The Hitchhiker’s Guide to the Linear Algebra Galaxy

Виктор Лопаткин

25 ноября 2025 г.

“...auf der Suche die Wahrheit sich tänzen lässt.”

F. Nietzsche[*]

Man wird nur schlauer,
wenn Man gegen schlauere Gegner spielt.

Die Grundlagen des Schachs, 1883[†]

Alles hervorragende ist ebenso
schwierig als selten.

B.Spinoza[‡]

Was man nachschlagen kann,
braucht Man sich nicht zu merken.

A.Einstein[§]

Wo viel Gefühl ist, ist auch viel Leid.

Leonardo da Vinci.[¶]

Jeder, der aufhört zu lernen, ist alt,
ob mit zwanzig oder achtzig.
Jeder, der weiter lernt, bleibt jung.

H. Ford.[‖]

---

[*] “В поисках правды можно танцевать”. Ф. Ницше.

[†] *“Единственный способ стать умнее это играть с более умным противником.”* Основы шахмат, 1883.

[‡] *“Всё превосходное столь же сложно, как и редко.”* Бенедикт Спиноза.

[§] *“Не нужно запоминать ту информацию, которую можно найти.”* Альберт Эйнштейн.

[¶] *“Там, где много чувств, много и страданий.”* Леонардо да Винчи.

[‖] *“Каждый, кто перестаёт учиться, — стареет, неважно, в 20 или 80 лет. А любой, кто продолжает учиться, остаётся молодым.”* Генри Форд.

Оглавление

Часть I

Теория

Глава 1

Системы линейных уравнений

1.1 Лекция #1. Линейные уравнения

Мы начнём с линейных уравнений и систем линейных уравнений и в конце секции предъявим метод Гаусса для решений линейных системы.

1.1.1 Основные понятия

Определение 1.1.1. *Линейное уравнение* с неизвестными $x_1, \dots, x_n$ это уравение вида

$$\alpha_1 x_1 + \cdots \alpha_n x_n = \beta,$$

где $\alpha_1, \dots, \alpha_n$,$\beta$ – какие-то заданные числа.

Пример 1.1 Уравнения

$$4x_1 - 5x_2 + 2 = x_3, \qquad x_2 = (\sqrt{15} - x_1) + x_3$$

являются линейными, в то время как уравнения

$$4x_1 - 5x_2 + 2 - x_1 x_3 = 10, \qquad x_1 - \sqrt{x_2} + 2x_5^3 = 1, \qquad \sin(x_1) + \cos(x_2) = x_3,$$

линейными не являются.

Определение 1.1.2. *Система линейных уравнений* или *линейная система* это набор одного или более чем одного линейных уравений.

Пример 1.2 Система

$$\begin{cases} 2x_1 - x_2 + 1.5x_3 = 8 \\ x_1 - 4x_3 = -7 \end{cases}$$

является линейной, а система

$$\begin{cases} 2x_1 - x_2 + 1.5x_3 = 8 \\ x_1^2 - 4x_3 = -7 \end{cases}$$

не является линейной.

Систему линейных уравнений с неизвестными $x_1, \dots, x_n$ удобно записывать следующим образом

$$\begin{cases} a_{11}x_1 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ \vdots & & \ddots & & \vdots & & \vdots \\ a_{m1}x_1 & + & \cdots & + & a_{mn}x_n & = & b_m \end{cases}$$

Таким образом, индекс i у коэффициента a_{ij} показывает номер уравнения, а второй индекс это номер неизвестной, в данном случае это x_j.

Определение 1.1.3. Пусть дана линейная система

$$\begin{cases} a_{11}x_1 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ \vdots & & \ddots & & \vdots & & \vdots \\ a_{m1}x_1 & + & \cdots & + & a_{mn}x_n & = & b_m \end{cases}$$

Решением системы называется набор чисел $(s_1, s_2, \dots, s_n)$, такой что, при подстановки

$$x_1 \mapsto s_1, \quad x_2 \mapsto s_2, \dots, x_n \mapsto s_n$$

в систему, мы получаем тождественные равенства.

Пример 1.3 Например, решением системы

$$\begin{cases} 2x_1 - x_2 + 1.5x_3 = 8 \\ x_1 - 4x_3 = -7 \end{cases}$$

является набор $(5, 6.5, 3)$. Действительно, подставляя получаем

$$\begin{cases} 2 \cdot 5 - 6.5 + 1.5 \cdot 3 = 8 \\ 5 - 4 \cdot 3 = -7. \end{cases}$$

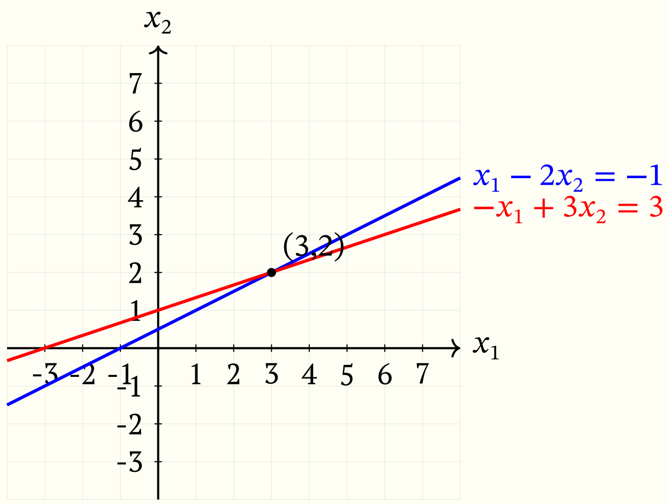
Напомним, что решения линейного уравнения $ax_1 + bx_2 = c$ можно геометрически понимать как прямую на плоскости, где введена прямоугольная система координат x_1Ox_2.

Рассмотрим следующие примеры.

Пример 1.4 Нетрудно проверить, что система

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

имеет решение $x_1 = 3$, $x_2 = 2$. При этом это решение единственно. Геометрически это означает, что соответствующие прямые пересекаются ровно в одной точке.

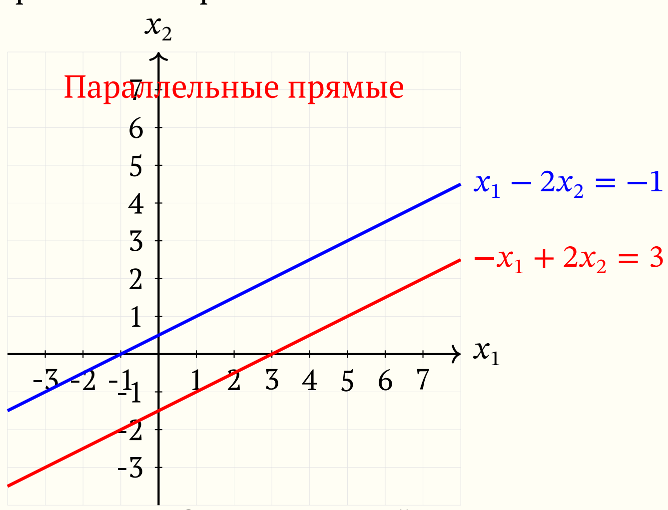


Пример 1: Единственное решение

Пример 1.5 Рассмотрим теперь систему уравнений

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

Если сложим эти два уравнения, то получим, что $0x_1 + 0x_2 = 2$, *т.е.*, $0 = 2$. Это значит, что система вообще не имеет решений, а геометрически это означает, что соответствующие прямые не пересекаются.

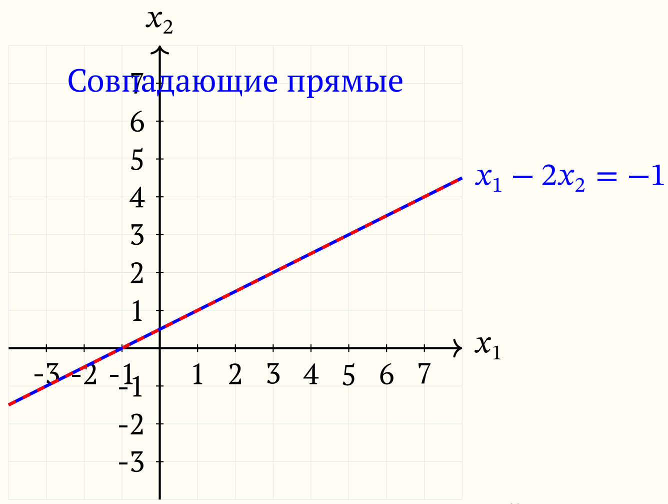


Пример 2: Нет решений

Пример 1.6 Наконец, рассмотрим систему

$$\begin{cases} x_1 - 2x_2 = -1 \\ -x_1 + 2x_2 = 1 \end{cases}$$

Видно, что это одно и то же уравнение, где второе – это просто первое с противоположным знаком. Эта система имеет бесконечное число решений, которые можно записать так: $x_2 = a$, $x_1 = 2a - 1$, где a – произвольное число. Геометрически это значит, что мы получили вместо двух прямых одну.



Пример 3: Бесконечно много решений

Важное замечание. Итак, мы получаем, что линейная система

1. может иметь только одно решение
2. вообще не иметь решений
3. иметь бесконечное число решений.

1.1.2 Матричное обозначение

Основную информацию о линейной системе можно записать в более компактном и удобном виде, используя таблицу, которую принято называть *матрицей*.

Пример 1.7 Расмотрим линейную систему

$$\begin{cases} x_1 - 2x_2 + x_3 = 0 \\ 2x_2 - 8x_3 = 8 \\ 5x_1 - 5x_3 = 10 \end{cases}$$

Итак, перед нами линейная система, в которой три уравнения и три неизвестных x_1,x_2,x_3. Чтобы лучше понять принцип построения матрицы линеной систему, мы запишем эту систему следующим образом

$$\begin{cases} x_1 - 2x_2 + x_3 = 0 \\ 0x_1 + 2x_2 - 8x_3 = 8 \\ 5x_1 + 0x_2 - 5x_3 = 10 \end{cases}$$

Прежде всего, запишем *матрицу коэффициентов* этой системы, где участвуют лишь коэффициенты при неизвестных

$$\begin{pmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ 5 & 0 & -5 \end{pmatrix}$$

а матрица самой системы, которая ещё называется *расширенной матрицей*, имеет вид

$$\begin{pmatrix} 1 & -2 & 1 & \vdots & 0 \\ 0 & 2 & -8 & \vdots & 8 \\ 5 & 0 & -5 & \vdots & 10 \end{pmatrix}$$

где точки символизируют знак = в системе.

1.1.3 Решение линейных систем

Мы покажем на примере метод решения линейных систем, который называется *методом Гаусса*.

Прежде всего заметим, что если у нас есть какое-то уравнение, то, умножив его на како-то число, мы не изменим первоначальное решение. Если у нас есть два уравнения и мы знаем решения каждого из них, то когда мы их сложим, то у полученного уравнения сумма решений будет совпадать с суммой предыдущих.

Пример 1.8 Рассмотрим линейную систему

$$\begin{cases} x_1 - 2x_2 + x_3 = 0 \\ 2x_2 - 8x_3 = 8 \\ 5x_1 - 5x_3 = 10 \end{cases}$$

Для удобства мы пронумеруем уравнения, нумеровать будем свехру вниз, *т.е.,* имеем

$$\begin{aligned} \text{уравнение №1 } &:= \{x_1 - 2x_2 + x_3 = 0\} \\ \text{уравнение №2 } &:= \{2x_2 - 8x_3 = 8\} \\ \text{уравнение №3 } &:= \{5x_1 - 5x_3 = 10\} \end{aligned}$$

Итак, мы хотим сохранить x_1 в первом уравнении и избавиться от него во втором и третьем уравнениях. Но во втором его и так нет, поэтому нужно избавиться от него только в третьем. Для этого мы умножим первое на -5 и прибавим к третьему

$$\begin{array}{c} -5 \cdot [\text{уравнение №1}] \\ +\,[\text{уравнение №3}] \\ \hline [\text{новое уравнение №3}] \end{array} \qquad \begin{array}{r} -5x_1 + 10x_2 - 5x_3 = 0 \\ +5x_1 + 0x_2 - 5x_3 = 10 \\ \hline 0x_1 + 10x_2 - 10x_3 = 10 \end{array}$$

Таким образом мы получили новую линейную систему

$$\begin{cases} x_1 - 2x_2 + x_3 = 0 \\ 2x_2 - 8x_3 = 8 \\ 10x_2 - 10x_3 = 10 \end{cases} \qquad \begin{pmatrix} 1 & -2 & 1 & \vdots & 0 \\ 0 & 2 & -8 & \vdots & 8 \\ 0 & 10 & -10 & \vdots & 10 \end{pmatrix}$$

Чтобы во втором уравнии при коэффициент при x_2 был равен 1 нужно всё второе уравнение умножить на $\frac{1}{2}$ и тогда мы получаем

$$\begin{cases} x_1 - 2x_2 + x_3 = 0 \\ x_2 - 4x_3 = 4 \\ 10x_2 - 10x_3 = 10 \end{cases} \qquad \begin{pmatrix} 1 & -2 & 1 & \vdots & 0 \\ 0 & 1 & -4 & \vdots & 4 \\ 0 & 10 & -10 & \vdots & 10 \end{pmatrix}$$

Теперь нам осталось избавиться от x_2 в третьем уравнеии. Для этого мы умножаем второе уравнение на -10 и прибавляем его к третьему

$$\begin{array}{c} -10 \cdot [\text{уравнение №2}] \\ +\,[\text{уравнение №3}] \\ \hline [\text{новое уравнение №3}] \end{array} \qquad \begin{array}{c} 0x_1 - 10x_2 + 40x_3 = -40 \\ 0x_1 + 10x_2 - 10x_3 = 10 \\ \hline 0x_1 + 0x_2 + 30x_3 = -30 \end{array}$$

Таким образом мы получаем следующую линейную систему

$$\begin{cases} x_1 - 2x_2 + x_3 = 0 \\ \qquad x_2 - 4x_3 = 4 \\ \qquad\quad 30x_3 = -30 \end{cases} \qquad \begin{pmatrix} 1 & -2 & 1 & \vdots & 0 \\ 0 & 1 & -4 & \vdots & 4 \\ 0 & 0 & -30 & \vdots & -30 \end{pmatrix}$$

и умножим третье уравнение на $\frac{1}{30}$ мы получаем

$$\begin{cases} x_1 - 2x_2 + x_3 = 0 \\ \qquad x_2 - 4x_3 = 4 \\ \qquad\qquad x_3 = -1 \end{cases} \qquad \begin{pmatrix} 1 & -2 & 1 & \vdots & 0 \\ 0 & 1 & -4 & \vdots & 4 \\ 0 & 0 & 1 & \vdots & -1 \end{pmatrix}$$

Полученная система называется *верхнетруегольной*, также называется и матрица. Мы уже можем решить, данную систему. Действительно, из третьего уравнения мы имеем $x_3 = -1$, тогда подставляя это значение в первое и второе уравнения вместо x_3 мы будем иметь

$$\begin{cases} x_1 - 2x_2 + -1 = 0 \\ \qquad x_2 + 4 = 4 \end{cases} \iff \begin{cases} x_1 - 2x_2 = 1 \\ \qquad x_2 = 0 \end{cases}$$

из второго мы получаем, что $x_2 = 0$ и подставляя в первое мы получаем

$$x_1 - 2 \cdot 0 = 1 \iff x_1 = 1.$$

Таким образом система имеем решение $x_1 = 1, x_2 = 0, x_3 = -1$.

На этом примере мы видим как операции с уравнениями системы выглядят на языке соответствующей расширенной матрицы. Мы, таким образом, имеем три операции на строках матицы.

Определение 1.1.4. *Элементарными преобразованиями строк* в произвольной матрице мы называем следующие операции

(1) Прибавить к любой строке матрице другую умноженную на произвольное число.

(2) Поменять местами строки.

(3) Умножить строку матрицы на произвольное число.

Две матрицы M,M' будем называть *эквивалентными* если матрица M' получается из матрицы M с помощью **конечной** комбинацией элементарных преобразований.

Важное замечание. Разумеется можно ввести аналогичные операции и над столбцами матрицы, в таком случае говорят либо о *столбцевой эквивалентности матриц*, либо о *строчной эквивалентности матрицы*. Но мы будем пока рассматривать элементарные преобразования строк. К тому же в терминах линейных систем эти операции вообще говоря смысла не имеют, лишь только перемена столбцов означает перенумеровку переменных.

Теорема 1.1.1. *Если расширенные матрицы двух линейных систем эквивалентны, то эти системы имеют одно и то же множество решений.*

Доказательство. Пусть дана линейная система

$$\begin{cases} a_{11}x_1 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ \vdots & & \ddots & & \vdots & & \vdots \\ a_{m1}x_1 & + & \cdots & + & a_{mn}x_n & = & b_m \end{cases}$$

Для удобства мы обозначим эту система следующим образом

$$\begin{cases} L_1(\mathbf{x}) & = & b_1 \\ \vdots & & \vdots \\ L_i(\mathbf{x}) & = & b_i \\ \vdots & & \vdots \\ L_j(\mathbf{x}) & = & b_j \\ \vdots & & \vdots \\ L_m(\mathbf{x}) & = & b_m \end{cases}$$

Рассмотрим первое элементарное преобразование, *т.е.,* пусть i-ое уравнение $L_i(\mathbf{x}) = b_i$ заменилось на $\lambda \cdot L_j(\mathbf{x}) + L_i(\mathbf{x}) = \lambda \cdot b_j + b_i$. Другими словами мы к i-му уравнению прибавили j-ое уравнение умноженное на число λ.

Имеем

$$\begin{cases} L_1(\mathbf{x}) & = & b_1 \\ \vdots & & \vdots \\ L_i(\mathbf{x}) & = & b_i \\ \vdots & & \vdots \\ L_j(\mathbf{x}) & = & b_j \\ \vdots & & \vdots \\ L_m(\mathbf{x}) & = & b_m \end{cases} \xrightarrow[]{\substack{[\text{уравнение № } i] \\ +\lambda\cdot[\text{уравнение № } j]}} \begin{cases} L_1(\mathbf{x}) & = & b_1 \\ \vdots & & \vdots \\ L_i(\mathbf{x}) + \lambda \cdot L_j(\mathbf{x}) & = & b_i + \lambda \cdot b_j \\ \vdots & & \vdots \\ L_j(\mathbf{x}) & = & b_j \\ \vdots & & \vdots \\ L_m(\mathbf{x}) & = & b_m \end{cases}$$

В новой системе такие же уравнение как и в первоначальной, но лишь уравнение № i имеет вид $L_i'(\mathbf{x}) = b_i'$, где $L_i'(\mathbf{x}) := L_i(\mathbf{s}) + \lambda \cdot L_j(\mathbf{s})$, и $b_i := b_i + \lambda \cdot b_j$

Пусть $\mathbf{s} = (s_1, \dots, s_n)$ – какое-то решение первоначальной системы. Тогда мы имеем тожества

$$L_1(\mathbf{s}) \equiv b_1, \dots, L_m(\mathbf{s}) \equiv b_m.$$

(1) Подставим в новую систему вместо $\mathbf{x}$ набор $\mathbf{s}$. Все уравнения кроме уравнения № i превратятся в такие же тождества, что и в первоначальной. Теперь если мы подставим этот же набор в уравнение № i, то получим

$$\begin{aligned} L_i'(\mathbf{s}) \quad &:= \quad L_i(\mathbf{s}) + \lambda \cdot L_j(\mathbf{s}) \\ &= \quad b_i + \lambda \cdot b_j \\ &=: \quad b_i', \end{aligned}$$

т.е., мы получили тождество $L_i'(\mathbf{s}) \equiv b_i'$, а это и показывает, что $\mathbf{s}$ тоже решение для новой системы.

Итак мы показали, что при первом элементарном преобразованием, решение первоначальной системы будет также решением и новой системы.

(2) Применим теперь к новой системе первое элементарное преобразование вида; прибавим к уравнению № i уравнение № j умноженное на $(-\lambda)$.

$$\begin{cases} L_1(\mathbf{x}) = b_1 \\ \vdots \\ L_i(\mathbf{x}) + \lambda L_j(\mathbf{x}) = b_i + \lambda b_j \\ \vdots \\ L_j(\mathbf{x}) = b_j \\ \vdots \\ L_m(\mathbf{x}) = b_m \end{cases} \xrightarrow{\substack{\left[\text{уравнение № } i\right] \\ +(-\lambda)\cdot\left[\text{уравнение № } j\right]}} \begin{cases} L_1(\mathbf{x}) = b_1 \\ \vdots \\ L_i(\mathbf{x}) + \lambda L_j(\mathbf{x}) + (-\lambda) L_j(\mathbf{x}) = b_i + \lambda b_j - \lambda b_j \\ \vdots \\ L_j(\mathbf{x}) = b_j \\ \vdots \\ L_m(\mathbf{x}) = b_m \end{cases}$$

В таком случаем, мы получаем первоначальную систему, а так как мы до этого уже показали, что при первом элементарном преобразовании решение сохраняется, то любое решение новой системы будет решением первоначальной.

Доказательство теоремы для остальных преобразований аналогично и более простое.

□

1.1.4 Ступенчатый вид матрицы

Здесь мы приводим основной алгоритм решения линейных систем. Прежде всего нам нужно ввести некоторые понятия и определения.

Пусть имеем некоторую матрицу, мы говорим что её строка *ненулевая* если в этой строке есть хотя бы один ненулевой элемент. Аналогично мы называем ненулевые столбцы.

Далее, самый левый ненулевой элемент в ненулевой строке называется *ведущим элементом этой строки.*

Определение 1.1.5. Говорят, что матрица имеет *ступенчатый вид* если выполняются следующие три условия:

1. Номера строк с ведущими элементами строго возрастают.

2. Ведущий элемент каждой ненулевой строки располагается строго правее ведущего элемента в строке, расположенной выше данной

Пример 1.9 Ниже показаны примеры матрицы у которых ступенчатый вид, ведущие элементы помещены в квадратики.

$$\begin{pmatrix} \boxed{1} & 4 & -5 & 0 \\ 0 & \boxed{4} & 0 & 1 \\ 0 & 0 & \boxed{1} & 3 \end{pmatrix}, \qquad \begin{pmatrix} 0 & \boxed{a_{12}} & a_{13} & a_{14} & a_{15} & a_{16} & a_{17} & a_{18} \\ 0 & 0 & 0 & 0 & \boxed{a_{25}} & a_{26} & a_{27} & a_{28} \\ 0 & 0 & 0 & 0 & 0 & \boxed{a_{36}} & a_{37} & a_{38} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \boxed{a_{48}} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Определение 1.1.6. Говорят, что матрица имеет *приведённый ступенчатый вид*, если

()1 она имеет ступенчатый вид

(2) каждый ведущий элемент ненулевой строки – это единица, и он является единственным ненулевым элементом в своём столбце.

Пример 1.10 Вот примеры матриц в приведённом ступенчатом виде

$$\begin{pmatrix} \boxed{1} & 0 & 0 & 0 \\ 0 & \boxed{1} & 0 & 1 \\ 0 & 0 & \boxed{1} & 2 \end{pmatrix}, \qquad \begin{pmatrix} 0 & \boxed{1} & a_{13} & a_{14} & 0 & 0 & a_{17} & 0 \\ 0 & 0 & 0 & 0 & \boxed{1} & 0 & a_{27} & 0 \\ 0 & 0 & 0 & 0 & 0 & \boxed{1} & a_{37} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Теорема 1.1.2. *Любую матрицу с помощью элементарных преобразований можно привести к ступенчатому виду, а также к редуцированному ступенчатому виду.*

Доказательство. Пусть дана матрица $A \in \mathrm{Mat}_{n\times m}(\mathbb{R})$,

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

Если A состоит из нулей, то мы уже имеем ступенчатый вид, поэтому мы будем считать что не все её элементы нулевые.

Пусть $\mathbf{c}_i(A)$ – самый левый ненулевой столбец, если $a_{1j} = 0$, то с помощью перестановки строк добиваемся того, чтобы первый столбец был ненулевой. Поэтому мы можем считать, что $a_{1j} \neq 0$.

Теперь с помощью a_{1j} мы зануляем все элементы в этом столбце которые находятся ниже его; делаем преобразования на строках

$$\mathbf{r}_2(A) \mapsto \mathbf{r}_2(A) - \frac{a_{2j}}{a_{1j}}\mathbf{r}_1(A), \qquad \mathbf{r}_m(A) \mapsto \mathbf{r}_m(A) - \frac{a_{mj}}{a_{1j}}\mathbf{r}_1(A),$$

здесь $\mathbf{r}_i(A)$ означает i-ую строку матрицу A.

Тогда, после таких преобразований наша матрица примет вид

$$\begin{pmatrix} 0 & \dots & 0 & a_{1j} & a_{1,j+1} & \dots & a_{1n} \\ 0 & \dots & 0 & 0 & a'_{2,j+1} & \dots & a'_{2n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & a'_{m,j+1} & \dots & a'_{mn} \end{pmatrix}.$$

Тогда проделаем такую же процедуру, но уже только с матрицей

$$A' := \begin{pmatrix} a'_{2,j+1} & \dots & a'_{2n} \\ \vdots & \ddots & \vdots \\ a'_{m,j+1} & \dots & a'_{mn} \end{pmatrix}.$$

В результате мы получим ступенчатый вид, обозначим эту матрицу через B, и пусть её элементы это b_{ij}.

Далее, пусть $b_{1,j_1}, b_{2,j_2}, \dots, b_{k,j_k}$ – ведущие элементы матрицы B. Тогда поделив первую строку на b_{1,j_1}, вторую на b_{2,j_2}, …, k-ую строку на b_{k,j_k} мы получаем, что в полученной матрице B' ведущие элементы все равны 1.

Наконец, с помощью этих элементов и сложением строк мы добиваемся того, чтобы выше этих ведущих элементов были нули. Это завершает доказательство. □

1.1.5 Метод Гаусса

Из доказанной теоремы вытекает метод решения линейных систем.

Пусть исходная система выглядит следующим образом:

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n &= b_1 \\ \dots \\ a_{m1}x_1 + \dots + a_{mn}x_n &= b_m \end{cases}$$

Тогда, согласно свойству элементарных преобразований над строками, основную матрицу этой системы можно привести к ступенчатому виду (эти же преобразования нужно применять к столбцу свободных членов):

$$\begin{cases} \alpha_{1j_1}x_{j_1} + \alpha_{1j_2}x_{j_2} + \dots + \alpha_{1j_r}x_{j_r} + \dots + \alpha_{1j_n}x_{j_n} &= \beta_1 \\ \alpha_{2j_2}x_{j_2} + \dots + \alpha_{2j_r}x_{j_r} + \dots + \alpha_{2j_n}x_{j_n} &= \beta_2 \\ & \dots \\ \alpha_{rj_r}x_{j_r} + \dots + \alpha_{rj_n}x_{j_n} &= \beta_r \\ 0 &= \beta_{r+1} \\ & \dots \\ 0 &= \beta_m \end{cases}$$

где$\alpha_{1j_1}, \dots, \alpha_{rj_r} \neq 0$.

Тогда переменные $x_{j_1}, \dots, x_{j_r}$ называются *главными переменными*. Все остальные называются *свободными*.

Если хотя бы одно число $\beta_i \neq 0$, где $i > r$, то говорят, что рассматриваемая система *несовместна*, *т.е.,* у неё нет ни одного решения.

Пусть $\beta_i = 0$ для любых $i > r$.

Перенесём свободные переменные за знаки равенств и поделим каждое из уравнений системы на свой коэффициент при самом левом x ($\alpha_{ij_i}, i = 1, \dots, r$ где i— номер строки):

$$\begin{cases} x_{j_1} + \widehat{\alpha}_{1j_2}x_{j_2} + \dots + \widehat{\alpha}_{1j_r}x_{j_r} & = & \widehat{\beta}_1 - \widehat{\alpha}_{1j_{r+1}}x_{j_{r+1}} - \dots - \widehat{\alpha}_{1j_n}x_{j_n} \\ x_{j_2} + \dots + \widehat{\alpha}_{2j_r}x_{j_r} & = & \widehat{\beta}_2 - \widehat{\alpha}_{2j_{r+1}}x_{j_{r+1}} - \dots - \widehat{\alpha}_{2j_n}x_{j_n} \\ & \dots & \\ x_{j_r} & = & \widehat{\beta}_r - \widehat{\alpha}_{rj_{r+1}}x_{j_{r+1}} - \dots - \widehat{\alpha}_{rj_n}x_{j_n} \end{cases}$$

где $\widehat{\beta}_i = \frac{\beta_i}{\alpha_{ij_i}}, \widehat{\alpha}_{ij_k} = \frac{\alpha_{ij_k}}{\alpha_{ij_i}}, i = 1, \dots, r, \quad k = i + 1, \dots, n.$

Если свободным переменным полученной системы придавать все возможные значения и решать новую систему относительно главных неизвестных снизу вверх (*т.е.,* от нижнего уравнения к верхнему), то мы получим все решения этой системы. Так как эта система получена путём элементарных преобразований над исходной системой, то по теореме об эквивалентности при элементарных преобразованиях системы эквивалентны, то есть множества их решений совпадают.

Глава 2

Векторные пространства

На прошлой лекции мы работали с системами линейных уравнений и наборами чисел — векторами-столбцами. Мы видели, что их можно складывать и умножать на числа. Оказывается, эти операции и их свойства — не просто особенность столбцов из чисел. Это универсальный язык, на котором говорят:

(1) Геометрия: направленные отрезки (векторы на плоскости и в пространстве).

(2) Data Science. Наборы данных (например, вектор признаков объекта: (рост, вес, возраст)).

(3) Компьютерная графика. Цвета часто представляют как векторы в пространстве RGB (Red, Green, Blue) или RGBA. Операции над ними – это и есть наложение фильтров, изменение яркости.

(4) Машинное обучение. Веса в нейронной сети, векторы эмбеддингов слов.

(5) Программирование. Массивы фиксированной длины, к которым применяются поэлементные операции.

Линейная алгебра — это наука, которая изучает свойства всех этих объектов единым образом, через аксиомы векторного пространства. Давайте их сформулируем.

2.1 Лекция #2.Понятие векторного пространства

2.1.1 Аксиомы векторного пространства

Определение 2.1.1. Пусть $\mathbf{V}$ – некоторое непустое множество, элементы которого мы будем называть *векторами*, хотя их природа может быть произвольной. Предположим, что любым векторам $\mathbf{a},\mathbf{b} \in \mathbf{V}$ как-то сопоставлен третий вектор, обозначаемый символом $\mathbf{a} + \mathbf{b}$ и называемый *суммой* векторов $\mathbf{a}$ и $\mathbf{b}$. Кроме, того, предположим, что любому числу $\lambda \in \mathbb{R}$ и любому вектору $\mathbf{a} \in \mathbf{V}$ как-то сопоставлен новый вектор, обозначаемый символом $\lambda\mathbf{a}$ и называемый *произведением* вектора $\mathbf{a}$ на число λ. Если при этом выполнены следующие свойства, то множество $\mathbf{V}$ называется *векторным (=линейным) пространством.*

(1) $\mathbf{a} + (\mathbf{b} + \mathbf{c}) = (\mathbf{a} + \mathbf{b}) + \mathbf{c}$, для любых $\mathbf{a},\mathbf{b},\mathbf{c} \in \mathbf{V}$.

(2) $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$, для любых $\mathbf{a}, \mathbf{b} \in \mathbf{V}$.

(3) Существует такой вектор, называемый *нулевым вектором*, $\mathbf{o} \in \mathbf{V}$, что $\mathbf{o} + \mathbf{a} = \mathbf{a}$, для любого $\mathbf{a} \in \mathbf{V}$.

(4) Для любого вектора $\mathbf{a} \in \mathbf{V}$ существует такой вектор $-\mathbf{a} \in \mathbf{V}$, что $\mathbf{a} + (-\mathbf{a}) = \mathbf{o}$.

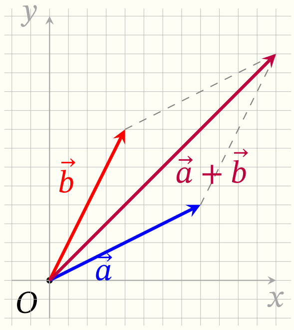
(5) $(\alpha + \beta) \cdot \mathbf{a} = \alpha \cdot \mathbf{a} + \beta \cdot \mathbf{a}$ для любых чисел $\alpha, \beta \in \mathbb{R}$ и любого вектора $\mathbf{a} \in \mathbf{V}$.

(6) $(\alpha\beta) \cdot \mathbf{a} = \alpha \cdot (\beta \cdot \mathbf{a})$ для любых чисел $\alpha, \beta \in \mathbb{R}$ и любого вектора $\mathbf{a} \in \mathbf{V}$.

(7) $\alpha \cdot (\mathbf{a} + \mathbf{b}) = \alpha \cdot \mathbf{a} + \alpha \cdot \mathbf{b}$ для любых векторов $\mathbf{a}$,$\mathbf{b} \in \mathbf{V}$ и любого числа $\alpha \in \mathbb{R}$.

(8) $1 \cdot \mathbf{a} = \mathbf{a}$ для любого вектора $\mathbf{a} \in \mathbf{V}$.

Пример 2.1 Интуитивно-геометрические векторы (=направленные отрезки), о которых говорилось в школе, позволяют простроить несколько различных линейных пространств. Именно, можно рассматривать либо всевозможные векторы в пространстве, либо в плоскости, либо на прямой, при этом они все должны начинаться в одной точке, которая и будет нулевым вектором.



Пример 2.2 Самый важный для пример это $\mathbb{R}^n$, которое есть множество все возможных наборов n чисел, который мы будем записывать так $(\alpha_1, \dots, \alpha_n)^\top$. После введения операций сложения и умножение на число λ следующим образом

$$\begin{aligned}(\alpha_1, \dots, \alpha_n)^\top + (\beta_1, \dots, \beta_n)^\top &:= (\alpha_1 + \beta_1, \dots, \alpha_n + \beta_n)^\top,\\ \lambda \cdot (\alpha_1, \dots, \alpha_n)^\top &:= (\lambda\alpha_1, \dots, \lambda\alpha_n)^\top,\end{aligned}$$

и пологая что $\mathbf{o} := (0, \dots, 0)^\top$, мы получаем векторное пространство.

Пример 2.3 В компьютерной графике цвет часто представляется в формате RGB как тройка чисел (R, G, B), где:

- R — интенсивность красного канала (от 0 до 255)
- G — интенсивность зеленого канала (от 0 до 255)
- B — интенсивность синего канала (от 0 до 255)

Определим операции:

$$(R_1, G_1, B_1) + (R_2, G_2, B_2) = (\min(R_1 + R_2, 255), \min(G_1 + G_2, 255), \min(B_1 + B_2, 255))$$
$$\lambda \cdot (R, G, B) = (\min(\lambda R, 255), \min(\lambda G, 255), \min(\lambda B, 255))$$

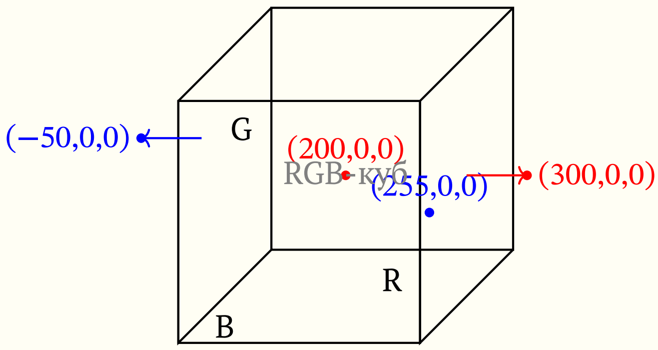
Нулевой вектор — чёрный цвет: $\mathbf{o} = (0, 0, 0)$.
Таким образом, эти операции моделируют:

- Смешение цветов (сложение векторов)
- Изменение яркости (умножение на скаляр)
- Затемнение изображений (умножение на $\lambda < 1$)

Однако, множество цветов с такими операциями не будет векторным векторным пространством, так как нарушается выполнение цветов.

1. **Нарушение замкнутости относительно сложения**:
   Рассмотрим два цвета: $\mathbf{c_1} = (255, 0, 0)$ и $\mathbf{c_2} = (255, 0, 0)$.
   Их сумма: $\mathbf{c_1} + \mathbf{c_2} = (510, 0, 0) \notin [0, 255]^3$
2. **Нарушение замкнутости относительно умножения на скаляр**:
   Возьмем цвет $\mathbf{c} = (200, 0, 0)$ и скаляр $\lambda = 1.5$.
   Их произведение: $\lambda \cdot \mathbf{c} = (300, 0, 0) \notin [0, 255]^3$
3. **Отсутствие обратных элементов**:
   Для цвета $\mathbf{c} = (50, 0, 0)$ обратный элемент должен быть $-\mathbf{c} = (-50, 0, 0)$, но $(-50, 0, 0) \notin [0, 255]^3$

Визуализация проблем



Хотя формально RGB-пространство не является векторным пространством, на практике:

- Для многих операций (смешивание цветов с коэффициентами, дающими сумму 1) можно использовать интуицую из линейной алгебры
- В компьютерной графике часто используются нормализованные значения $[0, 1]$ вместо $[0, 255]$
- Для сложных операций с цветами лучше работать в пространствах типа Lab или XYZ, которые лучше приближают линейное поведение

Пример 2.4 В цифровой обработке звука дискретный сигнал можно представить как вектор $(a_1, a_2, \dots, a_n)$, где:

- a_i — амплитуда сигнала в момент времени i
- n — количество отсчетов (длина сигнала)

Операции определяются покомпонентно:

$$(a_1, \dots, a_n) + (b_1, \dots, b_n) = (a_1 + b_1, \dots, a_n + b_n)$$
$$\lambda \cdot (a_1, \dots, a_n) = (\lambda a_1, \dots, \lambda a_n)$$

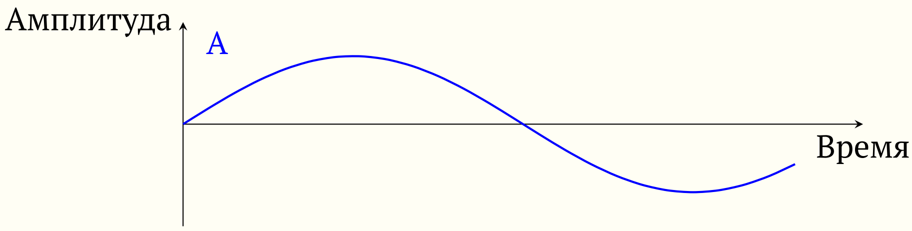
Это пространство используется для:

- Наложения звуковых дорожек (микширование)
- Регулировки громкости (умножение на скаляр)
- Фильтрации шумов (линейные преобразования)
- Создания эхо-эффектов (линейные операции)

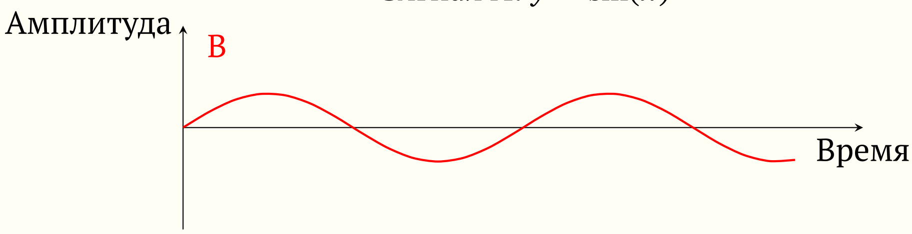
Разберём этот пример более подробно. Итак, звуковые сигналы образуют векторное пространство, где:

- **Векторы**: Дискретные последовательности амплитуд $(a_1, a_2, \dots, a_n)$
- **Сложение**: Поэлементное сложение амплитуд
- **Умножение на скаляр**: Умножение всех амплитуд на константу
- **Нулевой вектор**: Тишина $(0, 0, \dots, 0)$

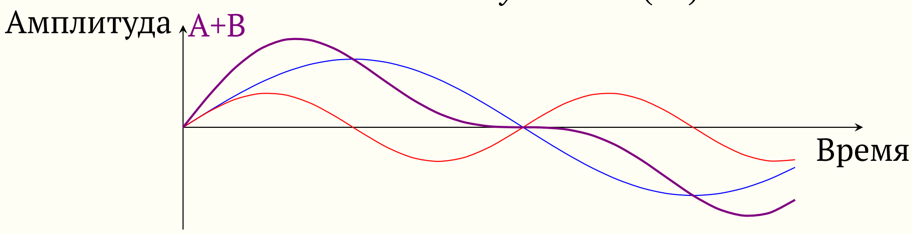
Приведём визуализацию операций с звуковыми сигналами.



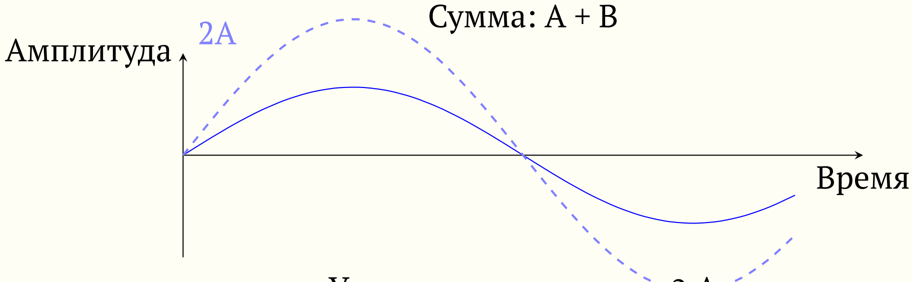
Сигнал A: $y = \sin(x)$



Сигнал B: $y = 0.5\sin(2x)$



Сумма: A + B

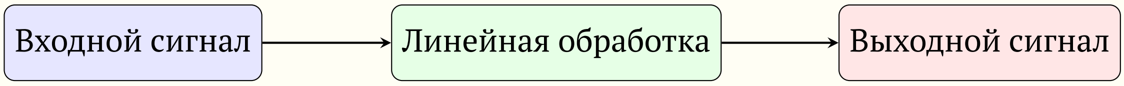


Умножение на скаляр: 2·A

Другими словами, операции (сложение векторов, умножение на скаляр) имеют следующую интерпретацию.

- **Замыкание относительно сложения**: Сумма двух сигналов даёт другой сигнал
- **Замыкание относительно умножения на скаляр**: Усиление/ослабление сигнала даёт другой сигнал
- **Ассоциативность и коммутативность**: Последовательность сложения сигналов не важна
- **Нулевой элемент**: Тишина (отсутствие сигнала)
- **Обратный элемент**: Инвертированный сигнал (противофаза)
- **Смешивание аудиодорожек**: Сложение сигналов
- **Регулировка громкости**: Умножение на скаляр

- **Фильтрация**: Линейные преобразования сигналов
- **Шумоподавление**: Вычитание шумового сигнала
- **Эхо-эффекты**: Линейные комбинации сигналов с задержкой



Применение операций векторного пространства

Итак, в цифровых системах звуковые сигналы представляются как векторы в $\mathbb{R}^n$, где n - количество отсчетов:

$$\mathbf{s} = (s_1, s_2, \dots, s_n)^\top$$

Операции определяются покомпонентно:

$$\begin{aligned} \mathbf{s} + \mathbf{t} &= (s_1 + t_1, s_2 + t_2, \dots, s_n + t_n)^\top \\ \alpha \cdot \mathbf{s} &= (\alpha s_1, \alpha s_2, \dots, \alpha s_n)^\top \end{aligned}$$

Такое представление позволяет применять весь аппарат линейной алгебры для обработки звука.

Пример 2.5 В машинном обучении каждый объект описывается вектором признаков (feature vector). Например, пользователь сайта может быть описан вектором:

$$(user\_age, purchases\_count, session\_duration, click\_rate)^\top$$

Операции определены как в $\mathbb{R}^n$:

$$\begin{aligned} \mathbf{x} + \mathbf{y} &= (x_1 + y_1, \dots, x_n + y_n) \\ \lambda \cdot \mathbf{x} &= (\lambda x_1, \dots, \lambda x_n) \end{aligned}$$

Это позволяет:

- Сравнивать объекты (расстояние между векторами)
- Кластеризовать пользователей (линейные преобразования)
- Нормализовать данные (умножение на скаляр)
- Строить линейные модели классификации

2.1.2 Некоторые следствия из аксиом

Наблюдение 2.1 В любом векторном пространстве $\mathbf{V}$ верны следующие утверждения

(1) вектор $\mathbf{o}$ – единствен,

(2) для любого вектора $\mathbf{a} \in \mathbf{V}$ имеется только один обратный к нему,

(3) $\alpha \cdot \mathbf{o} = \mathbf{o}$,

(4) $\alpha(-\mathbf{v}) = -\alpha\mathbf{v}$

Доказательство. (1) Пусть $\mathbf{o}$,$\mathbf{o}'$ – два нулевых вектора, тогда имеем

$$\mathbf{o}' = \mathbf{o} + \mathbf{o}',$$

потому что $\mathbf{o}$ – нулевой вектор. Но ведь и $\mathbf{o}'$ нулевой, значит $\mathbf{v} + \mathbf{o}' = \mathbf{v}$ для любого $\mathbf{v} \in \mathbf{V}$, в частности $\mathbf{o} + \mathbf{o}' = \mathbf{o}$, *т.е.,* мы получили

$$\mathbf{o}' = \mathbf{o} + \mathbf{o}' = \mathbf{o}.$$

(2) Пусть для вектора $\mathbf{a}$ есть два обратных $\mathbf{b}, \mathbf{c}$, *т.е.,* $\mathbf{a} + \mathbf{b} = \mathbf{a} + \mathbf{c} = \mathbf{o}$.
Имеем

$$\begin{aligned}
\mathbf{c} &= \mathbf{c} + \mathbf{o} \\
&= \mathbf{c} + (\mathbf{a} + \mathbf{b}) \\
&= (\mathbf{c} + \mathbf{a}) + \mathbf{b} \\
&= \mathbf{o} + \mathbf{b} \\
&= \mathbf{b},
\end{aligned}$$

т.е., $\mathbf{c} = \mathbf{b}$, что и требовалось.
(3) Имеем $\mathbf{v} = \mathbf{v} + \mathbf{o}$, с другой стороны существует единственный вектор $-(\alpha\mathbf{v})$ обратный к $\alpha\mathbf{v}$, *т.е.,* $\mathbf{o} = \alpha\mathbf{v} + (-(\alpha\mathbf{v}))$.
Так как $\mathbf{v} = \mathbf{v} + \mathbf{o}$, то $\alpha\mathbf{v} = \alpha\mathbf{v} + \alpha\mathbf{o}$. Далее, получаем

$$\begin{aligned}
\mathbf{o} &= \alpha\mathbf{v} + (-\alpha\mathbf{v}) \\
&= \alpha\mathbf{v} + \alpha\mathbf{o} + (-\alpha\mathbf{v}) \\
&= \alpha\mathbf{v} + (-\alpha\mathbf{v}) + \alpha\mathbf{o} \\
&= \mathbf{o} + \alpha\mathbf{o} \\
&= \alpha\mathbf{o},
\end{aligned}$$

что и требовалось доказать.
(4) Имеем

$$\begin{aligned}
\alpha\mathbf{v} + \alpha \cdot (-\mathbf{v}) + (-\alpha\mathbf{v}) &= \big(\alpha\mathbf{v} + \alpha \cdot (-\mathbf{v})\big) + (-\alpha\mathbf{v}) \\
&= \big(\alpha(\mathbf{v} + (-\mathbf{v}))\big) + (-\alpha\mathbf{v}) \\
&= \alpha\mathbf{o} + (-\alpha\mathbf{v}) \\
&= \mathbf{o} + (-\alpha\mathbf{v}) \\
&= -\alpha\mathbf{v}.
\end{aligned}$$

С другой стороны

$$\begin{aligned}\alpha\mathbf{v}+\alpha\cdot(-\mathbf{v})+(-\alpha\mathbf{v}) &= \big(\alpha\mathbf{v}+(-\alpha\mathbf{v})\big)+\alpha\cdot(-\mathbf{v})\\ &= \mathbf{o}+\alpha\cdot(-\mathbf{v})\\ &= \alpha\cdot(-\mathbf{v})\end{aligned}$$

откуда и следует, что

$$\alpha\cdot(-\mathbf{v})=-\alpha\mathbf{v}.$$

□

Важное замечание. В частности, обратный вектор к вектору $\mathbf{v}$ **всегда** имеет вид $(-1)\cdot\mathbf{v}$.

Определение 2.1.2. Пусть $\mathbf{W}\subseteq\mathbf{V}$ – некоторое непустое подмножество в векторном пространстве $\mathbf{V}$, тогда если

1. $\mathbf{0}_\mathbf{V}\in\mathbf{W}$,
2. для любых $\mathbf{w},\mathbf{w}'\in\mathbf{W}$, верно $\mathbf{w}+\mathbf{w}'\in\mathbf{W}$,
3. для любого $\mathbf{w}\in\mathbf{W}$ и любого числа $\alpha\in\mathbb{R}$ верно $\alpha\cdot\mathbf{w}\in\mathbf{W}$,

то такое подмножество называется *векторным подпространством* пространства $\mathbf{V}$.

Предложение 2.1.1. *Векторное подпространство является векторным пространством относительно тех же операций которые определены в изначальном пространстве.*

Доказательство. Пусть $\mathbf{V}$ – векторное пространство, а $\mathbf{W}$ – его подпространство. Так как $\mathbf{W}\subseteq\mathbf{V}$, то элементы из $\mathbf{W}$ это векторы из $\mathbf{V}$. Но, так как их сумма, согласно определению, должна вновь оказаться в $\mathbf{W}$, то условия (1), (2) определения 2.1.1 тогда выполнены. Далее, так как по условию $\mathbf{o}\in\mathbf{W}$, то выполнена аксиома (3). Наконец, так как для любого $\mathbf{w}\in\mathbf{W}$ и любого числа $\alpha\in\mathbb{R}$ должно быть $\alpha\cdot\mathbf{w}\in\mathbf{W}$, но $\mathbf{W}\subseteq\mathbf{V}$, *т.е.,* мы получаем опять вектор из $\mathbf{V}$, то оставшиеся аксиомы выполнены автоматически. □

Предложение 2.1.2. *Множество всех решений какой-то однородной линейной системы от n переменных есть подпространство в $\mathbb{R}^n$.*

Доказательство. Пусть дана однородная линейная система

$$\begin{cases} a_{11}x_1+\ldots+a_{1n}x_n &= 0\\ \ldots & \\ a_{m1}x_1+\ldots+a_{mn}x_n &= 0\end{cases}$$

и пусть S – множество всех её решений. Докажем, что S есть векторное подпространство в $\mathbb{R}^n$. Так как для однородной системы вектор $\mathbf{o}=(0,\ldots,0)^\top$ (=нулевое решение) является решением, то $\mathbf{o}\in S$.

