, то для любой элементарной матрицы значения этих функций совпадают.*

Предложение 2.8.7. *Для любой матрицы* $\mathbf{A}$ *и любой элементарной матрицы* $\mathbf{T}$ *выполнено равенство*

$$\det(\mathbf{TA}) = \det \mathbf{T} \det \mathbf{A}. \tag{2.2}$$

Действительно, матрица $\mathbf{S}_{k,\lambda}\mathbf{A}$ может быть получена из матрицы $\mathbf{A}$ умножением k-й строки на число λ, поэтому $\det \mathbf{S}_{k,\lambda}\mathbf{A} = \lambda \det \mathbf{A}$. Так как $\det \mathbf{S}_{k,\lambda} = \lambda$, то равенство $\det \mathbf{S}_{k,\lambda}\mathbf{A} = \det \mathbf{S}_{k,\lambda} \det \mathbf{A}$ выполнено.

Матрица $\mathbf{C}_{kl}\mathbf{A}$ может быть получена из матрицы $\mathbf{A}$ заменой k-й строки на сумму ее l-й строкой, поэтому $\det \mathbf{C}_{kl}\mathbf{A} = \det \mathbf{A}$. Так как $\det \mathbf{C}_{kl} = 1$, то равенство $\det \mathbf{C}_{kl}\mathbf{A} = \det \mathbf{C}_{kl} \det \mathbf{A}$ выполнено. Предложение доказано.

Единственность детерминанта устанавливает следующая теорема.

Теорема 2.8.1. *На множестве квадратных матриц данного порядка не может быть более одной функции, удовлетворяющей определению детерминанта.*

Доказательство. Предположим, что функции $f_1(\mathbf{A})$ и $f_2(\mathbf{A})$ удовлетворяют определению детерминанта. Покажем, что $f_1(\mathbf{A}) = f_2(\mathbf{A})$ для любой квадратной матрицы $\mathbf{A}$.

Если матрица $\mathbf{A}$ вырождена, то по определению $f_1(\mathbf{A}) = f_2(\mathbf{A}) = 0$.

Пусть $\mathbf{A}$ невырождена. Тогда она может быть разложена в произведение элементарных матриц:

$$\mathbf{A} = \mathbf{T}_1\mathbf{T}_2 \ldots \mathbf{T}_M.$$

Последовательно применяя соотношение (2.2), получим:

$$f_1(\mathbf{A}) = f_1(\mathbf{T}_1\mathbf{T}_2 \ldots \mathbf{T}_M) = f_1(\mathbf{T}_1) f_1(\mathbf{T}_2) \ldots f_1(\mathbf{T}_M);$$

аналогично

$$f_2(\mathbf{A}) = f_2(\mathbf{T}_1\mathbf{T}_2 \ldots \mathbf{T}_M) = f_2(\mathbf{T}_1) f_2(\mathbf{T}_2) \ldots f_2(\mathbf{T}_M).$$

Но из предложения 2.8.6 следует, что

$$f_1(\mathbf{T}_1) = f_2(\mathbf{T}_1),\ f_1(\mathbf{T}_2) = f_2(\mathbf{T}_2),\ \ldots,\ f_1(\mathbf{T}_M) = f_2(\mathbf{T}_M).$$

Поэтому $f_1(\mathbf{A}) = f_2(\mathbf{A})$.

Вместе с доказательством теоремы было установлено, что детерминант невырожденной матрицы равен произведению детерминантов элементарных сомножителей в ее разложении:

$$\det \mathbf{A} = \det \mathbf{T}_1 \det \mathbf{T}_2 \ldots \det \mathbf{T}_M . \qquad (2.3)$$

Так как детерминанты элементарных матриц

$$\det \mathbf{S}_{k,\lambda} = \lambda \neq 0, \ \det \mathbf{C}_{kl} = 1,$$

то из последнего равенства можно заключить, что *детерминант невырожденной матрицы отличен от нуля*. Как следствие, имеет место теорема:

Теорема 2.8.2. *Матрица вырождена тогда и только тогда, когда ее детерминант равен нулю.*

Обратимся теперь к доказательству *существования* детерминанта. Пусть дана матрица n-го порядка:

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

Если в матрице $\mathbf{A}$ вычеркнуть i-ю строку и j-й столбец, то останется подматрица $(n-1)$-го порядка $\mathbf{M}_{ij}$, которая называется *дополнительной подматрицей* элемента a_{ij}. Детерминант $M_{ij} = \det \mathbf{M}_{ij}$ дополнительной подматрицы называется *дополнительным минором* элемента a_{ij}.

Алгебраическим дополнением (или *адъюнктом*) A_{ij}яэлемента a_{ij} называется дополнительный минор, взятый со своим или противоположным знаком в зависимости от того, будет ли сумма индексов $i+j$ четной или нечетной:

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

Теорема 2.8.3. *На множестве квадратных матриц n-го порядка существует функция, удовлетворяющая определению детерминанта.*

Доказательство существования детерминанта произвольного порядка выполним *по индукции*: проверим существование для некоторого фиксированного *начального* порядка (говорят, что выполняется *база индукции*), предположим существование для произвольного порядка $n-1$ (говорят, что выполняется *индуктивное предположение*) и затем докажем существование для порядка n.

База индукции тривиальна: свойства, положенные в основу определения, были первоначально установлены для детерминанта второго порядка.

Предположим, что для матриц $(n-1)$-го порядка функция с требуемыми свойствами существует.

Фиксируем произвольно номер столбца j и поставим в соответствие матрице n-го порядка число

$$f_j(\mathbf{A}) = \sum_{i=1}^{n} a_{ij} A_{ij} = \sum_{i=1}^{n} a_{ij} (-1)^{i+j} M_{ij}\,, \tag{2.4}$$

где M_{ij} – дополнительный минор элемента a_{ij} (дополнительные миноры являются детерминантами $(n-1)$-го порядка и существуют в силу индуктивного предположения). Покажем, что функция (2.4) удовлетворяет определению детерминанта.

Выберем в матрице $\mathbf{A}$ произвольную k-ю строку. Тогда при $i=k$ слагаемое $a_{kj}A_{kj}$ содержит элемент k-й строки. Коэффициент A_{kj} при этом элементе не зависит от элементов k-й строки, так как эта строка не входит в дополнительную подматрицу $\mathbf{M}_{kj}$ элемента a_{kj}. Поэтому слагаемое $a_{kj}A_{kj}$ является линейным многочленом от элемента a_{kj}. Далее, при $i \neq k$ в слагаемых $a_{ij}A_{ij}$ множитель a_{ij} не принадлежит k-й строке, а множитель $A_{ij} = (-1)^{i+j} \det \mathbf{M}_{ij}$ является линейным многочленом от элементов k-й строки (по индуктивному предположению). Тогда функция (2.4) является суммой линейных многочленов от элементов k-й строки; следовательно, она сама является линейным многочленом от элементов этой строки. Свойство 1 выполнено.

Покажем, что для вырожденной матрицы $f_j(\mathbf{A}) = 0$. В силу предложения 2.8.4 и доказанной линейности по строке достаточно проверить, что $f_j(\mathbf{A}) = 0$ для произвольной матрицы, имеющей две одинаковые строки $\mathbf{a}^k = \mathbf{a}^l$ (без ограничения общности можно потребовать $l > k$). Тогда в сумме (2.4) отличны от нуля только два слагаемых, так как при $i \neq k$ и $i \neq l$ дополнительная подматрица $\mathbf{M}_{ij}$

содержит две одинаковые строки и ее детерминант равен нулю по индуктивному предположению. Следовательно, с учетом $a_{kj} = a_{lj}$:

$$f_j(\mathbf{A}) = a_{kj}(-1)^{k+j} \det \mathbf{M}_{kj} + a_{lj}(-1)^{l+j} \det \mathbf{M}_{lj} =$$

$$= (-1)^j a_{kj}\left((-1)^k \det \mathbf{M}_{kj} + (-1)^l \det \mathbf{M}_{lj}\right).$$

Дополнительные подматрицы $\mathbf{M}_{kj}$ и $\mathbf{M}_{lj}$ отличаются только порядком строк: в каждой из них остался отрезок одной из двух одинаковых строк, но в матрице $\mathbf{M}_{lj}$ он находится на k-м месте, а в матрице $\mathbf{M}_{kj}$ – на $(l-1)$-м.

Переставим в матрице $\mathbf{M}_{kj}$ строку с номером $l-1$ на k-е место, не нарушая порядка остальных строк. Для этого последовательно будем переставлять ее с $(l-2)$-й, $(l-3)$-й, …, k-й строкой. Всего потребуется $l-k-1$ перестановок; при каждой такой перестановке детерминант $\det \mathbf{M}_{kj}$ меняет знак (вновь, по индуктивному предположению). Следовательно,

$$\det \mathbf{M}_{lj} = (-1)^{l-k-1} \det \mathbf{M}_{kj},$$

$$f_j(\mathbf{A}) = (-1)^j a_{kj}\left((-1)^k \det \mathbf{M}_{kj} + (-1)^l (-1)^{l-k-1} \det \mathbf{M}_{kj}\right) =$$

$$= (-1)^{j+k} a_{kj} \det \mathbf{M}_{kj}\left(1 + (-1)^{2(l-k)-1}\right) = 0.$$

Свойство 2 выполнено.

Пусть $\mathbf{E}$ – единичная матрица n-го порядка. Тогда в j-м столбце все элементы кроме $a_{jj} = 1$ равны нулю, и в сумме (2.4) остается одно слагаемое:

$$f_j(\mathbf{E}) = (-1)^{j+j} \det \mathbf{M}_{jj}.$$

Но $\mathbf{M}_{jj}$ – единичная матрица порядка $n-1$, и ее детерминант равен единице по индуктивному предположению. Следовательно, $f_j(\mathbf{E}) = 1$; свойство 3 выполнено. Теорема доказана.

Таким образом, *детерминантом* матрицы **A** называется сумма произведений элементов произвольного j-го столбца на их алгебраические дополнения:

$$\det \mathbf{A} = \sum_{i=1}^{n} a_{ij} A_{ij} = \sum_{i=1}^{n} a_{ij} (-1)^{i+j} M_{ij} = \sum_{i=1}^{n} a_{ij} (-1)^{i+j} \det \mathbf{M}_{ij} . \qquad (2.5)$$

Соотношение (2.5) называется *разложением* детерминанта по j-му столбцу. Так как правая часть (2.5) является линейным многочленом от элементов j-го столбца, то имеет место предложение:

Предложение 2.8.8. *Детерминант обладает свойством линейности по столбцам.*

Соотношение (2.5) устанавливает возможность разложения детерминанта по любому столбцу. Естественно возникает вопрос о возможности разложения детерминанта по любой строке. Ответ на этот вопрос дает следующее предложение:

Предложение 2.8.9. *Для произвольной i-й строки детерминанта справедлива формула разложения по строке*:

$$\det \mathbf{A} = \sum_{j=1}^{n} a_{ij} A_{ij} = \sum_{j=1}^{n} a_{ij} (-1)^{i+j} M_{ij} = \sum_{j=1}^{n} a_{ij} (-1)^{i+j} \det \mathbf{M}_{ij} , \qquad (2.6)$$

где M_{ij} – дополнительный минор элемента a_{ij}.

