§1. Системы линейных уравнений. Основные понятия.

Система m линейных уравнений с n неизвестными имеет вид

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

Коэффициенты при неизвестных нумеруются двойными индексами. Через a_{ij} обозначен коэффициент в *i-м* уравнении при неизвестной x_j. Линейные уравнения можно рассматривать как частный случай систем, когда $m=1$.

С каждой системой линейных уравнений связывают следующие объекты.

1. *Матрица коэффициентов* системы — таблица из m строк и n столбцов, составленная из коэффициентов при неизвестных:

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

В случае, когда матрица состоит из m строк и n столбцов, её называют прямоугольной матрицей типа $m\times n$. Если количество строк совпадает с количеством столбцов, т. е. $m=n$, то говорят о квадратной матрице порядка n.

2. *Столбец неизвестных:* $X=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$.

3. *Столбец свободных членов:* $B=\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$.

Используя введённые обозначения, систему иногда записывают в символическом виде $AX=B$.

4. *Расширенная матрица:* $(A|B)=\left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ . & . & \dots & . & . \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right)$

5. *Решение* системы — упорядоченный набор $(\alpha_1, \alpha_2, \dots, \alpha_n)$ чисел, подстановка которых в систему вместо соответствующих неизвестных превращает *каждое* уравнение системы в верное равенство.

Отметим специально, что решением системы являются не сами числа α_i, а их *набор.*

Определение 1. Система уравнений называется *совместной,* если она имеет по крайней мере одно решение.

В противном случае система называется *несовместной.* Другими словами, каков бы ни был упорядоченный набор из *n* чисел, в несовместной системе найдётся хотя бы одно уравнение, подстановка в которое чисел набора на даёт верного равенства.

Пример 1. Система $\begin{cases} 2x_1+3x_2-x_3=-4, \\ x_1-x_2-3x_3=3 \end{cases}$ является совместной, поскольку упорядоченная тройка чисел $(1,-2,0)$ является её решением.

Пример 2. Очевидно, что система $\begin{cases} x_1+x_2=-1, \\ x_1+x_2=1 \end{cases}$ несовместная.

Определение 2. Система уравнений называется *определённой,* если она имеет единственное решение.

Таким образом определённая система является совместной. Если совместная система имеет более одного решения, то она называется *неопределённой.*

Пример 3. Система из примера 1 является неопределённой, так как набор чисел $(-1,-1,-1)$ также является решением.

Пример 4. Система $\begin{cases} 1\cdot x_1+0\cdot x_2=\pi, \\ 0\cdot x_1+1\cdot x_2=\sqrt{5} \end{cases}$ является определённой.

§2. Эквивалентные системы

Пусть имеются две системы уравнений, которые условно обозначим символами **(I)** и **(II).**

Определение 1. Система ***(II)*** называется *следствием* системы ***(I)***, если каждое решение системы ***(I)*** является решением системы ***(II)***.

Тот факт, что система ***(II)*** является следствием системы ***(I)*** коротко записывается в виде $(\boldsymbol{I})\Rightarrow(\boldsymbol{II})$.

Пример 1. Система ***(II)***, состоящая из уравнения $3x_1+2x_2=-1$ является следствием системы ***(I)*** $\begin{cases} 2x_1+3x_2=-4, \\ x_1-x_2=3. \end{cases}$ Это вытекает из того, что после почленного сложения уравнений системы ***(I)*** получается уравнение ***(II).***

Заметим, что в рассматриваемом примере система ***(I)*** *не* является следствием системы ***(II).*** Например, пара $(-1,1)$ является решением ***(II),*** но не ***(I).***

Определение 2. Системы ***(I)*** и ***(II)*** называются *эквивалентными,* если всякое решение одной из них является решением другой.

Другими словами, множества решений эквивалентных систем состоят из одних и тех же элементов. Можно также сказать, что каждая из эквивалентных систем является следствием другой.

Эквивалентность систем ***(I)*** и ***(II)*** кратко обозначается так: $(\boldsymbol{I})\sim(\boldsymbol{II})$.

Поиск решений системы линейных уравнений фактически сводится к поиску эквивалентной системы специального вида, более удобного для описания решений.

Теорема. Система ***(I)*** линейных уравнений (1) эквивалентна системе ***(II)***, получающейся из неё следующими преобразованиями:

1) умножение уравнения на ненулевое число;

2) перестановка двух уравнений;

3) прибавление к одному уравнению другого, умноженного на число.

Доказательство.

В случае преобразований 1) и 2) теорема тривиальна. Поэтому рассмотрим только случай, когда система ***(II)*** получается из системы ***(I)*** преобразованием 3).

Допустим, что *i*-е уравнение системы ***(II)*** является суммой *i*-го уравнения системы ***(I)*** и *j*-го уравнения той же системы, умноженного на число μ. Очевидно, каждое решение системы ***(I)*** является решением системы ***(II)***. С другой стороны, сумма *i*-го уравнения системы ***(II)*** и *j*-го уравнения той же системы, умноженного на число $-\mu$, совпадает с *i-м* уравнением системы ***(I)***. Поэтому каждое решение системы ***(II)*** является решением системы ***(I)***.

Таким образом, системы ***(I)*** и ***(II)*** эквивалентны. □

§3. Формулы Крамера[i]. Определитель второго порядка

Рассмотрим систему двух уравнений с двумя неизвестными общего вида

$$\begin{cases} a_{11}x_1+a_{12}x_2=b_1, \\ a_{21}x_1+a_{22}x_2=b_2. \end{cases} \tag{2}$$

Для нахождения решения умножим первое уравнение на a_{22}, второе — на $-a_{12}$ и сложим почленно полученные равенства. Из полученного соотношения

$$(a_{11}a_{22}-a_{21}a_{12})x_1=b_1a_{22}-b_2a_{12},$$

предполагая множитель $a_{11}a_{22}-a_{21}a_{12}$ ненулевым, находим значение переменной x_1:

$$x_1=\frac{b_1a_{22}-b_2a_{12}}{a_{11}a_{22}-a_{21}a_{12}}. \tag{3}$$

Аналогично поступаем для нахождения x_2: первое уравнение умножаем на $-a_{21}$, второе — на a_{11} и складываем почленно полученные равенства. После очевидных преобразований находим

$$x_2=\frac{a_{11}b_2-a_{21}b_1}{a_{11}a_{22}-a_{21}a_{12}}. \tag{4}$$

Формулы (3) и (4), выражающие решение системы через коэффициенты и свободные члены, называются *формулами Крамера.* Их можно записать в виде, более удобном для запоминания, если ввести понятие определителя.

Определение. *Определителем* матрицы $\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}$ второго порядка (или просто определителем второго порядка) называется число $\alpha\cdot\delta-\beta\cdot\gamma$.

Для обозначения определителя вместо круглых скобок используют прямые. Таким образом,

$$\begin{vmatrix}\alpha & \beta\\ \gamma & \delta\end{vmatrix}=\alpha\cdot\delta-\beta\cdot\gamma.$$

Также определитель матрицы A обозначают символом $\det A$.

Пример 1. $\begin{vmatrix}4 & 7\\ 5 & 8\end{vmatrix}=4\cdot 8-7\cdot 5=-3$.

Пример 2. $\begin{vmatrix}\cos\xi & \sin\xi\\ -\sin\xi & \cos\xi\end{vmatrix}=\cos^2\xi+\sin^2\xi=1$.

Ясно, что знаменатель Δ каждой дроби в формулах Крамера является определителем матрицы системы (2)

$$\Delta=\begin{vmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix}.$$

Числитель Δ_1 дроби в формуле (3) (Δ_2 в формуле (4)) также можно записать в виде определителя, который получается из Δ заменой первого (второго) столбца столбцом свободных членов:

$$\Delta_1 = \begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}, \Delta_2 = \begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}.$$

С учётом введённых обозначений формулы (3) и (4) можно записать в виде

$$x_1 = \frac{\Delta_1}{\Delta}, x_2 = \frac{\Delta_2}{\Delta}.$$

Ещё раз отметим, что формулы Крамера получены в предположении $\Delta \neq 0$. При выполнении этого условия значения переменных x_1, x_2 определяются однозначно. Другими словами, система (2) является определённой.

Пример 3. Найдём решения системы уравнений $\begin{cases} 2x_1 + 3x_2 = -4, \\ x_1 - x_2 = 3. \end{cases}$ Так как определитель этой системы $\Delta = \begin{vmatrix} 2 & 3 \\ 1 & -1 \end{vmatrix} = -5$ отличен от нуля, то решение единственно и может быть найдено по формулам Крамера:

$$\Delta_1 = \begin{vmatrix} -4 & 3 \\ 3 & -1 \end{vmatrix} = -5, \Delta_2 = \begin{vmatrix} 2 & -4 \\ 1 & 3 \end{vmatrix} = 10,$$

$$x_1 = \frac{-5}{-5} = 1, x_2 = \frac{10}{-5} = -2.$$

§4. Свойства определителей второго порядка

Пусть A – некоторая матрица. Построим матрицу (в общем случае отличную от A) по правилу: в i-ю строку поместим элементы i-го столбца матрицы A. Составленная таким образом матрица называется матрицей, *транспонированной* к матрице A и обозначается символом A^t. Например, если $A = \begin{pmatrix} 2 & -1 & 1 \\ -7 & 3 & 5 \end{pmatrix}$, то $A^t = \begin{pmatrix} 2 & -7 \\ -1 & 3 \\ 1 & 5 \end{pmatrix}$.

Свойство 1. Для любой матрицы второго порядка $det\, A = det\, A^t$:

$$\begin{vmatrix} \alpha & \beta \\ \gamma & \delta \end{vmatrix} = \alpha\delta - \beta\gamma = \begin{vmatrix} \alpha & \gamma \\ \beta & \delta \end{vmatrix}.$$

Данное свойство означает, что справедливость утверждения об определителе не нарушится, если в формулировке этого утверждения поменять ролями строки и столбцы.

Свойство 2. При умножении *любого* столбца определителя на число, определитель умножается на это число. Например,

$$\begin{vmatrix}\alpha & k\beta\\ \gamma & k\delta\end{vmatrix}=\alpha(k\delta)-(k\beta)\gamma=k(\alpha\delta-\beta\gamma)=k\begin{vmatrix}\alpha & \gamma\\ \beta & \delta\end{vmatrix}.$$

Свойство 3. Если в определителе поменять местами столбцы, то определитель изменит знак:

$$\begin{vmatrix}\beta & \alpha\\ \delta & \gamma\end{vmatrix}=\beta\gamma-\alpha\delta=-(\alpha\delta-\beta\gamma)=-\begin{vmatrix}\alpha & \beta\\ \gamma & \delta\end{vmatrix}.$$

Свойство 4. Определитель, элементы некоторого столбца которого представлены в виде суммы двух слагаемых, равен сумме двух определителей, каждый из которых получается из исходного заменой одноимённого столбца столбцом первых и вторых слагаемых соответственно. Например,

$$\begin{vmatrix}\alpha_1+\alpha_2 & \beta\\ \gamma_1+\gamma_2 & \delta\end{vmatrix}=(\alpha_1+\alpha_2)\delta-\beta(\gamma_1+\gamma_2)=(\alpha_1\delta-\beta\gamma_1)+(\alpha_1\delta-\beta\gamma_1)=\begin{vmatrix}\alpha_1 & \beta\\ \gamma_1 & \delta\end{vmatrix}+\begin{vmatrix}\alpha_2 & \beta\\ \gamma_2 & \delta\end{vmatrix}.$$

Свойство 5. Определитель с нулевым столбцом равен нулю.

Свойство 6. Определитель с одинаковыми столбцами равен нулю.

Из свойств 2, 4 и 6 вытекает

Свойство 7. Определитель не изменится, если к элементам некоторого столбца определителя прибавить соответственные элементы другого столбца, умноженные на одно и то же число.

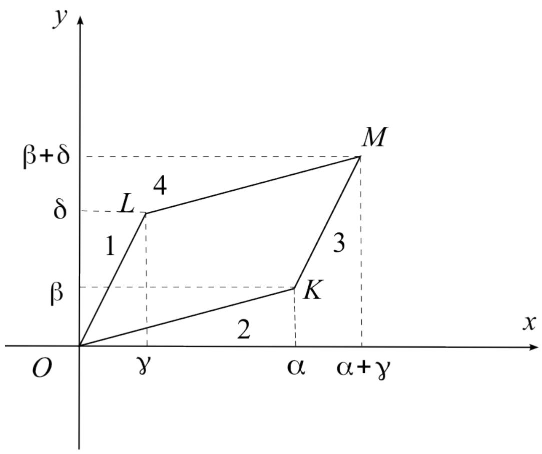
Свойство 8. Определитель равен нулю тогда и только тогда, когда его столбцы пропорциональны:

$$\begin{vmatrix}\alpha & \beta\\ \gamma & \delta\end{vmatrix}=0\Leftrightarrow\alpha\delta=\beta\gamma\Leftrightarrow\frac{\alpha}{\beta}=\frac{\gamma}{\delta}.$$

Разумеется, приведённые рассуждения имеют силу лишь в том случае, когда ни β, ни δ не равны нулю. Если это условие не выполнено, например $\beta=0$, то хотя бы одно из чисел α или δ равно нулю и свойство очевидно.

§5. Геометрический смысл определителя второго порядка

Для выяснения геометрического смысла определителя второго порядка вычислим площадь параллелограмма (см. рисунок), построенного на векторах $\overrightarrow{OK}=(\alpha,\beta)$ и $\overrightarrow{OL}=(\gamma,\delta)$. Проще всего это сделать, вычитая из площади прямоугольника с вершинами в точках $O,\alpha+\gamma,M,\beta+\delta$ площади «лишних» фигур: треугольников $1,2$ и трапеций $3,4$. Из чертежа видно, что сумма



площадей треугольника 1 и трапеции 3 равна площади прямоугольника со сторонами γ и $\beta+\delta$. Аналогичным образом, сумма площадей треугольника 2 и трапеции 4 равна площади прямоугольника со сторонами $\alpha+\gamma$ и β. Поэтому искомая площадь равна

$$(\alpha+\gamma)(\beta+\delta)-\gamma(\beta+\delta)-(\alpha+\gamma)\beta=\alpha(\beta+\delta)-(\alpha+\gamma)\beta=\alpha\delta-\gamma\beta=\begin{vmatrix}\alpha & \beta\\ \gamma & \delta\end{vmatrix}.$$

Итак, площадь параллелограмма $OKML$ равна определителю $\begin{vmatrix}\alpha & \beta\\ \gamma & \delta\end{vmatrix}$, составленному из координат векторов, порождающих параллелограмм.

На самом деле определитель может оказаться отрицательным числом и даже равняться нулю. В последнем случае, согласно свойствам 1 и 8, координаты векторов $\overrightarrow{OK}$ и $\overrightarrow{OL}$ пропорциональны, а значит сами векторы коллинеарны. Поэтому параллелограмм, порождаемый ими, является вырожденным и естественно считать его площадь равной нулю.

В случае, когда векторы $\overrightarrow{OK}=(\alpha,\beta)$ и $\overrightarrow{OL}=(\gamma,\delta)$ неколлинеарны, можно показать, что определитель $\begin{vmatrix}\alpha & \beta\\ \gamma & \delta\end{vmatrix}$ будет положительным если и только если кратчайший поворот от вектора $\overrightarrow{OK}$ к вектору $\overrightarrow{OL}$ происходит *против* часовой стрелки. Следовательно, в определителе заключена информация о расположении сторон параллелограмма или, как принято говорить, об *ориентации* параллелограмма.

Коротко этот факт выражают следующими словами: определитель $\begin{vmatrix}\alpha & \beta\\ \gamma & \delta\end{vmatrix}$ равен *ориентированной* площади параллелограмма, построенного на векторах с координатами (α,β) и (γ,δ).

Пример. Вычислим площадь параллелограмма, построенного на векторах $\vec{a}=(-4,7)$ и $\vec{b}=(5,-8)$. Так как $\begin{vmatrix}-4 & 7\\ 5 & -8\end{vmatrix}=-3$, то площадь параллелограмма равна 3 кв. ед., причём кратчайший поворот от $\vec{a}$ к $\vec{b}$ происходит *по* часовой стрелке.

§6. Однородные системы двух линейных уравнений с тремя неизвестными

Система (2) линейных уравнений называется *однородной,* если все свободные члены b_i равны нулю. Рассмотрим систему двух линейных уравнений с тремя неизвестными[1] c_1, c_2, c_3:

$$\begin{cases} a_{11}c_1 + a_{21}c_2 + a_{31}c_3 = 0, \\ a_{12}c_1 + a_{22}c_2 + a_{32}c_3 = 0 \end{cases} \qquad (5)$$

Естественно считать, что коэффициенты уравнений непропорциональны, так как иначе система (5) эквивалентна системе из одного уравнения. Согласно свойству 8, это означает, что один из определителей $\begin{vmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{vmatrix}$, $\begin{vmatrix} a_{21} & a_{31} \\ a_{22} & a_{32} \end{vmatrix}$ или $\begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix}$ отличен от нуля. Без ограничения общности можно считать, что

$$\begin{vmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{vmatrix} \neq 0.$$

Этого всегда можно добиться подходящей перенумерацией неизвестных.

Мы хотим описать все решения системы (5), удовлетворяющей перечисленным условиям. Заметим, прежде всего, что если набор $(c_1, c_2, 0)$ является решением, то $c_1 = c_2 = 0$. Это напрямую вытекает из формул Крамера.

Предположим теперь, что $c_3 \neq 0$. В этом случае систему (5) можно переписать в виде

$$\begin{cases} a_{11}\dfrac{c_1}{c_3} + a_{21}\dfrac{c_2}{c_3} = -a_{31}, \\ a_{12}\dfrac{c_1}{c_3} + a_{22}\dfrac{c_2}{c_3} = -a_{32} \end{cases}$$

и рассматривать последнюю систему как систему двух уравнений от *двух* неизвестных $\dfrac{c_1}{c_3}$ и $\dfrac{c_2}{c_3}$.

Используя формулы Крамера и свойства определителей, выпишем решение

1 В системе (5) мы намеренно отступили от принятого ранее правила индексации коэффициентов при неизвестных.

$$\frac{c_1}{c_3}=\frac{\begin{vmatrix} a_{21} & a_{31} \\ a_{22} & a_{32} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{vmatrix}},\quad \frac{c_2}{c_3}=\frac{\begin{vmatrix} a_{31} & a_{11} \\ a_{32} & a_{12} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{vmatrix}}.$$

Из найденных формул следует, что компоненты c_i решения системы (5) пропорциональны соответствующим определителям второго порядка:

$$c_1=\lambda\begin{vmatrix} a_{21} & a_{31} \\ a_{22} & a_{32} \end{vmatrix}, c_2=\lambda\begin{vmatrix} a_{31} & a_{11} \\ a_{32} & a_{12} \end{vmatrix}, c_3=\lambda\begin{vmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{vmatrix},$$

где λ — произвольное ненулевое число. Если допустить и нулевое значение параметра λ, то последние формулы определяют все решения системы (5).