Далее, пусть $\mathbf{s}=(s_1,\ldots,s_n)^\top$, $\mathbf{s}'=(s_1',\ldots,s_n')^\top$ – решения системы, *т.е.,*

$$\begin{cases} a_{11}s_1+\ldots+a_{1n}s_n &= 0\\ \ldots & \\ a_{m1}s_1+\ldots+a_{mn}s_n &= 0\end{cases}\qquad \begin{cases} a_{11}s_1'+\ldots+a_{1n}s_n' &= 0\\ \ldots & \\ a_{m1}s_1'+\ldots+a_{mn}s_n' &= 0\end{cases}$$

но тогда

$$\begin{cases} a_{11}(s_1+s_1')+\ldots+a_{1n}(s_n+s_n') &= 0\\ \ldots \\ a_{m1}(s_1+s_1')+\ldots+a_{mn}(s_n+s_n') &= 0\end{cases}$$

что означает, что $\mathbf{s}+\mathbf{s}'$ – решение системы. Наконец, для любого числа $\lambda\in\mathbb{R}$, получаем

$$\begin{cases} \lambda(a_{11}s_1+\ldots+a_{1n}s_n) &= 0\\ \ldots \\ \lambda(a_{m1}s_1+\ldots+a_{mn}s_n) &= 0\end{cases} \Longleftarrow \begin{cases} a_{11}(\lambda s_1)+\ldots+a_{1n}(\lambda s_n) &= 0\\ \ldots \\ a_{m1}(\lambda s_1)+\ldots+a_{mn}(\lambda s_n) &= 0\end{cases}$$

т.е., $\lambda\mathbf{s}=(\lambda s_1,\ldots,\lambda s_n)$ – решение. Таким образом, все условия определения 2.1.2 выполнены. Это доказывает предложения. □

2.2 Лекция #3.

2.2.1 Линейная зависимость

Пусть $\mathbf{V}$ – векторное пространство, и $\mathbf{a}_1,\ldots,\mathbf{a}_m$ – некоторое семейство его векторов.

Определение 2.2.1. *Линейной комбинацией векторов* $\mathbf{a}_1,\ldots,\mathbf{a}_m$ с коэффициентами $\alpha_1,\ldots,\alpha_m$ называется вектор

$$\mathbf{v} := \alpha_1\mathbf{a}_1+\cdots+\alpha_m\mathbf{a}_m.$$

Об этом векторе говоря также, что он *линейно выражается через* векторы $\mathbf{a}_1,\ldots,\mathbf{a}_m$.

Пример 2.6 Представьте, что вы описываете пользователя вашего приложения. Каждый пользователь – это вектор, где координаты – это его числовые характеристики (признаки):

- $\mathbf{a}_1$: время, проведённое в приложении (часы)
- $\mathbf{a}_2$: количество совершённых покупок
- $\mathbf{a}_3$: возраст

Тогда линейная комбинация, например, $0.5\cdot\mathbf{a}_1+2\cdot\mathbf{a}_2-0.1\cdot\mathbf{a}_3$, может представлять собой некий *«показатель вовлечённости»*, вычисленный по этим признакам. Все возможные такие показатели образуют линейную оболочку $\mathrm{Span}(\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3)$.

Пример 2.7 Рассмотрим векторное пространство $\mathbf{V}=\mathbb{R}[x]/(x^3)$ – все полиномы степени не выше 2, тогда любой такой полином можно записать так $a_0+a_1x+a_2x^2$, где $a_0,a_1,a_2\in\mathbb{R}$. Это значит, что любой такой полином, есть линейная комбинация векторов $1=x^0,x^1,x^2$.

Пример 2.8 Представьте простейший синтезатор, который генерирует звук. Его **базисные векторы** — это чистые тоны (синусоиды) разной частоты:

- **bass** $= \sin(2\pi \cdot 100 \cdot t)$ (низкая частота, бас)
- **mid** $= \sin(2\pi \cdot 1000 \cdot t)$ (средняя частота)
- **treble** $= \sin(2\pi \cdot 5000 \cdot t)$ (высокая частота, высокие ноты)

Тогда **линейная комбинация** этих базисных сигналов:

$$\mathbf{sound} = 0.8 \cdot \mathbf{bass} + 0.5 \cdot \mathbf{mid} + 0.2 \cdot \mathbf{treble}$$

позволит создать сложный, насыщенный звук, в котором будут слышны и басы, и средние, и высокие частоты. Коэффициенты (0.8, 0.5, 0.2) в этой комбинации — это, по сути, **ручки эквалайзера**, которые усиливают или ослабляют соответствующие частоты. Вся современная цифровая обработка сигналов (DSP) построена на подобных операциях в бесконечномерных пространствах.

Пример 2.9 Вернёмся к примеру с RGB-палитрой. Линейная комбинация базисных цветов позволяет получить любой оттенок.

Например, чтобы получить **ярко-розовый** цвет, нужно много красного, немного синего и совсем немного зелёного:

$$\mathbf{pink} = 0.9 \cdot \mathbf{r} + 0.2 \cdot \mathbf{g} + 0.6 \cdot \mathbf{b} \approx (0.9, 0.2, 0.6)^\top$$

Чтобы получить **тёмно-коричневый**, нужно умеренно и примерно поровну задействовать все каналы:

$$\mathbf{brown} = 0.4 \cdot \mathbf{r} + 0.2 \cdot \mathbf{g} + 0.1 \cdot \mathbf{b} \approx (0.4, 0.2, 0.1)^\top$$

Эти операции лежат в основе работы любого графического редактора. Фильтры, коррекция цвета, наложение слоёв — всё это, по большей части, операции над векторами в цветовом пространстве.

Определение 2.2.2. Множество состоящее из всех линейных комбинаций данных векторов $\mathbf{a}_1, \dots, \mathbf{a}_m$ называется *линейной оболочкой натянутой на векторы* $\mathbf{a}_1, \dots, \mathbf{a}_m$ и обозначается $\mathrm{Span}(\mathbf{a}_1, \dots, \mathbf{a}_m)$.

Таким образом можно записать

$$\mathrm{Span}_{\mathbb{R}}(\mathbf{a}_1, \dots, \mathbf{a}_m) := \{\alpha_1\mathbf{a}_1 + \dots + \alpha_m\mathbf{a}_m\},$$

где $\alpha_1, \dots, \alpha_m$ – все возможные числа из $\mathbb{R}$.

Определение 2.2.3. Подмножество $S \subset \mathbf{V}$ называется *линейно независимым* если для каждого конечного набора векторов из S, $\mathbf{s}_1, \dots, \mathbf{s}_p \in S$ из равенства

$$\alpha_1\mathbf{s}_1 + \dots + \alpha_p\mathbf{s}_p = 0$$

вытекает, что $\alpha_1 = \dots = \alpha_p = 0$.

В случае, если найдутся не все сразу нулевые $\alpha_1, \dots, \alpha_p$, такие что для некоторых векторов $\mathbf{s}_1, \dots, \mathbf{s}_p \in S$ выполняется равенство

$$\alpha_1\mathbf{s}_1 + \dots + \alpha_p\mathbf{s}_p = 0,$$

то, говорят, что S – *линейно зависимое подмножество.*

2.2.2 Понятие линейной оболочки

Пусть теперь $\mathcal{S} \subseteq \mathbf{V}$ – непустое подмножество векторного пространства $\mathbf{V}$, под *линейной комбинацией* элементов из $\mathcal{S}$ понимают сумму вида

$$\alpha_1 \mathbf{s}_1 + \cdots + \alpha_m \mathbf{s}_m$$

где все $\alpha_1, \dots, \alpha_m \in \mathbb{R}$ и при этом сумма берётся лишь у конечного числа векторов из S.

Множество всех таких линейных комбинаций называется *линейной оболочкой натянутой на множество* $\mathcal{S}$ или порождённой множеством $\mathcal{S}$ и обозначается так $\mathrm{Span}_{\mathbb{R}}(\mathcal{S})$.

Пример 2.10 Рассмотрим знакомое трёхмерное пространство. Пусть у нас есть два вектора:

$$\mathbf{v}_1 = (1, 0, 0)^\top, \quad \mathbf{v}_2 = (0, 1, 0)^\top.$$

Их линейная оболочка $\mathrm{Span}(\mathbf{v}_1, \mathbf{v}_2)$ — это множество всех векторов вида:

$$\alpha_1 \cdot \mathbf{v}_1 + \alpha_2 \cdot \mathbf{v}_2 = (\alpha_1, \alpha_2, 0)^\top.$$

Геометрически это представляет собой **плоскость XY** в трёхмерном пространстве. Любая точка на этой плоскости может быть достигнута соответствующей линейной комбинацией наших векторов.

Если же векторы $\mathbf{v}_1$ и $\mathbf{v}_2$ **коллинеарны** (лежат на одной прямой), например, $\mathbf{v}_1 = (1, 0, 0)^\top$ и $\mathbf{v}_2 = (2, 0, 0)^\top$, то их линейная оболочка $\mathrm{Span}(\mathbf{v}_1, \mathbf{v}_2)$ вырождается в **прямую** (ось X), так как все комбинации дадут векторы вида $(\alpha_1 + 2\alpha_2, 0, 0)^\top$.

Пример 2.11 В пространстве полиномов степени не выше 2 рассмотрим два вектора: $\mathbf{p}_1 = 1$ и $\mathbf{p}_2 = x^2$. Их линейная оболочка:

$$\mathrm{Span}(\mathbf{p}_1, \mathbf{p}_2) = \{a \cdot 1 + b \cdot x^2 \mid a, b \in \mathbb{R}\}$$

представляет собой всевозможные **чётные** полиномы степени не выше 2. Это подпространство исходного пространства полиномов. Мы не сможем получить полиномы с членом x (например, x или $2x + 1$), находясь в этой линейной оболочке.

Пример 2.12 В звуковом синтезе линейная оболочка, натянутая на два тона разной частоты, определяет всевозможные тембры, которые можно получить их смешиванием. Пусть:

- **bass** $= \sin(2\pi \cdot 100 \cdot t)$ (низкая частота)
- **lead** $= \sin(2\pi \cdot 800 \cdot t)$ (высокая частота)

Тогда $\mathrm{Span}(\mathbf{bass}, \mathbf{lead})$ — это все звуковые волны вида:

$$a \cdot \mathbf{bass} + b \cdot \mathbf{lead} = a \cdot \sin(2\pi \cdot 100 \cdot t) + b \cdot \sin(2\pi \cdot 800 \cdot t)$$

где $a, b \in \mathbb{R}$ задают громкость каждого тона. Это позволяет создавать различные **звуковые ландшафты —** от глубокого баса с примесью высоких нот до доминирующего высокого тона с лёгкой басовой подложкой.

Пример 2.13 В RGB-пространстве рассмотрим линейную оболочку, натянутую на два цвета:

- $\mathbf{red} = (255, 0, 0)^\top$
- $\mathbf{green} = (0, 255, 0)^\top$

Их линейная оболочка $\mathrm{Span}(\mathbf{red}, \mathbf{green})$ состоит из всех векторов вида:

$$a \cdot \mathbf{red} + b \cdot \mathbf{green} = (255a, 255b, 0)^\top$$

где $a, b \in [0, 1]$ для корректного представления цвета. Это множество представляет собой **все оттенки жёлтого цвета** — от чисто красного через оранжевый и жёлтый до чисто зелёного. Таким образом, меняя коэффициенты a и b, мы можем получить любой цвет на красно-зелёной оси.

Лемма 2.2.1. *Если $\mathcal{S}' \subseteq \mathcal{S}$, то $\mathrm{Span}_{\mathbb{R}}(\mathcal{S}') \subseteq \mathrm{Span}_{\mathbb{R}}(\mathcal{S})$.*

Доказательство. Действительно, так как $\mathcal{S}' \subseteq \mathcal{S}$, то любая линейная комбинация элементов из $\mathcal{S}'$ это линейная комбинация из $\mathcal{S}$. □

Предложение 2.2.2. *Если $S \neq \varnothing$, то $\mathrm{Span}_{\mathbb{R}}(S)$ – векторное пространство.*

Доказательство. Пусть $S = \{\mathbf{s}_1, \dots, \mathbf{s}_m\}$ – некоторое множество векторов из $\mathbf{V}$. Покажем, что $\mathrm{Span}_{\mathbb{R}}(S)$ – векторное подпространство в $\mathbf{V}$. Тогда из предложения 2.1.1 будет вытекать утверждение.

Рассмотрим $\mathbf{v}, \mathbf{w} \in \mathrm{Span}_{\mathbb{R}}(S)$, тогда $\mathbf{v} = \alpha_1\mathbf{s}_1 + \dots + \alpha_m\mathbf{s}_m$, $\mathbf{w} = \beta_1\mathbf{s}_1 + \dots + \beta_m\mathbf{s}_m$. Тогда, согласно определению векторного пространства, $\mathbf{v}, \mathbf{w}$ – тоже векторы пространства $\mathbf{V}$ и для любого числа $\alpha \in \mathbb{R}$, $\alpha \cdot \mathbf{v}$ тоже вектор.

Согласно аксиомам векторного пространства имеем

$$\begin{aligned}
\mathbf{v} + \mathbf{w} &= (\alpha_1\mathbf{s}_1 + \dots + \alpha_m\mathbf{s}_m) + (\beta_1\mathbf{s}_1 + \dots + \beta_m\mathbf{s}_m) \\
&= (\alpha_1\mathbf{s}_1 + \beta_1\mathbf{s}_1) + \dots + (\alpha_m\mathbf{s}_m + \beta_m\mathbf{s}_m) \\
&= (\alpha_1 + \beta_1)\mathbf{s}_1 + \dots + (\alpha_m + \beta_m)\mathbf{s}_m,
\end{aligned}$$

т.е., $\mathbf{v} + \mathbf{w} \in \mathrm{Span}_{\mathbb{R}}(S)$.

Далее, получаем

$$\begin{aligned}
\alpha \cdot \mathbf{v} &= \alpha \cdot (\alpha_1\mathbf{s}_1 + \dots + \alpha_m\mathbf{s}_m) \\
&= \alpha \cdot (\alpha_1\mathbf{s}_1) + \dots + \alpha \cdot (\alpha_m\mathbf{s}_m) \\
& \quad (\alpha\alpha_1)\mathbf{s}_1 + \dots + (\alpha\alpha_m)\mathbf{s}_m,
\end{aligned}$$

т.е., для любого $\mathbf{v} \in \mathrm{Span}_{\mathbb{R}}(S)$ и любого числа $\alpha \in \mathbb{R}$, получаем, что $\alpha \cdot \mathbf{v} \in \mathrm{Span}_{\mathbb{R}}(S)$.

Наконец, ясно что $\mathbf{o} = 0 \cdot \mathbf{s}_1 + \dots + 0 \cdot \mathbf{s}_m$, *т.е.*, $\mathbf{o} \in \mathrm{Span}_{\mathbb{R}}(S)$, а тогда согласно определению 2.1.2, $\mathrm{Span}_{\mathbb{R}}(S)$ – векторное подпространство в $\mathbf{V}$, а тогда согласно предложению 2.1.1, оно есть векторное пространство, что и требовалось доказать. □

Пример 2.14 В компьютерной графике цвет часто представляется в RGB-модели. Векторное пространство цветов натянуто на три «базовых» вектора-цвета:

- $\mathbf{r} = (1, 0, 0)^\top$ (красный)
- $\mathbf{g} = (0, 1, 0)^\top$ (зелёный)
- $\mathbf{b} = (0, 0, 1)^\top$ (синий)

Линейная оболочка $\mathrm{Span}(\mathbf{r}, \mathbf{g}, \mathbf{b})$ – это всё множество цветов, которые можно отобразить на экране. Любой цвет $\mathbf{c}$ представляется в виде линейной комбинации: $\mathbf{c} = r_i \cdot \mathbf{r} + g_i \cdot \mathbf{g} + b_i \cdot \mathbf{b}$, где r_i, g_i, b_i – значения интенсивности каждого канала (от 0 до 1 или от 0 до 255).

Лемма 2.2.3. *Векторы $\mathbf{v}_1, \dots, \mathbf{v}_k$ линейно зависимы если и только если один из них находится в линейной оболочке натянутой на остальные.*

Доказательство. Другими словами тут говориться, что если, например $\mathbf{v}_i \in \mathrm{Span}_\mathbb{R}(S_i)$, где S_i – это все векторы $\mathbf{v}_1, \dots, \mathbf{v}_k$ кроме $\mathbf{v}_i$, то эти векторы линейной зависимы.

(1) Действительно, пусть векторы $\mathbf{v}_1, \dots, \mathbf{v}_k$ линейно зависимы, тогда существуют не все равные нулю числа $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ такие, что

$$\alpha_1 \mathbf{v}_1 + \dots + \alpha_k \mathbf{v}_k = \mathbf{o}.$$

Пусть $\alpha_i \neq 0$, тогда можно записать

$$\mathbf{v}_i = -\frac{\alpha_1}{\alpha_i}\mathbf{v}_1 - \dots - \frac{\alpha_k}{\alpha_i}\mathbf{v}_k,$$

но это и означает, что $\mathbf{v}_i \in \mathrm{Span}_\mathbb{R}(S_i)$.

(2) Пусть $\mathbf{v}_i \in \mathrm{Span}_\mathbb{R}(S_i)$, тогда найдутся такие числа $\beta_1, \dots, \beta_k \in \mathbb{R}$, которые не все нулевые, что

$$\mathbf{v}_i = \beta_1 \mathbf{v}_1 + \dots + \beta_k \mathbf{v}_k,$$

но это можно записать и так

$$-\mathbf{v}_i + \beta_1 \mathbf{v}_1 + \dots + \beta_k \mathbf{v}_k = \mathbf{o},$$

что и означает линейную зависимость. □

Лемма 2.2.4. *Векторы $\mathbf{v}_1, \dots, \mathbf{v}_k$ линейно независимы тогда и только тогда, когда каждый вектор $\mathbf{v}_j$ не принадлежит* $\mathrm{Span}_\mathbb{R}(\mathbf{v}_1, \dots, \mathbf{v}_{j-1})$.

Доказательство.
(1) Пусть векторы $\mathbf{v}_1, \dots, \mathbf{v}_k$ линейно независимы, и пусть $\mathbf{v}_j \in \mathrm{Span}_\mathbb{R}(\mathbf{v}_1, \dots, \mathbf{v}_{j-1})$, но тогда согласно Лемме 2.2.3, векторы $\mathbf{v}_1, \dots, \mathbf{v}_j$ – линейно зависимы, *т.е.*, имеются не все равные числа $\alpha_1, \dots, \alpha_j \in \mathbb{R}$, такие, что

$$\alpha_1 \mathbf{v}_1 + \dots + \alpha_j \mathbf{v}_j = \mathbf{o},$$

но тогда мы получаем равенство

$$\alpha_1 \mathbf{v}_1 + \dots + \alpha_j \mathbf{v}_j + 0 \cdot \mathbf{v}_{j+1} + \dots + 0 \cdot \mathbf{v}_k = \mathbf{o}$$

что означает линейную зависимость векторов $\mathbf{v}_1, \dots, \mathbf{v}_k$, что влечёт противоречие.

(2) Пусть для каждого $1 < j \leq k$, $\mathbf{v}_j \in \mathrm{Span}_\mathbb{R}(\mathbf{v}_1, \dots, \mathbf{v}_{j-1})$, но векторы $\mathbf{v}_1, \dots, \mathbf{v}_k$ – линейно зависимы. Тогда согласно лемме 2.2.3, $\mathbf{v}_k \in \mathrm{Span}_\mathbb{R}(\mathbf{v}_1, \dots, \mathbf{v}_{j-1})$, что противоречит предположению. □

2.3 Лекция #4

2.3.1 Понятие базиса и координаты вектора

Определение 2.3.1. Если $\mathbf{W} = \mathrm{Span}_{\mathbb{R}}(S)$, то говорят, что векторное пространство $\mathbf{W}$ *порождено множеством* S. Если же множество S – линейно независимо, то в таком случае говорят, что S – *базис* векторного пространства $\mathbf{W}$.

Важное замечание. Различают два случая, когда S конечно и когда оно бесконечно, в первом случае говорят, что пространство S конечномерное и имеет размерность n, где n – число элементов в множестве S. Если S – бесконечно, то говорят, что пространство $\mathbf{W}$ бесконечномерно.

Пример 2.15 Вернёмся к примеру с пользователями. Если наши векторы-признаки $\mathbf{a}_1$ (время), $\mathbf{a}_2$ (покупки) и $\mathbf{a}_3$ (возраст) линейно независимы (что скорее всего так и есть, ведь время в приложении не строго определяет возраст или количество покупок), то они образуют базис пространства признаков пользователя. Размерность этого пространства равна 3. Любого пользователя можно однозначно описать координатами в этом базисе: (5 часов, 2 покупки, 30 лет).

Определение 2.3.2. Пусть $\mathbf{e} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ – некоторое конечное множество векторов в некотором векторном пространстве. Пусть $\mathbf{W} = \mathrm{Span}_{\mathbb{R}}(\mathbf{e}_1, \dots, \mathbf{e}_n)$ – векторное пространство натянутое на множество S. Тогда это значит, что для любого вектора $\mathbf{w} \in \mathbf{W}$ можно найти такие числа $\alpha_1, \dots, \alpha_n \in \mathbb{R}$, что

$$\mathbf{w} = \alpha_1\mathbf{e}_1 + \cdots \alpha_n\mathbf{e}_n.$$

Эти числа называются *координатами вектора* $\mathbf{w}$ *в базисе* E и принято это записывать в виде

$$\mathbf{w}_E = (\alpha_1, \dots, \alpha_n)^\top$$

У нулевого вектора $\mathbf{o}$ координаты состоят из нелуй, то есть

$$\mathbf{o} = (0, \dots, 0)^\top.$$

Пример 2.16 Рассмотрим пространство всех полиномов степени не выше 2: $\mathbb{R}_{\leq 2}[x]$.

В этом пространстве можно выбрать **разные базисы**. Самый очевидный — **стандартный базис** «по степеням x»:

$$\mathbf{e}_1 = 1, \quad \mathbf{e}_2 = x, \quad \mathbf{e}_3 = x^2.$$

Любой полином $P(x) = a_0 + a_1x + a_2x^2$ — это линейная комбинация: $a_0 \cdot \mathbf{e}_1 + a_1 \cdot \mathbf{e}_2 + a_2 \cdot \mathbf{e}_3$. Его координаты в этом базисе — $(a_0, a_1, a_2)^\top$.

Но ничто не мешает выбрать другой базис, более удобный для конкретной задачи. Например, базис из **полиномов Лагранжа** крайне полезен для интерполяции и построения кривых в компьютерной графике. Умение работать с разными базисами — ключевой навык.

Пример 2.17

Представьте сложный звуковой сигнал — например, аккорд, сыгранный на гитаре. Этот сигнал можно рассматривать как функцию $s(t)$, описывающую изменение давления воздуха со временем.

В пространстве звуковых сигналов существует специальный **базис**, состоящий из «чистых» тонов — синусоид разной частоты:

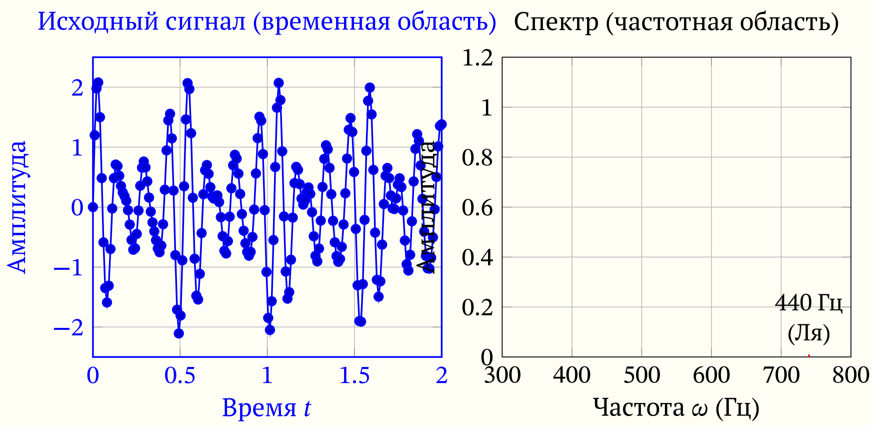
$$\mathbf{e}_\omega(t) = \sin(2\pi\omega t)$$

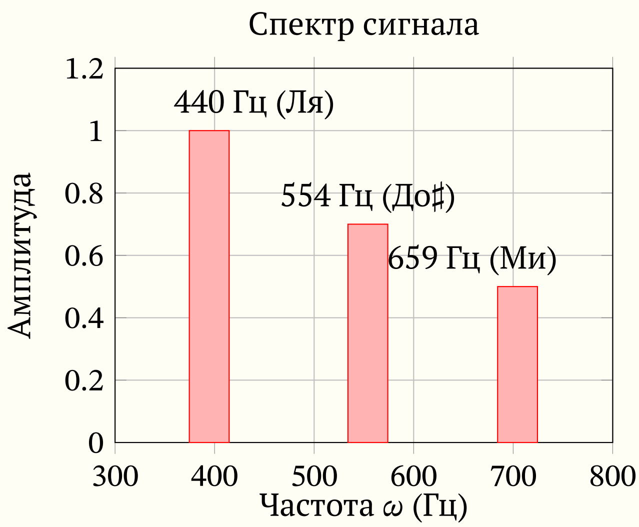
где ω — частота тона (например, 440 Гц для ноты «Ля» первой октавы).

Для работы с такими функциями применяют так называемое преобразование Фурье. Ключевая идея преобразования Фурье заключается в том, что **любой** звуковой сигнал $s(t)$ можно представить как бесконечную линейную комбинацию этих базисных функций:

$$s(t) = \alpha_1 \cdot \mathbf{e}_{\omega_1}(t) + \alpha_2 \cdot \mathbf{e}_{\omega_2}(t) + \alpha_3 \cdot \mathbf{e}_{\omega_3}(t) + \cdots$$

На практике мы работаем с дискретным представлением сигнала, и преобразование Фурье позволяет найти коэффициенты α_i этой линейной комбинации. Эти коэффициенты показывают, **насколько сильно** каждая частота представлена в исходном сигнале.





На рисунке выше:

- Слева — исходный сигнал $s(t)$ (аккорд из трёх нот)
- Справа — коэффициенты α_i разложения этого сигнала по базису синусоид
- Пики на правом графике соответствуют частотам нот, входящих в аккорд

Это представление чрезвычайно полезно на практике:

- **Сжатие audio (MP3)**: Мы можем отбросить коэффициенты α_i, соответствующие частотам, которые плохо воспринимаются человеческим ухом, значительно уменьшив размер файла
- **Фильтрация шума**: Шум обычно проявляется как случайные небольшие коэффициенты across many frequencies. Мы можем подавить их, установив маленькие коэффициенты в ноль
- **Выделение инструментов**: Разные музыкальные инструменты имеют характерные наборы частот. Анализируя коэффициенты, мы можем выделить партию отдельного инструмента из микса
- **Распознавание нот**: Пики на графике коэффициентов прямо указывают на ноты, входящие в аккорд

Таким образом, преобразование Фурье — это по сути **переход в другой базис** пространства сигналов, где многие практические задачи решаются значительно проще.

Лемма 2.3.1. *У данного вектора в данном базисе имеется только единственные координаты, то есть у одного и того же вектора в данном базисе не может быть две разные координаты.*

Доказательство. Пусть это не так, и пусть у вектора $\mathbf{w}$ имееются две координаты, скажем $(\alpha_1, \dots, \alpha_n)^\top$ и $(\alpha'_1, \dots, \alpha'_n)^\top$. Это значит, что имеют место равенства

$$\begin{aligned}\mathbf{w} &= \alpha_1\mathbf{e}_1 + \cdots \alpha_n\mathbf{e}_n \\ \mathbf{w} &= \alpha'_1\mathbf{e}_1 + \cdots \alpha'_n\mathbf{e}_n.\end{aligned}$$

Тогда получаем

$$\begin{aligned}\mathbf{w} - \mathbf{w} &= \alpha_1\mathbf{e}_1 + \cdots \alpha_n\mathbf{e}_n \\ &\quad -\alpha'_1\mathbf{e}_1 + \cdots \alpha'_n\mathbf{e}_n \\ &= (\alpha_1 - \alpha'_1)\mathbf{e}_1 + \cdots + (\alpha_n - \alpha'_n)\mathbf{e}_n\end{aligned}$$

Но $\mathbf{w} - \mathbf{w} = \mathbf{o} = (0, \dots, 0)^\top$, а тогда мы получаем, что

$$\alpha_1 - \alpha'_1 = 0, \dots, \alpha_n - \alpha'_n = 0,$$

т.е.,

$$\alpha_1 = \alpha'_1, \dots, \alpha_n = \alpha'_n,$$

что и доказывает лемму. □

Пусть $\mathbf{a} = (\alpha_1, \dots, \alpha_n)^\top$, $\mathbf{b} = (\beta_1, \dots, \beta_n)^\top$ – два произвольных вектора с координатами в каком-то базисе. Тогда получаем

$$\begin{aligned}\mathbf{a} + \mathbf{b} &= (\alpha_1 + \beta_1, \dots, \alpha_n + \beta_n)^\top \\ \lambda\mathbf{a} &= (\lambda\alpha_1, \dots, \lambda\alpha_n)^\top,\end{aligned}$$

для любого числа $\lambda \in \mathbb{R}$.

Важное замечание. Таким образом, базис можно охарактеризовать как такой набор векторов что каждый векторы однозначно выражается через них.

2.4 Лекция #5.

Итак мы ввели понятие векторного пространства, сейчас наша цель понять это понять когда мы два векторных пространства мы будем считать одинаковыми. Прежде всего нам нужно формально объяснить что значить то, что векторные пространства считаются одинаковыми. Для таких целей вводится особенное понятие под названием *изоморфизм*. Но прежде чем давать это определение нам необходимы некоторые понятия из общей теории множеств.

2.4.1 Иньекция, сюръекция и бекция

В этом отделе мы всегда под множеством понимаем непустое множество. Напомним, что под *отображением* f между двумя множествами A, B понимается правило, которое *любому* элементу $a \in A$ сопоставляет *один и только один* элементы из множества $b \in B$. Мы будем записывать это таким образом

$$f : A \to B, \qquad a \mapsto f(a).$$

Также говорят что задана функция, но для наших целей, нам удобно под функцией понимать отображение в числовое множество.

Пример 2.18 Отображение $f : \mathbb{R} \to \mathbb{R}$ заданное правилом $x \mapsto x^2$ очевидно является отображением, а вот правило $f : \mathbb{R}_+ \to \pm\mathbb{R}$, заданное так $x \mapsto \sqrt{x}$ отображением уже не будет, так как квадратный корень принимает уже два значения.

Определение 2.4.1. Пусть $f : A \to B$ – отображение из множества A в множество B.

(1) f называется *инъективным* (или *иньекцией*), если разные элементы из A переходят в разные элементы B, *если $a \neq a'$, то $f(a) \neq f(a')$, для всех $a,a' \in A$,*

(2) f называется сюръективным (или сюръекцией), если для каждого элемента $b \in B$, существует такое элементы $a \in A$, такой, что $f(a) = b$.

(3) f называется биективным (или биекцией), если оно одновременно инъективно и сюръективно.

Пример 2.19

1. Рассмотрим отображение $f : \mathbb{N} \to \mathbb{N}$, заданное так $f(n) = n^2$, тогда это отображение
   (a) оно иньективно потому что если $n^2 = m^2$ для натуральных n и m, то $n = m$ (так как $n, m > 0$),
   (b) оно оно не сюръективно, потому что, например, число 2 не является квадратом натурального числа.
2. Рассмотрим такое же отображение но уже между множествами $\mathbb{R}_+$ и $\mathbb{R}_+$, *т.е.* $f(x) = x^2$, тогда это отображение
   (a) иньективно, потому что если $x^2 = y^2$ означает что $x = y$ ибо $x,y > 0$,
   (b) сюръективно, так как уравнение $x^2 = b$ имеет единственное решение $\sqrt{b}$, так как $b > 0$,
   (c) таким образом это биекция.
3. Если же рассмотреть такое же отображение между множествами $\mathbb{R}_+$, $\mathbb{R}$, *т.е.* $f : \mathbb{R}_+ \to \mathbb{R}$ заданное также $x \mapsto x^2$, то это отображение
   (a) иньективно, потому что если $x^2 = y^2$ означает что $x = y$ ибо $x,y > 0$,
   (b) но оно не сюръективно, так как уравнение $x^2 = b$ не всегда имеет решение, например, если $b = -1$, то не существует числа $a \in \mathbb{R}_+$, такого, что $a^2 = -1$.

Наблюдение 2.2

Определение 2.4.2. Два множества A и B имеют **одинаковую мощность**, если существует биекция (взаимно-однозначное соответствие) $f : A \to B$.

Определение 2.4.3. Множество называется **счётным**, если оно имеет ту же мощность, что и множество натуральных чисел $\mathbb{N}$.

Определение 2.4.4. Множество называется **несчётным**, если оно не является счётным.

Пример 2.20 Покажем, что не существует биекции между множеством $\mathbb{N}$ и (0,1). Будет рассуждать от противного. Предположим, что существует биекция $f : \mathbb{N} \to (0{,}1) \subseteq \mathbb{R}$.

Каждое число $x \in [0, 1)$ можно представить в виде бесконечной десятичной дроби:

$$x = 0.x_1x_2x_3x_4 \dots$$

где $x_i \in \{0, 1, 2, \dots, 9\}$.

Важное замечание. Чтобы избежать неоднозначности (например, $0.1999\dots = 0.2000\dots$), договоримся всегда использовать запись **без бесконечной последовательности девяток**.содержимое...

Это объясняется тем, что, например $0.9999999\dots = 1$. Дейсвительно, пусть $x := 0.(9)$, тогда $10x = 9.(9)$, поэтому $10x - x = 9$, откуда $9x = 9$, поэтому $x = 1$.

Таким образом, если мы предположили что существует биекция f, то это значит, что мы можем составить полный список всех чисел из $(0, 1)$:

$$\begin{array}{rcl} f(1) & = & 0.a_{11}a_{12}a_{13}a_{14}a_{15}\dots \\ f(2) & = & 0.a_{21}a_{22}a_{23}a_{24}a_{25}\dots \\ f(3) & = & 0.a_{31}a_{32}a_{33}a_{34}a_{35}\dots \\ f(4) & = & 0.a_{41}a_{42}a_{43}a_{44}a_{45}\dots \\ f(5) & = & 0.a_{51}a_{52}a_{53}a_{54}a_{55}\dots \\ & \vdots & \end{array}$$

где каждый a_{ij} это цифра от 0 до 9.

Теперь сделаем следующее, мы выпишем **главную диагональ** этой таблицы:

$$a_{11}, a_{22}, a_{33}, a_{44}, a_{55}, \dots$$

и построим новое число $X = 0.b_1b_2b_3b_4b_5\dots$, где каждая цифра b_k выбирается по правилу:

$$b_k = \begin{cases} 1, & \text{если } a_{kk} \neq 1, \\ 2, & \text{если } a_{kk} = 1. \end{cases}$$

Наблюдение 2.3 И вот ключивой момент! Построенное число отсутствует в этом списке!

- $X \neq f(1)$, потому что $b_1 \neq a_{11}$ (они различаются в первом знаке)
- $X \neq f(2)$, потому что $b_2 \neq a_{22}$ (они различаются во втором знаке)
- $X \neq f(3)$, потому что $b_3 \neq a_{33}$ (они различаются в третьем знаке)
- ...
- $X \neq f(k)$ для любого $k \in \mathbb{N}$, потому что $b_k \neq a_{kk}$ (они различаются в k-ом знаке)

Таким образом, мы пришли к противоречию ибо:

- С одной стороны, по нашему предположению, список содержит **все** числа из $(0, 1)$,
- а с другой стороны, мы построили число $X \in (0, 1)$, которого **нет** в этом списке!

Следовательно, наше исходное предположение **неверно**. Множество $(0, 1)$ **несчётно**, а значит, и всё множество действительных чисел $\mathbb{R}$ также несчётно.

Историческое замечание. Рассужддение приведённое выше называют *диагональным аргументом Кантора*. Диагональный аргумент Кантора — одно из самых важных и элегантных доказательств в математике. Георг Кантор (1845-1918) с помощью этого метода показал, что существуют различные «уровни» бесконечности. В частности, он доказал, что множество действительных чисел **несчётно**, то есть его нельзя перенумеровать с помощью натуральных чисел.

Пример 2.21 Приведём наглядную иллюстрацию работы этого метода. Рассмотрим пример частичного списка:

$$
\begin{array}{rcl}
f(1) & = & 0.\mathbf{5}23140\ldots \\
f(2) & = & 0.3\mathbf{7}2159\ldots \\
f(3) & = & 0.12\mathbf{9}634\ldots \\
f(4) & = & 0.987\mathbf{6}32\ldots \\
f(5) & = & 0.1234\mathbf{5}67\ldots \\
& \vdots &
\end{array}
$$

Диагональные цифры: $\mathbf{5}, \mathbf{7}, \mathbf{9}, \mathbf{6}, \mathbf{5}, \ldots$
Строим число X:
Получаем $X = 0.11111\ldots$, которого заведомо нет в списке (поскольку мы отличаемся от каждого числа хотя бы в одном разряде).

Важное замечание. Говорят, что два множетсва X, Y имеют *одинаковую мощность*, если существует биекция между ними.

Мощность множества действительных чисел называется *мощностью континуума* и обозначается $\mathfrak{c}$ или $2^{\aleph_0}$.

Диагональный аргумент Кантора показывает, что бесконечности бывают разными:

- $\aleph_0$ — мощность счётных множеств ($\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$)
- $\mathfrak{c}$ — мощность континуума ($\mathbb{R}$, $[0,1]$, $\mathcal{P}(\mathbb{N})$)

При этом $\mathfrak{c} > \aleph_0$.

Диагональный аргумент можно обобщить: для любого множества A мощность его булеана (множества всех подмножеств) $\mathscr{P}(A)$ строго больше мощности самого A:

$$|\mathscr{P}(A)| > |A|$$

Это приводит к идее **иерархии бесконечностей**:

$$\aleph_0 < 2^{\aleph_0} < 2^{2^{\aleph_0}} < \cdots$$

Диагональный аргумент Кантора — фундаментальный результат, который изменил понимание бесконечности в математике. Он показывает, что не все бесконечности равны, и открывает путь к изучению различных мощностей множеств.

Пример 2.22 Рассмотрим отображение $f : \mathbb{N} \to \mathbb{Z}$, заданное формулой:

$$f(n) = \begin{cases} \frac{n}{2}, & \text{если } n \text{ чётно} \\ -\frac{n-1}{2}, & \text{если } n \text{ нечётно} \end{cases}$$

$$\begin{array}{ccc} \mathbb{N}: & & \mathbb{Z}: \\ 1 & \longrightarrow & 0 \\ 2 & \longrightarrow & 1 \\ 3 & \longrightarrow & -1 \\ 4 & \longrightarrow & 2 \\ 5 & \longrightarrow & -2 \\ 6 & \longrightarrow & 3 \end{array}$$

нетрудно видеть, что это биекция.

Пример 2.23 Используем диагональный метод Кантора для нумерации положительных рациональных чисел:

$$
\begin{tikzcd}
\frac{1}{1} \arrow[r] & \frac{1}{2} \arrow[dl] & \frac{1}{3} \arrow[r] & \frac{1}{4} \arrow[dl] \\
\frac{2}{1} \arrow[d] & & \frac{2}{3} \arrow[dl] & \\
\frac{3}{1} \arrow[uurr] & \frac{3}{2} \arrow[dl] & & \\
\frac{4}{1} & & &
\end{tikzcd}
$$

Порядок нумерации (пропуская повторяющиеся дроби):

$$1 \to \frac{1}{1}, \quad 2 \to \frac{2}{1}, \quad 3 \to \frac{1}{2}, \quad 4 \to \frac{1}{3}, \quad 5 \to \frac{2}{2}, \quad 6 \to \frac{3}{1}, \quad 7 \to \frac{2}{3}, \quad 8 \to \frac{1}{4}, \quad \dots$$

Это значит, что мы построили биекцию между $\mathbb{N}$ и $\mathbb{Q}_+$.

Пример 2.24
Комбинируем предыдущие идеи, мы можем показать биекцию между $\mathbb{N}$ и $\mathbb{Q}$. Сначала установим биекцию между $\mathbb{N}$ и $\mathbb{Q}_+$, затем расширим на все рациональные:

$$g(n) = \begin{cases} 0, & \text{если } n = 1 \\ f\left(\frac{n}{2}\right), & \text{если } n \text{ чётно} \\ -f\left(\frac{n-1}{2}\right), & \text{если } n \text{ нечётно и } n > 1 \end{cases}$$

где $f : \mathbb{N} \to \mathbb{Q}^+$ – биекция между натуральными и положительными рациональными.

$$
\begin{tikzcd}
\mathbb{N}: & \mathbb{Q}: \\
1 \arrow[r] & 0 \\
2 \arrow[r] & \frac{1}{1} \\
3 \arrow[r] & -\frac{1}{1} \\
4 \arrow[r] & \frac{2}{1} \\
5 \arrow[r] & -\frac{2}{1} \\
6 \arrow[r] & \frac{1}{2} \\
7 \arrow[r] & -\frac{1}{2}
\end{tikzcd}
$$

2.4.2 Понятие изоморфизма

Теперь у нас всё готово чтобы дать определение изоморфизма между векторными пространствами.

Определение 2.4.5. Пусть $\mathbf{V}$, $\mathbf{W}$ – два векторных пространства. Биективное отображение

$$\varphi : \mathbf{V} \to \mathbf{W}$$

называется *изоморфизмом*, если для любых векторов $\mathbf{a}, \mathbf{b} \in \mathbf{V}$ и любого числа $\lambda \in \mathbb{R}$

$$\varphi(\mathbf{a} + \mathbf{b}) = \varphi(\mathbf{a}) + \varphi(\mathbf{b})$$
$$\varphi(\lambda\mathbf{a}) = \lambda\varphi(\mathbf{a}).$$

Векторные пространства $\mathbf{V}$, $\mathbf{W}$ называются *изоморфными*, если существует хотя бы один изоморфизм $\mathbf{V} \to \mathbf{W}$. В таком случае пишут $\mathbf{V} \cong \mathbf{W}$.

Важное замечание. В линейной алгебре, как правило, интересуются лишь теми свойствами векторных пространств, которые сохраняются при изоморфизме. Поэтому изоморфные пространства считаются "одинаковыми".

Определение 2.4.6. Пусть $\mathbb{R}^n$ есть множество все возможных наборов n чисел, который мы будем записывать так $(\alpha_1, \dots, \alpha_n)^\top$. После введения операций сложения и умножение на число λ следующим образом

$$(\alpha_1, \dots, \alpha_n)^\top + (\beta_1, \dots, \beta_n)^\top := (\alpha_1 + \beta_1, \dots, \alpha_n + \beta_n)^\top, \tag{2.4.1}$$
$$\lambda \cdot (\alpha_1, \dots, \alpha_n)^\top := (\lambda\alpha_1, \dots, \lambda\alpha_n)^\top, \tag{2.4.2}$$

и пологая что $\mathbf{o} := (0, \dots, 0)^\top$, мы получаем векторное пространство размерности n.

Важное замечание. Особую роль в пространстве $\mathbb{R}^n$ играют следующие n векторов

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

Предложение 2.4.1. *Эти векторы образуют базис пространства $\mathbb{R}^n$ который принято называть стандартным базисом.*

Доказательство.

(1) Покажем, что $\mathbb{R}^n$ есть линейная оболочка натянутая на эти векторы, *т.е.,* любой вектор $(\alpha_1, \dots, \alpha_n)^\top \in \mathbb{R}^n$ как то линейно выражается через эти векторы.

Имеем

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = \alpha_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \dots + \alpha_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1. \end{pmatrix}$$

(2) Нам осталось показать, что эти векторы линейно независимы. Итак, пусть имеются некоторые числа $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ такие, что

$$\lambda_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \dots + \lambda_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

но тогда, согласно формулам (2.4.1), (2.4.2), получаем

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

откуда следует, что $\lambda_1 = \lambda_2 = \dots = \lambda_n = 0$. Это доказывает линейную независимость. Предложение доказано. □

2.4.3 Все конечномерные пространства это $\mathbb{R}^n$

Теорема 2.4.2. *Если* $\mathbf{V}$ *– векторное пространство размерности n, то оно изоморфно пространству* $\mathbb{R}^n$.

Доказательство. Пусть $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ – базис в $\mathbf{V}$, тогда положим

$$\begin{array}{rcl} \varphi(\mathbf{e}_1) & := & (1,0,0,\dots,0)^\top, \\ \varphi(\mathbf{e}_2) & := & (0,1,0,\dots,0)^\top, \\ \varphi(\mathbf{e}_3) & := & (0,0,1,\dots,0)^\top, \\ \dots & \dots & \dots\dots\dots\dots\dots \\ \varphi(\mathbf{e}_n) & := & (0,0,0,\dots,1)^\top. \end{array}$$

Далее, так как любой вектор $\mathbf{v} \in \mathbf{V}$ имеет какие-то координаты в этом базисе, скажем $\mathbf{v} = (\alpha_1, \dots, \alpha_n)^\top$, *т.е.*,

$$\mathbf{v} = \alpha_1 \mathbf{e}_1 + \dots + \alpha_n \mathbf{e}_n,$$

то пологая

$$\begin{array}{rcl} \varphi(\mathbf{v}) & := & \varphi(\alpha_1 \mathbf{e}_1 + \dots + \alpha_n \mathbf{e}_n) \\ & = & \alpha_1 \varphi(\mathbf{e}_1) + \dots + \alpha_n \varphi(\mathbf{e}_n), \end{array}$$

мы, тем самым, получаем

$$\begin{array}{rcl} \varphi(\mathbf{v}) & = & \alpha_1 \varphi(\mathbf{e}_1) + \dots + \alpha_n \varphi(\mathbf{e}_n) \\ & = & \alpha_1 \begin{pmatrix} 1 \\ \vdots \\ 0 \end{pmatrix} + \dots + \alpha_n \begin{pmatrix} 0 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \in \mathbb{R}^n. \end{array}$$

Таким образом, для любого вектора $(\alpha_1, \dots, \alpha_n)^\top \in \mathbb{R}^n$ можно найти вектор $\mathbf{v} \in \mathbf{V}$ у которого в базисе E координаты $\alpha_1, \dots, \alpha_n$, такой что $\varphi(\mathbf{v}) = (\alpha_1, \dots, \alpha_n)^\top$. Таким образом, φ – сюръективно.

Пусть теперь $\mathbf{v}, \mathbf{v}' \in \mathbf{V}$ две вектора, и пусть $\alpha_1, \dots, \alpha_n$ и $\alpha'_1, \dots, \alpha'_n$ – их координаты в базисе E.

Тогда,

$$\varphi(\mathbf{v}) = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}, \qquad \varphi(\mathbf{v}') = \begin{pmatrix} \alpha'_1 \\ \vdots \\ \alpha'_n \end{pmatrix}$$

и если $\varphi(\mathbf{v}) = \varphi)(\mathbf{v}')$, то

$$\varphi(\mathbf{v}) = \varphi(\mathbf{v}') \Longleftrightarrow \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \alpha'_1 \\ \vdots \\ \alpha'_n \end{pmatrix} \Longleftrightarrow \alpha_1 = \alpha'_1, \dots, \alpha_n = \alpha'_n,$$

поэтому $\mathbf{v} = \mathbf{v}'$, *т.е.,* φ – иньективно. Мы тем самым получаем линейное отображение $\varphi : \mathbf{V} \to \mathbb{R}^n$ которое биективно. То есть мы построили изоморфизм что и доказывает теорему. $\square$

Следствие 2.4.3. *Любые два векторных пространства конечной размерности изоморфны тогда и только тогда, когда у них одинаковая размерность.*

Когда мы будем говорить об $\mathbb{R}^n$ как о векторном пространстве, то каждый набор будем записывать как вектор, *т.е.,* в виде $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (x_1, \dots, x_n)^\top$.

Возьмём $\mathbf{x} = (x_1, \dots, x_n)^\top \in \mathbb{R}^n$, тогда ясно, что

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + x_n \begin{pmatrix} 0 \\ \vdots \\ 1 \end{pmatrix}$$

Множество $\mathbb{e} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$, где $\mathbf{e}_1 = (1,0,\dots,0)^\top, \dots, \mathbf{e}_n = (0,0,\dots,1)^\top$, называется *базисом* пространства $\mathbb{R}^n$.

Глава 3

Линейные отображения

3.1 Лекция #6. Линейные отображения

Раз мы уже поняли когда векторные пространства можно считать одинаковыми (=изоморфными), то теперь мы хотим изучать взаимодействия между ними. Для этого нам нужно просто научиться сохранять линейную структуру, поэтому мы будем расмматривать соответсвующие отображения.

3.1.1 Основные понятия

Итак, теперь мы ослабим понятие изоморфизма между векторными простраснвами следующим образом – мы не будем требовать биективности, но сохраним линейность.

