Математика для анализа данных, МФТИ 2023/2024
Линейная алгебра
Занятия #1-6

14 ноября 2023 г.

Содержание

1 Системы линейных уравнений

1.1 Несколько вводных примеров

Решение систем линейных уравнений (СЛАУ) является составной частью многих математических задач, встречающихся в этом курсе.

Пример. Найдите множество решений СЛАУ

$$\begin{cases} 2x + y = 3 \\ x + 3y = 4 \end{cases}$$

Решение. Выражая x из первого уравнения и подставляя полученное значение во второе, получаем

$$x = 4 - 3y$$
$$2(4 - 3y) + y = 3$$
$$8 - 6y + y = 3$$
$$y = 1, x = 1.$$

Фактически, при решении мы использовали эквивалентные преобразования линейных уравнений.

Пример. Решите СЛАУ

$$\begin{cases} x + 2y + 3z = 6 \\ 2x + y + z = 4 \\ x - y + z = 1 \end{cases}$$

Решение. Как и в прошлый раз, выражаем x и подставляем полученное выражение в 2-е и 3-е уравнения.

$$x = 6 - 2y - 3z,$$
$$\begin{cases} 2(6 - 2y - 3z) + y + z = 4 \\ 6 - 2y - 3z - y + z = 1 \end{cases}$$
$$\begin{cases} 12 - 3y - 5z = 4 \\ 6 - 3y - 2z = 1 \end{cases},$$
$$\begin{cases} 3y + 5z = 8 \\ 3y + 2z = 5 \end{cases}$$
$$3z = 3, z = 1, x = 1, y = 1.$$

Можно заметить, что проводимые операции затрагивают только коэффициенты перед x, y и z. Поэтому такое решение можно записать более компактно, используя матричную запись, в которой коэффициенты при

x, y и z, а также свободные члены уравнений записаны в таблицу

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

В общем случае, матрицы – прямоугольные таблицы значений; далее в тексте матрицы будут обозначаться заглавными латинскими буквами, например

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

где соответствующими маленькими буквами будут обозначаться элементы матрицы. В примере выше – квадратная матрица $n \times n$, у которой выделены элементы главной диагонали.

Аналогичное обозначение матрицы подразумевается записью $A = (a_{ij})$, где i и j – индексы, обозначающие позицию элемента a_{ij} в матрице. Элемент a_{ij} находится на пересечении i-й строки и j-го столбца таблицы.

1.2 Трансформация строк матриц

Далее будем использовать следующие типы трансформаций строк матриц, которые соответствуют действиям над компонентами системы линейных уравнений, и не приводят к изменениям множества решений, поскольку каждое из них приводит к созданию эквивалентной СЛАУ.

1. Строка j умножается на константу c и складывается с i-й строкой.

2. Строка i умножается на ненулевую константу c.

3. Строки i и j меняются местами.

1.3 Метод Гаусса

Метод Гаусса является основным алгоритмом решения СЛАУ. Идея метода заключается в обнулении коэффициентов под главной диагональю матрицы коэффициентов, приводя матрицу в т.н. верхнетреугольную форму.[1] После приведения матрицы коэффициентов к верхнетреугольному виду, корни СЛАУ найти очень легко.

Основные этапы метода Гаусса[2] проиллюстрированы следующими примерами.

Пример. Решите СЛАУ

$$\begin{cases} x + 2y + 3z = 6 \\ 2x + y + z = 4 \\ x - y + z = 1 \end{cases}$$

[1] Схожим образом определяются нижнетреугольные матрицы; квадратная матрица, являющаяся одновременно верхне- и нижнетреугольной, то есть имеющая ненулевые элементы только на главной диагонали, называется диагональной и обозначается $\operatorname{diag}(\lambda_1, \dots, \lambda_n)$.

[2] Точнее, его разновидности – метода Гаусса-Жордана

Решение.

$$\left(\begin{array}{rrr|r} 1 & 2 & 3 & 6 \\ 2 & 1 & 1 & 4 \\ 1 & -1 & 1 & 1 \end{array}\right) \xrightarrow{[2]\mapsto[2]-[1]\cdot 2} \left(\begin{array}{rrr|r} 1 & 2 & 3 & 6 \\ 0 & -3 & -5 & -8 \\ 1 & -1 & 1 & 1 \end{array}\right) \xrightarrow{[2]\mapsto[2]\cdot(-1)}$$

$$\left(\begin{array}{rrr|r} 1 & 2 & 3 & 6 \\ 0 & 3 & 5 & 8 \\ 1 & -1 & 1 & 1 \end{array}\right) \xrightarrow{[3]\mapsto[3]-[1]} \left(\begin{array}{rrr|r} 1 & 2 & 3 & 6 \\ 0 & 3 & 5 & 8 \\ 0 & -3 & -2 & -5 \end{array}\right) \xrightarrow{[3]\mapsto[3]+[2]}$$

$$\left(\begin{array}{rrr|r} 1 & 2 & 3 & 6 \\ 0 & 3 & 5 & 8 \\ 0 & 0 & 3 & 3 \end{array}\right) \xrightarrow{[3]\mapsto[3]/3} \left(\begin{array}{rrr|r} 1 & 2 & 3 & 6 \\ 0 & 3 & 5 & 8 \\ 0 & 0 & 1 & 1 \end{array}\right) \xrightarrow{[1]\mapsto[1]-[3]\cdot 3}$$

$$\left(\begin{array}{rrr|r} 1 & 2 & 0 & 3 \\ 0 & 3 & 5 & 8 \\ 0 & 0 & 1 & 1 \end{array}\right) \xrightarrow{[2]\mapsto[2]-[3]\cdot 5} \left(\begin{array}{rrr|r} 1 & 2 & 0 & 3 \\ 0 & 3 & 0 & 3 \\ 0 & 0 & 1 & 1 \end{array}\right) \xrightarrow{[2]\mapsto[2]/3}$$

$$\left(\begin{array}{rrr|r} 1 & 2 & 0 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right) \xrightarrow{[1]\mapsto[1]-[2]\cdot 2} \left(\begin{array}{rrr|r} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right).$$

Таким образом, алгоритм решения СЛАУ заключается в следующем:

- Глядя на первый столбец матрицы, при необходимости использовать перестановку строк (трансформацию 3-го типа) чтобы сделать элемент в верхней строке ненулевым; такой элемент называется ведущим или поворотным. Если все числа в первом стобце равны нулю, то аналогичную процедуру нужно осуществить со вторым столбцом, и т. д.

- Для каждой из строк снизу от вещущего элемента нужно использовать трансформацию 1-го типа, чтобы занулить все остальные коэффициенты в первом столбце.

- Переходим ко второму столбцу, при необходимости переставляя строки (кроме самой первой), чтобы сделать элемент на позиции $(2, 2)$ ненулевым. Этот элемент будет новым ведущим, и все элементы снизу от него нужно занулить с помощью трансформаций первого типа.

- Повторить процедуру до тех пор, пока матрица не будет приведена в верхнетреугольную форму.

Преобразованную матрицу легко использовать для нахождения решений СЛАУ, рассматривая уравнения, начиная с самого нижнего. При необходимости, также можно продолжить процедуру в обратной последовательности до приведения матрицы к диагональному виду (метод Гаусса-Жордана).

- Если все элементы на главной диагонали после преобразований оказались отличными от нуля, то с помощью преобразования второго типа можно сделать их равными 1. После этого можно использовать преобразования первого типа с подходящим коэффициентом λ, чтобы сделать все элементы сверху от ведущего элемента нулями.

- Аналогичным способом можно привести к нулю все элементы, кроме лежащих на главной диагонали (а их сделать равными единицам). Тогда решением системы уравнений будет являться вектор, получившийся в столбце свободных членов матрицы. Заметим, что всё вышеперечисленное верно только в том случае, когда в процессе преобразований мы не столкнулись с ситуацией, что какой-то из элементов на главной диагонали оказался равным нулю.

Следующий пример показывает, что ведущий элемент не всегда равен единице, хотя обычно удобно сделать его таким, чтобы не иметь дела с дробями в процессе вычислений.

Пример. Решите СЛАУ

$$\begin{cases} 2x+3y+z=3 \\ 3x-y+2z=5 \\ 2x+2y+z=3. \end{cases}$$

Решение.

$$\left(\begin{array}{ccc|c} 2 & 3 & 1 & 3 \\ 3 & -1 & 2 & 5 \\ 2 & 2 & 1 & 3 \end{array}\right) \xrightarrow{[2]\mapsto[2]-[1]\cdot\frac{3}{2}} \left(\begin{array}{ccc|c} 2 & 3 & 1 & 3 \\ 0 & -\frac{11}{2} & \frac{1}{2} & \frac{1}{2} \\ 2 & 2 & 1 & 3 \end{array}\right) \xrightarrow{[3]\mapsto[3]-[1]}$$

$$\left(\begin{array}{ccc|c} 2 & 3 & 1 & 3 \\ 0 & -\frac{11}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & -1 & 0 & 0 \end{array}\right) \xrightarrow{[2]\mapsto[3]} \left(\begin{array}{ccc|c} 2 & 3 & 1 & 3 \\ 0 & -1 & 0 & 0 \\ 0 & -\frac{11}{2} & \frac{1}{2} & \frac{1}{2} \end{array}\right) \xrightarrow{[2]\mapsto[2]\cdot(-1)}$$

$$\left(\begin{array}{ccc|c} 2 & 3 & 1 & 3 \\ 0 & 1 & 0 & 0 \\ 0 & -\frac{11}{2} & \frac{1}{2} & \frac{1}{2} \end{array}\right) \xrightarrow{[3]\mapsto[3]+[2]\cdot\frac{11}{2}} \left(\begin{array}{ccc|c} 2 & 3 & 1 & 3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{array}\right) \xrightarrow{[3\mapsto[3]\cdot 2}$$

$$\left(\begin{array}{ccc|c} 2 & 3 & 1 & 3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{array}\right) \xrightarrow{[1]\mapsto[1]-[3]} \left(\begin{array}{ccc|c} 2 & 3 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{array}\right) \xrightarrow{[1]\mapsto[1]-[2]\cdot 3}$$

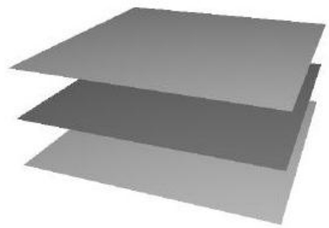
$$\left(\begin{array}{ccc|c} 2 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{array}\right) \xrightarrow{[1]\mapsto[1]/2} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{array}\right)$$

1.4 Геометрическая интерпретация СЛАУ

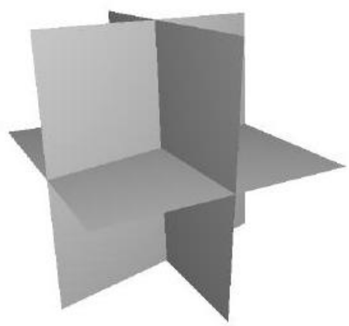
Геометрическая интерпретация решений СЛАУ проста и естественна. Рассмотрим систему из предыдущего примера. Уравнения $2x+y=3$ и $x+3y=4$ задают две прямые на числовой плоскости. Множество решений СЛАУ – это множество точек, удовлетворяющих обоим уравнениям, т.е. пересечение двух прямых.

Теоретически, возможны три разных случая взаимного расположения двух прямых: пересечение в одной точке (наиболее частый случай); параллельные прямые, или полное совпадение. Поэтому СЛАУ с двумя переменными может иметь одно, ноль, или бесконечно много решений. Так же выглядит множество решений СЛАУ в трёхмерном пространстве. Типичный случай – три плоскости, имеющие одну общую точку (две плоскости пересекаются по прямой линии, а она проходит через третью плоскость, имея одну общую точку с ней), но также возможны случаи бесконечного количества решенй (три плоскости имеют общую прямую линию, или совпадают полностью), и когда решений совсем нет (параллельные линии или плоскости). Некоторые из этих случаев показаны на рис.1.

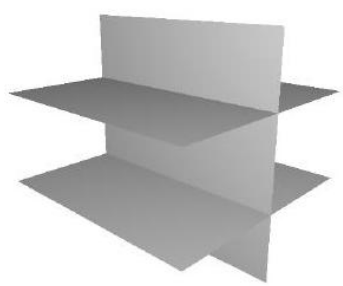
a) Бесконечно много решений



c) Нет решений



b) Ровно одно решение



d) Также нет решений

Рис. 1: Три плоскости в трёхмерном пространстве могут иметь ноль, одну или бесконечно много общих точек.

Повторяя ту же аргументацию для произвольного числа переменных или используя математическую индукцию, приходим к следующему соображению:

Теорема. Система линейных алгебраических уравнений может не иметь решений, иметь ровно одно решение, либо иметь бесконечно много решений.

Доказательство этого утверждения достаточно просто, однако для этого вначале нам понадобится ввести определение линейного (векторного) пространства.

В примерах ранее решение СЛАУ было единственным, поэтому в итоге матрицы коэффициентов приводились к диагональному виду. При решении вырожденных СЛАУ (т.е. не имеющих решения, либо имеющих бесконечно много решений) это не будет выполняться. Однако метод Гаусса можно эффективно применять и в этом случае.

Пример. Решите СЛАУ

$$\begin{cases} 2x+3y+z=3 \\ 2x+3y+2z=6 \end{cases}.$$

Решение.

$$\left(\begin{array}{ccc|c} 2 & 3 & 1 & 3 \\ 2 & 3 & 2 & 6 \end{array}\right) \xrightarrow{[2]\mapsto[2]-[1]} \left(\begin{array}{ccc|c} 2 & 3 & 1 & 3 \\ 0 & 0 & 1 & 3 \end{array}\right) \xrightarrow{[1]\mapsto[1]-[2]} \left(\begin{array}{ccc|c} 2 & 3 & 0 & 0 \\ 0 & 0 & 1 & 3 \end{array}\right)$$

Выполняется $x_3 = 3$ и $x_1 = -\frac{3}{2}x_2$. Здесь x_2 можно рассматривать, как параметр, и находить x_1 как функцию от x_2.

Заметим, что после второго шага матрица получилась верхнетреугольной, но на этот раз “шаг лестницы” оказался более длинным, чем в предыдущий примерах. Такой вид матрицы СЛАУ называется эшелонной формой.

1.5 Упражнения

Решите следующие СЛАУ

a) $\begin{cases} -3x_1+3x_3=-3 \\ -x_1+2x_2+2x_4=-8 \\ -2x_2+2x_3-2x_4=8 \\ 2x_1+4x_2-2x_3+2x_4=0 \end{cases}$ b) $\begin{cases} -5x_1-2x_2+4x_3+3x_4=-14 \\ -3x_1-4x_2-4x_3=10 \\ -5x_1+4x_2+4x_3-4x_4=-18 \\ -2x_2+3x_3+x_4=1 \end{cases}$

c) $\begin{cases} 4x_1-x_2+2x_3-3x_4-x_5=10 \\ -x_2+2x_3-3x_4+x_5=10 \\ -2x_1-5x_2+2x_3-4x_4+2x_5=14 \\ 3x_1+4x_2+2x_3-x_4+4x_5=-16 \\ x_1+4x_2-5x_3-x_4-4x_5=13 \end{cases}$ d) $\begin{cases} 4x_1+3x_2-4x_3+3x_4=-3 \\ -4x_1-3x_2+3x_4=21 \\ -x_1+2x_2+4x_3+3x_4=1 \\ -5x_1-4x_2-x_3+4x_4=28 \end{cases}$

e) $\begin{cases} 3x_1-5x_2+3x_3+4x_4=-27 \\ 4x_1-3x_2-2x_3-3x_4=0 \\ 4x_1-4x_2+3x_4=-18 \\ -x_1-3x_4=7 \end{cases}$ (f) $\begin{cases} -3x_1+3x_3-x_4=-3 \\ 3x_1-2x_2-2x_3+4x_4=20 \\ 4x_1+4x_2+2x_3+3x_4=25 \\ -4x_1+2x_2-4x_3-x_4=-39 \end{cases}$

g) $\begin{cases} x_1-2x_2+2x_3-5x_4=4 \\ 9x_1+2x_2-10x_3-6x_4=10 \\ 3x_1+5x_2-3x_3+3x_4=11 \\ 11x_1+5x_2+10x_4=0 \end{cases}$ (h) $\begin{cases} 3x_1+11x_2+7x_3-9x_4=-1 \\ 12x_1+10x_2-x_3+4x_4=-8 \\ 12x_1+3x_2-x_4=3 \\ 4x_1+2x_2+x_3+2x_4=-10 \end{cases}$

i) $\begin{cases} 5x_1+10x_2+11x_3+9x_4=-10 \\ 5x_1-4x_2+12x_3+6x_4=3 \\ 3x_1+5x_2-3x_3+3x_4=-6 \\ 6x_1+3x_2-5x_3+5x_4=-7 \end{cases}$ (j) $\begin{cases} x_1+12x_3-2x_4=-1 \\ 5x_1-8x_2+10x_3+x_4=8 \\ x_1-4x_2-4x_3+10x_4=-1 \\ 10x_1-8x_2-8x_3-x_4=-5 \end{cases}$

2 Линейное (векторное) пространство

Определение. Линейной пространство над полем вещественных чисел состоит из трёх объектов:

i) непустого множества V элементов произвольной природы, которые называются векторами;

ii) отображения $f : V \times V \to V$, сопоставляющего каждой паре элементов x, y множества V единственный элемент множества V, называемый их суммой и обозначаемый $x + y$.

iii) отображения $g : \mathbb{R} \times V \to V$, сопоставляющего каждому числу $\lambda \in \mathbb{R}$ и каждому элементу x множества V единственный элемент множества V, обозначаемый λx.