§7. Системы трёх линейных уравнений с тремя неизвестными

Попытаемся получить формулы, аналогичные (3) и (4) для случая трёх уравнений с тремя неизвестными

$$\begin{cases} a_{11}x_1+a_{12}x_2+a_{13}x_3=b_1, \\ a_{21}x_1+a_{22}x_2+a_{23}x_3=b_2, \\ a_{31}x_1+a_{32}x_2+a_{33}x_3=b_3. \end{cases} \tag{6}$$

Ограничимся выводом формулы для переменной x_3. С другими переменными рассуждения аналогичны. Воспользуемся тем же приёмом, что и при решении системы (2). Умножим i-е уравнение системы (6) на c_i и сложим их. Числа c_i выберем так, чтобы в полученном уравнении

$$(a_{11}c_1+a_{21}c_2+a_{31}c_3)x_1+(a_{12}c_1+a_{22}c_2+a_{32}c_3)x_2+(a_{13}c_1+a_{23}c_2+a_{33}c_3)x_3= \\ =b_1c_1+b_2c_2+b_3c_3$$

коэффициенты при x_1 и x_2 обратились в нуль. Для этого c_i должны удовлетворять системе (5). Если коэффициенты при x_1 и x_2 непропорциональны, то можно воспользоваться результатом предыдущего параграфа и искомое выражение для x_3 найдено:

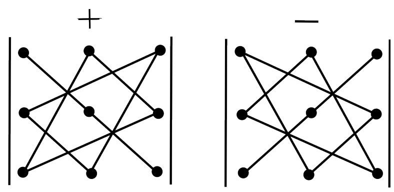
$$x_3=\frac{b_1\begin{vmatrix}a_{21} & a_{22}\\ a_{31} & a_{32}\end{vmatrix}+b_2\begin{vmatrix}a_{31} & a_{32}\\ a_{11} & a_{12}\end{vmatrix}+b_3\begin{vmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix}}{a_{13}\begin{vmatrix}a_{21} & a_{22}\\ a_{31} & a_{32}\end{vmatrix}+a_{23}\begin{vmatrix}a_{31} & a_{32}\\ a_{11} & a_{12}\end{vmatrix}+a_{33}\begin{vmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix}}.$$

Знаменатель[2] последней дроби, который обозначим через Δ, после очевидных преобразований можно записать в виде

$$\Delta=a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31}-a_{12}a_{21}a_{33}-a_{11}a_{23}a_{32}.$$

Определение. *Определителем* матрицы $\begin{pmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{pmatrix}$ третьего порядка (или просто определителем третьего порядка) называется число Δ, вычисленное по предыдущей формуле.

Для запоминания правой части равенства можно использовать так называемое правило треугольников. Схема этого правила изображена на рисунке.



Сумма в числителе дроби, определяющей x_3, отличается от суммы в знаменателе лишь первыми множителями. Поэтому числитель можно представить в виде определителя третьего порядка, получающегося из Δ заменой третьего столбца столбцом свободных членов. Если этот определитель обозначить через Δ_3, то

$$x_3=\frac{\Delta_3}{\Delta},\quad \text{где}\quad \Delta_3=\begin{vmatrix}a_{11} & a_{12} & b_1\\ a_{21} & a_{22} & b_2\\ a_{31} & a_{32} & b_3\end{vmatrix}.$$

Для других переменных имеют место аналогичные формулы:

$$x_1=\frac{\Delta_1}{\Delta},\, x_2=\frac{\Delta_2}{\Delta},$$

---

2 Рассматриваются лишь те системы, для которых знаменатель не равно нулю.

где $\Delta_1=\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}, \Delta_2=\begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}.$

Похожими рассуждениями, рассматривая системы четырёх уравнений с четырьмя неизвестными, можно прийти к понятию определителя четвёртого порядка, затем пятого и т. д.

Заметим, однако, что формулы Крамера определяют решения лишь тех систем, определитель которых отличен от нуля. Для систем с нулевым определителем и систем, у которых число уравнений не совпадает с числом переменных[3] требуются другие методы нахождения решений. Один из них мы обсудим ниже.

§8. Арифметическое *n*-мерное пространство

Определение. *Действительным арифметическим* *n*-мерным пространством $\mathbb{R}^n$ называется множество упорядоченных наборов *n* действительных чисел:

$$\mathbb{R}^n=\{(a_1, a_2, \ldots, a_n) \mid \forall\, i=1,2, \ldots, n \quad a_i \in \mathbb{R}\}.$$

Элементы пространства $\mathbb{R}^n$ называются *арифметическими векторами* и обычно обозначаются полужирными строчными буквами латинского алфавита. Для арифметического вектора $\boldsymbol{a}=(a_1, a_2, \ldots, a_n)$ число a_i называется его *i-й компонентой.* Два вектора $\boldsymbol{a}$ и $\boldsymbol{b}$ называются *равными,* если равны их соответствующие компоненты:

$$\boldsymbol{a}=\boldsymbol{b} \Leftrightarrow \forall\, i=1,2, \ldots, n \quad a_i=b_i.$$

В пространстве $\mathbb{R}^n$ естественным образом определены операции сложения векторов и умножения вектора на число. *Суммой* векторов $\boldsymbol{a}=(a_1, a_2, \ldots, a_n)$ и $\boldsymbol{b}=(b_1, b_2, \ldots, b_n)$ называется вектор с компонентами $(a_1+b_1, \ldots, a_n+b_n)$. Сумма векторов $\boldsymbol{a}$ и $\boldsymbol{b}$ обозначается $\boldsymbol{a}+\boldsymbol{b}$.

Произведением вектора $\boldsymbol{a}=(a_1, a_2, \ldots, a_n)$ на число λ называется вектор с компонентами $(\lambda a_1, \lambda a_2, \ldots, \lambda a_n)$. Произведение вектора $\boldsymbol{a}$ на число λ обозначается $\lambda \boldsymbol{a}$.

Вектор $\mathbf{0}=(0, 0, \ldots, 0)$ с нулевыми компонентами называется *нулевым.*

---

3 В принципе, всегда возможно данную систему вида (1) заменить эквивалентной системой с одинаковым количеством строк и столбцов. Если $m<n$, то можно продублировать какое-нибудь уравнение нужное число раз. Если $m>n$, то недостающие переменные можно добавить, взяв их с нулевыми коэффициентами.

Вектор $(-a_1, -a_2, \dots, -a_n)$ называется вектором, *противоположным* вектору $\boldsymbol{a}=(a_1, a_2, \dots, a_n)$. Противоположный к $\boldsymbol{a}$ вектор обозначается $-\boldsymbol{a}$.

Сумма $\boldsymbol{a}+(-\boldsymbol{b})$ векторов $\boldsymbol{a}$ и $-\boldsymbol{b}$ называется *разностью* векторов $\boldsymbol{a}$ и $\boldsymbol{b}$ и обозначается $\boldsymbol{a}-\boldsymbol{b}$.

Перечислим некоторые свойства операций сложения векторов и умножения вектора на число. Греческими буквами обозначаются действительные числа.

Сложение векторов:

1. $\forall \boldsymbol{a}, \boldsymbol{b} \quad \boldsymbol{a}+\boldsymbol{b}=\boldsymbol{b}+\boldsymbol{a}$;
2. $\forall \boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \quad (\boldsymbol{a}+\boldsymbol{b})+\boldsymbol{c}=\boldsymbol{a}+(\boldsymbol{b}+\boldsymbol{c})$;
3. $\exists \mathbf{0} | \forall \boldsymbol{a} \quad \boldsymbol{a}+\boldsymbol{o}=\mathbf{0}+\boldsymbol{a}=\boldsymbol{a}$;
4. $\forall \boldsymbol{a} \quad \exists(-\boldsymbol{a}) | \boldsymbol{a}+(-\boldsymbol{a})=(-\boldsymbol{a})+\boldsymbol{a}=\mathbf{0}$.

Умножение вектора на число:

1. $\forall \boldsymbol{a}, \forall \lambda, \mu \quad (\lambda\mu)\boldsymbol{a}=\lambda(\mu\boldsymbol{a})$;
2. $\forall \boldsymbol{a}, \forall \lambda, \mu \quad (\lambda+\mu)\boldsymbol{a}=\lambda\boldsymbol{a}+\mu\boldsymbol{a}$;
3. $\forall \boldsymbol{a}, \boldsymbol{b} \quad \lambda(\boldsymbol{a}+\boldsymbol{b})=\lambda\boldsymbol{a}+\lambda\boldsymbol{b}$;
4. $\forall \boldsymbol{a} \quad 1\boldsymbol{a}=\boldsymbol{a}$.

Замечание. Понятие векторного пространства играет важную роль в современной математике и её приложениях. Оно обобщает пространство направленных отрезков, изучаемое в школьном курсе математике. В самом общем случае векторным пространством называют всякое множество, элементы которого можно складывать между собой и умножать на число. При этом требуется, чтобы эти две операции (сложение и умножение на число) удовлетворяли перечисленным выше восьми свойствам.

Таким образом, с современной точки зрения векторное пространство образуют не только множество направленных отрезков, но и упорядоченные наборы *n* действительных чисел. Множество всех вещественнозначных функций, определённых, скажем, на отрезке $[0;1]$, также образуют векторное пространство относительно естественных операций сложения и умножения на число.

§9. Линейная зависимость векторов. Базис

Пусть дана некоторая *конечная* совокупность векторов $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_l$, среди которых могут быть и равные. Для любых чисел $\alpha_1, \alpha_2, \dots, \alpha_l$ вектор

$$\alpha_1\boldsymbol{a}_1+\alpha_2\boldsymbol{a}_2+\dots+\alpha_l\boldsymbol{a}_l$$

называется линейной комбинацией векторов $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_l$.

Определение. Конечная совокупность векторов $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_l$ называется *ли-*

нейно независимой, если равенство $\alpha_1 \boldsymbol{a_1} + \alpha_2 \boldsymbol{a_2} + \ldots + \alpha_l \boldsymbol{a_l} = \mathbf{0}$ возможно тогда и только тогда, когда $\alpha_1 = \alpha_2 = \ldots = \alpha_l = 0$.

Совокупность векторов, не являющаяся линейно независимой, называется *линейно зависимой.* Другими словами, система векторов $\boldsymbol{a_1}, \boldsymbol{a_2}, \ldots, \boldsymbol{a_l}$ линейно зависима, если найдутся числа $\alpha_1, \alpha_2, \ldots, \alpha_l$, *не все* из которых равны нулю, для которых верно равенство $\alpha_1 \boldsymbol{a_1} + \alpha_2 \boldsymbol{a_2} + \ldots + \alpha_l \boldsymbol{a_l} = \mathbf{0}$.

Пример 1. Рассмотрим совокупность трёх векторов $\boldsymbol{a_1} = (1, -2, 1, 3)$, $\boldsymbol{a_2} = (0, 0, 1, 1)$, $\boldsymbol{a_3} = (0, 0, 0, 2)$. Являются ли они линейно зависимыми? Для ответа составим линейную комбинацию этих векторов и выясним, при каких коэффициентах комбинация равна нулевому вектору. Из определения операций сложения векторов и умножения вектора на число следует, что

$$\alpha_1 \boldsymbol{a_1} + \alpha_2 \boldsymbol{a_2} + \alpha_3 \boldsymbol{a_3} = (\alpha_1, -2\alpha_1, \alpha_1 + \alpha_2, 3\alpha_1 + \alpha_2 + 2\alpha_3).$$

Из равенства этого вектора нулевому следует равенство нулю каждой его координаты. Отсюда без труда получаем, что все числа α_i равны нулю.

Итак, линейная комбинация данных векторов равна нулевому вектору *только* при нулевых коэффициентах. Следовательно векторы линейно независимы.

Пример 2. Векторы $\boldsymbol{a} = (1, 2, 0)$, $\boldsymbol{b} = (-2, -3, 1)$, $\boldsymbol{c} = (2, 0, -1)$ линейно зависимые, поскольку, например, $3\boldsymbol{a} + 2\boldsymbol{b} + \frac{1}{2}\boldsymbol{c} = \mathbf{0}$.

Пример 3. Рассмотрим систему, состоящую из одного вектора $\boldsymbol{a}$. Если этот вектор нулевой, то рассматривая система линейно зависима: $1 \cdot \boldsymbol{a} = 1 \cdot \mathbf{0} = \mathbf{0}$. В случае, когда вектор $\boldsymbol{a} \neq \mathbf{0}$, одна из его компонент a_i отлична от нуля. Поэтому из равенства $\alpha \boldsymbol{a} = \mathbf{0}$ следует равенство $\alpha a_i = 0$ и $\alpha = 0$. Т. е. система линейно независима.

Пример 4. Система, состоящая из двух одинаковых векторов линейно зависима: $1 \cdot \boldsymbol{a} + (-1) \cdot \boldsymbol{a} = \mathbf{0}$.

Теорема 1. Система векторов $\boldsymbol{a_1}, \boldsymbol{a_2}, \ldots, \boldsymbol{a_l}$ линейно зависима тогда и только тогда, когда один из векторов системы является линейной комбинацией других.

Доказательство.

Необходимость. Пусть система векторов $\boldsymbol{a_1}, \boldsymbol{a_2}, \ldots, \boldsymbol{a_l}$ линейно зависима. Это значит, что существуют числа $\alpha_1, \alpha_2, \ldots, \alpha_l$, для которых верно равенство $\alpha_1 \boldsymbol{a_1} + \alpha_2 \boldsymbol{a_2} + \ldots + \alpha_l \boldsymbol{a_l} = \mathbf{0}$ и хотя бы одно из них отлично от нуля. Перенумеруем, в случае необходимости, векторы так, чтобы $\alpha_1 \neq 0$. Перенеся в последнем равенстве все слагаемые, кроме первого, в правую часть и разделив обе части на α_1, получим, что

$$\boldsymbol{a_1} = -\frac{\alpha_2}{\alpha_1}\boldsymbol{a_2} - \frac{\alpha_3}{\alpha_1}\boldsymbol{a_3} - \ldots - \frac{\alpha_l}{\alpha_1}\boldsymbol{a_l},$$

т. е. $\boldsymbol{a}_1$ является линейной комбинацией векторов $\boldsymbol{a}_2, \boldsymbol{a}_3, \ldots, \boldsymbol{a}_l$.

<u>Достаточность.</u> Пусть один из векторов (без ограничения общности можно считать, что это $\boldsymbol{a}_1$) является линейной комбинацией других векторов: $\boldsymbol{a}_1 = \beta_2 \boldsymbol{a}_2 + \beta_2 \boldsymbol{a}_3 + \ldots + \beta_l \boldsymbol{a}_l$ для некоторых $\beta_2, \beta_3, \ldots, \beta_l$. Линейная зависимость векторов $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_l$ становится очевидной, если переписать последнее равенство в виде

$$1 \cdot \boldsymbol{a}_1 - \beta_2 \cdot \boldsymbol{a}_2 - \beta_3 \cdot \boldsymbol{a}_3 - \ldots - \beta_l \cdot \boldsymbol{a}_l = \mathbf{0},$$

поскольку коэффициент при $\boldsymbol{a}_1$ заведомо отличен от нуля. □

Следствие. Система из двух векторов линейно зависима тогда и только тогда, когда эти векторы пропорциональны.

Теорема 2. Если система векторов $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_l$ содержит линейно зависимую подсистему, то и вся система линейно зависима.

Доказательство.

Перенумеруем, в случае необходимости, данные векторы так, чтобы линейно зависимыми оказались первые r векторов. Тогда имеет место равенство

$$\alpha_1 \boldsymbol{a}_1 + \alpha_2 \boldsymbol{a}_2 + \ldots + \alpha_r \boldsymbol{a}_r = \mathbf{0},$$

в котором не все коэффициенты нулевые. Поэтому последнее равенство, переписанное в виде

$$\alpha_1 \cdot \boldsymbol{a}_1 + \alpha_2 \cdot \boldsymbol{a}_2 + \ldots + \alpha_r \cdot \boldsymbol{a}_r + 0 \cdot \boldsymbol{a}_{r+1} + \ldots + 0 \cdot \boldsymbol{a}_l = \mathbf{0}$$

и означает линейную зависимость всей системы. □

Следствие 1. Система векторов, содержащая нулевой вектор, является линейно зависимой.

Следствие 2. Система векторов, содержащая два равных вектора, является линейно зависимой.

Следствие 3. Если система векторов линейно независима, то и любая её подсистема линейно независима.

Понятие линейной зависимости можно распространить на системы, состоящие из бесконечного числа векторов. По определению, *бесконечная* система называется линейно зависимой, если существует конечная линейно зависимая подсистема.

Определение. *Базисом* данной (возможно бесконечной) системы векторов называется всякая линейно независимая подсистема, добавление к которой любого другого вектора системы превращает её в линейно зависимую.

Пример 5. Для системы векторов $\boldsymbol{a}_1 = (2, 1, 1), \boldsymbol{a}_2 = (-4, 3, 1), \boldsymbol{a}_3 = (-1, 2, 1)$ в качестве базиса можно выбрать векторы $\boldsymbol{a}_1$ и $\boldsymbol{a}_2$, так как они линейно независимы (см. следствие теоремы 1) , а вся система линейно зависима:

$$\boldsymbol{a}_1+\boldsymbol{a}_2-2\boldsymbol{a}_3=\boldsymbol{0}.$$

Замечание. При задании базиса, учитывается *порядок* базисных векторов. Так, в рассмотренном примере базисы, состоящие из $\boldsymbol{a}_1, \boldsymbol{a}_2$ и $\boldsymbol{a}_2, \boldsymbol{a}_1$ считаются различными.

Пример 6. Векторы $\boldsymbol{e}_1=(1,0,\ldots,0), \boldsymbol{e}_2=(0,1,\ldots,0),\ldots,\boldsymbol{e}_n=(0,0,\ldots,1)$ линейно независимы и для всякого $\boldsymbol{x}=(x_1,x_2,\ldots,x_n)$ справедливо равенство

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

которое означает, согласно теореме 1, линейную зависимость системы векторов $\boldsymbol{x},\boldsymbol{e}_1,\boldsymbol{e}_2,\ldots,\boldsymbol{e}_n$. Таким образом, векторы $\boldsymbol{e}_1,\boldsymbol{e}_2,\ldots,\boldsymbol{e}_n$ образуют базис всего пространства $\mathbb{R}^n$. Данный базис называется *стандартным* или *каноническим.*

В следующем параграфе мы покажем, что всякая система векторов из $\mathbb{R}^n$ содержит *конечный* базис и что любые два базиса системы содержат одинаковое количество векторов.

Теорема 3. Пусть $\boldsymbol{e}_1,\boldsymbol{e}_2,\ldots,\boldsymbol{e}_r$ — базис некоторой системы векторов. Тогда всякий вектор системы единственным образом представляется в виде линейной комбинации базисных векторов.

Доказательство.