Доказательство. Сгруппируем слагаемые в (2.5) следующим образом:

$$\det \mathbf{A} = \sum_{i=1}^{n} a_{ij} (-1)^{i+j} M_{ij} = a_{ij} (-1)^{i+j} M_{ij} + \sum_{\substack{k=1 \\ k \neq i}}^{n} a_{kj} (-1)^{k+j} M_{kj} = a_{ij} (-1)^{i+j} M_{ij} + r .$$

Но детерминант является линейным однородным многочленом от элементов i-й строки:

$$\det \mathbf{A} = \sum_{j=1}^{n} a_{ij} h_{ij} = a_{ij} h_{ij} + \sum_{\substack{k=1 \\ k \neq j}}^{n} a_{ik} h_{ik} = a_{ij} h_{ij} + q .$$

По определению, h_{ij} не зависит от элементов i-й строки, а q содержит все ее элементы, кроме a_{ij}. Аналогично, при любом i в дополнительный минор M_{ij} не входит j-й столбец, поэтому ни один из дополнительных миноров не зависит от элемента a_{ij}. Следовательно, сумма r также не зависит от этого элемента.

Заменим в матрице $\mathbf{A}$ элемент a_{ij} нулем (подобная замена не сказывается ни на значении r, ни на значении q) и обозначим полученную матрицу $\mathbf{A}_0$. Получим: $\det \mathbf{A}_0 = r$, $\det \mathbf{A}_0 = q$, откуда $q = r$. Заменим элемент a_{ij} единицей; тогда $\det \mathbf{A}_1 = (-1)^{i+j} M_{ij} + r$, $\det \mathbf{A}_1 = h_{ij} + q$, откуда с учетом $q = r$ получим

$$h_{ij} = (-1)^{i+j} M_{ij}.$$

Предложение доказано.

Линейность детерминанта по столбцам с учетом ранее доказанной единственности приводит к следующему предложению:

Предложение 2.8.10. *Детерминант не меняется при транспонировании:* $\det \mathbf{A} = \det \mathbf{A}^T$.

Действительно, рассмотрим функцию $f(\mathbf{A}) = \det \mathbf{A}^T$. Эта функция линейна по столбцам матрицы $\mathbf{A}^T$. Следовательно, она линейна по строкам матрицы $\mathbf{A}$ и входящее в определение свойство 1 для нее выполнено. Далее, если матрица $\mathbf{A}$ вырождена, то вырожденной будет также матрица $\mathbf{A}^T$. Поэтому для вырожденной матрицы $\mathbf{A}$

$$f(\mathbf{A}) = \det \mathbf{A}^T = 0;$$

свойство 2 выполнено.

Наконец, единичная матрица не меняется при транспонировании; поэтому $f(\mathbf{E}) = \det \mathbf{E}^T = \det \mathbf{E} = 1$, и свойство 3 также выполнено. Предложение доказано.

Доказательство этого предложения можно выполнить индуктивно. База выполнена:

$$\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc = ad - cb = \det\begin{pmatrix} a & c \\ b & d \end{pmatrix}$$

После индуктивного предположения следует разложить детерминант матрицы $\mathbf{A}$ по k-й строке, разложить детерминант матрицы $\mathbf{A}^T$ по k-му столбцу и сравнить полученные разложения.

Последнее предложение устанавливает равноправность строк и столбцов: любое утверждение о детерминантах, сформулированное для строк, остается верным и для столбцов. Поэтому в частности:

1. Столбцы матрицы линейно зависимы тогда и только тогда, когда детерминант матрицы равен нулю.

2. При перестановке двух столбцов детерминант меняет знак.

3. Если к одному из столбцов прибавить другой, умноженный на любое число, то детерминант не изменится.

Одним из важнейших следствий теоремы 2.8.3 и предложения 2.8.8 является:

Предложение 2.8.11. *Сумма произведений элементов любого ряда детерминанта на алгебраические дополнения элементов любого другого параллельного ряда равна нулю:*

$$\sum_{j=1}^{n} a_{sj} A_{kj} = \sum_{i=1}^{n} a_{is} A_{ik} = 0 .$$

Доказательство достаточно выполнить для строк. Раскладывая детерминант по k-й строке, получим:

$$\det\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & & & \\ a_{k1} & a_{k2} & \dots & a_{kn} \\ \dots & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} = \sum_{j=1}^{n} a_{kj} A_{kj} .$$

Так как алгебраические дополнения A_{kj} не зависят от элементов k-й строки, то равенство сохранится при замене (в левой и правой частях) всех чисел a_{kj} любыми другими n числами. Заменяя a_{kj} соответствующими элементами s-й строки ($s \neq k$), получим в левой части детерминант матрицы с двумя совпадающими строками:

$$\det\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & & & \\ a_{s1} & a_{s2} & \dots & a_{sn} \\ \dots & & & \\ a_{s1} & a_{s2} & \dots & a_{sn} \\ \dots & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} = \sum_{j=1}^{n} a_{sj} A_{kj} .$$

Но матрица с двумя совпадающими строками вырождена, поэтому такой детерминант равен нулю:

$$\sum_{j=1}^{n} a_{sj} A_{kj} = 0 .$$

Из предложения 2.8.7 и соотношения (2.3) следует предложение:

Предложение 2.8.12. *Детерминант произведения квадратных матриц равен произведению их детерминантов:*

$$\det \mathbf{AB} = \det \mathbf{A} \det \mathbf{B} .$$

Доказательство. Пусть матрица $\mathbf{A}$ невырождена. Тогда она может быть разложена в произведение элементарных матриц: $\mathbf{A} = \mathbf{T}_1\mathbf{T}_2...\mathbf{T}_M$. Последовательно применяя соотношение (2.2), получим

$$\det \mathbf{AB} = \det \mathbf{T}_1 \det \mathbf{T}_2 ... \det \mathbf{T}_M \det \mathbf{B} .$$

С другой стороны, $\det \mathbf{A} = \det \mathbf{T}_1 \det \mathbf{T}_2 ... \det \mathbf{T}_M$. Поэтому

$$\det \mathbf{AB} = \det \mathbf{A} \det \mathbf{B} .$$

Если матрица $\mathbf{A}$ вырождена, то ее ранг меньше ее порядка. Следовательно, ранг произведения $\mathbf{AB}$ также меньше его порядка, и

$$\det \mathbf{A} \det \mathbf{B} = 0 = \det \mathbf{AB} .$$

Предложение доказано.

Пусть дана прямоугольная матрица размера $m \times n$, имеющая ранг r (здесь $r \le n$, $r \le m$) Тогда в этой матрице можно выделить r линейно независимых строк и r линейно независимых столбцов, на пересечении которых находится базисная подматрица (см. п. 2.6) порядка r, и все подматрицы большего порядка вырождены. Детерминант базисной подматрицы называется *базисным минором*.

Из теоремы 2.8.2 сразу следует теорема:

Теорема 2.8.4. *Ранг матрицы равен наибольшему порядку отличных от нуля миноров*.

Контрольные вопросы и упражнения

1. Сформулируйте определения детерминантов от нулевого до второго порядков.

2. Сформулируйте необходимое и достаточное условие вырожденности матрицы. Исходя из этого условия установите, являются ли вырожденными матрицы

$$\begin{pmatrix} -2 & 1 \\ 3 & 2 \end{pmatrix}; \begin{pmatrix} 2 & 6 \\ 1 & 3 \end{pmatrix}; \begin{pmatrix} -4 & 2 \\ -2 & 1 \end{pmatrix}.$$

3. Как связан наибольший порядок отличных от нуля миноров и ранг матрицы?

В дальнейшем для обозначения детерминанта мы часто будем использовать одинарные прямые скобки:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}.$$

Детерминант можно вычислить, используя соотношение (2.5) или (2.6). Например:

$$\begin{vmatrix} 1 & 3 & -2 \\ 2 & -1 & 4 \\ 1 & -3 & -2 \end{vmatrix} = 1\cdot\begin{vmatrix} -1 & 4 \\ -3 & -2 \end{vmatrix} - 3\cdot\begin{vmatrix} 2 & 4 \\ 1 & -2 \end{vmatrix} - 2\cdot\begin{vmatrix} 2 & -1 \\ 1 & -3 \end{vmatrix} =$$

$$= 1\cdot\big((-1)\cdot(-2) + 4\cdot 3\big) - 3\cdot\big(2\cdot(-2) - 4\cdot 1\big) - 2\cdot\big(2\cdot(-3) + 1\cdot 1\big) = 14 + 24 + 10 = 48.$$

На практике разложение следует вести по ряду (строке или столбцу), содержащему наибольшее количество нулевых элементов. Так, детерминант

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

можно разложить по второй строке:

$$\begin{vmatrix} 1 & 2 & 0 \\ 2 & 0 & 0 \\ -1 & 1 & 3 \end{vmatrix} = -2\begin{vmatrix} 2 & 0 \\ 1 & 3 \end{vmatrix} = -12.$$

Соотношения (2.5) или (2.6) можно использовать для того, чтобы выразить детерминант непосредственно через его элементы, получив *формулу полного разложения*. Например, для матрицы третьего порядка:

$$\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11}(-1)^2\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + a_{12}(-1)^3\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}(-1)^4\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} =$$

$$= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}.$$

Порядок сомножителей в каждом слагаемом правой части найденной формулы полного разложения выражается *правилом Саррюса*. Первое слагаемое есть произведение элементов главной диагонали. Во втором и третьем слагаемых два сомножителя берутся «параллельно» главной диагонали, а третий по отношению к ним является угловым. Четвертое слагаемое есть произведение элементов побочной диагонали; в пятом и шестом слагаемых два сомножителя берутся «параллельно» побочной диагонали, а третий по отношению к ним является угловым. Это правило обычно иллюстрируют схемой, показанной на рис. 2.1.

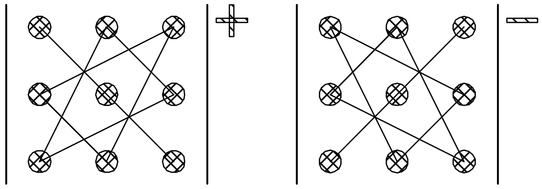

Рис. 2.1. Правило Саррюса

Для приведенного выше примера:

$$\begin{vmatrix} 1 & 3 & -2 \\ 2 & -1 & 4 \\ 1 & -3 & -2 \end{vmatrix} = 1\cdot(-1)\cdot(-2)+3\cdot 4\cdot 1+2\cdot(-3)\cdot(-2)-$$

$$-(-2)\cdot(-1)\cdot 1-3\cdot 2\cdot(-2)-4\cdot(-3)\cdot 1=2+12+12-2+12+12=48.$$

Несложно заметить, что ранее (при выполнении разложения по первой строке) для вычисления детерминанта потребовалось 14 арифметических действий, в то время как использование формулы полного разложения требует уже 17 действий.

Для вычисления детерминанта n-го порядка (разложением по какому-либо ряду) требуется вычислить n детерминантов $(n-1)$-го порядка, умножить их на элементы соответствующего ряда и сложить полученные произведения. Поэтому в общем случае применение соотношений (2.5) или (2.6) для вычисления детерминанта порядка n требует выполнения $K_n = n(K_{n-1}+2)-1$ арифметических операций, где K_{n-1} – число операций, необходимое для вычисления детерминанта $(n-1)$-го порядка. Число операций возрастает очень быстро (табл. 2.1).