Отображение f называется сложением; отображение g называется умножением на скаляр. Отображение $f(x, y)$ обычно обозначается знаком "+", т.е. $f(x, y) = x + y$, в то время, как x и y могут отличаться от объектов, которые мы привыкли суммировать друг с другом (например, векторами могут быть студенты или преподаватели). Аналогично, отображение $g(\lambda, x)$ часто обозначается как $g(\lambda, x) = \lambda x$, хотя оно может не иметь ничего общего с обычным умножением. Элементы V называются векторами, хотя они могут отличаться от геометрических векторов. Множество V с заданными операциями должно подчиняться следующим аксиомам:

1. $\forall x, y \in V : x + y = y + x$;
2. $\forall x, y, z \in V : (x + y) + z = x + (y + z)$;
3. $\exists 0 \in V : \forall x \in V : x + 0 = 0 + x = x$;
4. $\forall x \in V \exists (-x) \in V : x + (-x) = (-x) + x = 0$;
5. $\forall x, y \in V, \forall \lambda \in \mathbb{R} : \lambda(x + y) = \lambda x + \lambda y$;
6. $\forall x \in V, \forall \lambda, \mu \in \mathbb{R} : \lambda(\mu x) = (\lambda\mu)x$.

Следующие примеры позволяют лучше разобраться, какие наборы являются векторными пространствами, а какие – нет.

1. Множество $\mathbb{R}$ образует векторное пространство по отношению к стандартным операциям сложения и умножения на число.
2. Множество $\mathbb{R}^2 = \{(x, y)\}$ образует векторное пространство по отношению к покомпонентному сложению и умножению на скаляры, т.е., $(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2)$ и $\lambda (x_1, x_2) = (\lambda x_1, \lambda x_2)$.
3. Аналогично, множество $\mathbb{R}^n$ образует векторное пространство по отношению к покомпонентному сложению и умножению на скаляры, т.е.

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

   и

   $$\lambda (x_1, \ldots, x_n) = (\lambda x_1, \ldots, \lambda x_n);$$

4. Множество $P_{[n]}$ полиномов степени n не является векторным пространством, поскольку $0 \notin P_{[n]}$.
5. Множество $P_{[\leq n]}$ полиномов, степень которых не превышает n образует векторное пространство.

6. Множество $C[a,b]$ всех непрерывных функций, определённых на отрезке $[a,b]$ образует векторное пространство (поскольку сумма непрерывных функций является непрерывной функцией).

7. Множество $C^n[a,b]$ всех функций, определённых на $[a,b]$, и имеющих n-ю непрерывную производную (также, n может принимать значение ∞) образует векторное пространство.

Пример. Проверим, что множество функций вида $f(x) = A\cos(x+\varphi)$, где $A, \varphi \in \mathbb{R}$, является векторным пространством по отношению к "обычному" сложению функций и умножению на константу.

Пусть $f(x) = a_1\cos(x+\varphi_1)$ и $g(x) = a_2\cos(x+\varphi_2)$. Применяя тригонометрические равенства, получаем

$$f(x) = a_1\cos(\varphi_1)\cdot\cos(x) - a_1\sin(\varphi_1)\cdot\sin(x),$$

$$g(x) = a_2\cos(\varphi_2)\cdot\cos(x) - a_2\sin(\varphi_2)\cdot\sin(x),$$

$$f(x)+g(x) = \underbrace{\big(a_1\cos(\varphi_1)+a_2\cos(\varphi_2)\big)}_{b_1}\cos(x) - \underbrace{\big(a_1\sin(\varphi_1)+a_2\sin(\varphi_2)\big)}_{b_2}\sin(x) =$$

$$= \sqrt{b_1^2+b_2^2}\left(\frac{b_1}{\sqrt{b_1^2+b_2^2}}\cos(x) - \frac{b_2}{\sqrt{b_1^2+b_2^2}}\sin(x)\right) = \sqrt{b_1^2+b_2^2}\cos\left(x+\arccos\frac{b_1}{\sqrt{b_1^2+b_2^2}}\right)$$

Поскольку $\sqrt{b_1^2+b_1^2}$ и $\arccos\frac{b_1}{\sqrt{b_1^2+b_2^2}}$ – вещественные числа для любых $a_1, a_2, \varphi_1, \varphi_2$ за исключением случая, когда $b_1 = 0$ и $b_2 = 0$ одновременно, получающаяся функция относится к классу функций $A\cos(x+\varphi)$. Остальные свойства векторного пространства проверяются тривиально.

В частном случае $b_1 = b_2 = 0, f(x)+g(x) \equiv 0$, то есть, результат тоже принадлежит к классу функций вида $A\cos(x+\phi)$, при этом $A = 0$.

Заметим, что в большинстве случаев операции над векторами являются логическим продолжением привычных операций сложения и умножения чисел, и вопрос, является ли множество V векторным пространством, зависит от того, замкнуто ли оно относительно таких операций, т.е. верно ли, что применение заданных операций к элементам V даёт в результате всегда элемент, также принадлежащий V. Это свойство объясняет следующее определение.

Определение. Пусть $(V, +, \cdot)$ – векторное пространство. Подмножество $U \subseteq V$ называется подпространством V, если $(U, +, \cdot)$ образует векторное пространство по отношению к операциям "$+$" и $\cdot$, унаследованным от V (с помощью ограничения операций на подмножество исходного множества определения). Другими словами, U является подпространством V если $x, y \in U \Rightarrow x+y \in U$ и $\lambda \in \mathbb{R}, x \in U \Rightarrow \lambda x \in U$.

Примеры векторных подпространств.

1. $\mathbb{R} \subset \mathbb{R}^2$ может рассматриваться как подпространство исходного пространства.

2. $P_{[\leq n]}$ является линейным подпространством $C(-\infty, +\infty)$.

3. $C^{n+1}[a,b]$ является линейным подпространством $C^n[a,b]$ – поскольку не каждой функции $f(x)$, такой, что $f^{(n)}(x)$ непрерывна, есть непрерывная $n+1$-я производная, но если известно, что $f^{(n+1)}(x)$ существует и непрерывна, то $f(x) \in C^n$.

4. Множество $B_2 = \big\{(x,y) \mid x^2+y^2 \leq 1\big\}$ не является линейным подпространством $\mathbb{R}^2$.

2.1 Линейная независимость

Понятия линейной зависимости и независимости являются ключевыми в линейной алгебре.

Определение. Пусть V – линейное пространство и пусть $x_1, x_2, \ldots, x_n$ – векторы. Также пусть $\lambda_1, \lambda_2, \ldots, \lambda_n$ – набор (вещественных) чисел. Выражение $\lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_n x_n$ называется линейной комбинацией векторов $x_1, x_2, \ldots, x_n$ с множителями (весами) $\lambda_1, \lambda_2, \ldots, \lambda_n$. Линейная комбинация называется тривиальной, если $\lambda_1 = \lambda_2 = \ldots = \lambda_n = 0$.

Определение. Пусть $x_1, x_2, \ldots, x_n$ – набор векторов в линейном пространстве V. Вектора $x_1, x_2, \ldots, x_n$ называются линейно зависимыми, если существует нетривиальный набор чисел $\lambda_1, \lambda_2, \ldots, \lambda_n$, такой, что $\lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_n x_n = 0$. Фактически, если набор векторов линейно зависим, это значить что какой-то из них (не всегда заранее понятно, какой именно) можно выразить в виде линейной комбинации остальных векторов.

Набор векторов, не являющийся линейно зависимым, называется линейно независимым.

Пример. Рассмотрим следующие вектора: $x_1 = (1, 2, 3), x_2 = (0, 1, 2), x_3 = (1, 4, 7)$. Нужно проверить, являются ли они линейно зависимыми. Предположим, что существуют числа $\lambda_1, \lambda_2, \lambda_3$, такие что $\lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3 = 0$. То есть, $(\lambda_1, 2\lambda_1, 3\lambda_1) + (0, \lambda_2, \lambda_2) + (\lambda_3, 4\lambda_3, 7\lambda_3) = (0, 0, 0)$. Такое возможно тогда и только тогда, когда

$$\begin{cases} \lambda_1 + \lambda_3 = 0 \\ 2\lambda_1 + \lambda_2 + 4\lambda_3 = 0 \\ 3\lambda_1 + 2\lambda_2 + 7\lambda_3 = 0 \end{cases}$$

Решая полученную СЛАУ методом Гаусса, получаем

$$\left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 2 & 1 & 4 & 0 \\ 3 & 2 & 7 & 0 \end{array}\right) \xrightarrow{[2] \mapsto [2]-[1]\cdot 2} \left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 3 & 2 & 7 & 0 \end{array}\right) \xrightarrow{[3] \mapsto [3]-[1]\cdot 3}$$

$$\left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 2 & 4 & 0 \end{array}\right) \xrightarrow{[3] \mapsto [3]\cdot \frac{1}{2}} \left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 1 & 2 & 0 \end{array}\right) \xrightarrow{[3] \mapsto [3]-[2]} \left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Видно, что третье уравнение преобразовалось к виду $0 = 0$, т.е. фактически, осталось лишь два уравнения. Это подразумевает, что СЛАУ имеет бесконечно много решений, среди которых найдётся и нетривиальное. Поэтому можно заключить, что набор x_1, x_2, x_3 является линейно зависимым.

Определение. Вектор x линейно выражается с помощью векторов $x_1, x_2, \ldots, x_n$ если $x = \lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_n x_n$ для каких-то $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}$. Другими словами, x линейно выражается с помощью $x_1, x_2, \ldots, x_n$, если x равен какой-то линейной комбинации этих векторов.

Теорема. Набор $x_1, x_2, \ldots, x_n$ является линейно зависимым тогда и только тогда, когда какой-то из векторов этого набора можно линейно выразить с помощью остальных векторов.

Доказательство. В одну сторону: предположим, что x_i можно линейно выразить с помощью $x_1, \ldots, x_{i-1}$, $x_{i+1}, \ldots, x_n$. То есть, $x_i = \lambda_1 x_1 + \ldots + \lambda_{i-1} x_{i-1} + \lambda_{i+1} x_{i+1} + \ldots + \lambda_n x_n$. Но это означает, что

$$\lambda_1 x_1 + \ldots + \lambda_{i-1} x_{i-1} - x_i + \lambda_{i+1} x_{i+1} + \ldots + \lambda_n x_n = 0,$$

т.е. существует нетривиальная (поскольку множитель при x_i равен -1) линейная комбинация векторов набора $x_1, \ldots, x_n$, равная нулю.

В обратную сторону: пусть $\lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_n x_n = 0$ – нетривиальная линейная комбинация векторов, в которой $\lambda_i \neq 0$ для некоторого i. Тогда $-\lambda_i x = \lambda_1 x_1 + \ldots + \lambda_{i-1} x_{i-1} + \lambda_{i+1} x_{i+1} + \ldots + \lambda_n x_n$, и поэтому

$$x_i = \frac{\lambda_1}{-\lambda_i} x_1 + \ldots + \frac{\lambda_{i-1}}{-\lambda_i} x_{i-1} + \frac{\lambda_{i+1}}{-\lambda_i} x_{i+1} + \ldots + \frac{\lambda_n}{-\lambda_i} x_n,$$

т.е. x_i линейно выражается с помощью $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$.

Несколько комментариев на тему свойств линейно зависимых наборов векторов.

1. Если набор векторов $x_1, x_2, \ldots, x_{n-1}$ является линейно зависимым, то набор $x_1, x_2, \ldots, x_{n-1}, x_n$ тоже является линейно зависимым.

2. Если набор векторов $x_1, x_2, \ldots, x_n$ содержит нулевой вектор, то этот набор является линейно зависимым.

3. Если набор векторов $x_1, x_2, \ldots, x_n$ является линейно независимым, то и набор $x_1, x_2, \ldots, x_{n-1}$ также является линейно независимым.

2.2 Линейная оболочка и базисы

Определение.

Пусть $x_1, x_2, \ldots, x_n$ – набор векторов в линейном пространстве V. Множество

$$\mathcal{L}(x_1, x_2, \ldots, x_n) = \{\lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_n x_n \mid \lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}\}$$

называется линейной оболочкой $x_1, x_2, \ldots, x_n$. Другими словами, линейная оболочка $x_1, x_2, \ldots, x_n$ – это набор всех возможных линейных комбинаций этих векторов.

Определение.

Набор векторов $x_1, x_2, \ldots, x_n$ в линейном пространстве V порождает V если $\mathcal{L}(x_1, x_2, \ldots, x_n) = V$. Другими словами, $x_1, x_2, \ldots, x_n$ порождает V, если каждый вектор из V можно выразить как линейную комбинацию $x_1, x_2, \ldots, x_n$.

Замечание. Понятие линейной независимости набора $x_1, \ldots, x_n$ определяется безотносительно того, в каком объемлющем пространстве V рассматриваются вектора. В отличие от этого, свойство набора векторов порождать всё объемлющее пространство безусловно зависит от того, что берется в качестве него. Например, вектор $(1, 1)$ порождает прямую на плоскости, проходящую через центр координат под углом $45°$, но не всю плоскость.

Определение. Набор $x_1, x_2, \ldots, x_n$ векторов линейного пространства V называется базисом V, если он порождает V и является линейно независимым.

Теорема. Пусть наборы $x_1, x_2, \ldots, x_n$ и $y_1, y_2, \ldots, y_m$ – два базиса линейного пространства V. Тогда $n = m$.

Доказательство. От противного, предположим, что $n \neq m$. Без потери общности можно считать, что $n > m$. Тогда каждый вектор $x_1, x_2, \ldots, x_n$ выражается линейно с помощью $y_1, y_2, \ldots, y_m$.

$$\begin{aligned} x_1 &= c_{11} y_1 + c_{12} y_2 + \ldots + c_{1m} y_m, \\ x_2 &= c_{21} y_1 + c_{22} y_2 + \ldots + c_{2m} y_m, \\ &\vdots \\ x_n &= c_{n1} y_1 + c_{n2} y_2 + \ldots + c_{nm} y_m. \end{aligned}$$

Покажем, что нетривиальная линейная комбинация $x_1, x_2, \ldots, x_n$ может равняться нуль-вектору. Пусть для некоторых $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ выполнено

$$\lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_n x_n = \lambda_1 (c_{11} y_1 + c_{12} y_2 + \ldots + c_{1m} y_m) + \lambda_2 (c_{21} y_1 + c_{22} y_2 + \ldots + c_{2m} y_m) +$$

$$+ \ldots + \lambda_n (c_{n1} y_1 + c_{n2} y_2 + \ldots + c_{nm} y_m) = (c_{11}\lambda_1 + c_{21}\lambda_2 + \ldots + c_{n1}\lambda_n)\, y_1 +$$

$$+ (c_{12}\lambda_1 + c_{22}\lambda_2 + \ldots + c_{n2}\lambda_n)\, y_2 + \ldots + (c_{1m}\lambda_1 + c_{2m}\lambda_2 + \ldots + c_{nm}\lambda_n)\, y_m = 0.$$

Поскольку $y_1, \ldots, y_m$ является линейно независимым набором векторов, то полученное условие эквивалентно СЛАУ

$$\begin{cases} c_{11}\lambda_1 + c_{21}\lambda_2 + \ldots + c_{n1}\lambda_n = 0 \\ c_{12}\lambda_1 + c_{22}\lambda_2 + \ldots + c_{n2}\lambda_n = 0 \\ c_{1m}\lambda_1 + c_{2m}\lambda_2 + \ldots + c_{nm}\lambda_n = 0 \end{cases}$$

которая имеет как минимум одно ненулевое решение, поскольку $m < n$ по нашему предположению. Из этого следует, что $x_1, x_2, \ldots, x_n$ не могут быть линейно независимыми, и поэтому $m = n$.

Следствие. Пусть $x = \lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_n x_n = \mu_1 x_1 + \mu_2 x_2 + \ldots + \mu_n x_n$ – два представления одного и того же вектора x в базисе $x_1, x_2, \ldots, x_n$. Тогда $\forall i : \lambda_i = \mu_i$. То есть, коэффициенты линейного разложения вектора x по базису являются уникальными. Эти коэффициенты называются координатами вектора x в данном базисе. Заметим, что порядок у коэффициентов $\lambda_1, \ldots, \lambda_n$ важен, и зависит от заданного порядка векторов в наборе $x_1, x_2, \ldots, x_n$.

Пример. Найдите координаты вектора $x = (1, 2, 3)$ в базисе $e_1 = (1, 1, 0), e_2 = (1, 0, 1)$, и $e_3 = (0, 1, 1)$.

Решение. Здесь и далее будем следовать правилу, что произвольный вектор по умолчанию является вектором-столбцом (и является частными случаем прямоугольной матрицы).

По определению, выполнено

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \lambda_1 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \lambda_3 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \lambda_1 + \lambda_2 \\ \lambda_1 + \lambda_3 \\ \lambda_2 + \lambda_3 \end{pmatrix}.$$

Получили следующую СЛАУ:

$$\begin{cases} \lambda_1 + \lambda_2 = 1 \\ \lambda_1 + \lambda_3 = 2 \\ \lambda_2 + \lambda_3 = 3 \end{cases}$$

которую можно решить методом Гаусса:

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

$$\left(\begin{array}{ccc|c} 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{array}\right).$$

Таким образом, координаты x в базисе e_1, e_2, e_3 – это $(0, 1, 2)$. Действительно, $x = 0 \cdot e_1 + 1.\ e_2 + 2 \cdot e_3$.

Пример. Найдите координаты функции $f(x) = \frac{1}{x^2 - 5x + 6}$, рассматриваемой в качестве вектора в соответствующем линейном пространстве с базисом $e_1 = \frac{1}{x-2}$ и $e_2 = \frac{1}{x-3}$.

Решение.

$$\frac{1}{x^2-5x+6} = \frac{\lambda_1}{x-2} + \frac{\lambda_2}{x-3} = \frac{\lambda_1(x-3)+\lambda_2(x-2)}{x^2-5x+6}.$$

Для того, чтобы это выполнялось, сумма свободных членов полинома в числителе правой части уравнения должна быть равна свободному члену полинома в числителе левой части, и аналогично для компоненты, линейной по x. Поэтому

$$\begin{cases} \lambda_1+\lambda_2=0 \\ -3\lambda_1-2\lambda_2=1 \end{cases}$$

Получаем

$$\left(\begin{array}{rr|r} 1 & 1 & 0 \\ -3 & -2 & 1 \end{array}\right) \to \left(\begin{array}{rr|r} 1 & 1 & 0 \\ 0 & 1 & 1 \end{array}\right) \to \left(\begin{array}{rr|r} 1 & 0 & -1 \\ 0 & 1 & 1 \end{array}\right)$$

Таким образом, координаты $f(x)$ в базисе e_1, e_2 – это $(-1, 1)$.