Пусть $\boldsymbol{x}$ — произвольный вектор системы. По определению базиса, система векторов $\boldsymbol{x},\boldsymbol{e}_1,\boldsymbol{e}_2,\ldots,\boldsymbol{e}_r$ линейно зависима, т. е. имеет место равенство

$$\alpha\boldsymbol{x}+\alpha_1\boldsymbol{e}_1+\alpha_2\boldsymbol{e}_2+\ldots+\alpha_r\boldsymbol{e}_r=\boldsymbol{0},$$

где не все коэффициенты равны нулю. Коэффициент $\alpha\neq0$, иначе из линейной независимости векторов $\boldsymbol{e}_i$ следовало бы, что и все $\alpha_i=0$. Выражая из предпоследнего равенства $\boldsymbol{x}$, находим его искомое представление в виде линейной комбинации $\boldsymbol{e}_1,\boldsymbol{e}_2,\ldots,\boldsymbol{e}_r$:

$$\boldsymbol{x}=-\frac{\alpha_1}{\alpha}\boldsymbol{e}_1-\frac{\alpha_2}{\alpha}\boldsymbol{e}_2-\ldots-\frac{\alpha_r}{\alpha}\boldsymbol{e}_r.$$

Докажем теперь, что такое представление единственно. Допустим, предположив противное, что имеют места два соотношения

$$\boldsymbol{x}=\beta_1\boldsymbol{e}_1+\beta_2\boldsymbol{e}_2+\ldots+\beta_r\boldsymbol{e}_r,$$
$$\boldsymbol{x}=\gamma_1\boldsymbol{e}_1+\gamma_2\boldsymbol{e}_2+\ldots+\gamma_r\boldsymbol{e}_r.$$

Вычитая почленно эти равенства, получаем $\boldsymbol{0}=(\beta_1-\gamma_1)\boldsymbol{e}_1+\ldots+(\beta_r-\gamma_r)\boldsymbol{e}_r$. Из линейной независимости векторов $\boldsymbol{e}_1,\boldsymbol{e}_2,\ldots,\boldsymbol{e}_r$ следует, что $\beta_i=\gamma_i$. □

Отметим, что указанное в теореме представление вектора в виде линейной комбинации базисных часто называют *разложением* этого вектора по базису.

§10. Существование и конечность базиса

Пусть S — произвольная система векторов пространства $\mathbb{R}^n$. Мы хотим показать, что S содержит подсистему, являющуюся базисом и что любые два базиса состоят из одного и того же числа векторов. Для этого сначала докажем следующую

Теорема. Предположим, что каждый из векторов $\boldsymbol{a_1}, \boldsymbol{a_2}, \ldots, \boldsymbol{a_l}$ является линейной комбинацией векторов $\boldsymbol{b_1}, \boldsymbol{b_2}, \ldots, \boldsymbol{b_k}$. Если векторы $\boldsymbol{a_1}, \boldsymbol{a_2}, \ldots, \boldsymbol{a_l}$ линейно независимы, то $l \leqslant k$.

Доказательство проведём методом математической индукции (см. Добавление) по параметру k.

<u>База индукции.</u> При $k=1$ при некоторых числах $\alpha_i, i=1,\ldots l$ справедливы равенства $\boldsymbol{a_i}=\alpha_i \boldsymbol{b_1}$. Следствие 1 теоремы 2 предыдущего параграфа гарантирует, что все $\alpha_i \neq 0$. Если $l>1$, то имеем, например, $\boldsymbol{a_1}=\frac{\alpha_1}{\alpha_2}\boldsymbol{a_2}$, что противоречит линейной независимости векторов $\boldsymbol{a_i}$. Поэтому $l=1$ и в рассматриваемом случае теорема доказана.

<u>Индукционный шаг.</u> Предположим, что предложение верно при $k=m-1$. Докажем его справедливость при $k=m$.

По условию, имеют место равенства

$$\begin{aligned}
\boldsymbol{a_1} &= \alpha_{11}\boldsymbol{b_1} + \alpha_{12}\boldsymbol{b_2} + \ldots + \alpha_{1,m-1}\boldsymbol{b_{m-1}} + \alpha_{1m}\boldsymbol{b_m},\\
\boldsymbol{a_2} &= \alpha_{21}\boldsymbol{b_1} + \alpha_{22}\boldsymbol{b_2} + \ldots + \alpha_{2,m-1}\boldsymbol{b_{m-1}} + \alpha_{2m}\boldsymbol{b_m},\\
&\vdots\\
\boldsymbol{a_{l-1}} &= \alpha_{l-1,1}\boldsymbol{b_1} + \alpha_{l-1,2}\boldsymbol{b_2} + \ldots + \alpha_{l-1,m-1}\boldsymbol{b_{m-1}} + \alpha_{l-1,m}\boldsymbol{b_m},\\
\boldsymbol{a_l} &= \alpha_{l1}\boldsymbol{b_1} + \alpha_{l2}\boldsymbol{b_2} + \ldots + \alpha_{l,m-1}\boldsymbol{b_{m-1}} + \alpha_{lm}\boldsymbol{b_m}.
\end{aligned}$$

Если коэффициенты $\alpha_{1m}=\alpha_{2m}=\ldots=\alpha_{mm}=0$, то фактически векторы $\boldsymbol{a}$ раскладываются по $m-1$ вектору $\boldsymbol{b}$ и по предположению индукции справедливо неравенство $l \leqslant m-1$. Значит $l<m$ и в этом случае теорема доказана.

Теперь рассмотрим случай, когда один из коэффициентов при $\boldsymbol{b_m}$ отличен от нуля. Без ограничения общности можно считать, что $\alpha_{lm} \neq 0$ (этого всегда можно добиться, перенумеровав подходящим образом векторы $\boldsymbol{a_i}$).

Преобразуем выписанную систему равенств в эквивалентную таким образом, чтобы во всех уравнениях, кроме последнего, коэффициенты при $\boldsymbol{b_m}$ были равны нулю. Для этого к i-у $(i=1,2,\ldots,l-1)$ равенству прибавим l-е, умноженное на $-\frac{\alpha_{im}}{\alpha_{lm}}$. В итоге получим

$$
\begin{aligned}
&\boldsymbol{a}_1 - \frac{\alpha_{1m}}{\alpha_{lm}}\boldsymbol{a}_l = \beta_{11}\boldsymbol{b}_1 + \beta_{12}\boldsymbol{b}_2 + \ldots + \beta_{1,m-1}\boldsymbol{b}_{m-1},\\
&\boldsymbol{a}_2 - \frac{\alpha_{2m}}{\alpha_{lm}}\boldsymbol{a}_l = \beta_{21}\boldsymbol{b}_1 + \beta_{22}\boldsymbol{b}_2 + \ldots + \beta_{2,m-1}\boldsymbol{b}_{m-1},\\
&\vdots\\
&\boldsymbol{a}_{l-1} - \frac{\alpha_{l-1,m}}{\alpha_{lm}}\boldsymbol{a}_l = \beta_{l-1,1}\boldsymbol{b}_1 + \beta_{l-1,2}\boldsymbol{b}_2 + \ldots + \beta_{l-1,m-1}\boldsymbol{b}_{m-1},\\
&\boldsymbol{a}_l = \alpha_{l1}\boldsymbol{b}_1 + \alpha_{m2}\boldsymbol{b}_2 + \ldots + \alpha_{l,m-1}\boldsymbol{b}_{m-1} + \alpha_{lm}\boldsymbol{b}_m,
\end{aligned}
$$

где $\beta_{ij}=\alpha_{ij}-\frac{\alpha_{im}}{\alpha_{lm}}\alpha_{lj}$.

Таким образом, каждый из векторов $\hat{\boldsymbol{a}}_i=\boldsymbol{a}_i-\frac{\alpha_{im}}{\alpha_{lm}}\boldsymbol{a}_l$, $(i=1,2,\ldots,l-1)$ линейно выражается через $m-1$ вектор $\boldsymbol{b}_1,\boldsymbol{b}_2,\ldots,\boldsymbol{b}_{m-1}$. С другой стороны, векторы $\hat{\boldsymbol{a}}_i$ линейно независимы:

$$
\mu_1\hat{\boldsymbol{a}}_1+\mu_2\hat{\boldsymbol{a}}_2+\ldots+\mu_{l-1}\hat{\boldsymbol{a}}_{l-1}=\mathbf{0} \Leftrightarrow \mu_1\boldsymbol{a}_1+\mu_2\boldsymbol{a}_2+\ldots+\mu_{l-1}\boldsymbol{a}_{l-1}-\left(\mu_1\frac{\alpha_{1m}}{\alpha_{lm}}+\mu_2\frac{\alpha_{2m}}{\alpha_{lm}}+\ldots+\right.
$$
$$
\left.+\mu_{l-1}\frac{\alpha_{l-1,m}}{\alpha_{lm}}\right)\boldsymbol{a}_l=\mathbf{0} \Leftrightarrow \mu_1=\mu_2=\ldots=\mu_{m-1}=0.
$$

Последний вывод следует из линейной независимости векторов $\boldsymbol{a}_1,\boldsymbol{a}_2,\ldots,\boldsymbol{a}_l$. По предположению индукции справедливо неравенство $l-1\leqslant m-1$, эквивалентное неравенству $l\leqslant m$. □

Следствие 1. Если S — система векторов пространства $\mathbb{R}^n$, то всякая её линейно независимая подсистема состоит не более чем из n векторов.

Доказательство.

Пусть S_0 — линейно независимая подсистема системы S. В примере 6 предыдущего параграфа показано, что каждый вектор из $\mathbb{R}^n$ (а, следовательно, и из S_0) раскладывается по n векторам канонического базиса. Поэтому S_0 состоит не более чем из n векторов. □

Теперь мы в состоянии доказать *существование* базиса. Согласно доказанному следствию, в S существует подсистема $\tilde{S}$ состоящая из наибольшего числа линейно независимых векторов. Очевидно, что эта подсистема и является базисом.

Следствие 2. Любые два базиса системы S состоят из одного и того же числа векторов.

Доказательство.

Пусть $\boldsymbol{a}_1,\boldsymbol{a}_2,\ldots,\boldsymbol{a}_l$ и $\boldsymbol{b}_1,\boldsymbol{b}_2,\ldots,\boldsymbol{b}_k$ — два базиса в S. По теореме 3 предыдущего параграфа каждый вектор $\boldsymbol{a}_i$ раскладывается по векторам $\boldsymbol{b}_j$. Из доказанной теоремы следует равенство $l\leqslant k$. Но векторы $\boldsymbol{b}_i$ также раскладываются по векторам $\boldsymbol{a}_j$ и поэтому $l\geqslant k$. Объединяя полученные равенства, приходим к выводу, что $l=k$. □

§11. Ранг матрицы. Ступенчатая матрица

Пусть A — произвольная матрица типа $m\times n$

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

Строчечный ранг r_A^R матрицы A — это максимальное количество её линейно независимых строк, рассматриваемых как элементы пространства $\mathbb{R}^n$.

Столбцовый ранг r_A^C матрицы A — это максимальное количество линейно независимых столбцов, рассматриваемых как элементы пространства $\mathbb{R}^m$.

В этом параграфе будет доказано, что $r_A^R=r_A^C$. Действовать мы будем по следующей схеме. Сначала введём так называемые элементарные преобразования строк матрицы и установим, что при этих преобразованиях ни r_A^R, ни r_A^C не меняются. Далее, мы покажем, что элементарными преобразованиями всякую матрицу можно привести к специальному — ступенчатому — виду, в котором система ненулевых строк линейно независимая. И, наконец, будет доказано, что максимальное количество линейно независимых столбцов в матрице ступенчатого вида равно количеству её ненулевых строк.

Определение. *Элементарными преобразованиями* строк матрицы называются преобразования следующих типов:

1) умножение строки на ненулевое число;
2) перестановка двух строк;
3) прибавление к одной строке другой, умноженной на число.

Теорема 1. Строчечный ранг матрицы не меняется при элементарных преобразованиях строк.

Доказательство.

Обозначим строки матрицы A через $\boldsymbol{a}_{1*}, \boldsymbol{a}_{2*}, \dots, \boldsymbol{a}_{m*}$ и пусть строки $\boldsymbol{a}_{i_1*}$, $\boldsymbol{a}_{i_2*}, \dots, \boldsymbol{a}_{i_r*}$ образуют базис в системе строк. Нужно показать, что после каждого элементарного преобразования базис преобразованной системы по прежнему состоит из r элементов.

Преобразование 1). Пусть на ненулевое число λ умножается строка $\boldsymbol{a}_{i*}$, не входящая в набор $\boldsymbol{a}_{i_1*}, \boldsymbol{a}_{i_2*}, \dots, \boldsymbol{a}_{i_r*}$. Так как строка $\boldsymbol{a}_{i*}$ может быть разложена по базису, то это же верно и для строки $\lambda\boldsymbol{a}_{i*}$:

$$\boldsymbol{a}_{i*}=\alpha_1\boldsymbol{a}_{i_1*}+\alpha_2\boldsymbol{a}_{i_2*}+\ldots+\alpha_r\boldsymbol{a}_{i_r*} \Rightarrow \lambda\boldsymbol{a}_{i*}=\lambda\alpha_1\boldsymbol{a}_{i_1*}+\lambda\alpha_2\boldsymbol{a}_{i_2*}+\ldots+\lambda\alpha_r\boldsymbol{a}_{i_r*}.$$

Таким образом, строки $\boldsymbol{a}_{i_1*}, \boldsymbol{a}_{i_2*}, \ldots, \boldsymbol{a}_{i_r*}$ образуют базис и для преобразованной системы строк.

Если на ненулевое число λ умножается первая строка (для других строк рассуждения аналогичны) из набора $\boldsymbol{a}_{i_1*}, \boldsymbol{a}_{i_2*}, \ldots, \boldsymbol{a}_{i_r*},$ то в качестве базиса преобразованной системы можно выбрать строки $\lambda \boldsymbol{a}_{i_1*}, \boldsymbol{a}_{i_2*}, \ldots, \boldsymbol{a}_{i_r*}$. В самом деле, так как

$$\alpha_1(\lambda \boldsymbol{a}_{i_1*}) + \alpha_2 \boldsymbol{a}_{i_2*} + \ldots + \alpha_r \boldsymbol{a}_{i_r*} = \boldsymbol{0} \Leftrightarrow (\alpha_1 \lambda)\boldsymbol{a}_{i_1*} + \alpha_2 \boldsymbol{a}_{i_2*} + \ldots + \alpha_r \boldsymbol{a}_{i_r*} = \boldsymbol{0} \Leftrightarrow$$
$$\Leftrightarrow \alpha_1 \lambda = \alpha_2 = \ldots = \alpha_r = 0 \overset{\lambda \neq 0}{\Leftrightarrow} \alpha_1 = \alpha_2 = \ldots = \alpha_r = 0,$$

строки $\lambda \boldsymbol{a}_{i_1*}, \boldsymbol{a}_{i_2*}, \ldots, \boldsymbol{a}_{i_r*}$ линейно независимы. И поскольку для i-й строки $(i = 1, 2, \ldots, n)$ имеет место разложение $\boldsymbol{a}_{i*} = \alpha_1 \boldsymbol{a}_{i_1*} + \alpha_2 \boldsymbol{a}_{i_2*} + \ldots + \alpha_r \boldsymbol{a}_{i_r*}$ по строкам $\boldsymbol{a}_{i_1*}, \boldsymbol{a}_{i_2*}, \ldots, \boldsymbol{a}_{i_r*},$ то

$$\boldsymbol{a}_{i*} = \frac{\alpha_1}{\lambda}(\lambda \boldsymbol{a}_{i_1*}) + \alpha_2 \boldsymbol{a}_{i_2*} + \ldots + \alpha_r \boldsymbol{a}_{i_r*} \text{ —}$$

разложение $\boldsymbol{a}_{i*}$ по строкам $\lambda \boldsymbol{a}_{i_1*}, \boldsymbol{a}_{i_2*}, \ldots, \boldsymbol{a}_{i_r*}$.

<u>Преобразование 2).</u> В этом случае теорема очевидна, поскольку понятие линейной зависимости не зависит от порядка следования строк.

<u>Преобразование 3).</u> Пусть к i-й строке прибавлена j-я, умноженная на μ. Если строка $\boldsymbol{a}_{i*}$ не принадлежит базису $\boldsymbol{a}_{i_1*}, \boldsymbol{a}_{i_2*}, \ldots, \boldsymbol{a}_{i_r*}$ исходной системы, то строки $\boldsymbol{a}_{i_1*}, \boldsymbol{a}_{i_2*}, \ldots, \boldsymbol{a}_{i_r*}$ образуют базис и в преобразованной системе. Если же j-я строка, умноженная на μ, прибавляется к $\boldsymbol{a}_{i_1*}$ (для других строк рассуждения аналогичны), то в преобразованной системе в качестве базиса можно выбрать строки $\boldsymbol{a}_{i_1*} + \mu \boldsymbol{a}_{j*}, \boldsymbol{a}_{i_2*}, \ldots, \boldsymbol{a}_{i_r*}$.

Обоснование сформулированных утверждений проводится как и в случае преобразования 1). Недостающие детали предлагается восполнить читателю. □

Теорема 2. Столбцовый ранг матрицы не меняется при элементарных преобразованиях строк.

Доказательство.

Обозначим столбцы матрицы A через $\boldsymbol{a}_{*1}, \boldsymbol{a}_{*2}, \ldots, \boldsymbol{a}_{*n}$ и пусть столбцы $\boldsymbol{a}_{*i_1}, \boldsymbol{a}_{*i_2}, \ldots, \boldsymbol{a}_{*i_s}$ образуют базис в системе столбцов. Нужно показать, что после каждого элементарного преобразования базис преобразованной системы по прежнему состоит из s элементов.

Заметим, прежде всего, что однородная система линейных уравнений

$$
\begin{cases}
a_{1i_1}x_1+a_{1i_2}x_2+\ldots+a_{1i_s}x_s=0\,, \\
a_{2i_1}x_1+a_{2i_2}x_2+\ldots+a_{2i_s}x_s=0\,, \\
\ldots \\
a_{mi_1}x_1+a_{mi_2}x_2+\ldots+a_{mi_s}x_s=0\,.
\end{cases} \tag{7}
$$

является развёрнутой формой записи уравнения $x_1\boldsymbol{a}_{*i_1}+x_2\boldsymbol{a}_{*i_2}+\ldots+x_s\boldsymbol{a}_{*i_s}=\boldsymbol{0}$. Линейная независимость столбцов $\boldsymbol{a}_{*i_1},\boldsymbol{a}_{*i_2},\ldots,\boldsymbol{a}_{*i_s}$ означает, что система (7) имеет только нулевое решение.

Каждому элементарному преобразованию строк соответствует преобразование системы (7), переводящее её, согласно теореме из §2, в эквивалентную. Поэтому столбцы, полученные из $\boldsymbol{a}_{*i_1},\boldsymbol{a}_{*i_2},\ldots,\boldsymbol{a}_{*i_s}$ элементарными преобразований строк, являются линейно независимыми.

Точно такими же рассуждениями показывается, что всякая линейно зависимая система столбцов переводится в линейно зависимую. Всё вместе это означает, что базис $\boldsymbol{a}_{*i_1},\boldsymbol{a}_{*i_2},\ldots,\boldsymbol{a}_{*i_s}$ переводится элементарными преобразованиями в базис. □

Определение. Матрица A называется матрицей *ступенчатого* вида, если первый ненулевой элемент каждой строки, начиная со второй, расположен правее первого ненулевого элемента предыдущей строки.

Из определения следует, что если в матрице ступенчатого вида имеются нулевые строки, то все они расположены под ненулевыми.

Пример 1. Матрица $\begin{pmatrix} 0 & -3 & 2 & 0 & 5 \\ 0 & 0 & 4 & 1 & -2 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ имеет ступенчатый вид: первый ненулевой элемент второй строки $a_{23}=4$ расположен правее первого ненулевого элемента первой строки $a_{12}=-3$; первый ненулевой элемент третьей строки $a_{35}=1$ расположен правее первого ненулевого элемента второй строки $a_{23}=4$; в четвёртой строке нет ненулевых элементов.