Таблица 2.1.

Число арифметических операций, требуемых для вычисления детерминанта при помощи (2.5) и (2.6)

Порядок	3	4	10	50	150
Число операций	14	63	9864099	$8{,}3\cdot 10^{64}$	$1{,}6\cdot 10^{263}$

Формула полного разложения детерминанта n-го порядка содержит

$$n! = 1\cdot 2\cdot\ldots\cdot(n-1)\cdot n$$

слагаемых. Вычисление каждого из них требует выполнения $(n-1)$-го умножения. Поэтому при использовании формул полного разложения общее количество арифметических операций оказывается равным $K_n = n!n-1$ (табл. 2.2).

Таблица 2.2.

Число операций в формулах полного разложения

Порядок	3	4	10	50	150
Число операций	17	95	36287999	$1{,}5\cdot 10^{66}$	$8{,}6\cdot 10^{264}$

Из табл. 2.1 и 2.2 со всей очевидностью следует практическая нецелесообразность применения соотношений (2.5), (2.6) и формул полного разложения[1]. Все пригодные для практики методы вычисления детерминантов основаны на линейности по строке или столбцу.

Предложение 2.9.1. *Детерминант треугольной матрицы равен произведению ее диагональных элементов*.

Доказательство достаточно провести по индукции для верхней треугольной матрицы.

Детерминант треугольной матрицы второго порядка равен

$$\begin{vmatrix} a & b \\ 0 & c \end{vmatrix} = ac - 0\cdot b = ac\,;$$

база выполнена.

[1] Существует и другая причина, по которой эти соотношения применять *ни в коем случае нельзя*. Дело в том, что при их использовании погрешности исходных данных и ошибки округления машинных операций очень быстро приводят к порче результата.

Предположим, что утверждение справедливо для детерминанта $(n-1)$-го порядка и найдем детерминант

$$\det \mathbf{A} = \begin{vmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{22} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \dots \\ 0 & 0 & 0 & \dots & a_{nn} \end{vmatrix}.$$

Раскладывая его по первому столбцу, получим:

$$\det \mathbf{A} = \begin{vmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{22} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \dots \\ 0 & 0 & 0 & \dots & a_{nn} \end{vmatrix} = a_{11} M_{11} = a_{11} \begin{vmatrix} a_{22} & a_{23} & \dots & a_{2n} \\ 0 & a_{33} & \dots & a_{3n} \\ \dots \\ 0 & 0 & \dots & a_{nn} \end{vmatrix}.$$

Но дополнительный минор M_{11} является детерминантом верхней треугольной матрицы $(n-1)$-го порядка и по индуктивному предположению равен произведению диагональных элементов; следовательно

$$\det \mathbf{A} = a_{11} a_{22} \dots a_{nn}.$$

Совместное применение последнего предложения и линейных свойств приводит к методам из группы *методов Гаусса*.

Пусть, например, требуется вычислить детерминант:

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

Прибавим ко второй строке первую, умноженную на -2; прибавим к третьей строке первую, умноженную на -1. Получим:

$$\begin{vmatrix} 1 & 3 & -2 \\ 2 & -1 & 4 \\ 1 & -3 & -2 \end{vmatrix} = \begin{vmatrix} 1 & 3 & -2 \\ 0 & -7 & 8 \\ 0 & -6 & 0 \end{vmatrix}.$$

Для приведения детерминанта к верхнему треугольному виду необходимо исключить элемент a_{32}. Для этого можно было бы заменить третью строку суммой ее со второй строкой, умноженной на

$-\frac{6}{7}$. Однако это приводит к появлению дробных значений, что затрудняет ручные вычисления. Поэтому лучше переставить второй и третий столбцы:

$$\begin{vmatrix} 1 & 3 & -2 \\ 0 & -7 & 8 \\ 0 & -6 & 0 \end{vmatrix} = -\begin{vmatrix} 1 & -2 & 3 \\ 0 & 8 & -7 \\ 0 & 0 & -6 \end{vmatrix} = -(-48) = 48.$$

Число арифметических операций, необходимых для вычисления детерминанта методом Гаусса, при увеличении порядка возрастает по степенному закону (пропорционально третьей степени порядка). Поэтому в отличие от методов, основанных на разложениях (2.5) или (2.6), метод Гаусса сохраняет практическую пригодность вплоть до порядков $n \sim 1000$.

Контрольные вопросы и упражнения

1. Используя правило Саррюса, вычислите

$$\begin{vmatrix} -3 & 3 & 2 \\ -2 & 2 & 1 \\ 1 & 1 & -1 \end{vmatrix}; \quad \begin{vmatrix} 2 & -2 & -3 \\ -2 & -1 & -3 \\ -2 & -2 & 1 \end{vmatrix}; \quad \begin{vmatrix} -1 & 2 & -3 \\ -3 & -2 & -3 \\ 3 & -1 & 1 \end{vmatrix}.$$

2. Докажите

$$\begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix} = (a-b)(b-c)(c-a)$$

3 [10]. Вычислите

$$\begin{vmatrix} \sin\alpha & \cos\alpha \\ -\cos\alpha & \sin\alpha \end{vmatrix}; \quad \begin{vmatrix} \mathrm{tg}\,\alpha & -1 \\ 1 & \mathrm{tg}\,\alpha \end{vmatrix}; \quad \begin{vmatrix} \cos\varphi + i\sin\varphi & 1 \\ 1 & \cos\varphi - i\sin\varphi \end{vmatrix}.$$

4. Раскладывая детерминант по строке или столбцу, вычислите

$$\begin{vmatrix} 1 & 1 & -2 & -3 \\ 2 & -2 & 1 & -2 \\ 2 & -3 & -2 & -3 \\ 2 & 1 & -1 & 1 \end{vmatrix}; \quad \begin{vmatrix} 2 & -1 & 3 & 1 \\ 3 & -2 & -1 & -2 \\ 3 & 2 & -3 & -3 \\ 2 & 3 & -1 & -2 \end{vmatrix}.$$

5. Используя метод Гаусса, вычислите определители упражнений 1 и 4.

6 [9]. Решите уравнения

$$\begin{vmatrix} 3 & x & -x \\ 2 & -1 & 3 \\ x+10 & 1 & 1 \end{vmatrix} = 0; \quad \begin{vmatrix} x & x+1 & x+2 \\ x+3 & x+4 & x+5 \\ x+6 & x+7 & x+8 \end{vmatrix} = 0.$$

7*. Докажите, что определитель кососимметрической матрицы нечетного порядка равен нулю.

Глава 3. Системы линейных алгебраических уравнений

Одной из важнейших вычислительных задач линейной алгебры, наряду с вычислением определителей и обращением матриц, является решение *систем линейных алгебраических уравнений*.

Все методы решения таких систем принято разделять на две основные группы.

Прямые методы позволяют получить точное решение после конечного числа арифметических операций. Эти методы применяются в том случае, если порядок системы невелик (не более 10^3), или же если матрица коэффициентов имеет специальный вид (например, является ленточной, блочной или симметрической).

Итерационные методы позволяют после конечного числа операций получить лишь приближенное решение (для получения точного решения требуется выполнить бесконечное число операций). Итерационные методы обычно применяют для решения систем высокого порядка ($n \sim 10^3 ... 10^8$) или же для уточнения решения, полученного прямыми методами.

Далее рассматриваются только прямые методы.

3.1. Теорема Кронекера–Капелли

Неоднородной системой m линейных алгебраических уравнений с n неизвестными называется система

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n = b_2 \\ ... \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n = b_m \end{cases}, \tag{3.1}$$

или, в сокращенной записи:

$$\sum_{j=1}^{n} a_{ij}x_j = b_i, \ \ i = \overline{1,m}. \tag{3.2}$$

Здесь a_{ij} – *коэффициенты* при j-м неизвестном в i-м уравнении, b_i – *свободный член* в i-м уравнении.

Вектор-столбцы

$$\mathbf{x} = (x_1, x_2, ..., x_n)^T, \ \ \mathbf{b} = (b_1, b_2, ..., b_m)^T$$

носят название *столбца неизвестных* и *столбца свободных членов*.

Матрицей коэффициентов, или *основной матрицей системы* называется матрица

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

Из определения операции умножения матриц следует, что система (3.1) может быть записана в виде матричного уравнения

$$\mathbf{Ax} = \mathbf{b}.$$

Решением системы (3.1) называется столбец $\mathbf{x}$, обращающий все уравнения системы в тождества. Если столбцы матрицы коэффициентов обозначить как $\mathbf{a}_j$, то систему (3.2) можно записать в виде

$$\sum_{j=1}^{n} \mathbf{a}_j x_j = \mathbf{b}; \tag{3.3}$$

следовательно, *решение системы – это совокупность коэффициентов в разложении столбца свободных членов по столбцам матрицы коэффициентов*.

Система называется *совместной*, если она имеет хотя бы одно решение, и *несовместной*, если она не имеет решений. Например, система

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

является совместной: нетрудно проверить, что вектор-столбец $\mathbf{x} = (1,2)^T$ обращает оба ее уравнения в верные равенства. Система

$$\begin{cases} x_1 + x_2 = 0 \\ x_1 + x_2 = 1 \end{cases}$$

является несовместной, так как нельзя найти такой пары чисел, сумма которых была бы равна нулю и единице одновременно.

Совместная система называется *определенной*, если она имеет единственное решение, и *неопределенной*, если она имеет более одного решения.

Примером неопределенной системы является

$$\begin{cases} x_1 + x_2 - 2x_3 = 2 \\ x_1 - x_2 = 0 \end{cases}.$$

Нетрудно проверить, что любой вектор-столбец

$$\mathbf{x} = (\alpha + 1, \alpha + 1, \alpha)^T,$$

где α – произвольное число, является решением системы.

Вопрос определенности системы решает предложение:

Предложение 3.1.1. *Если столбцы матрицы коэффициентов линейно независимы, то система не может быть неопределенной.*

Действительно, пусть $\mathbf{x}' = (x'_j)$ и $\mathbf{x}'' = (x''_j)$ – решения системы:

$$\sum_{j=1}^{n} \mathbf{a}_j x'_j = \mathbf{b}, \ \sum_{j=1}^{n} \mathbf{a}_j x''_j = \mathbf{b}.$$

Вычитая второе разложение из первого, получим:

$$\sum_{j=1}^{n} \mathbf{a}_j (x'_j - x''_j) = \mathbf{0},$$

откуда в силу линейной независимости столбцов $\mathbf{a}_j$ следует

$$x'_j - x''_j = 0, \ j = \overline{1,n},$$

или $\mathbf{x}' = \mathbf{x}''$. Предложение доказано.

Расширенной матрицей системы (3.1) называется матрица коэффициентов, дополненная столбцом свободных членов:

$$\mathbf{A}^* = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & & & & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right).$$

Вопрос о совместности системы (3.1) решает

Теорема 3.1.1 (*теорема Кронекера–Капелли*). *Система линейных уравнений совместна тогда и только тогда, когда ранг основной матрицы системы равен рангу ее расширенной матрицы.*

Докажем необходимость. Пусть система совместна и столбец $\mathbf{x} = (x_j)$ является ее решением.