2.3 Фундаментальное множество решений СЛАУ

Понятие линейного пространства помогает сформулировать, как в наиболее общем случае (количество переменных и уравнений произвольно) устроено решение СЛАУ.

Рассмотрим СЛАУ с n переменными и m уравнениями:

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

которая в форме расширенной матрицы коэффициентов записывается следующим образом:

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

Определение. Система линейных алгебраических уравнений называется однородной, если $b_1 = b_2 = \ldots = b_n = 0$. В противном случае, СЛАУ называется неоднородной. Для сокращения, однородные и неоднородные СЛАУ будут обозначаться абревиатурами HSLE и NHSLE (Homogenous/Non-Homogenous System of Linear Equations).

Теорема. Множество решений однородной СЛАУ является линейным пространством.

Доказательство. Пусть $x = (x_1, \ldots, x_n)$ и $y = (y_1, \ldots, y_n)$ – два решения однородной СЛАУ вида (1) с $b_1 = b_2 = \ldots = b_m = 0$. Проверим, что $x + y$ также является решением:

$$a_{i1}(x_1+y_1) + \ldots + a_{in}(x_n+y_n) = (a_{i1}x_1 + \ldots + a_{in}x_n) + (a_{i1}y_1 + \ldots + a_{in}y_n) = 0 + 0 = 0$$

при любых i. Аналогично, $\lambda x = (\lambda x_1, \ldots, \lambda x_n)$ и $0 = (0, \ldots, 0)$ – также решения той же СЛАУ. Поэтому множество решений образует линейное пространство.

Определение. Фундаментальное множество решений СЛАУ, или общее решение СЛАУ – множество, включающее в себя все возможные решения заданной системы линейных алгебраических уравнений. В противоположность этому, частное решение СЛАУ – какое-то одно конкретное решение, которое может быть не единственным.

Теорема. Общее решение неоднородной СЛАУ можно выразить как общее решение соответствующей однородной СЛАУ плюс частное решение неоднородной.

Доказательство. Утверждение теоремы следует из того, что если $x = (x_1, \ldots, x_n)$ и $y = (y_1, \ldots, y_n)$ – два решения неоднородной СЛАУ, то $x - y$ является решением однородной СЛАУ. Действительно,

$$a_{i1}(x_1 - y_1) + \ldots + a_{in}(x_n - y_n) = (a_{i1}x_1 + \ldots + a_{in}x_n) -$$

$$-(a_{i1}y_1 + \ldots + a_{in}y_n) = b_i - b_i = 0$$

для всех i. Аналогично, если $x = (x_1, \ldots, x_n)$ является решением NHSLE и $y = (y_1, \ldots, y_n)$ – решение HSLE, то $x + y$ – решение NHSLE, потому что

$$a_{i1}(x_1 + y_1) + \ldots + a_{in}(x_n + y_n) = (a_{i1}x_1 + \ldots + a_{in}x_n) + (a_{i1}y_1 + \ldots + a_{in}y_n) = b_i + 0 = b_i.$$

2.4 Размерность линейного пространства

Понятие размерности играет важную роль во многих разделах математики. Вероятно, читатели сталкивались с обозначениями, что прямая линия имеет размерность 1, плоскость – размерность 2, пространство – размерность 3. Если добавить в трёхмерное пространство временную ось, то можно представить себе четырёхмерное пространство, но дальнейшие обобщения выглядят достаточно сложно. Однако, в каждом из этих случаев, количество измерений соответствует количеству независимых направлений, в которых можно перемещаться по нашему линейному пространству.

Определение. Размерность линейного пространства V – это количество векторов в его базисе. Обозначение: $\dim(V)$.

Размерность линейного пространства может быть конечной или бесконечной. В нашем курсе мы будем иметь дело только с конечномерными пространствами. Например, пространство полиномов степени, не превышающей 2, является трёхмерным, поскольку оно порождается функциями $\{1, x, x^2\}$. С другой стороны, пространство $C[a, b]$ является бесконечномерным, поскольку набор $\{1, x, \ldots, x^n\}$ является линейно независимым для любого $n \in \mathbb{N}$.

Теорема. Пусть U и V – конечномерные пространства, такие что $U \subset V$. Тогда $\dim(U) \leq \dim(V)$ и $\dim(U) = \dim(V)$ тогда и только тогда, когда $U = V$.

Доказательство. Используя предыдущую теорему, если $\dim(U) = \dim(V)$, то каждый базис $e_1, e_2, \ldots, e_n$ пространства U также является базисом пространства V. Тогда имеем $U \subset V$ и $V \subset U$, т.е. $U = V$.

Теперь рассмотрим случай, когда $U \subset V$ и $U \neq V$. Пусть $e_1, e_2, \ldots, e_n$ – базис пространства U. Поскольку $V \backslash U \neq \emptyset$, существует хотя бы один вектор $y \in V \backslash U$. Тогда $e_1, e_2, \ldots, e_n$ и y – линейно независимы, поскольку в противном случае либо $y \in U$, либо $e_1, e_2, \ldots, e_n$ являются зависимыми. Получили линейное пространство $U_1 = \mathcal{L}(e_1, e_2, \ldots, e_n, y)$. Аналогично, если $U_1 \neq V$, то существует вектор $z \in V$, не являющийся линейной комбинацией $e_1, e_2, \ldots, e_n$ и y, и тогда получаем $U_2 = \mathcal{L}(e_1, e_2, \ldots, e_n, y, z)$. Для некоторого $k \in \mathbb{N}$ мы получим $U_k = V$, поскольку $\dim(V) < \infty$ и возрастающая последовательность размерностей вложенных подпространств должна в какой-то момент остановиться. Поэтому базис V состоит из $e_1, e_2, \ldots, e_n$ и каких-то других векторов, а значит $\dim(U) < \dim(V)$.

Процедура, описанная в этой теореме, называется дополнением базиса. Из неё следует, что если одно линейное пространство содержится в другом (конечномерном), то любой базис меньшего пространства можно дополнить так, что он станет базисом большего пространства. Однако, в случае бесконечномерных пространств это выполняется не всегда; например, такая процедура не поможет дополнить базис пространства всевозможных полиномов, определённых на отрезке $[a, b]$ до базиса пространства всех функций на этом отрезке.

2.5 Упражнения

1. Определите, какие из заданных наборов векторов являются линейно независимыми. Если набор является линейно зависимым, выберите в нём базис и выразите остальные вектора с помощью выбранного базисы.

(a) $a_1 = \begin{pmatrix} 2 \\ 0 \\ -3 \\ 1 \end{pmatrix}, a_2 = \begin{pmatrix} -2 \\ 3 \\ 1 \\ -5 \end{pmatrix}, a_3 = \begin{pmatrix} -3 \\ -5 \\ 4 \\ -1 \end{pmatrix}, a_4 = \begin{pmatrix} 4 \\ 1 \\ 4 \\ 4 \end{pmatrix}$

(b) $a_1 = \begin{pmatrix} 3 \\ 0 \\ -4 \\ -1 \\ -1 \end{pmatrix}, a_2 = \begin{pmatrix} -3 \\ 1 \\ -5 \\ -3 \\ 3 \end{pmatrix}, a_3 = \begin{pmatrix} -1 \\ 1 \\ -4 \\ 1 \\ -1 \end{pmatrix}, a_4 = \begin{pmatrix} -2 \\ 0 \\ -5 \\ 3 \\ -4 \end{pmatrix}$

(c) $a_1 = \begin{pmatrix} -3 \\ -4 \\ 3 \end{pmatrix}, a_2 = \begin{pmatrix} -1 \\ -3 \\ 4 \end{pmatrix}, a_3 = \begin{pmatrix} -1 \\ 3 \\ -3 \end{pmatrix}, a_4 = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$

(d) $a_1 = \begin{pmatrix} 4 \\ 1 \\ 2 \\ -4 \end{pmatrix}, a_2 = \begin{pmatrix} -2 \\ -1 \\ -4 \\ 3 \end{pmatrix}, a_3 = \begin{pmatrix} 3 \\ 4 \\ 2 \\ 2 \end{pmatrix}, a_4 = \begin{pmatrix} -5 \\ 2 \\ 2 \\ -3 \end{pmatrix}, a_5 = \begin{pmatrix} -5 \\ 4 \\ 2 \\ 1 \end{pmatrix}.$

2. Найдите координаты вектора x в соответствующих базисах.

(a) $e_1 = \begin{pmatrix} 1 \\ 2 \\ 7 \end{pmatrix}, e_2 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, e_3 = \begin{pmatrix} -1 \\ -2 \\ -8 \end{pmatrix}, x = \begin{pmatrix} -3 \\ -5 \\ 1 \end{pmatrix}$

(b) $e_1 = \begin{pmatrix} 6 \\ 3 \\ 4 \end{pmatrix}, e_2 = \begin{pmatrix} 9 \\ 4 \\ 8 \end{pmatrix}, e_3 = \begin{pmatrix} -4 \\ -2 \\ -3 \end{pmatrix}, x = \begin{pmatrix} 0 \\ 1 \\ -4 \end{pmatrix}$

(c) $x = \begin{pmatrix} 3 \\ 4 \\ 6 \\ 5 \end{pmatrix}, e_1 = \begin{pmatrix} 5 \\ 1 \\ 2 \\ 2 \end{pmatrix}, e_2 = \begin{pmatrix} 1 \\ -4 \\ -5 \\ 0 \end{pmatrix}, e_3 = \begin{pmatrix} 4 \\ 2 \\ -1 \\ -1 \end{pmatrix}, e_4 = \begin{pmatrix} -7 \\ -1 \\ -6 \\ -2 \end{pmatrix}$

(d) $x = \begin{pmatrix} -6 \\ 2 \\ 1 \\ -1 \end{pmatrix}, e_1 = \begin{pmatrix} 3 \\ 1 \\ 4 \\ 4 \end{pmatrix}, e_2 = \begin{pmatrix} 3 \\ 1 \\ 1 \\ -3 \end{pmatrix}, e_3 = \begin{pmatrix} 7 \\ -2 \\ 0 \\ -5 \end{pmatrix}, e_4 = \begin{pmatrix} 4 \\ -1 \\ 2 \\ -2 \end{pmatrix}$

(e) $x = \begin{pmatrix} 1 \\ 2 \\ -4 \\ 3 \end{pmatrix}, e_1 = \begin{pmatrix} 3 \\ 5 \\ 2 \\ 7 \end{pmatrix}, e_2 = \begin{pmatrix} -5 \\ -6 \\ 2 \\ -7 \end{pmatrix}, e_3 = \begin{pmatrix} 3 \\ 4 \\ 5 \\ 8 \end{pmatrix}, e_4 = \begin{pmatrix} 1 \\ 1 \\ -7 \\ -3 \end{pmatrix}$

2.6 Ранг матрицы

Вернемся к теореме о том, что общее решение HSLE является линейным пространством – чтобы понять, как связана размерность этого пространства с размерностью пространства, порожденного матрицей коэффициентов.

Определение. Пусть A – $m \times n$ матрица. Максимальное число линейно независимых строк этой матрицы называется её рангом и обозначается $\mathrm{rk}(A)$.

Используя метод Гаусса, можно легко показать, что ранг матрицы равен числу ненулевых строк в матрице после приведения её в эшелонную форму.

Определение. Пусть A – $m \times n$ матрица. Транспонированная матрица A обозначается A^T, имеет размерность $n \times m$ и получается путём записи строк матрицы A в качестве столбцов A^T. То есть, если

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

Оказывается, что максимальное число линейно-независимых строк и линейно-независимых столбцов одинаково. Таким образом, $\mathrm{rk}(A) = \mathrm{rk}\left(A^T\right)$.

Рассмотрим набор векторов $a_1 = (a_{11}, a_{12}, \dots, a_{1n}), a_2 = (a_{21}, a_{22}, \dots, a_{2n}),\ a_m = (a_{m1}, a_{m2}, \dots, a_{mn})$ в $\mathbb{R}^n$, образованных строками матрицы A. Тогда $\mathrm{rk}(A) = \dim \mathcal{L}(a_1, a_2, \dots, a_m)$. Связь между рангом матрицы коэффициентов HSLE и размерностью пространства решений описывается следующей теоремой.

Теорема. Пусть V – линейное пространство решений HSLE, задаваемой матрицей A. Тогда $\dim(V) = n - \mathrm{rk}(A)$.

Прежде чем доказать эту теорему, рассмотрим следующий пример.

Пример. Найдите фундаментальное множество решений СЛАУ

$$\begin{cases} x_1 + 2x_2 + x_4 = 0 \\ x_1 + x_2 - x_3 = 0 \end{cases}$$

Решение. Используем метод Гаусса

$$\left(\begin{array}{cccc|c} 1 & 2 & 0 & 1 & 0 \\ 1 & 1 & -1 & 0 & 0 \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & 1 & -1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \end{array}\right) \to \left(\begin{array}{cccc|c} \underline{1} & 0 & -2 & -1 & 0 \\ 0 & \underline{1} & 1 & 1 & 0 \end{array}\right)$$

и получаем СЛАУ вида

$$\begin{cases} x_1 - 2x_3 - x_4 = 0 \\ x_2 + x_3 + x_4 = 0 \end{cases}$$

Видно, что переменные x_3 и x_4 являются “свободными”, то есть они могут принимать любые значения без каких-либо ограничений, тогда как x_1 и x_2 могут быть выражены с помощью переменных x_3 и x_4 (это связано с тем, что коэффициенты перед этими переменными, подчёркнутые в матрице, оказались ненулевыми). Поэтому можно сделать следующее: вначале используем второе уравнение, чтобы выразить $x_2 = -x_3 - x_4$, а затем решаем первое уравнение относительно x_1. Получается $x_1 = 2x_3 + x_4$. Теперь полученное решение можно представить в следующем виде:

$$\begin{pmatrix} 2x_3 + x_4 \\ -x_3 - x_4 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 2x_3 \\ -x_3 \\ x_3 \\ 0 \end{pmatrix} + \begin{pmatrix} x_4 \\ -x_4 \\ 0 \\ x_4 \end{pmatrix} = x_3 \begin{pmatrix} 2 \\ -1 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix}. \tag{2}$$

Поскольку свободные переменные x_3 и x_4 могут принимать любые значения, то фундаментальное множество решений СЛАУ, представленное в (2), можно переписать в виде

$$C_1 \begin{pmatrix} 2 \\ -1 \\ 1 \\ 0 \end{pmatrix} + C_2 \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix},$$

где C_1, C_2 – произвольные константы.

Теорема утверждает, что $\dim(V) = n - \mathrm{rk}(A) = 4 - 2 = 2$. С другой стороны, два найденных вектора $e_1 = (2, -1, 1, 0)$ и $e_2 = (1, -1, 0, 1)$ являются линейно независимыми, поэтому множество решений является линейной оболочкой e_1 и e_2, т.е., любое решение исходной системы можно представить в виде линейной комбинации e_1 и e_2,

$$x = \lambda_1 \begin{pmatrix} 2 \\ -1 \\ 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix}$$

Доказательство теоремы. Во-первых, элементарные преобразования строк не влияют на множество решений, потому что они приводят к эквивалентным СЛАУ. Достаточно рассмотреть недоопределённую систему (число уравнений m меньше числа переменных n), поскольку утверждение теоремы тривиально для определённых и переопределённых ($m > n$) систем. Фактически, идея доказательства аналогична предыдущему примеру, где мы использовали метод Гаусса для преобразования матрицы коэффициентов A к форме

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2m} & \dots & a_{2n} \\ a_{31} & a_{32} & \dots & a_{3m} & \dots & a_{3n} \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mm} & \dots & a_{mn} \end{pmatrix} \to \begin{pmatrix} \cdot & \cdot & \dots & \cdot & \cdot & \dots & \cdot \\ 0 & \cdot & \dots & \cdot & \cdot & \dots & \cdot \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \cdot & \cdot & \dots & \cdot \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix}$$

меняя порядок уравнений при необходимости. Тогда размерность пространства решений точно равна числу свободных переменных, которое, в свою очередь, равно числу переменных минус число уравнений в трансформированной матрице. Число переменных – n, число уравнений в трансформированной матрице равно её рангу, поэтому $\dim(V) = n - \mathrm{rk}(A)$.