Определение 3.1.1. Пусть даны два векторных пространства $\mathbf{V}, \mathbf{W}$. Отображение между множествами $\varphi : \mathbf{V} \to \mathbf{W}$ называется *линейным* если

$$\begin{aligned} \varphi(\mathbf{a} + \mathbf{b}) &= \varphi(\mathbf{a}) + \varphi(\mathbf{b}), \\ \varphi(\lambda \mathbf{a}) &= \lambda \varphi(\mathbf{a}), \end{aligned}$$

для любых векторов $\mathbf{a}, \mathbf{b} \in \mathbf{V}$ и любого числа $\lambda \in \mathbb{R}$.

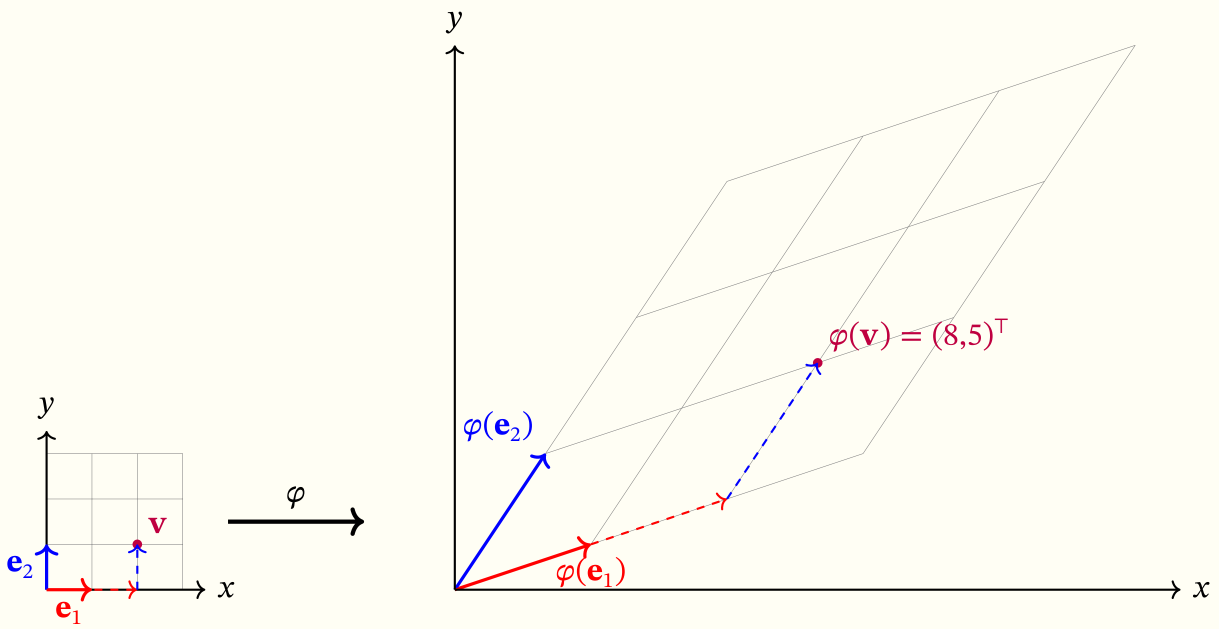
Важное замечание. В частности, если $\lambda = 0$, то мы получаем $\varphi(\mathbf{0}_\mathbf{V}) = \mathbf{0}_\mathbf{W}$, *т.е.,* линейное отображение нулевой вектор всегда переводит в нулевой вектор!

3.1.2 Пример: почему достаточно знать образы базисных векторов

Пример 3.1 Рассмотрим линейное отображение $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$, заданное на стандартном базисе:

$$\varphi(\mathbf{e}_1) = (3,1), \quad \varphi(\mathbf{e}_2) = (2,3)$$

где $\mathbf{e}_1 = (1,0)^\top$, $\mathbf{e}_2 = (0,1)^\top$.
Покажем, что этого достаточно, чтобы определить образ **любого** вектора $(x,y)^\top$.



Исходное пространство: $\mathbf{v} = 2\mathbf{e}_1 + 1\mathbf{e}_2$
Линейное отображение полностью определяется образами базисных векторов
$\varphi(\mathbf{v}) = 2 \cdot (3,1) + 1 \cdot (2,3) = (6 + 2, 2 + 3) = (8,5)$

Вычисление образа произвольного вектора. Любой вектор $\mathbf{v} = (x,y)^\top$ можно разложить по базису:

$$\mathbf{v} = x\mathbf{e}_1 + y\mathbf{e}_2 = x\begin{pmatrix}1\\0\end{pmatrix} + y\begin{pmatrix}0\\1\end{pmatrix}$$

В силу линейности отображения:

$$\begin{aligned}
\varphi(\mathbf{v}) &= \varphi(x\mathbf{e}_1 + y\mathbf{e}_2)\\
&= x\varphi(\mathbf{e}_1) + y\varphi(\mathbf{e}_2)\\
&= x\begin{pmatrix}3\\1\end{pmatrix} + y\begin{pmatrix}2\\3\end{pmatrix}\\
&= \begin{pmatrix}3x + 2y\\x + 3y\end{pmatrix}
\end{aligned}$$

Таким образом, для любого вектора $(x,y)^\top$ его образ однозначно определяется:

$$\varphi\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}3x + 2y\\x + 3y\end{pmatrix}$$

□

3.1.3 Почему квадраты становятся параллелограммами?

Комментарий. Рассмотрим единичный квадрат с вершинами:

$$O = (0,0), \quad A = (1,0), \quad B = (1,1), \quad C = (0,1)$$

Под действием отображения φ:

$$\begin{aligned}
&\varphi(O) = (0,0)\\
&\varphi(A) = \varphi(\mathbf{e}_1) = (3,1)\\
&\varphi(B) = \varphi(\mathbf{e}_1 + \mathbf{e}_2) = \varphi(\mathbf{e}_1) + \varphi(\mathbf{e}_2) = (3,1) + (2,3) = (5,4)\\
&\varphi(C) = \varphi(\mathbf{e}_2) = (2,3)
\end{aligned}$$

Получаем параллелограмм с вершинами (0,0), (3,1), (5,4), (2,3).
Почему это именно параллелограмм?

- Противоположные стороны равны и параллельны:

$$\begin{aligned}
\overrightarrow{\varphi(O)\varphi(A)} &= \varphi(\mathbf{e}_1) = (3,1)\\
\overrightarrow{\varphi(C)\varphi(B)} &= \varphi(\mathbf{e}_1) = (3,1)\\
\overrightarrow{\varphi(O)\varphi(C)} &= \varphi(\mathbf{e}_2) = (2,3)\\
\overrightarrow{\varphi(A)\varphi(B)} &= \varphi(\mathbf{e}_2) = (2,3)
\end{aligned}$$

- Все квадраты преобразуются в **одинаковые** параллелограммы, потому что каждый квадрат получается сдвигом единичного квадрата, а линейное отображение сохраняет эту структуру.

Замечание 3.1.1. Матрица этого линейного отображения в стандартном базисе:

$$A = \begin{pmatrix}\varphi(\mathbf{e}_1) & \varphi(\mathbf{e}_2)\end{pmatrix} = \begin{pmatrix}3 & 2\\ 1 & 3\end{pmatrix}$$

Действительно, для любого вектора $\mathbf{v} = (x,y)^\top$:

$$A\mathbf{v} = \begin{pmatrix}3 & 2\\ 1 & 3\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix} = \begin{pmatrix}3x + 2y\\ x + 3y\end{pmatrix} = \varphi(\mathbf{v})$$

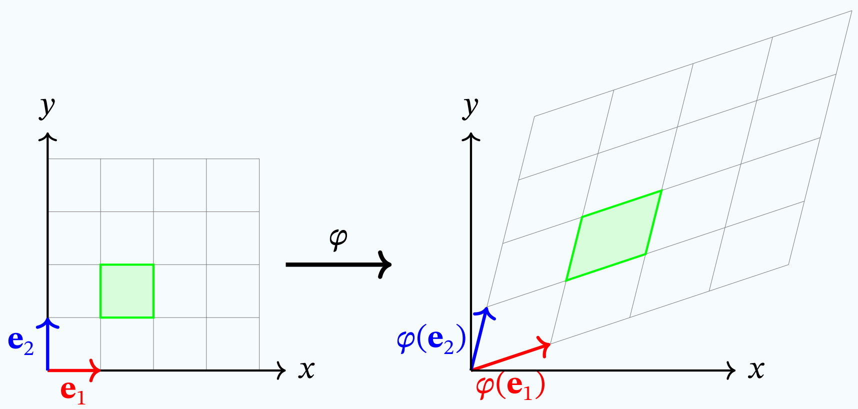
Этот пример наглядно показывает: 1. Линейное отображение полностью определяется образами базисных векторов 2. Координаты вектора сохраняются при отображении (в смысле коэффициентов разложения) 3. Квадраты превращаются в параллелограммы из-за линейности преобразования 4. Все параллелограммы одинаковы, так как отображение действует равномерно на всё пространство

Наблюдение 3.1 Геометрически, линейность отображения означает, что образ прямой – это опять прямая. Но если образ прямой это прямая, то вовсе это не значит что само отображение линейное!

Более точно: линейное отображение обладает свойством **сохранения структуры координатной сетки**. Представьте себе регулярную сетку из одинаковых квадратов. При линейном отображении:

- Все прямые линии остаются прямыми
- Параллельные линии остаются параллельными
- Относительные расстояния сохраняются (хотя масштаб может меняться)
- Квадраты превращаются в **одинаковые параллелограммы**

Это происходит потому, что линейное отображение равномерно действует на всё пространство — оно полностью определяется своим действием на базисные векторы.



Исходное пространство: одинаковые квадраты

Образ: одинаковые параллелограммы

Линейное отображение преобразует регулярную сетку квадратов в регулярную сетку одинаковых параллелограммов.

Пример 3.2 Рассмотрим отображение $f(x,y) = (x, y + x^2)$. Оно переводит прямые в прямые:

- Вертикальная прямая $x = c$ переходит в вертикальную прямую $x = c$
- Горизонтальная прямая $y = c$ переходит в прямую $y = c + x^2$ — это уже **не прямая**!

Хотя некоторые прямые переходят в прямые, отображение не является линейным, так как не сохраняет структуру всей сетки — квадраты превращаются в криволинейные фигуры разной формы.

Комментарий. Ключевое наблюдение: линейное отображение $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ полностью определяется тем, куда оно переводит базисные векторы $\mathbf{e}_1$ и $\mathbf{e}_2$.
Если исходная сетка состояла из квадратов $1{\times}1$, то после отображения:

- Квадрат с вершинами $(0,0), (1,0), (1,1), (0,1)$ переходит в параллелограмм с вершинами:

$$\begin{aligned}(0,0) &\mapsto (0,0)\\ (1,0) &\mapsto \varphi(\mathbf{e}_1)\\ (1,1) &\mapsto \varphi(\mathbf{e}_1)+\varphi(\mathbf{e}_2)\\ (0,1) &\mapsto \varphi(\mathbf{e}_2)\end{aligned}$$

- Все квадраты преобразуются одинаково — получаются **равные параллелограммы**

- Отношения площадей сохраняются (площадь каждого параллелограмма = $|\det(A)|$)

Это свойство **не выполняется** для нелинейных отображений, которые могут искривлять линии или превращать одинаковые квадраты в фигуры разной формы.

Пример 3.3 Рассмотрим отображение $f(x,y) = (x, y + x^2)$. Оно переводит прямые в прямые:

- Вертикальная прямая $x = c$ переходит в вертикальную прямую $x = c$

- Горизонтальная прямая $y = c$ переходит в прямую $y = c + x^2$ — это уже **не прямая**!

Хотя некоторые прямые переходят в прямые, отображение не является линейным, так как не сохраняет структуру всей сетки — квадраты превращаются в криволинейные фигуры разной формы.

Наблюдение 3.2 Геометрически, линейность отображения означает, что образ прямой – это опять прямая. Но если образ прямой это прямая, то вовсе это не значит что само отображение линейное!

Пример 3.4 В компьютерной графике линейные отображения повсеместно используются для преобразований:

- **Масштабирование**: $f(x,y) = (2x, 2y)$ - увеличивает изображение в 2 раза

- **Поворот**: $f(x,y) = (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta)$ - поворачивает изображение

- **Сдвиг**: $f(x,y) = (x + ay, y)$ - наклоняет изображение

Все эти преобразования представляются матрицами и сохраняют прямые линии, что делает их линейными.

Пример 3.5 В машинном обучении линейная регрессия — это линейное отображение из пространства признаков в пространство предсказаний:

$$f : \mathbb{R}^n \to \mathbb{R}, \quad f(\mathbf{x}) = w_1 x_1 + w_2 x_2 + \dots + w_n x_n + b$$

где $\mathbf{w}$ — веса модели, b — смещение. Это отображение линейно относительно входных признаков.

Важное замечание. В таком случае, линейное отображение $f : \mathbb{R}^n \to \mathbb{R}^m$ достаточно задать на базисных векторах и мы получаем что-то вроде

$$\begin{pmatrix} 1 \\ \vdots \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ \vdots \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix},$$

что и кодируется матрицей $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$

Возьмём теперь произвольный вектор $\mathbf{x} \in \mathbf{V}$, и пусть $\mathbb{e} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ – базис в $\mathbf{V}$, тогда его можно расписать следующим образом

$$\mathbf{x} = x_1 \mathbf{e}_1 + \dots + x_n \mathbf{e}_n.$$

Поэтому, если f – линейное отображение, то

$$\begin{aligned} f(\mathbf{x}) &= f(x_1 \mathbf{e}_1 + \dots + x_n \mathbf{e}_n) \\ &= x_1 f(\mathbf{e}_1) + \dots + x_n f(\mathbf{e}_n) \\ &= x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}, \end{aligned}$$

поэтому естественно записать в матричном виде это ображение следующим образом

$$f : \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} a_{11} x_1 + \dots + a_{1n} x_n \\ \vdots \\ a_{m1} x_1 + \dots + a_{mn} x_n. \end{pmatrix}$$

и мы можем тогда положить

$$\boxed{A \cdot \mathbf{x} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} a_{11} x_1 + \dots + a_{1n} x_n \\ \vdots \\ a_{m1} x_1 + \dots + a_{mn} x_n. \end{pmatrix}}. \tag{3.1.1}$$

Теорема 3.1.1. *Пусть даны два конечномерных векторных пространства* $\mathbf{V}$, $\mathbf{W}$ *над* $\mathbb{R}$ *с размерностями* n, m *соответственно. Пусть* $f : \mathbf{V} \to \mathbf{W}$ *– линейное отображение. Тогда существует единственная матрица* $A \in \mathrm{Mat}_{m \times n}(\mathbb{R})$*, такая что*

$$f(\mathbf{v}) = A \cdot \mathbf{v}, \qquad \text{для всех } \mathbf{v} \in \mathbf{V}.$$

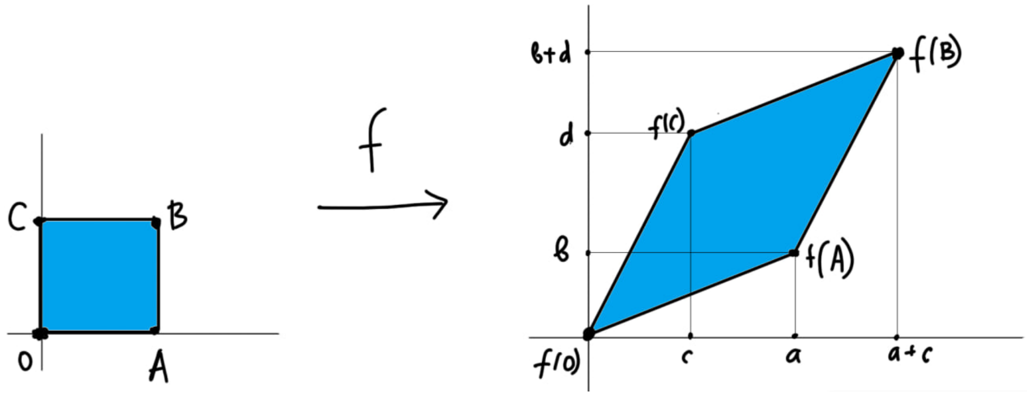


Рис. 3.1: Линейное отображение $f : \mathbb{R}^2 \to \mathbb{R}^2$, которое задаётся матрицей $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$

Более того, эта матрица имеет вид $A = \begin{pmatrix} f(\mathbf{e}_1) & \dots & f(\mathbf{e}_n) \end{pmatrix}$, *где* $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ *– какой-то базис в* $\mathbf{W}$.

Доказательство. Действительно, пусть $\mathbf{e} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ – какой-то базис в $\mathbf{V}$ и $\mathbf{f} = \{\mathbf{f}_1, \dots, \mathbf{f}_m\}$ – какой-то базис в $\mathbf{W}$. Тогда (см. Определение 2.3.2) для любого вектора $\mathbf{v} \in \mathbf{V}$ имеем $\mathbf{v} = v_1\mathbf{e}_1 + \dots + v_n\mathbf{e}_n$ и в силу линейности f (см. Определение 3.1), имеем

$$\begin{aligned} f(\mathbf{v}) &= f(v_1\mathbf{e}_1 + \dots + v_n\mathbf{e}_n) \\ &= v_1 f(\mathbf{e}_1) + \dots + v_n f(\mathbf{e}_n) \\ &= \begin{pmatrix} f(\mathbf{e}_1) & \dots & f(\mathbf{e}_n) \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = A \cdot \mathbf{v}, \end{aligned}$$

что и требовалось доказать. □

Таким образом, любая матрица $A \in \mathrm{Mat}_{n\times m}(\mathbb{R})$ задаёт линейное отображение $f : \mathbb{R}^n \to \mathbb{R}^m$, которое отправляет i-ый базисный вектор $\mathbf{e}_i$ в i-ый столбец этой матрицы.

Замечание 3.1.2. Если мы будем рассматривать линейное отображения в каких-то векторных пространствах $f : \mathbf{V} \to \mathbf{W}$ в которых заданы базисы $\mathbb{e} = \{\mathbf{e}_1, \dots, \mathbf{e}_m\}$, $\mathbb{f} = \{\mathbf{f}_1, \dots, \mathbf{f}_m\}$, соответственно, то, в таком случае, матрицу A линейного отображения записывают следующим образом $A(f; \mathbb{e}, \mathbb{f})$.

Понятно, что одними только линейными всё не ограничивается. Ведь вовсе не обязательно, например, что образ прямой будет всегда прямая при любом её отображении.

Можно рассмотреть, например, и что-то более экзотическое, как показано на рисунке 3.4.

Однако это не означает, что линейную алгебру не надо изучать. Как раз наоборот, в сущности, математический анализ изучает любые подобное отображения с помощью линейной алгебры; локально они устроены как раз таки линейно (см. Рис.3.3).

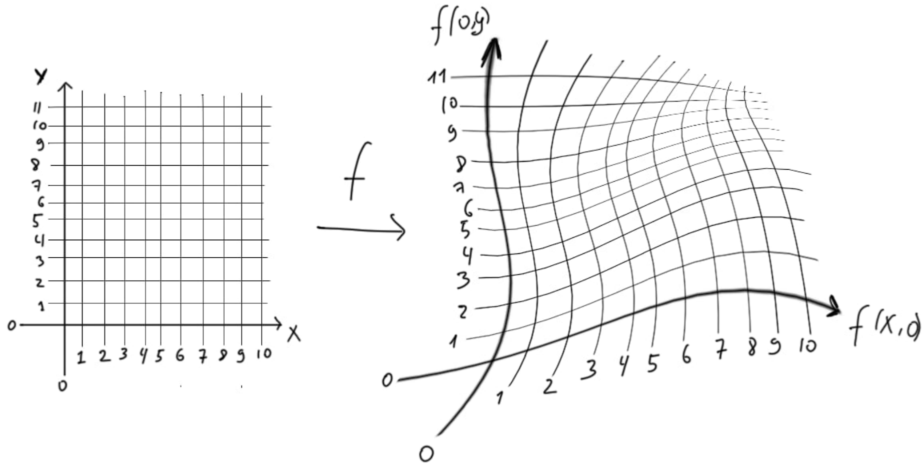


Рис. 3.2: Пример нелинейного отображения.

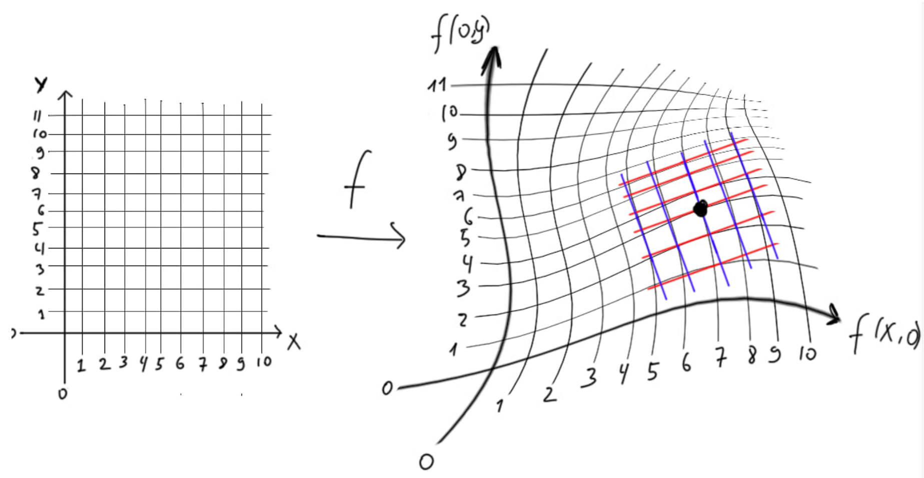


Рис. 3.3: Вблизи это трудное отображение очень похоже на линейное.

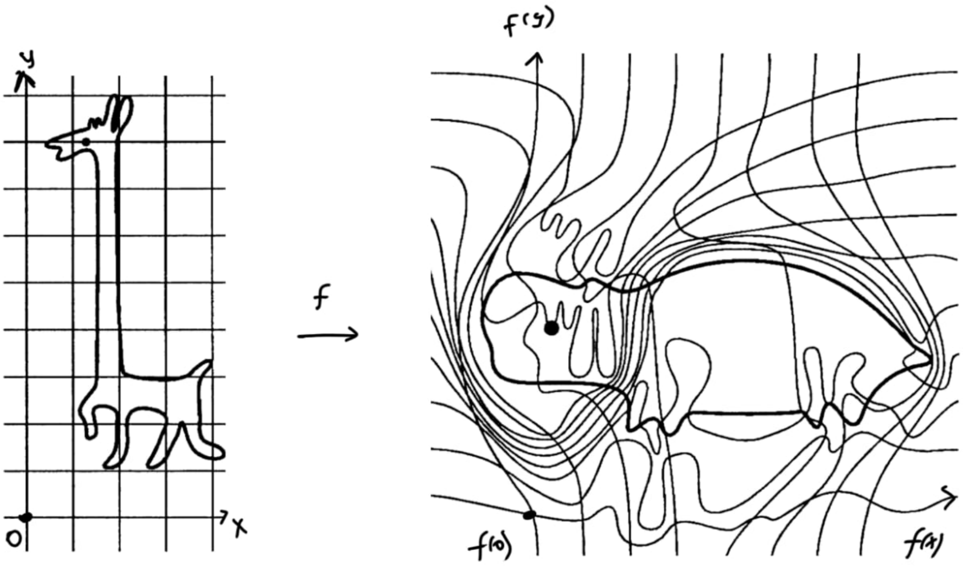


Рис. 3.4: Пример нелинейного отображения: из жирафа получается бегемот.

3.1.4 Решение матричного уравнения $A \cdot \mathbf{x} = \mathbf{b}$

Пусть имеем линейное отображение $f : \mathbb{R}^n \to \mathbb{R}^m$ с матрицей

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

и пусть дан какой-то конкретный вектор $\mathbf{b} \in \mathbb{R}$. Мы хотим найти (или показать что такого не может быть) все такие векторы $\mathbf{x} \in \mathbb{R}^n$, что $f(\mathbf{x}) = \mathbf{b}$.

Пусть $\mathbf{b} = (b_1, \dots, b_m)^\top$, тогда, согласно (3.1.1), мы получаем

$$A \cdot \mathbf{x} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n. \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

т.е., мы получаем линейную систему

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases}$$

Таким образом, данная задача сводится к решению линейной системы.

Более того, теперь любую линейную систему можно лаконично записать в виде

$$A \cdot \mathbf{x} = \mathbf{b},$$

где A – матрица коэффициентов, $\mathbf{b}$ – вектор координаты которого это числа, стоящие справа от равно, а $\mathbf{x} = (x_1, \dots, x_n)^\top$ это вектор координаты которого нужно найти.

3.1.5 Ядро, образ линейного отображения и понятие ранга матрицы

Пусть $f : \mathbf{V} \to \mathbf{W}$ – линейное отображение между двумя векторными пространствами.

Определение 3.1.2. Ядром отображения f называется множество

$$\operatorname{Ker}(f) := \{\mathbf{v} \in \mathbf{V} \ : \ f(\mathbf{v}) = \mathbf{0}_\mathbf{W}\}$$

Образ отображения f называется множество

$$\operatorname{Im}(f) := \{f(\mathbf{v}) \ : \ \mathbf{v} \in \mathbf{V}\}.$$

Предложение 3.1.2. *Множество* $\operatorname{Ker}(f)$ *– векторное подпространство в* $\mathbf{V}$*, а* $\operatorname{Im}(f)$ *– векторное подпространство в* $\mathbf{W}$.

Доказательство. Нужно проверить выполнения требований из Определения 2.1.1.

(1) Покажем что $\operatorname{Ker}(f)$ – подпространство в $\mathbf{V}$. Так как f – линейное, то (см. Замечание **??**) $f(\mathbf{0}_\mathbf{V}) = \mathbf{0}_\mathbf{W}$ поэтому $\mathbf{0}_\mathbf{V} \in \operatorname{Ker}(f)$. Далее, согласно определению 3.1, если $\mathbf{v}, \mathbf{v}' \in \operatorname{Ker}(f)$, то

$$f(\mathbf{v} + \mathbf{v}') = f(\mathbf{v}) + f(\mathbf{v}') = \mathbf{0}_\mathbf{W} + \mathbf{0}_\mathbf{W} = \mathbf{0}_\mathbf{W},$$

т.е., $\mathbf{v}+\mathbf{v} \in \mathrm{Ker}(f)$.

Далее, для любого $\alpha \in \mathbb{R}$, имеем

$$f(\alpha\mathbf{v}) = \alpha f(\mathbf{v}) = \alpha\mathbf{0}_{\mathbf{W}} = \mathbf{0}_{\mathbf{W}},$$

т.е., $\alpha\mathbf{v} \in \mathrm{Ker}(f)$ – подпространство.

(2) Покажем теперь, что $\mathrm{Im}(f)$ – подпространство в $\mathbf{W}$.

Так как $f(\mathbf{0}_{\mathbf{V}}) = \mathbf{0}_{\mathbf{W}}$, *т.е.,* $\mathbf{0}_{\mathbf{W}} \in \mathrm{Im}(f)$. Пусть $\mathbf{w}, \mathbf{w} \ \in \mathrm{Im}(f)$, покажем что $\mathbf{w}+\mathbf{w}'$ и $\alpha\mathbf{w} \in \mathrm{Im}(f)$ для любого $\alpha \in \mathbb{R}$.

Так как $\mathbf{w}, \mathbf{w} \ \in \mathrm{Im}(f)$, то найдутся такие $\mathbf{v}, \mathbf{v}' \in \mathbf{V}$, что $\mathbf{w} = f(\mathbf{v})$ и $\mathbf{w}' = f(\mathbf{v}')$. Тогда в силу линейности, имеем

$$\mathbf{w}+\mathbf{w}' = f(\mathbf{v}) + f(\mathbf{v}') = f(\mathbf{v}+\mathbf{v}'),$$

т.е., $\mathbf{w}+\mathbf{w}' = f(\mathbf{v}+\mathbf{v})$, что и означает, что $\mathbf{w}+\mathbf{w} \in \mathrm{Im}(f)$.

Наконец,

$$\alpha\mathbf{w} = \alpha f(\mathbf{v}) = f(\alpha\mathbf{v}),$$

т.е., $\alpha\mathbf{w} = f(\alpha\mathbf{v})$, что и означает, что $\alpha\mathbf{w} \in \mathrm{Im}(f)$. Это завершает доказательство. □

Определение 3.1.3. Размерность пространства $\mathrm{Im}(f)$ называется *рангом* отображения, в случае когда f задано матрицей A, то говорят про *ранг матрицы* A и обозначают его так $\mathrm{rk}(A)$. Таким образом, можно написать $\mathrm{rk}(A) := \dim(\mathrm{Im}(f))$.

3.2 Лекция #7. Композиция линейных отображений

Скажем также, что если у нас есть два линейных отображения $f : \mathbb{R}^n \to \mathbb{R}^k$, $g : \mathbb{R}^k \to \mathbb{R}^m$, закодированные матрицами B и A соответственно, то как закодировать с помощью матрицы отображение $g \circ f : \mathbb{R}^n \to \mathbb{R}^k$?

3.2.1 Вращение плоскости

Мы начнём с примера, где будем рассматривать вращение плоскости вокруг фиксированной точки.

Пример 3.2.1. Введём на плоскости $\mathbb{R}^2$ начало координат и будет вращать её вокруг этой точки. Тогда если мы делаем поворот на угол α, то как нетрудно видеть такое преобразование запишется так

$$F_\alpha : \begin{pmatrix}1\\0\end{pmatrix} \mapsto \begin{pmatrix}\cos\alpha\\ \sin\alpha\end{pmatrix}, \qquad F_\alpha : \begin{pmatrix}0\\1\end{pmatrix} \mapsto \begin{pmatrix}-\sin\alpha\\ \cos\alpha\end{pmatrix}$$

Таким образом можно написать $F_\alpha = \begin{pmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{pmatrix}$.

Рассмотрим теперь два поворота F_α, F_β. Ясно что последовательное применение сначала F_α а потом F_β приведёт к повороту на угол $\alpha+\beta$. Это можно записать так $F_\beta \circ F_\alpha = F_{\alpha+\beta}$. Значит мы можем записать $F_{\alpha+\beta} = \begin{pmatrix}\cos(\alpha+\beta) & -\sin(\alpha+\beta)\\ \sin(\alpha+\beta) & \cos(\alpha+\beta)\end{pmatrix}$

С другой стороны, посмотрим теперь более подробно что происходит.

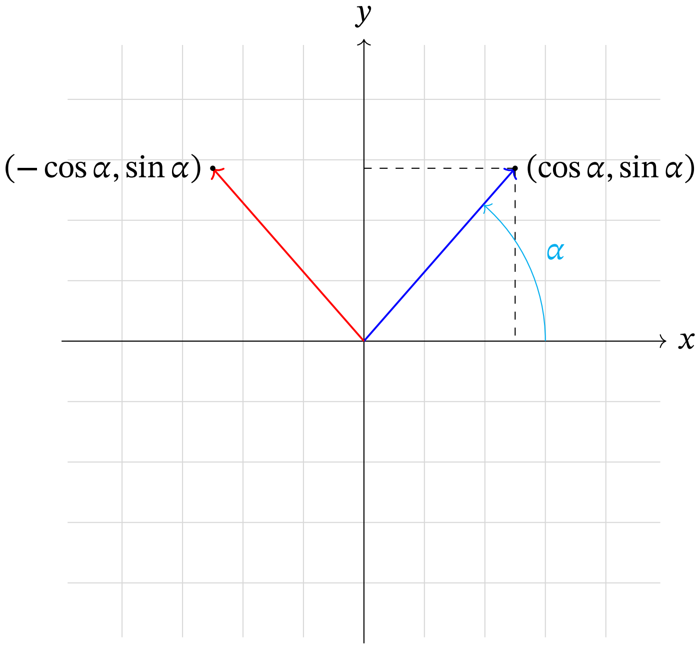


Рис. 3.5: Здесь показан обычный поворот плоскости вокруг начала координат на угол α.

Имеем

$$\begin{aligned} F_\beta\left(F_\alpha\begin{pmatrix}1\\0\end{pmatrix}\right) &= F_\beta\begin{pmatrix}\cos\alpha\\ \sin\alpha\end{pmatrix} = \begin{pmatrix}\cos\beta & -\sin\beta\\ \sin\beta & \cos\beta\end{pmatrix}\begin{pmatrix}\cos\alpha\\ \sin\alpha\end{pmatrix} \\ &= \begin{pmatrix}\cos\beta\cos\alpha - \sin\beta\sin\alpha\\ \sin\beta\cos\alpha + \cos\beta\sin\alpha\end{pmatrix}\end{aligned}$$

Аналогично, находим что

$$F_\beta\left(F_\alpha\begin{pmatrix}0\\1\end{pmatrix}\right) = \begin{pmatrix}-\cos\beta\sin\alpha - \sin\beta\cos\alpha\\ -\sin\beta\sin\alpha + \cos\beta\cos\alpha\end{pmatrix}.$$

Таким образом должно выполняться равенство

$$\begin{pmatrix}\cos(\alpha+\beta) & -\sin(\alpha+\beta)\\ \sin(\alpha+\beta) & \cos(\alpha+\beta)\end{pmatrix} = \begin{pmatrix}\cos\beta\cos\alpha - \sin\beta\sin\alpha & -\cos\beta\sin\alpha - \sin\beta\cos\alpha\\ \sin\beta\cos\alpha + \cos\beta\sin\alpha & -\sin\beta\sin\alpha + \cos\beta\cos\alpha\end{pmatrix}$$

но, видно, что матрицы совпадают. Таким образом, такое задание умножения матриц несёт себе простую идею - каждая матрица это некоторое преобразование векторного пространства. Тогда умножение матриц (если оно имеет смысл) соответствует композиции (=последовательному выполнению этих операций) этих преобразований.

3.2.2 Композиция отображений и его матричное представление

Теперь мы приступим к общему случаю, но сначала введём полезные для дальнейшего обозначения и соглашения

Пусть дана матрица $A \in \mathrm{Mat}_{n\times m}(\mathbb{R})$,

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

обозначим через $\mathbf{r}_i(A) := \begin{pmatrix} a_{i1} & a_{i2} & \dots & a_{in} \end{pmatrix}$ – её i-ую строку, а через $\mathbf{c}_j(A) := \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$ – её j-ый столбец.

Тогда саму матрицу A можно записать следующим образом

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} = \begin{pmatrix} \mathbf{r}_1(A) \\ \mathbf{r}_2(A) \\ \vdots \\ \mathbf{r}_m(A) \end{pmatrix} = \begin{pmatrix} \mathbf{c}_1(A) & \mathbf{c}_2(A) & \dots & \mathbf{c}_n(A) \end{pmatrix}.$$

Итак, пусть у нас есть три конечномерных векторных пространства $\mathbb{R}^n, \mathbb{R}^k, \mathbb{R}^m$ и два линейных отображения $g : \mathbb{R}^n \to \mathbb{R}^k$, $f : \mathbb{R}^k \to \mathbb{R}^m$. Это проще записать следующим образом

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, "g"] \arrow[dr, "f\circ g"'] & \mathbb{R}^k \arrow[d, "f"] \\ & \mathbb{R}^m \end{tikzcd}$$

Возьмём произвольный вектор $\mathbf{x} \in \mathbb{R}^n$, пусть $\mathbf{x} = x_1\mathbf{e}_1 + \dots + x_n\mathbf{e}_n$, тогда, получаем

$$\begin{aligned} g(\mathbf{x}) &= g(x_1\mathbf{e}_1 + \dots + x_n\mathbf{e}_n) \\ &= x_1 g(\mathbf{e}_1) + \dots + x_n g(\mathbf{e}_n) \\ &= x_1 \begin{pmatrix} b_{11} \\ \vdots \\ b_{m1} \end{pmatrix} + \dots + x_n \begin{pmatrix} b_{1n} \\ \vdots \\ b_{mn} \end{pmatrix} \\ &= x_1\mathbf{c}_1(B) + \dots + x_n\mathbf{c}_n(B). \end{aligned}$$

Далее, имеем

$$\begin{aligned} f(g(\mathbf{x})) &= f(x_1\mathbf{c}_1(B) + \dots + x_n\mathbf{c}_n(B)) \\ &= x_1 f(\mathbf{c}_1(B)) + \dots + x_n f(\mathbf{c}_n(B)) \\ &= x_1 A\mathbf{c}_1(B) + \dots + x_n A\mathbf{c}_n(B) \end{aligned}$$

Таким образом, мы получаем, что композиция отображений $f\circ g$ задаётся матрицей M, которая находится следующим образом

$$M := \begin{pmatrix} A\mathbf{c}_1(B) & \dots & A\mathbf{c}_n(B) \end{pmatrix}$$

Определение 3.2.1. Такую матрицу называют произвдением матриц A и B и обозначают её так AB.

Пример 3.6 Вычислим произведение AB, где $A = \begin{pmatrix} 2 & 3 \\ 1 & -5 \end{pmatrix}$ и $B = \begin{pmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{pmatrix}$.
Имеем

$$B = \begin{pmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{pmatrix} = \begin{pmatrix} \mathbf{c}_1(B) & \mathbf{c}_2(B) & \mathbf{c}_3(B) \end{pmatrix}$$

где

$$\mathbf{c}_1(B) = \begin{pmatrix} 4 \\ 1 \end{pmatrix}, \quad \mathbf{c}_2(B) = \begin{pmatrix} 3 \\ -2 \end{pmatrix}, \quad \mathbf{c}_3(B) = \begin{pmatrix} 6 \\ 3 \end{pmatrix}.$$

Тогда, согласно (3.1.1), находим

$$\begin{aligned} A \cdot \mathbf{c}_1(B) &= \begin{pmatrix} 2 & 3 \\ 1 & -5 \end{pmatrix} \cdot \begin{pmatrix} 4 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \cdot 4 + 3 \cdot 1 \\ 1 \cdot 4 + (-5) \cdot 1 \end{pmatrix} = \begin{pmatrix} 11 \\ -1 \end{pmatrix} \\ A \cdot \mathbf{c}_2(B) &= \begin{pmatrix} 2 & 3 \\ 1 & -5 \end{pmatrix} \cdot \begin{pmatrix} 3 \\ -2 \end{pmatrix} = \begin{pmatrix} 2 \cdot 3 + 3 \cdot (-2) \\ 1 \cdot 3 + (-5) \cdot (-2) \end{pmatrix} = \begin{pmatrix} 0 \\ 13 \end{pmatrix} \\ A \cdot \mathbf{c}_3(B) &= \begin{pmatrix} 2 & 3 \\ 1 & -5 \end{pmatrix} \cdot \begin{pmatrix} 6 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \cdot 6 + 3 \cdot 3 \\ 1 \cdot 6 + (-5) \cdot 3 \end{pmatrix} = \begin{pmatrix} 21 \\ -9 \end{pmatrix} \end{aligned}$$

Таким образом получили

$$AB = \begin{pmatrix} 11 & 0 & 21 \\ -1 & 13 & -9 \end{pmatrix}.$$

3.3 Лекция #8. Обратная матица

Особенный случай линейных отображений возникает в случае $\varphi : \mathbf{V} \to \mathbf{V}$. В таком случае, линейное отображение φ называется *линейным оператором на пространстве* $\mathbf{V}$.

Определение 3.3.1. В случае когда рассматривается линейное отображение $\varphi : \mathbf{V} \to \mathbf{V}$, то оно называется линейным оператором. Если $\psi : \mathbf{V} \to \mathbf{V}$ такой линейный оператор, что $\varphi \circ \psi = \psi \circ \varphi = \mathrm{id}_{\mathbf{V}}$, то ψ называется *обратным к* φ и обозначается так $\psi := \varphi^{-1}$.

Важное замечание. Мы ограничиваемся случаем, когда пространство $\mathbf{V}$ конечномерное, $\dim(\mathbf{V}) = n < \infty$, в таком случае, согласно Теореме 2.4.2, имеем $\mathbf{V} \cong \mathbb{R}^n$. Далее, согласно Теореме 3.1.1, наш оператор задаётся квадратной матрицей $A \in \mathrm{Mat}_{n \times n}(\mathbb{R})$.

Определение 3.3.2. Пусть дан линейный оператор $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ и A – его матрица, пусть у него имеется обратный φ^{-1}, тогда матрица обратного оператора называется обратной матрицей к A и обозначается так A^{-1}.

Наблюдение 3.3 В силу определения обратного линейного оператора, и так как матрица у тождественного оператора $\mathrm{id}_{\mathbb{R}^n}$ это единичная матрица, то обратную матрицу можно определить следующим образом, не прибегая к линейным операторам.

Определение 3.3.3. Говорят, что квадратная матрица A *обратима* если существует такая квадратная матрица A', того же размера, что $AA' = A'A = E$. При этом A' принято обозначать так A^{-1}.

Важное замечание. На самом деле одно из равенств $AA' = E$, $A'A = E$ вытекает из другого.

Пусть $\mathbb{R}^n$ — конечномерное векторное пространство, и пусть $\varphi, \psi : \mathbb{R}^n \to \mathbb{R}^n$ — линейные операторы с матрицами A и A' соответственно. Предположим, что $\varphi \circ \psi = \mathrm{id}$, что в матричной форме означает $AA' = E$. Требуется доказать, что $\psi \circ \varphi = \mathrm{id}$, т.е. $A'A = E$.

Лемма 3.3.1. *Пусть V — конечномерное векторное пространство, и $T : V \to V$ — линейный оператор. Тогда T инъективен тогда и только тогда, когда T сюръективен.*

Доказательство. Пусть $\dim V = n$ и $\{e_1, \dots, e_n\}$ — базис V.
Если T инъективен, то векторы $T(e_1), \dots, T(e_n)$ линейно независимы. Поскольку в n-мерном пространстве любая система из n линейно независимых векторов образует базис, то $\{T(e_1), \dots, T(e_n)\}$ — базис V. Следовательно, $\operatorname{Im} T = V$, и T сюръективен.
Обратно, если T сюръективен, то векторы $T(e_1), \dots, T(e_n)$ образуют порождающую систему V. Так как их количество равно n, они линейно независимы. Значит, T инъективен. □

Теорема 3.3.2. *Если $\varphi \circ \psi = \mathrm{id}$, то $\psi \circ \varphi = \mathrm{id}$.*

Доказательство. Предположим, что $\varphi \circ \psi = \mathrm{id}$.

1. Покажем, что ψ инъективен. Пусть $\psi(x) = \psi(y)$. Тогда
$$\varphi(\psi(x)) = \varphi(\psi(y)) \Rightarrow (\varphi \circ \psi)(x) = (\varphi \circ \psi)(y) \Rightarrow \mathrm{id}(x) = \mathrm{id}(y) \Rightarrow x = y.$$
Следовательно, ψ инъективен.

2. По лемме, из инъективности ψ следует его сюръективность. Значит, ψ биективен, и существует обратный оператор ψ^{-1}.

3. Умножим равенство $\varphi \circ \psi = \mathrm{id}$ справа на ψ^{-1}:
$$(\varphi \circ \psi) \circ \psi^{-1} = \mathrm{id} \circ \psi^{-1} \Rightarrow \varphi \circ (\psi \circ \psi^{-1}) = \psi^{-1} \Rightarrow \varphi \circ \mathrm{id} = \psi^{-1} \Rightarrow \varphi = \psi^{-1}.$$
Таким образом, φ — обратный оператор к ψ, и поэтому $\psi \circ \varphi = \mathrm{id}$.

В матричной форме это означает, что из $AA' = E$ следует $A'A = E$. □

3.3.1 Линейные операторы в малых размерностях

Мы рассмотрим линейные операторы в размерностях 1 и 2 и исследуем некоторые их свойства.

Линейные операторы в размерности 1

Наблюдение 3.4 В таком случае линейный оператор $\varphi : \mathbb{R} \to \mathbb{R}$ задаётся матрицей 1×1, *т.е.*, числом, скажем $a \in \mathbb{R}$. Другими словами он описывается следующим образом

$$\varphi(x) := ax,$$

где $x \in \mathbb{R}$ – произвольное число, которое мыслиться как вектор в $\mathbb{R}$.

> Геометрически это означает растяжение прямой (или сжатие) в a раз. Где a это коэффициент растяжения если $a > 1$ или сжатия если $a < 1$.

Тогда, если $a \neq 0$, то мы можем рассмотрем обратное к нему отображение

$$\varphi^{-1}(x) = a^{-1}x,$$

для любого $x \in \mathbb{R}$. Ясно что оно линейное и $\varphi(\varphi^{-1}(x)) = \varphi^{-1}(\varphi(x)) = x$ для любого $x \in \mathbb{R}$.‘

> **Важное замечание.** Таким образом, линейный оператор $\varphi : \mathbb{R}^1 \to \mathbb{R}^1$ заданный формулой $\varphi(x) := ax$ имеет обратный тогда и только тогда, когда $a \neq 0$.

Геометрически это означает растяжение прямой (или сжатие) в a раз. Где a это коэффициент растяжения если $a > 1$ или сжатия если $0 < a < 1$.

Тогда, если $a \neq 0$, то мы можем рассмотреть обратное к нему отображение

$$\varphi^{-1}(x) = a^{-1}x,$$

для любого $x \in \mathbb{R}$. Ясно что оно линейное и $\varphi(\varphi^{-1}(x)) = \varphi^{-1}(\varphi(x)) = x$ для любого $x \in \mathbb{R}$.

Линейные операторы в размерности 2

Итак в случае когда мы линейно отображали прямую на себя, то мы поняли что это либо растяжение либо сжатие или же это отображение прямой в точку 0 (случай когда $a = 0$). В последнем случае мы это отображение обратить не сможем.

Теперь посмотрим что происходит когдм мы будем отображать плоскость на саму себя.

Мы начнём с примера.

Пример 3.7

Дана матрица $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Требуется найти обратную матрицу A^{-1}, решив соответствующие системы линейных уравнений.

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

По определению обратной матрицы выполняется условие:

$$A \cdot X = I,$$

где I – единичная матрица $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Перепишем это матричное уравнение в виде системы:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Это эквивалентно двум независимым системам линейных уравнений (по столбцам):

(1) **Столбец 1: Находим x_{11} и x_{21}** Первое уравнение (из первого столбца правой части):

$$\begin{cases} 1 \cdot x_{11} + 2 \cdot x_{21} = 1 \\ 3 \cdot x_{11} + 4 \cdot x_{21} = 0 \end{cases}$$

Решим эту систему методом исключения.

Умножим первое уравнение на 3:

$$3x_{11} + 6x_{21} = 3$$

Вычтем из него второе уравнение:

$$(3x_{11} + 6x_{21}) - (3x_{11} + 4x_{21}) = 3 - 0$$

$$2x_{21} = 3$$

$$x_{21} = -\frac{3}{2}$$

Подставим найденное x_{21} в первое уравнение системы:

$$1 \cdot x_{11} + 2 \cdot \left(-\frac{3}{2}\right) = 1$$

$$x_{11} - 3 = 1$$

$$x_{11} = 4$$

Таким образом, первый столбец обратной матрицы: $x_{11} = 4$, $x_{21} = -\frac{3}{2}$.

(2) **Столбец 2: Находим x_{12} и x_{22}**. Второе уравнение (из второго столбца правой части):

$$\begin{cases} 1 \cdot x_{12} + 2 \cdot x_{22} = 0 \\ 3 \cdot x_{12} + 4 \cdot x_{22} = 1 \end{cases}$$

Решим эту систему.

Умножим первое уравнение на 3:

$$3x_{12} + 6x_{22} = 0$$

Вычтем из него второе уравнение:

$$(3x_{12} + 6x_{22}) - (3x_{12} + 4x_{22}) = 0 - 1$$

$$2x_{22} = -1$$

$$x_{22} = \frac{1}{2}$$

Подставим найденное x_{22} в первое уравнение системы:

$$1 \cdot x_{12} + 2 \cdot \left(\frac{1}{2}\right) = 0$$

$$x_{12} + 1 = 0$$

$$x_{12} = -1$$

Таким образом, второй столбец обратной матрицы: $x_{12} = -1$, $x_{22} = \frac{1}{2}$.

Собирая найденные элементы в матрицу, получаем обратную матрицу:

$$A^{-1} = \begin{pmatrix} 4 & -1 \\ -\frac{3}{2} & \frac{1}{2} \end{pmatrix}$$

Для проверки можно выполнить умножение:

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

Рассмотрим теперь такой пример

Пример 3.8 Пусть линейный оператор $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ задан матрицей A,

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

это значит, что базисный вектор $\mathbf{e}_1$ отображается в вектор $(1{,}0)^\top$ (=первый столбец), а второй $\mathbf{e}_2$ отображется в нулевой вектор $(0{,}0)^\top$ (=второй столбец).

Допустим, что эта матрица имеет обратную $X = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}$, тогда мы получаем

$$AX = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} = \begin{pmatrix} x_{11} & x_{12} \\ 0 & 0 \end{pmatrix},$$

и мы видим что полученная матица не может быть равна единичной, поэтому у матрицы A обратной не имеется.

3.3.2 Нахождение обратной матрицы

Итак как мы уже убиделись обратная матрица существует не всегда, сейчас мы поймём как находиь обратную матрицу и помогут нам в этом наши знания о решении систем линейных уравнений.

Теорема 3.3.3. *Если квадратная матрица $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ обратима, тогда для каждого вектора $\mathbf{b} \in \mathbb{R}^n$ уравнение $A\mathbf{x} = \mathbf{b}$ имеет единственное решение $\mathbf{x} = A^{-1}\mathbf{b}$.*

Доказательство. Возьмём произвольный вектор $\mathbf{b} \in \mathbb{R}^n$ и рассмотрим вектор $\mathbf{x}_0 := A^{-1}\mathbf{b}$. Тогда, подставляя $\mathbf{x}_0$ в уравнение $A\mathbf{x} = \mathbf{b}$, получаем:

$$A \cdot (A^{-1}\mathbf{b}) = (AA^{-1})\mathbf{b} = E\mathbf{b} = \mathbf{b},$$

т.е. $\mathbf{x}_0 = A^{-1}\mathbf{b}$ – решение уравнения.