Тогда последний столбец расширенной матрицы является линейной комбинацией ее первых n столбцов:

$$\mathbf{b} = \sum_{j=1}^{n} \mathbf{a}_j x_j .$$

Следовательно, элементарными преобразованиями столбцов расширенной матрицы можно обратить ее последний столбец в столбец из нулей. Но ни элементарные преобразования столбцов, ни приписывание нулевого столбца не меняют ранга матрицы; поэтому $\operatorname{rang}\mathbf{A} = \operatorname{rang}\mathbf{A}^*$.

Докажем достаточность. Пусть $\operatorname{rang}\mathbf{A} = \operatorname{rang}\mathbf{A}^* = r$; тогда базисная подматрица матрицы коэффициентов $\mathbf{A}$ одновременно является базисной и для расширенной матрицы $\mathbf{A}^*$. Следовательно, последний столбец расширенной матрицы может быть разложен по r столбцам $\mathbf{a}_{j_1}, \mathbf{a}_{j_2}, ..., \mathbf{a}_{j_r}$ базисной подматрицы:

$$\mathbf{b} = \sum_{k=1}^{r} x_{j_k} \mathbf{a}_{j_k} .$$

Но тогда его можно представить в виде линейной комбинации всех столбцов матрицы $\mathbf{A}$, добавив недостающие столбцы с нулевыми коэффициентами. Теорема доказана.

Контрольные вопросы и упражнения

1. Что называют системой линейных алгебраических уравнений? Что называют основной матрицей системы? Что называют столбцом свободных членов? Столбцом неизвестных?

2. Что называют решением системы? Какая система называется совместной? Какая система называется определенной?

3. Что называют расширенной матрицей системы?

4. Сформулируйте условие теоремы Кронекера–Капелли.

5. Используя теорему Кронекера–Капелли и предложение 3.3.1, решите вопрос о совместности и определенности системы

$$\begin{cases} 2x + y + z = 7 \\ x - y + 2z = 5 \end{cases}.$$

3.2. Метод Гаусса (общий случай)

Пусть дана матрица

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

Главными минорами матрицы **A** называются детерминанты

$$A_1 = |a_{11}|;\ A_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix};\ \dots;\ A_k = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \dots & & & \\ a_{k1} & a_{k2} & \dots & a_{kk} \end{vmatrix},$$

где k − наименьшее среди m и n.

Пусть имеется система (3.1) из m линейных алгебраических уравнений с n неизвестными, причем ранг матрицы коэффициентов совпадает с рангом расширенной матрицы:

$$\operatorname{rang} \mathbf{A} = \operatorname{rang} \mathbf{A}^* = r.$$

Без ограничения общности будем считать, что первые r главных миноров матрицы коэффициентов отличны от нуля.

Умножение любого уравнения на отличное от нуля число и замена любого уравнения суммой его с другим уравнением не меняют решения системы. Поэтому *решен, еяс, стемыянеяменяетсяяпр, яэлементарныэяпреобразован, яэястрокярасш, реннойяматр, цы.*

Для нахождения решения расширенной матрице следует элементарными преобразованиями строк придать приведенно-ступенчатую форму:

$$\left(\begin{array}{cccccccc|c} 1 & 0 & \dots & 0 & \alpha_{1\,r+1} & \alpha_{1\,r+2} & \dots & \alpha_{1\,n} & \beta_1 \\ 0 & 1 & \dots & 0 & \alpha_{2\,r+1} & \alpha_{2\,r+2} & \dots & \alpha_{2\,n} & \beta_2 \\ \dots & & & & & & & & \\ 0 & 0 & \dots & 1 & \alpha_{r\,r+1} & \alpha_{r\,r+2} & \dots & \alpha_{r\,n} & \beta_r \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ \dots & & & & & & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \end{array}\right).$$

Последней матрице соответствует система из r уравнений:

$$x_i + \sum_{j=r+1}^{n} \alpha_{ij} x_j = \beta_i, \ i = \overline{1,r},$$

или

$$x_i = \beta_i - \sum_{j=r+1}^{n} \alpha_{ij} x_j, \ i = \overline{1,r}. \tag{3.4}$$

Первые r неизвестных, которые соответствуют r базисным столбцам матрицы коэффициентов и оставлены в левых частях уравнений системы (3.4), называют *базисными*. Остальные $n-r$ неизвестных называют *свободными*. Значения свободных неизвестных можно выбирать произвольно; при этом по формулам (3.4) значения базисных неизвестных определяются так, что вместе со значениями свободных неизвестных образуют решение системы. Следовательно, формулы (3.4) являются описанием всего множества решений.

Приведение расширенной матрицы к ступенчатой форме осуществляется так же, как и при доказательстве теоремы 2.6.3 – методом Гаусса. Поэтому и сам метод решения системы линейных уравнений называют *методом Гаусса*.

Пусть, например, требуется решить систему

$$\begin{cases} x_1 + x_2 - 2x_3 = 2 \\ x_1 - x_2 = 0 \\ x_1 - x_3 = 1 \end{cases}.$$

Совместность системы (равенство рангов матрицы коэффициентов и расширенной матрицы) может быть установлена в процессе приведения расширенной матрицы к ступенчатой форме:

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

Подматрица, расположенная в первых двух строках и первых двух столбцах, является базисной как для матрицы коэффициентов, так и для расширенной матрицы преобразованной системы. Следовательно, ранги этих матриц совпадают и условия теоремы Кронекера–Капелли (для исходной и преобразованной систем) выполнены. Система является совместной, ее решение дается формулами (3.4) и в данном примере имеет вид:

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

Контрольные вопросы и упражнения

1. Пусть ранг матрицы коэффициентов и расширенной матрицы системы равен r, но среди первых r главных миноров есть нулевые. Можно ли в этом случае элементарными преобразованиями строк привести расширенную матрицу к ступенчатому виду?

2. Найдите решения или докажите несовместность систем:

$$\begin{cases} x - y + z = -4 \\ x + y + z = -2 \\ x - 2y - 2z = -2 \end{cases}; \quad \begin{cases} 2x - y + z = 5 \\ -x + y + z = -2 \\ x - 2y + z = 6 \end{cases}; \quad \begin{cases} x + 2y + z = 6 \\ x + y - 2z = -1 \\ 2x + 3y - z = 5 \end{cases}; \quad \begin{cases} x + 2y - z = 4 \\ x - 3y + 2z = 1 \\ 2x - y + z = 0 \end{cases}.$$

Выполните проверку полученных результатов.

3.3. Однородные системы

Система линейных алгебраических уравнений называется *однородной*, если в ней все свободные члены равны нулю:

$$\sum_{j=1}^{n} a_{ij} x_j = 0, \quad i = \overline{1,m}. \tag{3.5}$$

Так как приписывание нулевого столбца к матрице коэффициентов не меняет ее ранга, то для однородной системы условия теоремы Кронекера–Капелли выполнены всегда. Поэтому однородная система всегда совместна – одним из ее решений является нулевой вектор. Это решение называют *тривиальным*; все остальные (ненулевые) решения называют *нетривиальными*.

Важнейшее свойство однородной системы устанавливает предложение:

Предложение 3.3.1. *Линейная комбинация решений однородной системы также является ее решением.*

Доказательство. Пусть $\mathbf{x}_1$, $\mathbf{x}_2$, …, $\mathbf{x}_k$ – решения системы (3.5):

$$\mathbf{A}\mathbf{x}_s = \mathbf{0}, \quad s = \overline{1,k}.$$

Тогда

$$\mathbf{A}\sum_{s=1}^{k}\alpha_s\mathbf{x}_s=\sum_{s=1}^{k}\alpha_s\mathbf{A}\mathbf{x}_s=\sum_{s=1}^{k}\alpha_s\mathbf{0}=\mathbf{0}.$$

Предложение доказано.

Если столбцы $\mathbf{a}_j$ матрицы коэффициентов системы (3.5) линейно независимы, то из предложения 3.1.1 сразу следует, что однородная система не имеет нетривиальных решений. Действительно, в этом случае равенство

$$\sum_{j=1}^{n}\mathbf{a}_j x_j=\mathbf{0}$$

возможно только с нулевыми коэффициентами x_j.

Предложение 3.3.1 позволяет сделать вывод о *замкнутости* множества решений однородной системы (3.5) относительно операций сложения и умножения на число. Каждое решение системы (3.5) является набором из n чисел, поэтому для множества таких наборов выполнены все аксиомы линейного пространства. Следовательно, множество решений однородной системы является линейным пространством. Базис этого пространства называют *фундаментальным набором решений*; любое другое решение является линейной комбинацией решений фундаментального набора. Матрицу $\mathbf{F}$, в столбцах которой находятся решения из фундаментального набора, называют *фундаментальной матрицей системы*.

Для нахождения фундаментального набора можно по формулам (3.4) выразить r базисных неизвестных через $n-r$ свободных:

$$x_j=\sum_{j=r+1}^{n}\alpha_{ij}x_j,\ \ i=\overline{1,r}.$$

Затем вектор-столбец свободных неизвестных последовательно полагается равным векторам естественного базиса $(n-r)$-мерного координатного пространства.

Например, пусть требуется найти фундаментальный набор решений системы

$$\begin{cases} x_1+x_2-2x_3+2x_4=0\\ x_1-x_2-2x_4=0\\ x_1-x_3=0 \end{cases}.$$

Приводя матрицу коэффициентов к ступенчатому виду, будем иметь:

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

Последней матрице соответствует система из двух уравнений с двумя свободными неизвестными:

$$\begin{cases} x_1 = x_3 \\ x_2 = x_3 - 2x_4 \end{cases}.$$

Полагая вектор-столбец $(x_3, x_4)^T$ последовательно равным векторам $(1,0)^T$ и $(0,1)^T$, получим два вектора, составляющих базис пространства решений:

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \ \mathbf{e}_2 = \begin{pmatrix} 0 \\ -2 \\ 0 \\ 1 \end{pmatrix}.$$

Любое другое решение системы является линейной комбинацией векторов фундаментального набора:

$$\mathbf{x} = \alpha_1 \mathbf{e}_1 + \alpha_2 \mathbf{e}_2 = \begin{pmatrix} \alpha_1 \\ \alpha_1 - 2\alpha_2 \\ \alpha_1 \\ \alpha_2 \end{pmatrix}.$$

Записав фундаментальную матрицу системы

$$\mathbf{F} = \begin{pmatrix} 1 & 0 \\ 1 & -2 \\ 1 & 0 \\ 0 & 1 \end{pmatrix},$$

произвольное – *общее* – решение системы можно представить в виде

$$\mathbf{x} = \mathbf{F}\mathbf{c},$$

где $\mathbf{c} = (\alpha_1, \alpha_2)^T$ – произвольный вектор.

Пусть дана неоднородная система (3.1) из m уравнений с n неизвестными:

$$\mathbf{Ax} = \mathbf{b}.$$

Тогда однородная система (3.5) с той же матрицей коэффициентов, что и (3.1):

$$\mathbf{Ax} = \mathbf{0}$$

называется *приведенной* по отношению к (3.1). Связь между решениями исходной (неоднородной) и приведенной систем устанавливают следующие два предложения.