Покажем, что всякая матрица элементарными преобразованиями может быть приведена к ступенчатому виду за конечное число шагов.

Нулевая матрица по определению имеет ступенчатый вид. Если матрица ненулевая, то в ней имеется хотя бы один ненулевой столбец. Пусть первый из них имеет номер i_1. Перестановкой строк можно добиться того, чтобы $a_{1i_1}\neq 0$. Элементарное преобразование 3 позволяет добиться равенств $a_{2i_1}=a_{3i_1}=\ldots=0$. Для этого нужно прибавить к каждой строке, начиная со второй, первую строку, умноженную на подходящее число. В итоге получим матрицу вида

$$\left(\begin{array}{c|c} 0 \quad \dots \quad 0 & a_{1i_1} \quad \dots\dots\dots\dots\dots\dots \\ \hline \mathbf{0} & \boxed{A_1} \end{array}\right)$$

Поступая таким же образом с матрицей A_1, мы в конце концов получим матрицу ступенчатого вида.

Проиллюстрируем приведённые рассуждения на конкретном примере.

Пример 2. Приведём к ступенчатому виду матрицу

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

Прибавляя ко 2-й, 3-й и 4-й строкам 1-ю строку, умноженную на числа $-1, -2$ и -2 соответственно, получаем матрицу

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

Далее, прибавляя к 3-й и 4-й строкам 2-ю строку, умноженную на 3 и 4 соответственно, получаем матрицу

$$\begin{pmatrix} 1 & 2 & 1 & 0 & 2 \\ 0 & 1 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 5 \end{pmatrix}.$$

Наконец, переставляя 3-ю и 4-ю строки, получаем матрицу ступенчатого вида

$$\begin{pmatrix} 1 & 2 & 1 & 0 & 2 \\ 0 & 1 & 1 & -1 & 2 \\ 0 & 0 & 0 & -1 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Теорема 3. Система ненулевых строк матрицы ступенчатого вида линейно независима.

Доказательство.

Ступенчатая матрица, как это следует из её определения может быть записана в виде

$$\begin{pmatrix} a_{1i_1}\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \qquad a_{2i_2}\ldots\ldots\ldots\ldots\ldots \\ \qquad\qquad \ldots\ldots\ldots\ldots \\ \qquad\qquad\qquad a_{ri_r}\ldots \\ \\ \end{pmatrix}, \tag{8}$$

где элементы[4] $a_{1i_1}, a_{2i_2}, \ldots, a_{ri_r}$ отличны от нуля, а все элементы, находящиеся слева от них и ниже них, равны нулю. При этом $i_1 < i_2 < \ldots < i_r$.

Как и при доказательстве теоремы 1, обозначим через $\boldsymbol{a}_{k*}$ k-ю строку матрицы. Рассмотрим векторное равенство $\alpha_1 \boldsymbol{a}_{1*} + \alpha_2 \boldsymbol{a}_{2*} + \ldots + \alpha_r \boldsymbol{a}_{r*} = \boldsymbol{0}$ и выпишем явно равенства компонент с номерами $i_1, i_2, \ldots, i_r$:

$$\begin{aligned} &\alpha_1 a_{1i_1} = 0, \\ &\alpha_1 a_{1i_1} + \alpha_2 a_{2i_2} = 0, \\ &\ldots \\ &\alpha_1 a_{1i_1} + \alpha_2 a_{2i_2} + \ldots + \alpha_r a_{ri_r} = 0, \end{aligned}$$

из которых следует, что $\alpha_1 = \alpha_2 = \ldots = \alpha_r = 0$, что и доказывает теорему. □

Очевидно, что ненулевые строки ступенчатой матрицы образуют максимальную систему линейно независимых строк. Другими словами, строчечный ранг матрицы (8) равен r.

Теорема 4. Столбцовый ранг ступенчатой матрицы равен числу её ненулевых строк.

Доказательство.

Пусть матрица имеет вид (8). Тривиально проверяется, что r столбцов с номерами $i_1, i_2, \ldots, i_r$ линейно независимы. Докажем, что эти столбцы образуют максимальную линейно независимую подсистему. Для этого достаточно показать (см. §9, теорема 1), что всякий столбец $\boldsymbol{a}_{*i}$ матрицы линейно выражается через столбцы $\boldsymbol{a}_{*i_1}, \boldsymbol{a}_{*i_2}, \ldots, \boldsymbol{a}_{*i_r}$. Как и при доказательстве теоремы 2, замечаем, что векторное уравнение $\alpha_1 \boldsymbol{a}_{*i_1} + \alpha_2 \boldsymbol{a}_{*i_2} + \ldots + \alpha_r \boldsymbol{a}_{*i_r} = \boldsymbol{a}_{*i}$ с неизвестными коэффициентами $\alpha_1, \alpha_2, \ldots, \alpha_r$ эквивалентно системе линейных уравнений

4 Эти элементы матрицы ступенчатого вида называются *ведущими.*

$$\begin{cases} \alpha_1 a_{1i_1}+\alpha_2 a_{1i_2}+\ldots+\alpha_r a_{1i_r}=a_{1i}, \\ \qquad\quad \alpha_2 a_{2i_2}+\ldots+\alpha_r a_{2i_r}=a_{2i}, \\ \qquad\qquad\qquad \ldots \\ \qquad\qquad\qquad\qquad \alpha_r a_{ri_r}=a_{ri}. \end{cases}$$

Так как числа $a_{1i_1}, a_{2i_2}, \ldots, a_{ri_r}$ отличны от нуля, то выписанная система очевидно совместная. Следовательно столбец $\boldsymbol{a}_{*i}$ является линейной комбинацией столбцов $\boldsymbol{a}_{*i_1}, \boldsymbol{a}_{*i_2}, \ldots, \boldsymbol{a}_{*i_r}$. □

Итак, всякая матрица элементарными преобразованиями строк может быть приведена к ступенчатой форме. Строчечный и столбцовый ранги ступенчатой матрицы равны между собой. А поскольку при элементарных преобразованиях ни строчечный, ни столбцовый ранг не меняется, то их равенство справедливо для произвольной матрицы.

Ранг матрицы A обозначается символом $\operatorname{rank} A$. Так, например, для матрицы A из примера 2, $\operatorname{rank} A=3$.

§12. Метод Гаусса[ii] решения систем линейных уравнений

Метод Гаусса решения систем линейных уравнений заключается в приведении данной системы линейных уравнений к *эквивалентной системе,* расширенная матрица которой имеет ступенчатый вид.

Пусть $(A|B)$ — расширенная матрица системы (1). Приведём её к ступенчатому виду и полученную матрицу обозначим через $(T\,|\,D)$. Очевидно, что элементарные преобразования строк матрицы $(A|B)$ равносильны преобразованиям уравнений системы (1), описанным в формулировке теоремы §2. Согласно упомянутой теореме, система линейных уравнений, матрица которой совпадает с $(T\,|\,D)$, эквивалентна системе (1).

Пусть $\operatorname{rank} A=r$, а $\operatorname{rank}(A\,|\,B)=\tilde{r}$. Если столбец B свободных членов линейно выражается через столбцы матрицы A, то $\tilde{r}=r$. В противном случае $\tilde{r}=r+1$. Соотношения между числами r и $\tilde{r}$ следуют также из вида матрицы $(T\,|\,D)$, которая, напомним, имеет ступенчатый вид.

Возможны следующие три принципиально различных случая.

1. $\tilde{r}=r+1$. В этом случае система, матрица которой имеет ступенчатый вид, содержит уравнение

$$0\cdot x_1+0\cdot x_2+\ldots+0\cdot x_n=d,$$

где $d\neq 0$, и, следовательно, несовместна. По теореме из §2 исходная система также несовместна.

2. $\tilde{r}=r=n$. В этом случае, после отбрасывания нулевых уравнений получается так называемая *строго треугольная* система

$$\begin{cases} t_{11}x_1+t_{12}x_2+t_{13}x_3+\ldots+t_{1n}x_n=d_1, \\ \qquad t_{22}x_2+t_{23}x_3+\ldots+t_{2n}x_n=d_2, \\ \qquad\qquad t_{33}x_3+\ldots+t_{3n}x_n=d_3, \\ \qquad\qquad\qquad \ldots \\ \qquad\qquad\qquad\qquad t_{nn}x_n=d_n, \end{cases}$$

где коэффициенты $t_{11}, t_{22}, \ldots, t_{nn}$ отличны от нуля. Из последнего уравнения этой системы однозначно определяется x_n, затем из предпоследнего уравнения определяется x_{n-1} и т. д. Следовательно треугольная система, а значит и система (1), имеет единственное решение.

3. $\tilde{r}=r<n$. Пусть в этом случае $t_{1i_1}, t_{2i_2}, \ldots, t_{ri_r}$ — ведущие элементы матрицы $(T|D)$. Неизвестные $x_{i_1}, x_{i_2}, \ldots, x_{i_r}$ назовём *главными,* а остальные — *свободными.* После отбрасывания нулевых уравнений и перенесения членов со свободными неизвестными в правую часть получается строго треугольная система относительно главных неизвестных. Решая её, как в предыдущем случае, найдём выражения главных неизвестных через свободные. Все решения системы получаются из этих выражений подстановкой каких-либо значений свободных неизвестных. Поскольку эти значения можно выбирать произвольно, то система имеет бесконечно много решений.

Пример. Решим систему уравнений

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

расширенной матрицей которой служит матрица из примера 2 предыдущего параграфа. Вычисления, проведённые в этом примере и теорема из §2 показывают, что данная система эквивалентна ступенчатой системе

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

Считая неизвестные x_1, x_2, x_4 главными, а неизвестное x_3 — свободным, перепишем систему в виде

$$\begin{cases} x_1+2x_2 \qquad = 2-x_3, \\ \qquad x_2- \ x_4 = 2-x_3, \\ \qquad\quad - \ x_4 = 5. \end{cases}$$

Придавая свободной неизвестной x_3 произвольное значение (обозначим его через c), получаем, что рассматриваемая система имеет бесконечно много решений, каждое из которых можно записать в виде

$$(8+c, -3-c, c, -5)$$

для подходящего значения c.

Из предыдущих рассуждений вытекают утверждения, позволяющие исследовать систему линейных уравнений на совместность и определённость, не решая самой системы.

Теорема 1. (Кронекер[iii], Капелли[iv]) Система линейных уравнений является совместной тогда и только тогда, когда ранг матрицы коэффициентов системы равен рангу расширенной матрицы.

Теорема 2. Совместная система линейных уравнений является определенной тогда и только тогда, когда ранг матрицы коэффициентов системы равен числу неизвестных.

§13. Однородные системы линейных уравнений

Напомним (см. §6), что система линейных уравнений называется однородной, если столбец свободных членов является нулевым. Всякая однородная система является совместной, поскольку она имеет нулевое решение. Если система определенна, то она имеет *только* нулевое решение, если неопределенна, то имеет бесконечно много решений. По теореме Кронекера — Капелли последнее имеет место лишь в том случае, когда $r<n$, где r — ранг матрицы коэффициентов, а n — число неизвестных. Допустим, что система состоит из m уравнений. Тогда из очевидного равенства $r\leqslant m$ вытекает следующая

Теорема 1. Всякая однородная система линейных уравнений, число уравнений которой меньше числа неизвестных, имеет ненулевой решение.

В следующей теореме описываются свойства решений (рассматриваемых как элементы пространства $\mathbb{R}^n$) однородной системы линейных уравнений от n неизвестных.

Теорема 2. Пусть $\boldsymbol{u}, \boldsymbol{v}$ — произвольные решения данной системы линейных однородных уравнений и α — произвольное число. Тогда элементы $\boldsymbol{u}+\boldsymbol{v}$ и $\alpha\boldsymbol{u}$ являются решениями той же системы.

Доказательство.

Пусть $\boldsymbol{u}=(u_1, u_2, \ldots, u_n)$ и $\boldsymbol{v}=(v_1, v_2, \ldots, v_n)$ являются решениями системы

$$\begin{cases} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=0\,, \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=0\,, \\ \ldots \\ a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n=0\,. \end{cases} \tag{9}$$

Это означает, что для всех $i=1,2,\ldots,m$ равенства

$$a_{i1}u_1+a_{i2}u_2+\ldots+a_{in}u_n=0,$$
$$a_{i1}v_1+a_{i2}v_2+\ldots+a_{in}v_n=0$$

являются верными. Поэтому верными будут и равенства

$$a_{i1}(\alpha u_1)+a_{i2}(\alpha u_2)+\ldots+a_{in}(\alpha u_n)=0,$$
$$a_{i1}(u_1+v_1)+a_{i2}(u_2+v_2)+\ldots+a_{in}(u_n+v_n)=0,$$

которые означают, что векторы $\boldsymbol{u}+\boldsymbol{v}$ и $\alpha\boldsymbol{u}$ также являются решениями рассматриваемой системы. □

Замечание. Теорема 2 означает, что на множестве решений однородной системы линейных уравнений корректно определены операции сложения и умножения на число. Непосредственно проверяется, что для этих операций справедливы все свойства, перечисленные в §8. Таким образом, множество решений однородной системы линейных уравнений образуют векторное пространство (см. замечание в §8) относительно указанных операций. По этой причине в дальнейшем мы будем иногда говорить не о множестве, а о *пространстве решений* однородной системы.

Предположим, что матрицы коэффициентов систем уравнений (1) и (9) совпадают. Если система (1) совместная, то множество её решений описывает

Теорема 3. Пусть $\hat{\boldsymbol{u}}$ — *фиксированное* решение неоднородной системы (1). Тогда множество *всех* её решений совпадает со множеством векторов из $\mathbb{R}^n$ вида $\hat{\boldsymbol{u}}+\boldsymbol{v}$, где $\boldsymbol{v}$ пробегает все решения однородной системы (9).

Доказательство.

Пусть $\hat{\boldsymbol{u}}=(\hat{u}_1,\hat{u}_2,\ldots,\hat{u}_n), \boldsymbol{v}=(v_1,v_2,\ldots,v_n)$. По условию теоремы всех i справедливы равенства

$$a_{i1}\hat{u}_1+a_{i2}\hat{u}_2+\ldots+a_{in}\hat{u}_n=b_i\,,$$
$$a_{i1}v_1+a_{i2}v_2+\ldots+a_{in}v_n=0\,,$$

из которых вытекает справедливость равенств

$$(\hat{u}_1+v_1)+a_{i2}(\hat{u}_2+v_2)+\ldots+a_{in}(\hat{u}_n+v_n)=b_i\,,$$

означающих, что вектор $\hat{\boldsymbol{u}}+\boldsymbol{v}$ является решением системы (1).

Предположим теперь, что $\boldsymbol{u}$ — произвольное решение системы (1) и покажем, что оно может быть представлено в виде суммы решения $\hat{\boldsymbol{u}}$ неоднородной системы и некоторого решения $\boldsymbol{v}$ однородной системы. Аналогично предыдущему устанавливается, что вектор $\boldsymbol{u}-\hat{\boldsymbol{u}}$ является решением системы (9). Обозначая его через $\boldsymbol{v}$, получаем искомое представление $\boldsymbol{u}=\hat{\boldsymbol{u}}+\boldsymbol{v}$. □

В заключение рассмотрим вопрос о базисе пространства решений однородной системы.

Теорема 4. Для заданной однородной системы линейных уравнений с n неизвестными и матрицей коэффициентов A найдутся такие $n-\operatorname{rank} A$ линейно независимых решений, что произвольное решение системы является их линейной комбинацией.

Доказательство.

Рассмотрим систему уравнений (9) и с помощью элементарных преобразований приведём её к ступенчатому виду. Полученная система тоже является однородной и содержит $r=\operatorname{rank} A$ ненулевых уравнений. Поэтому с точностью до перенумерации неизвестных её можно записать в виде

$$\begin{cases} x_1=c_{11}x_{r+1}+c_{12}x_{r+2}+\ldots+c_{1,n-r}x_n, \\ x_2=c_{21}x_{r+1}+c_{22}x_{r+2}+\ldots+c_{2,n-r}x_n, \\ \qquad\ldots \\ x_r=c_{r1}x_{r+1}+c_{r2}x_{r+2}+\ldots+c_{r,n-r}x_n. \end{cases} \tag{10}$$

Придавая по очереди одной из свободных неизвестных $x_{r+1}, x_{r+2}, \ldots, x_n$ значение 1, а остальным — значения 0, мы получим следующие решения системы (9):

$$\begin{gathered} \boldsymbol{u_1}=(c_{11}, c_{21}, \ldots, c_{r,1}, 1, 0, \ldots, 0), \\ \boldsymbol{u_2}=(c_{12}, c_{22}, \ldots, c_{r,2}, 0, 1, \ldots, 0), \\ \cdots \\ \boldsymbol{u_{n-r}}=(c_{1,n-r}, c_{2,n-r}, \ldots, c_{r,n-r}, 0, 0, \ldots, 1). \end{gathered}$$

Покажем, что они удовлетворяют всем перечисленным в теореме требованиям.

Для любых чисел $\lambda_1, \lambda_2, \ldots, \lambda_{n-r} \in \mathbb{R}$ линейная комбинация

$$\boldsymbol{u}=\lambda_1\boldsymbol{u_1}+\lambda_2\boldsymbol{u_2}+\ldots+\lambda_{n-r}\boldsymbol{u_{n-r}}$$

является решением (см. теорему 2) системы (9), в котором свободные неизвестные $x_{r+1}, x_{r+2}, \ldots, x_n$ равны $\lambda_1, \lambda_2, \ldots, \lambda_{n-r}$ соответственно. Заметим, что значения главных неизвестных *однозначно* определяются значениями свободных неизвестных по формулам (10). Поэтому, придавая числам λ_i все возможные значения, мы получим все возможные решения системы (9). Следовательно, каждое решение этой системы является комбинацией решений $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{n-r}$.

С другой стороны равенство $\boldsymbol{u}=\boldsymbol{0}$ выполняется только при $\lambda_i=0$, т. е. решения $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{n-r}$ линейно независимые. □

Всякий базис пространства решений однородной системы линейных уравнений называется *фундаментальной системой решений.* Доказательство предыдущей теоремы даёт практический способ построения фундаментальной системы решений.

§14. Линейные операторы

Определение. Отображение A из арифметического пространства $\mathbb{R}^n$ в пространство $\mathbb{R}^m$ называется *линейным оператором,* если выполняются следующие условия:

1. *Аддитивность*: $\forall\, \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n \quad A(\boldsymbol{x}+\boldsymbol{y})=A\boldsymbol{x}+A\boldsymbol{y}$.
2. *Однородность*: $\forall\, \boldsymbol{x} \in \mathbb{R}^n, \forall\, \alpha \in \mathbb{R} \quad A(\alpha \boldsymbol{x})=\alpha A\boldsymbol{x}$.

Заметим, что образ вектора $\boldsymbol{x}\in\mathbb{R}^n$ при действии линейного оператора A обозначается через $A\boldsymbol{x}$, а не $A(\boldsymbol{x})$. Это общепринятое соглашение будет использоваться всякий раз, когда отсутствие скобок не может привести к недоразумениям.