Докажем теперь его единственность. Пусть $\mathbf{x}_1$ – ещё одно решение уравнения: $A\mathbf{x}_1 = \mathbf{b}$. Умножим это равенство слева на A^{-1}:

$$A^{-1}(A\mathbf{x}_1) = A^{-1}\mathbf{b} \quad \Rightarrow \quad (A^{-1}A)\mathbf{x}_1 = A^{-1}\mathbf{b} \quad \Rightarrow \quad E\mathbf{x}_1 = A^{-1}\mathbf{b} \quad \Rightarrow \quad \mathbf{x}_1 = A^{-1}\mathbf{b} = \mathbf{x}_0,$$

что и доказывает единственность. □

3.3.3 Элементарные матрицы

С другой же стороны мы решаем что линую систему можно решеить методом Гаусса просто преобразую матрицу коэффициентов к ступенчатому виду. Оказывается что элементарные преобразования над строками матриц можно представить через умножения на специальные матрицы, которые называются *элементарными*.

Интуитивно понятно, что у элементарных преобразований есть обратные. Например, если $L_{3,1}(-5)A$ означает "прибавить к третьей строке первую строку, умноженную на -5 то обратная операция – "прибавить к третьей строке первую, умноженную на 5". То есть, $L_{3,1}^{-1}(-5) = L_{3,1}(5)$.

Определение 3.3.4. *Элементарные матрицы* – это квадратные матрицы размера $n \times n$ ($n \geq 1$) вида $L_{i,j}(\lambda)$, $T_{i,j}$, $D_i(\lambda)$, где $\lambda \in \mathbb{R}$. Они описываются следующим образом:

1. $L_{i,j}(\lambda)$ получается из единичной матрицы E путём прибавления к j-ой строке i-ой, умноженной на λ.

2. $T_{i,j}$ получается из единичной матрицы путём перемены местами i-ой и j-ой строк.

3. $D_i(\lambda)$ получается из единичной матрицы путём умножения её i-ой строки на число $\lambda \neq 0$.

Предложение 3.3.4. *Элементарные матрицы обратимы, и их обратные находятся по формулам:*

$$L_{i,j}(\lambda)^{-1} = L_{i,j}(-\lambda), \qquad T_{i,j}^{-1} = T_{i,j}, \qquad D_i^{-1}(\lambda) = D_i(\lambda^{-1}).$$

Доказательство. Это сразу следует из того, что обратное к элементарному преобразованию строк есть элементарное преобразование, отменяющее исходное. □

Здесь мы показываем, что элементарные преобразования строк можно выразить в терминах умножения на матрицы слева, а преобразования столбцов – умножением справа.

Рассмотрим пример.

Пример 3.9 Рассмотрим матрицы:

$$L_{1,3}(-4) := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{pmatrix}, \quad T_{1,2} := \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad D_3(5) := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

и возьмём произвольную матрицу

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}.$$

- **Умножение слева:**

$$L_{1,3}(-4)A = \begin{pmatrix} a & b & c \\ d & e & f \\ g-4a & h-4b & i-4c \end{pmatrix}.$$

Умножение матрицы A слева на $L_{1,3}(-4)$ эквивалентно прибавлению к третьей строке A первой строки, умноженной на -4.

$$T_{1,2}A = \begin{pmatrix} d & e & f \\ a & b & c \\ g & h & i \end{pmatrix}.$$

Умножение матрицы A слева на $T_{1,2}$ эквивалентно перестановке первой и второй строк.

$$D_3(5)A = \begin{pmatrix} a & b & c \\ d & e & f \\ 5g & 5h & 5i \end{pmatrix}.$$

Умножение матрицы A слева на $D_3(5)$ эквивалентно умножению третьей строки на 5.

- **Умножение справа:**

$$AL_{1,3}(-4) = \begin{pmatrix} a-4c & b & c \\ d-4f & e & f \\ g-4i & h & i \end{pmatrix}.$$

Умножение матрицы A справа на $L_{1,3}(-4)$ эквивалентно прибавлению к первому столбцу A третьего столбца, умноженного на -4.

$$AT_{1,2} = \begin{pmatrix} b & a & c \\ e & d & f \\ h & g & i \end{pmatrix}.$$

Умножение матрицы A справа на $T_{1,2}$ эквивалентно перестановке первого и второго столбцов.

$$AD_3(5) = \begin{pmatrix} a & b & 5c \\ d & e & 5f \\ g & h & 5i \end{pmatrix}.$$

Умножение матрицы A справа на $D_3(5)$ эквивалентно умножению третьего столбца на 5.

3.3.4 Критерий обратимости матрицы

Напоминание.

Определение 3.3.5. *Элементарными преобразованиями строк* в произвольной матрице называются следующие операции:

(1) Прибавить к одной строке другую строку, умноженную на произвольное число.

(2) Поменять местами две строки.

(3) Умножить строку на произвольное ненулевое число.

Две матрицы M, M' называются *эквивалентными по строкам*, если M' получается из M с помощью конечной последовательности элементарных преобразований строк.

Теорема 3.3.5. *Матрица $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ обратима тогда и только тогда, когда она эквивалентна по строкам единичной матрице E.*

Доказательство.

($\Rightarrow$) Пусть A обратима. Тогда, согласно Теореме 3.3.3, система $A\mathbf{x} = \mathbf{b}$ имеет единственное решение для любого $\mathbf{b}$. Это означает, что метод Гаусса (состоящий из элементарных преобразований строк над расширенной матрицей) приводит систему к виду $E\mathbf{x} = \mathbf{b}'$. Следовательно, элементарные преобразования строк переводят матрицу A в единичную матрицу E, то есть A эквивалентна E.

($\Leftarrow$) Пусть A эквивалентна единичной матрице E. Это значит, что существует последовательность элементарных преобразований строк, переводящая A в E. Каждое элементарное преобразование строк соответствует умножению слева на соответствующую элементарную матрицу. Значит, существует такая последовательность элементарных матриц $U_1, U_2, \dots, U_p$, что:

$$U_p U_{p-1} \cdots U_1 A = E.$$

Обозначим $U = U_p U_{p-1} \cdots U_1$. Тогда $UA = E$, откуда $A = U^{-1}$. Следовательно, A обратима, и $A^{-1} = U$.

□

Следствие 3.3.6 (Алгоритм нахождения обратной матрицы). *Чтобы найти обратную матрицу к матрице A (или установить, что её не существует), нужно:*

1. *Составить расширенную матрицу $(A \mid E)$.*
2. *С помощью элементарных преобразований* ***строк*** *привести левый блок (A) к единичной матрице E.*
3. *Если это удалось, то правый блок автоматически станет матрицей A^{-1}.*
4. *Если левый блок привести к E невозможно (например, получится нулевая строка), то матрица A необратима.*

Обоснование алгоритма. Элементарные преобразования строк над расширенной матрицей $(A \mid E)$ эквивалентны умножению слева на некоторую матрицу U:

$$U(A \mid E) = (UA \mid UE).$$

Если нам удалось привести A к E, то $UA = E$, значит $U = A^{-1}$. Но тогда $UE = A^{-1}E = A^{-1}$. Таким образом, в правой части получается A^{-1}. □

Пример 3.10 Найдём обратную матрицу для $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$.

Составим расширенную матрицу и будем выполнять преобразования:

$$(A \mid E) = \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array}\right)$$

1. Преобразование: $R_2 \to R_2 - 3R_1$ (прибавляем к второй строке первую, умноженную на -3).
$$\left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -2 & -3 & 1 \end{array}\right)$$
2. Преобразование: $R_2 \to R_2 \cdot (-\frac{1}{2})$ (умножаем вторую строку на $-\frac{1}{2}$).
$$\left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & 1.5 & -0.5 \end{array}\right)$$
3. Преобразование: $R_1 \to R_1 - 2R_2$ (прибавляем к первой строке вторую, умноженную на -2).
$$\left(\begin{array}{cc|cc} 1 & 0 & -2 & 1 \\ 0 & 1 & 1.5 & -0.5 \end{array}\right)$$

Левая часть стала единичной матрицей. Значит, обратная матрица находится справа:

$$A^{-1} = \begin{pmatrix} -2 & 1 \\ 1.5 & -0.5 \end{pmatrix}.$$

Убедимся, что это совпадает с формулой для обратной матрицы 2x2:

$$\det(A) = 1 \cdot 4 - 2 \cdot 3 = -2, \quad A^{-1} = \frac{1}{-2}\begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ 1.5 & -0.5 \end{pmatrix}.$$

Глава 4

Детерминант матрицы

В этой главе мы позкомимся к констукцией бивекторов и не только, но для начала нам нужна важная техника, с которой мы и начнём.

4.1 Перестановки и конфигурации Тёрстона

Мы здесь говорим про перестановки но с более геометрической стороны дела. Для определения знака перестановки существует много способов, но на мой взгляд самый понятный это геометрический которых можно понимать себе как пересечения линий которые схематично показывают биекцию (=перестановку). Чтобы всё это аккуратно сформулировать мы будем следовать одному подходу который принадлежит Уильяму Тёрстону и описанному в Chapter 9 книги Word Processing in Groups.

4.1.1 Основные понятия

Мы начнём с напоминания что такое перестановка и сразу перейдём на геометрический язык.

Определение 4.1.1. *Перестановкой* множества $\{1, 2, \dots, n\}$ называется биекция

$$\pi : \{1, 2, \dots, n\} \to \{1, 2, \dots, n\}.$$

Множество всех перестановок на множестве из n элментнов мы обозначаем через $\mathfrak{S}_n$.

Под *знаком* перестановки $\sigma \in \mathfrak{S}_n$ мы понимаем чётность числа её всех *инверсий*. Инверсия это такая пара (i,j), где $1 \leq i < j \leq n$, но $\sigma(i) > \sigma(j)$.

Пример 4.1 Пусть $X = \{1,2,3,4,5\}$, тогда биекция $\pi : X \to X$ заданное правилами

$$\pi(1) = 3, \quad ,\pi(2) = 4, \quad \pi(3) = 5, \quad \pi(4) = 2, \quad \pi(5) = 1$$

и задаёт перестановку. Это правило удобно записать так

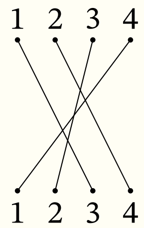
$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 2 & 1 \end{pmatrix}$$

Наблюдение 4.1 Таким образом, любую перестановку π на множестве из n элементов можно записывать в матричной форме

$$\pi = \begin{pmatrix} 1 & 2 & \dots & n \\ \pi(1) & \pi(2) & \dots & \pi(n) \end{pmatrix}$$

Такую биекцию удобно представлять себе геометрически: расположим числа $1, 2, \dots, n$ в верхнем ряду и те же числа — в нижнем ряду, а затем соединим каждое число i в верхнем ряду с числом $\pi(i)$ в нижнем ряду непрерывной линией.

Пример 4.2



Геометрическое представление перестановки $\pi = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 2 & 1 \end{pmatrix}$. Каждое пересечение линий соответствует инверсии в перестановке.

Важное замечание. При таком представлении линии, соединяющие элементы, могут пересекаться. Оказывается, что эти пересечения не случайны — они кодируют важные комбинаторные свойства перестановки. В частности, каждое пересечение соответствует *инверсии* — паре (i, j) такой, что $i < j$, но $\pi(i) > \pi(j)$.

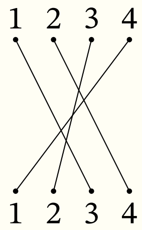
Историческое замечание. У. Тёрстон предложил рассматривать такие конфигурации линий как “плоские тени” более сложных объектов — *кос*. В теории кос естественно возникают так называемые *простые косы*, в которых любые две нити переплетаются не более одного раза. Нас будет интересовать “плоский” аналог этих кос, и, как мы увидим далее, этот подход окажется очень полезным для изучения перестановок.

Определение 4.1.2. Для произвольной перестановки $\pi \in \mathfrak{S}_n$ рассмотрим множество пар $R(\pi) \subseteq \{1, \dots, n\} \times \{1, \dots, n\}$, определённое следующим образом

$$R(\pi) := \{(i, j) \ : \ i < j, \ \pi(i) > \pi(j)\}.$$

Важное замечание. Другими словами множество $R(\pi)$ – это множество всех инверсий в перестановке.

Пример 4.3
Например, для перестановки, изображённой ниже,



находим

$$R(\pi) = \{(1,3),(1,4),(2,3),(2,4),(3,4)\}$$

это означает что все эти пары образуют множество всех инверсий в этой перестановке.

Нетрудно видеть, что каждой паре из $R(\pi)$ соответствует пара пересекающихся нитей в геометрическом представлении перестановки.

Лемма 4.1.1. *Множество $R \subseteq \{1,\dots,n\}\times\{1,\dots,n\}$ пар (i,j), где $i < j$, будет множеством вида $R(\pi)$ для некоторой перестановки $\pi \in \mathfrak{S}_n$ если и только если выполняются следующие два условия:*

(1) если $(i,j) \in R$ и $(j,k) \in R$, то $(i,k) \in R$,

(2) если $(i,k) \in R$, то $(i,j) \in R$ или $(j,k) \in R$ для каждого $i < j < k$.

Приведённое ниже доказательство является очень техническим и большую ценность для дальнейшего не представляет, тем не менее для полноты подачи материала я решил его привести.

Доказательство. Докажем эквивалентность: множество R является множеством инверсий некоторой перестановки π тогда и только тогда, когда выполняются условия 1 и 2. **Необходимость.** Пусть $R = R(\pi)$ для некоторой перестановки $\pi \in \mathfrak{S}_n$. Тогда:

1. Если $(i,j) \in R$ и $(j,k) \in R$, то $\pi(i) > \pi(j)$ и $\pi(j) > \pi(k)$, откуда $\pi(i) > \pi(k)$, так что $(i,k) \in R$.

2. Если $(i,k) \in R$, то $\pi(i) > \pi(k)$. Рассмотрим j такой, что $i < j < k$. Если бы ни $(i,j) \in R$, ни $(j,k) \in R$, то было бы $\pi(i) < \pi(j)$ и $\pi(j) < \pi(k)$, откуда $\pi(i) < \pi(k)$, что противоречит $(i,k) \in R$. Значит, либо $(i,j) \in R$, либо $(j,k) \in R$.

Достаточность. Пусть R удовлетворяет условиям 1 и 2. Определим бинарное отношение $\succ$ на множестве $\{1,\dots,n\}$ следующим образом: для любых $i \neq j$ полагаем

$$i \succ j \quad \text{тогда и только тогда, когда} \quad (\min(i,j), \max(i,j)) \in R.$$

Более подробно:

- Если $i < j$, то $i \succ j$ означает $(i,j) \in R$, а $j \succ i$ означает $(i,j) \notin R$.
- Если $i > j$, то $i \succ j$ означает $(j,i) \notin R$, а $j \succ i$ означает $(j,i) \in R$.

Отношение $\succ$ тотально по построению. Покажем, что оно транзитивно.

Рассмотрим i, j, k такие, что $i \succ j$ и $j \succ k$. Нужно показать, что $i \succ k$. Разберем все возможные случаи взаимного порядка i, j, k.

1. **Случай 1:** $i < j < k$. Тогда $i \succ j$ означает $(i,j) \in R$, $j \succ k$ означает $(j,k) \in R$. По условию 1, $(i,k) \in R$, значит $i \succ k$.

2. **Случай 2:** $i < k < j$. Тогда $i \succ j$ означает $(i,j) \in R$, $j \succ k$ означает $(k,j) \notin R$ (так как $k < j$). Применим условие 2 к тройке (i,k,j): из $(i,j) \in R$ следует, что $(i,k) \in R$ или $(k,j) \in R$. Но $(k,j) \notin R$, поэтому $(i,k) \in R$, значит $i \succ k$.

3. **Случай 3:** $j < i < k$. Тогда $i \succ j$ означает $(j,i) \notin R$, $j \succ k$ означает $(j,k) \in R$. Применим условие 2 к тройке (j,i,k): из $(j,k) \in R$ следует, что $(j,i) \in R$ или $(i,k) \in R$. Но $(j,i) \notin R$, поэтому $(i,k) \in R$, значит $i \succ k$.

4. **Случай 4:** $j < k < i$. Тогда $i \succ j$ означает $(j,i) \notin R$, $j \succ k$ означает $(j,k) \in R$. Предположим, что $k \succ i$, т.е. $(k,i) \in R$ (так как $k < i$). Тогда из $(j,k) \in R$ и $(k,i) \in R$ по условию 1 следует $(j,i) \in R$, что противоречит $(j,i) \notin R$. Значит, $i \succ k$.

5. **Случай 5:** $k < i < j$. Тогда $i \succ j$ означает $(i,j) \in R$, $j \succ k$ означает $(k,j) \notin R$. Предположим, что $k \succ i$, т.е. $(k,i) \in R$ (так как $k < i$). Тогда из $(k,i) \in R$ и $(i,j) \in R$ по условию 1 следует $(k,j) \in R$, что противоречит $(k,j) \notin R$. Значит, $i \succ k$.

6. **Случай 6:** $k < j < i$. Тогда $i \succ j$ означает $(j,i) \notin R$, $j \succ k$ означает $(k,j) \notin R$. Предположим, что $k \succ i$, т.е. $(k,i) \in R$. Применим условие 2 к тройке (k,j,i): из $(k,i) \in R$ следует, что $(k,j) \in R$ или $(j,i) \in R$. Но оба $(k,j) \notin R$ и $(j,i) \notin R$, противоречие. Значит, $i \succ k$.

Таким образом, $\succ$ — строгий линейный порядок на $\{1, \ldots, n\}$. Пусть $a_1, a_2, \ldots, a_n$ — единственная перестановка такая, что $a_1 \prec a_2 \prec \cdots \prec a_n$, где $\prec$ — отношение, обратное к $\succ$. Определим $\pi(a_i) = i$. Тогда для любых $i < j$ имеем:

$$(i,j) \in R \quad \Leftrightarrow \quad i \succ j \quad \Leftrightarrow \quad \pi(i) > \pi(j),$$

так что $R = R(\pi)$. Это завершает доказательство. □

Эти условия позволяют нам дать следующее геометрическое определение:

Конструкция 4.1 Для любой перестановки $\pi \in \mathfrak{S}_n$ рассмотрим конфигурацию n прямых на плоскости $\mathbb{R}^2$, пронумерованных числами $1, \ldots, n$. Обозначим эти прямые $\ell_1, \ldots, \ell_n$ соответственно. Положим, что прямые ℓ_i и ℓ_j пересекаются в одной точке (*соотв.* не пересекаются) если и только если $(i,j) \in R(\pi)$ (*соотв.* $(i,j) \notin R(\pi)$). Полученную конфигурацию и любой её диффеоморфный образ будем называть *конфигурацией Тёрстона перестановки π*.

Важное замечание. Согласно Лемме 4.1.1 эта конструкция корректна — условия на множество $R(\pi)$ гарантируют, что соответствующая конфигурация прямых действительно существует.

4.1.2 Умножение перестановок

Прежде чем говорить об умножении, важно договориться о порядке.

Важное замечание. В теории групп перестановки почти всегда действуют **справа налево**. То есть, для элементов x и перестановок σ, τ композиция $\sigma\tau$ определяется как $(\sigma\tau)(x) = \sigma(\tau(x))$. Сначала применяется та перестановка, которая стоит справа (τ), а затем — та, что стоит слева (σ).

Определение 4.1.3. Для перестановок $\sigma, \tau \in \mathfrak{S}_n$ их *произведение* $\sigma\tau$ — это композиция функций:

$$(\sigma\tau)(i) = \sigma(\tau(i)) \quad \text{для всех } i \in \{1, \dots, n\}.$$

Пример 4.4 Рассмотрим перестановки $\tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 2 & 1 \end{pmatrix}$ и $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix}$.

Найдём их произведение $\sigma\tau$, применяя их справа налево:

$$\begin{aligned}
(\sigma\tau)(1) &= \sigma(\tau(1)) = \sigma(3) = 1, \\
(\sigma\tau)(2) &= \sigma(\tau(2)) = \sigma(4) = 3, \\
(\sigma\tau)(3) &= \sigma(\tau(3)) = \sigma(2) = 4, \\
(\sigma\tau)(4) &= \sigma(\tau(4)) = \sigma(1) = 2.
\end{aligned}$$

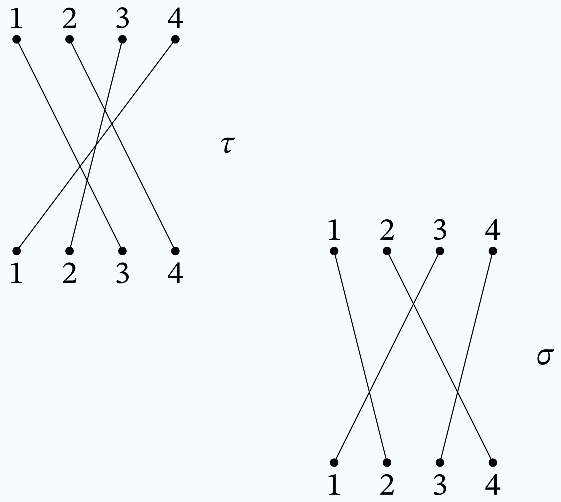
Таким образом,

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

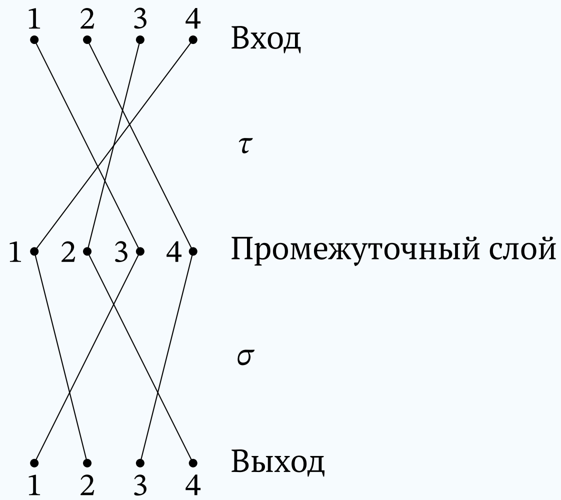
Теперь мы подходим к самому интересному: как выглядит множество инверсий произведения $\sigma\tau$ с геометрической точки зрения? Подход через конфигурации Тёрстона даёт изящный ответ.

Наблюдение 4.2 *Как геометрически интерпритировать умножения перестановок?*

Представим перестановки τ и σ их конфигурациями Тёрстона. Расположим эти конфигурации друг под другом, совместив их “нижний” ряд точек первой конфигурации с “верхним” рядом точек второй.



Чтобы получить конфигурацию для $\sigma\tau$, мы можем мысленно “склеить” эти две картинки, отождествив нижний ряд первой с верхним рядом второй. При этом линии из обеих конфигураций объединятся в непрерывные кривые, идущие от самого верха (чисел $1, \dots, n$) до самого низа (чисел $\sigma(\tau(1)), \dots, \sigma(\tau(n))$).



Теперь ключевое наблюдение: **пересечение в итоговой конфигурации для** $\sigma\tau$ (между верхом и низом) происходит в двух случаях:

1. **Пересечение было в** τ (в верхней половине), но **не было “распутано” в** σ (в нижней половине) для соответствующих нитей.
2. **Пересечения не было в** τ, но **оно появилось в** σ для соответствующих нитей.

А что значит “для соответствующих нитей”?

Обратите внимание: чтобы понять, как две нити i и j ведут себя в нижней конфигурации (σ), нам нужно знать, из каких “портов” промежуточного слоя они выходят. А выходят они из портов $\tau(i)$ и $\tau(j)$.

Таким образом, пересечение между нитями i и j в нижней половине существует тогда и только тогда, когда $(\tau(i), \tau(j)) \in R(\sigma)$, при условии $\tau(i) < \tau(j)$. Но в нашем множестве $R(\sigma)$ пары всегда упорядочены по возрастанию.

Поэтому условие $(\tau(i), \tau(j)) \in R(\sigma)$ эквивалентно тому, что $\tau(i) > \tau(j)$, то есть $(i,j) \in R(\tau)$. Это ровно то, что кодирует множество $\tau^{-1}R(\sigma)$.

Определение 4.1.4. Для множества пар $R \subseteq \{1, \dots, n\} \times \{1, \dots, n\}$ и перестановки τ определим *прообраз* $\tau^{-1}R$ как множество всех пар (i,j) с $i < j$, для которых $(\tau(i), \tau(j)) \in R$.

Пример 4.5 Для σ из нашего примера $R(\sigma) = \{(1,3), (2,3), (2,4)\}$. Для τ имеем $\tau(1) = 3, \tau(2) = 4, \tau(3) = 2, \tau(4) = 1$. Вычислим $\tau^{-1}R(\sigma)$:

- Пара $(1,2) : (\tau(1), \tau(2)) = (3,4)$, $(3,4) \notin R(\sigma)$. Значит, $(1,2) \notin \tau^{-1}R(\sigma)$.
- Пара $(1,3) : (\tau(1), \tau(3)) = (3,2)$. Поскольку $3 > 2$, мы смотрим на упорядоченную пару $(2,3)$. $(2,3) \in R(\sigma)$. Значит, $(1,3) \in \tau^{-1}R(\sigma)$.
- Пара $(1,4)$: $(\tau(1), \tau(4)) = (3,1) \to (1,3) \in R(\sigma)$. Значит, $(1,4) \in \tau^{-1}R(\sigma)$.
- Пара $(2,3)$: $(\tau(2), \tau(3)) = (4,2) \to (2,4) \in R(\sigma)$. Значит, $(2,3) \in \tau^{-1}R(\sigma)$.
- Пара $(2,4)$: $(\tau(2), \tau(4)) = (4,1) \to (1,4) \notin R(\sigma)$. Значит, $(2,4) \notin \tau^{-1}R(\sigma)$.
- Пара $(3,4) : (\tau(3), \tau(4)) = (2,1) \to (1,2) \notin R(\sigma)$. Значит, $(3,4) \notin \tau^{-1}R(\sigma)$.

Итак, $\tau^{-1}R(\sigma) = \{(1,3), (1,4), (2,3)\}$.

Комментарий. Другими словами множество $\tau^{-1}R(\sigma)$ — это в точности те пары нитей на входе, которые **пересекаются в нижней конфигурации (σ)**.

Множество $R(\tau)$ — это те пары нитей на входе, которые **пересекаются в верхней конфигурации (τ)**.

Итоговое пересечение в полной конфигурации $\sigma\tau$ происходит тогда и только тогда, когда пересечение есть ровно в одной из двух половин: либо только вверху, либо только внизу. Это в точности операция **симметрической разности** множеств.

Это рассуждение приводит нас к следующему важному результату.

Теорема 4.1.2 (Формула Тёрстона для инверсий)**.** *Для любых перестановок $\tau, \sigma \in \mathfrak{S}_n$ множество инверсий их произведения задаётся формулой:*

$$R(\sigma\tau) = \left(R(\tau) \setminus \tau^{-1}R(\sigma)\right) \cup \left(\tau^{-1}R(\sigma) \setminus R(\tau)\right).$$

Иными словами, $R(\sigma\tau)$ — это симметрическая разность множеств $R(\tau)$ и $\tau^{-1}R(\sigma)$:

$$R(\sigma\tau) = R(\tau) \bigtriangleup \tau^{-1}R(\sigma).$$

Пример 4.6 В нашем примере:

- $R(\tau) = \{(1,3), (1,4), (2,3), (2,4), (3,4)\}$
- $\tau^{-1}R(\sigma) = \{(1,3), (1,4), (2,3)\}$
- $R(\tau) \setminus \tau^{-1}R(\sigma) = \{(2,4), (3,4)\}$
- $\tau^{-1}R(\sigma) \setminus R(\tau) = \varnothing$
- $R(\sigma\tau) = \{(2,4), (3,4)\}$

Давайте проверим по определению для $\sigma\tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 2 \end{pmatrix}$:

- (1,2): 1 < 3? Да. Не инверсия.
- (1,3): 1 < 4? Да. Не инверсия.
- (1,4): 1 < 2? Да. Не инверсия.
- (2,3): 3 < 4? Да. Не инверсия.
- (2,4): 3 < 2? Нет. Инверсия (2,4).
- (3,4): 4 < 2? Нет. Инверсия (3,4).

Действительно, $R(\sigma\tau) = \{(2,4), (3,4)\}$. Формула работает.

Важное замечание. Формула Тёрстона не только элегантна, но и крайне полезна. Заметьте, что мощность симметрической разности связана с мощностями исходных множеств:

$$|A \triangle B| = |A| + |B| - 2|A \cap B|.$$

Применяя это к нашей формуле, получаем:

$$|R(\sigma\tau)| = |R(\tau)| + |\tau^{-1}R(\sigma)| - 2|R(\tau) \cap \tau^{-1}R(\sigma)|.$$

Но $|\tau^{-1}R(\sigma)| = |R(\sigma)|$, так как τ — биекция. Следовательно,

$$|R(\sigma\tau)| = |R(\tau)| + |R(\sigma)| - 2|R(\tau) \cap \tau^{-1}R(\sigma)|.$$

Это показывает, что число инверсий в произведении может быть как больше, так и меньше суммы чисел инверсий сомножителей, в зависимости от размера пересечения $|R(\tau) \cap \tau^{-1}R(\sigma)|$.

Доказательство теоремы. (1) Пусть $(i,j) \in R(\sigma\tau)$ тогда $i < j$ и $\sigma\tau(i) > \sigma\tau(j)$.

i) Если $\tau(i) > \tau(j)$, то $(i,j) \in R(\tau)$ и $(\tau(i), \tau(j)) \notin R(\sigma)$. Тогда $(i,j) \notin \tau^{-1}R(\sigma)$.

ii) Если $\tau(i) < \tau(j)$, то $(i,j) \notin R(\tau)$ и $(\tau(i), \tau(j) \in R(\sigma)$, и тогда $(i,j) \in \tau^{-1}R(\sigma)$.

Это значит, что если $(i,j) \in R(\sigma\tau)$ то $(i,j) \in R(\tau) \cup \tau^{-1}R(\sigma) \setminus R(\tau) \cap \tau^{-1}R(\sigma)$, *т.е.*, $R(\sigma\tau) \subseteq \big(\tau^{-1}R(\sigma) \setminus R(\tau)\big) \cup \big(R(\tau) \setminus \tau^{-1}R(\sigma)\big)$.

(2) Пусть $(i,j) \in \big(R(\tau) \setminus \tau^{-1}R(\sigma)\big) \cup \big(\tau^{-1}R(\sigma) \setminus R(\tau)\big)$.

i) Если $(i,j) \in R(\tau) \setminus \tau^{-1}R(\sigma)$, то $\tau(i) > \tau(j)$ и $(i,j) \notin \tau^{-1}R(\sigma)$. Тогда $\sigma\tau(i) > \sigma\tau(j)$, а это значит, что $(i,j) \in R(\sigma\tau)$.

ii) Если $(i,j) \in \tau^{-1}R(\sigma) \setminus R(\tau)$, то $\tau(i) < \tau(j)$ и $(\sigma\tau(i), \sigma\tau(j)) \in R(\sigma\tau)$, следовательно $(i,j) \in R(\sigma\tau)$.

Тогда получаем, что $\big(\tau^{-1}R(\sigma) \setminus R(\tau)\big) \cup \big(R(\tau) \setminus \tau^{-1}R(\sigma)\big) \subseteq R(\sigma\tau)$. Это завершает доказательство. □

Следствие 4.1.3. *$R(\pi^{-1}) = \pi R(\pi)$, для любой перестановки $\pi \in \mathfrak{S}_n$.*

Доказательство. Действительно, пусть $\sigma = \pi$, $\tau = \pi^{-1}$, тогда

$$R(\pi\pi^{-1}) = \big(\pi R(\pi) \setminus R(\pi^{-1})\big) \cup \big(R(\pi^{-1}) \setminus \pi R(\pi)\big)$$

и так как $R(\pi\pi^{-1}) = \varnothing$, то мы получаем требуемое. □

Определение 4.1.5. Для перестановки

$$\omega = \begin{pmatrix} 1 & 2 & \dots & n-1 & n \\ n & n-1 & \dots & 2 & 1 \end{pmatrix},$$

соответствующее множество $R(\omega)$ будем называть *элементом Гарсайда* и обозначать через Δ_n.

Нетрудно видеть, что $\Delta_n = \{(i,j) \;:\; 1 \leq i < j \leq n\}$. Тогда, для любой перестановки $\pi \in \mathfrak{S}_n$, из Леммы 4.1.2 вытекает

$$\begin{aligned} R(\omega\pi) &= \big(\pi^{-1}\Delta_n \setminus R(\pi)\big) \cup \big(R(\pi) \setminus \pi^{-1}\Delta_n\big) \\ &= (\Delta_n \setminus R(\pi)) \cup (R(\pi) \setminus \Delta_n) \\ &= (\Delta_n \setminus R(\pi)) \cup \varnothing \\ &= \Delta_n \setminus R(\pi). \end{aligned}$$

Для любой перестановки $\pi \in \mathfrak{S}_n$, мы положим

$$\neg R(\pi) := R(\omega\pi) = \Delta_n \setminus R(\pi).$$

Важное замечание. Ясно, что для данной перестановки $\pi \in \mathfrak{S}_n$ множество $\neg R(\pi)$ может быть также описано следующим образом

$$\neg R(\pi) = \{(i,j) \mid i < j,\ \pi(i) < \pi(j)\}.$$

Пример 4.7 Рассмотрим перестановку

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

Тогда получаем

$$\begin{aligned} R(\pi) &= \{(2{,}3), (3{,}6), (3{,}4)\}, \\ \neg R(\pi) &= \{(1{,}2), (1{,}3), (1{,}4), (1{,}5), (1{,}6), (2{,}5), (2{,}6), (3{,}5), (3{,}6), (4{,}5), (4{,}6), (5{,}6)\}. \end{aligned}$$

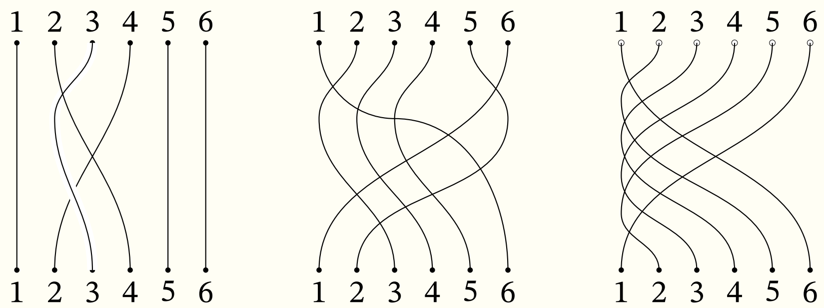
Далее, имеем

$$\omega\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 5 & 4 & 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 4 & 3 & 2 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 3 & 4 & 5 & 6 & 1 \end{pmatrix}$$

Тогда

$$R(\omega\pi) = \{(1{,}2), (1{,}3), (1{,}4), (1{,}5), (1{,}6), (2{,}5), (2{,}6), (3{,}5), (3{,}6), (4{,}5), (4{,}6), (5{,}6)\},$$

т.е., $R(\omega\pi) = \neg R(\pi)$. Образ соответствующей конфигурации Тёртсона для перестановки π показан ниже



$a)$ $b)$ $c)$

На этом рисунке, для перестановки $\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 4 & 3 & 2 & 5 & 6 \end{pmatrix}$ изображены:

a) $R(\pi)$

b) $\neg R(\pi) = R(\omega\pi)$

c) элемент Гарсайда Δ_6.

4.1.3 Знак перестановки

Как известно, *знак перестановки* $\pi \in \mathfrak{S}_n$ определяется как

$$\operatorname{sgn}(\pi) = (-1)^{|R(\pi)|},$$

где $|R(\pi)|$ — число инверсий в перестановке π.

С помощью формулы Тёрстона (Лемма 4.1.2) мы можем дать простое доказательство того, что знак произведения перестановок равен произведению их знаков.

Теорема 4.1.4. *Для любых перестановок $\sigma, \tau \in \mathfrak{S}_n$ выполняется равенство*

$$\operatorname{sgn}(\sigma\tau) = \operatorname{sgn}(\sigma) \operatorname{sgn}(\tau).$$

Доказательство. Из формулы Тёрстона (Лемма 4.1.2) имеем:

$$R(\sigma\tau) = \left(\tau^{-1}R(\sigma) \setminus R(\tau)\right) \cup \left(R(\tau) \setminus \tau^{-1}R(\sigma)\right).$$

Заметим, что объединение в правой части дизъюнктно. Поэтому

$$|R(\sigma\tau)| = |\tau^{-1}R(\sigma) \setminus R(\tau)| + |R(\tau) \setminus \tau^{-1}R(\sigma)|.$$

Обозначим $A = \tau^{-1}R(\sigma)$ и $B = R(\tau)$. Тогда

$$|R(\sigma\tau)| = |A \setminus B| + |B \setminus A| = |A| + |B| - 2|A \cap B|.$$

Но $|A| = |\tau^{-1}R(\sigma)| = |R(\sigma)|$, так как действие перестановки τ^{-1} на пары (i,j) является биекцией. Следовательно,

$$|R(\sigma\tau)| = |R(\sigma)| + |R(\tau)| - 2|A \cap B|.$$

Теперь вычислим знак:

$$\operatorname{sgn}(\sigma\tau) = (-1)^{|R(\sigma\tau)|} = (-1)^{|R(\sigma)|+|R(\tau)|-2|A\cap B|} = (-1)^{|R(\sigma)|+|R(\tau)|} = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau),$$

поскольку $-2|A \cap B|$ — чётное число и не влияет на знак. □

Важное замечание. Это доказательство демонстрирует, как конфигурации Тёрстона и связанные с ними комбинаторные структуры позволяют естественным образом выводить классические свойства перестановок. В частности, формула Тёрстона даёт явное описание инверсий в произведении перестановок, что непосредственно приводит к мультипликативности знака.

4.1.4 Циклическая запись перестановки

Геометрическое представление перестановок и формула Тёрстона дают нам глубокое понимание структуры перестановок. Теперь рассмотрим другой мощный способ описания перестановок — *циклическую запись*, которая часто оказывается более компактной и удобной для вычислений.

Определение 4.1.6. *Циклом* длины k называется перестановка π, для которой существуют различные элементы $a_1, a_2, \dots, a_k$ такие, что

$$\pi(a_1) = a_2, \quad \pi(a_2) = a_3, \quad \dots, \quad \pi(a_{k-1}) = a_k, \quad \pi(a_k) = a_1,$$

а все остальные элементы остаются на месте. Такой цикл обозначается $(a_1 a_2 \dots a_k)$.

Пример 4.8 Рассмотрим перестановку $\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 2 & 1 \end{pmatrix}$. Найдём её циклическое разложение.
Начнем с элемента 1:

$$\pi(1) = 3, \quad \pi(3) = 5, \quad \pi(5) = 1.$$

Получили цикл $(1\,3\,5)$. Теперь возьмем следующий непросмотренный элемент — 2:

$$\pi(2) = 4, \quad \pi(4) = 2.$$

Получили цикл $(2\,4)$. Таким образом,

$$\pi = (1\,3\,5)(2\,4).$$

Важное замечание. Циклы, не имеющие общих элементов, называются *непересекающимися*. Непересекающиеся циклы коммутируют. В циклической записи обычно опускают циклы длины 1 (неподвижные элементы). Таким образом, перестановка записывается как произведение непересекающихся циклов.

Лемма 4.1.5. *Любая перестановка может быть единственным образом (с точностью до порядка циклов) разложена в произведение непересекающихся циклов.*

Доказательство. Доказательство конструктивное: мы можем построить циклы, как в примере выше, начиная с произвольного элемента и последовательно применяя перестановку, пока не вернемся к исходному элементу. Затем берем следующий непросмотренный элемент и повторяем. Так как множество конечно, процесс завершится, и мы получим разложение на непересекающиеся циклы.

Единственность (с точностью до порядка циклов) следует из того, что циклы определяются орбитами действия перестановки на множестве. □

Циклическая запись полезна, например, для нахождения порядка перестановки и для вычисления знака перестановки.

Лемма 4.1.6. *Знак перестановки π равен $(-1)^{n-c}$, где c — количество циклов в разложении π (включая циклы длины 1).*

Доказательство. Заметим, что знак цикла длины k равен $(-1)^{k-1}$. Действительно, цикл длины k можно представить как произведение $k-1$ транспозиций (например, $(a_1 a_2 \ldots a_k) = (a_1 a_2)(a_2 a_3)\ldots(a_{k-1} a_k)$). Тогда знак цикла равен $(-1)^{k-1}$.

Пусть перестановка π разложена в произведение непересекающихся циклов с длинами $k_1, k_2, \ldots, k_c$. Тогда

$$\text{знак}(\pi) = \prod_{i=1}^{c} (-1)^{k_i - 1} = (-1)^{\sum_{i=1}^{c}(k_i - 1)} = (-1)^{n-c},$$

поскольку $\sum_{i=1}^{c} k_i = n$. □

Пример 4.9 Для перестановки $\pi = (1\,3\,5)(2\,4)$ имеем $n = 5$, $c = 2$, поэтому знак равен $(-1)^{5-2} = (-1)^3 = -1$. С другой стороны, можно посчитать число инверсий в перестановке $\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 2 & 1 \end{pmatrix}$:

Пары (i,j) с $i < j$:

- (1,2): 3 и 4 $\to$ нет инверсии

- (1,3): 3 и 5 $\to$ нет
- (1,4): 3 и 2 $\to$ да
- (1,5): 3 и 1 $\to$ да
- (2,3): 4 и 5 $\to$ нет
- (2,4): 4 и 2 $\to$ да
- (2,5): 4 и 1 $\to$ да
- (3,4): 5 и 2 $\to$ да
- (3,5): 5 и 1 $\to$ да
- (4,5): 2 и 1 $\to$ да

Итого: 7 инверсий. Знак $= (-1)^7 = -1$. Совпадает с вычислением по циклам.

Циклическая запись также удобна для вычисления степеней перестановки. Чтобы возвести перестановку в степень, достаточно возвести в эту степень каждый цикл.

Пример 4.10 Пусть $\pi = (1\,3\,5)(2\,4)$. Тогда:

$$\begin{aligned}
\pi^2 &= (1\,3\,5)^2(2\,4)^2 = (1\,5\,3)(2)(4) = (1\,5\,3),\\
\pi^3 &= (1\,3\,5)^3(2\,4)^3 = (1)(3)(5)(2\,4) = (2\,4),\\
\pi^4 &= (1\,3\,5)^4(2\,4)^4 = (1\,3\,5)(2)(4) = (1\,3\,5),\\
\pi^5 &= (1\,3\,5)^5(2\,4)^5 = (1\,5\,3)(2\,4),\\
\pi^6 &= (1\,3\,5)^6(2\,4)^6 = \text{тождественная перестановка.}
\end{aligned}$$

Порядок перестановки π равен 6 — наименьшему общему кратному длин циклов: НОК$(3,2) = 6$.

Важное замечание. Геометрическая интерпретация циклов также очень наглядна. Если рассмотреть конфигурацию Тёрстона для циклической перестановки, то можно увидеть, как нити “закручиваются” в циклы. Каждый цикл соответствует замкнутой траектории в динамической системе, заданной перестановкой.

4.2 От вектора к детерминанту: геометрическая мотивация

Обратимся к наивному пониманию вектора. Это направленный отрезок. Важнейшее его свойство – он “плавающий”: мы можем перемещать его в пространстве, не вращая и не отражая. Это очень хорошо согласуется с нашей интуицией. Существенны только длина и направление. Начальная точка не важна. Два направленных отрезка представляют один и тот же вектор, если один можно точно совместить с другим путём параллельного переноса.

Наблюдение 4.3 Пусть всё происходит на прямой, тогда под вектором мы понимаем направленный кусок этой прямой. Давайте постараемся понять что это значит.

(1) Во-первых, под куском мы понимаем отрезок, у которого есть концы AB. И нам удобно считать что этот отрезок имеет конечную и ненулевую длину. Итак, мы разобрались что значит слово "кусок"прямой.

(2) Теперь давайте поймём что значит слово "направленный". Это означает, что мы говорим какую точку мы будем считать началам, а какую – концом. Для этого мы примем обозначение

! $\vec{AB}$ – отрезок на прямой $\mathbb{R}$ конечной длины с концами A,B где точка A – объявлена его началом, а точка B – концом.

Важное замечание. Таким образом, мы можем теперь записать равенство $\vec{AB} = -\vec{BA}$ которая говорит о том, что это один и тот же отрезок, но направления у него разные.

4.2.1 Понятие бивектора

Естественно возникает вопрос: можно ли рассмотреть не кусок отрезка котоофый перемещается на прямой, а кусок плоскости который перемещается уже в плоскости?

Давайте так и сделаем. Будем рассматривать плоский элемент (маленький квадратик, кружочек или что то сложной формы) и не будем делать разницы между ним и тем, который можно получить совмещением путём параллельного переноса (движения без вращения и отражения).

Наблюдение 4.4 Прежде всего мы должны понять что мы можем перенести с одномерного случай на случай двумерный случай, иными словами мы должны спросить себя следующим образом

? Какие характеристики этого плоского элемента остаются неизменными и являются существенными?

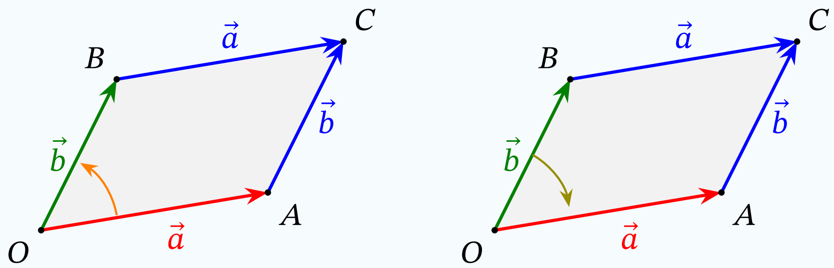
(1) Во-первых, под "куском"плоскости можно наивно понимать произвольную плоскую фигуру, но становится сразу ясно, что это понятие является не очень чётким. Более того, нам хочется прямого обобщения с направленными отрезками, поэтому логично предположить, что под куском плоскости мы будем понимать параллелограмм построенный по двум направленным отрезкам выходящим из одной точки.

(2) В таком случае, мы можем говорить о его площади, что и будет аналогм длины для направленного отрезка, итак "кусок"плоскости это параллелограмм с

конечной, но не нулевой площадью.

В этом подходе замечательно то, что если один из образующих векторов это точка или же совпадает с другим, то мы получаем, что наш параллелелграмм это просто направленный отрезок.

(3) Теперь, мы также можем дать понятие направленного параллелограмма! А именно просто указав порядок следований его образующих векторов.



Важное замечание. Таким образом, мы можем полученный направленный параллелограмм обозначать так $(\vec{\mathbf{a}}, \vec{\mathbf{b}})$. И как в случае с отрезками, мы можем теперь написать равенство

$$(\vec{\mathbf{a}}, \vec{\mathbf{b}}) = -(\vec{\mathbf{b}}, \vec{\mathbf{a}}),$$

показывающее тот факт, что это один и тот же параллелограмм, но начала и конец у которого поменяны местами.

Таким образом, мы готовы теперь дать следующее определение

Определение 4.2.1. Мы будем считать две пары $(\vec{\mathbf{a}},\vec{\mathbf{b}})$, $(\vec{\mathbf{a}}', \vec{\mathbf{b}}')$ векторов *эквивалентными*, если для построенных на них параллелограммах выполняются условия:

(1) Они имеют одинаковую площадь.

(2) Направления вращения у них совпадают.

Класс эквивалентности таких пар мы называем *бивектором* и обозначаем так $\vec{\mathbf{a}} \wedge \vec{\mathbf{b}}$.

Наблюдение 4.5 Из геометрических соображений, нетрудно установить следующие свойства:

$$\begin{aligned} \vec{\mathbf{a}} \wedge (\vec{\mathbf{b}} + \vec{\mathbf{c}}) &= \vec{\mathbf{a}} \wedge \vec{\mathbf{b}} + \vec{\mathbf{a}} \wedge \vec{\mathbf{c}}, \\ \vec{\mathbf{a}} \wedge \vec{\mathbf{a}} &= 0, \\ \vec{\mathbf{a}} \wedge \vec{\mathbf{b}} &= -\vec{\mathbf{b}} \wedge \vec{\mathbf{a}}, \\ (\alpha\vec{\mathbf{a}}) \wedge (\beta\vec{\mathbf{b}}) &= \alpha\beta(\vec{\mathbf{a}} \wedge \vec{\mathbf{b}}), \end{aligned}$$

здесь $\alpha, \beta \in \mathbb{R}$.

Эти свойства, наряду с дистрибутивностью по первому аргументу $(\vec{\mathbf{a}} + \vec{\mathbf{b}}) \wedge \vec{\mathbf{c}} = \vec{\mathbf{a}} \wedge \vec{\mathbf{c}} + \vec{\mathbf{b}} \wedge \vec{\mathbf{c}}$ (которая доказывается аналогично) и ассоциативностью относительно умножения на скаляр, являются аксиомами **внешней** (или **грассмановой**) **алгебры**.

Поскольку при этой операции мы выходим за пределы векторов, она и называется внешним умножением.

Итак, бивектор $\vec{\mathbf{a}} \wedge \vec{\mathbf{b}}$ — это базовый объект в этой алгебре, обобщающий понятие вектора на двумерные ориентированные элементы площади.

4.2.2 Появление детерминанта матрицы 2×2

Если мы вернёмся к случаю прямой, то можно сделать следующее

Наблюдение 4.6 Все направленные отрезки пропорциональны единичному вектору $\vec{\mathbf{e}}_1$, при этом коэффициент пропорциональности и есть длина этого вектора, аеё знак говорит о сононаправленности этого вектора с $\vec{\mathbf{e}}_1$.