Пример. Аналогичный приём можно использовать для нахождения общего решения NHSLE.

$$\begin{cases} x_1 + 2x_2 + x_3 - x_4 + x_5 = 4 \\ x_1 + x_2 - x_3 + 2x_4 - x_5 = 2 \end{cases}$$

Решение.

$$\left(\begin{array}{rrrrr|r} 1 & 2 & 1 & -1 & 1 & 4 \\ 1 & 1 & -1 & 2 & -1 & 2 \end{array}\right) \to \left(\begin{array}{rrrrr|r} 1 & 2 & 1 & -1 & 1 & 4 \\ 0 & 1 & 2 & -3 & 2 & 2 \end{array}\right)$$

В полученной матрице эшелонной формы поворотными элементам соответствуют x_1, x_2. Соответственно, x_3, x_4, x_5 – свободные переменные. Выразим x_2 из второго уравнения через свободные переменные, после этого сделаем то же самое с x_1:

$$x_2 = 2 - 2x_3 + 3x_4 - 2x_5, \quad x_1 + 2(2 - 2x_3 + 3x_4 - 2x_5) + x_3 - x_4 + x_5 = 4 \Rightarrow x_1 = 3x_3 - 5x_4 + 3x_5.$$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 3x_3 - 5x_4 + 3x_5 \\ 2 - 2x_3 + 3x_4 - 2x_5 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = C_1 \begin{pmatrix} 3 \\ -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + C_2 \begin{pmatrix} -5 \\ 3 \\ 0 \\ 1 \\ 0 \end{pmatrix} + C_3 \begin{pmatrix} 3 \\ -2 \\ 0 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 2 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

2.7 Упражнения

1. Найдите ранг следующих матриц

a) $\begin{pmatrix} 4 & 8 & -4 & 16 & -8 \\ -1 & -2 & 1 & -4 & 2 \\ 3 & 6 & -3 & 12 & -6 \\ 2 & 4 & -2 & 8 & -4 \end{pmatrix}$ b) $\begin{pmatrix} -2 & -10 & -10 & 2 & 4 \\ -2 & -10 & -10 & 2 & 4 \\ 1 & 5 & 5 & -1 & -2 \\ 2 & 10 & 10 & -2 & -4 \end{pmatrix}$ c) $\begin{pmatrix} 9 & 10 & -16 & 11 & 3 \\ 9 & 8 & -8 & 7 & -1 \\ -22 & -18 & -20 & 15 & 11 \\ 4 & -3 & -1 & 12 & -10 \end{pmatrix}$

d) $\begin{pmatrix} 12 & -2 & -3 & 12 & -2 \\ 12 & 8 & 3 & 12 & -4 \\ -20 & 0 & 3 & -20 & 4 \\ 12 & -12 & -9 & 12 & 0 \end{pmatrix}$ e) $\begin{pmatrix} 4 & -8 & -4 & -16 & -12 \\ -4 & 8 & 4 & 16 & 12 \\ -3 & 6 & 3 & 12 & 9 \\ 3 & -6 & -3 & -12 & -9 \end{pmatrix}$ f) $\begin{pmatrix} 0 & 1 & -3 & -1 & -4 \\ -16 & 16 & 4 & 0 & -8 \\ 16 & -13 & -13 & -3 & -4 \\ -8 & 6 & 8 & 2 & 4 \end{pmatrix}$

g) $\begin{pmatrix} 14 & -14 & 4 & 22 & 14 \\ -8 & -16 & 20 & -4 & -8 \\ 0 & -14 & 13 & 5 & 0 \\ 6 & -16 & 11 & 13 & 6 \end{pmatrix}$ h) $\begin{pmatrix} -3 & -3 & 9 & -15 & -9 \\ 3 & 3 & -9 & 15 & 9 \\ 5 & 5 & -15 & 25 & 15 \\ -2 & -2 & 6 & -10 & -6 \end{pmatrix}$ i) $\begin{pmatrix} 5 & -2 & 4 & -4 & -4 \\ 12 & -9 & 16 & -8 & -14 \\ -1 & -8 & 12 & 4 & -8 \\ -9 & 12 & -20 & 4 & 16 \end{pmatrix}$

j) $\begin{pmatrix} 10 & -14 & -2 & -6 & -4 \\ -19 & 11 & 5 & 21 & 4 \\ -15 & 8 & 4 & 17 & 3 \\ 13 & -13 & -3 & -11 & -4 \end{pmatrix}$

2. Найдите фундаментальное множество решений у следующих СЛАУ

a) $\begin{cases} x_1 - x_2 + 5x_3 - 13x_4 = 3 \\ x_1 + x_2 - x_3 + x_4 = -3 \end{cases}$ b) $\begin{cases} 3x_1 - 5x_2 - 3x_3 + 13x_4 = 10 \\ 2x_1 - x_2 + 5x_3 + 4x_4 = -5 \end{cases}$

c) $\begin{cases} 2x_1 - x_2 - 2x_3 - 5x_4 = -7 \\ 2x_1 + x_2 - 6x_3 + x_4 = -1 \end{cases}$ d) $\begin{cases} 3x_1 + x_2 + x_3 + 6x_5 = -6 \\ 2x_1 + 3x_2 + 2x_3 - 4x_4 + 5x_5 = -4 \\ 3x_1 + 3x_2 + 5x_3 + 8x_4 - 16x_5 = -6 \end{cases}$

e) $\begin{cases} -2x_2 + x_3 - 7x_4 - 4x_5 = 7 \\ x_1 + 2x_2 - x_3 + x_4 - 2x_5 = -7 \\ 3x_1 + 4x_2 - 2x_3 - 4x_4 - 10x_5 = -14 \end{cases}$ f) $\begin{cases} x_1 - 6x_2 - 3x_3 - 17x_4 - 17x_5 = -1 \\ 5x_2 + 7x_3 - 8x_4 - 17x_5 = -5 \\ 3x_2 + 5x_3 - 8x_4 - 15x_5 = -3 \end{cases}$

g) $\begin{cases} 2x_1 + x_2 + 9x_3 - 15x_4 = 9 \\ x_1 + 2x_3 - 6x_4 + x_5 = 1 \end{cases}$ h) $\begin{cases} x_1 + x_2 + 2x_3 + 2x_4 - 2x_5 = 0 \\ x_2 + 4x_3 + 7x_4 - 5x_5 = -2 \end{cases}$

i) $\begin{cases} 3x_1 - 5x_2 + 18x_3 + x_4 - 3x_5 = 2 \\ x_1 + x_2 - 2x_3 + 11x_4 - x_5 = 6 \end{cases}$ j) $\begin{cases} x_1 - x_2 - 3x_3 + 7x_5 = 0 \\ x_1 + 2x_3 - x_4 + 7x_5 = 1 \end{cases}$

3 Опеределитель и его свойства

Определитель квадратной матрицы A – это число, получаемое как функция от элементов матрицы, и имеющее некоторый геометрический смысл, который мы обсудим в этом разделе.

3.1 Матрицы 2×2 и 3×3

Определитель $\det(A)$ матрицы размером 2×2 задаётся следующим образом:

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

Векторы (a_{11}, a_{12}) и (a_{21}, a_{22}) образуют параллелограм на плоскости $x - y$ с вершинами в точках

$$(0,0),\ (a_{11}, a_{12}),\ (a_{21}, a_{22},\ (a_{11} + a_{21}, a_{21} + a_{22}).$$

Такой параллелограм называется натянутым на вектора (a_{11}, a_{12}) и (a_{21}, a_{22}).

Абсолютное значение выражения $a_{11}a_{22} - a_{12}a_{21}$ является площадью параллелограмма. Таким образом, определитель 2×2 матрицы является ориентированной площадью параллелограмма, натянутого на вектора-столбцы матрицы. Ориентированная площадь по модулю совпадает с обычной площадью фигуры, но меняет знак, когда два вектора меняются местами.

Определитель матрицы $\det(A)$ размером 3×3

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

вычисляется следующим образом:

$$\det(A) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{21}a_{32}a_{13} - a_{31}a_{22}a_{13} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}.$$

Аналогично случаю с 2×2 матрицей, геометрическая интерпретация определителя – это ориентированный объем параллелепипеда, натянутого на вектора-столбцы матрицы $(a_{11}, a_{12}, a_{13})\,, (a_{21}, a_{22}, a_{23})$, и (a_{31}, a_{32}, a_{33}). Абсолютное значение определителя равно объему заданной фигуры.

Аналитическое выражение для определителя становится более сложным с ростом размерности матрицы (выражение для матрицы $n\times n$ содержит $n!$ слагаемых). В следующем разделе мы зададим формальное определение $\det(A)$ для квадратных матриц произвольного размера, которое для 2×2 и 3×3 матриц совпадает с рассмотренными ранее случаями. Пока же просто перечислим свойства определителя, которые уникальным образом задают его как функцию от матрицы; при этом геометрическая интерпретация определителя матрицы размера $n \times n$ – ориентированный n-мерный объём n-мерного параллелепипеда, натянутого на вектора-столбцы. Часто определитель матрицы обозначают с помощью вертикальных “стенок” у матрицы

Пример. $\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = 1 \cdot 4 - 2 \cdot 3 = 4 - 6 = -2.$

Пример. $\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = 1\cdot 5\cdot 9+4\cdot 8\cdot 3+2\cdot 6\cdot 7-7\cdot 5\cdot 3-4\cdot 2\cdot 9-1\cdot 8\cdot 6 = 45+96+84-105-72-48 = 0.$

3.2 Свойства определителя

Напомним три вида операций со строками матрицы, которые мы использовали для решения СЛАУ методом Гаусса. Для преобразований первого типа j-я строка умножается на произвольную константу c и складывается с i-й. Для преобразований II типа i-я строка умножается на ненулевую константу c. Преобразование III типа – это перестановка i-й и j-й строк матрицы местами. Свойства определителя матрицы также связаны с преобразованиями таких типов.

1. $\det(A)$ не изменяется при преобразованиях I-го типа, причем как со строками, так и со столбцами
2. $\det(A)$ умножается на c при преобразованиях II-го типа, т.е., если строка или столбец матрицы умножаются на число c, $\det(A)$ также увеличивается в c раз.
3. $\det(A)$ умножается на -1 при преобразованиях III-го типа, т.е., когда строки или столбцы матрицы меняются местами.

Эти три свойства выполнены для значения площади параллелограмма и объема параллелепипеда, натянутых на вектора-столбцы или вектора-строки матрицы A. Они также выполняются для аналогичных геометрических объектов в $\mathbb{R}^n$ и они уникальным образом задают функцию $\det(A)$ с точностью до множителя. Если дополнительно потребовать $\det(I) = 1$, где

$$I = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

– единичная матрица размером $n \times n$, то свойства 1-3 можно использовать в качестве определения детерминанта (определителя) матрицы.

3.3 Формальное определение

Перестановка $(i_1, i_2, \dots, i_n)$ чисел $(1, 2, \dots, n)$ – это отображение $f : (1, 2, \dots, n) \to (1, 2, \dots, n)$, такое что никакая пара чисел i и j не отображаются в одно и то же значение, т.е. из $f(i) = f(j)$ следует $i = j$. Это означает, что $i_1, i_2, \dots, i_n$ – это тот же набор чисел, что и $1, 2, \dots, n$, но записанный в другом порядка. Например, перестановка $(2, 3, 1, 4)$ чисел $(1, 2, 3, 4)$ – это отображение f, такое, что $f(1) = 2, f(2) = 3, f(3) = 1$, и $f(4) = 4$.

Любая перестановка $(i_1, i_2, \dots, i_n)$ допускает разложение в произведение циклов, а далее – в произведение транспозиций, т.е. перестановок между парами элементов.

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

где первое представление (в виде цикла) интерпретируется так: “1 переходит в 2, 2 в 3, 3 в 1; кроме того, 4 переходит в 4”, а второе (в виде произведения транспозиций) – “1-й элемент меняем со вторым, потом новый 1-й элемент с 3-м, потом новый 1-й элемент с 4-м, потом 4-й элемент с новым 1-м”.

Таким образом, любая перестановка может быть разложена в произведение транспозиций. Хотя это можно сделать многими способами, чётность числа транспозиций во всех таких разложениях одинакова. Это позволяет определить знак перестановки (также называемый чётностью или сигнатурой перестановки) σ как:

$$\varepsilon_\sigma = (-1)^t,$$

где t – число транспозиций в каком-то разложении перестановки σ. При этом σ называют чётной перестановкой, если $\varepsilon_\pi = 1$, и нечётной перестановкой, если $\varepsilon_\pi = -1$. Эквивалентно, знак перестановки определяется её цикловой структурой: знак перестановки σ из n элементов, состоящий из k циклов, равен

$$\varepsilon_\sigma = (-1)^{n-k}$$

Знак перестановки σ также может быть определён через число инверсий $N(\sigma)$ в σ :

$$\varepsilon_\sigma = (-1)^{N(\sigma)}.$$

Инверсия в перестановке σ – это число таких пар $\{i, j\}$, что $i < j$, но $\sigma(i) > \sigma(j)$. Чётность перестановки $\sigma-$ это чётность числа всех её инверсий. Например, для перестановки

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

все инверсии – это $\{1,4\}, \{1,5\}, \{2,3\}, \{2,4\}, \{2,5\}, \{3,4\}, \{3,5\}$. Всего 7 штук, поэтому такая перестановка нечётная.

Пример

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 9 & 8 & 6 & 1 & 4 & 7 & 5 & 2 \end{pmatrix} = (1385)(29)(46) = (1,3)(1,8)(1,5)(2,9)(4,6)$$

Определитель произвольной $n \times n$ матрицы задаётся следующим образом:

$$\det \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} = \sum_{(i_1, i_2, \dots, i_n)} \sigma\left(i_1, i_2, \dots, i_n\right) a_{1i_1} a_{2i_2} \dots a_{ni_n},$$

где суммирование происходит по всем возможным перестановкам множества $(1, 2, \dots, n)$. Таких перестановок ровно $n! = 1 \cdot 2 \dots n$, и поэтому выражение для определителя состоит из $n!$ слагаемых. Также можно сказать, что определитель матрицы $A = (a_{ij})$ – это сумма произведений элементов матрицы, таких, что в каждом слагаемом берется ровно по одному элементу из каждой строки и столбца матрицы.

Теорема. Определитель верхнетреугольной квадратной матрицы равен произведению диагональных элементов.

Доказательство. Пусть A – верхнетреугольная матрица вида

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

В произведениях элементов матрицы, таких, что ровно по одному элементу берется из каждой строки и столбца всегда будет присутствовать элемент снизу от диагонали матрицы, за исключением случая, когда произведение полностью состоит из элементов главной диагонали. Поэтому $\det(A) = \prod_{i=1}^{n} a_{ii}$.

3.4 Теорема Лапласа

Минор $M_{i,j}$ матрицы

$$A = \begin{pmatrix} a_{11} & \dots & a_{1j-1} & a_{1j} & a_{1j+1} & \dots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & & \vdots \\ a_{i-11} & \dots & a_{i-1j-1} & a_{i-1j} & a_{i-1j+1} & \dots & a_{i-1n} \\ a_{i1} & \dots & a_{ij-1} & a_{ij} & a_{ij+1} & \dots & a_{in} \\ a_{i+11} & \dots & a_{i+1j-1} & a_{i+1j} & a_{i+1j+1} & \dots & a_{i+1n} \\ \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nj-1} & a_{nj} & a_{nj+1} & \dots & a_{nn} \end{pmatrix}$$

– это определитель матрицы размером $(n-1)\times(n-1)$, которая получается из A при удалении i-й строки и j-го столбца:

$$M_{ij} = \det \begin{pmatrix} a_{11} & \dots & a_{1j-1} & a_{1j+1} & \dots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ a_{i-11} & \dots & a_{i-1j-1} & a_{i-1j+1} & \dots & a_{i-1n} \\ a_{i+11} & \dots & a_{i+1j-1} & a_{i+1j+1} & \dots & a_{i+1n} \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nj-1} & a_{nj+1} & \dots & a_{nn} \end{pmatrix}$$

Минор M_{ij} также называют алгебраическим дополнением элемента a_{ij}. Определитель A можно найти с использованием разложения Лапласа.

Теорема (без доказательства).

$$\det(A) = \textstyle\sum_{i=1}^{n}(-1)^{i+j}a_{ij}M_{ij} \text{ (разложение по } j\text{-му столбцу)},$$
$$\det(A) = \textstyle\sum_{j=1}^{n}(-1)^{i+j}a_{ij}M_{ij} \text{ (разложение по } i\text{-й строке)}.$$

Теорема Лапласа предлагает удобный способ для вычисления определителей, при котором определитель матрицы $n \times n$ выражается рекурсивно с помощью определителей меньшего порядка.

Пример (разложение определителя по второму столбцу).

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

$$+(-1) \cdot (-1)^{3+2} \begin{vmatrix} 1 & 1 & 2 \\ 2 & 1 & 1 \\ 2 & 1 & 2 \end{vmatrix} + 0 \cdot (-1)^{4+2} \begin{vmatrix} 1 & 1 & 2 \\ 2 & 1 & 1 \\ -1 & 3 & 1 \end{vmatrix} = 2 + 2 + 4 - 4 - 4 - 1 = -1.$$

Этот пример показывает, что формула Лапласа наиболее удобна для вычисления определителя, если строка или столбец матрицы содержат много нулей. Если в исходной матрице нет таких столбцов или строк, то их можно получить при помощи трансформаций 1-го типа, так как они не меняют значения определителя.

Пример.

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

$$+0 \times (-1)^{2+3} \begin{vmatrix} 1 & 1 & -1 \\ 3 & -1 & 1 \\ 5 & 2 & -1 \end{vmatrix} + 3 \times (-1)^{2+4} \begin{vmatrix} 1 & 1 & 1 \\ 3 & -1 & 1 \\ 5 & 2 & 1 \end{vmatrix} = -2 + 30 = 28.$$

Пример. Найдите интеграл Пуассона $\int_{-\infty}^{\infty} e^{-x^2}\, dx$.