Пример 1. Рассмотрим отображение $A:\mathbb{R}^2\to\mathbb{R}$, определённое по правилу

$$\forall\, \boldsymbol{x}, \boldsymbol{x}=(x_1, x_2)\in\mathbb{R}^2 \quad A(x_1, x_2)=x_1+x_2.$$

Пусть $\boldsymbol{x}=(x_1, x_2), \boldsymbol{y}=(y_1, y_2)$ — произвольные векторы из $\mathbb{R}^2$. Поскольку $\boldsymbol{x}+\boldsymbol{y}=(x_1+y_1, x_2+y_2)$, то $A(\boldsymbol{x}+\boldsymbol{y})=(x_1+y_1)+(x_2+y_2)$, что, очевидно, совпадает с суммой $A\boldsymbol{x}+A\boldsymbol{y}=(x_1+x_2)+(y_1+y_2)$. Следовательно условие аддитивности для отображения A выполняется.

Точно также, из равенства $\alpha\boldsymbol{x}=(\alpha x_1, \alpha x_2)$ следует, что $A(\alpha\boldsymbol{x})=\alpha x_1+\alpha x_2$. Правая часть последнего соотношения равна с $\alpha(x_1+x_2)=\alpha A\boldsymbol{x}$. Значит условие однородности для отображения A также выполняется.

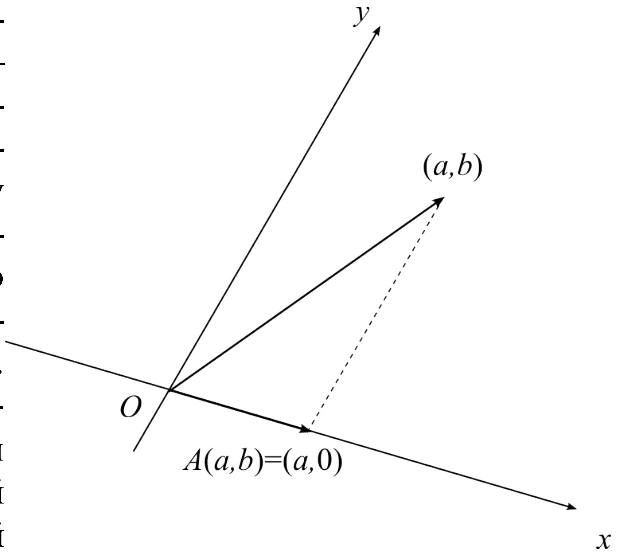
Таким образом, отображение A является линейным оператором.

Пример 2. Отображение $A:\mathbb{R}^2\to\mathbb{R}^2$, определённое равенством

$$A(a, b)=(a, 0)$$

также является линейным оператором. Проверка этого факта предоставляется читателю.

Линейный оператор из последнего примера имеет простой геометрический смысл. Если на плоскости зафиксировать систему координат xOy, то каждому геометрическому вектору можно сопоставить упорядоченную пару чисел — координаты этого вектора в рассматриваемой системе координат. Другими словами, каждому геометрическому вектору можно поставить в соответствие арифметический. Из курса геометрии читателю должно быть известно, что это соответствие является взаимно однозначным. При этом координаты $(a,0)$ имеют геометрические векторы, параллельные оси Ox и только они. Поэтому линейный оператор A каждый геометрический вектор отображает в его проекцию на ось Ox вдоль направления, параллельного оси Oy.



Замечание 1. Отметим ещё раз, что соответствие между геометрическими и арифметическими векторами *зависит от выбора системы координат.*

Пример 3. Отображения $A:\mathbb{R}\to\mathbb{R}$ и $B:\mathbb{R}^2\to\mathbb{R}^3$, определённые соответственно равенствами

$$A(a)=a^2,\quad B(x_1,x_2)=(x_1,x_2+1,x_1),$$

не являются линейными операторами. Например, $A(a+b)=(a+b)^2$, $Aa+Ab=$ $=a^2+b^2$. Если оба числа a и b отличны от нуля, то $A(a+b)\neq Aa+Ab$.

Пример 4. Пусть $\boldsymbol{a_1},\boldsymbol{a_2},\boldsymbol{a_3}$ — произвольный набор трёх векторов из $\mathbb{R}^m$. Тогда отображение $A:\mathbb{R}^3\to\mathbb{R}^m$, определённое равенством

$$A(x_1,x_2,x_3)=x_1\boldsymbol{a_1}+x_2\boldsymbol{a_2}+x_3\boldsymbol{a_3}$$

является линейным оператором.

Из определения (например, требования однородности) следует, что для всякого линейного оператора $A:\mathbb{R}^n\to\mathbb{R}^m$ справедливы следующие свойства:

$$1.\ \ A\mathbf{0}=\mathbf{0};\ \ 2.\ \ \forall\,\boldsymbol{x}\in\mathbb{R}^n\quad A(-\boldsymbol{x})=-A\boldsymbol{x}.$$

Замечание (по настоятельному требованию формалиста-зануды). В первом свойстве аргумент оператора A является нулевым вектором арифметического пространства $\mathbb{R}^n$, а его образ (см. правую часть равенства) также является нулевым вектором, но уже пространства $\mathbb{R}^m$.

§15. Координатная запись линейного оператора

В этом параграфе мы установим общий вид произвольного линейного оператора в координатной форме. Естественно, что говорить о координатной форме имеет смысл лишь в том случае, когда выбран базис.

Начнём с рассмотрения частного случая, когда линейный оператор λ действует из $\mathbb{R}$ в $\mathbb{R}$. В этом случае λ обычно называют линейной функцией. Из определения линейного оператора получаем следующую цепочку равенств

$$\lambda(h)=\lambda(h\cdot 1)=h\cdot\lambda(1),$$

справедливую для любого $h\in\mathbb{R}$. Таким образом, функция λ полностью определяется значением при $h=1$. Если $\lambda(1)$ обозначить через a, то получим, что всякая линейная функция $\lambda:\mathbb{R}\to\mathbb{R}$ имеет вид $\lambda(h)=a\cdot h$.

С этим результатом вы столкнётесь в курсе математического анализа при изучении понятия дифференциала функции *одной* переменной.

Далее рассмотрим случай линейного оператора $A:\mathbb{R}^n\to\mathbb{R}$. Пусть $\boldsymbol{e_1},\boldsymbol{e_2},\ldots,$ $\boldsymbol{e_n}$ — канонический базис пространства $\mathbb{R}^n$ (см. §9, пример 6). Всякий вектор $\boldsymbol{h}=(h_1,h_2,\ldots,h_n)\in\mathbb{R}^n$ может быть записан в виде суммы $h_1\boldsymbol{e_1}+\ldots+h_n\boldsymbol{e_n}$. И снова из определения линейного оператора получаем

$$A\boldsymbol{h}=A(h_1\boldsymbol{e_1}+\ldots+h_n\boldsymbol{e_n})=h_1A\boldsymbol{e_1}+\ldots+h_nA\boldsymbol{e_n}.$$

Если числа $A\boldsymbol{e_i}$ обозначить через a_i, то можно записать

$$A\boldsymbol{h}=a_1h_1+a_2h_2+\ldots+a_nh_n.$$

Таким образом, последнее равенство является самой общей формой записи линейного оператора $A:\mathbb{R}^n\to\mathbb{R}$ в координатах.

С этим результатом вы столкнётесь в курсе математического анализа при изучении понятия дифференциала функции *n* переменных.

В случае произвольного линейного оператора $A:\mathbb{R}^n\to\mathbb{R}^m$ рассуждения аналогичны. Прежде всего выберем в $\mathbb{R}^n$ базис (не обязательно канонический):

$$\boldsymbol{e_1},\boldsymbol{e_2},\ldots,\boldsymbol{e_n}.$$

Всякий вектор $\boldsymbol{x}\in\mathbb{R}^n$ может быть записан единственным способом (см. теорему 3, §9) в виде суммы $x_1\boldsymbol{e_1}+x_2\boldsymbol{e_2}+\ldots+x_n\boldsymbol{e_n}$. Поэтому образ $A\boldsymbol{x}$ вектора $\boldsymbol{x}$

можно представить в виде

$$A\boldsymbol{x}=A(x_1\boldsymbol{e}_1+\ldots+x_n\boldsymbol{e}_n)=x_1A\boldsymbol{e}_1+\ldots+x_nA\boldsymbol{e}_n=x_1\boldsymbol{a}_1+\ldots+x_n\boldsymbol{a}_n,$$

где $\boldsymbol{a}_i=A\boldsymbol{e}_i$ являются *векторами* пространства $\mathbb{R}^m$.

Итак, самая общая запись линейного оператора $A:\mathbb{R}^n\to\mathbb{R}^m$ в координатах (относительно базиса $\boldsymbol{e}_i\in\mathbb{R}^n$) имеет вид

$$A\boldsymbol{x}=x_1\boldsymbol{a}_1+\ldots+x_n\boldsymbol{a}_n,$$

где $\boldsymbol{a}_i$ — произвольные векторы из $\mathbb{R}^m$, являющиеся образами базисных.

§16. Матрица линейного оператора. Вычисление координат образа

Выберем в пространствах $\mathbb{R}^n$ и $\mathbb{R}^m$ базисы $\boldsymbol{e}_1,\boldsymbol{e}_2,\ldots,\boldsymbol{e}_n$ и $\boldsymbol{f}_1,\boldsymbol{f}_2,\ldots,\boldsymbol{f}_n$ соответственно и рассмотрим линейный оператор $A:\mathbb{R}^n\to\mathbb{R}^m$. Векторы $A\boldsymbol{e}_1,A\boldsymbol{e}_2,\ldots,A\boldsymbol{e}_n$ принадлежат пространству $\mathbb{R}^m$ и поэтому каждый из этих векторов можно разложить по базису $\boldsymbol{f}_1,\boldsymbol{f}_2,\ldots,\boldsymbol{f}_n$. Если i-ю координату вектора $A\boldsymbol{e}_j$ обозначить через a_{ij}, то получим следующие разложения:

$$\begin{aligned}A\boldsymbol{e}_1&=a_{11}\boldsymbol{f}_1+a_{21}\boldsymbol{f}_2+\ldots+a_{m1}\boldsymbol{f}_m,\\ A\boldsymbol{e}_2&=a_{12}\boldsymbol{f}_1+a_{22}\boldsymbol{f}_2+\ldots+a_{m2}\boldsymbol{f}_m,\\ &\ldots\\ A\boldsymbol{e}_n&=a_{1n}\boldsymbol{f}_1+a_{2n}\boldsymbol{f}_2+\ldots+a_{mn}\boldsymbol{f}_m.\end{aligned}\tag{11}$$

Определение. *Матрицей линейного оператора* $A:\mathbb{R}^n\to\mathbb{R}^m$ относительно базисов $\boldsymbol{e}_1,\boldsymbol{e}_2,\ldots,\boldsymbol{e}_n$ и $\boldsymbol{f}_1,\boldsymbol{f}_2,\ldots,\boldsymbol{f}_m$ называется матрица, i-й столбец которой состоит из координат вектора $A\boldsymbol{e}_i$ в базисе $\boldsymbol{f}_1,\boldsymbol{f}_2,\ldots,\boldsymbol{f}_m$.

Таким образом каждому линейному оператору A соответствует некоторая матрица, которую кратко мы будем обозначать (a_{ij}) . Подчеркнём особо, что это соответствие зависит от выбора базисов. При замене базиса матрица *того же* оператора может измениться.

Выясним, как вычислить координаты образа произвольного вектора $\boldsymbol{x}$ при действии оператора A. Предположим, что числа $\alpha_1,\alpha_2,\ldots,\alpha_n$ являются координатами вектора $\boldsymbol{x}$ в базисе $\boldsymbol{e}_1,\boldsymbol{e}_2,\ldots\boldsymbol{e}_n$. Другими словами имеет место равенство

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

из которого следует, что $A\boldsymbol{x}=\alpha_1 A\boldsymbol{e_1}+\alpha_2 A\boldsymbol{e_2}+\ldots+\alpha_n A\boldsymbol{e_n}$. Используя равенства (11), получаем

$$\begin{aligned} A\boldsymbol{x}=&\alpha_1(a_{11}\boldsymbol{f_1}+a_{21}\boldsymbol{f_2}+\ldots+a_{m1}\boldsymbol{f_m})+\\ &+\alpha_2(a_{12}\boldsymbol{f_1}+a_{22}\boldsymbol{f_2}+\ldots+a_{m2}\boldsymbol{f_m})+\\ &\ldots\\ &+\alpha_n(a_{1n}\boldsymbol{f_1}+a_{2n}\boldsymbol{f_2}+\ldots+a_{mn}\boldsymbol{f_m}) \end{aligned}$$

или, после перегруппировки слагаемых,

$$\begin{aligned} A\boldsymbol{x}=&(a_{11}\alpha_1+a_{12}\alpha_2+\ldots+a_{1n}\alpha_n)\boldsymbol{f_1}+\\ &+(a_{21}\alpha_1+a_{22}\alpha_2+\ldots+a_{2n}\alpha_n)\boldsymbol{f_2}+\\ &\ldots\\ &+(a_{m1}\alpha_1+a_{m2}\alpha_2+\ldots+a_{mn}\alpha_n)\boldsymbol{f_m}. \end{aligned} \tag{12}$$

Если координаты вектора $A\boldsymbol{x}$ относительно базиса $\boldsymbol{f_1},\boldsymbol{f_2},\ldots,\boldsymbol{f_m}$ обозначить через $\beta_1,\beta_2,\ldots,\beta_m$, то из формулы (12), в силу однозначности разложения вектора по базису, следуют равенства

$$\begin{aligned} \beta_1&=a_{11}\alpha_1+a_{12}\alpha_2+\ldots+a_{1n}\alpha_n,\\ \beta_2&=a_{21}\alpha_1+a_{22}\alpha_2+\ldots+a_{2n}\alpha_n,\\ &\ldots\\ \beta_m&=a_{m1}\alpha_1+a_{m2}\alpha_2+\ldots+a_{mn}\alpha_n. \end{aligned} \tag{13}$$

Полученные формулы можно легко запомнить, если вспомнить известное из курса геометрии правило вычисления скалярного произведения векторов как суммы произведений одноимённых координат и распространить его на n-мерные арифметические векторы. Тогда координата β_i вектора $A\boldsymbol{x}$ равна скалярному произведению i-й строки матрицы оператора на вектор, образованный координатами $\alpha_1,\alpha_2,\ldots,\alpha_n$ вектора $\boldsymbol{x}$.

Для равенств (13) часто используют матричную форму записи:

$$\begin{pmatrix}\beta_1\\ \beta_2\\ \vdots\\ \beta_m\end{pmatrix}=\begin{pmatrix}a_{11} & a_{12} & \ldots & a_{1n}\\ a_{21} & a_{22} & \ldots & a_{2n}\\ . & . & \ldots & .\\ a_{m1} & a_{m2} & \ldots & a_{mn}\end{pmatrix}\cdot\begin{pmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{pmatrix}. \tag{14}$$

Замечание. Равенство (14) можно рассматривать как определение *умножения* матрицы типа $m\times n$ на n-элементный столбец, т. е. матрицу типа $n\times 1$. Результатом такого умножения является матрица типа $m\times 1$. При этом i-й элемент произведения вычисляется как скалярное произведение i-й строки матрицы на n-элементный столбец.

Более общую ситуацию мы рассмотрим в следующем параграфе.

§17. Действия над линейными операторами и их матрицами

Сложение линейных операторов

Пусть A и B — линейные операторы, действующие из $\mathbb{R}^n$ в $\mathbb{R}^m$.

Определение 1. *Суммой* линейных операторов A и B называется линейный оператор из $\mathbb{R}^n$ в $\mathbb{R}^m$, обозначаемый $A+B$ и действующий на $\mathbb{R}^n$ по правилу:

$$\forall\, \boldsymbol{x}\in\mathbb{R}^n \ (A+B)\boldsymbol{x}=A\boldsymbol{x}+B\boldsymbol{x}.$$

Чтобы определение суммы линейных операторов A и B было корректно, нужно проверить, что определяемый оператор $A+B$ также является линейным. Эта несложная проверка оставляется читателю.

Если в пространствах $\mathbb{R}^n$ и $\mathbb{R}^m$ зафиксировать соответственно базисы $\boldsymbol{e}_1,\boldsymbol{e}_2,\ldots,\boldsymbol{e}_n$ и $\boldsymbol{f}_1,\boldsymbol{f}_2,\ldots,\boldsymbol{f}_n$, то можно вычислить элементы матриц операторов A и B. Пусть эти матрицы (относительно выбранных базисов) имеют вид (a_{ij}) и (b_{ij}) соответственно. Это означает, что

$$A\boldsymbol{e}_j=\sum_{i=1}^{m} a_{ij}\boldsymbol{f}_i, B\boldsymbol{e}_j=\sum_{i=1}^{m} b_{ij}\boldsymbol{f}_i.$$

Складывая почленно последние соотношения, находим

$$(A+B)\boldsymbol{e}_j=A\boldsymbol{e}_j+B\boldsymbol{e}_j=\sum_{i=1}^{m} a_{ij}\boldsymbol{f}_i+\sum_{i=1}^{m} b_{ij}\boldsymbol{f}_i=\sum_{i=1}^{m}(a_{ij}+b_{ij})\boldsymbol{f}_i.$$

Полученное равенство $(A+B)\boldsymbol{e}_j=\sum_{i=1}^{m}(a_{ij}+b_{ij})\boldsymbol{f}_i$ означает, что матрица оператора $A+B$ получается из матриц операторов A и B покомпонентным сложением. Т. е. чтобы вычислить элемент матрицы оператора $A+B$, расположенный в позиции $(i\,j)$, нужно сложить элементы матриц операторов A и B, расположенные в тех же позициях.

Матрицу линейного оператора $A+B$ естественно называть суммой матриц (a_{ij}) и (b_{ij}).

Определение 1'. *Суммой* двух матриц (a_{ij}) и (b_{ij}) типа $m\times n$ называется матрица (c_{ij}) того же типа $m\times n$, элементы которой определяются равенством $c_{ij}=a_{ij}+b_{ij}$.

Умножение линейных операторов

Обозначим через A линейный оператор, действующий из $\mathbb{R}^n$ в $\mathbb{R}^l$, а через B — линейный оператор, действующий из $\mathbb{R}^l$ в $\mathbb{R}^m$.

Определение 2. *Произведением* линейных операторов A и B (в указанном порядке) называется линейный оператор из $\mathbb{R}^n$ в $\mathbb{R}^m$, обозначаемый BA и действующий на $\mathbb{R}^n$ по правилу:

$$\forall\, \boldsymbol{x}\in\mathbb{R}^n\ (BA)\boldsymbol{x}=B(A\boldsymbol{x}).$$

То, что отображение BA действует из $\mathbb{R}^n$ в $\mathbb{R}^m$ следует из определения. Остаётся только проверить линейность этого отображения, что предоставляется читателю.

Пусть $\boldsymbol{e_1},\boldsymbol{e_2},\ldots,\boldsymbol{e_n}$, $\boldsymbol{f_1},\boldsymbol{f_2},\ldots,\boldsymbol{f_l}$ и $\boldsymbol{g_1},\boldsymbol{g_2},\ldots,\boldsymbol{g_m}$ — базисы в пространствах $\mathbb{R}^n,\mathbb{R}^l$ и $\mathbb{R}^m$, а $(a_{ij}),(b_{pq})$ — матрицы операторов A и B относительно соответствующих базисов.