Теперь мы хотим понять, как произвольный бивектор $\vec{\mathbf{a}} \wedge \vec{\mathbf{b}}$ связан с бивектором $\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2$ который построен на бизисных векторах $\mathbf{e}_1 = (1{,}0)^\top$, $\mathbf{e}_2 = (0{,}1)^\top$.

Наблюдение 4.7 Пусть векторы $\vec{\mathbf{a}}$ и $\vec{\mathbf{b}}$ выражены через векторы $\vec{\mathbf{e}}_1$ и $\vec{\mathbf{e}}_2$ следующим образом:

$$\begin{aligned} \vec{\mathbf{a}} &= a_1\vec{\mathbf{e}}_1 + a_2\vec{\mathbf{e}}_2, \\ \vec{\mathbf{b}} &= b_1\vec{\mathbf{e}}_1 + b_2\vec{\mathbf{e}}_2. \end{aligned}$$

Тогда, используя свойства операции $\wedge$, имеем:

$$\begin{aligned} \vec{\mathbf{a}} \wedge \vec{\mathbf{b}} &= (a_1\vec{\mathbf{e}}_1 + a_2\vec{\mathbf{e}}_2) \wedge (b_1\vec{\mathbf{e}}_1 + b_2\vec{\mathbf{e}}_2) \\ &= a_1b_1(\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_1) + a_1b_2(\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2) + a_2b_1(\vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_1) + a_2b_2(\vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_2). \end{aligned}$$

Так как $\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_1 = \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_2 = 0$ и $\vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_1 = -\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2$, то мы получаем:

$$\vec{\mathbf{a}} \wedge \vec{\mathbf{b}} = (a_1b_2 - a_2b_1)(\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2).$$

Важное замечание. Выражение в скобках — это в точности детерминант матрицы коэффициентов

$$A = \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix}$$

Таким образом мы можем записать

$$\vec{\mathbf{a}} \wedge \vec{\mathbf{b}} = \det(A) \cdot (\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2)$$

Вот так мы получили детерминант матрицы размера 2×2! То есть это просто коэффициент пропорциональности бивекторов на плоскости! Аналогично тому, как коэффициент пропорциональности векторов на прямой.

4.2.3 Понятие тривектора и детерминант матрицы 3×3

Естественным образом возникает вопрос: можно ли обобщить нашу конструкцию на трёхмерное пространство? Давайте рассмотрим “плавающий” кусочек трёхмерного пространства.

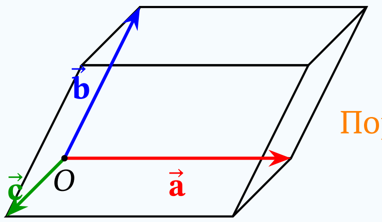
Наблюдение 4.8 Аналогично предыдущим случаям:

(1) На прямой: “кусок” — направленный отрезок (1-мерный объём)

(2) На плоскости: “кусок” — направленный параллелограмм (2-мерный объём = площадь)

(3) В пространстве: “кусок” — направленный параллелепипед (3-мерный объём)

Под направленным параллелепипедом мы понимаем параллелепипед, построенный на трёх векторах $\vec{\mathbf{a}}, \vec{\mathbf{b}}, \vec{\mathbf{c}}$, выходящих из одной точки, с указанием порядка этих векторов.



Определение 4.2.2. Мы будем считать две тройки $(\vec{\mathbf{a}}, \vec{\mathbf{b}}, \vec{\mathbf{c}})$, $(\vec{\mathbf{a}'}, \vec{\mathbf{b}'}, \vec{\mathbf{c}'})$ векторов *эквивалентными*, если для построенных на них параллелепипедах выполняются условия:

(1) Они имеют одинаковый объём.

(2) Порядок векторов у них совпадает (т.е. ориентации параллелепипедов одинаковы).

Класс эквивалентности таких троек мы называем *тривектором* и обозначаем так $\vec{\mathbf{a}} \wedge \vec{\mathbf{b}} \wedge \vec{\mathbf{c}}$.

Наблюдение 4.9 Из геометрических соображений устанавливаем свойства:

$$\begin{aligned}
\vec{\mathbf{a}} \wedge \vec{\mathbf{b}} \wedge \vec{\mathbf{c}} &= -\vec{\mathbf{b}} \wedge \vec{\mathbf{a}} \wedge \vec{\mathbf{c}}, \\
\vec{\mathbf{a}} \wedge \vec{\mathbf{a}} \wedge \vec{\mathbf{c}} &= 0, \\
\vec{\mathbf{a}} \wedge \vec{\mathbf{b}} \wedge (\vec{\mathbf{c}} + \vec{\mathbf{d}}) &= \vec{\mathbf{a}} \wedge \vec{\mathbf{b}} \wedge \vec{\mathbf{c}} + \vec{\mathbf{a}} \wedge \vec{\mathbf{b}} \wedge \vec{\mathbf{d}}, \\
(\alpha\vec{\mathbf{a}}) \wedge (\beta\vec{\mathbf{b}}) \wedge (\gamma\vec{\mathbf{c}}) &= \alpha\beta\gamma(\vec{\mathbf{a}} \wedge \vec{\mathbf{b}} \wedge \vec{\mathbf{c}}).
\end{aligned}$$

Важное замечание. Операция внешнего умножения **ассоциативна**! Это означает, что

$$(\vec{\mathbf{a}} \wedge \vec{\mathbf{b}}) \wedge \vec{\mathbf{c}} = \vec{\mathbf{a}} \wedge (\vec{\mathbf{b}} \wedge \vec{\mathbf{c}}).$$

Это очень важное свойство: порядок, в котором мы “строим” параллелепипед, не важен. Мы можем сначала построить параллелограмм на векторах $\vec{\mathbf{a}}$ и $\vec{\mathbf{b}}$, а затем “натянуть” его на вектор $\vec{\mathbf{c}}$, или наоборот — результат будет один и тот же!

4.2.4 Появление детерминанта матрицы 3×3

Теперь мы хотим понять, как произвольный тривектор $\vec{\mathbf{a}} \wedge \vec{\mathbf{b}} \wedge \vec{\mathbf{c}}$ связан с тривектором $\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3$, построенным на базисных векторах.

Наблюдение 4.10 Пусть векторы выражены через базис:

$$\begin{aligned}
\vec{\mathbf{a}} &= a_1\vec{\mathbf{e}}_1 + a_2\vec{\mathbf{e}}_2 + a_3\vec{\mathbf{e}}_3,\\
\vec{\mathbf{b}} &= b_1\vec{\mathbf{e}}_1 + b_2\vec{\mathbf{e}}_2 + b_3\vec{\mathbf{e}}_3,\\
\vec{\mathbf{c}} &= c_1\vec{\mathbf{e}}_1 + c_2\vec{\mathbf{e}}_2 + c_3\vec{\mathbf{e}}_3.
\end{aligned}$$

Используя свойства внешнего умножения, получаем:

$$\begin{aligned}
\vec{\mathbf{a}} \wedge \vec{\mathbf{b}} \wedge \vec{\mathbf{c}} =&(a_1\vec{\mathbf{e}}_1 + a_2\vec{\mathbf{e}}_2 + a_3\vec{\mathbf{e}}_3)\wedge\\
&(b_1\vec{\mathbf{e}}_1 + b_2\vec{\mathbf{e}}_2 + b_3\vec{\mathbf{e}}_3)\wedge\\
&(c_1\vec{\mathbf{e}}_1 + c_2\vec{\mathbf{e}}_2 + c_3\vec{\mathbf{e}}_3)
\end{aligned}$$

Раскрывая скобки и учитывая, что слагаемые с повторяющимися базисными векторами равны нулю, получим сумму по всем перестановкам индексов:

$$\vec{\mathbf{a}} \wedge \vec{\mathbf{b}} \wedge \vec{\mathbf{c}} = \left(\sum_{\sigma \in S_3} \operatorname{sign}(\sigma) a_{1\sigma(1)} b_{2\sigma(2)} c_{3\sigma(3)}\right)(\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3)$$

Важное замечание. Сумма в скобках — это в точности определение детерминанта матрицы:

$$A = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$$

Таким образом:

$$\vec{\mathbf{a}} \wedge \vec{\mathbf{b}} \wedge \vec{\mathbf{c}} = \det(A) \cdot (\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3)$$

Детерминант матрицы 3×3 — это коэффициент пропорциональности тривекторов в пространстве!

Сейчас мы более подробно опишем процедуру получения формулы для детерминанта матрицы 3×3.

Пример 4.11 Для трёх векторов:

$$\begin{aligned}
\vec{\mathbf{a}}_1 &= a_{11}\vec{\mathbf{e}}_1 + a_{12}\vec{\mathbf{e}}_2 + a_{13}\vec{\mathbf{e}}_3\\
\vec{\mathbf{a}}_2 &= a_{21}\vec{\mathbf{e}}_1 + a_{22}\vec{\mathbf{e}}_2 + a_{23}\vec{\mathbf{e}}_3\\
\vec{\mathbf{a}}_3 &= a_{31}\vec{\mathbf{e}}_1 + a_{32}\vec{\mathbf{e}}_2 + a_{33}\vec{\mathbf{e}}_3
\end{aligned}$$

Вычислим внешнее произведение, раскрывая скобки:

$$\begin{aligned}\vec{\mathbf{a}}_1 \wedge \vec{\mathbf{a}}_2 \wedge \vec{\mathbf{a}}_3 = &(a_{11}\vec{\mathbf{e}}_1 + a_{12}\vec{\mathbf{e}}_2 + a_{13}\vec{\mathbf{e}}_3)\wedge \\ &(a_{21}\vec{\mathbf{e}}_1 + a_{22}\vec{\mathbf{e}}_2 + a_{23}\vec{\mathbf{e}}_3)\wedge \\ &(a_{31}\vec{\mathbf{e}}_1 + a_{32}\vec{\mathbf{e}}_2 + a_{33}\vec{\mathbf{e}}_3)\end{aligned}$$

При раскрытии получим $3^3 = 27$ слагаемых вида:

$$a_{1i}a_{2j}a_{3k}(\vec{\mathbf{e}}_i \wedge \vec{\mathbf{e}}_j \wedge \vec{\mathbf{e}}_k)$$

где $i,j,k \in \{1,2,3\}$.

Важное замечание. Если среди индексов i,j,k есть повторения, то $\vec{\mathbf{e}}_i \wedge \vec{\mathbf{e}}_j \wedge \vec{\mathbf{e}}_k = 0$. Поэтому остаются только слагаемые, где i,j,k — различные индексы, то есть перестановки множества $\{1,2,3\}$.

Выпишем все ненулевые слагаемые:

$$\begin{aligned}&a_{11}a_{22}a_{33}(\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3)\\ &a_{11}a_{23}a_{32}(\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_3 \wedge \vec{\mathbf{e}}_2)\\ &a_{12}a_{21}a_{33}(\vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_3)\\ &a_{12}a_{23}a_{31}(\vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3 \wedge \vec{\mathbf{e}}_1)\\ &a_{13}a_{21}a_{32}(\vec{\mathbf{e}}_3 \wedge \vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2)\\ &a_{13}a_{22}a_{31}(\vec{\mathbf{e}}_3 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_1)\end{aligned}$$

Упростим выражения с базисными векторами:

- $\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3 = +1 \cdot (\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3)$
- $\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_3 \wedge \vec{\mathbf{e}}_2 = -\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3 = -1 \cdot (\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3)$
- $\vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_3 = -\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3 = -1 \cdot (\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3)$
- $\vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3 \wedge \vec{\mathbf{e}}_1 = \vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3 = +1 \cdot (\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3)$
- $\vec{\mathbf{e}}_3 \wedge \vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 = \vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3 = +1 \cdot (\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3)$
- $\vec{\mathbf{e}}_3 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_1 = -\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3 = -1 \cdot (\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3)$

Подставим обратно:

$$\begin{aligned}\vec{\mathbf{a}}_1 \wedge \vec{\mathbf{a}}_2 \wedge \vec{\mathbf{a}}_3 = &+a_{11}a_{22}a_{33}(\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3)\\ &-a_{11}a_{23}a_{32}(\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3)\\ &-a_{12}a_{21}a_{33}(\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3)\\ &+a_{12}a_{23}a_{31}(\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3)\\ &+a_{13}a_{21}a_{32}(\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3)\\ &-a_{13}a_{22}a_{31}(\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3)\end{aligned}$$

Сгруппируем коэффициенты:

$$\vec{\mathbf{a}}_1 \wedge \vec{\mathbf{a}}_2 \wedge \vec{\mathbf{a}}_3 = (a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}) \\ \times (\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3)$$

Связь с перестановками:
Заметим, что каждое слагаемое соответствует определённой перестановке σ множества $\{1,2,3\}$:

- $(1,2,3)$: знак $+$, произведение $a_{11}a_{22}a_{33}$
- $(1,3,2)$: знак $-$, произведение $a_{11}a_{23}a_{32}$
- $(2,1,3)$: знак $-$, произведение $a_{12}a_{21}a_{33}$
- $(2,3,1)$: знак $+$, произведение $a_{12}a_{23}a_{31}$
- $(3,1,2)$: знак $+$, произведение $a_{13}a_{21}a_{32}$
- $(3,2,1)$: знак $-$, произведение $a_{13}a_{22}a_{31}$

Таким образом, мы получаем формулу:

$$\vec{\mathbf{a}}_1 \wedge \vec{\mathbf{a}}_2 \wedge \vec{\mathbf{a}}_3 = \left(\sum_{\sigma \in S_3} \operatorname{sign}(\sigma)\, a_{1\sigma(1)} a_{2\sigma(2)} a_{3\sigma(3)} \right) (\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3)$$

Сумма в скобках — это в точности определение детерминанта матрицы 3×3:

$$\vec{\mathbf{a}}_1 \wedge \vec{\mathbf{a}}_2 \wedge \vec{\mathbf{a}}_3 = \det(A) \cdot (\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3)$$

4.2.5 Обобщение на n-мерное пространство: n-векторы

Теперь мы готовы обобщить нашу конструкцию на произвольную размерность.

Наблюдение 4.11 Рассмотрим n-мерное векторное пространство с базисом $\{\vec{\mathbf{e}}_1, \vec{\mathbf{e}}_2, \dots, \vec{\mathbf{e}}_n\}$. Аналогично предыдущим случаям:

- На прямой: 1-вектор — направленный отрезок (1-мерный объём)
- На плоскости: 2-вектор — направленный параллелограмм (2-мерный объём)
- В пространстве: 3-вектор — направленный параллелепипед (3-мерный объём)
- В n-мерном пространстве: n-вектор — направленный n-мерный параллелепипед (n-мерный объём)

Определение 4.2.3. Рассмотрим n векторов в n-мерном пространстве:

$$\begin{aligned}\vec{\mathbf{a}}_1 &= a_{11}\vec{\mathbf{e}}_1 + a_{12}\vec{\mathbf{e}}_2 + \dots + a_{1n}\vec{\mathbf{e}}_n\\ \vec{\mathbf{a}}_2 &= a_{21}\vec{\mathbf{e}}_1 + a_{22}\vec{\mathbf{e}}_2 + \dots + a_{2n}\vec{\mathbf{e}}_n\\ &\vdots\\ \vec{\mathbf{a}}_n &= a_{n1}\vec{\mathbf{e}}_1 + a_{n2}\vec{\mathbf{e}}_2 + \dots + a_{nn}\vec{\mathbf{e}}_n\end{aligned}$$

Мы будем считать две системы векторов *эквивалентными*, если для построенных на них n-мерных параллелепипедах выполняются условия:

(1) Они имеют одинаковый n-мерный объём.

(2) Порядок векторов у них совпадает.

Класс эквивалентности таких систем мы называем *n-вектором* и обозначаем так

$$\vec{\mathbf{a}}_1 \wedge \vec{\mathbf{a}}_2 \wedge \dots \wedge \vec{\mathbf{a}}_n.$$

Пример 4.12 Вычислим внешнее произведение n векторов:

$$\begin{aligned}\vec{\mathbf{a}}_1 \wedge \vec{\mathbf{a}}_2 \wedge \dots \wedge \vec{\mathbf{a}}_n =&(a_{11}\vec{\mathbf{e}}_1 + \dots + a_{1n}\vec{\mathbf{e}}_n)\wedge\\ &(a_{21}\vec{\mathbf{e}}_1 + \dots + a_{2n}\vec{\mathbf{e}}_n)\wedge\\ &\vdots\\ &(a_{n1}\vec{\mathbf{e}}_1 + \dots + a_{nn}\vec{\mathbf{e}}_n)\end{aligned}$$

При раскрытии получим n^n слагаемых вида:

$$a_{1i_1}a_{2i_2}\cdots a_{ni_n}(\vec{\mathbf{e}}_{i_1} \wedge \vec{\mathbf{e}}_{i_2} \wedge \dots \wedge \vec{\mathbf{e}}_{i_n})$$

где $i_1, i_2, \dots, i_n \in \{1,2,\dots,n\}$.

Важное замечание. Если среди индексов $i_1, i_2, \dots, i_n$ есть повторения, то $\vec{\mathbf{e}}_{i_1} \wedge \vec{\mathbf{e}}_{i_2} \wedge \dots \wedge \vec{\mathbf{e}}_{i_n} = 0$.
Поэтому остаются только слагаемые, где $i_1, i_2, \dots, i_n$ — различные индексы, то есть перестановки множества $\{1,2,\dots,n\}$.
Всего таких перестановок $n!$.

Для каждой перестановки $\sigma \in S_n$ получаем слагаемое:

$$a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)}(\vec{\mathbf{e}}_{\sigma(1)} \wedge \vec{\mathbf{e}}_{\sigma(2)} \wedge \dots \wedge \vec{\mathbf{e}}_{\sigma(n)})$$

Упростим выражения с базисными векторами:
Для любой перестановки σ:

$$\vec{\mathbf{e}}_{\sigma(1)} \wedge \vec{\mathbf{e}}_{\sigma(2)} \wedge \dots \wedge \vec{\mathbf{e}}_{\sigma(n)} = \operatorname{sign}(\sigma) \cdot (\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \dots \wedge \vec{\mathbf{e}}_n)$$

где $\operatorname{sign}(\sigma)$ — знак перестановки σ.
Подставим обратно:

$$\vec{\mathbf{a}}_1 \wedge \vec{\mathbf{a}}_2 \wedge \cdots \wedge \vec{\mathbf{a}}_n = \left(\sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \cdot a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)} \right) \cdot (\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \cdots \wedge \vec{\mathbf{e}}_n)$$

Сумма в скобках — это в точности определение детерминанта матрицы:

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

Важное замечание. Таким образом, мы получаем фундаментальную формулу:

$$\vec{\mathbf{a}}_1 \wedge \vec{\mathbf{a}}_2 \wedge \cdots \wedge \vec{\mathbf{a}}_n = \det(A) \cdot (\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \cdots \wedge \vec{\mathbf{e}}_n)$$

где коэффициент пропорциональности бивектора $\vec{\mathbf{a}}_1 \wedge \vec{\mathbf{a}}_2 \wedge \cdots \wedge \vec{\mathbf{a}}_n$ бивектору $\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \cdots \wedge \vec{\mathbf{e}}_n$ называется *детерминантом матрицы* A,

$$\det(A) := \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \cdot a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}.$$

Наблюдение 4.12 Свойства n-векторов естественным образом обобщают свойства бивекторов и тривекторов:

- **Антикоммутативность**: Перестановка любых двух соседних векторов меняет знак:

$$\vec{\mathbf{a}}_1 \wedge \cdots \wedge \vec{\mathbf{a}}_i \wedge \vec{\mathbf{a}}_{i+1} \wedge \cdots \wedge \vec{\mathbf{a}}_n = -(\vec{\mathbf{a}}_1 \wedge \cdots \wedge \vec{\mathbf{a}}_{i+1} \wedge \vec{\mathbf{a}}_i \wedge \cdots \wedge \vec{\mathbf{a}}_n)$$

- **Нулевой результат при линейной зависимости**: Если векторы $\vec{\mathbf{a}}_1, \dots, \vec{\mathbf{a}}_n$ линейно зависимы, то
$$\vec{\mathbf{a}}_1 \wedge \vec{\mathbf{a}}_2 \wedge \cdots \wedge \vec{\mathbf{a}}_n = 0$$

- **Мультилинейность**: Линейность по каждому аргументу.

- **Ассоциативность**: Порядок расстановки скобок не важен.

Важное замечание. Мы завершили наш путь от геометрической интуиции к общему понятию детерминанта:

Направленные отрезки → Направленные параллелограммы → Направленные параллелепипеды → Направленные n-мерные параллелепипеды

Детерминант матрицы $n \times n$ — это коэффициент, показывающий:

- Во сколько раз ориентированный n-мерный объём, построенный на векторах-столбцах матрицы, больше (или меньше) объёма, построенного на базисных векторах

- Как изменяется ориентированный объём при линейном преобразовании, заданном этой матрицей

- Являются ли векторы-столбцы линейно независимыми (детерминант$\neq 0$) или зависимыми (детерминант $= 0$)

Таким образом, детерминант — это не просто абстрактная формула с перестановками, а естественный геометрический объект, возникающий из понятия ориентированного объёма в n-мерном пространстве!

4.3 Свойства Детерминанта

Наблюдение 4.13 *Для любой матрицы A, имеет место равенство* $\det A^\top = \det A$.

Доказательство. Пусть $A = (a_{ij})$ — квадратная матрица размера $n \times n$, и $A^\top = (b_{ij})$ — её транспонированная матрица, где $b_{ij} = a_{ji}$. По определению:

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \cdot a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}.$$

Для транспонированной матрицы:

$$\det(A^\top) = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \cdot b_{1\sigma(1)} b_{2\sigma(2)} \cdots b_{n\sigma(n)} = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \cdot a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}.$$

Рассмотрим обратную перестановку $\tau = \sigma^{-1}$. Тогда:

$$a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} = a_{1\tau(1)} a_{2\tau(2)} \cdots a_{n\tau(n)},$$

так как для каждого i элемент $a_{\sigma(i)i}$ соответствует $a_{j\tau(j)}$, где $j = \sigma(i)$, и тогда $\tau(j) = i$. Учитывая, что $\operatorname{sign}(\sigma) = \operatorname{sign}(\sigma^{-1}) = \operatorname{sign}(\tau)$ и что отображение $\sigma \mapsto \sigma^{-1}$ является биекцией на S_n.
Тогда, это значит, что каждому слагаемому в сумме для $\det(A^\top)$, соответствующему перестановке σ, однозначно соответствует слагаемое в сумме для $\det(A)$, соответствующее перестановке $\tau = \sigma^{-1}$, и наоборот. При этом:

- множители $a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}$ и $a_{1\tau(1)} a_{2\tau(2)} \cdots a_{n\tau(n)}$ равны,
- знаки $\operatorname{sign}(\sigma)$ и $\operatorname{sign}(\tau)$ равны.

Таким образом,

$$\det(A^\top) = \sum_{\tau \in S_n} \operatorname{sign}(\tau) \cdot a_{1\tau(1)} a_{2\tau(2)} \cdots a_{n\tau(n)} = \det(A),$$

другими слоами все слагаемые в обеих суммах совпадают, и поэтому $\det(A^T) = \det(A)$. □

Наблюдение 4.14 *Детерминант матрицы с нулевой строкой или столбцом равен нулю*

Доказательство. Используем геометрическое определение определителя через внешнее произведение:

$$\vec{\mathbf{a}}_1 \wedge \vec{\mathbf{a}}_2 \wedge \cdots \wedge \vec{\mathbf{a}}_n = \det(A) \cdot (\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \cdots \wedge \vec{\mathbf{e}}_n)$$

Пусть i-я строка матрицы A нулевая. Это означает, что вектор $\vec{\mathbf{a}}_i = 0$. Тогда:

$$\vec{\mathbf{a}}_1 \wedge \cdots \wedge \vec{\mathbf{a}}_i \wedge \cdots \wedge \vec{\mathbf{a}}_n = \vec{\mathbf{a}}_1 \wedge \cdots \wedge 0 \wedge \cdots \wedge \vec{\mathbf{a}}_n = 0$$

С другой стороны, по определению:

$$\vec{\mathbf{a}}_1 \wedge \cdots \wedge \vec{\mathbf{a}}_n = \det(A) \cdot (\vec{\mathbf{e}}_1 \wedge \cdots \wedge \vec{\mathbf{e}}_n)$$

Поскольку базисные векторы $\vec{\mathbf{e}}_1, \ldots, \vec{\mathbf{e}}_n$ линейно независимы, их внешнее произведение $\vec{\mathbf{e}}_1 \wedge \cdots \wedge \vec{\mathbf{e}}_n \neq 0$. Следовательно:

$$\det(A) = 0$$

Наконец, если в матрице A есть нулевой столбец, то в матрице $A^\top$ этот столбец становится нулевой строкой, и тогда в силу равенства $\det A = \det A^\top$, получаем $\det A = 0$. □

Наблюдение 4.15 *Если произвольную строку матрицы умножить на число, то детерминант матрицы умножится на это число.*

Доказательство. Пусть мы умножаем i-ю строку матрицы A на число λ. Это означает, что вектор $\vec{\mathbf{a}}_i$ заменяется на $\lambda\vec{\mathbf{a}}_i$. Тогда новое внешнее произведение:

$$\vec{\mathbf{a}}_1 \wedge \cdots \wedge (\lambda\vec{\mathbf{a}}_i) \wedge \cdots \wedge \vec{\mathbf{a}}_n = \lambda \cdot (\vec{\mathbf{a}}_1 \wedge \cdots \wedge \vec{\mathbf{a}}_i \wedge \cdots \wedge \vec{\mathbf{a}}_n)$$

По определению определителя:

$$\vec{\mathbf{a}}_1 \wedge \cdots \wedge \vec{\mathbf{a}}_n = \det(A) \cdot (\vec{\mathbf{e}}_1 \wedge \cdots \wedge \vec{\mathbf{e}}_n)$$

$$\vec{\mathbf{a}}_1 \wedge \cdots \wedge (\lambda\vec{\mathbf{a}}_i) \wedge \cdots \wedge \vec{\mathbf{a}}_n = \det(A') \cdot (\vec{\mathbf{e}}_1 \wedge \cdots \wedge \vec{\mathbf{e}}_n)$$

где A' — новая матрица. Сравнивая оба выражения, получаем:

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

□

Наблюдение 4.16 *Если строку матрицы умножить на число и прибавить к любой строке, то детерминант матрицы не изменится.*

Доказательство. Пусть мы к i-й строке прибавляем j-ю строку, умноженную на число λ (где $i \neq j$). Это означает, что вектор $\vec{\mathbf{a}}_i$ заменяется на $\vec{\mathbf{a}}_i + \lambda\vec{\mathbf{a}}_j$. Тогда новое внешнее произведение:

$$\begin{aligned}
&\vec{\mathbf{a}}_1 \wedge \cdots \wedge (\vec{\mathbf{a}}_i + \lambda\vec{\mathbf{a}}_j) \wedge \cdots \wedge \vec{\mathbf{a}}_j \wedge \cdots \wedge \vec{\mathbf{a}}_n \\
&= \vec{\mathbf{a}}_1 \wedge \cdots \wedge \vec{\mathbf{a}}_i \wedge \cdots \wedge \vec{\mathbf{a}}_j \wedge \cdots \wedge \vec{\mathbf{a}}_n \\
&\quad + \lambda \cdot (\vec{\mathbf{a}}_1 \wedge \cdots \wedge \vec{\mathbf{a}}_j \wedge \cdots \wedge \vec{\mathbf{a}}_j \wedge \cdots \wedge \vec{\mathbf{a}}_n)
\end{aligned}$$

Во втором слагаемом вектор $\vec{\mathbf{a}}_j$ встречается дважды. По свойству внешнего произведения, если какой-либо вектор повторяется, то всё произведение равно нулю:

$$\vec{\mathbf{a}}_1 \wedge \cdots \wedge \vec{\mathbf{a}}_j \wedge \cdots \wedge \vec{\mathbf{a}}_j \wedge \cdots \wedge \vec{\mathbf{a}}_n = 0$$

Поэтому:

$$\vec{\mathbf{a}}_1 \wedge \cdots \wedge (\vec{\mathbf{a}}_i + \lambda\vec{\mathbf{a}}_j) \wedge \cdots \wedge \vec{\mathbf{a}}_j \wedge \cdots \wedge \vec{\mathbf{a}}_n = \vec{\mathbf{a}}_1 \wedge \cdots \wedge \vec{\mathbf{a}}_i \wedge \cdots \wedge \vec{\mathbf{a}}_j \wedge \cdots \wedge \vec{\mathbf{a}}_n$$

Следовательно, определитель не изменяется:

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

□

Наблюдение 4.17 *Если в матрице есть две одинаковые строки, то её детерминант равен нулю.*

Доказательство. Пусть в матрице A строки с номерами i и j одинаковы, то есть $\vec{\mathbf{a}}_i = \vec{\mathbf{a}}_j$ для $i \neq j$. Рассмотрим внешнее произведение:

$$\vec{\mathbf{a}}_1 \wedge \cdots \wedge \vec{\mathbf{a}}_i \wedge \cdots \wedge \vec{\mathbf{a}}_j \wedge \cdots \wedge \vec{\mathbf{a}}_n$$

Поскольку $\vec{\mathbf{a}}_i = \vec{\mathbf{a}}_j$, в этом произведении два одинаковых вектора. По свойству внешнего произведения, если какой-либо вектор повторяется, то всё произведение равно нулю:

$$\vec{\mathbf{a}}_1 \wedge \cdots \wedge \vec{\mathbf{a}}_i \wedge \cdots \wedge \vec{\mathbf{a}}_j \wedge \cdots \wedge \vec{\mathbf{a}}_n = 0$$

С другой стороны, по определению определителя:

$$\vec{\mathbf{a}}_1 \wedge \cdots \wedge \vec{\mathbf{a}}_n = \det(A) \cdot (\vec{\mathbf{e}}_1 \wedge \cdots \wedge \vec{\mathbf{e}}_n)$$

Поскольку базисные векторы линейно независимы, их внешнее произведение не равно нулю. Следовательно:

$$\det(A) = 0$$

Что и требовалось доказать. □

Наблюдение 4.18 *Если поменять местами две строки матрицы, то определитель меняет знак.*

Доказательство. Пусть мы меняем местами строки с номерами i и j ($i \neq j$). Рассмотрим внешнее произведение исходной матрицы:

$$P = \vec{\mathbf{a}}_1 \wedge \cdots \wedge \vec{\mathbf{a}}_i \wedge \cdots \wedge \vec{\mathbf{a}}_j \wedge \cdots \wedge \vec{\mathbf{a}}_n$$

После перестановки строк получаем:

$$P' = \vec{\mathbf{a}}_1 \wedge \cdots \wedge \vec{\mathbf{a}}_j \wedge \cdots \wedge \vec{\mathbf{a}}_i \wedge \cdots \wedge \vec{\mathbf{a}}_n$$

По свойству антикоммутативности внешнего произведения, перестановка двух соседних векторов меняет знак произведения. Чтобы поменять местами векторы $\vec{\mathbf{a}}_i$ и $\vec{\mathbf{a}}_j$, нам потребуется нечётное число перестановок соседних векторов (ровно $2|j - i| - 1$ перестановок, что нечётно). Поэтому:

$$P' = -P$$

С другой стороны, по определению определителя:

$$P = \det(A) \cdot (\vec{\mathbf{e}}_1 \wedge \cdots \wedge \vec{\mathbf{e}}_n)$$
$$P' = \det(A') \cdot (\vec{\mathbf{e}}_1 \wedge \cdots \wedge \vec{\mathbf{e}}_n)$$

где A' — матрица после перестановки строк. Сравнивая эти выражения, получаем:

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

Что и требовалось доказать. □

Наблюдение 4.19 *Детерминант произведения двух квадратных матриц равен произведению их детерминантов:* $\det(AB) = \det(A)\det(B)$.

Доказательство. Доказательство проведём в несколько этапов, используя свойства элементарных преобразований.

Этап 1: Случай, когда A — элементарная матрица

Рассмотрим три типа элементарных матриц:

1. **Матрица перестановки строк** P: $\det(P) = -1$, и $\det(PA) = -\det(A) = \det(P)\det(A)$
2. **Матрица умножения строки на скаляр** D: $\det(D) = \lambda$, и $\det(DA) = \lambda\det(A) = \det(D)\det(A)$
3. **Матрица прибавления к строке другой строки** E: $\det(E) = 1$, и $\det(EA) = \det(A) = \det(E)\det(A)$

Таким образом, для любой элементарной матрицы U и любой матрицы A выполняется:

$$\det(UA) = \det(U)\det(A)$$

Этап 2: Случай, когда A — обратимая матрица

Если A обратима, то её можно представить как произведение элементарных матриц:

$$A = U_1 U_2 \cdots U_k$$

где U_i — элементарные матрицы.

Тогда:

$$\begin{aligned}
\det(AB) &= \det(U_1 U_2 \cdots U_k B) \\
&= \det(U_1)\det(U_2 \cdots U_k B) \quad \text{(по шагу 1)} \\
&= \det(U_1)\det(U_2)\cdots\det(U_k)\det(B) \quad \text{(по индукции)} \\
&= \det(U_1 U_2 \cdots U_k)\det(B) \quad \text{(по шагу 1)} \\
&= \det(A)\det(B)
\end{aligned}$$

Этап 3: Случай, когда A — необратимая матрица

Если A необратима, то $\det(A) = 0$. Также матрица AB необратима (иначе A была бы обратима), поэтому $\det(AB) = 0$. Следовательно:

$$\det(AB) = 0 = 0 \cdot \det(B) = \det(A)\det(B)$$

Таким образом, во всех случаях выполняется равенство $\det(AB) = \det(A)\det(B)$. □

Наблюдение 4.20 *Если строку матрицы представить в виде суммы двух векторов, то детерминант исходной матрицы равен сумме детерминантов двух матриц,* *в которых эта строка заменена на каждое из слагаемых.*

Доказательство. Пусть матрица A имеет строки $\vec{\mathbf{a}}_1, \vec{\mathbf{a}}_2, \dots, \vec{\mathbf{a}}_n$, и пусть i-я строка представлена в виде суммы: $\vec{\mathbf{a}}_i = \vec{\mathbf{b}} + \vec{\mathbf{c}}$.
Рассмотрим внешнее произведение векторов-строк исходной матрицы:

$$\vec{\mathbf{a}}_1 \wedge \dots \wedge \vec{\mathbf{a}}_{i-1} \wedge (\vec{\mathbf{b}} + \vec{\mathbf{c}}) \wedge \vec{\mathbf{a}}_{i+1} \wedge \dots \wedge \vec{\mathbf{a}}_n$$

По свойству линейности внешнего произведения (мультилинейности), оно линейно по каждому аргументу. Поэтому:

$$\begin{aligned}&\vec{\mathbf{a}}_1 \wedge \dots \wedge \vec{\mathbf{a}}_{i-1} \wedge (\vec{\mathbf{b}} + \vec{\mathbf{c}}) \wedge \vec{\mathbf{a}}_{i+1} \wedge \dots \wedge \vec{\mathbf{a}}_n \\ &= \vec{\mathbf{a}}_1 \wedge \dots \wedge \vec{\mathbf{a}}_{i-1} \wedge \vec{\mathbf{b}} \wedge \vec{\mathbf{a}}_{i+1} \wedge \dots \wedge \vec{\mathbf{a}}_n \\ &\quad + \vec{\mathbf{a}}_1 \wedge \dots \wedge \vec{\mathbf{a}}_{i-1} \wedge \vec{\mathbf{c}} \wedge \vec{\mathbf{a}}_{i+1} \wedge \dots \wedge \vec{\mathbf{a}}_n\end{aligned}$$

С другой стороны, по определению детерминанта через внешнее произведение:

$$\begin{aligned}\vec{\mathbf{a}}_1 \wedge \dots \wedge (\vec{\mathbf{b}} + \vec{\mathbf{c}}) \wedge \dots \wedge \vec{\mathbf{a}}_n &= \det(A) \cdot (\vec{\mathbf{e}}_1 \wedge \dots \wedge \vec{\mathbf{e}}_n) \\ \vec{\mathbf{a}}_1 \wedge \dots \wedge \vec{\mathbf{b}} \wedge \dots \wedge \vec{\mathbf{a}}_n &= \det(A_b) \cdot (\vec{\mathbf{e}}_1 \wedge \dots \wedge \vec{\mathbf{e}}_n) \\ \vec{\mathbf{a}}_1 \wedge \dots \wedge \vec{\mathbf{c}} \wedge \dots \wedge \vec{\mathbf{a}}_n &= \det(A_c) \cdot (\vec{\mathbf{e}}_1 \wedge \dots \wedge \vec{\mathbf{e}}_n)\end{aligned}$$

где:

- A — исходная матрица с i-й строкой $\vec{\mathbf{b}} + \vec{\mathbf{c}}$
- A_b — матрица, где i-я строка заменена на $\vec{\mathbf{b}}$
- A_c — матрица, где i-я строка заменена на $\vec{\mathbf{c}}$

Подставляя эти выражения в предыдущее равенство, получаем:

$$\det(A) \cdot (\vec{\mathbf{e}}_1 \wedge \dots \wedge \vec{\mathbf{e}}_n) = [\det(A_b) + \det(A_c)] \cdot (\vec{\mathbf{e}}_1 \wedge \dots \wedge \vec{\mathbf{e}}_n)$$

Поскольку базисные векторы $\vec{\mathbf{e}}_1, \dots, \vec{\mathbf{e}}_n$ линейно независимы, их внешнее произведение $\vec{\mathbf{e}}_1 \wedge \dots \wedge \vec{\mathbf{e}}_n \neq 0$. Следовательно, мы можем сократить на этот ненулевой множитель:

$$\det(A) = \det(A_b) + \det(A_c)$$

Что и требовалось доказать. □

Наблюдение 4.21 *Если матрица является нижнетреугольной, то её детерминант равен произведению элементов на главной диагонали.*

Доказательство. Пусть A — нижнетреугольная матрица размера $n \times n$. Это означает, что её элементы удовлетворяют условию $a_{ij} = 0$ при $i < j$.

Запишем векторы-строки матрицы A через базисные векторы $\vec{\mathbf{e}}_1, \vec{\mathbf{e}}_2, \dots, \vec{\mathbf{e}}_n$:

$$
\begin{aligned}
\vec{\mathbf{a}}_1 &= a_{11}\vec{\mathbf{e}}_1 \\
\vec{\mathbf{a}}_2 &= a_{21}\vec{\mathbf{e}}_1 + a_{22}\vec{\mathbf{e}}_2 \\
\vec{\mathbf{a}}_3 &= a_{31}\vec{\mathbf{e}}_1 + a_{32}\vec{\mathbf{e}}_2 + a_{33}\vec{\mathbf{e}}_3 \\
&\vdots \\
\vec{\mathbf{a}}_n &= a_{n1}\vec{\mathbf{e}}_1 + a_{n2}\vec{\mathbf{e}}_2 + \cdots + a_{nn}\vec{\mathbf{e}}_n
\end{aligned}
$$

Вычислим внешнее произведение этих векторов:

$$\vec{\mathbf{a}}_1 \wedge \vec{\mathbf{a}}_2 \wedge \cdots \wedge \vec{\mathbf{a}}_n$$

Подставим выражение для $\vec{\mathbf{a}}_1$:

$$= (a_{11}\vec{\mathbf{e}}_1) \wedge \vec{\mathbf{a}}_2 \wedge \cdots \wedge \vec{\mathbf{a}}_n = a_{11}(\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{a}}_2 \wedge \cdots \wedge \vec{\mathbf{a}}_n)$$

Теперь подставим выражение для $\vec{\mathbf{a}}_2$:

$$
\begin{aligned}
&= a_{11}(\vec{\mathbf{e}}_1 \wedge (a_{21}\vec{\mathbf{e}}_1 + a_{22}\vec{\mathbf{e}}_2) \wedge \vec{\mathbf{a}}_3 \wedge \cdots \wedge \vec{\mathbf{a}}_n) \\
&= a_{11}[a_{21}(\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_1) + a_{22}(\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2)] \wedge \vec{\mathbf{a}}_3 \wedge \cdots \wedge \vec{\mathbf{a}}_n
\end{aligned}
$$

Поскольку $\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_1 = 0$, остаётся только:

$$= a_{11}a_{22}(\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2) \wedge \vec{\mathbf{a}}_3 \wedge \cdots \wedge \vec{\mathbf{a}}_n$$

Продолжим этот процесс. На k-м шаге мы получим:

$$a_{11}a_{22}\cdots a_{kk}(\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \cdots \wedge \vec{\mathbf{e}}_k) \wedge \vec{\mathbf{a}}_{k+1} \wedge \cdots \wedge \vec{\mathbf{a}}_n$$

Подставим выражение для $\vec{\mathbf{a}}_{k+1}$:

$$
\begin{aligned}
&= a_{11}a_{22}\cdots a_{kk}(\vec{\mathbf{e}}_1 \wedge \cdots \wedge \vec{\mathbf{e}}_k)\wedge \\
&\quad (a_{k+1,1}\vec{\mathbf{e}}_1 + \cdots + a_{k+1,k}\vec{\mathbf{e}}_k + a_{k+1,k+1}\vec{\mathbf{e}}_{k+1} + \cdots + a_{k+1,n}\vec{\mathbf{e}}_n) \wedge \vec{\mathbf{a}}_{k+2} \wedge \cdots \wedge \vec{\mathbf{a}}_n
\end{aligned}
$$

Заметим, что при внешнем умножении на $\vec{\mathbf{e}}_1 \wedge \cdots \wedge \vec{\mathbf{e}}_k$ все слагаемые, содержащие $\vec{\mathbf{e}}_1, \dots, \vec{\mathbf{e}}_k$, дадут ноль из-за повторяющихся векторов. Остаётся только слагаемое с $\vec{\mathbf{e}}_{k+1}$:

$$= a_{11}a_{22}\cdots a_{kk}a_{k+1,k+1}(\vec{\mathbf{e}}_1 \wedge \cdots \wedge \vec{\mathbf{e}}_k \wedge \vec{\mathbf{e}}_{k+1}) \wedge \vec{\mathbf{a}}_{k+2} \wedge \cdots \wedge \vec{\mathbf{a}}_n$$

Продолжая этот процесс до n-го шага, получим:

$$\vec{\mathbf{a}}_1 \wedge \vec{\mathbf{a}}_2 \wedge \cdots \wedge \vec{\mathbf{a}}_n = (a_{11}a_{22}\cdots a_{nn})(\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \cdots \wedge \vec{\mathbf{e}}_n)$$

С другой стороны, по определению детерминанта:

$$\vec{\mathbf{a}}_1 \wedge \vec{\mathbf{a}}_2 \wedge \cdots \wedge \vec{\mathbf{a}}_n = \det(A) \cdot (\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \cdots \wedge \vec{\mathbf{e}}_n)$$

Сравнивая эти выражения и учитывая, что $\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \cdots \wedge \vec{\mathbf{e}}_n \neq 0$, получаем:

$$\det(A) = a_{11}a_{22}\cdots a_{nn}$$

Что и требовалось доказать. □

4.4 Алегбраическое дополнение

Мы уже знаем, что детерминант матрицы A определяется через внешнее произведение её векторов-строк:

$$\vec{\mathbf{a}}_1 \wedge \vec{\mathbf{a}}_2 \wedge \cdots \wedge \vec{\mathbf{a}}_n = \det(A) \cdot (\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \cdots \wedge \vec{\mathbf{e}}_n)$$

Но как практически вычислить детерминант? Рассмотрим стратегию разложения по первой строке.

Наблюдение 4.22
Представим первый вектор $\vec{\mathbf{a}}_1$ в виде линейной комбинации базисных векторов:

$$\vec{\mathbf{a}}_1 = a_{11}\vec{\mathbf{e}}_1 + a_{12}\vec{\mathbf{e}}_2 + \cdots + a_{1n}\vec{\mathbf{e}}_n$$

Подставим это разложение в внешнее произведение:

$$\begin{aligned}\vec{\mathbf{a}}_1 \wedge \vec{\mathbf{a}}_2 \wedge \cdots \wedge \vec{\mathbf{a}}_n &= (a_{11}\vec{\mathbf{e}}_1 + a_{12}\vec{\mathbf{e}}_2 + \cdots + a_{1n}\vec{\mathbf{e}}_n) \wedge \vec{\mathbf{a}}_2 \wedge \cdots \wedge \vec{\mathbf{a}}_n \\ &= a_{11}(\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{a}}_2 \wedge \cdots \wedge \vec{\mathbf{a}}_n) \\ &\quad + a_{12}(\vec{\mathbf{e}}_2 \wedge \vec{\mathbf{a}}_2 \wedge \cdots \wedge \vec{\mathbf{a}}_n) \\ &\quad + \cdots \\ &\quad + a_{1n}(\vec{\mathbf{e}}_n \wedge \vec{\mathbf{a}}_2 \wedge \cdots \wedge \vec{\mathbf{a}}_n)\end{aligned}$$

Пример 4.13 Рассмотрим случай тривекторов $\vec{\mathbf{a}}_1, \vec{\mathbf{a}}_2, \vec{\mathbf{a}}_3$ и пусть $A = (a_{ij})$ – соответствующая матрица строки которой это координаты этих векторов.

Тогда детерминант этой матрицы — это коэффициент в выражении:

$$\vec{\mathbf{a}}_1 \wedge \vec{\mathbf{a}}_2 \wedge \vec{\mathbf{a}}_3 = \det(A) \cdot (\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3)$$

Разложим вектор $\vec{\mathbf{a}}_1$ по базису:

$$\vec{\mathbf{a}}_1 = a_{11}\vec{\mathbf{e}}_1 + a_{12}\vec{\mathbf{e}}_2 + a_{13}\vec{\mathbf{e}}_3$$

Подставим это разложение в внешнее произведение:

$$\begin{aligned}\vec{\mathbf{a}}_1 \wedge \vec{\mathbf{a}}_2 \wedge \vec{\mathbf{a}}_3 &= (a_{11}\vec{\mathbf{e}}_1 + a_{12}\vec{\mathbf{e}}_2 + a_{13}\vec{\mathbf{e}}_3) \wedge \vec{\mathbf{a}}_2 \wedge \vec{\mathbf{a}}_3 \\ &= a_{11}(\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{a}}_2 \wedge \vec{\mathbf{a}}_3) + a_{12}(\vec{\mathbf{e}}_2 \wedge \vec{\mathbf{a}}_2 \wedge \vec{\mathbf{a}}_3) + a_{13}(\vec{\mathbf{e}}_3 \wedge \vec{\mathbf{a}}_2 \wedge \vec{\mathbf{a}}_3).\end{aligned}$$

Наблюдение 4.23
Теперь давайте более внимательно рассмотрим каждое слагаемое.