Решение. Обозначим $I = \int_{-\infty}^{\infty} e^{-x^2}\, dx$. Наиболее простой способ сделать это – перейти к I^2 и использовать полярную систему координат:

$$I^2 = \left(\int_{-\infty}^{\infty} e^{-x^2}\, dx \right)^2 = \int_{-\infty}^{\infty} e^{-x^2}\, dx \cdot \int_{-\infty}^{\infty} e^{-y^2}\, dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-x^2-y^2}\, dxdy. \tag{3}$$

Делаем замену $x = r\cos\phi$, $y = r\sin\phi$. Для преобразования элементов площади $dxdy$ в $drd\phi$ запишем координаты векторов $\mathbf{dx} = (dx, 0)$ и $\mathbf{dy} = (0, dy)$ в базисе $(dr,\ d\phi)$. Поскольку $dx = \cos\phi\, dr - r\sin\phi\, d\phi$, а $dy = \sin\phi\, dr + r\cos\phi\, d\phi$, то

$$\mathbf{dx} = \begin{pmatrix} \cos\phi \\ -r\sin\phi \end{pmatrix}, \quad \mathbf{dy} = \begin{pmatrix} \sin\phi \\ r\cos\phi \end{pmatrix}.$$

Площадь параллелограмма, натянутого на вектора $\mathbf{dx}$, $\mathbf{dy}$, равна детерминанту матрицы, составленной из них:

$$\begin{vmatrix} \cos\phi & \sin\phi \\ -r\sin\phi & r\cos\phi \end{vmatrix} = r \quad \text{(Якобиан перехода к полярной системе координат).}$$

Таким образом, элемент площади $dxdy$ в декартовой (прямоугольной) системе координат соответствует $rdrd\phi$ в полярной системе координат. Подставим полученный результат в (3):

$$I^2 = \int_0^{\infty} \int_0^{2\pi} e^{-r^2}\, rdrd\phi = \pi \int_0^{\infty} e^{-r^2}\, d\left(r^2\right) \xrightarrow{z=-r^2} \pi \int_{-\infty}^{0} e^z\, dz = \pi \Rightarrow I = \sqrt{\pi}.$$

3.5 Миноры и ранги

Понятие минора также связано с рангом матрицы. Рассмотрим два набора индексов у матрицы размером $n \times m$: $1 \le i_1 \le i_2 \le \ldots \le i_k \le n$ и $1 \le j_1 \le j_2 \le \ldots \le j_k \le m$, и рассмотрим подматрицу A, состоящую из элементов, находящихся на пересечении строк $i_1, i_2, \ldots i_k$ и столбцов $j_1, j_2, \ldots j_k$:

$$\begin{pmatrix} a_{i_1 j_1} & a_{i_1 j_2} & \ldots & a_{i_1 j_k} \\ a_{i_2 j_1} & a_{i_2 j_2} & \ldots & a_{i_2 j_k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i_k j_1} & a_{i_k j_2} & \ldots & a_{i_k j_k} \end{pmatrix}.$$

Её определитель называется минором k-го порядка матрицы A. Например, если

$$A = \begin{pmatrix} 1 & -1 & -8 & 1 & 2 \\ 0 & 8 & -4 & -5 & 6 \\ -6 & 9 & 0 & 6 & -1 \\ 2 & 1 & 2 & 3 & 6 \end{pmatrix}$$

$(i_1, i_2, i_3) = (1, 2, 4)$, и $(j_1, j_2, j_3) = (2, 3, 5)$,то соответствующий минор 3-го порядка будет равен

$$\begin{vmatrix} -1 & -8 & 2 \\ 8 & -4 & 6 \\ 1 & 2 & 6 \end{vmatrix} = 24 - 48 + 32 + 8 + 6 + 384 + 12 = 418.$$

Теорема. Пусть A – матрица размером $n \times m$. Если хотя бы один её минор k-го порядка не равен нулю, то $\mathrm{rk}(A) \ge k$. Если все её миноры k-го порядка равны нулю, то $\mathrm{rk}(A) < k$.

Пример.

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

Решение. Матрица содержит минор $\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1 \neq 0$, а значит её ранг не меньше 3. Поскольку он не может быть больше 3 (матрица содержит лишь три строки), то он равен трём.

4 Матричная алгебра

4.1 Матричные операции

Множество $n \times m$ матриц $\mathbb{R}^{n \times m}$ имеет естественную структуру линейного пространства по отношению к операциям матричного (покоординатного) сложения и умножения на скаляр. Пусть $A, B \in M_{n,m}(\mathbb{R})$,

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nm} \end{pmatrix}, B = \begin{pmatrix} b_{11} & b_{12} & \ldots & b_{1m} \\ b_{21} & b_{22} & \ldots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \ldots & b_{nm} \end{pmatrix}$$

Матрицы $A+B$ и λA определяются покомпонентным сложением элементов A и B и покомпонентным умножением элементов A на число λ, т.е.,

$$A+B=\begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & a_{1m}+b_{1m} \\ a_{21}+b_{21} & a_{22}+b_{22} & \dots & a_{2m}+b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}+b_{n1} & a_{n2}+b_{n2} & \dots & a_{nm}+b_{nm} \end{pmatrix},$$

$$\lambda A=\begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \dots & \lambda a_{1m} \\ \lambda a_{21} & \lambda a_{22} & \dots & \lambda a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda a_{n1} & \lambda a_{n2} & \dots & \lambda a_{nm} \end{pmatrix}$$

Пример.

$$A=\begin{pmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \end{pmatrix}, \quad B=\begin{pmatrix} 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad A+B=\begin{pmatrix} 2 & 2 & 2 \\ 3 & 4 & 6 \end{pmatrix}$$

Следующие свойства матричного сложения являются очевидными.

1. $\forall A,B,C \in M_{n,m}(\mathbb{R})(A+B)+C=A+(B+C)$
2. $\exists 0 \in M_{n,m}(\mathbb{R}) : \forall A \in M_{n,m}(\mathbb{R})A+0=0+A=A$
3. $\forall A \in M_{n,m}(\mathbb{R})\exists(-A) \in M_{n,m}(\mathbb{R}) : A+(-A)=(-A)+A=0$
4. $A+B=B+A$ for $\forall A,B \in M_{n,m}(\mathbb{R})$
5. $\forall \lambda \in \mathbb{R} \forall A,B \in M_{n,m}(\mathbb{R})\lambda(A+B)=\lambda A+\lambda B$
6. $\forall \lambda,\mu \in \mathbb{R} \forall A \in M_{n,m}(\mathbb{R})(\lambda\mu)A=\lambda(\mu)A$

Поэтому $\mathbb{R}^{n\times m}$ – линейное пространство[3] с $\dim(\mathbb{R}^{n\times m})=nm$. В отличие от матричного сложения, матричное умножение определено для матриц разных размеров.

Пусть $A \in \mathbb{R}^{n\times k}$ и $B \in \mathbb{R}^{k\times m}$. Заметим, что число столбцов в левой матрице должно совпадать с числом строк правой матрицы. Такие матрицы совместимы для умножения, и в этом случае матрица $C=A\times B=(c_{ij})$ определена следующим образом:

$$c_{ij}=\sum_{s=1}^{k} a_{is}\cdot b_{sj}$$

[3]Что является естественным базисом в этом пространстве?

Пример.

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

$$c_{11} = \sum_{s=1}^{2} a_{1s} b_{s1} = a_{11} b_{11} + a_{12} b_{21} = 1 \cdot 1 + 2 \cdot 4 = 9,$$

$$c_{12} = \sum_{s=1}^{2} a_{1s} b_{s2} = a_{11} b_{12} + a_{12} b_{22} = 1 \cdot 2 + 2 \cdot 5 = 12,$$

$$\vdots$$

$$c_{23} = \sum_{s=1}^{2} a_{2s} b_{s3} = a_{21} b_{13} + a_{22} b_{23} = 3 \cdot 3 + 4 \cdot 6 = 33,$$

$$A \cdot B = \begin{pmatrix} 9 & 12 & 15 \\ 19 & 26 & 33 \end{pmatrix}.$$

Матричное умножение – это бинарная операция, которую можно считать отображением $f : \mathbb{R}^{n \times k} \times \mathbb{R}^{k \times m} \to \mathbb{R}^{n \times m}$ со следующими свойствами:

1. $\forall A \in M_{n,k}(\mathbb{R}) \forall B \in M_{k,l}(\mathbb{R}) \forall C \in M_{l,m}(\mathbb{R}) (A \cdot B) \cdot C = A \cdot (B \cdot C)$
2. $\forall A \in M_{n,k}(\mathbb{R}) \forall B, C \in M_{k,m}(\mathbb{R}) A \cdot (B + C) = A \cdot B + A \cdot C$
3. $\forall \lambda \in \mathbb{R} \forall A \in M_{n,k}(\mathbb{R}) \forall B \in M_{k,m}(\mathbb{R}) (\lambda A) \cdot B = \lambda (A \cdot B)$

Заметим, что в общем случае $A \cdot B \neq B \cdot A$, то есть матричное умножение не является коммутативным. В этом заключается наиболее сильное отличие матричного умножения от умножения чисел. Поэтому при умножении матриц важно следить за порядком операций.

Пример.

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

Лемма 13. $(A \cdot B)^T = B^T \cdot A^T$

Доказательство. Рассмотрим $A \cdot B = C = (c_{ij})$. Тогда $c_{ij} = \sum_k a_{ik} b_{kj}$, и $C^T = (c_{ji})$, а также $c_{ji} = \sum_k a_{jk} b_{ki} = \sum_k b_{ki} a_{jk} = \sum_k \tilde{b}_{ik} \tilde{a}_{kj}$, где $\tilde{b}_{ij}$ и $\tilde{a}_{ij}$ являются элементами B^T и A^T соответственно. Поэтому $C^T = B^T \cdot A^T$.

4.2 Обратные матрицы

Понятие обратной матрицы применимо только к квадратным матрицам, и во многом похоже на понятие обратного числа. Напомним: если $x \in \mathbb{R}$, то x^{-1} определяется как число со свойством $x \cdot x^{-1} = 1$. Для матриц роль числа 1 играет единичная матрица I

$$I_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} \in \mathbb{R}^{n \times n},$$

которая является нейтральным элементом в матричном умножении, т.е. $A \cdot I = I \cdot A = A$ для любой $A \in \mathbb{R}^{n \times n}$.

Определение. Пусть A – матрица размером $n \times n$. Матрица $A^{-1} \in \mathbb{R}^{n \times n}$ называется обратной к A, если $A \cdot A^{-1} = A^{-1} \cdot A = I$, где I – единичная матрица.

Лемма 14. $(A \cdot B)^{-1} = B^{-1} \times A^{-1}$.

Доказательство. $\left(B^{-1} \cdot A^{-1}\right) \cdot (A \cdot B) = B^{-1} \cdot \left(A^{-1} \cdot A\right) \cdot B = B^{-1} \cdot B = I$. Аналогично, $(A \cdot B) \cdot \left(B^{-1} \cdot A^{-1}\right) = A \cdot \left(B \cdot B^{-1}\right) \cdot A^{-1} = A \cdot A^{-1} = I$.

Лемма. $\left(A^T\right)^{-1} = \left(A^{-1}\right)^T$.

Доказательство. В соответствии с леммой, $\left(A^{-1}\right)^T \cdot A^T = \left(A \cdot A^{-1}\right)^T = I$ и $A^T \cdot \left(A^{-1}\right)^T = \left(A^{-1} \cdot A\right)^T = I$. Поэтому, $\left(A^{-1}\right)^T$ является обратной к A^T.

4.3 Обращение матрицы с помощью метода Гаусса

Следующая теорема предлагает необходимое и достаточное условия существования обратной матрицы.

Теорема. A^{-1} существует тогда и только тогда, когда $\det(A) \neq 0$.

Доказательство. Необходимость следует по определению: $\det\left(A \cdot A^{-1}\right) = \det(A) \det\left(A^{-1}\right) = 1$ и, следовательно, $\det(A)$ не может быть равен нулю.

Теперь докажем условие достаточности. Предположим, что

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}, \quad A^{-1} = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{pmatrix}$$

и $C = A \cdot B$. Нам нужно $C = I$. Поэтому для первого столбца C нужно взять

$$c_{11} = a_{11}b_{11} + a_{12}b_{21} + \ldots + a_{1n}b_{n1} = 1,$$
$$c_{21} = a_{21}b_{11} + a_{22}b_{21} + \ldots + a_{2n}b_{n1} = 0$$
$$c_{n1} = a_{n1}b_{11} + a_{n2}b_{21} + \ldots + a_{nn}b_{n1} = 0$$

То есть, чтобы вычислить $b_{11}, \ldots, b_{n1}$, нужно решить СЛАУ

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & 1 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 \\ \vdots & \vdots & \ddots & \vdots & 0 \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 \end{array}\right)$$

Аналогично, для второго стобца C имеем

$$c_{12} = a_{11}b_{12} + a_{12}b_{22} + \ldots + a_{1n}b_{n2} = 0$$
$$c_{22} = a_{21}b_{12} + a_{22}b_{22} + \ldots + a_{2n}b_{n2} = 1$$
$$c_{n2} = a_{n1}b_{12} + a_{n2}b_{22} + \ldots + a_{nn}b_{n2} = 0$$

т.е., чтобы вычислить $b_{12}, \dots, b_{n2}$, нужно решить следующее СЛАУ:

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 1 \\ \vdots & \vdots & \ddots & \vdots & 0 \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 \end{array}\right)$$

Аналогичное СЛАУ можно записать для $b_{1n}, \dots, b_{nn}$. Заметим, что все эти СЛАУ имеют одинаковую матрицу коэффициентов и отличаются только свободными членами. Поэтому их можно решить одновременно, поскольку последовательность преобразований будет одной и той же для каждого из столбцов в правой части расширенной матрицы. Другими словами, можно взять матрицу вида

$$\left(\begin{array}{cccc|cccc} a_{11} & a_{12} & \dots & a_{1n} & 1 & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & 0 \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 & 0 & \dots & 1 \end{array}\right)$$

которую будем обозначать $(A \mid I)$, и использовать метод Гаусса для преобразования A в I. При этом матрица в правой части преобразуется в A^{-1}. В соответствии с теоремой, если $\det(A) \neq 0$, то решение для рассматриваемой СЛАУ существует и является уникальным.

Замечание. Доказательство этой теоремы можно использовать для численного нахождения A^{-1}.

Пример.

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

Решение.

$$\left(\begin{array}{cc|cc} 2 & 3 & 1 & 0 \\ 3 & 5 & 0 & 1 \end{array}\right) \to \left(\begin{array}{cc|cc} 1 & 2 & -1 & 1 \\ 2 & 3 & 1 & 0 \end{array}\right) \to \left(\begin{array}{cc|cc} 1 & 2 & -1 & 1 \\ 0 & -1 & 3 & -2 \end{array}\right) \to$$

$$\left(\begin{array}{cc|cc} 1 & 2 & -1 & 1 \\ 0 & 1 & -3 & 2 \end{array}\right) \to \left(\begin{array}{cc|cc} 2 & 3 & 1 & 0 \\ 3 & 5 & 0 & 1 \end{array}\right) \to \left(\begin{array}{cc|cc} 1 & 0 & 5 & -3 \\ 0 & 1 & -3 & 2 \end{array}\right)$$

Проверка умножением: $\begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix} \times \begin{pmatrix} 5 & -3 \\ -3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

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

Пример. Пусть $A = \begin{pmatrix} 2 & 1 & -1 \\ -1 & 1 & 2 \\ 0 & 3 & 1 \end{pmatrix}$, найдите A^{-1}.

Решение.

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

$$\left(\begin{array}{ccc|ccc} 1 & -1 & -2 & 0 & -1 & 0 \\ 0 & 1 & 1 & \frac{1}{3} & \frac{2}{3} & 0 \\ 0 & 0 & 1 & \frac{1}{2} & 1 & -\frac{1}{2} \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 1 & -1 \\ 0 & 1 & 0 & -\frac{1}{6} & -\frac{1}{3} & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{2} & 1 & -\frac{1}{2} \end{array}\right) \to$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{5}{6} & \frac{2}{3} & -\frac{1}{2} \\ 0 & 1 & 0 & -\frac{1}{6} & -\frac{1}{3} & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{2} & 1 & -\frac{1}{2} \end{array}\right)$$

Проверка:

$$\begin{pmatrix} 2 & 1 & -1 \\ -1 & 1 & 2 \\ 0 & 3 & 1 \end{pmatrix} \times \begin{pmatrix} \frac{5}{6} & \frac{2}{3} & -\frac{1}{2} \\ -\frac{1}{6} & -\frac{1}{3} & \frac{1}{2} \\ \frac{1}{2} & 1 & -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.\cdot$$

4.4 Матрицы элементарных преобразований

Трансформации строк матриц каждого из трёх типов можно выразить в форме матричного умножения с помощью специальных матриц, называемых матрицами элементарных преобразований.

Трансоформация I типа, в которой j-я строка умножается на константу λ и прибавляется к i-й строке эквивалентна умножению (слева) исходной матрицы на матрицу

$$U_{ij}(\lambda) = \begin{pmatrix} 1 & 0 & \lambda & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

где λ находится на пересечении i-й строки и j-го столбца; например, в случае матриц 3×3,

$$U_{13}(\lambda) = \begin{pmatrix} 1 & 0 & \lambda \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Трансформации II типа, в которых i-я строка умножается на ненулевую константу λ, эквивалентна на умножение (слева) исходной матрицы на

$$P_i(\lambda) = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & \lambda & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

где λ находится на i-й строке главной диагонали; например,

$$P_2(\lambda) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Трансформация III типа, в которой i-я и j-я строки исходной матрицы меняются местами, может быть выражена умножением (снова слева) на матрицу

$$T_{ij}(\lambda) = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & 1 & \dots & 0 \\ 0 & 0 & \dots & 1 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & 1 \end{pmatrix}$$

где единицы вне главной диагонали находятся на пересечении i-й строки и j-го столбца, а также i-го столбца и j-й строки, например

$$T_{12} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

<u>Пример</u>. Пусть A – матрица размером 3×3 с произвольными элементами

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

тогда $U_{13}(\lambda) \cdot A =$

$$= \begin{pmatrix} 1 & 0 & \lambda \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{11} + \lambda a_{31} & a_{12} + \lambda a_{32} & a_{13} + \lambda a_{33} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},$$

$$P_2(\lambda) \times A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ \lambda a_{21} & \lambda a_{22} & \lambda a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \text{ and}$$

$$T_{12} \times A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Аналогично, можно показать, что умножение справа на $U_{ji}(\lambda)$, $P_i(\lambda)$, и T_{ij} приводят к трансофрмациям I-III типов, применённым к столбцам матрицы. Полезно заметить, что умножение на $U_{ij}(\lambda)$ не изменяет $\det(A)$, тогда как умножение на T_{ij} приводит к изменению знака $\det(A)$. Поэтому иногда используется матрица

$$R_{ij} = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & -1 & \dots & 0 \\ 0 & 0 & \dots & 1 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & 1 \end{pmatrix}$$

где 1 и -1 назодятся на пересечении i-й и j-й строки и i-го и j-го столбца. Умножение слева или справа на такую матрицу не меняет определитель матрицы. Элементарные преобразования используются для доказательства следующей теоремы:

Теорема.

$$\det(A \times B) = \det(A)\det(B)$$

Доказательство. Во первых, доказательство тривиально для вырожденных матриц. Если какая-то из A или B является вырожденной, то $A \times B$ тоже является ывронжденной, поскольку её строки являются линейными комбинациями линейно зависимого набора векторов, так что $\det(A \times B)$ и $\det(A)\det(B)$ оба равны нулю.