Вопрос для самоконтроля.
Какие значения могут принимать индексы i,j,p,q?

В силу определений и принятых обозначений, имеют место равенства

$$A\boldsymbol{e_j}=\sum_{i=1}^{l} a_{ij}\boldsymbol{f_i},\ B\boldsymbol{f_q}=\sum_{i=1}^{m} b_{pq}\boldsymbol{g_p}.$$

Чтобы найти элементы матрицы оператора BA, нужно вычислить $(BA)\boldsymbol{e_j}$:

$$(BA)\boldsymbol{e_j}=B(A\boldsymbol{e_j})=B(\sum_{i=1}^{l} a_{ij}\boldsymbol{f_i})=\sum_{i=1}^{l} a_{ij}B\boldsymbol{f_i}=\sum_{i=1}^{l} a_{ij}\sum_{k=1}^{m} b_{ki}\boldsymbol{g_k}.$$

В силу дистрибутивности операции умножения относительно сложения, произведение $a_{ij}\sum_{k=1}^{m} b_{ki}\boldsymbol{g_k}$ можно переписать в виде $\sum_{k=1}^{m} a_{ij}b_{ki}\boldsymbol{g_k}=\sum_{k=1}^{m} b_{ki}a_{ij}\boldsymbol{g_k}$. Следовательно $(BA)\boldsymbol{e_j}=\sum_{i=1}^{l}\sum_{k=1}^{m} b_{ki}a_{ij}\boldsymbol{g_k}$. Поскольку операция сложения чисел является коммутативной, в последнем равенстве можно изменить порядок суммирования, что приведёт к соотношению

$$(B\,A)\boldsymbol{e_j}=\sum_{k=1}^{m}\sum_{i=1}^{l}b_{k\,i}a_{i\,j}\,\boldsymbol{g_k}=\sum_{k=1}^{m}\left(\sum_{i=1}^{l}b_{k\,i}a_{i\,j}\right)\boldsymbol{g_k}.$$

Полученное соотношение означает, что элемент матрицы оператора $B\,A$, расположенный в позиции $(k\,j)$, равен $\sum_{i=1}^{l}b_{k\,i}a_{i\,j}$.

Матрицу линейного оператора $B\,A$ естественно назвать произведением матриц $(a_{i\,j})$ и $(b_{p\,q})$.

Определение 2'. *Произведением* матрицы $(a_{i\,j})$ типа $l\times n$ на матрицу $(b_{p\,q})$ типа $m\times l$ (в указанном порядке) называется матрица $(c_{r\,s})$ типа $m\times n$, элементы которой определяются равенством

$$c_{r\,s}=\sum_{i=1}^{l}b_{r\,i}a_{i\,s}. \tag{15}$$

Более подробно, элемент матрицы оператора $B\,A$, расположенный на позиции $(r\,s)$ равен, выражаясь геометрическим языком, скалярному произведению элементов r-й строки матрицы $(b_{p\,q})$ и s-го столбца матрицы $(a_{i\,j})$.

Заметим, что в случае $n=1$ равенства (15) совпадают с равенствами (13).

Внимательный читатель (слушатель) должно быть заметил, что операция умножения линейных операторов является частным случаем более общей операции — композиции функций. Следовательно, *операция умножения линейных операторов* является *ассоциативной*, т. е. для любых трёх линейных операторов $A:\mathbb{R}^n\to\mathbb{R}^l$, $B:\mathbb{R}^l\to\mathbb{R}^k$, $C:\mathbb{R}^k\to\mathbb{R}^m$ справедливо равенство

$$(C\,B)\,A=C\,(B\,A).$$

Поскольку при фиксированном базисе каждая матрица однозначно определяет некоторый линейный оператор, то операция *умножения матриц* также является *ассоциативной*. Рекомендуется проверить последнее утверждение напрямую, исходя из (15).

В заключение обсудим вопрос о коммутативности умножения операторов и матриц. Для линейных операторов $A:\mathbb{R}^n\to\mathbb{R}^l$, $B:\mathbb{R}^l\to\mathbb{R}^m$ определено произведение $B\,A$. Для того, чтобы было определено произведение $A\,B$, необходимо выполнение равенства $m=n$. В этом случае $B\,A$ действует из $\mathbb{R}^n$ в $\mathbb{R}^n$, а $A\,B$ — из $\mathbb{R}^l$ в $\mathbb{R}^l$. Если $l\neq n$, то, очевидно, $B\,A\neq A\,B$ (хотя бы потому, что у произведений разные области определения). Но даже в случае $l=n$ равенство $B\,A=A\,B$ может не выполняться.

Пример. Пусть линейные операторы A и B действуют из $\mathbb{R}^2$ в $\mathbb{R}^2$ и определяются в некотором базисе (одинаковом для обоих операторов) матрицами $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ и $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ соответственно. Тогда оператор BA имеет в том же базисе матрицу $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, а оператор AB — матрицу $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Следовательно, операторы BA и AB не совпадают.

Таким образом, в общем случае произведение операторов не коммутативно. То же высказывание справедливо и для матриц.

Умножение линейного оператора на число

Пусть A — линейный оператор, действующий из $\mathbb{R}^n$ в $\mathbb{R}^m$, а α — произвольное число.

Определение 3. *Произведением* линейного оператора A на число α называется линейный оператор из $\mathbb{R}^n$ в $\mathbb{R}^m$, обозначаемый αA и действующий на $\mathbb{R}^n$ по правилу:

$$\forall\, \boldsymbol{x} \in \mathbb{R}^n \ (\alpha A)\boldsymbol{x} = \alpha(A\boldsymbol{x}).$$

Как обычно, проверка линейности оператора αA оставляется читателю.

Также читателю предоставляется найти матрицу оператора αA, если матрица оператора A известна. После выполнения этого упражнения становится естественным

Определение 3'. *Произведением* матрицы (a_{ij}) типа $m \times n$ на число α называется матрица (c_{ij}) того же типа $m \times n$, элементы которой определяются равенством $c_{ij} = \alpha a_{ij}$.

Задача. Пусть A, B и C — линейные операторы, действующие из $\mathbb{R}^n$ в $\mathbb{R}^m$ и λ, μ — произвольные числа. Докажите следующие свойства операций сложения операторов и умножения оператора на число.

Сложение операторов:

1. $\forall\, A, B \quad A+B=B+A;$
2. $\forall\, A, B, C \quad (A+B)+C=A+(B+C);$
3. $\exists O \mid \forall\, A \quad A+O=O+A=A;$
4. $\forall\, A \quad \exists(-A) \mid A+(-A)=(-A)+A=O.$

В свойстве 3 нужно указать, из каких элементов состоит матрица O и проверить справедливость написанных равенств. Аналогично, в свойстве 4 требуется определить, как по известным элементам матрицы A найти элементы матрицы $-A$ и, конечно же, проверить справедливость написанных равенств.

Умножение оператора на число:

1. $\forall A, \forall \lambda, \mu \quad (\lambda\mu)A = \lambda(\mu A)$;
2. $\forall A, \forall \lambda, \mu \quad (\lambda+\mu)A = \lambda A + \mu A$;
3. $\forall A, B \quad \lambda(A+B) = \lambda A + \lambda B$;
4. $\forall A \quad 1A = A$.

Вопрос для самоконтроля.

Прочитайте ещё раз замечание из §8. Что Вы можете сказать о множестве всех линейных операторов из $\mathbb{R}^n$ в $\mathbb{R}^m$ с точки зрения современной[5] математики? А о множестве всех матриц типа $m \times n$?

§18. Квадратные матрицы. Единичная матрица

Рассмотрим множество матриц типа $n \times n$, т. е. матриц, у которых количество строк совпадает с количеством столбцов. Такие матрицы называются *квадратными,* а число *n* — *порядком* матрицы. Нетрудно видеть, что любые две квадратные матрицы порядка *n* можно умножать и что их произведением будет являться квадратная матрица того же порядка *n.* На протяжении всего параграфа будут рассматриваться только квадратные матрицы фиксированного порядка.

Коль скоро на множестве квадратных матриц задана операция умножения, можно поставить задачу о нахождении специальной матрицы, обладающей свойством единицы.

Определение. Квадратная матрица E называется *единичной,* если для всякой матрицы A того же порядка справедливы равенства

$$AE = EA = A.$$

Элементы единичной матрицы порядка *n* обозначим через δ_{ij}. Для вычисления δ_{ij} воспользуемся формулой (15). Если $A = (a_{ij})$, то равенство $AE = A$ в развёрнутой форме примет вид

$$\sum_{k=1}^{n} a_{ik}\delta_{kj} = a_{ij}. \tag{16}$$

Подчеркнём, что последнее соотношение представляет собой на самом деле *систему* n^2 равенств, так как оба индекса *i* и *j* независимо друг от друга принимают значения от 1 до *n.*

Зафиксируем какие-нибудь значения индексов $i = i_0$ и $j = j_0$. Матричное ра-

5 Некоторые закоренелые формалисты утверждают, что при такой постановке вопроса лучше говорить не о современной, а формальной математике.

венство $AE=A$ должно выполняться для *любой* матрицы A, в частности, для матрицы, у которой все элементы, кроме одного, расположенного на позиции $(i_0 j_0)$, равны нулю. Допустим, что $a_{i_0 j_0}=1$. Для такой матрицы все уравнения системы (16), у которых $i\neq i_0$, будут иметь вид $0=0$. Если же $i=i_0$, то уравнения системы (16) примут вид $\delta_{j_0 j}=a_{i_0 j}$, так как в левой части (16) все слагаемые с $k\neq j_0$ заведомо равны нулю и $a_{i_0 j_0}=1$. Следовательно $\delta_{j_0 j}=0$ в случае $j\neq j_0$ и $\delta_{j_0 j}=1$ в случае $j=j_0$.

Таким образом, все недиагональные элементы единичной матрицы равны нулю, а диагональные — единице. Например, единичная матрица порядка 4 имеет вид

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

Замечание 1. Символ $\delta_{ij}=\begin{cases}0, i\neq j,\\ 1, i=j\end{cases}$ довольно часто используется в сокращённых записях сумм. Символ δ_{ij} называется *дельта-символом Кронекера.*

Замечание 2. Понятие единичной матрицы можно ввести иным способом. Можно рассмотреть оператор $I:\mathbb{R}^n\to\mathbb{R}^n$, $I(\boldsymbol{x})=\boldsymbol{x}$ и заметить, что для любого оператора $A:\mathbb{R}^n\to\mathbb{R}^n$ справедливо равенство $AI=IA=A$. Остаётся только заметить (а читателю проверить), что во всяком базисе пространства $\mathbb{R}^n$ элементы линейного оператора I равны δ_{ij}.

§19. Обратимые матрицы. Понятие обратной матрицы

Здесь, как и в предыдущем параграфе, рассматриваются только квадратные матрицы фиксированного порядка. Единичная матрица обозначается через E.

Определение. Матрица A порядка n называется *обратимой,* если существует такая матрица B, что справедливы равенства

$$AB=BA=E.$$

Если матрица A обратима, то матрица B из предыдущего определения называется матрицей, *обратной* к A. Обратная к A матрица обычно обозначается символом A^{-1}.

Метод вычисления обратной матрицы поясним на примере.

Пример. Найдём матрицу, обратную к $A=\begin{pmatrix}2&3\\1&2\end{pmatrix}$. Положим

$$A^{-1}=\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}.$$

Числа x_{ij} нужно подобрать таким образом, чтобы выполнялось двойное матричное равенство $AA^{-1}=A^{-1}A=E$. Рассмотрим отдельно соотношение $A^{-1}A=E$, которое эквивалентно системе четырёх линейных уравнений с четырьмя неизвестными x_{ij} :

$$\begin{cases} 2x_{11}+3x_{21}=1, \\ x_{11}+2x_{21}=0, \\ 2x_{12}+3x_{22}=0, \\ x_{12}+2x_{22}=1. \end{cases}$$

Решая эту систему, например, методом Гаусса, находим, что $A^{-1}=\begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix}$.

Замечание. В рассмотренном примере мы использовали только одно из равенств, определяющих A^{-1}, а именно $AA^{-1}=E$. Можно доказать, что во всех случаях из справедливости равенства $AB=E$ вытекает справедливость равенства $BA=E$.

§20. Перестановки из *n* элементов

Рассмотрим множество $X=\{1,2,\ldots,n\}$. *Перестановкой* из *n* элементов называется любое биективное отображение $\sigma : X \to X$. Множество всех перестановок из *n* элементов будем обозначать через S_n.

Имеется два основных способа записи перестановок. Один из них заключается в том, что перестановку σ задают с помощью матрицы из двух строк. В первой строке перечисляются элементы X, а во второй — их образы:

$$\sigma=\begin{pmatrix} 1 & 2 & \ldots & n \\ \sigma(1) & \sigma(2) & \ldots & \sigma(n) \end{pmatrix}.$$

Задача. Сколько элементов содержит множество S_n?

Второй способ задания называется *представлением перестановки в виде произведения*[6] *независимых циклов.* Если для всякого $i\in X$ имеет место равенство $\sigma(i)=i$, то такая перестановка называется *единичной* и она записывается в виде $\sigma=(1)$. Для любой другой перестановки обязательно найдётся i, для

6 Слово «произведение» пока нужно понимать формально.

которого $\sigma(i)\neq i$. В этом случае представление σ в виде произведения независимых циклов начинается так:

$$\sigma=(i\ \sigma(i)\ \sigma^2(i)\ldots)\ \ldots,$$

где σ^k — k-я итерация отображения σ. Выписанная строчка (*цикл*) состоит из конечного числа *попарно различных* элементов множества X. (*Почему?*)

Цикл $(i\ \sigma(i)\ \sigma^2(i)\ldots)$ полностью определяет перестановку, если в выписанной строчке встречаются все элементы множества X. Если же имеется элемент $j\in X$, такой что $\sigma(j)\neq j$ и j не входит в построенный цикл, то для этого элемента строим аналогичный цикл из итераций и приписываем его к первому. Так поступаем до тех пор, пока не будут исчерпаны все элементы множества X.

При таком способе задания перестановки договариваются не включать в запись *неподвижные* элементы перестановки, т. е. те элементы множества X, для которых $\sigma(i)=i$. Таким образом, циклы любой перестановки (кроме единичной) содержат не менее двух элементов.

Из алгоритма разложения перестановки на циклы видно, что различные циклы не имеют общих элементов. Такие циклы называют *независимыми.*

Пример 1. Перестановка $\sigma=\begin{pmatrix}1&2&3&4&5\\4&5&1&2&3\end{pmatrix}$ представляется единственным циклом $\sigma=(1\,4\,2\,5\,3)$, так как $\sigma(1)=4,\sigma(2)=5,\sigma(5)=3,\sigma(3)=1$.

Пример 2. Для перестановки $\tau=\begin{pmatrix}1&2&3&4&5&6&7\\3&2&4&1&6&5&7\end{pmatrix}$ $\tau(1)=3,\tau(3)=4$, $\tau(4)=1$, поэтому цикл $(1\,3\,4)$ не исчерпывает всех элементов τ и построение циклов нужно продолжить: $\tau(5)=6,\tau(6)=5$. Поскольку элементы 2 и 7 являются неподвижными для τ, то $\tau=(1\,3\,4)(5\,6)$.

Пример 3. Пусть перестановка из S_7 задана в виде произведения независимых циклов $\sigma=(1\,6)(2\,5)(3\,4)$. Тогда её можно записать в виде матрицы

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

Определение. Перестановка состоящая из одного двухэлементного цикла называется *транспозицией.*

В примере 3 каждый независимый цикл является транспозицией.

§21. Умножение перестановок

Из курса математического анализа читателю должно быть известно, что композиция биекций сама является биекцией. Поэтому для любых двух перестановок $\sigma,\tau\in S_n$ композиция $\tau\sigma$ также является перестановкой из S_n. Перестановка $\tau\sigma$ называется *произведением*[7] σ и τ (в указанном порядке).

Определённая операция умножения перестановок является ассоциативной операцией (*почему?*), но, в общем случае, не коммутативной.

Пример 1. Для двух перестановок $\sigma=(214),\tau=(13)\in S_4$ произведение $\tau\sigma=(1423)$, но $\sigma\tau=(1342)$. Ясно, что $\tau\sigma\neq\sigma\tau$, поскольку, например, $\tau\sigma(2)=3,\sigma\tau(2)=1$.

Из того же курса математического анализа известно, что всякое биективное отображение обратимо, причём обратное отображение само является биективным. Применительно ко множеству S_n этот факт означает, что для всякой перестановки σ найдётся перестановка τ, такая что оба произведения $\tau\sigma$ и $\sigma\tau$ являются единичными перестановками. Как обычно перестановка, обратная к σ, обозначается через σ^{-1}.

Очевидно также, что если единичную перестановку обозначить через ε, то для всякой перестановки σ справедливы равенства $\sigma\varepsilon=\varepsilon\sigma=\sigma$.

Итак, операция умножения перестановок обладает следующими свойствами.

1. $\forall\sigma,\tau,\rho\in S_n \quad (\sigma\tau)\rho=\sigma(\tau\rho)$;
2. $\exists\varepsilon\in S_n|\forall\sigma\in S_n \quad \sigma\varepsilon=\varepsilon\sigma=\sigma$;
3. $\forall\sigma \quad \exists(\sigma^{-1})|\sigma\sigma^{-1}=\sigma^{-1}\sigma=\varepsilon$.

Теорема 1. Всякая перестановка может быть представлена в виде произведения транспозиций.

Доказательство.

Достаточно рассмотреть случай, когда перестановка представлена одним циклом, т. е. когда $\sigma=(i_1 i_2\dots i_k)$. В целях наглядности будем считать, что $k=5$. Доказательство теоремы завершает проверка соотношения

$$(i_1 i_2 i_3 i_4 i_5)=(i_1 i_2)(i_2 i_3)(i_3 i_4)(i_4 i_5). \ \square$$

Пример 2. $(134)(56)=(13)(34)(56)$.

Представление перестановки в виде произведения транспозиций не единственно. Например, в S_n, $n\geqslant 3$, является верным равенство

$$(123)=(12)(23)=(13)(23)(13)(23).$$

Т. е. перестановка (123) представлена в виде произведения транспозиций дву-

---

7 Введённая операция позволяет рассматривать произведение независимых циклов из §20 в буквальном смысле, т. е. как результат умножения.

мя различными способами. Однако имеет место

Теорема 2. Если заданы два представления перестановки в виде произведения s и r транпозиций, то числа s и r имеют одинаковую чётность.

В примере 2 числа s и r равны 2 и 4.

Прежде чем доказывать теорему 2, введём понятие инверсии.

Определение 1. Упорядоченная пара различных целых чисел $(i\,j)$ образует *инверсию,* если $i>j$. Числа $(i\,j)$ образуют *правильную* пару, если $i<j$.

Пусть перестановка $\sigma\in S_n$ задана в виде таблицы

$$\sigma=\begin{pmatrix} i_1 & i_2 & \dots & i_n \\ \sigma(i_1) & \sigma(i_2) & \dots & \sigma(i_n)\end{pmatrix}.$$

Обозначим через $I(\sigma)$ количество упорядоченных пар в обеих строках, образующих инверсию.