(1) Рассмотрим слагаемое $\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{a}}_2 \wedge \vec{\mathbf{a}}_3$. С одной стороны:

$$\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{a}}_2 \wedge \vec{\mathbf{a}}_3 = \det\begin{pmatrix} 1 & 0 & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \cdot (\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3)$$

С другой стороны, разложим векторы $\vec{\mathbf{a}}_2$ и $\vec{\mathbf{a}}_3$:

$$\vec{\mathbf{a}}_2 = a_{21}\vec{\mathbf{e}}_1 + a_{22}\vec{\mathbf{e}}_2 + a_{23}\vec{\mathbf{e}}_3$$
$$\vec{\mathbf{a}}_3 = a_{31}\vec{\mathbf{e}}_1 + a_{32}\vec{\mathbf{e}}_2 + a_{33}\vec{\mathbf{e}}_3$$

Тогда:

$$\begin{aligned}
\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{a}}_2 \wedge \vec{\mathbf{a}}_3 &= \vec{\mathbf{e}}_1 \wedge (a_{21}\vec{\mathbf{e}}_1 + a_{22}\vec{\mathbf{e}}_2 + a_{23}\vec{\mathbf{e}}_3) \wedge (a_{31}\vec{\mathbf{e}}_1 + a_{32}\vec{\mathbf{e}}_2 + a_{33}\vec{\mathbf{e}}_3) \\
&= a_{22}a_{33}(\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3) + a_{23}a_{32}(\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_3 \wedge \vec{\mathbf{e}}_2) \\
&= (a_{22}a_{33} - a_{23}a_{32})(\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3) \\
&= \det\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} \cdot (\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3)
\end{aligned}$$

(2) Рассмотрим слагаемое $\vec{\mathbf{e}}_2 \wedge \vec{\mathbf{a}}_2 \wedge \vec{\mathbf{a}}_3$. С одной стороны:

$$\vec{\mathbf{e}}_2 \wedge \vec{\mathbf{a}}_2 \wedge \vec{\mathbf{a}}_3 = \det\begin{pmatrix} 0 & 1 & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \cdot (\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3)$$

С другой стороны, разложим:

$$\begin{aligned}
\vec{\mathbf{e}}_2 \wedge \vec{\mathbf{a}}_2 \wedge \vec{\mathbf{a}}_3 &= \vec{\mathbf{e}}_2 \wedge (a_{21}\vec{\mathbf{e}}_1 + a_{22}\vec{\mathbf{e}}_2 + a_{23}\vec{\mathbf{e}}_3) \wedge (a_{31}\vec{\mathbf{e}}_1 + a_{32}\vec{\mathbf{e}}_2 + a_{33}\vec{\mathbf{e}}_3) \\
&= a_{21}a_{33}(\vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_3) + a_{23}a_{31}(\vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3 \wedge \vec{\mathbf{e}}_1) \\
&= (-a_{21}a_{33} + a_{23}a_{31})(\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3) \\
&= -\det\begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} \cdot (\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3)
\end{aligned}$$

(3) Рассмотрим слагаемое $\vec{\mathbf{e}}_3 \wedge \vec{\mathbf{a}}_2 \wedge \vec{\mathbf{a}}_3$. С одной стороны:

$$\vec{\mathbf{e}}_3 \wedge \vec{\mathbf{a}}_2 \wedge \vec{\mathbf{a}}_3 = \det\begin{pmatrix} 0 & 0 & 1 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \cdot (\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3)$$

С другой стороны, разложим:

$$\begin{aligned}
\vec{\mathbf{e}}_3 \wedge \vec{\mathbf{a}}_2 \wedge \vec{\mathbf{a}}_3 &= \vec{\mathbf{e}}_3 \wedge (a_{21}\vec{\mathbf{e}}_1 + a_{22}\vec{\mathbf{e}}_2 + a_{23}\vec{\mathbf{e}}_3) \wedge (a_{31}\vec{\mathbf{e}}_1 + a_{32}\vec{\mathbf{e}}_2 + a_{33}\vec{\mathbf{e}}_3) \\
&= a_{21}a_{32}(\vec{\mathbf{e}}_3 \wedge \vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2) + a_{22}a_{31}(\vec{\mathbf{e}}_3 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_1) \\
&= (a_{21}a_{32} - a_{22}a_{31})(\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3) \\
&= \det\begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \cdot (\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3)
\end{aligned}$$

Теперь соберём все слагаемые вместе:

$$\vec{\mathbf{a}}_1 \wedge \vec{\mathbf{a}}_2 \wedge \vec{\mathbf{a}}_3 = \left[a_{11}\det\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{12}\det\begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} + a_{13}\det\begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}\right] \cdot (\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_3)$$

Откуда, получаем формулу

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

Разложение детерминанта по строке в общем случае

Рассмотрим матрицу A размера $n \times n$ с векторами-строками $\vec{\mathbf{a}}_1, \dots, \vec{\mathbf{a}}_n$. По определению:

$$\vec{\mathbf{a}}_1 \wedge \vec{\mathbf{a}}_2 \wedge \dots \wedge \vec{\mathbf{a}}_n = \det(A) \cdot (\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \dots \wedge \vec{\mathbf{e}}_n)$$

Разложим первый вектор по базису:

$$\vec{\mathbf{a}}_1 = a_{11}\vec{\mathbf{e}}_1 + a_{12}\vec{\mathbf{e}}_2 + \dots + a_{1n}\vec{\mathbf{e}}_n$$

Подставим в внешнее произведение:

$$\begin{aligned} \vec{\mathbf{a}}_1 \wedge \vec{\mathbf{a}}_2 \wedge \dots \wedge \vec{\mathbf{a}}_n &= (a_{11}\vec{\mathbf{e}}_1 + a_{12}\vec{\mathbf{e}}_2 + \dots + a_{1n}\vec{\mathbf{e}}_n) \wedge \vec{\mathbf{a}}_2 \wedge \dots \wedge \vec{\mathbf{a}}_n \\ &= a_{11}(\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{a}}_2 \wedge \dots \wedge \vec{\mathbf{a}}_n) \\ &\quad + a_{12}(\vec{\mathbf{e}}_2 \wedge \vec{\mathbf{a}}_2 \wedge \dots \wedge \vec{\mathbf{a}}_n) \\ &\quad + \cdots \\ &\quad + a_{1n}(\vec{\mathbf{e}}_n \wedge \vec{\mathbf{a}}_2 \wedge \dots \wedge \vec{\mathbf{a}}_n) \end{aligned}$$

Теперь рассмотрим каждое слагаемое отдельно. Для фиксированного j рассмотрим:

$$\vec{\mathbf{e}}_j \wedge \vec{\mathbf{a}}_2 \wedge \dots \wedge \vec{\mathbf{a}}_n$$

Разложим векторы $\vec{\mathbf{a}}_2, \dots, \vec{\mathbf{a}}_n$ по базису:

$$\begin{aligned} \vec{\mathbf{a}}_2 &= a_{21}\vec{\mathbf{e}}_1 + a_{22}\vec{\mathbf{e}}_2 + \dots + a_{2n}\vec{\mathbf{e}}_n \\ \vec{\mathbf{a}}_3 &= a_{31}\vec{\mathbf{e}}_1 + a_{32}\vec{\mathbf{e}}_2 + \dots + a_{3n}\vec{\mathbf{e}}_n \\ &\vdots \\ \vec{\mathbf{a}}_n &= a_{n1}\vec{\mathbf{e}}_1 + a_{n2}\vec{\mathbf{e}}_2 + \dots + a_{nn}\vec{\mathbf{e}}_n \end{aligned}$$

Тогда:

$$\vec{\mathbf{e}}_j \wedge \vec{\mathbf{a}}_2 \wedge \dots \wedge \vec{\mathbf{a}}_n = \vec{\mathbf{e}}_j \wedge (a_{21}\vec{\mathbf{e}}_1 + \dots + a_{2n}\vec{\mathbf{e}}_n) \wedge \dots \wedge (a_{n1}\vec{\mathbf{e}}_1 + \dots + a_{nn}\vec{\mathbf{e}}_n)$$

Раскрывая скобки, получим сумму слагаемых вида:

$$a_{2i_2}a_{3i_3}\cdots a_{ni_n} \cdot (\vec{\mathbf{e}}_j \wedge \vec{\mathbf{e}}_{i_2} \wedge \vec{\mathbf{e}}_{i_3} \wedge \dots \wedge \vec{\mathbf{e}}_{i_n})$$

где $i_2, i_3, \dots, i_n \in \{1,2,\dots,n\}$.

Наблюдение 4.24 Заметим, что внешнее произведение $\vec{\mathbf{e}}_j \wedge \vec{\mathbf{e}}_{i_2} \wedge \dots \wedge \vec{\mathbf{e}}_{i_n}$ равно нулю, если среди индексов $j, i_2, \dots, i_n$ есть повторения. Ненулевые слагаемые соответствуют случаю, когда $\{j, i_2, \dots, i_n\} = \{1,2,\dots,n\}$, то есть когда $(i_2, \dots, i_n)$ — перестановка множества $\{1,\dots,n\} \setminus \{j\}$.

Для таких ненулевых слагаемых:

$$\vec{\mathbf{e}}_j \wedge \vec{\mathbf{e}}_{i_2} \wedge \dots \wedge \vec{\mathbf{e}}_{i_n} = \operatorname{sign}(\sigma) \cdot (\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \dots \wedge \vec{\mathbf{e}}_n)$$

где σ — перестановка, переводящая $(1,2,\dots,n)$ в $(j, i_2, \dots, i_n)$

.

Знак $\mathrm{sign}(\sigma)$ можно вычислить, заметив, что чтобы перевести порядок $(j, i_2, \dots, i_n)$ в стандартный порядок $(1,2,\dots,n)$, нужно сделать $j-1$ транспозиций, чтобы переместить j на первое место, а затем привести оставшиеся индексы к стандартному порядку. Общий знак будет $(-1)^{j-1} \cdot \mathrm{sign}(\tau)$, где τ — перестановка множества $\{1,\dots,n\} \setminus \{j\}$.

Таким образом:

$$\vec{\mathbf{e}}_j \wedge \vec{\mathbf{a}}_2 \wedge \cdots \wedge \vec{\mathbf{a}}_n = (-1)^{j-1} \left(\sum_{\tau} \mathrm{sign}(\tau) \cdot a_{2,\tau(2)} a_{3,\tau(3)} \cdots a_{n,\tau(n)} \right) \cdot (\vec{\mathbf{e}}_1 \wedge \cdots \wedge \vec{\mathbf{e}}_n)$$

где сумма берётся по всем перестановкам τ множества $\{1,\dots,n\} \setminus \{j\}$.

Выражение в скобках — это в точности детерминант матрицы A_{1j}, полученной из A вычёркиванием первой строки и j-го столбца:

$$\det(A_{1j}) = \sum_{\tau} \mathrm{sign}(\tau) \cdot a_{2,\tau(2)} a_{3,\tau(3)} \cdots a_{n,\tau(n)}$$

Следовательно:

$$\vec{\mathbf{e}}_j \wedge \vec{\mathbf{a}}_2 \wedge \cdots \wedge \vec{\mathbf{a}}_n = (-1)^{j-1} \det(A_{1j}) \cdot (\vec{\mathbf{e}}_1 \wedge \cdots \wedge \vec{\mathbf{e}}_n)$$

Теперь соберём все слагаемые:

$$\begin{aligned} \vec{\mathbf{a}}_1 \wedge \cdots \wedge \vec{\mathbf{a}}_n &= \left[\sum_{j=1}^{n} a_{1j} \cdot (-1)^{j-1} \det(A_{1j}) \right] \cdot (\vec{\mathbf{e}}_1 \wedge \cdots \wedge \vec{\mathbf{e}}_n) \\ &= \left[\sum_{j=1}^{n} a_{1j} \cdot (-1)^{1+j} \det(A_{1j}) \right] \cdot (\vec{\mathbf{e}}_1 \wedge \cdots \wedge \vec{\mathbf{e}}_n) \end{aligned}$$

Сравнивая с определением детерминанта, получаем:

$$\det(A) = \sum_{j=1}^{n} a_{1j} \cdot (-1)^{1+j} \det(A_{1j})$$

4.4.1 Формулы Крамера

Одно из приложений детерминанта матрицы это явные формулы для решения систем линейных уравнений в котоом число неизвестных совпадает с числом уравнений, другими словами когда матрица коэффициентов является квадратной.

Теорема 4.4.1. *Пусть дана система линейныйх уравнений* $A\cdot\mathbf{x} = \mathbf{b}$*, тогда если* $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$*, то для каждого* $1 \le i \le n$*, при условии* $\det(A) \neq 0$*, имеют место равенства*

$$x_i = \frac{\det A_i(\mathbf{b})}{\det(A)},$$

где $A_i(\mathbf{b})$ *– матрица полученная из матрицы* A *заменой её* i*-ого столбца на столбец* $\mathbf{b}$.

Доказательство. Так как $A\mathbf{x} = \mathbf{b}$, то имеем для единичной матрицы E равенство

$$\begin{aligned} A \cdot E_i(\mathbf{x}) &= A \cdot \begin{pmatrix} \mathbf{e}_1 & \dots & \mathbf{x} & \dots & \mathbf{e}_n \end{pmatrix} = \begin{pmatrix} A\mathbf{e}_1 & \dots & A\mathbf{x} & \dots & A\mathbf{e}_n \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{c}_1(A) & \dots & \mathbf{b} & \dots & \mathbf{c}_n(A) \end{pmatrix} \\ &= A_i(\mathbf{b}) \end{aligned}$$

Так как детерминант от произведения матриц равен произведению детерминантов, получаем

$$\det(A \cdot E_i(\mathbf{x})) = \det(A) \cdot \det)(E_i(\mathbf{x})) = \det A_i(\mathbf{b}),$$

и так как $\det E_i(\mathbf{x}) = x_i$ мы получаем необходимые формулы.

□

Пример 4.14 Решить систему уравнений методом Крамера:

$$\begin{cases} 3x_1 - 2x_2 = 6 \\ -5x_1 + 4x_2 = 8 \end{cases}$$

Решение. Матрица системы:

$$A = \begin{pmatrix} 3 & -2 \\ -5 & 4 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 6 \\ 8 \end{pmatrix}$$

Вычислим детерминант матрицы A:

$$\det A = \begin{vmatrix} 3 & -2 \\ -5 & 4 \end{vmatrix} = 3 \cdot 4 - (-2) \cdot (-5) = 12 - 10 = 2$$

Заменим первый столбец матрицы A на вектор $\mathbf{b}$:

$$A_1(\mathbf{b}) = \begin{pmatrix} 6 & -2 \\ 8 & 4 \end{pmatrix}$$

$$\det A_1(\mathbf{b}) = \begin{vmatrix} 6 & -2 \\ 8 & 4 \end{vmatrix} = 6 \cdot 4 - (-2) \cdot 8 = 24 + 16 = 40$$

Заменим второй столбец матрицы A на вектор $\mathbf{b}$:

$$A_2(\mathbf{b}) = \begin{pmatrix} 3 & 6 \\ -5 & 8 \end{pmatrix}$$

$$\det A_2(\mathbf{b}) = \begin{vmatrix} 3 & 6 \\ -5 & 8 \end{vmatrix} = 3 \cdot 8 - 6 \cdot (-5) = 24 + 30 = 54$$

По формулам Крамера:

$$x_1 = \frac{\det A_1(\mathbf{b})}{\det A} = \frac{40}{2} = 20, \quad x_2 = \frac{\det A_2(\mathbf{b})}{\det A} = \frac{54}{2} = 27$$

Ответ: $x_1 = 20$, $x_2 = 27$ □

Проверим решение подстановкой в исходную систему:

$$3 \cdot 20 - 2 \cdot 27 = 60 - 54 = 6$$
$$-5 \cdot 20 + 4 \cdot 27 = -100 + 108 = 8$$

Условия выполняются.

Часть II

Семинары

4.5 Семинар # 1.6

На этом семинаре мы рассмотрим некоторые примеры умножения матриц, Для нас очень важно сохранить геометрическую сущность этой операции, ведь каждая матрица кодирует линейное отображение где её столбцы это просто образы базисных векторов. Умножение же соответсвует композизции линейных отображений.

4.5.1 Умножение матриц

Напоминание. Пусть у нас есть три конечномерных векторных пространства $\mathbb{R}^n, \mathbb{R}^k, \mathbb{R}^m$ и два линейных отображения $g : \mathbb{R}^n \to \mathbb{R}^k$, $f : \mathbb{R}^k \to \mathbb{R}^m$. Это проще записать следующим образом

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, "g"] \arrow[dr, "f \circ g"'] & \mathbb{R}^k \arrow[d, "f"] \\ & \mathbb{R}^m \end{tikzcd}$$

Возьмём произвольный вектор $\mathbf{x} \in \mathbb{R}^n$, пусть $\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$, тогда, получаем

$$\begin{aligned} g(\mathbf{x}) &= g(x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n) \\ &= x_1 g(\mathbf{e}_1) + \cdots + x_n g(\mathbf{e}_n) \\ &= x_1 \begin{pmatrix} b_{11} \\ \vdots \\ b_{m1} \end{pmatrix} + \cdots + x_n \begin{pmatrix} b_{1n} \\ \vdots \\ b_{mn} \end{pmatrix} \\ &= x_1\mathbf{c}_1(B) + \cdots + x_n\mathbf{c}_n(B). \end{aligned}$$

Далее, имеем

$$\begin{aligned} f(g(\mathbf{x})) &= f(x_1\mathbf{c}_1(B) + \cdots + x_n\mathbf{c}_n(B)) \\ &= x_1 f(\mathbf{c}_1(B)) + \cdots + x_n f(\mathbf{c}_n(B)) \\ &= x_1 A\mathbf{c}_1(B) + \cdots + x_n A\mathbf{c}_n(B) \end{aligned}$$

Таким образом, мы получаем, что композиция отображений $f \circ g$ задаётся матрицей M, которая находится следующим образом

$$M := \begin{pmatrix} A\mathbf{c}_1(B) & \dots & A\mathbf{c}_n(B) \end{pmatrix}$$

Определение 4.5.1. Такую матрицу называют произвдением матриц A и B и обозначают её так AB.

Важное замечание. При умножении матриц, очень важны размеры матриц, так они кодируют линейные отображения, и этому важно как эти отображения действуют.

Умножение описывается правилом “строка на столбец”. Таким образом для матриц A,B, их умножение AB существует если и только если число строк матрицы A совпадает с числом столбцов матрицы B.

Другими словами должно быть такое $A \in \mathrm{Mat}_{m\times p}(\mathbb{R})$, $B \in \mathrm{Mat}_{p\times n}(\mathbb{R})$, тогда $C := AB \in \mathrm{Mat}_{m\times n}(\mathbb{R})$, при этом C определяется следующим образом $C = (c_{i,j})$, где $c_{i,j} := a_{i,1}b_{1,j} + a_{i,2}b_{2,j} + \cdots a_{i,p}b_{p,j}$.

Пример 4.15 Пусть $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ и $B = \begin{pmatrix} 0 & -1 \\ -3 & 10 \\ -7 & 0 \end{pmatrix}$. Находим

$$
\begin{aligned}
A \cdot B &= \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \cdot \begin{pmatrix} 0 & -1 \\ -3 & 10 \\ -7 & 0 \end{pmatrix} \\
&= \begin{pmatrix} \boxed{1 \cdot 0 + 2 \cdot (-3) + 3 \cdot (-7)} & \boxed{1 \cdot (-1) + 2 \cdot 10 + 3 \cdot 0} \\ \boxed{4 \cdot 0 + 5 \cdot (-3) + 6 \cdot (-7)} & \boxed{4 \cdot (-1) + 5 \cdot 10 + 6 \cdot 0} \end{pmatrix} \\
&= \begin{pmatrix} \boxed{0 - 6 - 21} & \boxed{-1 + 20 + 0} \\ \boxed{0 - 15 - 42} & \boxed{-4 + 50 + 0} \end{pmatrix} \\
&= \begin{pmatrix} -27 & 19 \\ -57 & 45 \end{pmatrix}.
\end{aligned}
$$

Эти матрицы кодируют линейные отображения

$$
\begin{tikzcd}
\mathbb{R}^3 \arrow[r, "\text{матрица } B"] \arrow[dr, dashed, "\text{матрица } AB"'] & \mathbb{R}^2 \arrow[d, "\text{матрица } A"] \\
& \mathbb{R}^2
\end{tikzcd}
$$

Пример 4.16 Пусть $B = \begin{pmatrix} 0 & -1 \\ -3 & 10 \\ -7 & 0 \end{pmatrix}$ и $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$. Находим

$$\begin{aligned}
B \cdot A &= \begin{pmatrix} 0 & -1 \\ -3 & 10 \\ -7 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \\
&= \begin{pmatrix} \boxed{0 \cdot 1 + (-1) \cdot 4} & \boxed{0 \cdot 2 + (-1) \cdot 5} & \boxed{0 \cdot 3 + (-1) \cdot 6} \\ \boxed{(-3) \cdot 1 + 10 \cdot 4} & \boxed{(-3) \cdot 2 + 10 \cdot 5} & \boxed{(-3) \cdot 3 + 10 \cdot 6} \\ \boxed{(-7) \cdot 1 + 0 \cdot 4} & \boxed{(-7) \cdot 2 + 0 \cdot 5} & \boxed{(-7) \cdot 3 + 0 \cdot 6} \end{pmatrix} \\
&= \begin{pmatrix} \boxed{0 - 4} & \boxed{0 - 5} & \boxed{0 - 6} \\ \boxed{-3 + 40} & \boxed{-6 + 50} & \boxed{-9 + 60} \\ \boxed{-7 + 0} & \boxed{-14 + 0} & \boxed{-21 + 0} \end{pmatrix} \\
&= \begin{pmatrix} -4 & -5 & -6 \\ 37 & 44 & 51 \\ -7 & -14 & -21 \end{pmatrix}
\end{aligned}$$

Эти матрицы кодируют линейные отображения как на диаграмме

$$\begin{tikzcd}
\mathbb{R}^2 \arrow[r, "\text{матрица } A"] \arrow[dr, dashed, "\text{матрица } BA"'] & \mathbb{R}^3 \arrow[d, "\text{матрица } B"] \\
& \mathbb{R}^3
\end{tikzcd}$$

4.5.2 Степень матрицы

Если же матрицы квадратные и одного размера, то их произведение всегда определено. В частности, это означает что можно матрицу возводить в степень. Пусть $A \in \mathrm{Mat}_{n \times n}(\mathbb{R})$, тогда положим

$$A^1 := A, \qquad A^2 := A \cdot A, \qquad A^3 := A \cdot A \cdot A, \dots, A^n := \underbrace{A \cdots A}_{n},$$

С другой стороны, для нахождения степени, можно воспользоваться следующим приёмом, $A^{2k} = (A^2)^k$, $A^{2k+1} = A \cdot (A^2)^k$.

Пример 4.17 Пусть $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, найдём A^{12}. Имеем, $A^{12} = (A^3)^4$.

Находим
$$A^3 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 37 & 54 \\ 81 & 118 \end{pmatrix}$$
далее,
$$A^3 \cdot A^3 = \begin{pmatrix} 37 & 54 \\ 81 & 118 \end{pmatrix}\begin{pmatrix} 37 & 54 \\ 81 & 118 \end{pmatrix} = \begin{pmatrix} 5743 & 8370 \\ 12555 & 18298 \end{pmatrix}$$
и наконец
$$A^{12} = A^6 \cdot A^6 = \begin{pmatrix} 5743 & 8370 \\ 12555 & 18298 \end{pmatrix}\begin{pmatrix} 5743 & 8370 \\ 12555 & 18298 \end{pmatrix} = \begin{pmatrix} 138067399 & 201223170 \\ 301834755 & 439902154 \end{pmatrix}.$$

Введём в рассмотрение очень важную матрицу $E_n := \mathrm{diag}(\underbrace{1, \dots, 1}_{n}) \in \mathrm{Mat}_{n\times n}(\mathbb{R})$, которая определяется как матрица на диагонали которой стоят единицы, а в остальных местах нули. Например,
$$E_1 = (1), \qquad E_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad E_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \dots$$

Пример 4.18 Пусть $A = \begin{pmatrix} 17 & -6 \\ 35 & -12 \end{pmatrix}$, нужно найти A^5 если при этом известно, что
$$A = \begin{pmatrix} 2 & 3 \\ 5 & 7 \end{pmatrix} \cdot \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} -7 & 3 \\ 5 & -2 \end{pmatrix}$$
Пусть $B = \begin{pmatrix} 2 & 3 \\ 5 & 7 \end{pmatrix}$ и $C = \begin{pmatrix} -7 & 3 \\ 5 & -2 \end{pmatrix}$. Тогда заметим, что $BC = CB = E = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, *т.е.,* $C = B^{-1}$. Таким образом, $A = BDB^{-1}$, и тогда,
$$A^n = \underbrace{(BDB^{-1})(BDB^{-1}) \cdots BDB^{-1}}_{n} = BD^nB^{-1},$$
где $D = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ – диагональная матрица. Нетрудно видеть что $D^5 = \begin{pmatrix} 2^5 & 0 \\ 0 & 3^5 \end{pmatrix} = \begin{pmatrix} 32 & 0 \\ 0 & 243 \end{pmatrix}$. Тогда, получаем
$$A^5 = \begin{pmatrix} 2 & 3 \\ 5 & 7 \end{pmatrix} \cdot \begin{pmatrix} 32 & 0 \\ 0 & 243 \end{pmatrix}\begin{pmatrix} -7 & 3 \\ 5 & -2 \end{pmatrix} = \begin{pmatrix} 64 & 729 \\ 224 & 1215 \end{pmatrix}\begin{pmatrix} -7 & 3 \\ 5 & -2 \end{pmatrix} = \begin{pmatrix} 3197 & -1266 \\ 4504 & -1758 \end{pmatrix}.$$

Определение 4.5.2. Матрица N называется **нильтреугольной**, если:

(1) Она **нильпотентна**: существует $k \in \mathbb{N}$ такое, что $N^k = 0$

(2) Она **верхнетреугольна**: все элементы под главной диагональю равны нулю

(3) Все **диагональные элементы равны нулю**: $n_{ii} = 0$ для всех i

Пример 4.19 Простейшие примеры нильтреугольных матриц в $\mathbb{R}^2$ и $\mathbb{R}^3$:

- В $\mathbb{R}^2$: $N_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$
- В $\mathbb{R}^3$: $N_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$

Проверим нильпотентность:

$$N_2^2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$N_3^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad N_3^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Пример 4.20 Рассмотрим нильтреугольную матрицу $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Отображение N действует следующим образом:

- $N\mathbf{e}_1 = 0$ - первый базисный вектор переходит в ноль
- $N\mathbf{e}_2 = \mathbf{e}_1$ - второй базисный вектор "сдвигается" на первый
- Для произвольного вектора $\mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix}$:

$$N\mathbf{v} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ 0 \end{pmatrix}$$

Геометрически: отображение "проецирует" всю плоскость на ось x, сохраняя только y-координату.

Пример 4.21 Рассмотрим матрицу $N = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$.

Действие отображения N в $\mathbb{R}^3$:

- $N\mathbf{e}_1 = 0$
- $N\mathbf{e}_2 = \mathbf{e}_1$

- $N\mathbf{e}_3 = \mathbf{e}_2$

- Для произвольного вектора $\mathbf{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$:

$$N\mathbf{v} = \begin{pmatrix} y \\ z \\ 0 \end{pmatrix}, \quad N^2\mathbf{v} = \begin{pmatrix} z \\ 0 \\ 0 \end{pmatrix}, \quad N^3\mathbf{v} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Геометрически: отображение осуществляет "сдвиг" координат - каждая координата перемещается на позицию предыдущей, а последняя обнуляется.

4.5.3 Симметрические и кососимметрические матрицы

Рассмотрим ещё одну операцию на множестве матриц, которая называется *транспонирование*, для каждой матрицы $A \in \mathrm{Mat}_{n\times m}(\mathbb{R})$, положим $A^\top \in \mathrm{Mat}_{m\times n}(\mathbb{R})$, строки которой есть столбцы матрицы A.

Пример 4.22 Пусть

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & -1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 & 2 \\ -3 & 2 & 0 \end{pmatrix}, \qquad C = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}.$$

Вычислим $AB^\top + 3C$.
Находим

$$\begin{aligned} B^\top &= = \begin{pmatrix} 1 & -3 \\ 0 & 2 \\ 2 & 0 \end{pmatrix} \\ AB^\top &= \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & -3 \\ 0 & 2 \\ 2 & 0 \end{pmatrix} = \begin{pmatrix} 2 & -4 \\ -2 & 4 \end{pmatrix} \\ AB^\top + 3C &= \begin{pmatrix} 2 & -4 \\ -2 & 4 \end{pmatrix} + \begin{pmatrix} 0 & 3 \\ 3 & 6 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 1 & 10 \end{pmatrix}. \end{aligned}$$

Определение 4.5.3. Квадратная матрица A называется **симметрической**, если она равна своей транспонированной:

$$A = A^\top \quad \text{или} \quad a_{ij} = a_{ji} \quad \text{для всех } i,j$$

Определение 4.5.4. Квадратная матрица A называется **кососимметрической** (или антисимметрической), если она равна минус своей транспонированной:

$$A = -A^\top \quad \text{или} \quad a_{ij} = -a_{ji} \quad \text{для всех } i,j$$

При этом диагональные элементы кососимметрической матрицы равны нулю.

Пример 4.23 Симметрические матрицы в $\mathbb{R}^2$ имеют вид:

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

Конкретные примеры:

$$A_1 = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & -1 \\ -1 & 5 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 4 & 0 \\ 0 & -2 \end{pmatrix}$$

Пример 4.24 Симметрические матрицы в $\mathbb{R}^3$ имеют вид:

$$A = \begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}$$

Конкретные примеры:

$$B_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 0 & -1 & 2 \\ -1 & 3 & 0 \\ 2 & 0 & -4 \end{pmatrix}$$

Пример 4.25 Кососимметрические матрицы в $\mathbb{R}^2$ имеют вид:

$$A = \begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix}$$

Конкретные примеры:

$$C_1 = \begin{pmatrix} 0 & 3 \\ -3 & 0 \end{pmatrix}, \quad C_2 = \begin{pmatrix} 0 & -2 \\ 2 & 0 \end{pmatrix}$$

Пример 4.26 Кососимметрические матрицы в $\mathbb{R}^3$ имеют вид:

$$A = \begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix}$$

Конкретные примеры:

$$D_1 = \begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & 3 \\ -2 & -3 & 0 \end{pmatrix}, \quad D_2 = \begin{pmatrix} 0 & -4 & 5 \\ 4 & 0 & -6 \\ -5 & 6 & 0 \end{pmatrix}$$

4.5.4 Сложение матриц

Множество всех матриц с фиксированным размером, скажем $m \times n$, мы обозначим через $\mathrm{Mat}_{m\times n}(\mathbb{R})$, где $\mathbb{R}$ – множество действительных чисел, тем самым мы подразумеваем что элементы множества $\mathrm{Mat}_{m\times n}(\mathbb{R})$ это матрицы элементы которых действительные числа.

Мы также можем рассматривать матрицы элементами которых могут быть целые или рациональные числа, тогда такие множества будут обозначаться через $\mathrm{Mat}_{m\times n}(\mathbb{Z})$, $\mathrm{Mat}_{m\times n}(\mathbb{Q})$, соответственно.

Наблюдение 4.25 Матрицы можно складывать. При этом сложение матриц определенно только для матриц одинакового размера. Более того множество матриц $\mathrm{Mat}_{m\times n}(\mathbb{R})$ образует векторное пространство, где матрица состоящая из одних только нулей будет соответствовать нулевому вектору.

Обозначение. Любую матрицу, которую можно в общем виде записать так

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

можно более короче записать следующим образом $A = (a_{i,j})$, где нужно сказать что $1 \le i \le m$, $1 \le j \le n$.

Конструкция 4.2 Пусть $A = (a_{i,j}), B = (b_{i,j}) \in \mathrm{Mat}_{m\times n}(\mathbb{R})$, тогда их сумма это матрица $C = (c_{i,j})$, где каждый $c_{i,j} := a_{i,j} + b_{i,j}$. В таких случаях говорят что сложение покомпонентное.

Пример 4.27 Рассмотрим $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, B = \begin{pmatrix} -2 & 1 & -4 \\ 7 & -5 & 108 \end{pmatrix} \in \mathrm{Mat}_{2\times 3}(\mathbb{R})$.

Находим

$$\begin{aligned} A + B &= \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} + \begin{pmatrix} -2 & 1 & -4 \\ 7 & -5 & 108 \end{pmatrix} \\ &= \begin{pmatrix} 1+(-2) & 2+1 & 3+(-4) \\ 4+7 & 5+(-5) & 6+108 \end{pmatrix} \\ &= \begin{pmatrix} -1 & 3 & -1 \\ 11 & 0 & 114 \end{pmatrix}. \end{aligned}$$

Конструкция 4.3 Матрицы можно умножать на числа. Например, если $A = (a_{i,j}) \in \mathrm{Mat}_{m\times n}(\mathbb{R})$, то для любого числа $\lambda \in \mathbb{R}$, мы полагаем $\lambda A := \lambda \cdot A := (\lambda a_{i,j})$. В частности, $-1A$ просто обозначается через $-A$.

Пример 4.28 Рассмотрим матрицу

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix},$$

тогда находим

$$2A := \begin{pmatrix} 2\cdot 1 & 2\cdot 2 & 2\cdot 3 \\ 2\cdot 4 & 2\cdot 5 & 2\cdot 6 \end{pmatrix} = \begin{pmatrix} 2 & 4 & 6 \\ 8 & 10 & 12 \end{pmatrix}$$

и

$$-1A := \begin{pmatrix} -1\cdot 1 & -1\cdot 2 & -1\cdot 3 \\ -1\cdot 4 & 2\cdot 5 & -1\cdot 6 \end{pmatrix} = \begin{pmatrix} -1 & -2 & -3 \\ -4 & -5 & -6 \end{pmatrix}$$

Нетрудно видеть что

$$A + (-1A) = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} + \begin{pmatrix} -1 & -2 & -3 \\ -4 & -5 & -6 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Матрицу состоящую из нулей мы обозначим через O и будем называть её *нулевой матрицей*. Отсюда вытекает определение равенства матриц (конечно же одно размера); мы говорим $A = B$ если $A - B = O$, или что равносильно тому что все $a_{i,j} = b_{i,j}$.

Предложение 4.5.1. *для любых трёх матриц* $A,B,C \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ *имеют место равенства* $A + (B + C) = (A + B) + C$*,* $A + B = B + A$.

Доказательство.
(1) Докажем* что $A + (B + C) = (A + B) + C$ Пусть $A = (a_{i,j})$, $B = (b_{i,j})$, и $C = (c_{i,j})$. Пусть $D = B + C$, $F = A + D$, то есть $A + (B + C) = A + D = F$.

По определению сложения матриц, находим $D = (d_{i,j})$, где $d_{i,j} := b_{i,j} + c_{i,j}$. Далее, $F = (f_{i,j})$, где $f_{i,j} := a_{i,j} + d_{i,j} = a_{i,j} + (b_{i,j} + c_{i,j})$.

С другой стороны, пусть $G := A + B$, $H := G + C$, то есть $H = (A + B) + C$. Нам нужно показать что $H = F$. Получаем, $G = (g_{i,j})$, где $g_{i,j} := a_{i,j} + b_{i,j}$. Далее, $H = (h_{i,j})$, где $h_{i,j} = g_{i,j} + c_{i,j} = (a_{i,j} + b_{i,j}) + c_{i,j}$.

Таким образом, **матричное** равенство $A + (B + C) = (A + B) + C$ выполнено если и только если в множестве действительных чисел выполнено равенство $a_{i,j} + (b_{i,j} + c_{i,j}) = (a_{i,j} + b_{i,j}) + c_{i,j}$ **для чисел**, но ассоциативность в $\mathbb{R}$ постулируется что и доказывает требуемое.

(2) Докажем $A + B = B + A$. Пусть $C := A + B$, $D := B + A$. Тогда, получаем $C = (c_{i,j})$, где $c_{i,j} := a_{i,j} + b_{i,j}$. Далее, $D = (d_{i,j})$, где $d_{i,j} = b_{i,j} + a_{i,j}$. Тогда равенство $C = D$ равносильно равенствам $a_{i,j} + b_{i,j} = b_{i,j} + a_{i,j}$ для чисел. Но для действительных чисел коммутативность постулируется, что и доказывает требуемое. □

*это свойство называется *ассоциативность*.

Теорема 4.5.2 (Разложение на симметрическую и кососимметрическую части)**.** *Любую квадратную матрицу A можно единственным образом представить в виде суммы симметрической и кососимметрической матриц:*

$$A = \frac{A + A^T}{2} + \frac{A - A^T}{2}$$

где $\frac{A+A^T}{2}$ *- симметрическая часть, а* $\frac{A-A^T}{2}$ *- кососимметрическая часть.*

Доказательство. Проверим непосредственно:

$$\left(\frac{A + A^T}{2}\right)^T = \frac{A^T + A}{2} = \frac{A + A^T}{2}$$

$$\left(\frac{A - A^T}{2}\right)^T = \frac{A^T - A}{2} = -\frac{A - A^T}{2}$$

□

Пример 4.29 Разложим матрицу $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$:

$$A_{\text{sym}} = \frac{A + A^T}{2} = \frac{1}{2}\begin{pmatrix} 2 & 6 & 10 \\ 6 & 10 & 14 \\ 10 & 14 & 18 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 5 \\ 3 & 5 & 7 \\ 5 & 7 & 9 \end{pmatrix}$$

$$A_{\text{skew}} = \frac{A - A^T}{2} = \frac{1}{2}\begin{pmatrix} 0 & -2 & -4 \\ 2 & 0 & -2 \\ 4 & 2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 1 & 0 \end{pmatrix}$$

Проверка: $A_{\text{sym}} + A_{\text{skew}} = A$.

Теорема 4.5.3. *Пространство* $\mathrm{Mat}_{m\times n}(\mathbb{R})$ *имеет размерность* $m \cdot n$.

Доказательство. Построим стандартный базис этого пространства. Для каждого $i = 1, \ldots, m$ и $j = 1, \ldots, n$ определим матрицу E_{ij}, у которой на позиции (i,j) стоит 1, а все остальные элементы равны 0.

Всего таких матриц $m \cdot n$, и они линейно независимы. Любая матрица $A = (a_{ij})$ единственным образом представляется в виде:

$$A = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} E_{ij}$$

Следовательно, $\{E_{ij}\}$ образует базис пространства $\mathrm{Mat}_{m\times n}(\mathbb{R})$, и его размерность равна $m \cdot n$. □

Пример 4.30 Пространство $\mathrm{Mat}_{2\times 2}(\mathbb{R})$ имеет размерность 4. Стандартный базис:

$$E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Любая матрица $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ представляется как:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = aE_{11} + bE_{12} + cE_{21} + dE_{22}$$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = a\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + c\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + d\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Пример 4.31 Пространство $\mathrm{Mat}_{2\times 3}(\mathbb{R})$ имеет размерность 6. Стандартный базис состоит из 6 матриц:

$$E_{11} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad E_{12} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad E_{13} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$
$$E_{21} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad E_{22} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad E_{23} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Любая матрица $\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}$ представляется как:

$$aE_{11} + bE_{12} + cE_{13} + dE_{21} + eE_{22} + fE_{23}$$

Пространство матриц $\mathrm{Mat}_{m\times n}(\mathbb{R})$ можно представлять себе как $m \cdot n$-мерное векторное пространство, где каждая матрица - это вектор, а базисные матрицы E_{ij} - это координатные оси.

Теорема 4.5.4. *Пространство* $\mathrm{Mat}_{m\times n}(\mathbb{R})$ *изоморфно пространству* $\mathbb{R}^{m\cdot n}$.

Доказательство. Построим изоморфизм $\varphi : \mathrm{Mat}_{m\times n}(\mathbb{R}) \to \mathbb{R}^{m\cdot n}$ следующим образом:

Для матрицы $A = (a_{ij})$ определим:

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

Это отображение:

- Линейно: $\varphi(\alpha A + \beta B) = \alpha\varphi(A) + \beta\varphi(B)$
- Биективно: у каждого вектора в $\mathbb{R}^{m\cdot n}$ есть единственный прообраз в $\mathrm{Mat}_{m\times n}(\mathbb{R})$
- Сохраняет операции: $\varphi(A + B) = \varphi(A) + \varphi(B)$ и $\varphi(\alpha A) = \alpha\varphi(A)$

Следовательно, φ - изоморфизм векторных пространств. □

4.5.5 Домашнее Задание # 1.6

Упражнение 4.1 Рассмотрим множество всех симметричных матриц $\mathrm{Mat}_n^{(+)}(\mathbb{R})$ и множество всех кососимметричных матриц $\mathrm{Mat}_n^{(-)}(\mathbb{R})$.

1. Покажите что они являются векторными подпространствами в пространтсве $\mathrm{Mat}_n(\mathbb{R})$.
2. Опишите базисы этих подпространств.
3. Найдите размерности этих подпространств.
4. Существует ли матрица которая одновременно и симметричная и кососимметричная? Другими словами опишите множество $\mathrm{Mat}_n^{(+)}(\mathbb{R}) \cap \mathrm{Mat}_n^{(-)}(\mathbb{R})$.

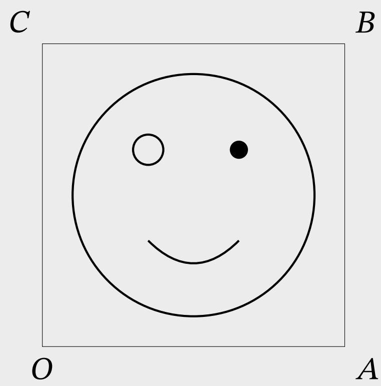
Упражнение 4.2 Используя метод математической индукции докажите что произведение нильтреуольних матриц произозольного размера $n \times n$ тоже будет нильтреуольной[a].

[a] Квадратная матрица называется *нильтреугольной* если её главная диогональ нулевая и всё что ниже этой диагонали тоже нулевое. Нулевая матрица считается нильтреугольной.

Упражнение 4.3 Найдите все (2×2)-матрицы B, коммутирующие с матрицей $A = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}$, то есть удовлетворяющие условию $AB = BA$.

Упражнение 4.4 Рассмотрите матрицу $A = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$ и найдите для любого[a] натурального $n \geq 1$ её степень A^n.

Давайте опять вернёмся к нашему смайлику из предыдущего ДЗ. На плоскости $\mathbb{R}^2$ с фиксированной декартовой системой координат, рассмотрим квадрат $OABC$, где $O = (0,0)$, $A = (1,0)$, $B = (1,1)$, $C = (0,1)$.



Исследуйте и изобразите образ квадрата со смайликом при линейном отображении с этой матрицей A^n (достаточно рассмотреть A, A^2, A^3 и сказать что будет при росте

n), рассмотрите случаи когда $0 < \lambda < 1, \lambda \geq 1$.

[a]Вам опять понадобиться метод математической индукции

Упражнение 4.5 Зная теперь что матрица $A_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ кодирует поворот плоскости вокруг начало координат на угол θ, а также используя Ваши знание о том что композиция линейных отображений соответствует умножению матриц, получите формулы для $\sin(n\theta)$, $\cos(n\theta)$ для любого[a] натурального $n \geq 1$.

[a]Вам опять понадобиться метод математической индукции

Упражнение 4.6 Вычислите

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

воспользововашись равенством

$$\begin{pmatrix} 4 & 3 & -3 \\ 2 & 3 & -2 \\ 4 & 4 & -3 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 1 \\ 2 & 2 & 1 \\ 3 & 4 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 2 & -1 \\ 1 & 1 & -1 \\ -2 & -5 & 4 \end{pmatrix}.$$

4.6 Семинар # 1.7

Этот семинар посвящён изучению обратных матриц, а также некоторым смежным вопросам.

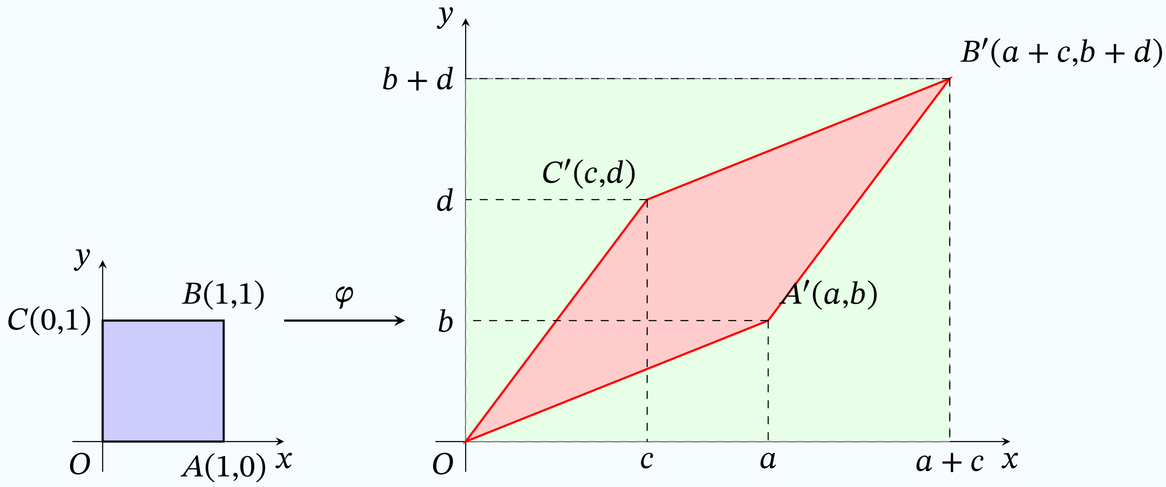
Вернёмся ещё раз к примеру рассмотренному ранее и рассмотрим его более подробно.

Наблюдение 4.26
Пусть линейный оператор $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ задаётся матрицей A, где

$$A = \begin{pmatrix} a & c \\ b & d \end{pmatrix},$$

Где нужно помнить что первый столбец это образ вектора $\mathbf{e}_1 = (1,0)^\top$, а второй столбец это образ вектора $\mathbf{e}_2 = (0,1)^\top$.

Тогда мы получаем следующую схему этого отображения, где O', A', B', C' – образы соответствующих вершин при отображении φ.

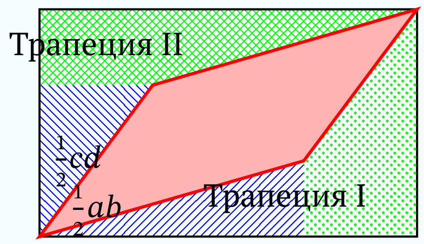


Пусть квадрат $ABCD$ имел площадь равную 1, тогда вычислим площадь полученного параллелограмма, это нам покажет величину увелечиние нашего изначального квадрата.

> **Важное замечание.** Наша цель заключается в том, что когда мы можем наше отображение обратить. Используя геометрические рассуждения это можно будет сделать если наш квадрат не отображается в отрезок *т.е.,* когда точки A', C' не лежат на одной прямой.

Вычислим площадь параллелограмма, вычитая из площади окружающего прямоугольника площади двух треугольников и двух трапеций:

$$
\begin{aligned}
\text{Площадь прямоугольника} &= (a+c)(b+d) \\
\text{Площадь треугольника I} &= \frac{1}{2}ab \\
\text{Площадь треугольника II} &= \frac{1}{2}cd \\
\text{Площадь трапеции I} &= \frac{1}{2}(b+(b+d))\cdot c = \frac{1}{2}(2b+d)c = bc + \frac{1}{2}cd \\
\text{Площадь трапеции II} &= \frac{1}{2}(d+(b+d))\cdot a = \frac{1}{2}(b+2d)a = \frac{1}{2}ab + ad
\end{aligned}
$$



Теперь вычтем все лишние площади:

$$
\begin{aligned}
S_{\text{паралл}} &= (a+c)(b+d) - \left[\tfrac{1}{2}ab + \tfrac{1}{2}cd + (bc + \tfrac{1}{2}cd) + (\tfrac{1}{2}ab + ad)\right] \\
&= (a+c)(b+d) - [ab + cd + bc + ad] \\
&= ab + ad + bc + cd - ab - cd - bc - ad \\
&= ad - bc
\end{aligned}
$$

Определение 4.6.1. Величина $ad - bc$ называется *детерминантом матрицы* $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$ и обозначается следующим образом $\det(A)$.

Как видно, эта величина может быть как положжительной так и отрицательной, а также быть равной нулю.

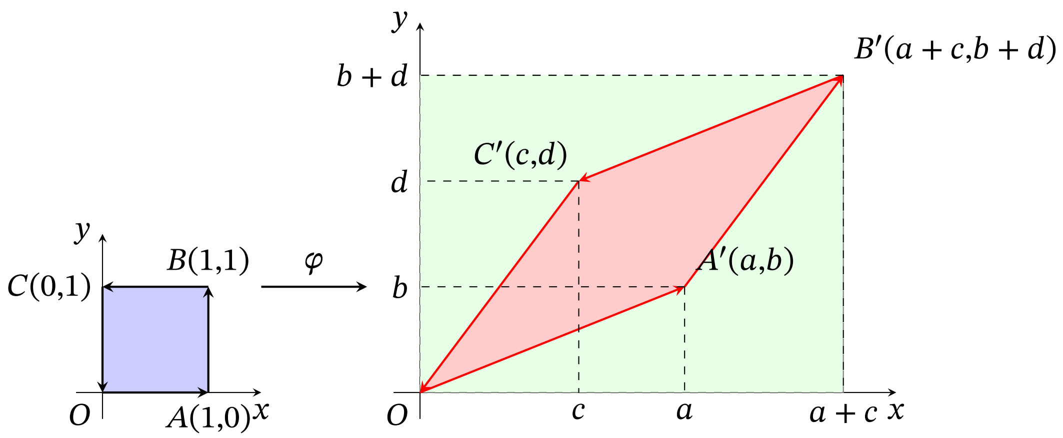
4.6.1 Геометрический смысл детерминанта матрицы 2×2

Пусть линейный оператор $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ задаётся матрицей A, где

$$A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}.$$

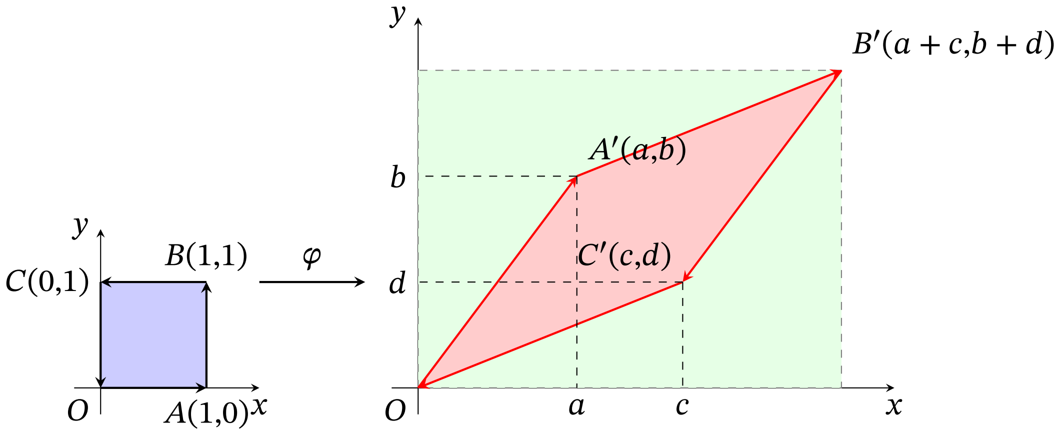
Первый столбец — образ вектора $\vec{e}_1 = (1,0)^\top$, второй столбец — образ вектора $\vec{e}_2 = (0,1)^\top$.

Случай 1: $\det A > 0$ (ориентация сохраняется)



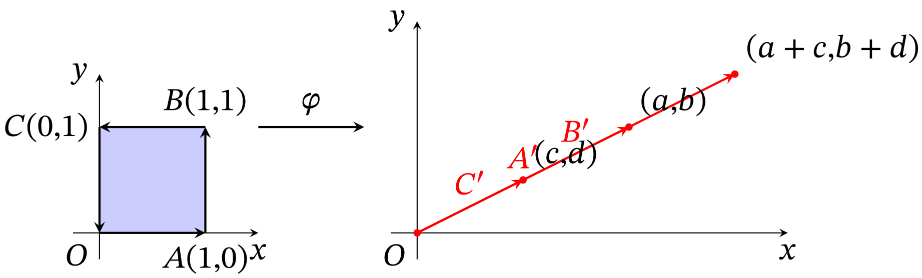
$\det A = ad - bc > 0$, ориентация сохраняется

Случай 2: $\det A < 0$ **(ориентация меняется)**



$\det A = ad - bc < 0$, ориентация меняется

Случай 3: $\det A = 0$ **(вырожденный случай)**



$\det A = ad - bc = 0$, образ — отрезок

Вычисление площади

Площадь параллелограмма равна модулю определителя:

$$S = |\det A| = |ad - bc|$$

В вырожденном случае ($\det A = 0$) площадь равна 0.

4.6.2 Нахождение обратной матрицы

Напоминание. Напомним практический способ нахождения обратной матрицы: она является решением матричного уравнения $AX = E$ (если решение существует). Для решения этого уравнения нужно взять матрицу $(A\,|E)$ и элементарными преобразованиями строк привести её к виду $(E\,|B)$, тогда $B = A^{-1}$.