Предложение 3.3.2. *Сумма любого решения* $\mathbf{x}$ *неоднородной системы с любым решением* $\mathbf{x}_0$ *приведенной системы является решением неоднородной системы.*

Действительно, так как $\mathbf{Ax} = \mathbf{b}$ и $\mathbf{Ax}_0 = \mathbf{0}$, то

$$\mathbf{A}(\mathbf{x} + \mathbf{x}_0) = \mathbf{Ax} + \mathbf{Ax}_0 = \mathbf{b} + \mathbf{0} = \mathbf{b}.$$

Предложение 3.3.3. *Разность двух любых решений* $\mathbf{x}_1$ *и* $\mathbf{x}_2$ *неоднородной системы является решением приведенной системы.*

Действительно,

$$\mathbf{A}(\mathbf{x}_1 - \mathbf{x}_2) = \mathbf{Ax}_1 - \mathbf{Ax}_2 = \mathbf{b} - \mathbf{b} = \mathbf{0}.$$

Пусть $\mathbf{F}$ – фундаментальная матрица приведенной системы. Из доказанных предложений следует, что найдя какое-либо *частное* решение $\mathbf{x}'$ неоднородной системы и складывая его с каждым решением приведенной системы, мы найдем все решения неоднородной системы

$$\mathbf{x} = \mathbf{x}' + \mathbf{Fc}, \tag{3.6}$$

где $\mathbf{c}$ – произвольный вектор-столбец, высота которого равна разности между числом неизвестных и рангом матрицы коэффициентов.

Соотношение (3.6) называют *общим решением неоднородной системы*; следует подчеркнуть, что оно отличается от (3.4) только по форме. Действительно, полагая

$$\mathbf{x}' = (\beta_1, \beta_2, \ldots, \beta_r, 0, \ldots, 0)^T,$$

$$\mathbf{c} = (x_{r+1}, x_{r+2}, \ldots, x_n)^T,$$

$$\mathbf{F} = \begin{pmatrix} -\alpha_{1\,r+1} & -\alpha_{1\,r+2} & \dots & -\alpha_{1\,n} \\ -\alpha_{2\,r+1} & -\alpha_{2\,r+2} & \dots & -\alpha_{2\,n} \\ \dots & & & \\ -\alpha_{r\,r+1} & -\alpha_{r\,r+2} & \dots & -\alpha_{r\,n} \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & & & \\ 0 & 0 & \dots & 1 \end{pmatrix},$$

вновь придем к соотношению (3.4).

Контрольные вопросы и упражнения

1. Какую систему называют однородной?
2. Что называют тривиальным решением однородной системы?
3. Что называют фундаментальным набором решений?
4. Найдите фундаментальные наборы решений и укажите размерность пространств решений однородных систем:

$$\begin{cases} x - y + z = 0 \\ x - 2y - z = 0 \\ x + y + z = 0 \end{cases}; \quad \begin{cases} x - y + z + 2t = 0 \\ x - 2y - z + t = 0 \\ x + y + z - t = 0 \end{cases}; \quad \begin{cases} x + 2y - z + t = 0 \\ x + 3y + z - t = 0 \\ 3x + 8y + z - t = 0 \end{cases}.$$

3.4. Разновидности метода Гаусса

Пусть требуется решить систему, в которой число уравнений n совпадает с числом неизвестных:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n \end{cases}, \tag{3.7}$$

Пусть ранг матрицы коэффициентов совпадает с числом неизвестных. Тогда ранг расширенной матрицы также совпадает с числом неизвестных (матрица коэффициентов для расширенной матрицы является базисной). Поэтому из предложения 3.1.1 и теоремы Кронекера–Капелли следует, что система (3.7) совместна и определенна. Как и ранее, ее решение можно получить по формулам (3.4):

$$x_i = \beta_i\,,\ \ i = \overline{1,n}\,,$$

где β_i – числа, находящиеся в последнем столбце расширенной матрицы после приведения ее к ступенчатому виду:

$$\left(\begin{array}{cccc|c} 1 & 0 & \dots & 0 & \beta_1 \\ 0 & 1 & \dots & 0 & \beta_2 \\ \dots & & & & \\ 0 & 0 & \dots & 1 & \beta_n \end{array}\right).$$

Таким образом, *для решения системы* (3.7) *левую часть ее расширенной матрицы* (*матрицу коэффициентов*) *следует элементарными преобразованиями строк привести к единичной матрице; тогда последний столбец расширенной матрицы будет являться решением системы*.

Элементарные преобразования строк соответствуют умножению матрицы справа на соответствующие элементарные матрицы. Пусть преобразования $\mathbf{T}_1$, $\mathbf{T}_2$, …, $\mathbf{T}_s$ переводят матрицу коэффициентов в единичную:

$$\mathbf{T}_s\mathbf{T}_{s-1}\dots\mathbf{T}_1\mathbf{A} = \mathbf{E}\,.$$

Тогда эта же последовательность преобразований переведет столбец свободных членов в решение системы:

$$\mathbf{T}_s\mathbf{T}_{s-1}\dots\mathbf{T}_1\mathbf{b} = \mathbf{x}\,.$$

Но из изложенного в п. 2.7 следует

$$\mathbf{T}_s\mathbf{T}_{s-1}\dots\mathbf{T}_1 = \mathbf{A}^{-1},$$

поэтому решение можно записать в виде

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}\,. \tag{3.8}$$

Последнее соотношение можно получить иначе. Достаточно умножить обе части системы $\mathbf{Ax} = \mathbf{b}$ на матрицу, обратную к матрице коэффициентов:

$$\mathbf{A}^{-1}\mathbf{Ax} = \mathbf{A}^{-1}\mathbf{b}\,,\;\; \mathbf{Ex} = \mathbf{A}^{-1}\mathbf{b}\,,\;\; \mathbf{x} = \mathbf{A}^{-1}\mathbf{b}\,.$$

Формулой (3.8) решение системы явно выражается через матрицу коэффициентов и столбец свободных членов.

Пусть, например, требуется решить систему

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

Обратим матрицу коэффициентов:

$$\left(\begin{array}{cc|cc} 1 & -1 & 1 & 0 \\ 2 & 1 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{cc|cc} 1 & -1 & 1 & 0 \\ 0 & 3 & -2 & 1 \end{array}\right) \sim \left(\begin{array}{cc|cc} 1 & -1 & 1 & 0 \\ 0 & 1 & -2/3 & 1/3 \end{array}\right) \sim \left(\begin{array}{cc|cc} 1 & 0 & 1/3 & 1/3 \\ 0 & 1 & -2/3 & 1/3 \end{array}\right),$$

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

Решение системы:

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} = \begin{pmatrix} 1/3 & 1/3 \\ -2/3 & 1/3 \end{pmatrix}\begin{pmatrix} -1 \\ 4 \end{pmatrix} = \begin{pmatrix} -1/3 + 4/3 \\ 2/3 + 4/3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Обращение матрицы методом Гаусса соответствует одновременному решению n систем линейных уравнений, правые части которых являются столбцами единичной матрицы. Поэтому использование соотношения (3.8) для решения единственной системы неэффективно. Это соотношение могло бы оказаться полезным, если бы требовалось решить большое число систем с одинаковой матрицей коэффициентов (отличающихся только столбцом свободных членов). Однако и для этого случая разработаны методы, сравнимые с соотношением (3.8) по вычислительной эффективности и в меньшей степени подверженные влиянию ошибок округления и погрешностей исходных данных.

Выбор конкретной последовательности элементарных преобразований, превращающей матрицу коэффициентов в единичную, порождает соответствующую реализацию метода Гаусса.

Пусть дана система линейных уравнений (3.7), причем все главные миноры ее матрицы коэффициентов

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

отличны от нуля (выполнения этого требования всегда можно добиться перестановкой строк).

Разновидность метода Гаусса, называемая *схемой единственного деления*, состоит в преобразовании исходной системы в систему с верхней треугольной матрицей коэффициентов (*прямой ход*) и последовательном определении неизвестных, начиная с n-го (*обратная подстановка*).

Замечание 3.1. Наряду с эквивалентными преобразованиями строк расширенной матрицы возможно изменение порядка ее столбцов, однако такое изменение должно сопровождаться перенумерацией неизвестных.

Как и ранее, все преобразования удобно применять не к исходной системе, а к ее расширенной матрице.

Первый главный минор отличен от нуля: $A_1 = a_{11} \neq 0$. Преобразования *прямого хода* начинаются с деления первой строки расширенной матрицы на элемент a_{11} (*единственное деление*), за которым следует замена строк со второй по n-ю на разности их с первой строкой, умноженной на $-a_{i1}$ (здесь $i = \overline{2, n}$).

В результате расширенная матрица примет вид

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & & & & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} & b_n \end{array}\right) \sim \left(\begin{array}{cccc|c} 1 & a'_{12} & \dots & a'_{1n} & b'_1 \\ 0 & a'_{22} & \dots & a'_{2n} & b'_2 \\ \dots & & & & \dots \\ 0 & a'_{n2} & \dots & a'_{nn} & b'_n \end{array}\right);$$

говорят, что в матрице *исключены* элементы a_{21}, a_{31}, …, a_{n1}.

Так как второй главный минор исходной расширенной матрицы был отличен от нуля, то в преобразованной матрице минор

$$\begin{vmatrix} 1 & a'_{12} \\ 0 & a'_{22} \end{vmatrix} = 1 \cdot a'_{22} - 0 \cdot a'_{12} = a'_{22}$$

также отличен от нуля; следовательно, $a'_{22} \neq 0$. Второй шаг схемы единственного деления состоит в применении преобразований первого шага к подматрице $(n-1)$-го порядка, расположенной в столбцах и строках с номерами от 2 до n. Разделим вторую строку расширенной матрицы на a'_{22} и заменим строки с третьей по n-ю на разности их со второй строкой, умноженной на $-a'_{i2}$, $i = \overline{3, n}$. Получим:

$$\left(\begin{array}{cccc|c} 1 & a'_{12} & \dots & a'_{1n} & b'_1 \\ 0 & a'_{22} & \dots & a'_{2n} & b'_2 \\ \dots & & & & \dots \\ 0 & a'_{n2} & \dots & a'_{nn} & b'_n \end{array}\right) \sim \left(\begin{array}{ccccc|c} 1 & a'_{12} & a'_{13} & \dots & a'_{1n} & b'_1 \\ 0 & 1 & a''_{23} & \dots & a''_{2n} & b''_2 \\ 0 & 0 & a''_{33} & \dots & a''_{3n} & b''_3 \\ \dots & & & & & \\ 0 & 0 & a''_{n3} & \dots & a''_{nn} & b''_n \end{array}\right).$$

Выполнив в соответствии с данным алгоритмом $n-1$ шаг, преобразуем расширенную матрицу в верхнюю треугольную с единицами на главной диагонали.

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & & & & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} & b_n \end{array}\right) \sim \left(\begin{array}{ccccc|c} 1 & \alpha_{12} & \alpha_{13} & \dots & \alpha_{1n} & \beta_1 \\ 0 & 1 & \alpha_{23} & \dots & \alpha_{2n} & \beta_2 \\ 0 & 0 & 1 & \dots & \alpha_{3n} & \beta_3 \\ \dots & & & & & \\ 0 & 0 & 0 & \dots & 1 & \beta_n \end{array}\right) \qquad (3.9)$$

Соответствующая этой матрице система уравнений

$$x_i + \sum_{j=i+1}^{n} \alpha_{ij} x_j = \beta_i \qquad (3.10)$$

допускает последовательное определение неизвестных, начиная с последнего (*обратная подстановка*).