Далее, докажем утверждение теоремы для особого случая, когда A и B являются верхнетреугольными. Используем следующие обозначения:

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

Тогда, используя определение матричного умножения, получаем

$$AB = \begin{pmatrix} a_{11}b_{11} & \cdot & \cdot & \dots & \cdot \\ 0 & a_{22}b_{22} & \cdot & \dots & \cdot \\ 0 & 0 & a_{33}b_{33} & \dots & \cdot \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_{nn}b_{nn} \end{pmatrix}$$

и поскольку определитель верхнетреугольной матрицы равен произведению диагональных элементов, то $\det(A \times B) = a_{11}b_{11}a_{22}b_{22}\dots a_{nn}b_{nn} = a_{11}a_{22}\dots a_{nn}b_{11}b_{22}\dots b_{nn} = \det(A)\det(B)$.

Общий случай произведения произвольных квадратных матриц сводится к произведению верхнетреугольных матриц с помощью элементарных преобразований строк матрицы A и с помощью элементарных преобразований столбцой матрицы B. В таком случае A можно представить в виде произведения $A = A_1 \cdot A_2 \cdot \ldots \cdot A_k \cdot X$, где A_i – одна из матриц $U_{ij}(\lambda)$ или R_{ij}, которые не меняют определитель, тогда как X – верхнетреугольная матрица со свойством $\det(A) = \det(X)$. Аналогично мы можем преобразовать B к форме $Y \cdot B_l \cdot \ldots \cdot B_2 \cdot B_1$, где B_j – это матрицы $U_{ij}(\lambda)$ или R_{ij}, а Y – верхнетреугольная матрица с $\det(B) = \det(Y)$. Тогда $\det(A \cdot B) = \det(A_1 \cdot A_2 \cdot \ldots \cdot A_k \cdot X \cdot Y \cdot B_l \cdot \ldots \cdot B_2 \cdot B_1) = \det(X \cdot Y) = \det(X)\det(Y) = \det(A)\det(B)$.

4.5 Упражнения

1. Найдите обратную матрицу для каждой из следующих.

(a) $\begin{pmatrix} -1 & -2 & 4 & 11 & -2 \\ -4 & 0 & 9 & -7 & -3 \\ 0 & 1 & -1 & -8 & 0 \\ -5 & -1 & 12 & -3 & -5 \\ 4 & 0 & -9 & 9 & 4 \end{pmatrix}$ (b) $\begin{pmatrix} 2 & 0 & -7 & -1 \\ 0 & 0 & 1 & 0 \\ -5 & 3 & -9 & 2 \\ -12 & 5 & -3 & 5 \end{pmatrix}$ (c) $\begin{pmatrix} -6 & 7 & 7 & -6 \\ -1 & 2 & 0 & 0 \\ -6 & 7 & 8 & -6 \\ 2 & -3 & -1 & 1 \end{pmatrix}$

(d) $\begin{pmatrix} 7 & 4 & 6 & 3 \\ 4 & -2 & -2 & 1 \\ 12 & 6 & 9 & 5 \\ 10 & 3 & 5 & 4 \end{pmatrix}$ (e) $\begin{pmatrix} 3 & -2 & -4 \\ -2 & 1 & 3 \\ -6 & 3 & 8 \end{pmatrix}$ (f) $\begin{pmatrix} 11 & -11 & 10 & 10 \\ 0 & -1 & 1 & 1 \\ -3 & 5 & -4 & -5 \\ 7 & -4 & 4 & 3 \end{pmatrix}$

(g) $\begin{pmatrix} -11 & -10 & -6 \\ -6 & -5 & -3 \\ -10 & -8 & -5 \end{pmatrix}$ (h) $\begin{pmatrix} -3 & 10 & -1 & 0 & -1 \\ 4 & -12 & 1 & 0 & 1 \\ 1 & -2 & 3 & 0 & -2 \\ -1 & 0 & 5 & 1 & -2 \\ 0 & -3 & 7 & 1 & -3 \end{pmatrix}$ (i) $\begin{pmatrix} 3 & -7 & -11 & -7 & 5 \\ 7 & -1 & 11 & -8 & 3 \\ -5 & 2 & -5 & 6 & -3 \\ 5 & 1 & 12 & -5 & 1 \\ 1 & 4 & 12 & 1 & -2 \end{pmatrix}$

(j) $\begin{pmatrix} 1 & 2 & -1 & -1 & -2 \\ -1 & -8 & 10 & 0 & 5 \\ 1 & 1 & -2 & -4 & -2 \\ -1 & 0 & 2 & 6 & 2 \\ 1 & 2 & -5 & -6 & -3 \end{pmatrix}$

2. Решите матричные уравнения

(a) $\begin{pmatrix} -3 & 3 & 5 \\ 1 & -1 & -2 \\ -7 & 6 & 10 \end{pmatrix} X = \begin{pmatrix} -1 & -5 & -4 \\ 3 & 2 & -2 \\ -1 & 2 & 2 \end{pmatrix}$

(b) $\begin{pmatrix} 6 & 4 & 1 & -6 \\ -5 & -4 & -1 & 6 \\ 8 & 3 & 1 & -4 \\ 10 & 0 & 1 & 1 \end{pmatrix} X = \begin{pmatrix} 1 & -1 & 0 & -4 \\ -4 & -5 & -4 & -2 \\ -2 & -2 & 2 & 1 \\ -1 & 2 & 1 & 3 \end{pmatrix}$

(c) $X \begin{pmatrix} 0 & -3 & 11 & 0 & 9 \\ 1 & 0 & -5 & -1 & -3 \\ 1 & 0 & 5 & 2 & 5 \\ -1 & 4 & -2 & 3 & -3 \\ -1 & -5 & 2 & -5 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 3 & -1 & 3 \\ 4 & 1 & -3 & 2 & 3 \\ -4 & 1 & -4 & -1 & -4 \\ -2 & 4 & 1 & 0 & 3 \end{pmatrix}$

(d) $X \begin{pmatrix} 7 & -6 & -8 \\ -1 & 1 & 1 \\ 3 & -2 & -3 \end{pmatrix} = \begin{pmatrix} 1 & -2 & -3 \\ 1 & 2 & -3 \\ -4 & 3 & 2 \\ 4 & 1 & -3 \\ 1 & -5 & -5 \end{pmatrix}$

(e) $\begin{pmatrix} 3 & -3 & -1 \\ 2 & 1 & 0 \\ 4 & -3 & -1 \end{pmatrix} X = \begin{pmatrix} -4 & -2 & 1 \\ 2 & 0 & -2 \\ -3 & -3 & -1 \end{pmatrix}$

(f) $\begin{pmatrix} 8 & -12 & -1 & 0 \\ -3 & 7 & 2 & -1 \\ -7 & 6 & -2 & 2 \\ 0 & 5 & 3 & -2 \end{pmatrix} X = \begin{pmatrix} 2 & 3 & 2 & -2 & -4 \\ 4 & -3 & 1 & 1 & 4 \\ -5 & -3 & 2 & 2 & 4 \\ 3 & 4 & -2 & -5 & -2 \end{pmatrix}$

5 Линейные операторы

Понятие линейных операторов является очень широким, и имеет множество применений в математике, включая анализ, статистику и оптимизационные задачи. Синонимом понятия “линейный оператор” являются “линейное отображение”, “линейная трансформация” и “линейное функция”.

Определение. Пусть V и W – линейные (векторные) пространств. Отображение $f : V \to W$ называется линейным оператором, если

1. $\forall x, y \in V \quad f(x+y) = f(x) + f(y)$
2. $\forall x \in V, \forall \lambda \in \mathbb{R} f(\lambda x) = \lambda f(x)$,

Несколько примеров отображений, являющихся (или не являющихся) линейными:

1. $f : \mathbb{R} \to \mathbb{R}, f(x) = cx$ – линейный оператор, так как $c(x+y) = cx + cy$ и $c(\lambda x) = (c\lambda)x$.
2. $f : \mathbb{R} \to \mathbb{R}, f(x) = x^2$ НЕ линейный оператор, поскольку $(x+y)^2 \neq x^2 + y^2$.
3. $f : \mathbb{R} \to \mathbb{R}, f(x) = x + 1$ НЕ линейный оператор, поскольку $(x+1) + (y+1) \neq (x+y) + 1$.
4. Оператор дифференциирования по x, то есть, $\frac{\partial}{\partial x} : C^\infty[a,b] \to C^\infty[a,b]$, заданный как $\frac{\partial}{\partial x}(f) = f'(x)$, и определённый на пространстве бесконечно гладких функций, является линейным оператором.
5. Оператор интегрирования на отрезке $[a,b]$, то есть, $I : C[a,b] \to \mathbb{R}$, такой, что $I(f) = \int_a^b f(x)\,dx$ является линейным оператором, поскольку $\int_a^b (f+g)(x)\,dx = \int_a^b f(x)dx + \int_a^b g(x)\,dx$ и $\int_a^b (\lambda f)(x)\,dx = \lambda \int_a^b f(x)\,dx$.
6. Оператор взятия математического ожидания $\mathbb{E}$ задаётся на линейном пространстве случайных величин (заданных на фиксированном вероятностном пространстве), и обладает свойствами $\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$ and $\mathbb{E}(\lambda X) = \lambda \mathbb{E}(X)$.
7. $f : \mathbb{R}^2 \to \mathbb{R}^2, f(x,y) = (y,x)$ – линейный оператор симметрии на плоскости относительно прямой, проходящей через центр координат под 45°, поскольку вектор, симметричный сумме векторов, равен сумме их симметричных отображений; то же самое выполняется для растяжений.
8. $f : \mathbb{R}^2 \to \mathbb{R}^2, f(x,y) = (y,-x)$ – линейный оператор поворота на 90° против часовой стрелки
9. $f : \mathbb{R}^2 \to \mathbb{R}, f(x,y) = x$ – линейный оператор проекции на ось Ox, поскольку проекция суммы векторов равна сумме их проекций.
10. $f : \mathbb{R}^2 \to \mathbb{R}, f(x,y) = \sqrt{x^2+y^2}$ НЕ является линейным, так как длина суммы векторов не равна сумме длин.

5.1 Матрица линейного оператора

Можно показать, что любой линейный оператор в конечномерном пространстве можно представить в виде матрицы, но это представление неоднозначно и зависит от выбранного в линейном пространстве базиса.

Рассмотрим линейное пространство V с базисом $e_1, e_2, \ldots, e_n$. Поскольку каждый вектор $x \in V$ можно представить в виде линейной комбинации $e_1, e_2, \ldots, e_n$, получаем $x = x_1e_1 + x_2e_2 + \ldots + x_ne_n$, где $x_1, x_2, \ldots, x_n$ – координаты x в базисе $e_1, e_2, \ldots, e_n$. Если $A : V \to W$ – линейный оператор, то

$$A(x) = A\left(x_1e_1 + x_2e_2 + \ldots + x_ne_n\right) = x_1A\left(e_1\right) + x_2A\left(e_2\right) + \ldots + x_nA\left(e_n\right).$$

Здесь $A\left(e_1\right), A\left(e_2\right), \ldots, A\left(e_n\right)$ – образы векторов при отображении f.

Аналогичные выражения верны для любого x. Выберем базис $f_1, f_2, \ldots, f_m$ в векторном пространстве W и разложим вектора $A(e_1), A(e_2), \ldots, A(e_n)$ по этому базису[4]

$$A(e_1) = a_{11}f_1 + a_{21}f_2 + \ldots + a_{m1}f_m$$
$$A(e_2) = a_{12}f_1 + a_{22}f_2 + \ldots + a_{m2}f_m$$
$$A(e_n) = a_{1n}f_1 + a_{2n}f_2 + \ldots + a_{mn}f_m$$

Получили

$$A(x) = x_1 A(e_1) + x_2 A(e_2) + \ldots + x_n A(e_n) =$$

$$= x_1(a_{11}f_1 + a_{21}f_2 + \ldots + a_{m1}f_m) + x_2(a_{12}f_1 + a_{22}f_2 + \ldots + a_{m2}f_m) + \ldots + x_n(a_{1n}f_1 + a_{2n}f_2 + \ldots + a_{mn}f_m) =$$

$$= (a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n) f_1 + (a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n) f_2 + \ldots + (a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n) f_m.$$

Поэтому преобразование $y = A(x)$ можно записать, используя матричное умножение

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{1mn} \end{pmatrix} \times \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

где $x_1, x_2, \ldots, x_n$ – координаты x в базисе $e_1, e_2, \ldots, e_n$ и $y_1, y_2, \ldots, y_m$ являются координатами $A(x)$ в базисе $f_1, f_2, \ldots, f_m$. Заметим, что это тот же базис, который приводил к матрице (a_{ij}) координат $f(e_i)$. Матрица, составленная из (a_{ij}) называется матрицей линейного оператора A и обычно обозначается так же, как сам оператор. Элементы этой матрицы уникальны для базисов $e_1, e_2, \ldots, e_n$ и $f_1, f_2, \ldots, f_m$. Для того, чтобы построить матрицу какого-то линейного оператора, достаточно построить образы всех базисных векторов пространства V, и найти их координаты в базисе f пространства W. Столбцами матрицы A в этом случае будут состоять из полученных координат векторов-образов. Построенная матрица является матрицей линейного оператора.

Пример. Пусть $A : \mathbb{R}^2 \to \mathbb{R}^2$ – линейный оператор поворота на $90°$ против часовой стрелки. Рассмотрим $e_1 = f_1 = (1,0)^T$ и $e_2 = f_2 = (0,1)^T$. Тогда $A(e_1) = f_2 = 0 \cdot f_1 + 1 \cdot f_2$ и $A(e_2) = -f_1 = (-1) \cdot f_1 + 0 \cdot f_2$, поэтому матрица A в этих базисах имеет вид

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

Пример. Пусть V – векторное пространство полиномов степени, не превышающей 2 и пусть $A : V \to V$ – оператор дифференциирования по x. Рассмотрим базисы e_1, e_2, e_3 и f_1, f_2, f_3, такие, что $e_1 = f_1 = 1, e_2 = f_2 = x$ и $e_3 = f_3 = x^2$. Тогда $A(e_1) = 0 = 0 \cdot f_1 + 0 \cdot f_2 + 0 \cdot f_3, A(e_2) = f_1 = 1 \cdot f_1 + 0 \cdot f_2 + 0 \cdot f_3$, и $A(e_3) = 2f_2 = 0 \cdot f_1 + 2 \cdot f_2 + 0 \cdot f_3$. Поэтому матрица A имеет вид

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

Пример. Пусть $A : \mathbb{R}^3 \to \mathbb{R}^3$ – линейный оператор поворота в трёхмерном пространстве, который заклю-

[4]Обратите внимание на нестандартные обозначения индексов, которые использованы, чтобы получить более удобное выражение для итогового результата.

чается в изменении осей Ox, Oy, и Oz по циклу, – то есть, это поворот вокруг прямой линии $x = y = z$ на 120° по часовой стрелке, если смотреть на центр координат из первого октанта пространства.[5] Рассмотрим $e_1 = f_1 = (1,0,0)^T, e_2 = f_2 = (0,1,0)^T$, и $e_3 = f_3 = (0,0,1)^T$. Тогда $A(e_1) = f_2 = 0 \cdot f_1 + 1 \cdot f_2 + 0 \cdot f_3$, $A(e_2) = f_3 = 0 \cdot f_1 + 0 \cdot f_2 + 1 \cdot f_3$, и $A(e_3) = f_1 = 1 \cdot f_1 + 0 \cdot f_2 + 0 \cdot f_3$, поэтому матрица A в этих базисах имеет вид

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

5.2 Изменение базиса

В прошлой части мы рассмотривали два разных линейных пространства V и W, каждое со своим базисом. Рассмотрим теперь линейное пространство V с двумя различными базисами, $e_1, e_2, \ldots, e_n$ и $e'_1, e'_2, \ldots, e'_n$. Далее они будут обозначаться как "старый" и "новый" базисы соответственно, то есть мы рассмотрим задачу замены базиса. Пусть $x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n = x'_1 e'_1 + x'_2 e'_2 + \ldots + x'_n e'_n$ – два набора координат вектора x в старом и новом базисах. Обозначим координаты векторов старого базиса в новом базисе следующим образом:

$$\begin{aligned} e_1 &= c_{11} e'_1 + c_{21} e'_2 + \ldots + c_{n1} e'_n, \\ e_2 &= c_{12} e'_1 + c_{22} e'_2 + \ldots + c_{n2} e'_n \\ e_n &= c_{1n} e'_1 + c_{2n} e'_2 + \ldots + c_{nn} e'_n \end{aligned}$$

Тогда, имеем

$$x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n =$$

$$= x_1 \left(c_{11} e'_1 + c_{21} e'_2 + \ldots + c_{n1} e'_n\right) + x_2 \left(c_{12} e'_1 + c_{22} e'_2 + \ldots + c_{n2} e'_n\right) + \ldots + x_n \left(c_{1n} e'_1 + c_{2n} e'_2 + \ldots + c_{nn} e'_n\right) =$$

$$= \left(x_1 c_{11} + x_2 c_{12} + \ldots + x_n c_{1n}\right) e'_1 + \left(x_1 c_{21} + x_2 c_{22} + \ldots + x_n c_{2n}\right) e'_2 + \ldots + \left(x_1 c_{n1} + x_2 c_{n2} + \ldots + x_n c_{nn}\right) e'_n =$$

$$= x'_1 e'_1 + x'_2 e'_2 + \ldots + x'_n e'_n.$$

Поскольку координаты вектора в базисе определены уникальным образом, получаем

$$\begin{cases} x'_1 = c_{11} x_1 + c_{12} x_2 + \ldots + c_{1n} x_n \\ x'_2 = c_{12} x_1 + c_{22} x_2 + \ldots + c_{2n} x_n \\ x'_n = c_{1n} x_1 + c_{n2} x_2 + \ldots + c_{nn} x_n \end{cases}$$

Таким образом, замена базиса приводит к замене координат вектора в соответствии с правилом

$$x' = Cx,$$

где x' – вектор-столбец координат в новом базисе, а x – вектор-столбец координат в старом базисе. Преобразование координат в направлении $e' \to e$ можно произвести используя обратную матрицу, то есть

$$x = C^{-1} x'.$$

Матрица C называется матрицей преобразования базиса из $e_1, e_2, \ldots, e_n$ в $e'_1, e'_2, \ldots, e'_n$.

Теперь рассмотрим линейный оператор $A : V \to V$ с матрицей A в старом базисе (который играет роль

[5] Докажите, что такая операция действительно является поворотом на 120°.