Пример 2. Если $\sigma=\begin{pmatrix} 3 & 2 & 4 & 1 & 6 & 5 \\ 6 & 5 & 3 & 4 & 2 & 1\end{pmatrix}$, то в верхней строке инверсию образуют пары $(3\,1),(2\,1),(4\,1),(3\,2),(6\,5)$, а в нижней — $(6\,1),(5\,1),(3\,1),(4\,1)$, $(2\,1),(6\,2),(5\,2),(3\,2),(4\,2),(6\,3),(5\,3),(6\,4),(5\,4),(6\,5)$. Значит $I(\sigma)=19$.

Лемма 1. Для любой перестановки $\sigma\in S_n$ $\quad I(\sigma\cdot(i\ \ i+1))=I(\sigma)\pm 1$.

Доказательство.

Если $\sigma=\begin{pmatrix} \dots & i & i+1 & \dots \\ \dots & \sigma(i) & \sigma(i+1) & \dots\end{pmatrix}$, то $\sigma\cdot(i\ \ i+1)=\begin{pmatrix} \dots & i & i+1 & \dots \\ \dots & \sigma(i+1) & \sigma(i) & \dots\end{pmatrix}$.

Ясно, что если пара $(\sigma(i)\ \sigma(i+1))$ образует инверсию, то $(\sigma(i+1)\ \sigma(i))$ — правильная пара и наоборот. Кроме того, порядок других пар в каждой перестановке σ и $\sigma\cdot(i\ \ i+1)$ одинаковый. Поэтому $I(\sigma\cdot(i\ \ i+1))=I(\sigma)\pm 1$. □

Определение 2. Перестановка σ называется *чётной,* если $I(\sigma)$ чётно, и *нечётной,* если $I(\sigma)$ нечётно.

Задача 1. Проверьте, что перестановка $(i\ \ i+1)$ является нечётной.

Задача 2. Проверьте справедливость равенства

$$(i\ \ i+3)=(i\ \ i+1)(i+1\ \ i+2)(i+2\ \ i+3)(i+1\ \ i+2)(i\ \ i+1).$$

Равенство из задачи 2 означает, что транспозиция $(i\ \ i+3)$ может быть представлена в виде произведения пяти транспозиций *соседних* элементов.

Задача 3. Запишите транспозицию $(i\ \ i+p)$ в виде произведения $2p-1$ транспозиций соседних элементов.

Указание. Проанализируйте соотношение из задачи 2 и попытайтесь выяснить его происхождение. Поэкспериментируйте с перекладыванием монеток, ручек или др. подручных материалов.

Лемма 2. Всякая транспозиция является нечётной перестановкой.

Доказательство вытекает из результатов задач 1, 3 и леммы 1. □

Задача 4. Произведение чётного числа транспозиций является чётной перестановкой, а произведение нечётного числа транспозиций — нечётной перестановкой.

Доказательство теоремы 2.

Пусть перестановка σ представлена в виде произведения s транспозиций, каждая из которых согласно лемме 2 является нечётной перестановкой. Допустим, что σ является чётной перестановкой. Из задачи 4 следует, что это возможно только при чётном s. Аналогично, если перестановка σ — нечётная, то число s может быть только нечётным. □

Задача 5. Произведение двух чётных (двух нечётных) перестановок является чётной перестановкой. Произведение чётной и нечётной перестановки есть перестановка нечётная.

Задача 6. При $n \geqslant 2$ во множестве S_n чётных перестановок столько же, сколько нечётных.

§22. Определители *n*-го порядка

Определение 1. Знаком перестановки σ называют число $\operatorname{sgn}\sigma = (-1)^{I(\sigma)}$.

Рассмотрим квадратную матрицу порядка n

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

Определение 2. *Определителем* матрицы A называется число

$$\det A = \sum_{\sigma \in S_n} \operatorname{sgn}\sigma\, a_{1\sigma(1)} a_{2\sigma(2)} \dots a_{n\sigma(n)}. \tag{17}$$

Таким образом, определитель матрицы порядка n есть сумма $n!$ слагаемых, каждое из которых является произведением n элементов матрицы, взятых из каждой строки и каждого столбца по одному и только одному разу; если множители упорядочить по возрастанию номеров строк, то слагаемое берётся со знаком $+$, если перестановка номеров столбцов четная и со знаком $-$ в противном случае.

Свойство 1. Для любой матрицы *n*-го порядка $\det A^t = \det A$.

Доказательство.

По определению, *i*-я строка матрицы A^t состоит из элементов *i*-го столбца матрицы A. Поэтому $\det A^t = \sum_{\sigma\in S_n} \operatorname{sgn}\sigma\, a_{\sigma(1)1} a_{\sigma(2)2} \dots a_{\sigma(n)n}$. Заметим, что если перестановка σ пробегает всё множество S_n, то обратная перестановка σ^{-1} также пробегает всё множество S_n. Следовательно замена σ на σ^{-1} в сумме, определяющей число $\det A^t$, лишь изменит порядок слагаемых, что конечно не изменит значения самой суммы:

$$\det A^t = \sum_{\sigma\in S_n} \operatorname{sgn}\sigma^{-1} a_{\sigma^{-1}(1)1} a_{\sigma^{-1}(2)2} \dots a_{\sigma^{-1}(n)n}.$$

Теперь в каждом слагаемом $\operatorname{sgn}\sigma^{-1} a_{\sigma^{-1}(1)1} a_{\sigma^{-1}(2)2} \dots a_{\sigma^{-1}(n)n}$ изменим порядок множителей таким образом, чтобы номера строк (первые индексы у элементов a_{ij} матрицы A) располагались в возрастающем порядке. Нетрудно видеть[8], что после такой перестановки произведение перепишется в виде $\operatorname{sgn}\sigma^{-1} a_{1\sigma(1)} a_{2\sigma(2)} \dots a_{n\sigma(n)}$. И поскольку $\operatorname{sgn}\sigma^{-1} = \operatorname{sgn}\sigma$,

$$\det A^t = \sum_{\sigma\in S_n} \operatorname{sgn}\sigma\, a_{1\sigma(1)} a_{2\sigma(2)} \dots a_{n\sigma(n)} = \det A. \quad \square$$

Свойство 2. При умножении *любой* строки определителя на число, определитель умножается на это число.

Свойство 3. Если в определителе поменять местами строки, то определитель изменит знак.

Доказательство.

Будем для определённости считать, что переставляются 1-я и 2-я строки. Обозначая матрицу с переставленными строками через A_{12}, запишем равенства

$$\det A_{12} = \sum_{\sigma\in S_n} \operatorname{sgn}\sigma\, a_{2\sigma(1)} a_{1\sigma(2)} \dots a_{n\sigma(n)} = \sum_{\sigma\in S_n} \operatorname{sgn}\sigma\, a_{1\sigma(2)} a_{2\sigma(1)} \dots a_{n\sigma(n)}.$$

Перестановку τ номеров столбцов в последней сумме можно представить в виде произведения $\sigma\cdot(1\,2)$ (см. доказательство леммы 1 в §21). Из результатов предыдущего параграфа следует, что $\operatorname{sgn}\tau = -\operatorname{sgn}\sigma$. Для завершения доказательства нужно заметить, что если σ пробегает всё множество S_n, то τ также пробегает всё множество S_n:

$$\det A_{12} = -\sum_{\tau\in S_n} \operatorname{sgn}\tau\, a_{1\tau(1)} a_{2\tau(2)} \dots a_{n\tau(n)} = -\det A. \quad \square$$

8 Это не стандартная фраза, оправдывающая отсутствие обоснования. Формулируемое далее утверждение на самом деле очевидно.

Свойство 4. Определитель, элементы некоторой строки которого представлены в виде суммы двух слагаемых, равен сумме двух определителей, каждый из которых получается из исходного заменой одноимённой строки строкой первых и вторых слагаемых соответственно.

Свойство 5. Определитель с нулевой строкой равен нулю.

Свойство 6. Определитель с одинаковыми строками равен нулю.

Доказательство.

Если в матрице A поменять местами две одинаковые строки, то её определитель не изменится. С другой стороны, согласно свойству 3, определитель матрицы A изменит знак. Поэтому $\det A = -\det A$, т. е. $\det A = 0$. □

Из свойств 2, 4 и 6 вытекает

Свойство 7. Определитель не изменится, если к элементам некоторой строки определителя прибавить соответственные элементы другой строки, умноженные на одно и то же число.

Свойство 8. Определитель равен нулю тогда и только тогда, когда его строки линейно зависимы.

Доказательство.

Если в матрице A система строк является линейно зависимой, то равенство $\det A = 0$ вытекает из свойств 2, 4 и 6.

Доказательство обратного утверждения, т. е. что из равенства $\det A = 0$ следует линейная зависимость строк матрицы можно провести следующим образом. Согласно свойствам 2, 3 и 7 элементарные преобразования строк матрицы не меняют абсолютной величины её определителя. Напомним, что всякую матрицу с помощью элементарных преобразований можно привести к ступенчатому виду и что при элементарных преобразованиях ранг матрицы не меняется не меняется.

Приведём матрицу A порядка n к ступенчатому виду. Если строки матрицы линейно независимы, т. е. если $\operatorname{rank} A = n$, то преобразованная матрица будет иметь строго треугольную форму (см. §12). Определитель такой матрицы равен произведению её диагональных элементов, а значит отличен от нуля, что противоречит исходному равенству $\det A = 0$. □

§23. Разложение определителя по элементам строки (столбца)

Рассмотрим определитель матрицы A n-го порядка. Выберем в этой матрице произвольным образом k строк и k столбцов $(k \leqslant n)$. Элементы матрицы A, расположенные на пересечении выбранных строк и столбцов, образуют матрицу порядка k, определитель которой называется *минором* k-го порядка матрицы A. Если $i_1, i_2, \ldots, i_k$ — номера строк, а $j_1, j_2, \ldots, j_k$ — номера столбцов, то соответствующий минор k-го порядка обозначим через $M_{j_1 j_2 \ldots j_k}^{i_1 i_2 \ldots i_k}$.

Задача. Сколькими способами можно составить минор k-го порядка из матрицы n-го порядка?

Элементы матрицы A, не принадлежащие выбранным строкам и столбцам, образуют матрицу порядка $n-k$, определитель которой называется *дополнительным минором* и обозначается символом $\overline{M}{}^{i_1 i_2 \dots i_k}_{j_1 j_2 \dots j_k}$. Число $A^{i_1 i_2 \dots i_k}_{j_1 j_2 \dots j_k}$, равное по определению числу $(-1)^{i_1+i_2+\dots+i_k+j_1+j_2+\dots+j_k}\overline{M}{}^{i_1 i_2 \dots i_k}_{j_1 j_2 \dots j_k}$, называется *алгебраическим дополнением* минора $M^{i_1 i_2 \dots i_k}_{j_1 j_2 \dots j_k}$.

Пример 1. Пусть $A=\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{pmatrix}, i_1=1, i_2=4, j_1=3, j_2=5.$ Тогда

$$M^{14}_{35}=\begin{vmatrix} a_{13} & a_{15} \\ a_{43} & a_{45} \end{vmatrix},\ \overline{M}{}^{14}_{35}=\begin{vmatrix} a_{21} & a_{22} & a_{24} \\ a_{31} & a_{32} & a_{34} \\ a_{51} & a_{52} & a_{54} \end{vmatrix},\ A^{14}_{35}=-\begin{vmatrix} a_{21} & a_{22} & a_{24} \\ a_{31} & a_{32} & a_{34} \\ a_{51} & a_{52} & a_{54} \end{vmatrix}.$$

Пример 2. Каждый элемент a_{ij} матрицы A является минором первого порядка, а определитель $\det A$ — минором n-го порядка.

Пример 3. Рассмотрим матрицу, у которой все элементы первой строки, за исключением может быть первого, равны нулю. Например, для матрицы пятого порядка

$$A=\begin{pmatrix} a_{11} & 0 & 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{pmatrix}.$$

По определению $\det A=\sum\limits_{\sigma\in S_n} \operatorname{sgn}\sigma\, a_{1\sigma(1)}a_{2\sigma(2)}\dots a_{n\sigma(n)}$, однако все слагаемые отвечающие перестановкам $\sigma\in S_n$ с $\sigma(1)\neq 1$ будут нулевыми. Поэтому при вычислении определителя достаточно рассмотреть лишь те $\sigma\in S_n$, для которых $\sigma(1)=1$. Такие перестановки образуют множество *всех* биекций множества $\{2,3,\dots,n\}$, состоящего из $n-1$ элемента. Отсюда получаем, что

$$\det A=a_{11}\sum_{\sigma\in S_{n-1}} \operatorname{sgn}\sigma\, a_{2\sigma(2)}\dots a_{n\sigma(n)}.$$

Последняя сумма равна минору $\overline{M}{}^1_1$, дополнительному к элементу a_{11}. Следовательно, для матрицы рассматриваемого типа $\det A=a_{11}\overline{M}{}^1_1$.

Рассмотренный пример обобщается в теоремах 1 и 2.

Теорема 1. Определитель матрицы, все элементы i-й строки которой, кроме может быть расположенного в j-м столбце, равны нулю, равен произведению элемента a_{ij} на его алгебраическое дополнение.

Доказательство.

В случае, когда $i=j=1$ утверждение теоремы доказано, так как $A_1^1=\overline{M}_1^1$. К этому случаю можно свести и общий. С этой целью переставим i-ю строку с $(i-1)$-й, затем с $(i-2)$-й и так до тех пор, пока i-я строка не станет первой. Аналогичным образом переместим j-й столбец полученной матрицы в первый. В результате этих $(i-1)+(j-1)$ перестановок определитель матрицы умножится на $(-1)^{i+j}$, а относительное расположение строк (отличных от i-й) и столбцов (отличных от j-й) останется прежним. Поэтому $(-1)^{i+j}\det A=a_{ij}\overline{M}_j^i$ и $\det A=a_{ij}(-1)^{i+j}\overline{M}_j^i=a_{ij}A_j^i$. □

Теорема 2. (Разложение определителя по элементам строки) Определитель матрицы равен сумме произведений элементов i-й строки на их алгебраические дополнения.

Доказательство.

Строку $(a_{i1},a_{i2},\ldots,a_{in})$ можно представить в виде суммы

$$(a_{i1},0,\ldots,0)+(0,a_{i2},\ldots,0)+\ldots+(0,0,\ldots,a_{in}).$$

Для завершения доказательства остаётся воспользоваться свойством определителей №4 и уже доказанной теоремой 1:

$$\det A=a_{i1}A_1^i+a_{i2}A_2^i+\ldots+a_{in}A_n^i. \quad \square$$

Пример 4. Рассмотрим матрицу n-го порядка, у которой имеются две одинаковые строки. Например, при $n=5$ в матрице

$$A=\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \end{pmatrix}$$

совпадают вторая и пятая строки. По свойству определителей №5 $\det A=0$. Разложив определитель матрицы A по элементам пятой строки, получим равенство

$$a_{21}A_1^5+a_{22}A_2^5+a_{23}A_3^5+a_{24}A_4^5+a_{25}A_5^5=0,$$

левую часть которого можно рассматривать как сумму произведений элементов *второй* строки на алгебраические дополнения соответствующих элементов *пятой* строки. Это рассуждение очевидным образом преобразуется в доказательство следующей теоремы.

Теорема 3. Сумма произведений элементов строки определителя на алгебраические дополнения соответствующих элементов *другой* строки равна нулю.

Теорема 2 является частным случаем более общего утверждения — теоремы Лапласа[v]. Прежде чем сформулировать теорему Лапласа, рассмотрим матрицу, в которой верхний правый блок размера $p\times(n-p)$ представляет собой нулевую матрицу. Такую матрицу удобно изображать в так называемом блочном виде

$$A=\begin{pmatrix} A_1 & O \\ A_2 & A_3 \end{pmatrix}, \tag{18}$$

где O — нулевая матрица типа $p\times(n-p)$, а A_1, A_2, A_3 — произвольные матрицы типа $p\times p, (n-p)\times p, (n-p)\times(n-p)$ соответственно. По определению имеет место равенство

$$\det A=\sum_{\sigma\in S_n}\operatorname{sgn}\sigma\, a_{1\sigma(1)}a_{2\sigma(2)}\dots a_{n\sigma(n)}.$$

Произведение $a_{1\sigma(1)}a_{2\sigma(2)}\dots a_{n\sigma(n)}$ будет заведомо нулевым, если хотя бы для одного элемента $i\in\{1,2,\dots,p\}$ $\sigma(i)\notin\{1,2,\dots,p\}$. Поэтому в сумме, определяющей $\det A$, достаточно рассмотреть только те перестановки из S_n, для которых $\sigma(i)\in\{1,2,\dots,p\}$ для всех $i\in\{1,2,\dots,p\}$. Нетрудно понять, что *каждая* такая перестановка может быть единственным способом представлена в виде произведения $\tau\rho$, где $\tau\in S_p, \rho\in S_{n-p}$. При этом ρ является биекцией множества $\{p+1, p+2,\dots,n\}$ на себя. Поскольку $\operatorname{sgn}(\tau\rho)=\operatorname{sgn}\tau\operatorname{sgn}\rho$, то

$$\begin{aligned}\det A&=\sum_{\tau\in S_p,\rho\in S_{n-p}}\operatorname{sgn}(\tau)\operatorname{sgn}(\rho)a_{1\tau(1)}\dots a_{p\tau(p)}a_{p+1\,\rho(p+1)}\dots a_{n\,\rho(n)}=\\&=\sum_{\tau\in S_p}\operatorname{sgn}\tau\, a_{1\tau(1)}\dots a_{p\tau(p)}\sum_{\rho\in S_{n-p}}\operatorname{sgn}\rho\, a_{p+1\,\rho(p+1)}\dots a_{n\,\rho(n)}=\det A_1\det A_3.\end{aligned}$$

Таким образом, определитель матрицы вида $\begin{pmatrix} A_1 & O \\ A_2 & A_3 \end{pmatrix}$, где A_1, A_3 — квадратные матрицы, равен произведению определителей матриц A_1 и A_3:

$$\det\begin{pmatrix} A_1 & O \\ A_2 & A_3 \end{pmatrix}=\det A_1\det A_3. \tag{19}$$

Формула (19) заслуживает того, чтобы её запомнили. И не только из-за простоты. Это соотношение часто используется в вопросах алгебраической независимости функций, при изучении свойств неявных функций, нелинейных заменах координат и т. п.

Теорема 4. (Лаплас) Фиксируем p строк матрицы A. Тогда сумма произведений миноров порядка p, принадлежащих этим строкам, на их алгебраические дополнения равна определителю матрицы A.

Доказательство.

В частном случае, когда матрица имеет вид (18), теорема уже доказана: равенство (19) можно переписать в виде $\det A = M_{12\ldots p}^{12\ldots p}\overline{M}_{12\ldots p}^{12\ldots p} = M_{12\ldots p}^{12\ldots p}A_{12\ldots p}^{12\ldots p}$.