Положим $i = n$, тогда (3.10) примет вид

$$x_n = \beta_n.$$

Полагая затем $i = n-1, n-2, ..., 1$, получим

$$x_i = \beta_i - \sum_{j=i+1}^{n} \alpha_{ij} x_j .$$

Вместо выполнения обратной подстановки можно элементарными преобразованиями строк довести матрицу (3.9) до единичной (*обратный ход* метода Гаусса). В этой форме схема единственного деления переходит в схему *Жордана–Гаусса*, которая уже была использована выше для решения систем линейных уравнений и обращения матриц.

Пусть, например, требуется решить систему линейных уравнений (переменные обозначены буквами x, y и z):

$$\begin{cases} x + y + 2z = -4 \\ 3x - 2y - z = -3 \\ x - 2y - 2z = 1 \end{cases}. \qquad \text{(a)}$$

Приведем расширенную матрицу к верхнему треугольному виду:

$$\left(\begin{array}{ccc|c} 1 & 1 & 2 & -4 \\ 3 & -2 & -1 & -3 \\ 1 & -2 & -2 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 1 & 2 & -4 \\ 0 & -5 & -7 & 9 \\ 0 & -3 & -4 & 5 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 1 & 2 & -4 \\ 0 & 15 & 21 & -27 \\ 0 & -15 & -20 & 25 \end{array}\right) \sim$$

$$\sim \left(\begin{array}{ccc|c} 1 & 1 & 2 & -4 \\ 0 & 5 & 7 & -9 \\ 0 & 0 & 1 & -2 \end{array}\right).$$

Здесь еще не все диагональные элементы равны единице, однако для нахождения решения этого и не требуется. Полученная матрица соответствует системе

$$\begin{cases} x + y + 2z = -4 \\ 5y + 7z = -9 \\ z = -2 \end{cases}.$$

Из третьего уравнения $z = -2$. Подставляя во второе уравнение, найдем:

$$5y - 14 = -9, \; 5y = 5, \; y = 1.$$

Подставляя найденные y и z в первое уравнение, получим:

$$x+1-4=-4\,,\ \ x=-1\,.$$

Вместо выполнения обратной подстановки можно выполнить обратный ход (фактически, выполняются те же самые действия; различие состоит только в форме записи):

$$\left(\begin{array}{ccc|c}1&1&2&-4\\0&5&7&-9\\0&0&1&-2\end{array}\right)\sim\left(\begin{array}{ccc|c}1&1&2&-4\\0&1&7/5&-9/5\\0&0&1&-2\end{array}\right)\sim\left(\begin{array}{ccc|c}1&1&0&0\\0&1&0&1\\0&0&1&-2\end{array}\right)\sim\left(\begin{array}{ccc|c}1&0&0&-1\\0&1&0&1\\0&0&1&-2\end{array}\right);$$

как и ранее, решением является столбец $(-1,1,-2)^T$.

При использовании *метода оптимального исключения* поочередно выполняются операции прямого и обратного хода метода Гаусса. Решение системы также начинается с деления первой строки расширенной матрицы на элемент a_{11}; полученная строка используется для исключения первого элемента a_{21} второй строки. Затем новая вторая строка делится на элемент a'_{22} и используется для исключения второго элемента a'_{12} первой строки. На втором шаге при помощи двух первых строк исключают первые два элемента третьей строки; затем при помощи третьей строки исключают третий элемент в первой и второй строках. После выполнения k шагов данного алгоритма первые k строк и k столбцов расширенной матрицы являются строками и столбцами единичной матрицы k-го порядка.

Пусть, например, требуется решить систему (а). Имеем:

$$\left(\begin{array}{ccc|c}1&1&2&-4\\3&-2&-1&-3\\1&-2&-2&1\end{array}\right)\sim\left(\begin{array}{ccc|c}1&1&2&-4\\0&-5&-7&9\\1&-2&-2&1\end{array}\right)\sim\left(\begin{array}{ccc|c}1&1&2&-4\\0&1&7/5&-9/5\\1&-2&-2&1\end{array}\right)\sim$$

$$\sim\left(\begin{array}{ccc|c}1&0&3/5&-11/5\\0&1&7/5&-9/5\\1&-2&-2&1\end{array}\right)\sim\left(\begin{array}{ccc|c}1&0&3/5&-11/5\\0&1&7/5&-9/5\\0&0&1/5&-2/5\end{array}\right)\sim$$

$$\sim\left(\begin{array}{ccc|c}1&0&3/5&-11/5\\0&1&7/5&-9/5\\0&0&1&-2\end{array}\right)\sim\left(\begin{array}{ccc|c}1&0&0&-1\\0&1&0&1\\0&0&1&-2\end{array}\right);$$

решением является столбец $(-1,1,-2)^T$.

Контрольные вопросы и упражнения

1. Используя схему единственного деления, решите системы:

$$\begin{cases} x+2y-2z=-3 \\ 2x-y+2z=5 \\ x+2y+2z=1 \end{cases}; \quad \begin{cases} 2x-y-z=4 \\ 3x+4y-2z=11 \\ 3x-2y+4z=11 \end{cases}; \quad \begin{cases} 2x-2y-z-t=-3 \\ -x+y-z-t=-3 \\ -x+2y-z+t=2 \\ x-2y+z+t=2 \end{cases}.$$

Выполните проверку полученных результатов.

2. Решите системы первого упражнения, используя метод оптимального исключения.

3. Найдите решения первых двух систем первого упражнения, обращая матрицу коэффициентов (решения находятся как произведения матриц, обратных к матрицам коэффициентов, на столбец свободных членов).

3.5. LU-разложение

Прежде всего вновь подчеркнем сходство операций обращения матрицы и решения системы линейных уравнений. При обращении матрица справа дополняется единичной, а при решении системы уравнений – столбцом свободных членов, однако характер последующих преобразований определяется *только самой матрицей* и в том и другом случаях один и тот же.

Естественно возникает вопрос – можно ли представить матрицу коэффициентов в такой форме, которая позволяла бы получить решение наиболее просто и / или с минимальной погрешностью (при фиксированных погрешностях округления и погрешностях исходных данных). Оказывается, что таких представлений (*разложений*) матрицы коэффициентов достаточно много. Здесь мы остановимся лишь на том из них, которое в некотором роде является «матричной записью» алгоритма единственного деления.

Как указано ранее, целью прямого хода метода единственного деления является приведение (элементарными преобразованиями строк) матрицы коэффициентов к верхней треугольной матрице (как и ранее, мы требуем, чтобы ее диагональные элементы были равны единице):

$$\mathbf{U} = \begin{pmatrix} 1 & u_{12} & u_{13} & \dots & u_{1n} \\ 0 & 1 & u_{23} & \dots & u_{2n} \\ 0 & 0 & 1 & \dots & u_{3n} \\ \dots & & & & \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}.$$

Каждое из элементарных преобразований прямого хода соответствует умножению матрицы коэффициентов на соответствующую элементарную матрицу:

$$\mathbf{T}_k \mathbf{T}_{k-1} \dots \mathbf{T}_1 \mathbf{A} = \mathbf{U}. \tag{b}$$

Важно, что при этом все матрицы $\mathbf{T}_1$, $\mathbf{T}_2$, …, $\mathbf{T}_k$ сами являются *нижними треугольными* (в частности, диагональными). Действительно, первый шаг прямого хода начинается с деления первой строки на элемент a_{11}. Этому преобразованию соответствует умножение на диагональную матрицу, полученную из единичной заменой первой единицы на диагонали числом, обратным к a_{11}:

$$\mathbf{T}_1 = \begin{pmatrix} 1/a_{11} & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & & & \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

После этого новая первая строка используется для исключения элемента a_{21}, находящегося во второй строке и первом столбце. Этому преобразованию соответствует умножение на матрицу, полученную из единичной заменой нуля во второй строке и первом столбце элементом, противоположным к a_{21}. Данная матрица является произведением трех элементарных матриц:

$$\mathbf{T}_4 \mathbf{T}_3 \mathbf{T}_2 = \begin{pmatrix} -1/a_{21} & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & & & \\ 0 & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & \dots & 0 \\ 1 & 1 & \dots & 0 \\ \dots & & & \\ 0 & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} -a_{21} & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & & & \\ 0 & 0 & \dots & 1 \end{pmatrix} =$$

$$= \begin{pmatrix} 1 & 0 & \dots & 0 \\ -a_{21} & 1 & \dots & 0 \\ \dots & & & \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

Нетрудно проверить, что все остальные матрицы $\mathbf{T}_5$, $\mathbf{T}_6$, …, $\mathbf{T}_k$ также будут нижними треугольными. Но произведение нижних треугольных матриц само является нижней треугольной матрицей. Обозначая эту матрицу $\mathbf{L}^{-1}$, запишем (b) в виде

$$\mathbf{L}^{-1}\mathbf{A} = \mathbf{U}. \tag{c}$$

Матрица $\mathbf{L}$, обратная к $\mathbf{L}^{-1}$, также нижняя треугольная. Умножив обе части (c) на матрицу $\mathbf{L}$, получим:

$$\mathbf{A} = \mathbf{LU}. \tag{3.11}$$

Соотношение (3.11) называется *LU-разложением*[1] матрицы $\mathbf{A}$. Если данное разложение найдено, то решение системы линейных уравнений сводится к последовательному выполнению прямой и обратной подстановок (и требует порядка n^2 арифметических операций; для прямого хода схемы единственного деления это число имеет порядок n^3).

Пусть, например, необходимо решить систему

$$\begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 5 \\ 4 \end{pmatrix}.$$

Введем промежуточные неизвестные z и t так, чтобы

$$\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} z \\ t \end{pmatrix}.$$

Для нахождения z и t получим систему

$$\begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} z \\ t \end{pmatrix} = \begin{pmatrix} 5 \\ 4 \end{pmatrix}.$$

В развернутой форме:

$$\begin{cases} z = 5 \\ 2z + 3t = 4 \end{cases},$$

откуда прямой подстановкой найдем $z = 5$, $10 + 3t = 4$, $3t = -6$, $t = -2$.

[1] От англ. *lower* – нижний, *upper* – верхний.

Возвращаясь к переменным x и y, будем иметь

$$\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 5 \\ -2 \end{pmatrix},$$

$$\begin{cases} x - y = 5 \\ y = -2 \end{cases},$$

откуда обратной подстановкой получим $y = -2$, $x + 2 = 5$, $x = 3$.

Простейший способ нахождения LU-разложения состоит в умножении обеих частей (3.11) на матрицу $\mathbf{U}^{-1}$:

$$\mathbf{AU}^{-1} = \mathbf{LUU}^{-1},$$

$$\mathbf{L} = \mathbf{AU}^{-1}.$$

Так как матрица $\mathbf{U}$ – верхняя треугольная, то ее обращение выполняется сравнительно просто.