и e_i, и f_i в обозначениях, использовавшихся при определении матрицы линейного оператора) и с матрицей A' в новом базисе. То есть, $y = Ax$ и $y' = A'x'$. Кроме того, $x' = Cx$ и $y' = Cy$ для некоторой матрицы преобразования базиса C. Поэтому $y' = Cy = A'x' = A'Cx$, и тогда $y = C^{-1}A'Cx = Ax$. Поскольку матрица линейного оператора однозначно определяется построением, мы получаем

$$A = C^{-1}A'C, \text{ or } A' = CAC^{-1}.$$

Таким образом получили правило преобразования матрицы линейного оператора при смене базиса.

Определение. Две матрицы A и B размером $n \times n$ называются сопряженными (или подобными), если существует невырожденная матрица C, такая, что $B = CAC^{-1}$.

Таким образом, матрица линейного оператора может выглядеть по-разному в разных базисах. Две матрицы одного и того же линейного оператора в разных базисах являются сопряженными. В таком случае возникает вопрос, какова "стандартная", т.е. наиболее простая форма матрицы линейного оператора? Для ответа на этот вопрос, рассмотрим понятие собственных величин и собственных векторов.

5.3 Собственные величины и собственные вектора

Определение. Пусть V – линейное пространство и пусть $A : V \to V$ – линейный оператор. Вектор $x \neq 0$ назывется собственным вектором оператора A с собственным числом $\lambda \in \mathbb{R}$, если $A(x) = \lambda x$. Другими словами, A действует на x, как умножение на скаляр.

Определение. Пусть A является матрицей $n \times n$. Ненулевой вектор $x \in \mathbb{R}^n$ называется собственным вектором A с собственным значением λ, если $Ax = \lambda x$.

Теорема. Число $\lambda \in \mathbb{R}$ – собственное число матрицы A тогда и только тогда, когда $\det(A - \lambda I) = 0$.

Доказательство. Пусть λ – собственное число матрицы A. Тогда для некоторого вектора $x \neq 0$ имеем $Ax = \lambda x$, то есть,

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = \lambda x_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = \lambda x_2 \\ \ldots + a_{n2}x_2 + \ldots + a_{nn}x_n = \lambda x_n \\ a_{n1}x_1 + \ldots + 1. \end{cases}$$

Таким образом

$$\begin{cases} (a_{11} - \lambda)\, x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0 \\ a_{21}x_1 + (a_{22} - \lambda)\, x_2 + \ldots + a_{2n}x_n = 0 \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + (a_{nn} - \lambda)\, x_n = 0 \\ \end{cases}$$

что в матричной форме имеет вид $(A - \lambda I)x = 0$. Это выражение задаёт однородную СЛАУ, которая гарантированно имеет решение $x = 0$. Чтобы она имела какое-то другое решение, матрица коэффициентов должна быть вырожденной, поэтому $\det(A - \lambda I) = 0$.

В обратную сторону, если $\det(A - \lambda I) = 0$ то $(A - \lambda I)x = 0$ – вырожденная однородная СЛАУ и имеет ненулевое решение x. Для этого вектор выполнено $Ax - \lambda x = 0$, поэтому x является собственным вектором A с собственным числом λ.

Заметим, что $\det(A-\lambda I)$ является многочленом:

$$\begin{vmatrix} (a_{11}-\lambda) & a_{12} & \dots & a_{1n} \\ a_{21} & (a_{22}-\lambda) & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & (a_{nn}-\lambda) \end{vmatrix} =$$

$$= (a_{11}-\lambda)(a_{22}-\lambda)(a_{33}-\lambda)\dots(a_{nn}-\lambda) - a_{12}a_{21}(a_{33}-\lambda)\dots(a_{nn}-\lambda) + \dots =$$

$$(-1)^n\lambda^n + c_{n-1}\lambda^{n-1} + c_{n-2}\lambda^{n-2} + c_1\lambda^1 + c_0.$$

Определение. Полином $f_A(\lambda) = \det(A-\lambda I)$ называется характеристическим полиномом матрицы A.

Теорема. Характеристические полиномы сопряженных матрицы одинаковы.

Доказательство. Пусть B – матрица, сопряженная к A, то есть, $B = CAC^{-1}$. Тогда используя свойства матричного умножения,

$$f_B(\lambda) = \det(CAC^{-1} - \lambda I) = \det\left(CAC^{-1} - \lambda CIC^{-1}\right) = \det\bigl(C(A-\lambda I)C^{-1}\bigr) =$$

$$= \det(C)\det(A-\lambda I)\det\left(C^{-1}\right) = \det(A-\lambda I) = f_A(\lambda).$$

Следствие Сумма диагональных элементов сопряженных квадратных матрицы одинакова.

Доказательство. Используем доказанное свойство, что характеристические полиномы не изменяются при сопряжении. В частности, не изменяется коэффициент c_{n-1} при λ^{n-1}. В характеристическом полиноме произведение $n-1$ λ возникает только в компоненте, получающейся за счёт произведения элементов на главной диагонали

$$(a_{11}-\lambda)(a_{22}-\lambda)(a_{33}-\lambda)\dots(a_{nn}-\lambda),$$

поскольку все остальные компоненты определителя содержат λ в степени $n-2$ или меньше. Тогда $c_{n-1}\lambda^{n-1} = (-1)^{n-1}\lambda^{n-1}(a_{11}+a_{22}+\dots+a_{nn})$. Поэтому сумма диагональных элементов матрицы не меняется при сопряжении.

Определение. Сумма элементов на главной диагонали квадратной матрицы называется её следом и обозначается $\mathrm{tr}(A)$.

Пример. Найдите собственные вектора и собственные числа матрицы

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

Решение. $A - \lambda I = \begin{pmatrix} 3 & 4 \\ 5 & 2 \end{pmatrix} - \lambda\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 5 & 2 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} = \begin{pmatrix} 3-\lambda & 4 \\ 5 & 2-\lambda \end{pmatrix}$. Поэтому,

$$\begin{vmatrix} 3-\lambda & 4 \\ 5 & 2-\lambda \end{vmatrix} = (3-\lambda)(2-\lambda) - 20 = \lambda^2 - 5\lambda - 14 = 0.$$

Решая квадратное уравнение, получаем $\lambda_1 = -2$ и $\lambda_2 = 7$.

Случай 1: $\lambda_1 = 7$.

$$A - 7I = \begin{pmatrix} -4 & 4 \\ 5 & -5 \end{pmatrix}$$

$$\begin{pmatrix} -4 & 4 \\ 5 & -5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\left(\begin{array}{cc|c} -4 & 4 & 0 \\ 5 & -5 & 0 \end{array}\right) \xrightarrow{[1]\mapsto[1]/(-4)} \left(\begin{array}{cc|c} 1 & -1 & 0 \\ 5 & -5 & 0 \end{array}\right) \xrightarrow{[2]\mapsto[2]-[1]\times 5} \left(\begin{array}{cc|c} 1 & -1 & 0 \end{array}\right)$$

Существует ненулевое решение, то есть вектор $x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ (и все вектора, пропорциональные ему) являются собственными векторами с собственным числом $\lambda_1 = 7$.

Случай 2: $\lambda_2 = 2$.

$$A - \lambda I = \begin{pmatrix} 5 & 4 \\ 5 & 4 \end{pmatrix}$$

Аналогично, получаем

$$\begin{pmatrix} 5 & 4 \\ 5 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

и вектора вида $y = \begin{pmatrix} 4 \\ -5 \end{pmatrix}$ являются собственными с собственным числом $\lambda_2 = 2$.

Пример. Найдите собственные вектора и собственные числа матрицы

$$\begin{pmatrix} 5 & -7 & 7 \\ -8 & 3 & -2 \\ -8 & -1 & 2 \end{pmatrix}$$

Решение.

$$\det \begin{pmatrix} 5-\lambda & -7 & 7 \\ -8 & 3-\lambda & -2 \\ -8 & -1 & 2-\lambda \end{pmatrix} =$$

$$= (5-\lambda) \times (3-\lambda) \times (2-\lambda) - 7 \times 8 \times 2 + 8 \times 1 \times 7 + 8 \times 7 \times (3-\lambda) -$$
$$8 \times 7 \times (2-\lambda) - 1 \times 2 \times (5-\lambda) = (5-\lambda) \times (1-\lambda) \times (4-\lambda) = 0.$$

Case 1: $\lambda_1 = 5$.

$$A - \lambda I = \begin{pmatrix} 0 & -7 & 7 \\ -8 & -2 & -2 \\ -8 & -1 & -3 \end{pmatrix}$$

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

Взяв $x_3 = -2$ получаем $x_2 = -2$ и $x_1 = 1$. Первый собственный вектор – это $x = \begin{pmatrix} 1 \\ -2 \\ -2 \end{pmatrix}$.

Случай 2: $\lambda_1 = 1$.

$$A - \lambda I = \begin{pmatrix} 4 & -7 & 7 \\ -8 & 2 & -2 \\ -8 & -1 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 4 & -7 & 7 \\ -8 & 2 & -2 \\ -8 & -1 & 1 \end{pmatrix} \to \begin{pmatrix} 4 & -7 & 7 \\ -8 & 2 & -2 \\ 1 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 0 & -7 & 7 \\ 0 & 1 & -1 \\ 1 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \end{pmatrix}$$

Взяв $x_3 = 1$, получаем $x_2 = 1$ и $x_1 = 0$. Второй собственный вектор – это $y = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$.

Случай 3: $\lambda_1 = 4$.

$$A - \lambda I = \begin{pmatrix} 1 & -7 & 7 \\ -8 & -1 & -2 \\ -8 & -1 & -2 \end{pmatrix}$$

$$\begin{pmatrix} 1 & -7 & 7 \\ -8 & -1 & -2 \\ -8 & -1 & -2 \end{pmatrix} \to \begin{pmatrix} 1 & -7 & 7 \\ 8 & 1 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & -7 & 7 \\ 0 & -57 & 54 \end{pmatrix} \to \begin{pmatrix} 1 & -7 & 7 \\ 0 & 19 & -18 \end{pmatrix}$$

Взяв $x_3 = 19$, получаем $x_2 = 18$ и $x_1 = -7$. Тогда последний собственный вектор – это $z = \begin{pmatrix} -7 \\ 18 \\ 19 \end{pmatrix}$.

В этом примере можно заметить, что матрица трансформации

$$C = \begin{pmatrix} 1 & 0 & -7 \\ -2 & 1 & 18 \\ -2 & 1 & 19 \end{pmatrix}$$

которая получается объединением в матрицу векторов x, y и z, преобразует A в диагональную:

$$C^{-1}AC = \begin{pmatrix} 1 & -7 & 7 \\ 2 & 5 & -4 \\ 0 & -1 & 1 \end{pmatrix} \times \begin{pmatrix} 5 & -7 & 7 \\ -8 & 3 & -2 \\ -8 & -1 & 2 \end{pmatrix} \times \begin{pmatrix} 1 & 0 & -7 \\ -2 & 1 & 18 \\ -2 & 1 & 19 \end{pmatrix} = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix},$$

где элементы на главной диагонали – это собственные числа A. Это происходит из-за того, что x, y и z линейно независимы и образуют базис в $\mathbb{R}^3$, тогда как матрица линейного оператора в базисе собственных векторов всегда является диагональной.

Теорема. Пусть A – матрица размером $n \times n$. Собственные вектора A, соответствующие разным собственным числам являются линейно независимыми.

Доказательство. Обозначим за $x_1, x_2, \ldots, x_k$ – собственные вектора, соответствующие собственным числам $\lambda_1, \lambda_2, \ldots, \lambda_k$. Требуется доказать, что если $\forall i \neq j,\ \lambda_i \neq \lambda_j$, то $x_1, x_2, \ldots, x_k$ линейно независимы.

Доказательство заключается в индукции по k. База индукции ($k = 1$) очевидна, поскольку каждый собственный вектор является ненулевым, и поэтмоу образует линейно независимый набор из одного элемента. Предположим, что верным является утверждение для набора из $k-1$ собственных векторов с разными собственными числами, и теперь хотим доказать его для k собственных векторов.

От противного, пусть при добавлении x_k к набору $x_1, x_2, \ldots, x_{k-1}$, он становится линейно зависимым, то есть существуют нетривиальные коэффициенты $\mu_1, \mu_2, \ldots, \mu_k$, такие, что

$$\mu_1 x_1 + \mu_2 x_2 + \ldots + \mu_k x_k = 0. \tag{4}$$

Умножая A на обе части уравнения, получаем

$$A(\mu_1 x_1 + \mu_2 x_2 + \ldots + \mu_k x_k) = A(\mu_1 x_1) + A(\mu_2 x_2) + \ldots + A(\mu_k x_k) =$$

$$= \mu_1 A(x_1) + \mu_2 A(x_2) + \ldots + \mu_k A(x_k) = \mu_1 \lambda_1 x_1 + \mu_2 \lambda_2 x_2 + \ldots + \mu_k \lambda_k x_k = 0. \tag{5}$$

Рассмотрим в виде системы уравнение (5) и (4), умноженное на λ_k.

$$\begin{cases} \mu_1 \lambda_1 x_1 + \mu_2 \lambda_2 x_2 + \ldots + \mu_k \lambda_k x_k = 0 \\ \mu_1 \lambda_k x_1 + \mu_2 \lambda_k x_2 + \ldots + \mu_k \lambda_k x_k = 0 \end{cases}$$

Из этой СЛАУ следует, что

$$\mu_1 (\lambda_1 - \lambda_k) x_1 + \mu_2 (\lambda_2 - \lambda_k) x_2 + \ldots + \mu_{k-1} (\lambda_{k-1} - \lambda_k) x_{k-1} = 0.$$

По предположению, линейная комбинация $x_1, \ldots, x_{k-1}$ может равняться нулю только если все коэффициенты перед x_i равны нулям, то есть $\mu_1 (\lambda_1 - \lambda_k) = \mu_2 (\lambda_2 - \lambda_k) = \ldots = \mu_{k-1} (\lambda_{k-1} - \lambda_k) = 0$. Так как все λ_i различны, должно выполняться $\lambda_i - \lambda_k \neq 0$ и, соответственно, $\mu_1 = \mu_2 = \ldots = \mu_{k-1} = 0$.

Повторяя ту же самую процедуру, только с умножением (4) на λ_1 вместо λ_k, получаем, что также выполняется $\mu_k = 0$. Поэтому $\mu_1 = \mu_2 = \ldots = \mu_n = 0$, а значит, только тривиальный набор коэффициентов делает линейную комбинацию $x_1, x_2, \ldots, x_k$ равной нулю, – а значит, эти вектора линейно независимы.

5.4 Диагонализуемые матрицы

Линейные операторы, имеющие базис из собственных векторов, очень удобны. Такие операторы играют важную роль во многих приложениях, в частности для решения линейных дифференциальных и разностных уравнений.

Определение. Линейный оператор $A : V \to V$ называется диагонализуемым, если его матрица является диагональной в некотором базисе пространства V. Естественно, вектора такого базиса являются собственными для A.

Определение. Матрица A размером $n \times n$ называется диагонализуемой, если существует невырожденная матрица C и диагональная матрица D, такие, что $CAC^{-1} = D$.

Очевидно, линейный оператор диагонализуем тогда и только тогда, когда его матрица (в некотором базисе) диагонализуема. Учитывая, что характеристический многочлен $f_A(\lambda)$ имеет степень n, а значит, не более n вещественных корней, следующие утверждения являются следствиями ранее доказанной теоремы.

Теорема. Матрица размером $n \times n$, имеющая n различныз собственных чисел диагонализуема.

Однако, это требование является достаточным, но не необходимым, то есть существуют диагонализуемые матрицы с совпадающими собственными числами.

Пример.

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

$\det(A - \lambda I) = (1-\lambda)^3$, то есть $\lambda = 1$. Однако $A - \lambda I = 0$, так что любой вектор в $\mathbb{R}^3$ – собственный для A, и конечно из них можно выбрать собственный базис A. Поэтому A является диагонализуемой (она уже имеет диагональную форму).

Пример. Матрица $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$ не является диагонализуемой. Действительно, $\det(A-\lambda I) = (2-\lambda)^3$, то есть, $\lambda = 2$. Тогда все собственные вектора A должны удовлетворять СЛАУ

$$\left(\begin{array}{ccc|c} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right)$$

Все решения такой СЛАУ пропорциональны вектору $x = (1,0,0)^T$ и поэтому не могут образовать базис в $\mathbb{R}^3$. Значит матрица A не диагонализуема.

Проблема с матрицей A в этом примере заключается в том, что кратность корня характеристического многочлена равна трём, то есть он представим в виде $(2-\lambda)^3$, тогда как решения соответствующей СЛАУ образуют лишь одномерное подпространство в $\mathbb{R}^3$. В предыдущем примере алгебраическая кратность корня характеристического многочлена также была равна трём, но соответствующая СЛАУ имела трёхмерное пространство решений.

Для каждого собственного числа λ_0 матрицы существуют два параметра, характеризующих число собственных векторов, относящихся к λ_0.

Определение. Пусть λ_0 – корень характеристического многочлена матрицы A размером $n \times n$. Алгебраическая кратность λ_0 – это наибольшая степень k, такая, что $f_A(\lambda)$ делится на $(\lambda - \lambda_0)^k$ без остатка. Геометрическая кратность λ_0 – это число линейно независимых решений у системы $(A - \lambda_0 I)\, x = 0$, то есть она равна $n - \operatorname{rk}(A - \lambda_0 I)$.

Лемма. Алгебраическая кратность собственного числа больше, либо равна геометрической.

Теорема (без доказательства). Матрица $n \times n$ является диагонализуемой тогда и только тогда, когда все корни $f_A(\lambda)$ – вещественны и их алгебраические кратности равны геометрическим.

Диагонализируемые матрицы являются особенными, поскольку они сопряжены с диагональными матрицами.