Рассмотрим теперь общий случай. Пусть $i_1, i_2, \ldots, i_p$, где $i_1 < i_2 < \ldots < i_p$, — номера фиксированных строк. В разложение (17) определителя матрицы A входят произведения элементов минора $M_{j_1 j_2 \ldots j_p}^{i_1 i_2 \ldots i_p}$ на элементы дополнительного минора $\overline{M}_{j_1 j_2 \ldots j_p}^{i_1 i_2 \ldots i_p}$, где $j_1 < j_2 < \ldots < j_p$, причём никаких других членов в разложении определителя матрицы A нет. Чтобы вычислить знак при этих произведениях, переставим строки и столбцы так, чтобы минор $M_{j_1 j_2 \ldots j_p}^{i_1 i_2 \ldots i_p}$ оказался в левом верхнем углу матрицы и относительное расположение строк и столбцов дополнительного минора $\overline{M}_{j_1 j_2 \ldots j_p}^{i_1 i_2 \ldots i_p}$ осталось неизменным. Для этого нужно совершить $(i_1-1)+(i_2-1)+\ldots+(i_p-1)+(j_1-1)+(j_2-1)+\ldots+(j_p-1)$ перестановок, в результате которых определитель матрицы умножится на $(-1)^{i_1+\ldots+i_p+j_1+\ldots+j_p}$. Следовательно для каждого набора $j_1, j_2, \ldots, j_p$, где $j_1 < j_2 < \ldots < j_p$, разложение выражения $(-1)^{i_1+\ldots+i_p+j_1+\ldots+j_p}\det A$ содержит слагаемое $M_{j_1 j_2 \ldots j_p}^{i_1 i_2 \ldots i_p}\overline{M}_{j_1 j_2 \ldots j_p}^{i_1 i_2 \ldots i_p}$ (см. текст перед теоремой 4), а потому разложение (17) определителя матрицы A содержит слагаемое

$$(-1)^{i_1+\ldots+i_p+j_1+\ldots+j_p} M_{j_1 j_2 \ldots j_p}^{i_1 i_2 \ldots i_p}\overline{M}_{j_1 j_2 \ldots j_p}^{i_1 i_2 \ldots i_p} = M_{j_1 j_2 \ldots j_p}^{i_1 i_2 \ldots i_p} A_{j_1 j_2 \ldots j_p}^{i_1 i_2 \ldots i_p}.$$

Для завершения доказательства остаётся только перегруппировать слагаемые в сумме (17) так, чтобы в каждую группу входили произведения элементов матрицы с одним и тем же набором индексов $j_1, j_2, \ldots, j_p$. □

Замечание. Разумеется все теоремы этого параграфа останутся верными, если вместо строк определителя рассматривать его столбцы.

§24. Вычисление обратной матрицы. Формулы Крамера

Рассмотрим квадратную матрицу A порядка n и построим по ней другую квадратную матрицу $\widehat{A}$ того же порядка n, поместив в позицию $(i\,j)$ алгебраическое дополнение элемента a_{ji} исходной матрицы. Т. е., если

$$A=\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ . & . & \dots & . \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}, \text{ то } \widehat{A}=\begin{pmatrix} A_1^1 & A_1^2 & \dots & A_1^n \\ A_2^1 & A_2^2 & \dots & A_2^n \\ . & . & \dots & . \\ A_n^1 & A_n^2 & \dots & A_n^n \end{pmatrix}.$$

Матрица $\widehat{A}$ называется матрицей, *присоединённой* к матрице A.

Пример 1. Для матрицы третьего порядка $A=\begin{pmatrix} -1 & 2 & 1 \\ 3 & 1 & 1 \\ -2 & -2 & -3 \end{pmatrix}$ присоединённая матрица $\widehat{A}=\begin{pmatrix} -1 & 4 & 1 \\ 7 & 5 & 4 \\ -4 & -6 & -7 \end{pmatrix}$.

Вычислим произведение матриц A и $\widehat{A}$. Элемент матрицы $A\widehat{A}$, расположенный на позиции $(i\,j)$ равен $\sum_{k=1}^{n} a_{ik} A_k^j$. В предыдущем параграфе было показано, что при $i=j$ значение суммы равно $\det A$, а при $i \neq j$ — нулю. Таким образом, $A\widehat{A}=(\det A)\cdot E$, где E — единичная матрица порядка *n*. К такому же выводу мы придём, рассматривая произведение $\widehat{A}A$.

Рассмотрим некоторые следствия равенств $A\widehat{A}=\widehat{A}A=(\det A)\cdot E$.

1. $\det A=0$

В этом случае $\widehat{A}A=O$. Нетрудно видеть, что равенство, например, первых строк матриц $\widehat{A}A$ и O можно переписать в виде векторного равенства

$$A_1^1 \boldsymbol{a}_{1*} + A_1^2 \boldsymbol{a}_{2*} + \ldots + A_1^n \boldsymbol{a}_{n*} = \mathbf{0},$$

где $\boldsymbol{a}_{i*}$ — i-я строка матрицы A. Если алгебраическое дополнение хотя бы одного элемента первой строки матрицы A отлично от нуля, то последнее векторное равенство означает линейную зависимость строк матрицы A.

Итак, в случае ненулевой матрицы $\widehat{A}$ мы получили ещё одно доказательство трудной части свойства определителей №8, т. е. что из равенства $\det A=0$ следует линейная зависимость строк матрицы A.

2. $\det A \neq 0$

Равенства $A\widehat{A}=\widehat{A}A=(\det A)\cdot E$ означают, что матрица $\frac{1}{\det A}\cdot\widehat{A}$ является обратной к A. Таким образом, если $\det A \neq 0$, то матрица A является обратимой и

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

В следующем параграфе будет доказана *мультипликативность* определителя (с. 52), т. е. что для любых квадратных матриц A и B одного и того же порядка справедливо равенство

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

Применяя это свойство к равенству $AA^{-1} = E$, получим, что $\det A \det A^{-1} = 1$ и, в частности, $\det A \neq 0$.

В итоге мы можем считать доказанным следующий критерий существования матрицы, обратной к данной.

Теорема 1. Матрица A обратима тогда и только тогда, когда $\det A \neq 0$.

Квадратную матрицу с ненулевым определителем называют также *невырожденной.* Теорема 1 означает, что все невырожденные матрицы обратимы и только они.

В §3 (§7) мы получили так называемые формулы Крамера, выражающие решение системы двух (трёх) линейных уравнений от двух (трёх) неизвестных через определители. Обобщим этот результат.

Рассмотрим систему *n* линейных уравнений с *n* неизвестными

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

матрица коэффициентов которой невырождена. Эта система в матричной записи имеет вид $AX = B$. В силу теоремы 1 матрица A обратима, поэтому матричное равенство $AX = B$ эквивалентно равенству $X = A^{-1}B$. Первая компонента левого вектора равна x_1, а правого — выражению

$$\frac{A_1^1 b_1 + A_1^2 b_2 + \ldots + A_1^n b_n}{det\ A}.$$

Но сумма $A_1^1 b_1 + A_1^2 b_2 + \ldots + A_1^n b_n$ равна (см. теорему 2 и замечание из §23) определителю матрицы, полученной из матрицы A заменой первого столбца столбцом свободных членов. Обозначив эту матрицу через A_1, получим равенство

$$x_1 = \frac{\det A_1}{\det A}.$$

Аналогичные формулы получатся, если приравнивать другие компоненты равенства $X = A^{-1}B$. Итак, доказана

Теорема 2. (Формулы Крамера) Пусть дана система *n* линейных уравнений с *n* неизвестными x_i, матрица A коэффициентов которой невырождена. Если через A_i обозначить матрицу, полученную из A заменой *i*-го столбца столбцом свободных членов, то $x_i = \dfrac{\det A_i}{\det A}$.

§25. Мультипликативность определителя

Каждой квадратной матрице может быть сопоставлено число — её определитель. Мы хотим доказать, что это соответствие мультипликативно, т. е. для любых двух квадратных матриц A и B одного и того же порядка имеет место равенство

$$\det(AB) = \det A \det B. \qquad (20)$$

Соотношение (20) может быть доказано несколькими способами. Можно, например, используя (17), установить требуемое соотношение прямыми выкладками. Другой путь основан на вычислении определителя вспомогательной матрицы порядка $2n$

$$\begin{pmatrix} A & O \\ -E & B \end{pmatrix},$$

имеющей блочный вид, двумя способами. С одной стороны определитель этой матрицы (по теореме Лапласа) равен $\det A \det B$. С другой стороны элементарными преобразованиями столбцов, *не* изменяющих значения определителя, рассматриваемую блочную матрицу можно привести к матрице

$$\begin{pmatrix} A & (A \cdot B) \\ -E & O \end{pmatrix},$$

определитель которой (по той же теореме Лапласа) равен $\det(AB)$.

В настоящем параграфе мы изложим *схему* доказательства мультипликативности определителя, отличное от двух упомянутых. Для того, чтобы читатель мог самостоятельно воспроизвести все недостающие детали, дальнейший текст представлен в виде набора определений и задач. При решении каждой из последних не должно возникать непреодолимых трудностей.

В §15 обсуждалось понятие *линейной* функции $f : \mathbb{R}^n \to \mathbb{R}$. В следующем определении вводится более общее понятие *полилинейной* функции.

Определение 1. Функция f, зависящая от m векторов пространства $\mathbb{R}^n$, называется m-линейной, если она линейна по каждому аргументу.

Например, линейность по первому аргументу означает, что

$$f(\boldsymbol{a}_1'+\boldsymbol{a}_1'',\boldsymbol{a}_2,\dots,\boldsymbol{a}_m)=f(\boldsymbol{a}_1',\boldsymbol{a}_2,\dots,\boldsymbol{a}_m)+f(\boldsymbol{a}_1'',\boldsymbol{a}_2,\dots,\boldsymbol{a}_m),$$
$$f(\lambda\boldsymbol{a}_1,\boldsymbol{a}_2,\dots,\boldsymbol{a}_m)=\lambda f(\boldsymbol{a}_1,\boldsymbol{a}_2,\dots,\boldsymbol{a}_m).$$

Задача 1. Пусть $\boldsymbol{a},\boldsymbol{b}$ — произвольные векторы из $\mathbb{R}^n$. Докажите, что функция $f(\boldsymbol{a},\boldsymbol{b})=a_1b_1+a_2b_2+\dots+a_nb_n$ является 2-линейной (*билинейной*) функцией.

Определение 2. Полилинейная функция $f(\boldsymbol{a}_1,\boldsymbol{a}_2,\dots,\boldsymbol{a}_m)$ называется *кососимметрической,* если при перестановке любых двух аргументов она умножается на -1.

Задача 2. Докажите, что значение кососимметрической функции равно нулю всякий раз, когда какие-либо два аргумента принимают одинаковые значения.

Задача 3. Докажите, что если значение полилинейной функции равно нулю всякий раз, когда какие-либо два аргумента принимают одинаковые значения, то эта функция является кососимметрической.

Указание. Рассмотрите сначала случай билинейной функции. По условию имеет место равенство $f(\boldsymbol{a}+\boldsymbol{b},\boldsymbol{a}+\boldsymbol{b})=0$. Затем воспользуйтесь линейностью функции f по каждому аргументу.

Задача 4. Пусть $\boldsymbol{a},\boldsymbol{b}$ — произвольные векторы из $\mathbb{R}^2$. Докажите, что функция $f(\boldsymbol{a},\boldsymbol{b})=a_1b_2-a_2b_1$ является кососимметрической билинейной функцией.

Всякую матрицу типа $m\times n$ можно рассматривать как упорядоченный набор m строк, рассматриваемых как арифметические векторы из $\mathbb{R}^n$. По этой причине функцию матричного аргумента можно рассматривать как функцию от m векторов из $\mathbb{R}^n$.

Задача 5. Докажите, что определитель является кососимметрической полилинейной функцией строк матрицы.

Пусть $f(\boldsymbol{a}_1,\boldsymbol{a}_2,\dots,\boldsymbol{a}_m)$ — произвольная функция от m векторов из $\mathbb{R}^n$, а σ — произвольная перестановка из S_m. Определим функцию σf равенством

$$(\sigma f)(\boldsymbol{a}_1,\boldsymbol{a}_2,\dots,\boldsymbol{a}_m)=f(\boldsymbol{a}_{\sigma(1)},\boldsymbol{a}_{\sigma(2)},\dots,\boldsymbol{a}_{\sigma(m)}).$$

Задача 6. Докажите, что если $f(\boldsymbol{a}_1,\boldsymbol{a}_2,\dots,\boldsymbol{a}_m)$ — кососимметрическая полилинейная функция, то $(\sigma f)(\boldsymbol{a}_1,\boldsymbol{a}_2,\dots,\boldsymbol{a}_m)=\operatorname{sgn}\sigma\, f(\boldsymbol{a}_1,\boldsymbol{a}_2,\dots,\boldsymbol{a}_m)$.

Указание. Воспользуйтесь теоремой 1, леммой 2 и результатом задачи 4 из §21.

Задача 7. а) Пусть $f(A)=f(\boldsymbol{a}_{1*},\boldsymbol{a}_{2*},\dots,\boldsymbol{a}_{n*})$ — полилинейная функция строк квадратной матрицы A, $\boldsymbol{e}_1,\boldsymbol{e}_2,\dots,\boldsymbol{e}_n$ — канонический базис $\mathbb{R}^n$. Докажите равенство

$$f(A)=\sum_{k_1,k_1,\ldots,k_n} a_{1k_1}a_{2k_2}\ldots a_{nk_n}f(\boldsymbol{e}_{\boldsymbol{k}_1},\boldsymbol{e}_{\boldsymbol{k}_2},\ldots,\boldsymbol{e}_{\boldsymbol{k}_n}),$$

где индексы суммирования $k_1,k_2,\ldots,k_n$ независимо друг от друга пробегают все значения от 1 до *n*.

Указание. Разложите каждую строку матрицы по векторам канонического базиса и воспользуйтесь полилинейностью функции.

б) Пусть функция f из предыдущего пункта является кососиммтерической. Докажите, что $f(A)=\det A\cdot f(E)$, где E — единичная матрица.

Указание. Воспользуйтесь результатами задач 2, 6, 7а и формулой (17).

Задача 8. Пусть B — фиксированная матрица порядка *n*. Докажите, что функция $f(A)=\det(AB)$ является кососимметрической полилинейной функцией строк матрицы A.

Задача 9. Чему равно $f(E)$ для функции f из предыдущей задачи?

Задача 10. Докажите равенство $\det(AB)=\det A\det B$.

Добавление. Принцип математической индукции

В курсе математического анализа множество $\mathbb{N}$ натуральных чисел определяется как пересечение всех индуктивных подмножеств множества $\mathbb{R}$, содержащих единицу. Другими словами $\mathbb{N}$ является *наименьшим* индуктивным подмножеством в $\mathbb{R}$. Напомним, что непустое подмножество M в $\mathbb{R}$ называется индуктивным, если из включения $x \in M$ следует, что $x+1 \in M$.

Итак, согласно определению, множество $\mathbb{N}$ натуральных чисел содержит единицу и вместе с каждым элементом n содержит $n+1$. Это означает, что элементами множества $\mathbb{N}$ являются числа $1, 1+1=2, 2+1=3$ и т. д., что согласуется с интуитивным представлением о множестве натуральных чисел.

Пусть M — подмножество множества $\mathbb{N}$. *Принцип математической индукции* утверждает, что если

1) $1 \in M$;

2) $n \in M \Rightarrow n+1 \in M$,

то M совпадает с $\mathbb{N}$.

Справедливость сформулированного утверждения очевидна. В самом деле, по условию, имеет место включение $M \subset \mathbb{N}$. С другой стороны, множество $\mathbb{N}$, являясь пересечением всех индуктивных подмножеств в $\mathbb{R}$, содержится в каждом из последних. В частности, справедливо включение $\mathbb{N} \subset M$. Из соотношений $M \subset \mathbb{N}$ и $\mathbb{N} \subset M$ следует, что $M = \mathbb{N}$.

Принцип математической индукции применяется к доказательству математических утверждений следующим образом. Предположим, что в формулировке предложения P используется натуральный параметр n. Множество значений этого параметра, при котором предложение P верно, обозначим через M. Если P верно при $n=1$ и из справедливости P при $n=k$ вытекает его справедливость при $n=k+1$ (т. е. M — индуктивное множество), то M совпадает с $\mathbb{N}$. Другими словами, предложение P справедливо при любом значении параметра n.

Пример 1. Покажем, что при любых $n \in \mathbb{N}$ число $10^{2n-1}+1$ делится на 11 без остатка. Утверждение, очевидно, верно при $n=1$. Предположим, что число $10^{2k-1}+1$ делится на 11. Имеем следующую цепочку равенств

$$10^{2(k+1)-1}+1=10^{2k+1}+1=10^{(2k-1)+2}+1=100\cdot 10^{2k-1}+1=99\cdot 10^{2k-1}+\left(10^{2k-1}+1\right),$$

из которой вытекает, что $10^{2(k+1)-1}+1$ также делится на 11.

Пример 2. Покажем, что для всех $n \in \mathbb{N}$ справедливо неравенство $2^n > n$. Рассуждение проведём методом математической индукции по параметру n.

При $n=1$ неравенство, очевидно, выполняется. Допустим, что справедливо неравенство $2^k > k$ и из этого допущения выведем справедливость неравенства $2^{k+1} > k+1$:

$$2^{k+1}=2\cdot 2^k>2k\geqslant k+1\,.$$

Неравенство $2k\geqslant k+1$ справедливо постольку, поскольку оно эквивалентно неравенству $k\geqslant 1$, справедливому при всех натуральных k.

Часто используется несколько иная формулировка принципа математической индукции. Именно, если $1\in M\subset\mathbb{N}$ и из включения $\{1,2,\dots,n-1\}\subset M$ следует, что $n\in M$, то $M=\mathbb{N}$.

Данная формулировка была использована в §10 при доказательстве теоремы.

Габриэль Крамер

Иоганн Карл Фридрих Гаусс

Пьер Симон,
маркиз де Лаплас

i Габриэль Крамер (31.07.1704-4.011752) — швейцарский математик, один из создателей линейной алгебры.

ii Иоганн Карл Фридрих Гаусс (30.04.1777-23.02.1855) — немецкий математик, механик, физик и астроном. Считается одним из величайших математиков всех времён. Иностранный член Шведской, Российской Академий наук и английского Королевского общества.

iii Леопольд Кронекер (7.12.1823-29.12.1891) — немецкий математик, член Берлинской Академии наук, иностранный член Петербургской Академии наук. Основные труды по алгебре и теории чисел. Был сторонником «арифметизации» математики, которая по его мнению, должна быть сведена к арифметике целых чисел. Следующее его выражение стало знаменитым: *Die ganzen Zahlen hat der liebe Gott gemacht, alles andere ist Menschenwerk.*

iv Альфредо Капелли (5.08.1855-28.01.1910) — итальянский математик, член Национальной академии наук деи Линчеи.

v Пьер-Симóн, маркиз де Лаплáс (23.03.1749-5.03.1827) — французский математик, механик, физик и астроном; известен работами в области небесной механики, дифференциальных уравнений, один из создателей теории вероятностей. Занимая пост министра внутренних дел, внёс в управление, по словам Наполеона, «дух бесконечно малых». Был приверженцем абсолютного детерминизма, утверждая, что если бы какое-нибудь разумное существо (демон Лапласа) смогло узнать положения и скорости всех частиц в мире в некий момент, оно могло бы совершенно точно предсказать все будущие и прошедшие мировые события. Философские взгляды Лапласа выразительно характеризует следующий диалог с Наполеоном:

— Вы написали такую огромную книгу [Изложение системы мира] о системе мира и ни разу не упомянули о его Творце!

— Сир, я не нуждался в этой гипотезе.