Пусть, например, требуется разложить матрицу

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

Приведем ее к верхней треугольной:

$$\begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 \\ 0 & 3 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}; \ \mathbf{U} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

Обращая матрицу $\mathbf{U}$, получим:

$$\left(\begin{array}{cc|cc} 1 & -1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{cc|cc} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{array}\right); \ \mathbf{U}^{-1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Тогда

$$\mathbf{L} = \mathbf{AU}^{-1} = \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix}.$$

Разложение имеет вид

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

На практике применяются другие, существенно более эффективные способы разложения.

Контрольные вопросы и упражнения

1. Что называют LU-разложением матрицы $\mathbf{A}$?
2. Найдите LU-разложения матриц

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

Выполните проверку полученных результатов.

3. Используя результаты предыдущего упражнения, решите системы (упражнение 1 п. 3.4):

$$\begin{cases} x+2y-2z=-3 \\ 2x-y+2z=5 \\ x+2y+2z=1 \end{cases}; \begin{cases} 2x-y-z=4 \\ 3x+4y-2z=11 \\ 3x-2y+4z=11 \end{cases}.$$

3.6. Формулы Крамера

Используя теорию определителей, можно явно выразить решение системы (3.7) через элементы матрицы коэффициентов и свободные члены.

Для нахождения решения $\mathbf{x}$ умножим i-е уравнение системы (3.7) на алгебраическое дополнение A_{ik} элемента a_{ik} матрицы коэффициентов и сложим полученные уравнения:

$$\sum_{i=1}^{n} A_{ik} \sum_{j=1}^{n} a_{ij} x_j = \sum_{i=1}^{n} b_i A_{ik} ,$$

что после изменения порядка суммирования дает

$$\sum_{j=1}^{n} x_j \sum_{i=1}^{n} a_{ij} A_{ik} = \sum_{i=1}^{n} b_i A_{ik} .$$

Так как сумма произведений элементов любого ряда детерминанта на алгебраические дополнения элементов любого другого параллельного ряда равна нулю, то среди всех n внутренних сумм в левой части только одна – соответствующая $j=k$ – будет отлична от нуля. Следовательно:

$$x_k \sum_{i=1}^{n} a_{ik} A_{ik} = \sum_{i=1}^{n} b_i A_{ik} .$$

Сумма в левой части при любом $k=\overline{1,n}$ равна детерминанту матрицы коэффициентов исходной системы:

$$\sum_{i=1}^{n} a_{ik} A_{ik} = |\mathbf{A}| = \Delta .$$

Сумма в правой части равна детерминанту матрицы, полученной из матрицы коэффициентов заменой k-го столбца на столбец свободных членов:

$$\sum_{i=1}^{n} b_i A_{ik} = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1k-1} & b_1 & a_{1k+1} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2k-1} & b_2 & a_{2k+1} & \dots & a_{2n} \\ \dots \\ a_{n1} & a_{n2} & \dots & a_{nk-1} & b_n & a_{nk+1} & \dots & a_{nn} \end{vmatrix} = \Delta_k .$$

Окончательно:

$$x_k \Delta = \Delta_k , \;\; k=\overline{1,n} . \qquad (3.12)$$

Соотношения (3.12) называются *формулами Крамера*.

Пусть ранг матрицы коэффициентов совпадает с числом неизвестных (и по необходимости равен рангу расширенной матрицы). Тогда для системы выполнены условия теоремы Кронекера–Капелли. В этом случае детерминант Δ отличен от нуля, и решение системы дается формулами

$$x_k = \frac{\Delta_k}{\Delta} , \;\; k=\overline{1,n} .$$

Если детерминант матрицы коэффициентов равен нулю (ранг матрицы коэффициентов меньше числа неизвестных), то система может быть либо несовместной (если хотя бы один из детерминантов Δ_k отличен от нуля – в этом случае ранг расширенной матрицы больше ранга матрицы коэффициентов и условия теоремы Кронекера–Капелли нарушены), либо неопределенной (если все детерминанты Δ_k равны нулю).

Решим систему

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

методом Кремера. Детерминант матрицы коэффициентов равен

$$\Delta = \begin{vmatrix} 1 & 1 & 2 \\ 3 & -2 & -1 \\ 1 & -2 & -2 \end{vmatrix} = \begin{vmatrix} -2 & -1 \\ -2 & -2 \end{vmatrix} - \begin{vmatrix} 3 & -1 \\ 1 & -2 \end{vmatrix} + 2 \cdot \begin{vmatrix} 3 & -2 \\ 1 & -2 \end{vmatrix} = 4 - 2 + 6 - 1 - 12 + 4 = -1 .$$

Он отличен от нуля, поэтому система совместна и определенна. Далее,

$$\Delta_1 = \begin{vmatrix} -4 & 1 & 2 \\ -3 & -2 & -1 \\ 1 & -2 & -2 \end{vmatrix} = 1 ; \ \Delta_2 = \begin{vmatrix} 1 & -4 & 2 \\ 3 & -3 & -1 \\ 1 & 1 & -2 \end{vmatrix} = -1 ; \ \Delta_3 = \begin{vmatrix} 1 & 1 & -4 \\ 3 & -2 & -3 \\ 1 & -2 & 1 \end{vmatrix} = 2 .$$

Следовательно:

$$x = \frac{\Delta_1}{\Delta} = \frac{1}{-1} = -1 ; \ y = \frac{\Delta_2}{\Delta} = \frac{-1}{-1} = 1 ; \ z = \frac{\Delta_3}{\Delta} = \frac{2}{-1} = -2 .$$

Формулы Крамера позволяют явно выразить элементы матрицы $\mathbf{A}^{-1} = \left(\alpha_{ij}\right)$, обратной к матрице $\mathbf{A}$. Обращение матрицы соответствует одновременному решению n систем линейных уравнений, правые части которых являются столбцами единичной матрицы.

Пусть отыскивается j-й столбец $\left(\alpha_{1j}, \alpha_{2j}, ..., \alpha_{nj}\right)^T$ обратной матрицы; тогда для его нахождения следует решить систему

$$\begin{cases} a_{11}\alpha_{1j} + a_{12}\alpha_{2j} + ... + a_{1n}\alpha_{nj} = 0 \\ a_{21}\alpha_{1j} + a_{22}\alpha_{2j} + ... + a_{2n}\alpha_{nj} = 0 \\ ... \\ a_{k1}\alpha_{1j} + a_{k2}\alpha_{2j} + ... + a_{kn}\alpha_{nj} = 1 \\ ... \\ a_{n1}\alpha_{1j} + a_{n2}\alpha_{2j} + ... + a_{nn}\alpha_{nj} = 0 \end{cases} . \qquad \text{(a)}$$

Детерминант, полученный из детерминанта $|\mathbf{A}|$ матрицы $\mathbf{A}$ заменой i-го столбца на j-й столбец единичной матрицы, равен:

$$\Delta_i = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1\,i-1} & 0 & a_{1\,i+1} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2\,i-1} & 0 & a_{2\,i+1} & \ldots & a_{2n} \\ \ldots & & & & \ldots & & & \\ a_{j1} & a_{j2} & \ldots & a_{j\,i-1} & 1 & a_{j\,i+1} & \ldots & a_{jn} \\ \ldots & & & & & & & \\ a_{n1} & a_{n2} & \ldots & a_{n\,i-1} & 0 & a_{n\,i+1} & \ldots & a_{nn} \end{vmatrix}.$$

Раскладывая его по i-му столбцу, получим только одно ненулевое слагаемое

$$\Delta_i = A_{ji}.$$

Решение системы (а) можно записать, используя формулы Крамера:

$$\alpha_{ij} = \frac{\Delta_i}{|\mathbf{A}|} = \frac{A_{ji}}{|\mathbf{A}|}, \; i = \overline{1,n}.$$

Следовательно, *j-й столбец обратной матрицы $\mathbf{A}^{-1}$ образован частными от деления алгебраических дополнений j-й строки матрицы* **A** *на определитель этой матрицы*.

Обратная матрица равна:

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|}\begin{pmatrix} A_{11} & A_{21} & \ldots & A_{n1} \\ A_{12} & A_{22} & \ldots & A_{n2} \\ \ldots & & & \\ A_{1n} & A_{2n} & \ldots & A_{nn} \end{pmatrix}. \qquad (3.13)$$

Матрица в правой части (3.13) по отношению к матрице **A** называется *союзной*, или *присоединенной*. Союзная матрица – результат замены в матрице **A** каждого элемента его алгебраическим дополнением и последующего транспонирования.

Пусть, например, требуется обратить матрицу

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

Решение целесообразно начать с нахождения алгебраических дополнений каждого элемента.

$$A_{11} = \begin{vmatrix} 2 & -1 \\ 1 & -2 \end{vmatrix} = -3; \; A_{12} = -\begin{vmatrix} -1 & -1 \\ 2 & -2 \end{vmatrix} = -4; \; A_{13} = \begin{vmatrix} -1 & 2 \\ 2 & 1 \end{vmatrix} = -5;$$

$$A_{21} = -\begin{vmatrix} -2 & 2 \\ 1 & -2 \end{vmatrix} = -2\,;\ A_{22} = \begin{vmatrix} 1 & 2 \\ 2 & -2 \end{vmatrix} = -6\,;\ A_{23} = -\begin{vmatrix} 1 & -2 \\ 2 & 1 \end{vmatrix} = -5\,;$$

$$A_{31} = \begin{vmatrix} -2 & 2 \\ 2 & -1 \end{vmatrix} = -2\,;\ A_{32} = -\begin{vmatrix} 1 & 2 \\ -1 & -1 \end{vmatrix} = -1\,;\ A_{33} = \begin{vmatrix} 1 & -2 \\ -1 & 2 \end{vmatrix} = 0\,.$$

Определитель найдем, выполняя разложение по первой строке:

$$|\mathbf{A}| = \sum_{j=1}^{3} a_{1j} A_{1j} = 1 \cdot (-3) + (-2) \cdot (-4) + 2 \cdot (-5) = -3 + 8 - 10 = -5\,.$$

Обратная матрица равна:

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|}\left(A_{ji}\right) = -\frac{1}{5}\begin{pmatrix} -3 & -2 & -2 \\ -4 & -6 & -1 \\ -5 & -5 & 0 \end{pmatrix} = \begin{pmatrix} 3/5 & 2/5 & 2/5 \\ 4/5 & 6/5 & 1/5 \\ 1 & 1 & 0 \end{pmatrix}.$$

И формулы Крамера (3.12), и соотношение (3.13) имеют лишь теоретическое значение и непригодны для практического применения.

Действительно, для решения системы n-го порядка методом Крамера необходимо вычислить $n+1$ детерминант n-го порядка. Если детерминанты вычисляются разложением по какому-либо ряду (так, как это было сделано в приведенном примере), то метод полностью теряет пригодность уже при $n \sim 10$. Даже тогда, когда для вычисления детерминантов используется метод Гаусса, решение по формулам Крамера требует примерно в n раз большего числа действий, нежели схема единственного деления. Сходные рассуждения свидетельствуют о практической нецелесообразности использования (3.13) для обращения матриц (обращение матрицы n-го порядка требует вычисления n^2 детерминантов $(n-1)$-го порядка).

Контрольные вопросы и упражнения

1. Что называют формулами Крамера?