Пример 4.32 Рассмотрим матрицу $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ и найдём A^{-1}, в предположении, что $a \neq 0$, $c \neq 0$, $ad - bc \neq 0$.
Приведём расширенную матрицу $(A|E)$ к виду $(E|A^{-1})$ элементарными преобразо-

ваниями строк.

$$\left(\begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array}\right) \xrightarrow[\substack{R_1:=\frac{1}{a}\cdot R_1 \\ R_2:=\frac{1}{c}\cdot R_2}]{} \left(\begin{array}{cc|cc} 1 & \frac{b}{a} & \frac{1}{a} & 0 \\ 1 & \frac{d}{c} & 0 & \frac{1}{c} \end{array}\right) \xrightarrow{R_2:=R_2-R_1} \left(\begin{array}{cc|cc} 1 & \frac{b}{a} & \frac{1}{a} & 0 \\ 0 & \frac{ad-bc}{ac} & -\frac{1}{a} & \frac{1}{c} \end{array}\right)$$

$$\xrightarrow{R_2:=\frac{ac}{ad-bc}\cdot R_2} \left(\begin{array}{cc|cc} 1 & \frac{b}{a} & \frac{1}{a} & 0 \\ 0 & 1 & -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{array}\right) \xrightarrow{R_1:=R_1-\frac{b}{a}\cdot R_2} \left(\begin{array}{cc|cc} 1 & 0 & \frac{d}{ad-bc} & -\frac{b}{ad-bc} \\ 0 & 1 & -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{array}\right)$$

Таким образом,

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

Пример 4.33 Пусть $A = \begin{pmatrix} 1 & 4 & -2 \\ 2 & -2 & 1 \\ 5 & -1 & -1 \end{pmatrix}$. Найдём обратную к ней.

Решение.

$$\begin{pmatrix} 1 & 4 & -2 & | & 1 & 0 & 0 \\ 2 & -2 & 1 & | & 0 & 1 & 0 \\ 5 & -1 & -1 & | & 0 & 0 & 1 \end{pmatrix} \xrightarrow{\substack{R_2:=R_2-2\cdot R_1 \\ R_3:=R_3-5\cdot R_1}} \begin{pmatrix} 1 & 4 & -2 & | & 1 & 0 & 0 \\ 0 & -10 & 5 & | & -2 & 1 & 0 \\ 0 & -21 & 9 & | & -5 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow{R_2:=\frac{-1}{10}\cdot R_2} \begin{pmatrix} 1 & 4 & -2 & | & 1 & 0 & 0 \\ 0 & 1 & \frac{-1}{2} & | & \frac{1}{5} & \frac{-1}{10} & 0 \\ 0 & -21 & 9 & | & -5 & 0 & 1 \end{pmatrix} \xrightarrow{R_3:=R_3-(-21)\cdot R_2} \begin{pmatrix} 1 & 4 & -2 & | & 1 & 0 & 0 \\ 0 & 1 & \frac{-1}{2} & | & \frac{1}{5} & \frac{-1}{10} & 0 \\ 0 & 0 & \frac{-3}{2} & | & \frac{-4}{5} & \frac{-21}{10} & 1 \end{pmatrix}$$

$$\xrightarrow{R_3:=\frac{-2}{3}\cdot R_3} \begin{pmatrix} 1 & 4 & -2 & | & 1 & 0 & 0 \\ 0 & 1 & \frac{-1}{2} & | & \frac{1}{5} & \frac{-1}{10} & 0 \\ 0 & 0 & 1 & | & \frac{8}{15} & \frac{7}{5} & \frac{-2}{3} \end{pmatrix} \xrightarrow{R_2:=R_2-\frac{-1}{2}\cdot R_3} \begin{pmatrix} 1 & 4 & -2 & | & 1 & 0 & 0 \\ 0 & 1 & 0 & | & \frac{7}{15} & \frac{3}{5} & \frac{-1}{3} \\ 0 & 0 & 1 & | & \frac{8}{15} & \frac{7}{5} & \frac{-2}{3} \end{pmatrix}$$

$$\xrightarrow{R_1:=R_1-(-2)\cdot R-3} \begin{pmatrix} 1 & 4 & 0 & | & \frac{31}{15} & \frac{14}{5} & \frac{-4}{3} \\ 0 & 1 & 0 & | & \frac{7}{15} & \frac{3}{5} & \frac{-1}{3} \\ 0 & 0 & 1 & | & \frac{8}{15} & \frac{7}{5} & \frac{-2}{3} \end{pmatrix} \xrightarrow{R_1:=R_1-4\cdot R_2} \begin{pmatrix} 1 & 0 & 0 & | & \frac{1}{15} & \frac{2}{5} & 0 \\ 0 & 1 & 0 & | & \frac{7}{15} & \frac{3}{5} & \frac{-1}{3} \\ 0 & 0 & 1 & | & \frac{8}{15} & \frac{7}{5} & \frac{-2}{3} \end{pmatrix}$$

□

Таким образом

$$A^{-1} = \begin{pmatrix} \frac{1}{15} & \frac{2}{5} & 0 \\ \frac{7}{15} & \frac{3}{5} & \frac{-1}{3} \\ \frac{8}{15} & \frac{7}{5} & \frac{-2}{3} \end{pmatrix} = \frac{1}{15}\begin{pmatrix} 1 & 6 & 0 \\ 7 & 9 & -5 \\ 8 & 21 & -10 \end{pmatrix}.$$

Пример 4.34 Решим уравнение

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} X = \begin{pmatrix} 3 & 5 \\ 5 & 9 \end{pmatrix}.$$

Решение. Если в уравнении $AX = B$ матрица A имеет обратную, то $X = A^{-1}B$, так как равенство $AX = B$ можно умножить слева на A^{-1}.

Найдём обратную матрицу A^{-1} для $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$.

Детерминат этой матриц равен $\det(A) = 1 \cdot 4 - 2 \cdot 3 = -2$. Тогда

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

Теперь вычислим $X = A^{-1}B$:

$$X = \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 3 & 5 \\ 5 & 9 \end{pmatrix} = \begin{pmatrix} (-2) \cdot 3 + 1 \cdot 5 & (-2) \cdot 5 + 1 \cdot 9 \\ \frac{3}{2} \cdot 3 + \left(-\frac{1}{2}\right) \cdot 5 & \frac{3}{2} \cdot 5 + \left(-\frac{1}{2}\right) \cdot 9 \end{pmatrix} = \begin{pmatrix} -1 & -1 \\ 2 & 3 \end{pmatrix}.$$

Проверка:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} -1 & -1 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 \cdot (-1) + 2 \cdot 2 & 1 \cdot (-1) + 2 \cdot 3 \\ 3 \cdot (-1) + 4 \cdot 2 & 3 \cdot (-1) + 4 \cdot 3 \end{pmatrix} = \begin{pmatrix} 3 & 5 \\ 5 & 9 \end{pmatrix}.$$

Полученная матрица совпадает с правой частью исходного уравнения. □

Наблюдение 4.27 Рассмотрим решение большого количества систем $Ax = \mathbf{b}_1$, $Ax = \mathbf{b}_2, \ldots, Ax = \mathbf{b}_s$ с одной и той же матрицей коэффициентов.

Нетрудно видеть, что их не нужно решать по отдельности, часть вычислений можно провести одновременно, выполняя элементарные преобразования строк в матрице $(A \mid B)$, где B составлена из столбцов $\mathbf{b}_1, \ldots, \mathbf{b}_s$.

Тогда после приведения матрицы A к улучшенному ступенчатому виду уже можно выполнять дальнейший анализ (сводящийся к выписыванию ответа) каждой из систем по отдельности.

Пример 4.35 Решим системы уравнений:

$$\begin{cases} x_1 + 2x_2 = 3 \\ 3x_1 + 4x_2 = 5 \end{cases} \quad \text{и} \quad \begin{cases} x_1 + 2x_2 = 5 \\ 3x_1 + 4x_2 = 9 \end{cases}$$

Решение. Запишем объединённую расширенную матрицу, содержащую матрицу коэффициентов и обе правые части:

$$(A \mid B) = \left(\begin{array}{cc|cc} 1 & 2 & 3 & 5 \\ 3 & 4 & 5 & 9 \end{array}\right)$$

Выполним элементарные преобразования строк:

$$\left(\begin{array}{cc|cc} 1 & 2 & 3 & 5 \\ 3 & 4 & 5 & 9 \end{array}\right) \xrightarrow{R_2:=R_2-3R_1} \left(\begin{array}{cc|cc} 1 & 2 & 3 & 5 \\ 0 & -2 & -4 & -6 \end{array}\right) \xrightarrow{R_2:=-\frac{1}{2}R_2} \left(\begin{array}{cc|cc} 1 & 2 & 3 & 5 \\ 0 & 1 & 2 & 3 \end{array}\right)$$

$$\xrightarrow{R_1:=R_1-2R_2} \left(\begin{array}{cc|cc} 1 & 0 & -1 & -1 \\ 0 & 1 & 2 & 3 \end{array}\right)$$

Получили улучшенный ступенчатый вид. Теперь выпишем решения для каждой системы отдельно:

- **Первая система** (столбец 1): $x_1 = -1$, $x_2 = 2$
- **Вторая система** (столбец 2): $x_1 = -1$, $x_2 = 3$

Проверка:

- Для первой системы: $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$ — верно
- Для второй системы: $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} -1 \\ 3 \end{pmatrix} = \begin{pmatrix} 5 \\ 9 \end{pmatrix}$ — верно

□

Наблюдение 4.28 Рассмотрим теперь очень похожую, но другую ситуацию: дана матрица коэффициентов A, и нам нужно решать большое количество систем вида $Ax = b$, где столбцы b поступают к нам в режиме реального времени (и их полный список неизвестен).

В этом случае удобно заранее «запомнить», каким образом матрица A приводится к улучшенному ступенчатому виду.

Для этого вспомним, что всякое элементарное преобразование строк матрицы реализуется при помощи умножения слева на подходящую «элементарную» матрицу.

Тогда, если A приводится к улучшенному ступенчатому виду A' при помощи элементарных преобразований, соответствующих умножению слева на матрицы $U_1, \dots, U_k$, то $A' = U_k \dots U_1 A$.

Если обозначить $U = U_k \dots U_1$, то любая система $Ax = b$ после умножения слева на

U переходит в эквивалентную систему $A'x = Ub$. Таким образом, при поступлении очередного столбца b для решения системы $Ax = b$ достаточно вычислить Ub, и уже можно выписывать ответ.

Для оптимального вычисления обеих матриц A' и U нужно записать матрицу $(A \mid E)$ и, выполняя в ней элементарные преобразования строк, привести A к улучшенному ступенчатому виду.

Важное замечание. Тут E это единичная матрица подходящего размера число строк матрицы $A \times$ число строк матрицы A.

Тогда справа от черты будет стоять U, то есть результирующая матрица как раз и будет $(A' \mid U)$. (Обратите внимание, что должно выполняться соотношение $UA = A'$).

Пример 4.36 Рассмотрим квадратную матрицу $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$.

Решение. Запишем расширенную матрицу $(A \mid E)$:

$$\left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array}\right)$$

Выполним элементарные преобразования:

$$\left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array}\right) \xrightarrow{R_2:=R_2-3R_1} \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -2 & -3 & 1 \end{array}\right) \xrightarrow{R_2:=-\frac{1}{2}R_2} \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & \frac{3}{2} & -\frac{1}{2} \end{array}\right)$$

$$\xrightarrow{R_1:=R_1-2R_2} \left(\begin{array}{cc|cc} 1 & 0 & -2 & 1 \\ 0 & 1 & \frac{3}{2} & -\frac{1}{2} \end{array}\right)$$

Получили: $A' = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ (улучшенный ступенчатый вид) и $U = \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix}$.

Проверим: $UA = \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = A'$.

Теперь для решения любой системы $Ax = b$ достаточно вычислить Ub. $\square$

Пример 4.37 Рассмотрим прямоугольную матрицу $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$.

Решение. Запишем расширенную матрицу $(A \mid E)$, где E — единичная матрица подходящего размера 2×2:

$$\left(\begin{array}{ccc|cc} 1 & 2 & 3 & 1 & 0 \\ 4 & 5 & 6 & 0 & 1 \end{array}\right)$$

Выполним элементарные преобразования:

$$\left(\begin{array}{ccc|cc} 1 & 2 & 3 & 1 & 0 \\ 4 & 5 & 6 & 0 & 1 \end{array}\right) \xrightarrow{R_2:=R_2-4R_1} \left(\begin{array}{ccc|cc} 1 & 2 & 3 & 1 & 0 \\ 0 & -3 & -6 & -4 & 1 \end{array}\right) \xrightarrow{R_2:=-\frac{1}{3}R_2} \left(\begin{array}{ccc|cc} 1 & 2 & 3 & 1 & 0 \\ 0 & 1 & 2 & \frac{4}{3} & -\frac{1}{3} \end{array}\right)$$

$$\xrightarrow{R_1:=R_1-2R_2} \left(\begin{array}{ccc|cc} 1 & 0 & -1 & -\frac{5}{3} & \frac{2}{3} \\ 0 & 1 & 2 & \frac{4}{3} & -\frac{1}{3} \end{array}\right)$$

Получили: $A' = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{pmatrix}$ (улучшенный ступенчатый вид) и $U = \begin{pmatrix} -\frac{5}{3} & \frac{2}{3} \\ \frac{4}{3} & -\frac{1}{3} \end{pmatrix}$.

Проверим: $UA = \begin{pmatrix} -\frac{5}{3} & \frac{2}{3} \\ \frac{4}{3} & -\frac{1}{3} \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{pmatrix} = A'$.

Теперь для любой системы $Ax = b$ достаточно вычислить Ub и решить систему $A'x = Ub$. □

Обоснование метода:

Пусть A — матрица размера $m \times n$. Рассматриваем расширенную матрицу $(A \mid E)$, где E — единичная матрица размера $m \times m$.

Каждое элементарное преобразование строк реализуется умножением слева на элементарную матрицу. Если $U_k \cdots U_1 A = A'$, то $U = U_k \cdots U_1$.

Записывая преобразования над $(A \mid E)$, получаем:

$$U_k \cdots U_1(A \mid E) = (U_k \cdots U_1 A \mid U_k \cdots U_1 E) = (A' \mid U)$$

Таким образом, после приведения A к улучшенному ступенчатому виду A', справа от черты автоматически получается матрица U, удовлетворяющая $UA = A'$.

Для решения системы $Ax = b$ при поступлении нового b:

1. Вычисляем Ub (применяем те же преобразования к b)
2. Решаем систему $A'x = Ub$ (что проще, так как A' в улучшенном ступенчатом виде)

4.6.3 Домашнее задание # 1.7

Упражнение 4.7 Опишите все такие линейные отображения $f : \mathbb{R}^2 \to \mathbb{R}^2$ для которых $f \circ f = \mathrm{id}$ *т.е.,* $(f \circ f)(\mathbf{v}) = \mathbf{v}$ для любого $\mathbf{v} \in \mathbb{R}^2$. Как эти отображения действуют на смайлик из предыдущего ДЗ?

Упражнение 4.8 Опишите все такие линейные отображения $f : \mathbb{R}^2 \to \mathbb{R}^2$ для которых $f \circ f = 0$ *т.е.,* $(f \circ f)(\mathbf{v}) = \mathbf{0}$ для любого $\mathbf{v} \in \mathbb{R}^2$. Как эти отображения действуют на смайлик из предыдущего ДЗ?

Упражнение 4.9 Для данной матрицы $A = \begin{pmatrix} 1 & 1 & 0 & 2 \\ 2 & -1 & -1 & 1 \\ -3 & 3 & 2 & 0 \end{pmatrix}$ найдите такую квадратную матрицу U, что матрица UA имеет диагональный вид, а также выпишите этот вид.

Упражнение 4.10 Для матрицы $\begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}$ найдите обратную матрицу, если она существует.

Упражнение 4.11 Решите матричное уравнение

$$\begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 0 & 1 & 1 & \dots & 1 \\ 0 & 1 & 1 & \dots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \cdot X = \begin{pmatrix} 1 & 2 & 3 & \dots & n \\ 0 & 1 & 2 & \dots & n-1 \\ 0 & 0 & 1 & \dots & n-2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

4.7 Семинар # 1.8

Напомним основные определения:

Определение 4.7.1. *Перестановкой* множества $\{1, 2, \dots, n\}$ называется биекция $\pi : \{1, 2, \dots, n\} \to \{1, 2, \dots, n\}$.

Определение 4.7.2. *Инверсией* в перестановке π называется пара (i,j) такая, что $i < j$ но $\pi(i) > \pi(j)$.

Определение 4.7.3. *Знаком перестановки* называется число $\varepsilon(\pi) = (-1)^{\text{число инверсий в } \pi}$.

Пример 4.38 Рассмотрим перестановку $\pi = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 2 & 1 \end{pmatrix}$. Её инверсии:

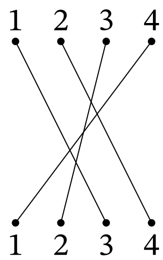
- (1,3): 3 > 2

- (1,4): $3 > 1$
- (2,3): $4 > 2$
- (2,4): $4 > 1$
- (3,4): $2 > 1$

Всего 5 инверсий, знак $\varepsilon(\pi) = (-1)^5 = -1$.

Геометрическое представление Тёрстона

Любую перестановку можно представить геометрически:



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

Каждое пересечение линий соответствует одной инверсии. В данном примере мы видим 5 пересечений.

4.7.1 Умножение перестановок

Определение 4.7.4. Для перестановок $\sigma, \tau \in \mathfrak{S}_n$ их *произведение* определяется как композиция:

$$(\sigma\tau)(i) = \sigma(\tau(i))$$

Важное замечание. Умножение перестановок **не коммутативно**!

Пример 4.39 Пусть $\sigma = (1\ 2)$ и $\tau = (2\ 3)$. Тогда:

$$\sigma\tau = (1\ 2)(2\ 3) = (1\ 3\ 2)$$
$$\tau\sigma = (2\ 3)(1\ 2) = (1\ 2\ 3)$$

Видим, что $\sigma\tau \neq \tau\sigma$.

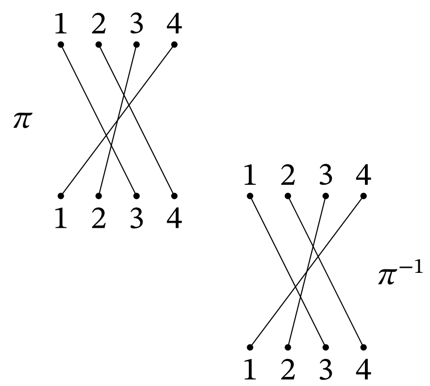
Обратная перестановка

Определение 4.7.5. *Обратной* к перестановке π называется перестановка π^{-1} такая, что $\pi\pi^{-1} = \pi^{-1}\pi = \mathrm{id}$.

Пример 4.40 Для $\pi = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 2 & 1 \end{pmatrix}$ находим:

$$\pi^{-1} = \begin{pmatrix} 3 & 4 & 2 & 1 \\ 1 & 2 & 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 1 & 2 \end{pmatrix}$$

Геометрически обратная перестановка получается "переворачиванием"диаграммы:



4.7.2 Вычисление знака через число пересечений

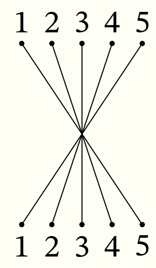
Из геометрической интерпретации следует, что знак перестановки можно вычислить как:

$$\varepsilon(\pi) = (-1)^{\text{число пересечений в диаграмме Тёрстона}}$$

Пример 4.41 Рассмотрим перестановку $\Delta_n \in \mathfrak{S}_n$, определяемую как:

$$\Delta_n(i) = n + 1 - i$$

Её геометрическое представление:



Число пересечений равно $1 + 2 + \cdots + (n-1) = \frac{n(n-1)}{2}$. Поэтому:

$$\varepsilon(\Delta_n) = (-1)^{\frac{n(n-1)}{2}} = \begin{cases} 1, & \text{если } n \equiv 0{,}1 \pmod 4 \\ -1, & \text{если } n \equiv 2{,}3 \pmod 4 \end{cases}$$

4.7.3 Решение уравнений с перестановками

Уравнение (1): $\sigma x = \pi$

Решение: $x = \sigma^{-1}\pi$

Пример 4.42 Пусть $\sigma = (1\ 2\ 3)$ и $\pi = (1\ 3\ 2)$. Найти x такое, что $\sigma x = \pi$.
Решение: $x = \sigma^{-1}\pi = (1\ 3\ 2)^{-1}(1\ 3\ 2) = (1\ 2\ 3)(1\ 3\ 2) = \mathrm{id}$

Уравнение (2): $x\sigma = \pi$

Решение: $x = \pi\sigma^{-1}$

Пример 4.43 Пусть $\sigma = (1\ 2\ 3)$ и $\pi = (1\ 3\ 2)$. Найти x такое, что $x\sigma = \pi$.
Решение: $x = \pi\sigma^{-1} = (1\ 3\ 2)(1\ 3\ 2)^{-1} = (1\ 3\ 2)(1\ 2\ 3) = (2\ 3)$

Уравнение (3): $\sigma x\tau = \pi$

Решение: $x = \sigma^{-1}\pi\tau^{-1}$

Пример 4.44 Пусть $\sigma = (1\ 2)$, $\tau = (2\ 3)$, $\pi = (1\ 3\ 2)$. Найти x.
Решение:

$$\begin{aligned} x &= \sigma^{-1}\pi\tau^{-1} = (1\ 2)^{-1}(1\ 3\ 2)(2\ 3)^{-1} \\ &= (1\ 2)(1\ 3\ 2)(2\ 3) = (1\ 3) \end{aligned}$$

4.7.4 Циклическая запись перестановок

Определение 4.7.6. *Циклом* длины k называется перестановка $(a_1 a_2 \dots a_k)$, где

$$\pi(a_1) = a_2, \pi(a_2) = a_3, \ \dots, \pi(a_k) = a_1$$

Теорема 4.7.1. *Любая перестановка единственным образом (с точностью до порядка циклов) раскладывается в произведение непересекающихся циклов.*

Пример 4.45 Разложим перестановку $\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 1 & 4 & 2 \end{pmatrix}$ на циклы:

- Начинаем с 1: $1 \to 3 \to 1$ получаем цикл (13).
- Следующий непросмотренный элемент 2: $2 \to 5 \to 2$ получаем цикл (25),
- Элемент 4 неподвижен поэтому имеем цикл (4).

Итого: $\pi = (1\ 3)(2\ 5)(4)$

Вычисление степеней перестановки

Чтобы вычислить π^k, достаточно возвести каждый цикл в степень k.

Пример 4.46 Рассмотрим перестановку

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 5 & 8 & 9 & 2 & 1 & 4 & 3 & 6 & 7 \end{pmatrix}$$

Разложим на циклы:

$$\begin{aligned} &1 \to 5 \to 1 \quad \text{цикл } (1\ 5) \\ &2 \to 8 \to 6 \to 4 \to 2 \quad \text{цикл } (2\ 8\ 6\ 4) \\ &3 \to 9 \to 7 \to 3 \quad \text{цикл } (3\ 9\ 7) \end{aligned}$$

Таким образом, $\pi = (1\ 5)(2\ 8\ 6\ 4)(3\ 9\ 7)$
Найдём $\pi^{100}, \pi^{101}, \pi^{102}$:

- Цикл $(1\ 5)$ длины 2: $(1\ 5)^{100} = \mathrm{id}$, $(1\ 5)^{101} = (1\ 5)$, $(1\ 5)^{102} = \mathrm{id}$
- Цикл $(2\ 8\ 6\ 4)$ длины 4: $(2\ 8\ 6\ 4)^{100} = (2\ 8\ 6\ 4)^{100 \bmod 4} = (2\ 8\ 6\ 4)^0 = \mathrm{id}$, $(2\ 8\ 6\ 4)^{101} = (2\ 8\ 6\ 4)^1 = (2\ 8\ 6\ 4)$, $(2\ 8\ 6\ 4)^{102} = (2\ 8\ 6\ 4)^2 = (2\ 6)(8\ 4)$
- Цикл $(3\ 9\ 7)$ длины 3: $(3\ 9\ 7)^{100} = (3\ 9\ 7)^{100 \bmod 3} = (3\ 9\ 7)^1 = (3\ 9\ 7)$, $(3\ 9\ 7)^{101} = (3\ 9\ 7)^2 = (3\ 7\ 9)$, $(3\ 9\ 7)^{102} = (3\ 9\ 7)^0 = \mathrm{id}$

Теперь перемножим результаты:

$$\begin{aligned} \pi^{100} &= \mathrm{id} \cdot \mathrm{id} \cdot (3\ 9\ 7) = (3\ 9\ 7) \\ \pi^{101} &= (1\ 5) \cdot (2\ 8\ 6\ 4) \cdot (3\ 7\ 9) = (1\ 5)(2\ 8\ 6\ 4)(3\ 7\ 9) \\ \pi^{102} &= \mathrm{id} \cdot (2\ 6)(8\ 4) \cdot \mathrm{id} = (2\ 6)(8\ 4) \end{aligned}$$

Важное замечание. Порядок перестановки равен НОК длин циклов в её разложении. В нашем примере порядок π равен НОК(2,4,3) = 12.

4.7.5 Домашнее задание # 1.8

Упражнение 4.12 Перемножить перестановки в указаном и обратном порядке

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

Упражнение 4.13 Записать в виде произведения независимых циклов перестановки:

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

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

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

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

Упражнение 4.14 Запишите в матричном виде перестановки:

(1) (136)(247)(5)

(2) (1654237)

Упражнение 4.15 Определите чётнос перестановки

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

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

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

Упражнение 4.16 Решите следующее уравнение относительно неизвестной перестановки X:

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

Упражнение 4.17 Вычислите σ^{77}, σ^{90}, σ^{119}, σ^{148}, если

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\ 5 & 9 & 11 & 3 & 8 & 10 & 2 & 12 & 7 & 4 & 6 & 1 \end{pmatrix}$$

Упражнение 4.18 На сколько может измениться число инверсий в перестановке

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & \dots & n \\ \sigma(1) & \sigma(2) & \sigma(3) & \dots & \sigma(n) \end{pmatrix},$$

если в её нижней строке поменять местами два соседних элемента? А как при этом изменится её знак?

Упражнение 4.19 Пусть ρ — перестановка, полученная из перестановки σ предыдущей задачи путём перестановки элементов $\sigma(i)$ и $\sigma(j)$ в нижней строке. Опишите перестановку $\sigma^{-1}\rho$.

4.8 Семинар # 1.9

Напомним основные определения

Напоминание. Бивекторы и операция внешнего произведения обладают следующими свойствами:

$$\begin{aligned}
\vec{\mathbf{a}} \wedge (\vec{\mathbf{b}} + \vec{\mathbf{c}}) &= \vec{\mathbf{a}} \wedge \vec{\mathbf{b}} + \vec{\mathbf{a}} \wedge \vec{\mathbf{c}} \\
\vec{\mathbf{a}} \wedge \vec{\mathbf{a}} &= 0 \\
\vec{\mathbf{a}} \wedge \vec{\mathbf{b}} &= -\vec{\mathbf{b}} \wedge \vec{\mathbf{a}} \\
(\alpha\vec{\mathbf{a}}) \wedge (\beta\vec{\mathbf{b}}) &= \alpha\beta(\vec{\mathbf{a}} \wedge \vec{\mathbf{b}})
\end{aligned}$$

Эти свойства, наряду с дистрибутивностью по первому аргументу $(\vec{\mathbf{a}}+\vec{\mathbf{b}})\wedge\vec{\mathbf{c}} = \vec{\mathbf{a}}\wedge\vec{\mathbf{c}}+\vec{\mathbf{b}}\wedge\vec{\mathbf{c}}$ (которая доказывается аналогично) и ассоциативностью относительно умножения на скаляр, являются аксиомами **внешней** (или **грассмановой**) **алгебры**. Поскольку при этой операции мы выходим за пределы векторов, она и называется внешним умножением.

Пример 4.47 Даны два вектора в $\mathbb{R}^2$: $\vec{\mathbf{a}} = (2, 1)$, $\vec{\mathbf{b}} = (1, 3)$.

1. Найдите внешнее произведение $\vec{\mathbf{a}} \wedge \vec{\mathbf{b}}$.
2. Вычислите площадь параллелограмма, построенного на векторах $\vec{\mathbf{a}}$ и $\vec{\mathbf{b}}$.
3. Проверьте, что коэффициент в выражении $\vec{\mathbf{a}} \wedge \vec{\mathbf{b}} = k\,(\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2)$ равен детерминанту матрицы, составленной из координат векторов $\vec{\mathbf{a}}$ и $\vec{\mathbf{b}}$.

Решение.

1. Внешнее произведение вычисляется как:

   $\vec{\mathbf{a}}\wedge\vec{\mathbf{b}} = (2\vec{\mathbf{e}}_1+1\vec{\mathbf{e}}_2)\wedge(1\vec{\mathbf{e}}_1+3\vec{\mathbf{e}}_2) = 2\cdot1(\vec{\mathbf{e}}_1\wedge\vec{\mathbf{e}}_1)+2\cdot3(\vec{\mathbf{e}}_1\wedge\vec{\mathbf{e}}_2)+1\cdot1(\vec{\mathbf{e}}_2\wedge\vec{\mathbf{e}}_1)+1\cdot3(\vec{\mathbf{e}}_2\wedge\vec{\mathbf{e}}_2)$.

   Учитывая, что $\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_1 = \vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_2 = 0$ и $\vec{\mathbf{e}}_2 \wedge \vec{\mathbf{e}}_1 = -\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2$, получаем:

   $$\vec{\mathbf{a}} \wedge \vec{\mathbf{b}} = (6 - 1)(\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2) = 5\,(\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2).$$

2. Площадь параллелограмма равна модулю коэффициента при $\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2$, то есть $|5| = 5$.

3. Матрица из координат векторов $\vec{\mathbf{a}}$ и $\vec{\mathbf{b}}$ (в столбцах):

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

Её детерминант равен $2 \cdot 3 - 1 \cdot 1 = 5$, что совпадает с коэффициентом $k = 5$.

□

Напоминание. Имеет место формул для n-векторов

$$\vec{\mathbf{a}}_1 \wedge \vec{\mathbf{a}}_2 \wedge \cdots \wedge \vec{\mathbf{a}}_n = \det(A) \cdot (\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \cdots \wedge \vec{\mathbf{e}}_n)$$

где коэффициент пропорциональности бивектора $\vec{\mathbf{a}}_1 \wedge \vec{\mathbf{a}}_2 \wedge \cdots \wedge \vec{\mathbf{a}}_n$ бивектору $\vec{\mathbf{e}}_1 \wedge \vec{\mathbf{e}}_2 \wedge \cdots \wedge \vec{\mathbf{e}}_n$ называется *детерминантом матрицы* A,

$$\det(A) := \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \cdot a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}.$$

4.8.1 Правило Саррюса для вычисления детерминанта 3×3

Для матрицы 3×3 существует удобное мнемоническое правило — **правило Саррюса**. Рассмотрим матрицу:

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

Правило Саррюса гласит:

$$\det(A) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$$

Вывод из общей формулы. Из общей формулы детерминанта для $n = 3$ имеем:

$$\det(A) = \sum_{\sigma \in S_3} \operatorname{sign}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} a_{3\sigma(3)}$$

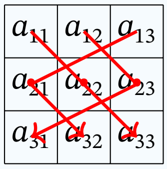
Выпишем все 6 перестановок и их знаки:

- (1,2,3): знак $+$, произведение $a_{11}a_{22}a_{33}$
- (1,3,2): знак $-$, произведение $a_{11}a_{23}a_{32}$
- (2,1,3): знак $-$, произведение $a_{12}a_{21}a_{33}$
- (2,3,1): знак $+$, произведение $a_{12}a_{23}a_{31}$
- (3,1,2): знак $+$, произведение $a_{13}a_{21}a_{32}$
- (3,2,1): знак $-$, произведение $a_{13}a_{22}a_{31}$

Суммируя, получаем в точности правило Саррюса. □

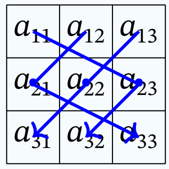
Наблюдение 4.29
Правило Саррюса можно наглядно представить с помощью двух диаграмм:

Положительные слагаемые (знак +)



$a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}$

Отрицательные слагаемые (знак: −)



$a_{13}a_{22}a_{31} + a_{11}a_{23}a_{32} + a_{12}a_{21}a_{33}$

Объяснение диаграмм:

- Левый рисунок показывает три диагонали, параллельные главной диагонали:
  - Главная диагональ: $a_{11} \to a_{22} \to a_{33}$
  - Первая параллельная: $a_{12} \to a_{23} \to a_{31}$
  - Вторая параллельная: $a_{13} \to a_{21} \to a_{32}$
- Правый рисунок показывает три диагонали, параллельные побочной диагонали:
  - Побочная диагональ: $a_{13} \to a_{22} \to a_{31}$
  - Первая параллельная: $a_{12} \to a_{21} \to a_{33}$
  - Вторая параллельная: $a_{11} \to a_{23} \to a_{32}$

Важное замечание. Полный детерминант вычисляется как сумма произведений элементов на красных диагоналях минус сумма произведений элементов на синих диагоналях:

$$\det(A) = (a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}) - (a_{13}a_{22}a_{31} + a_{11}a_{23}a_{32} + a_{12}a_{21}a_{33})$$

Пример 4.48 Даны три вектора в пространстве $\mathbb{R}^3$:

$$\vec{\mathbf{a}} = (1, 2, -1)$$
$$\vec{\mathbf{b}} = (3, -1, 2)$$
$$\vec{\mathbf{c}} = (2, 5, \lambda)$$

Найдите все значения параметра λ, при которых векторы $\vec{\mathbf{a}}, \vec{\mathbf{b}}, \vec{\mathbf{c}}$ лежат в одной плоскости.

Решение. Векторы лежат в одной плоскости тогда и только тогда, когда их внешнее произведение равно нулю:

$$\vec{\mathbf{a}} \wedge \vec{\mathbf{b}} \wedge \vec{\mathbf{c}} = 0$$

Это эквивалентно тому, что детерминант матрицы, составленной из координат этих векторов, равен нулю:

$$\det\begin{pmatrix}1 & 2 & -1\\ 3 & -1 & 2\\ 2 & 5 & \lambda\end{pmatrix} = 0$$

Вычислим детерминант по правилу Саррюса:

$$\begin{aligned}\det(A) &= 1\cdot(-1)\cdot\lambda + 2\cdot 2\cdot 2 + (-1)\cdot 3\cdot 5\\ &\quad - (-1)\cdot(-1)\cdot 2 - 1\cdot 2\cdot 5 - 2\cdot 3\cdot\lambda\\ &= (-\lambda) + 8 + (-15) - (2) - (10) - (6\lambda)\\ &= -\lambda + 8 - 15 - 2 - 10 - 6\lambda\\ &= -7\lambda - 19\end{aligned}$$

Приравниваем к нулю:

$$-7\lambda - 19 = 0 \quad\Rightarrow\quad \lambda = -\frac{19}{7}$$

Ответ: $\lambda = -\dfrac{19}{7}$ □

Пример 4.49 Выяснить, какие знаки из приведённых ниже произведений входят в детерминанти соответствующих порядков и с какими знаками:

(1) $a_{43}a_{21}a_{35}a_{12}a_{54}$,

(2) $a_{61}a_{23}a_{45}a_{36}a_{12}a_{54}$.

Решение. Так как общая формула для детерминанта матрицы A выглядит так

$$\det(A) := \sum_{\sigma\in S_n} \operatorname{sign}(\sigma)\cdot a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)}.$$

то нам нужно записать эти слагаемые в порядке возрастания левых индексов

(1) $a_{43}a_{21}a_{35}a_{12}a_{54} = a_{12}a_{21}a_{35}a_{43}a_{54}$,

(2) $a_{61}a_{23}a_{45}a_{36}a_{12}a_{54} = a_{12}a_{23}a_{36}a_{45}a_{54}a_{61}$.

Таким образом, для первого и второго слагаемого имеем следующие перестановки

(1)

□

Пример 4.50 Выяснить, какие знаки из приведённых ниже произведений входят в детерминанты соответствующих порядков и с какими знаками:

(1) $a_{43}a_{21}a_{35}a_{12}a_{54}$,

(2) $a_{61}a_{23}a_{45}a_{36}a_{12}a_{54}$.

Решение. Так как общая формула для детерминанта матрицы A выглядит так

$$\det(A) := \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \cdot a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}.$$

то нам нужно записать эти слагаемые в порядке возрастания левых индексов

(1) $a_{43}a_{21}a_{35}a_{12}a_{54} = a_{12}a_{21}a_{35}a_{43}a_{54}$,

(2) $a_{61}a_{23}a_{45}a_{36}a_{12}a_{54} = a_{12}a_{23}a_{36}a_{45}a_{54}a_{61}$.

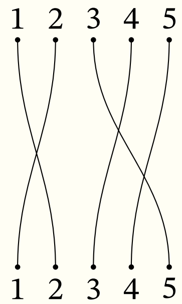
Таким образом, для первого и второго слагаемого имеем следующие перестановки:

(1) Перестановка σ_1 для $n = 5$: $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 5 & 3 & 4 \end{pmatrix}$, то есть $\sigma_1 = (2,1,5,3,4)$

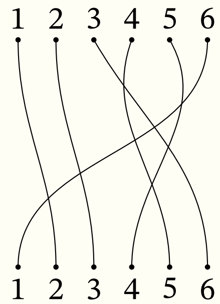
(2) Перестановка σ_2 для $n = 6$: $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 6 & 5 & 4 & 1 \end{pmatrix}$, то есть $\sigma_2 = (2,3,6,5,4,1)$

Построим диаграммы этих перестановок и посчитаем число пересечений:

Перестановка $\sigma_1 = (2,1,5,3,4)$



Перестановка $\sigma_2 = (2,3,6,5,4,1)$



Подсчёт числа пересечений:

(1) Для перестановки $\sigma_1 = (2,1,5,3,4)$ имеем всего 4 пересечения. Так как число пересечений чётное (4), то перестановка чётная и знак $+1$.

(2) Для перестановки $\sigma_2 = (2,3,6,5,4,1)$ имеем всего 8 пересечений. Так как число пересечений чётное (8), то перестановка чётная и знак $+1$.

Ответ:

(1) Произведение $a_{43}a_{21}a_{35}a_{12}a_{54}$ входит в детерминант со знаком $+1$

(2) Произведение $a_{61}a_{23}a_{45}a_{36}a_{12}a_{54}$ входит в детерминант со знаком $+1$

□

Пример 4.51 Вычислить знаки следующих произведений в детерминанте порядка n:

(1) Произведение элементов на главной диагонали: $a_{11}a_{22}\cdots a_{nn}$

(2) Произведение элементов на побочной диагонали: $a_{1n}a_{2,n-1}\cdots a_{n1}$

Решение. Вспомним, что каждое слагаемое в детерминанте соответствует некоторой перестановке $\sigma \in S_n$ и имеет знак $\operatorname{sign}(\sigma) = (-1)^{\text{число инверсий в } \sigma}$.

(1) Для главной диагонали перестановка тождественная: $\sigma(i) = i$ для всех i.

Число инверсий в тождественной перестановке равно 0, поэтому знак равен +1.

(2) Для побочной диагонали перестановка имеет вид: $\sigma(i) = n + 1 - i$.

Подсчитаем число инверсий в этой перестановке. Пара (i,j) с $i < j$ даёт инверсию, если $\sigma(i) > \sigma(j)$, то есть:

$$n + 1 - i > n + 1 - j \quad \Rightarrow \quad -i > -j \quad \Rightarrow \quad i < j$$

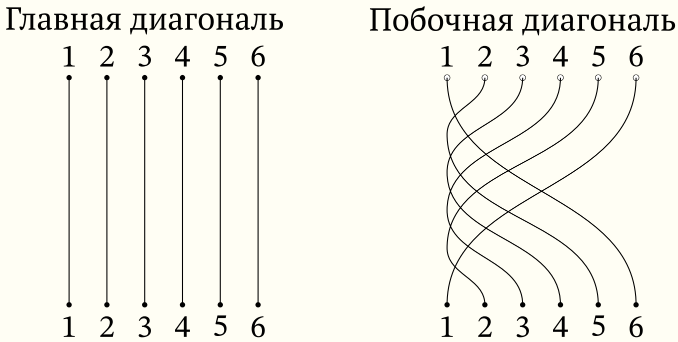
Это выполняется для всех пар (i,j) с $i < j$, так как $\sigma(i) = n+1-i$ является строго убывающей последовательностью.

Таким образом, каждая из $\binom{n}{2} = \frac{n(n-1)}{2}$ пар (i,j) с $i < j$ даёт инверсию.

Знак перестановки равен:

$$\operatorname{sign}(\sigma) = (-1)^{\frac{n(n-1)}{2}}$$

Геометрическая интерпретация с помощью диаграммы нитей:



Подсчёт числа пересечений для побочной диагонали:
Рассмотрим перестановку $\sigma(i) = n + 1 - i$. Подсчитаем число пересечений нитей:

- Первая нить (из 1 в n) пересекает все остальные $n - 1$ нити
- Вторая нить (из 2 в $n - 1$) пересекает все нити, кроме первой, то есть $n - 2$ нитей
- Третья нить (из 3 в $n - 2$) пересекает $n - 3$ нитей
- …
- k-я нить пересекает $n - k$ нитей

Общее число пересечений:

$$(n-1)+(n-2)+\cdots+1=\frac{n(n-1)}{2}$$

Так как каждое пересечение соответствует инверсии в перестановке, получаем тот же результат.
Ответ:

(1) Произведение на главной диагонали имеет знак $+1$

(2) Произведение на побочной диагонали имеет знак $(-1)^{\frac{n(n-1)}{2}}$

Важное замечание. Знак произведения на побочной диагонали зависит от n:

- Если $n \equiv 0 \pmod 4$ или $n \equiv 1 \pmod 4$, то $\frac{n(n-1)}{2}$ чётно и знак $+1$
- Если $n \equiv 2 \pmod 4$ или $n \equiv 3 \pmod 4$, то $\frac{n(n-1)}{2}$ нечётно и знак -1

Например:

- $n=2$: знак -1 (так как $\frac{2\cdot1}{2}=1$ - нечётно)
- $n=3$: знак -1 (так как $\frac{3\cdot2}{2}=3$ - нечётно)
- $n=4$: знак $+1$ (так как $\frac{4\cdot3}{2}=6$ - чётно)
- $n=5$: знак $+1$ (так как $\frac{5\cdot4}{2}=10$ - чётно)

□

Пример 4.52 Рассмотрим полином от переменной x который получается в результате раскрытия детерминанта у матрицы

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

Найдём коэффициенты при x^4, x^3, x^0.

Решение. Детерминант матрицы A представляет собой полином от x. Каждое слагаемое в определителе соответствует выбору по одному элементу из каждой строки и каждого столбца, умноженному на знак перестановки.

Отметим степени x в каждом элементе матрицы:

$$\begin{pmatrix} 5x & 1 & 2 & 3 \\ x & x & 1 & 2 \\ 1 & 2 & x & 3 \\ x & 1 & 2 & 2x \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \text{(степени } x\text{)}$$

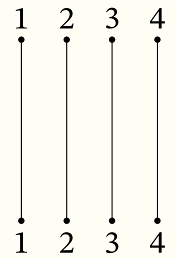
Коэффициент при x^4: Чтобы получить x^4, необходимо выбрать элементы со степенью x во всех четырёх строках. Рассмотрим возможные перестановки:

- Строка 1: единственный элемент со степенью x — это $5x$ в столбце 1.
- Строка 2: элементы со степенью x находятся в столбцах 1 и 2.
- Строка 3: единственный элемента со степенью x — это x в столбце 3.
- Строка 4: элементы со степенью x находятся в столбцах 1 и 4.

Поскольку в строке 1 мы обязаны выбрать столбец 1, а в строке 3 — столбец 3, то для строки 2 остаётся только столбец 2, а для строки 4 — столбец 4. Таким образом, единственная перестановка, дающая x^4, — это тождественная перестановка $\sigma = (1,2,3,4)$.

$$\begin{pmatrix} 5x & 1 & 2 & 3 \\ x & x & 1 & 2 \\ 1 & 2 & x & 3 \\ x & 1 & 2 & 2x \end{pmatrix}$$

Тождественная перестановка $\sigma = (1,2,3,4)$



Число пересечений: 0

Число пересечений: 0 (чётное), знак +1.
Произведение: $a_{11} \cdot a_{22} \cdot a_{33} \cdot a_{44} = (5x) \cdot x \cdot x \cdot (2x) = 10x^4$.
Вклад: $10x^4$.

Коэффициент при x^3: Чтобы получить x^3, необходимо выбрать элементы со степенью x ровно в трёх строках, а в одной строке — элемент без x. Рассмотрим случай, когда строка 1 выбирает элемент без x (так как если строка 1 выбирает элемент с x, то возникают противоречия при распределении столбцов).

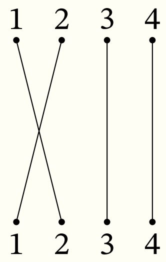
- Строка 1 без x: возможные столбцы — 2, 3, 4.
- Строки 2, 3, 4 должны выбирать элементы с x.

Возможные перестановки:

Перестановка $\sigma_1 = (2{,}1{,}3{,}4)$:

Перестановка $\sigma_1 = (2{,}1{,}3{,}4)$

$$\begin{pmatrix} 5x & 1 & 2 & 3 \\ x & x & 1 & 2 \\ 1 & 2 & x & 3 \\ x & 1 & 2 & 2x \end{pmatrix}$$



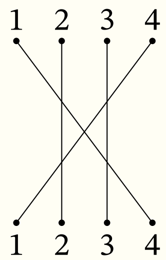
Число пересечений: 1

Число пересечений: 1 (нечётное), знак -1.
Произведение: $a_{12} \cdot a_{21} \cdot a_{33} \cdot a_{44} = 1 \cdot x \cdot x \cdot (2x) = 2x^3$.
Вклад: $-2x^3$.

Перестановка $\sigma_2 = (4{,}2{,}3{,}1)$:

Перестановка $\sigma_2 = (4{,}2{,}3{,}1)$

$$\begin{pmatrix} 5x & 1 & 2 & 3 \\ x & x & 1 & 2 \\ 1 & 2 & x & 3 \\ x & 1 & 2 & 2x \end{pmatrix}$$



Число пересечений: 5

Число пересечений: 5 (нечётное), знак -1.
Произведение: $a_{14} \cdot a_{22} \cdot a_{33} \cdot a_{41} = 3 \cdot x \cdot x \cdot x = 3x^3$.
Вклад: $-3x^3$.
Суммарный вклад для x^3: $-2x^3 - 3x^3 = -5x^3$.

Коэффициент при x^0: Чтобы получить x^0, необходимо во всех строках выбрать элементы без x. Возможные столбцы для каждой строки:

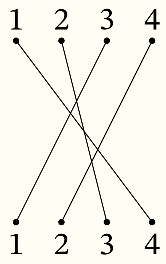
- Строка 1: столбцы 2, 3, 4
- Строка 2: столбцы 3, 4
- Строка 3: столбцы 1, 2, 4
- Строка 4: столбцы 2, 3

Возможные перестановки:

Перестановка $\sigma_3 = (4,3,1,2)$**:**

Перестановка $\sigma_3 = (4,3,1,2)$

$$\begin{pmatrix} 5x & 1 & 2 & 3 \\ x & x & 1 & 2 \\ 1 & 2 & x & 3 \\ x & 1 & 2 & 2x \end{pmatrix}$$



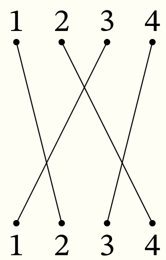
Число пересечений: 5

Число пересечений: 5 (нечётное), знак −1.
Произведение: $a_{14} \cdot a_{23} \cdot a_{31} \cdot a_{42} = 3 \cdot 1 \cdot 1 \cdot 1 = 3$.
Вклад: −3.

Перестановка $\sigma_4 = (2,4,1,3)$**:**

Перестановка $\sigma_4 = (2,4,1,3)$

$$\begin{pmatrix} 5x & 1 & 2 & 3 \\ x & x & 1 & 2 \\ 1 & 2 & x & 3 \\ x & 1 & 2 & 2x \end{pmatrix}$$



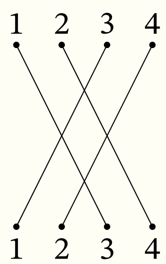
Число пересечений: 3

Число пересечений: 3 (нечётное), знак −1.
Произведение: $a_{12} \cdot a_{24} \cdot a_{31} \cdot a_{43} = 1 \cdot 2 \cdot 1 \cdot 2 = 4$.
Вклад: −4.

Перестановка $\sigma_5 = (3,4,1,2)$**:**

Перестановка $\sigma_5 = (3,4,1,2)$

$$\begin{pmatrix} 5x & 1 & 2 & 3 \\ x & x & 1 & 2 \\ 1 & 2 & x & 3 \\ x & 1 & 2 & 2x \end{pmatrix}$$



Число пересечений: 4

Число пересечений: 4 (чётное), знак +1.
Произведение: $a_{13} \cdot a_{24} \cdot a_{31} \cdot a_{42} = 2 \cdot 2 \cdot 1 \cdot 1 = 4$.
Вклад: +4.
Суммарный вклад для x^0: $-3 - 4 + 4 = -3$.

Ответ:

- Коэффициент при x^4: 10
- Коэффициент при x^3: -5
- Коэффициент при x^0: -3

□

4.8.2 Особенные типы матриц и их детерминанты

Наблюдение 4.30
Для матрицы определённой структуры детерминанты можно вычислять более удобном образом. Сейчас мы покажем один из таких случаев.

Теорема 4.8.1. *Пусть A — квадратная матрица размера $n \times n$, имеющая блочно-треугольный вид:*

$$A = \begin{pmatrix} P & Q \\ 0 & R \end{pmatrix} \quad \text{или} \quad A = \begin{pmatrix} P & 0 \\ Q & R \end{pmatrix}$$

где P и R — квадратные матрицы порядков k и m соответственно ($n = k + m$), а 0 — нулевая матрица соответствующего размера. Тогда

$$\det(A) = \det(P) \cdot \det(R).$$

Доказательство. Рассмотрим первый случай: $A = \begin{pmatrix} P & Q \\ 0 & R \end{pmatrix}$. Детерминант матрицы A выражается формулой:

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \cdot a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}.$$

Заметим, что матрица A имеет следующую структуру:

- В строках $1, \dots, k$ (соответствующих блоку P и Q) элементы в столбцах $1, \dots, k$ образуют матрицу P, а в столбцах $k + 1, \dots, n$ — матрицу Q.
- В строках $k+1, \dots, n$ (соответствующих блоку 0 и R) элементы в столбцах $1, \dots, k$ равны нулю (блок 0), а в столбцах $k + 1, \dots, n$ — матрица R.

Рассмотрим произвольное слагаемое детерминанта, соответствующий перестановке $\sigma \in S_n$. Если существует индекс $i \in \{1, \dots, k\}$ такой, что $\sigma(i) \in \{k + 1, \dots, n\}$ (т.е. в одной из первых k строк выбран элемент из блока Q), то в строках $k + 1, \dots, n$ будет хотя бы одна строка j, в которой придётся выбрать элемент из блока 0, потому что столбцы $k+1, \dots, n$ уже частично заняты.

Более формально: если в первых k строках выбрано r элементов из блоков Q (т.е. из столбцов $k+1, \dots, n$), то в оставшихся m строках доступно только $m-r$ столбцов из блока R, следовательно, в r строках из последних m придётся выбрать элементы из блока 0, которые равны нулю.

Таким образом, любое слагаемое детерминанта, в котором есть выбор из Q, равен нулю. Следовательно, ненулевые слагамемые детерминанта соответствуют таким перестановкам σ, для которых:

$$\sigma(\{1, \dots, k\}) \subseteq \{1, \dots, k\} \quad \text{и} \quad \sigma(\{k+1, \dots, n\}) \subseteq \{k+1, \dots, n\}.$$

Такие перестановки распадаются в прямое произведение перестановок $\sigma_1 \in S_k$ и $\sigma_2 \in S_m$, где σ_1 действует на $\{1, \dots, k\}$, а σ_2 — на $\{k+1, \dots, n\}$.

При этом знак перестановки σ равен $\operatorname{sign}(\sigma_1) \cdot \operatorname{sign}(\sigma_2)$, так как любые обмены между элементами из разных блоков не происходят.

Тогда ненулевые слагаемые детерминанта имеют вид:

$$\operatorname{sign}(\sigma_1) \cdot \operatorname{sign}(\sigma_2) \cdot \left(\prod_{i=1}^{k} a_{i\sigma_1(i)}\right) \cdot \left(\prod_{j=1}^{m} a_{k+j,k+\sigma_2(j)}\right).$$

Заметим, что $\prod_{i=1}^{k} a_{i\sigma_1(i)}$ — это слагаемое детерминанта матрицы P, соответствующий перестановке σ_1, а $\prod_{j=1}^{m} a_{k+j,k+\sigma_2(j)}$ — слагаемое детерминанта матрицы R, соответствующий перестановке σ_2. Суммируя по всем таким перестановкам, получаем:

$$\det(A) = \left(\sum_{\sigma_1 \in S_k} \operatorname{sign}(\sigma_1) \prod_{i=1}^{k} a_{i\sigma_1(i)}\right) \cdot \left(\sum_{\sigma_2 \in S_m} \operatorname{sign}(\sigma_2) \prod_{j=1}^{m} a_{k+j,k+\sigma_2(j)}\right) = \det(P) \cdot \det(R).$$

Второй случай: $A = \begin{pmatrix} P & 0 \\ Q & R \end{pmatrix}$ доказывается аналогично. Если в последних m строках (блок Q и R) выбрать элемент из блока Q (т.е. из столбцов $1, \dots, k$), то в первых k строках (блок P и 0) будет хотя бы одна строка, в которой придётся выбрать элемент из блока 0. Поэтому ненулевые слагаемые соответствуют перестановкам, для которых:

$$\sigma(\{1, \dots, k\}) \subseteq \{1, \dots, k\} \quad \text{и} \quad \sigma(\{k+1, \dots, n\}) \subseteq \{k+1, \dots, n\},$$

и снова получаем $\det(A) = \det(P) \cdot \det(R)$. □

Пример 4.53 Рассмотрим матрицу

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

Она имеет блочно-треугольный вид с $P = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}$ и $R = (5)$. Тогда

$$\det(A) = \det(P) \cdot \det(R) = (2 \cdot 1 - 1 \cdot 0) \cdot 5 = 2 \cdot 5 = 10.$$

Несложно проверить, что непосредственное вычисление даёт тот же результат.

Пример 4.54
Для вычисления детерминанта матрицы

$$A = \begin{pmatrix} a & 3 & 0 & 5 \\ 0 & b & 0 & 2 \\ a & 2 & c & 3 \\ 0 & 0 & 0 & d \end{pmatrix}$$

заметим, что её можно разбить на блоки следующим образом:

$$A = \left(\begin{array}{ccc|c} a & 3 & 0 & 5 \\ 0 & b & 0 & 2 \\ a & 2 & c & 3 \\ \hline 0 & 0 & 0 & d \end{array}\right).$$

Это соответствует представлению

$$A = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix},$$

где

$$B = \begin{pmatrix} a & 3 & 0 \\ 0 & b & 0 \\ a & 2 & c \end{pmatrix}, \quad C = \begin{pmatrix} 5 \\ 2 \\ 3 \end{pmatrix}, \quad D = \left(d\right).$$

По свойству детерминанта блочно-треугольной матрицы:

$$\det(A) = \det(B) \cdot \det(D) = \det(B) \cdot d.$$

Теперь вычислим $\det(B)$. Матрицу B также можно разбить на блоки:

$$B = \left(\begin{array}{cc|c} a & 3 & 0 \\ 0 & b & 0 \\ \hline a & 2 & c \end{array}\right).$$

Это соответствует представлению

$$B = \begin{pmatrix} E & F \\ G & H \end{pmatrix},$$

где

$$E = \begin{pmatrix} a & 3 \\ 0 & b \end{pmatrix}, \quad F = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad G = \begin{pmatrix} a & 2 \end{pmatrix}, \quad H = \begin{pmatrix} c \end{pmatrix}.$$

Заметим, что F — нулевой блок, поэтому по свойству детерминанта блочной матрицы:

$$\det(B) = \det(E) \cdot \det(H) = \det \begin{pmatrix} a & 3 \\ 0 & b \end{pmatrix} \cdot c.$$

Детерминант матрицы E равен:

$$\det(E) = a \cdot b - 3 \cdot 0 = ab.$$

Следовательно,

$$\det(B) = ab \cdot c = abc.$$

Окончательно:

$$\det(A) = d \cdot abc = abcd.$$

Таким образом, детерминант матрицы A равен $abcd$.

$$\boxed{abcd}$$

4.8.3 Появление детерминанта в процессе исключения Гаусса

Наблюдение 4.31
Рассмотрим матрицу

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

размера 3×3 и попытаемся понять, когда она обратима. Для этого мы обратимся к методу Гаусса приведения к ступенчатому виду. Мы можем считать, что $a_{11} \neq 0$ (иначе мы поменяем строки так, чтобы в позиции (1,1) стоял ненулевой элемент. Этого всегда можно добиться, так как в противном случае у матрицы A первый столбец нулевой, а это значит, что её нельзя будет привести к диагональному виду, где на диагонали стоят ненулевые элементы).
Умножим теперь вторую и третью строки на a_{11}, в результате получаем

$$A \sim \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{11}a_{21} & a_{11}a_{22} & a_{11}a_{23} \\ a_{11}a_{31} & a_{11}a_{32} & a_{11}a_{33} \end{pmatrix}.$$

Теперь мы умножаем первую строку на a_{21} и вычитаем из второй строки, далее, мы умножаем первую строку на a_{31} и вычитаем из третьей строки, в результате

получаем
$$A \sim \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{11}a_{22} - a_{12}a_{21} & a_{11}a_{23} - a_{13}a_{21} \\ 0 & a_{11}a_{32} - a_{12}a_{31} & a_{11}a_{33} - a_{13}a_{31} \end{pmatrix}.$$
Обозначим:
$$\begin{aligned} b_{22} &= a_{11}a_{22} - a_{12}a_{21}, \\ b_{23} &= a_{11}a_{23} - a_{13}a_{21}, \\ b_{32} &= a_{11}a_{32} - a_{12}a_{31}, \\ b_{33} &= a_{11}a_{33} - a_{13}a_{31}. \end{aligned}$$
Тогда матрица принимает вид:
$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & b_{22} & b_{23} \\ 0 & b_{32} & b_{33} \end{pmatrix}.$$
Теперь обнулим элемент в позиции (3,2). Для этого выполним преобразование:
$$\text{Новая третья строка} = b_{22} \cdot (\text{третья строка}) - b_{32} \cdot (\text{вторая строка}).$$
Элемент (3,2) становится:
$$b_{22} \cdot b_{32} - b_{32} \cdot b_{22} = 0.$$
Элемент (3,3) становится:
$$b_{22} \cdot b_{33} - b_{32} \cdot b_{23}.$$
Таким образом, матрица принимает вид:
$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & b_{22} & b_{23} \\ 0 & 0 & b_{22}b_{33} - b_{32}b_{23} \end{pmatrix}.$$
Вычислим выражение для элемента (3,3):
$$\begin{aligned} b_{22}b_{33} - b_{32}b_{23} = {} & (a_{11}a_{22} - a_{12}a_{21})(a_{11}a_{33} - a_{13}a_{31}) \\ & - (a_{11}a_{32} - a_{12}a_{31})(a_{11}a_{23} - a_{13}a_{21}). \end{aligned}$$
Раскроем скобки:
$$\begin{aligned} = {} & a_{11}^2 a_{22}a_{33} - a_{11}a_{22}a_{13}a_{31} - a_{11}a_{12}a_{21}a_{33} + a_{12}a_{21}a_{13}a_{31} \\ & - a_{11}^2 a_{32}a_{23} + a_{11}a_{32}a_{13}a_{21} + a_{11}a_{12}a_{31}a_{23} - a_{12}a_{31}a_{13}a_{21}. \end{aligned}$$
Сгруппируем слагаемые:
$$\begin{aligned} = {} & a_{11}^2(a_{22}a_{33} - a_{32}a_{23}) \\ & + a_{11}(-a_{22}a_{13}a_{31} + a_{32}a_{13}a_{21} - a_{12}a_{21}a_{33} + a_{12}a_{31}a_{23}) \\ & + (a_{12}a_{21}a_{13}a_{31} - a_{12}a_{31}a_{13}a_{21}). \end{aligned}$$

Заметим, что последняя скобка равна нулю. Вынесем a_{11} за скобку:

$$= a_{11}\left[a_{11}(a_{22}a_{33} - a_{32}a_{23}) - a_{22}a_{13}a_{31} + a_{32}a_{13}a_{21} - a_{12}a_{21}a_{33} + a_{12}a_{31}a_{23}\right].$$

Теперь распознаём в квадратной скобке формулу для детерминанта матрицы A:

$$\det(A) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}.$$

Перепишем наше выражение в том же порядке:

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

Таким образом:

$$b_{22}b_{33} - b_{32}b_{23} = a_{11} \cdot \det(A).$$

Итоговая ступенчатая форма матрицы:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & b_{22} & b_{23} \\ 0 & 0 & a_{11}\det(A) \end{pmatrix}.$$

Важное замечание. Это показывает, что процесс исключения Гаусса естественным образом приводит к появлению детерминанта! Матрица обратима тогда и только тогда, когда все диагональные элементы ступенчатой формы ненулевые, то есть когда:

$$a_{11} \neq 0, \quad b_{22} \neq 0, \quad \det(A) \neq 0.$$

В частности, условие $\det(A) \neq 0$ является необходимым и достаточным для обратимости матрицы.

Рассмотрим конкретный пример матрицы 3×3 и покажем, как в процессе исключения Гаусса детерминант появляется в угловом элементе.

Пример 4.55 Вычислим детерминант матрицы

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{pmatrix}$$

методом, демонстрирующим появление детерминанта в угловом элементе.
Шаг 1: Умножим вторую и третью строки на $a_{11} = 1$ (это не меняет матрицу):

$$A \sim \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{pmatrix}$$

Шаг 2: Обнулим элементы под (1,1). Для этого:

$$R_2 \leftarrow R_2 - 4R_1$$
$$R_3 \leftarrow R_3 - 7R_1$$

Получаем:

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -11 \end{pmatrix}$$

Обозначим:

$$\begin{aligned} b_{22} &= 1 \cdot 5 - 2 \cdot 4 = -3 \\ b_{23} &= 1 \cdot 6 - 3 \cdot 4 = -6 \\ b_{32} &= 1 \cdot 8 - 2 \cdot 7 = -6 \\ b_{33} &= 1 \cdot 10 - 3 \cdot 7 = -11 \end{aligned}$$

Шаг 3: Теперь обнулим элемент (3,2) с помощью преобразования:

$$R_3 \leftarrow b_{22} \cdot R_3 - b_{32} \cdot R_2$$

то есть:

$$R_3 \leftarrow (-3) \cdot R_3 - (-6) \cdot R_2 = -3R_3 + 6R_2$$

Вычисляем элементы новой третьей строки:

$$\begin{aligned} \text{Элемент (3,1)} &: -3 \cdot 0 + 6 \cdot 0 = 0 \\ \text{Элемент (3,2)} &: -3 \cdot (-6) + 6 \cdot (-3) = 18 - 18 = 0 \\ \text{Элемент (3,3)} &: -3 \cdot (-11) + 6 \cdot (-6) = 33 - 36 = -3 \end{aligned}$$

Получаем ступенчатую форму:

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

Согласно нашему общему выводу, элемент в позиции (3,3) должен равняться $a_{11} \cdot \det(A)$. В нашем случае:

$$a_{11} \cdot \det(A) = 1 \cdot \det(A) = -3$$

Следовательно, $\det(A) = -3$.

Проверка по правилу Саррюса. Вычислим детерминант непосредственно:

$$\begin{aligned} \det(A) &= 1 \cdot 5 \cdot 10 + 2 \cdot 6 \cdot 7 + 3 \cdot 4 \cdot 8 \\ &\quad - 3 \cdot 5 \cdot 7 - 1 \cdot 6 \cdot 8 - 2 \cdot 4 \cdot 10 \\ &= (50) + (84) + (96) - (105) - (48) - (80) \\ &= 230 - 233 = -3 \end{aligned}$$

Результат совпадает. □

Важное замечание. Этот пример наглядно демонстрирует, как в процессе исключения Гаусса детерминант матрицы появляется в угловом элементе ступенчатой формы. В общем случае для матрицы 3×3 после выполнения аналогичных преобразований элемент в позиции (3,3) будет равен $a_{11} \cdot \det(A)$.
Такой подход показывает, что детерминант — это не просто абстрактная формула, а естественная величина, возникающая в процессе решения систем линейных уравнений методом Гаусса.

4.8.4 Домашнее задание # 1.9

Упражнение 4.20 Даны три вектора в $\mathbb{R}^3$:

$$\vec{\mathbf{a}} = (1, -1, 2)$$
$$\vec{\mathbf{b}} = (3, 0, 1)$$
$$\vec{\mathbf{c}} = (2, 1, \lambda)$$

Найдите значение параметра λ, при котором векторы компланарны (лежат в одной плоскости).

Упражнение 4.21 Вычислите детерминант следующих матриц, используя правило Саррюса:

1. $A = \begin{pmatrix} 2 & -1 & 3 \\ 1 & 0 & 4 \\ -2 & 1 & 5 \end{pmatrix}$
2. $B = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}$

Упражнение 4.22 Определите, с каким знаком входят следующие произведения в детерминанты соответствующих порядков:

1. $a_{32}a_{13}a_{21}a_{44}a_{55}$ (порядок 5)
2. $a_{72}a_{34}a_{16}a_{53}a_{25}a_{41}a_{67}$ (порядок 7)

Для каждого случая запишите перестановку в двустрочной записи и вычислите её знак с помощью диаграммы нитей.

Упражнение 4.23 Рассмотрим матрицу

$$C = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 5 & 6 \\ 0 & 0 & 2 & 7 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Вычислите $\det(C)$, используя теорему о детерминанте блочно-треугольной матрицы.

Упражнение 4.24 Пусть дана матрица

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

Вычислите $\det(D)$, разбивая её на блоки.

Упражнение 4.25 Покажите, что если в матрице A размера $n \times n$ есть строка (столбец), состоящая из нулей, то $\det(A) = 0$. Объясните это с точки зрения внешнего произведения.

Упражнение 4.26 Даны векторы $\vec{\mathbf{a}} = (1, 2, 3)$, $\vec{\mathbf{b}} = (4, 5, 6)$, $\vec{\mathbf{c}} = (7, 8, 9)$.

1. Вычислите $\vec{\mathbf{a}} \wedge \vec{\mathbf{b}} \wedge \vec{\mathbf{c}}$.
2. Что можно сказать о линейной зависимости этих векторов?
3. Объясните геометрический смысл полученного результата.

Упражнение 4.27 Найдите коэффициент при x^4 в выражении детерминанта

$$\begin{vmatrix} 1 & 3 & x & -2 & x \\ x & 2 & 0 & 4 & 5 \\ x & 1 & x & 5 & 2 \\ 3 & x & -1 & 2 & -3 \\ -1 & -2 & 4 & x & -2 \end{vmatrix}.$$

4.9 Семинар 1-10. Основные свойства детерминанта

Напоминание.

(1) *Для любой матрицы A, имеет место равенство* $\det A^\top = \det A$.

(2) *Детерминант матрицы с нулевой строкой или столбцом равен нулю.*

(3) *Если произвольную строку матрицы умножить на число, то детерминант матрицы умножится на это число.*

(4) *Если строку матрицы умножить на число и прибавить к любой строке, то детерминант матрицы не изменится.*

(5) *Если в матрице есть две одинаковые строки, то её детерминант равен нулю.*

(6) *Если поменять местами две строки матрицы, то определитель меняет знак.*

(7) *Если строку матрицы представить в виде суммы двух векторов, то детерминант исходной матрицы равен сумме детерминантов двух матриц, в которых эта строка заменена на каждое из слагаемых.*

(8) *Если матрица является нижнетреугольной, то её детерминант равен произведению элементов на главной диагонали.*

(8) *Детерминант произведения двух квадратных матриц равен произведению их детерминантов:* $\det(AB) = \det(A)\det(B)$.

Для матрицы $A = (a_{ij})$ и для любого её элемента a_{ij} мы ввели понятие *минора* A_{ij} это детерминант матрицы которая получаеся из A вычёркиванием i-строки и j-ого столбца, тогда для любой строки, имеем

$$\det A = (-1)^{i+1} a_{i1} \det A_{i1} + (-1)^{i+2} a_{i2} \det A_{i2} + \cdots + (-1)^{i+n} a_{i+n} \det A_{in}.$$

Пример 4.56 Вычислить $\det A$ для матрицы $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{pmatrix}$ разложением по первой строке.

Решение. Разложим по первой строке:

$$\det A = (-1)^{1+1} a_{11} \det A_{11} + (-1)^{1+2} a_{12} \det A_{12} + (-1)^{1+3} a_{13} \det A_{13} + (-1)^{1+4} a_{14} \det A_{14}$$

$$= (-1)^2 \cdot 1 \cdot \det A_{11} + (-1)^3 \cdot 1 \cdot \det A_{12} + (-1)^4 \cdot 1 \cdot \det A_{13} + (-1)^5 \cdot 1 \cdot \det A_{14}$$

Вычислим миноры по правилу Саррюса:

$\det A_{11} = \begin{vmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{vmatrix} = (-1)(-1)(-1) + 1\cdot 1\cdot 1 + 1\cdot 1\cdot 1 - 1\cdot(-1)\cdot 1 - 1\cdot 1\cdot(-1) - (-1)\cdot 1\cdot 1 =$
$-1 + 1 + 1 + 1 + 1 + 1 = 4$

$\det A_{12} = \begin{vmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{vmatrix} = 1\cdot(-1)\cdot(-1)+1\cdot1\cdot1+1\cdot1\cdot1-1\cdot(-1)\cdot1-1\cdot1\cdot(-1)-1\cdot1\cdot1 =$
$1+1+1+1+1-1=4$

$\det A_{13} = \begin{vmatrix} 1 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & -1 \end{vmatrix} = 1\cdot1\cdot(-1)+(-1)\cdot1\cdot1+1\cdot1\cdot1-1\cdot1\cdot1-(-1)\cdot1\cdot(-1)-1\cdot1\cdot1 =$
$-1-1+1-1-1-1=-4$

$\det A_{14} = \begin{vmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{vmatrix} = 1\cdot1\cdot1+(-1)\cdot(-1)\cdot1+1\cdot1\cdot1-1\cdot1\cdot1-(-1)\cdot1\cdot1-1\cdot(-1)\cdot1 =$
$1+1+1-1+1+1=4$

Подставляем:

$$\det A = 1\cdot4+(-1)\cdot1\cdot4+1\cdot(-4)+(-1)\cdot1\cdot4 = 4-4-4-4=-8$$

□

Пример 4.57 Вычислить $\det A$ для матрицы $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{pmatrix}$ с помощью элементарных преобразований.

Решение. К первой строке прибавим вторую, третью и четвёртую строки:

$$\begin{pmatrix} 1+1+1+1 & 1+(-1)+1+1 & 1+1+(-1)+1 & 1+1+1+(-1) \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{pmatrix} = \begin{pmatrix} 4 & 2 & 2 & 2 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{pmatrix}$$

Разложим по первой строке:

$$\det A = (-1)^{1+1}\cdot4\cdot\det A_{11} + (-1)^{1+2}\cdot2\cdot\det A_{12} + (-1)^{1+3}\cdot2\cdot\det A_{13} + (-1)^{1+4}\cdot2\cdot\det A_{14}$$

Теперь вычислим $\det A_{11}$ с помощью преобразований:

$$A_{11} = \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}$$

К первой строке прибавим вторую и третью:

$$\begin{pmatrix} -1+1+1 & 1+(-1)+1 & 1+1+(-1) \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}$$

Из второй строки вычтем первую, из третьей строки вычтем первую:

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

Детерминант треугольной матрицы равен произведению диагональных элементов: $\det A_{11} = 1 \cdot (-2) \cdot (-2) = 4$
Аналогично вычисляем остальные миноры и получаем:

$$\det A = 4 \cdot 4 + (-1) \cdot 2 \cdot 4 + 2 \cdot (-4) + (-1) \cdot 2 \cdot 4 = 16 - 8 - 8 - 8 = -8$$

$\square$

Пример 4.58 Вычислить $\det A$ для матрицы $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{pmatrix}$ приведением к диагональному виду.

Решение. Выполним элементарные преобразования:
1. Вычтем первую строку из второй, третьей и четвёртой:

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

2. Обнулим элементы первого столбца кроме a_{11}: К первой строке прибавим $\frac{1}{2}$ второй строки, $\frac{1}{2}$ третьей строки и $\frac{1}{2}$ четвёртой строки:

$$\begin{pmatrix} 1+0+0+0 & 1+(-1)+0+0 & 1+0+(-1)+0 & 1+0+0+(-1) \\ 0 & -2 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix}$$

Детерминант диагональной матрицы равен произведению диагональных элементов:

$$\det A = 1 \cdot (-2) \cdot (-2) \cdot (-2) = -8$$

$\square$

Важное замечание. Все три метода дали одинаковый результат $\det A = -8$, что подтверждает правильность вычислений.

Напоминание. Напомним формулы Крамера, пусть дана система линеных уравнений $A\mathbf{x} = \mathbf{b}$, где $\mathbf{x} = (x_1, \dots, x_n)^\top$, $\mathbf{b} = (b_1, \dots, b_n)^\top$ и $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$.

Тогда неизвесные, при условии что $\det A \neq 0$ находятся так

$$x_i = \frac{\det A_i(\mathbf{b})}{\det A}, \qquad 1 \leq i \leq n,$$

где $A_i(\mathbf{b})$ – матрица полученная из A заменой его i-го столбца на столбец $\mathbf{b}$.

Пример 4.59 Найти коэффициенты a_0, a_1, a_2 полинома $\mathscr{L}(x) = a_0 + a_1 x + a_2 x^2$, если $\mathscr{L}(1) = 12$, $\mathscr{L}(2) = 13$, $\mathscr{L}(3) = 20$.

Решение. Из условий получаем систему уравнений:

$$\begin{cases} a_0 + a_1 + a_2 = 12 \\ a_0 + 2a_1 + 4a_2 = 13 \\ a_0 + 3a_1 + 9a_1 = 20. \end{cases}$$

Матрица системы: $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} 12 \\ 13 \\ 20 \end{pmatrix}$

Вычислим $\det A$:

$$\det A = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{vmatrix} = 1 \cdot (2 \cdot 9 - 4 \cdot 3) - 1 \cdot (1 \cdot 9 - 4 \cdot 1) + 1 \cdot (1 \cdot 3 - 2 \cdot 1)$$

$$= 1 \cdot (18 - 12) - 1 \cdot (9 - 4) + 1 \cdot (3 - 2) = 6 - 5 + 1 = 2$$

По правилу Крамера:

$$a_0 = \frac{\det A_1(\mathbf{b})}{\det A}, \quad a_1 = \frac{\det A_2(\mathbf{b})}{\det A}, \quad a_2 = \frac{\det A_3(\mathbf{b})}{\det A}$$

Вычислим $\det A_1(\mathbf{b})$:

$$A_1(\mathbf{b}) = \begin{pmatrix} 12 & 1 & 1 \\ 13 & 2 & 4 \\ 20 & 3 & 9 \end{pmatrix}$$

$$\det A_1(\mathbf{b}) = 12 \cdot (2 \cdot 9 - 4 \cdot 3) - 1 \cdot (13 \cdot 9 - 4 \cdot 20) + 1 \cdot (13 \cdot 3 - 2 \cdot 20)$$

$$= 12 \cdot (18 - 12) - 1 \cdot (117 - 80) + 1 \cdot (39 - 40) = 12 \cdot 6 - 37 - 1 = 72 - 38 = 34$$

Вычислим $\det A_2(\mathbf{b})$:

$$A_2(\mathbf{b}) = \begin{pmatrix} 1 & 12 & 1 \\ 1 & 13 & 4 \\ 1 & 20 & 9 \end{pmatrix}$$

$$\det A_2(\mathbf{b}) = 1 \cdot (13 \cdot 9 - 4 \cdot 20) - 12 \cdot (1 \cdot 9 - 4 \cdot 1) + 1 \cdot (1 \cdot 20 - 13 \cdot 1)$$

$$= 1 \cdot (117 - 80) - 12 \cdot (9 - 4) + 1 \cdot (20 - 13) = 37 - 12 \cdot 5 + 7 = 37 - 60 + 7 = -16$$

Вычислим $\det A_3(\mathbf{b})$:

$$A_3(\mathbf{b}) = \begin{pmatrix} 1 & 1 & 12 \\ 1 & 2 & 13 \\ 1 & 3 & 20 \end{pmatrix}$$

$$\det A_3(\mathbf{b}) = 1 \cdot (2 \cdot 20 - 13 \cdot 3) - 1 \cdot (1 \cdot 20 - 13 \cdot 1) + 12 \cdot (1 \cdot 3 - 2 \cdot 1)$$

$$= 1 \cdot (40 - 39) - 1 \cdot (20 - 13) + 12 \cdot (3 - 2) = 1 - 7 + 12 = 6$$

Находим коэффициенты:

$$a_0 = \frac{34}{2} = 17, \quad a_1 = \frac{-16}{2} = -8, \quad a_2 = \frac{6}{2} = 3$$

Искомый полином: $\mathscr{L}(x) = 17 - 8x + 3x^2$ □

Проверим правильность решения:

$$\mathscr{L}(1) = 17 - 8 \cdot 1 + 3 \cdot 1^2 = 17 - 8 + 3 = 12$$
$$\mathscr{L}(2) = 17 - 8 \cdot 2 + 3 \cdot 4 = 17 - 16 + 12 = 13$$
$$\mathscr{L}(3) = 17 - 8 \cdot 3 + 3 \cdot 9 = 17 - 24 + 27 = 20$$

Условия выполняются.

Предыдущая задача решается в общем виде с помощью так называется интерполяционного полинома Лагранжа $\mathscr{L}(x)$ который в имеет степень $n-1$ и принимает заданные значения $\mathscr{L}(x_1) = y_1, \dots, \mathscr{L}(x_n) = y_n$.

Пример 4.60 Вычислить детерминант Вандермонда:

$$V(x_1, \dots, x_n) = \begin{vmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{vmatrix}$$

Решение. Будем последовательно применять элементарные преобразования столбцов.

Вычтем из последнего столбца предпоследний, умноженный на x_1:

$$\begin{vmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{vmatrix} = \begin{vmatrix} 1 & x_1 & x_1^2 & \cdots & 0 \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} - x_1 x_2^{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} - x_1 x_n^{n-2} \end{vmatrix}$$

Заметим, что для $i = 2, \dots, n$:

$$x_i^{n-1} - x_1 x_i^{n-2} = x_i^{n-2}(x_i - x_1)$$

Теперь вычтем из предпоследнего столбца предыдущий, умноженный на x_1:

$$= \begin{vmatrix} 1 & x_1 & x_1^2 & \cdots & 0 \\ 1 & x_2 & x_2^2 - x_1 x_2 & \cdots & x_2^{n-2}(x_2 - x_1) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 - x_1 x_n & \cdots & x_n^{n-2}(x_n - x_1) \end{vmatrix}$$

Продолжая этот процесс для всех столбцов справа налево, в конечном итоге получим:

$$V(x_1, \dots, x_n) = \begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & x_2 - x_1 & x_2(x_2 - x_1) & \cdots & x_2^{n-2}(x_2 - x_1) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n - x_1 & x_n(x_n - x_1) & \cdots & x_n^{n-2}(x_n - x_1) \end{vmatrix}$$

Разложим по первой строке. Единственный ненулевой элемент в первой строке — это 1 в позиции (1,1), поэтому:

$$V(x_1, \dots, x_n) = 1 \cdot (-1)^{1+1} \cdot \begin{vmatrix} x_2 - x_1 & x_2(x_2 - x_1) & \cdots & x_2^{n-2}(x_2 - x_1) \\ x_3 - x_1 & x_3(x_3 - x_1) & \cdots & x_3^{n-2}(x_3 - x_1) \\ \vdots & \vdots & \ddots & \vdots \\ x_n - x_1 & x_n(x_n - x_1) & \cdots & x_n^{n-2}(x_n - x_1) \end{vmatrix}$$

В каждой строке полученного определителя вынесем общие множители:

$$= (x_2 - x_1)(x_3 - x_1) \cdots (x_n - x_1) \cdot \begin{vmatrix} 1 & x_2 & x_2^2 & \cdots & x_2^{n-2} \\ 1 & x_3 & x_3^2 & \cdots & x_3^{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-2} \end{vmatrix}$$

Но последний определитель — это $V(x_2, \dots, x_n)$. Таким образом, мы получили рекуррентное соотношение:

$$V(x_1, \dots, x_n) = \prod_{i=2}^{n} (x_i - x_1) \cdot V(x_2, \dots, x_n)$$

Применяя это соотношение последовательно, получаем:

$$V(x_1, \dots, x_n) = \prod_{1 \le i < j \le n} (x_j - x_i)$$

Важное замечание. Детерминант Вандермонда отличен от нуля тогда и только тогда, когда все числа $x_1, \dots, x_n$ попарно различны. Это имеет важное значение в интерполяции: система уравнений для нахождения коэффициентов полинома имеет единственное решение, если все узлы интерполяции различны.

□

Пример 4.61 Найти коэффициенты $a_0, a_1, \dots, a_{n-1}$ полинома $\mathcal{L}(x) = a_0 + a_1 x + \cdots + a_{n-1} x^{n-1}$, удовлетворяющего условиям $\mathcal{L}(x_1) = y_1$, $\mathcal{L}(x_2) = y_2$, ..., $\mathcal{L}(x_n) = y_n$.

Решение. Получаем систему линейных уравнений:

$$\begin{aligned} a_0 + a_1 x_1 + a_2 x_1^2 + \cdots + a_{n-1} x_1^{n-1} &= y_1 \\ a_0 + a_1 x_2 + a_2 x_2^2 + \cdots + a_{n-1} x_2^{n-1} &= y_2 \\ &\vdots \\ a_0 + a_1 x_n + a_2 x_n^2 + \cdots + a_{n-1} x_n^{n-1} &= y_n \end{aligned}$$

Матрица системы — матрица Вандермонда:

$$A = \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

Детерминант матрицы A равен:

$$\det A = V(x_1, \dots, x_n) = \prod_{1 \le i < j \le n} (x_j - x_i)$$

По правилу Крамера, для каждого $k = 0, 1, \dots, n-1$:

$$a_k = \frac{\det A_{k+1}(\mathbf{b})}{\det A}$$

где $A_{k+1}(\mathbf{b})$ — матрица, полученная из A заменой $(k+1)$-го столбца на столбец $\mathbf{b}$. Таким образом, коэффициенты полинома Лагранжа единственным образом определяются по формулам:

$$a_k = \frac{\begin{vmatrix} 1 & x_1 & \cdots & x_1^{k-1} & y_1 & x_1^{k+1} & \cdots & x_1^{n-1} \\ 1 & x_2 & \cdots & x_2^{k-1} & y_2 & x_2^{k+1} & \cdots & x_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \cdots & x_n^{k-1} & y_n & x_n^{k+1} & \cdots & x_n^{n-1} \end{vmatrix}}{\prod_{1 \le i < j \le n} (x_j - x_i)}$$

Наблюдение 4.32 Если все x_i попарно различны, то $\det A \neq 0$, и решение существует и единственно.

□

Важное замечание. Правило Крамера даёт явные формулы для коэффициентов интерполяционного полинома через детерминанты. Однако на практике для вычисления интерполяционного полинома чаще используют либо метод непосредственного решения системы, либо построение полинома в форме Лагранжа или Ньютона.

Пример 4.62 В частном случае $n = 3$ (квадратичный полином) с узлами x_1, x_2, x_3 и значениями y_1, y_2, y_3 получаем:

$$a_0 = \frac{\begin{vmatrix} y_1 & x_1 & x_1^2 \\ y_2 & x_2 & x_2^2 \\ y_3 & x_3 & x_3^2 \end{vmatrix}}{(x_2 - x_1)(x_3 - x_1)(x_3 - x_2)}, \quad a_1 = \frac{\begin{vmatrix} 1 & y_1 & x_1^2 \\ 1 & y_2 & x_2^2 \\ 1 & y_3 & x_3^2 \end{vmatrix}}{(x_2 - x_1)(x_3 - x_1)(x_3 - x_2)},$$

и

$$a_2 = \frac{\begin{vmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{vmatrix}}{(x_2 - x_1)(x_3 - x_1)(x_3 - x_2)}.$$

В итоге получаем

$$\mathscr{L}(x) = y_1 \frac{(x - x_2)(x - x_3)}{(x_1 - x_2)(x_1 - x_3)} + y_2 \frac{(x - x_1)(x - x_3)}{(x_2 - x_1)(x_2 - x_3)} + y_3 \frac{(x - x_1)(x - x_2)}{(x_3 - x_1)(x_3 - x_2)}.$$

Пример 4.63 Показать, что интерполяционный полином Лагранжа можно записать в виде:

$$\mathscr{L}(x) = y_1 \ell_1(x) + \cdots + y_n \ell_n(x)$$

где

$$\ell_i(x) = \prod_{\substack{j=1 \\ j \neq i}}^{n} \frac{x - x_j}{x_i - x_j}$$

Решение. Рассмотрим систему уравнений для коэффициентов полинома $\mathscr{L}(x) = a_0 + a_1 x + \cdots + a_{n-1} x^{n-1}$:

$$\begin{aligned} a_0 + a_1 x_1 + \cdots + a_{n-1} x_1^{n-1} &= y_1 \\ a_0 + a_1 x_2 + \cdots + a_{n-1} x_2^{n-1} &= y_2 \\ &\vdots \\ a_0 + a_1 x_n + \cdots + a_{n-1} x_n^{n-1} &= y_n \end{aligned}$$

Матрица системы — матрица Вандермонда:

$$A = \begin{pmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ 1 & x_2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \cdots & x_n^{n-1} \end{pmatrix}$$

По правилу Крамера, коэффициент a_k выражается как:

$$a_k = \frac{\det A_{k+1}(\mathbf{b})}{\det A}$$

где $A_{k+1}(\mathbf{b})$ — матрица, полученная из A заменой $(k+1)$-го столбца на столбец $\mathbf{b} = (y_1, \dots, y_n)^T$.
Рассмотрим значение полинома в точке x:

$$\mathscr{L}(x) = a_0 + a_1 x + \cdots + a_{n-1} x^{n-1} = \sum_{k=0}^{n-1} \frac{\det A_{k+1}(\mathbf{b})}{\det A} x^k$$

Это выражение можно переписать как разложение определителя по последнему столбцу следующей матрицы:

$$B(x) = \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} & y_1 \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} & y_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} & y_n \\ 1 & x & x^2 & \cdots & x^{n-1} & \mathscr{L}(x) \end{pmatrix}$$

Заметим, что если $\mathscr{L}(x)$ — интерполяционный полином, то последняя строка является линейной комбинацией предыдущих строк, поэтому $\det B(x) = 0$.
Разложим $\det B(x)$ по последней строке:

$$0 = \sum_{j=1}^{n+1} (-1)^{n+1+j} b_{n+1,j} M_{n+1,j}$$

где $M_{n+1,j}$ — минор элемента $b_{n+1,j}$.
Заметим, что:

$$b_{n+1,1} = 1, \qquad M_{n+1,1} = \begin{vmatrix} x_1 & x_1^2 & \cdots & x_1^{n-1} & y_1 \\ x_2 & x_2^2 & \cdots & x_2^{n-1} & y_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_n & x_n^2 & \cdots & x_n^{n-1} & y_n \end{vmatrix}$$

$$b_{n+1,2} = x, \qquad M_{n+1,2} = \begin{vmatrix} 1 & x_1^2 & \cdots & x_1^{n-1} & y_1 \\ 1 & x_2^2 & \cdots & x_2^{n-1} & y_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & x_n^2 & \cdots & x_n^{n-1} & y_n \end{vmatrix}$$

$$\vdots$$

$$b_{n+1,n} = x^{n-1}, \qquad M_{n+1,n} = \begin{vmatrix} 1 & x_1 & \cdots & x_1^{n-2} & y_1 \\ 1 & x_2 & \cdots & x_2^{n-2} & y_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & x_n & \cdots & x_n^{n-2} & y_n \end{vmatrix}$$

$$b_{n+1,n+1} = \mathscr{L}(x), \qquad M_{n+1,n+1} = \begin{vmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ 1 & x_2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \cdots & x_n^{n-1} \end{vmatrix} = \det A$$

Таким образом:

$$0 = \sum_{k=0}^{n-1} (-1)^{n+1+k+1} x^k \cdot \det A_{k+1}(\mathbf{b}) + (-1)^{2n+2} \mathscr{L}(x) \cdot \det A$$

Упростим знаки:

$$0 = \sum_{k=0}^{n-1} (-1)^{n+k} x^k \cdot \det A_{k+1}(\mathbf{b}) + \mathscr{L}(x) \cdot \det A$$

Отсюда:

$$\mathscr{L}(x) = \sum_{k=0}^{n-1} (-1)^{n+k+1} \frac{\det A_{k+1}(\mathbf{b})}{\det A} x^k$$

Сравнивая с исходным выражением для $\mathscr{L}(x)$, видим, что знаки должны согласовываться. Более простой подход — рассмотреть определитель Вандермонда с заменой одного столбца.
Рассмотрим определитель:

$$D_i(x) = \begin{vmatrix} 1 & x_1 & \cdots & x_1^{i-1} & y_1 & x_1^{i+1} & \cdots & x_1^{n-1} \\ 1 & x_2 & \cdots & x_2^{i-1} & y_2 & x_2^{i+1} & \cdots & x_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \cdots & x_n^{i-1} & y_n & x_n^{i+1} & \cdots & x_n^{n-1} \\ 1 & x & \cdots & x^{i-1} & \mathscr{L}(x) & x^{i+1} & \cdots & x^{n-1} \end{vmatrix}$$

Этот определитель равен нулю, так как последняя строка является линейной комбинацией предыдущих. Разложим его по последнему столбцу, содержащему $y_1, \dots, y_n, \mathscr{L}(x)$:

$$0 = \sum_{j=1}^{n} (-1)^{n+1+j} y_j M_j + (-1)^{2n+2} \mathscr{L}(x) M_{n+1}$$

где M_j — минор элемента y_j, а M_{n+1} — минор элемента $\mathscr{L}(x)$.
Заметим, что $M_{n+1} = V(x_1, \dots, x_n) = \prod_{1 \le i < j \le n} (x_j - x_i)$.
Аналогично, $M_j = (-1)^{n-j} V(x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n, x)$.
Таким образом:

$$\mathscr{L}(x) = \sum_{j=1}^{n} y_j \frac{V(x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n, x)}{V(x_1, \dots, x_n)}$$

Вычислим отношение определителей Вандермонда:

$$\frac{V(x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n, x)}{V(x_1, \dots, x_n)} = \frac{\prod_{1 \le k < l \le n, k, l \neq j} (x_l - x_k) \cdot \prod_{i \neq j} (x - x_i)}{\prod_{1 \le k < l \le n} (x_l - x_k)}$$

Заметим, что:

$$\prod_{1 \le k < l \le n} (x_l - x_k) = \left(\prod_{i \neq j} (x_j - x_i) \right) \cdot \prod_{1 \le k < l \le n, k, l \neq j} (x_l - x_k)$$

Поэтому:

$$\frac{V(x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n, x)}{V(x_1, \dots, x_n)} = \frac{\prod_{i \neq j} (x - x_i)}{\prod_{i \neq j} (x_j - x_i)} = \ell_j(x)$$

Таким образом:

$$\mathscr{L}(x) = \sum_{j=1}^{n} y_j \ell_j(x)$$

где

$$\ell_j(x) = \prod_{\substack{i=1 \\ i\neq j}}^{n} \frac{x - x_i}{x_j - x_i}$$

$\square$

Важное замечание. Полученная формула называется интерполяционной формулой Лагранжа. Она явно выражает интерполяционный полином через заданные значения y_i в узлах x_i.

Наблюдение 4.33 В частном случае $n = 3$ получаем:

$$\mathscr{L}(x) = y_1 \frac{(x-x_2)(x-x_3)}{(x_1-x_2)(x_1-x_3)} + y_2 \frac{(x-x_1)(x-x_3)}{(x_2-x_1)(x_2-x_3)} + y_3 \frac{(x-x_1)(x-x_2)}{(x_3-x_1)(x_3-x_2)}$$

Пример 4.64 Сравнить детерминанты матрицы $A = (a_{ij})$ и матрицы A', где: 1. $A' = (-a_{ij})$

Решение. Матрица A' получается из A умножением каждой строки на -1. Поскольку матрица имеет размер $n \times n$, получаем:

$$\det A' = (-1)^n \det A$$

$\square$

Пример 4.65 2. $A' = \begin{pmatrix} \mathbf{c}_2(A) & \dots & \mathbf{c}_n(A) & \mathbf{c}_1(A) \end{pmatrix}$

Решение. Матрица A' получается из A циклическим сдвигом столбцов влево. Такой сдвиг можно реализовать с помощью $n-1$ транспозиций соседних столбцов. Поэтому:

$$\det A' = (-1)^{n-1} \det A$$

$\square$

Пример 4.66 3. $A' = \begin{pmatrix} \mathbf{c}_{\pi(1)}(A) & \dots & \mathbf{c}_{\pi(n)}(A) \end{pmatrix}$

Решение. Матрица A' получается из A перестановкой столбцов согласно перестановке π. Тогда:

$$\det A' = \operatorname{sign}(\pi) \cdot \det A$$

где $\operatorname{sign}(\pi)$ — знак перестановки π. $\square$

Пример 4.67 4. $A' = \begin{pmatrix} \mathbf{r}_1(A) + \mathbf{r}_2(A) \\ \mathbf{r}_2(A) + \mathbf{r}_3(A) \\ \vdots \\ \mathbf{r}_{n-1}(A) + \mathbf{r}_n(A) \\ \mathbf{r}_n(A) \end{pmatrix}$

Решение. Рассмотрим последовательность элементарных преобразований:

$$\begin{aligned} &\mathbf{r}'_1 = \mathbf{r}_1 + \mathbf{r}_2 \\ &\mathbf{r}'_2 = \mathbf{r}_2 + \mathbf{r}_3 \\ &\vdots \\ &\mathbf{r}'_{n-1} = \mathbf{r}_{n-1} + \mathbf{r}_n \\ &\mathbf{r}'_n = \mathbf{r}_n \end{aligned}$$

Выполним обратные преобразования, начиная с последней строки:

$$\begin{aligned} &\mathbf{r}_{n-1} = \mathbf{r}'_{n-1} - \mathbf{r}'_n \\ &\mathbf{r}_{n-2} = \mathbf{r}'_{n-2} - \mathbf{r}_{n-1} \\ &\vdots \\ &\mathbf{r}_1 = \mathbf{r}'_1 - \mathbf{r}_2 \end{aligned}$$

Поскольку каждое преобразование является элементарной операцией сложения строки с другой строкой, оно не меняет детерминант. Поэтому:

$$\det A' = \det A$$

□

Пример 4.68 5. $A' = \begin{pmatrix} \mathbf{r}_1(A) + \mathbf{r}_2(A) \\ \mathbf{r}_2(A) + \mathbf{r}_3(A) \\ \vdots \\ \mathbf{r}_{n-1}(A) + \mathbf{r}_n(A) \\ \mathbf{r}_n(A) + \mathbf{r}_1(A) \end{pmatrix}$

Решение. Используем многолинейность детерминанта по строкам. Обозначим исходные строки матрицы A как $R_1, R_2, \ldots, R_n$. Тогда:

$$\det A' = \det(R_1 + R_2, R_2 + R_3, \ldots, R_{n-1} + R_n, R_n + R_1)$$

Разложим детерминант по линейности в каждой строке. Это даст сумму 2^n слагаемых вида:

$$\det(X_1, X_2, \ldots, X_n)$$

где каждый X_i равен либо R_i, либо R_{i+1} (причём для $i = n$: либо R_n, либо R_1).
Заметим, что если в наборе $X_1, X_2, \ldots, X_n$ какая-либо строка повторяется, то детерминант равен нулю. Ненулевой вклад дают только те наборы, где все строки различны.

Рассмотрим два случая:

1. $X_1 = R_1, X_2 = R_2, \dots, X_n = R_n$. Тогда:
$$\det(R_1, R_2, \dots, R_n) = \det A$$

2. $X_1 = R_2, X_2 = R_3, \dots, X_{n-1} = R_n, X_n = R_1$. Это циклический сдвиг строк. Перестановка строк $(1,2,\dots,n) \to (2,3,\dots,n,1)$ является циклом длины n, её знак равен $(-1)^{n-1}$. Поэтому:
$$\det(R_2, R_3, \dots, R_n, R_1) = (-1)^{n-1}\det A$$

Любой другой выбор приведёт к повторению строки. Например, если в какой-то позиции выбрать «левый» элемент, а в соседней — «правый», то возникнет повторение. Таким образом, других ненулевых слагаемых нет.
Суммируя два ненулевых слагаемых, получаем:
$$\det A' = \det A + (-1)^{n-1}\det A = [1 + (-1)^{n-1}]\det A$$

Упростим выражение:
$$1 + (-1)^{n-1} = 1 - (-1)^n = \begin{cases} 2, & \text{если } n \text{ нечётное} \\ 0, & \text{если } n \text{ чётное} \end{cases}$$

Окончательно:
$$\det A' = \begin{cases} 2\det A, & \text{если } n \text{ нечётное} \\ 0, & \text{если } n \text{ чётное} \end{cases}$$

□

Важное замечание. В случае 5 при чётном n строки матрицы A' линейно зависимы, поэтому её детерминант равен нулю. При нечётном n детерминант удваивается.

4.10 Домашнее Задание # 1.10

Упражнение 4.28 Как изменится детерминант матрицы порядка n, если его строки написать в обратном порядке?

Упражнение 4.29 Доказать что детерминант матрицы не изменится, если;

(1) из каждой строки, кроме последней вычесть все последующие строки,

(2) к каждому столбцу, начиная со второго, прибавить все предыдущие столбцы.

Упражнение 4.30 Даны матрицы $A,B \in \mathrm{M}_4(\mathbb{R})$. Известно, что $\det A = 1$ и

$$\begin{aligned}
\mathbf{c}_1(B) &= 3\mathbf{c}_3 - 2\mathbf{c}_4(A),\\
\mathbf{c}_2(B) &= 2\mathbf{c}_1(A) + 3\mathbf{c}_2(A),\\
\mathbf{c}_3(B) &= -2\mathbf{c}_1(A) + 2\mathbf{c}_3(A) + 2\mathbf{c}_4(A),\\
\mathbf{c}_4(B) &= 2\mathbf{c}_2 + 3\mathbf{c}_4(A),
\end{aligned}$$

где $\mathbf{c}_i(A), \mathbf{c}_j(A)$ обозначают i-й столбец матрицы A и j-й столбец матрицы B соответственно. Найдите $\det B$.

Упражнение 4.31 Пусть A,B — квадратные матрицы одинакового порядка, причём все строки в A одинаковы и все столбцы в B одинаковы. Чему может быть равен определитель матрицы $A + B$?