2. Может ли быть совместной система, детерминант матрицы коэффициентов которой равен нулю? Может ли такая система быть совместной и определенной? Может ли она быть несовместной?

3. Исходя из значений детерминантов, входящих в формулы Крамера, сформулируйте необходимое и достаточное условие совместности и определенности системы.

4. Используя формулы Крамера, решите системы

$$\begin{cases} 2x - y = -1 \\ x + y = 4 \end{cases}; \quad \begin{cases} x + 2y - 2z = -3 \\ 2x - y + 2z = 5 \\ x + 2y + 2z = 1 \end{cases}.$$

5. Используя методы теории определителей (соотношение 3.13), обратите основные матрицы систем предыдущего упражнения. Выполните проверку полученных результатов.

3.7. Задача собственных значений

Пусть заданы квадратная матрица $\mathbf{A} = (a_{ij})$ порядка n и ненулевой вектор-столбец $\mathbf{x} = (x_i)$ той же высоты. Если при этом вектор $\mathbf{x}$ и произведение $\mathbf{Ax}$ оказываются линейно зависимы:

$$\mathbf{Ax} = \lambda \mathbf{x}, \ \lambda \neq 0,$$

то вектор $\mathbf{x}$ называют *собственным* (иначе, *характеристическим*) *вектором*, а соответствующее число λ – *собственным значением* (или *характеристическим числом*) матрицы $\mathbf{A}$.

Найдем собственные значения и собственные векторы матрицы $\mathbf{A}$. Так как i-й элемент вектора-столбца $\mathbf{Ax}$ равен $\sum_{j=1}^{n} a_{ij} x_j$, то для нахождения собственных значений следует решить систему линейных уравнений

$$\sum_{j=1}^{n} a_{ij} x_j = \lambda x_i, \ i = \overline{1,n},$$

или

$$(a_{ii} - \lambda) x_i + \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij} x_j = 0, \ i = \overline{1,n}.$$

В развернутой форме:

$$\begin{cases} (a_{11} - \lambda)x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0 \\ a_{21}x_1 + (a_{22} - \lambda)x_2 + \ldots + a_{2n}x_n = 0 \\ \ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + (a_{nn} - \lambda)x_n = 0 \end{cases}. \qquad (3.14)$$

Для того чтобы эта однородная система имела нетривиальные решения, необходимо, чтобы ее столбцы были линейно зависимы. Но тогда должен быть равен нулю ее детерминант:

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ \dots & & & \\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda \end{vmatrix} = 0. \tag{3.15}$$

Уравнение (3.15) представляет собой алгебраическое уравнение n-й степени относительно λ. Его называют *характеристическим уравнением* матрицы $\mathbf{A}$. Характеристическое уравнение можно представить в форме

$$|\mathbf{A}-\lambda\mathbf{E}| = 0.$$

Матрица, находящаяся в правой части (3.15) под знаком детерминанта, может быть записана в виде $\mathbf{A}-\lambda\mathbf{E}$; ее называют *характеристической матрицей*. Левую часть характеристического уравнения (детерминант характеристической матрицы) называют *характеристическим многочленом*.

Матрица n-го порядка имеет не более n действительных собственных значений. В комплексном (координатном) n-мерном линейном пространстве матрица n-го порядка ровно n собственных значений (считая с кратными).

Теорема 3.7.1. *Собственные векторы* $\mathbf{x}_1$, $\mathbf{x}_2$, …, $\mathbf{x}_n$, *соответствующие различным собственным значениям* λ_1, λ_2, …, λ_n *матрицы* $\mathbf{A}$, *линейно независимы.*

Доказательство выполним индуктивно, по числу собственных векторов. Пусть собственный вектор $\mathbf{x}_1$ – единственный. Так как по определению $\mathbf{x}_1 \neq \mathbf{0}$, то теорема справедлива.

Пусть утверждение теоремы доказано для системы из n векторов $\mathbf{x}_1$, $\mathbf{x}_2$, …, $\mathbf{x}_n$. Присоединим к системе вектор $\mathbf{x}_{n+1}$ и найдем коэффициенты нулевой линейной комбинации

$$\sum_{k=1}^{n+1} \alpha_k \mathbf{x}_k = \mathbf{0}. \tag{a}$$

Умножим обе части равенства (a) справа на матрицу **A**. С учетом свойств операции умножения матриц получим

$$\mathbf{A}\left(\sum_{k=1}^{n+1}\alpha_k \mathbf{x}_k\right)=\mathbf{0},\ \sum_{k=1}^{n+1}\mathbf{A}(\alpha_k \mathbf{x}_k)=\mathbf{0},\ \sum_{k=1}^{n+1}\alpha_k \mathbf{A}\mathbf{x}_k=\mathbf{0}.$$

Так как $\mathbf{x}_k$ – собственные векторы, то $\mathbf{A}\mathbf{x}_k=\lambda_k \mathbf{x}_k$. Следовательно

$$\sum_{k=1}^{n+1}\alpha_k\lambda_k \mathbf{x}_k=\mathbf{0}. \tag{b}$$

Умножим обе части (a) на число λ_{n+1}:

$$\sum_{k=1}^{n+1}\lambda_{n+1}\alpha_k \mathbf{x}_k=\mathbf{0}. \tag{c}$$

Вычтем (c) из (b). Тогда в левой части последнее слагаемое будет нулевым: $(\lambda_{n+1}-\lambda_{n+1})\alpha_k \mathbf{x}_k=\mathbf{0}$, откуда

$$\sum_{k=1}^{n}(\lambda_k-\lambda_{n+1})\alpha_k \mathbf{x}_k=\mathbf{0}. \tag{d}$$

Векторы $\mathbf{x}_1$, $\mathbf{x}_2$, …, $\mathbf{x}_n$ линейно независимы по индуктивному предположению. Поэтому все коэффициенты линейной комбинации (d) должны быть нулевыми. Но собственные значения λ_k различны; следовательно $\lambda_k-\lambda_{n+1}\neq 0$, откуда $\alpha_1=\alpha_2=\ldots=\alpha_n=0$. Соотношение (a) переходит в равенство

$$\alpha_{n+1}\mathbf{x}_{n+1}=\mathbf{0}.$$

Вектор $\mathbf{x}_{n+1}$ – ненулевой, поэтому $\alpha_{n+1}=0$. Следовательно, линейная комбинация (a) является тривиальной, и векторы $\mathbf{x}_1$, $\mathbf{x}_2$, …, $\mathbf{x}_{n+1}$ линейно независимы. Теорема доказана.

Для отыскания собственных значений можно использовать непосредственно соотношение (3.15). Однако в этом случае основная сложность заключается в отсутствии т.н. *методов решения в радикалах* для алгебраических уравнений степени выше 4 (впрочем, уже для третьей степени соответствующий метод весьма громоздок).

Пусть требуется найти собственные значения и собственные векторы матрицы

$$\mathbf{A} = \begin{vmatrix} 1 & -1 & 0 \\ 1 & 3 & 0 \\ 2 & -1 & 3 \end{vmatrix}.$$

Характеристическое уравнение имеет вид

$$(3-\lambda)\left(\lambda^2 - 4\lambda + 4\right) = 0,$$

или

$$(3-\lambda)(\lambda-2)^2 = 0.$$

Оно имеет три действительных корня: однократный $\lambda_1 = 3$ и один двукратный $\lambda_2 = 2$.

Для нахождения собственных векторов $\mathbf{x} = \left(x_1, x_2, x_3\right)^T$ подставим найденные корни характеристического уравнения в систему (3.14). Для корня $\lambda_1 = 3$:

$$\begin{cases} -2x_1 - x_2 = 0 \\ x_1 = 0 \\ 2x_1 - x_2 = 0 \end{cases}.$$

Решением этой системы является вектор $\mathbf{x} = (0,0,\alpha)$, где α – произвольное число.

Для корня $\lambda_2 = 2$:

$$\begin{cases} -x_1 - x_2 = 0 \\ x_1 + x_2 = 0 \\ 2x_1 - x_2 + x_3 = 0 \end{cases}.$$

Выбирая в качестве свободной неизвестной $x_1 = \beta$, найдем:

$$\begin{cases} x_2 = -\beta \\ x_3 = -3\beta \end{cases}.$$

Решением является вектор $\mathbf{x} = (\beta,-\beta,-3\beta)$, где β – произвольное число.

Пусть требуется найти собственные значения матрицы

$$\mathbf{A} = \begin{vmatrix} 1 & -1 \\ 1 & 1 \end{vmatrix}.$$

В данном случае характеристическое уравнение

$$\lambda^2 - 2\lambda + 2 = 0$$

имеет мнимые корни $\lambda_1 = 1 + i$, $\lambda_2 = 1 - i$, где i – мнимая единица. Поэтому в действительном (двумерном координатном) линейном пространстве ни один вектор $\mathbf{x}$ не будет являться собственным вектором данной матрицы (можно проверить, что собственными векторами данной матрицы являются $\mathbf{x}_1 = (\alpha i, \alpha)$ и $\mathbf{x}_2 = (-\beta i, \beta)$, где α и β – произвольные числа).

Контрольные вопросы и упражнения

1. Что называют собственным вектором матрицы? Что называют собственным значением матрицы?

2 [9]. Найдите собственные значения и собственные векторы матриц:

$$\begin{pmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{pmatrix}; \begin{pmatrix} 2 & -1 & 2 \\ 5 & -3 & 3 \\ -1 & 0 & -2 \end{pmatrix}; \begin{pmatrix} 4 & -5 & 2 \\ 5 & -7 & 3 \\ 6 & -9 & 4 \end{pmatrix}.$$

Рекомендуемая литература

[1]. Александров П.С. Курс аналитической геометрии и линейной алгебры. – М.: Наука, 1979. – 511 с.
[2]. Беклемишев Д.В. Курс аналитической геометрии и линейной алгебры. – М.: ФИЗМАТЛИТ, 2003. – 303 с.
[3]. Беклемишев Д.В. Дополнительные главы линейной алгебры. – М.: Наука, 1983. – 335 с.
[4]. Бугров Я.С., Никольский С.М. Высшая математика. Элементы линейной алгебры и аналитической геометрии. – М.: Наука, 1988. – 224 с.
[5]. Верещагин Н.К., Шень А. Лекции по математической логике и теории алгоритмов. Часть 1. Начала теории множеств. – М.: МЦНМО, 1999. – 128 с.
[6]. Воеводин В.В. Линейная алгебра. – М.: Наука, 1980. – 400 с.
[7]. Ефимов Н.В., Розендорн Э.Р. Линейная алгебра и многомерная геометрия. – М.: Наука, 1970. – 528 с.
[8]. Рублев А.Н. Курс линейной алгебры и аналитической геометрии. – М.: Высшая школа, 1972. – 424 с.
[9]. Сборник задач по математике для втузов. Линейная алгебра и основы математического анализа. / Под ред. А.В. Ефимова, Б.П. Демидовича. – М.: Наука, 1981. – 464 с.
[10]. Фадеев Д.К., Соминский И.С. Сборник задач по высшей алгебре. – М.: Наука, 1977. – 288 с.
[11]. Федорчук В.В. Курс Аналитической геометрии и линейной алгебры. – М.: Изд-во МГУ, 1990. – 328 с.