Диагонализация может использоваться для нахождения степеней матрицы A. Предположим, что мы нашли, что

$$A = CDC^{-1}$$

и хотим найти A^k. Несложно заметить, что

$$A^k = CD\left(C^{-1}C\right)D\left(C^{-1}\dots C\right)DC^{-1} = CD^kC^{-1}$$

То есть, чтобы возвести диагонализируемую матрицу в степень, можно найти её диагональную форму, возвести её в соответствующую степень (она также будет диагональной матрицей, с элементами $\lambda_1^k, \dots, \lambda_n^k$ на главной диагонали), а затем преобразовать её обратно используя ту же самую матрицу трансформации.

Пример. Найдите A^{20} для $A = \begin{pmatrix} 5 & -5 & -8 \\ -8 & 2 & 8 \\ -1 & -5 & -2 \end{pmatrix}$.

Решение. Во-первых, поскольку A имеет три различных собственных числа ($\lambda_1 = -3, \lambda_2 = 2, \lambda_3 = 6$), она является дигонализуемой. Далее нужно найти матрицу собственных векторов и обратную к ней, чтобы провести диагонализацию:

$$\begin{pmatrix} 5 & -5 & -8 \\ -8 & 2 & 8 \\ -1 & -5 & -2 \end{pmatrix} = \begin{pmatrix} 1 & -1 & -2 \\ 0 & 1 & 2 \\ 1 & -1 & -1 \end{pmatrix} \times \begin{pmatrix} -3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 6 \end{pmatrix} \times \begin{pmatrix} 1 & 1 & 0 \\ 2 & 1 & -2 \\ -1 & 0 & 1 \end{pmatrix}.$$

В итоге

$$A^{20} = \begin{pmatrix} 1 & -1 & -2 \\ 0 & 1 & 2 \\ 1 & -1 & -1 \end{pmatrix} \times \begin{pmatrix} 3^{20} & 0 & 0 \\ 0 & 2^{20} & 0 \\ 0 & 0 & 6^{20} \end{pmatrix} \times \begin{pmatrix} 1 & 1 & 0 \\ 2 & 1 & -2 \\ -1 & 0 & 1 \end{pmatrix}$$

Процедура нахождения степени матрицы позволяет также определить понятие функции от матриц для любой аналитической функции, т.е. такой, которая совпадает со своим степенным рядом (то есть, если ряд Тейлора функции является абсолютно сходящимся). Пусть

$$\varphi(x) = \sum_{k=0}^{\infty} \frac{\varphi^{(k)}(0)}{k!} x^k$$

– аналитическая функция. Тогда мы можем определить

$$\varphi(A) = \sum_{k=0}^{\infty} \frac{\varphi^{(k)}(0)}{k!} A^k,$$

поскольку матричное сложение и умножение корректно определны. Также, если A диагонализуема, то $A = CDC^{-1}$ и

$$\varphi(A) = \varphi\left(CDC^{-1}\right) = \sum_{k=0}^{\infty} \frac{\varphi^{(k)}(0)}{k!} \left(CDC^{-1}\right)^k = C\left(\sum_{k=0}^{\infty} \frac{\varphi^{(k)}(0)}{k!} D^k\right) C^{-1}$$

Если $D = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ то $\varphi(D) = \operatorname{diag}(\varphi(\lambda_1), \varphi(\lambda_2), \ldots, \varphi(\lambda_n))$ и

$$\varphi(A) = C \times \operatorname{diag}(\varphi(\lambda_1), \varphi(\lambda_2), \ldots, \varphi(\lambda_n)) \times C^{-1}$$

Пример. Найдите $\exp(A)$ для $A = \begin{pmatrix} -7 & -8 & -8 \\ 4 & 5 & 4 \\ 2 & 4 & -1 \end{pmatrix}$.

Решение. В начале нужно найти собственные числа матрицы; это $\lambda_1 = -1, \lambda_2 = -3, \lambda_3 = 1$. A является диагонализуемой, поскольку все три собственные числа – разные. Далее,

$$A = \begin{pmatrix} -4 & -2 & 3 \\ 2 & 1 & -2 \\ 1 & 0 & -1 \end{pmatrix} \times \begin{pmatrix} -1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} -1 & -2 & 1 \\ 0 & 1 & -2 \\ -1 & -2 & 0 \end{pmatrix}, \quad \text{и}$$

$$\exp(A) = \begin{pmatrix} -4 & -2 & 3 \\ 2 & 1 & -2 \\ 1 & 0 & -1 \end{pmatrix} \times \begin{pmatrix} \frac{1}{e} & 0 & 0 \\ 0 & \frac{1}{e^3} & 0 \\ 0 & 0 & e \end{pmatrix} \times \begin{pmatrix} -1 & -2 & 1 \\ 0 & 1 & -2 \\ -1 & -2 & 0 \end{pmatrix}.$$

5.5 Упражнения

1. Найдите собственные числа и собственные вектора следующих матриц

(a) $\begin{pmatrix} 5 & -2 \\ 1 & 8 \end{pmatrix}$ (b) $\begin{pmatrix} 7 & 6 \\ -8 & -7 \end{pmatrix}$ (c) $\begin{pmatrix} 4 & -8 \\ 4 & -8 \end{pmatrix}$ (d) $\begin{pmatrix} -8 & -1 \\ 6 & -3 \end{pmatrix}$ (e) $\begin{pmatrix} 1 & -7 & -7 \\ 7 & -5 & 1 \\ -7 & 7 & 1 \end{pmatrix}$

(f) $\begin{pmatrix} -6 & 2 & -4 \\ -8 & -3 & 2 \\ -4 & 2 & -6 \end{pmatrix}$ (g) $\begin{pmatrix} -5 & -2 & -2 \\ -8 & 1 & -2 \\ 8 & -8 & -5 \end{pmatrix}$ (h) $\begin{pmatrix} -4 & -4 & -4 & -4 \\ 0 & -2 & 5 & 5 \\ 2 & 4 & -7 & -5 \\ -2 & -4 & 6 & 4 \end{pmatrix}$

(i) $\begin{pmatrix} -8 & 7 & -1 & -3 \\ -2 & 1 & 2 & -4 \\ 4 & -4 & 1 & -1 \\ 4 & -4 & 6 & -6 \end{pmatrix}$ (j) $\begin{pmatrix} 6 & 0 & -4 \\ 24 & 6 & -44 \\ 8 & 0 & -6 \end{pmatrix}$

2. Найдите собственные числа и собственные вектора, а такде проведите диагонализацию матриц, если это возможно

(a) $\begin{pmatrix} 3 & -6 \\ 4 & -7 \end{pmatrix}$ (b) $\begin{pmatrix} 5 & 4 \\ 0 & 7 \end{pmatrix}$ (c) $\begin{pmatrix} -6 & -2 \\ 3 & -1 \end{pmatrix}$ (d) $\begin{pmatrix} 6 & 14 \\ 0 & -8 \end{pmatrix}$ (e) $\begin{pmatrix} -2 & -6 & -2 \\ 2 & 2 & -2 \\ -2 & -6 & -2 \end{pmatrix}$

(f) $\begin{pmatrix} 3 & -8 & 3 \\ -3 & -2 & 3 \\ 2 & -8 & 4 \end{pmatrix}$ (g) $\begin{pmatrix} -7 & 0 & -10 \\ 0 & 7 & 4 \\ 0 & 0 & 3 \end{pmatrix}$ (h) $\begin{pmatrix} -2 & 5 & -8 \\ 2 & 1 & -8 \\ -6 & 6 & -1 \end{pmatrix}$ (i) $\begin{pmatrix} -4 & 2 & -2 \\ -1 & -7 & 2 \\ -1 & -2 & -3 \end{pmatrix}$

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

6 Квадратичные формы

Пусть $x_1, x_2, \ldots, x_n$ – набор переменных. Мономом (одночленом) называется произведением неотрицательных целых степеней $x_1, x_2, \ldots, x_n$. Например, мономом является $x_1^2 x_2 x_3^4$. Естественным образом определяется понятие степени монома по отношению к переменной x_i; например, степень x_1 в мономе $x_1^2 x_2 x_3^4$ равна 2. Общая степень монома – это сумма степеней всех его переменных; так, степень рассматриваемого монома $x_1^2 x_2 x_3^4$ равна 7.

Однородный полином – это сумма мономов одинаковой степени. Примером является $x_1^5 + 2x_1^3 x_2^2 + 9x_2^5$ – однородный полином степени 5; другой пример – $x_1^3 + 3x_1^2 x_2^3 + x_2^7$ не является однородным, так как степени разных слагаемых здесь не совпадают.

Квадратичная форма – это однородны полином степени 2 от n переменных. Они играют важную роль в многих областях математики, включая анализ и теорию оптимизации.

6.1 Матрица квадратичной формы

Каждую квадратичную форму от n переменных можно записать в канонической форме в матричной записи. Например, $Q(x_1, x_2) = x_1^2 + 6x_1x_2 + 5x_2^2$ можно записать как произведение

$$(x_1, x_2) \begin{pmatrix} 1 & 3 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (x_1, x_2) \begin{pmatrix} x_1 + 3x_2 \\ 3x_1 + 5x_2 \end{pmatrix} = x_1^2 + 6x_1x_2 + 5x_2^2$$

Можно заметить, что произведение содержит две копии монома $3x_1x_2$, суммирующиеся в $6x_1x_2$. В общем случае, матричное произведение такого типа трансформируется в

$$(x_1 x_2 \dots x_n) \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n \end{pmatrix} =$$

$$= x_1 (x_{11}x_1 + a_{12}x_2 + \ldots + x_n) \, x_n(a_{n1}x_1 + a_{n2}x_2 + \ldots + x_{nn} (a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n) + \ldots + a_{11}x_1^2 +$$

$$+ (a_{12} + a_{21}) \, x_1x_2 + \ldots + a_{22}x_2^2 + \ldots + (a_{1n} + a_{n1})x_1x_n + \ldots + a_{nn}x_n^2).$$

Поскольку коэффициенты a_{ij} еще не определены, удобно считать, что $a_{ij} = a_{ji}$:

$$Q(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} a_{ii}x_i^2 + 2\sum_{i<j} a_{ij}x_ix_j$$

Определение. Матрица A размером $n \times n$, состоящая из элементов (a_{ij}) называется симметричной (антисимметричной), если $\forall i, j a_{ij} = a_{ji}$ $(\forall i, j a_{ij} = -a_{ji})$. Другими словами, A симметрична (антисимметрична), если $A = A^T$ $\left(A = -A^T\right)$.

Таким образом, каждая симметричная матрица определяет квадратичную форму и наоборот, каждая квадратичная форма определяет симметричную матрицу A со свойством

$$Q(x) = x^T A x$$

где $x = (x_1 x_2 \ldots x_n)^T$ – вектор-столбец. Как правило, на главной диагонали A находятся коэффициенты при полных квадратах выражения $Q(x)$, а коэффициенты при перекрёстных произведениях делются на два и помещаются симметрично над и под главной диагональю.

Пример. $Q(x_1, x_2, x_3) = x_1^2 + 2x_2^2 + 3x_3^2 + 4x_1x_3 = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 3 \end{pmatrix}$

Пример. $Q(x_1, x_2, x_3) = 4x_1x_2 + 2x_1x_3 + 10x_2x_3 = \begin{pmatrix} 0 & 2 & 1 \\ 2 & 0 & 5 \\ 1 & 5 & 0 \end{pmatrix}$

Определение.

Квадратичная форма $Q(x_1, x_2, \ldots, x_n)$ называется положительно определённой, если $Q(x_1, x_2, \ldots, x_n) > 0$ для любого ненулевого вектора $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$. Аналогично, $Q(x_1, x_2, \ldots, x_n)$ называется отрицательно определённой (положительно полуопределённой, отрицательно полуопределённой), если $Q(x_1, x_2, \ldots, x_n) < 0$ $(Q(x_1, x_2, \ldots, x_n) \geq 0$, $Q(x_1, x_2, \ldots, x_n) \leq 0$, соответственно). В обратном случае квадратичная форма называется неопределённой.

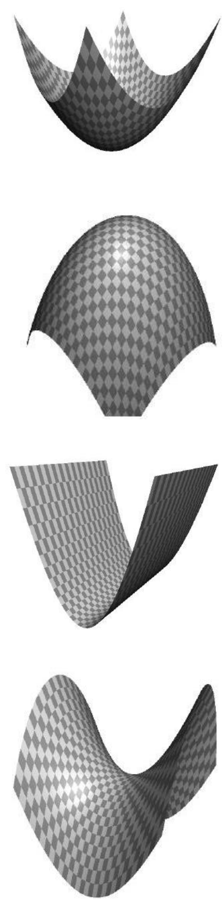
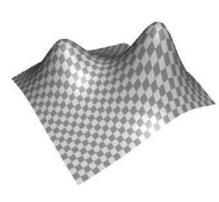


Рис. 2: 3D форма (a) положительно определённой, (b) неотрицательно определённой, (c) отрицательно определённой, и (d) неопределённой квадратичной функции.

На рисунке 2 представлены простые примеры поверхностей, задаваемых квадратичными формами. Во-первых, если $z = f(x, y)$ – квадратичная форма, то её сечения $z(y) = f(0, y)$ и $z(x) = f(x, 0)$ должны быть квадратичными функциями.

Например, 3D-график функции $z = x^2 + y^2$, приведённый на рис. a) – это поверхность имеющая кросс-секции (сечения плоскостями) $x = 0, \left(z = y^2\right)$ и $y = 0, \left(z = x^2\right)$. Квадратичная форма $z = x^2 + y^2$ является положительно определённой, поскольку она принимает только положительные значения при ненулевых аргументах (x, y), и равна нулю только если $x = y = 0$. Если z может принимать ненулевые значения при ненулевых (x, y), то такая квадратичная форма называется неотрицательно определенной, рис. (b). В качестве примера такой функции можно взять $z = x^2 - 2xy + y^2 = (x - y)^2 \geq 0$, поскольку она равна нулю при $x = y = 1$. Аналогично отрицательно-определёнаня квадратичная функция $z = -x^2 - y^2$ имеет сечения $x = 0 \left(z = -y^2\right)$ и $y = 0 \left(z = -x^2\right)$, но в этот раз кривые являются вогнутыми, рис. c). Неопределённая квадратичная форма – такая, что некоторые её сечения выпуклы, а некоторые – вогнуты. Примером является $z = x^2 - y^2$ – поскольку $z = x^2$ при $y = 0$ и $z = -y^2$ при $x = 0$.



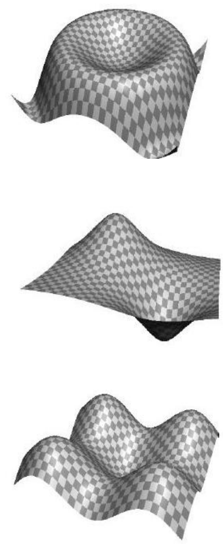


Рис. 3: Примеры экстремальных точек на обычных поверхностях. (а) два локальных максимума, разделенных седловой точкой, (б) локальный максимум и локальный минимум, без седловой точки между ними, (в) один локальный минимум и вырожденные локальные максимумы и минимумы, (г) четыре локальных максимума, один локальный минимум и четыре седловые точки.

Одной из характеристических точек на поверхности неопределенной квадратичной функции является седловая точка. Седловая точка имеет нулевой наклон, но поверхность выпукла в одном направлении и вогнута в другом. Это точка, расположенная прямо в середине поверхности, показанной на рисунке 2(d). Седловую точку можно рассматривать как нестабильное равновесие: она является локальным минимумом в одном сечении, но в то же время локальным максимумом в другом сечении. На обычных поверхностях может быть различное количество максимумов, минимумов и седловых точек, и квадратичные формы часто используются в качестве инструмента для приближенного описания локального поведения обычных поверхностей. Например, поверхность, показанная на рисунке (а), может быть описана как два холма и две соответствующие долины. “Кратчайший” путь от одной долины к другой проходит через седловую точку. Седловые точки играют важную роль в многомерном исчислении и в теории обыкновенных дифференциальных уравнений.

6.2 Замена базиса

В предыдущем разделе мы видели, как матрица линейного оператора изменялась в соответствии с правилом $A \mapsto CAC^{-1}$ при линейной трансформации векторного пространства под действием матрицы C. Давайте рассмотрим, как матрица квадратичной формы изменяется в аналогичном случае.

Рассмотрим два базиса, $e_1, e_2, \ldots, e_n$ и $e'_1, e'_2, \ldots, e'_n$ линейного пространства V и обозначим за $x = x_1e_1 + x_2e_2 + \ldots + x_ne_n = x'_1e'_1 + x'_2e'_2 + \ldots + x'_ne'_n$ два разложения по этим базисам вектора x. Новые координаты выражаются через старые следующим образом:

$$\begin{cases} x'_1 = c_{11}x_1 + c_{12}x_2 + \ldots + c_{1n}x_n \\ x'_2 = c_{12}x_1 + c_{22}x_2 + \ldots + c_{2n}x_n \\ x'_n = c_{1n}x_1 + c_{n2}x_2 + \ldots + c_{nn}x_n \end{cases}$$

Здесь матрица C состоит из координат старых векторов в новом базисе, т.е., $x' = Cx$.

Пусть A и A' – матрицы квадратичной формы в старом и новом базисе, соответственно. Имеем

$$Q(x) = (x')^T A'x' = (Cx)^T A'Cx = x^T \left(C^T A'C\right) x = x^T Ax.$$

Таким образом, матрица квадратичной формы преобразуется как $A \to C^T AC$, где C – матрица перехода.

6.3 Каноническая диагональная форма

Теорема. Пусть $Q(x)$ – квадратичная форма с n переменными. Существует замена базиса в пространстве, такая, что

$$Q(x) = x_1^2 + x_2^2 + \ldots + x_k^2 - x_{k+1}^2 - x_{k+2}^2 - \ldots - x_{k+l}^2$$

Доказательство. Предположим, что $Q(x_1, x_2, \ldots, x_n) = \sum_{i=1}^n a_{ii}x_i^2 + 2\sum_{i<j} a_{ij}x_ix_j$, а также $a_{11} \neq 0$. Будем искать замену координат в форме $x_1 = p_1y_1 + p_2y_2 + \ldots + p_ny_n$ и $x_i = y_i$ для $i \geq 2$. Получаем $Q(x_1, x_2, \ldots, x_n) = a_{11}x_1^2 + 2a_{12}x_1x_2 + \ldots + 2a_{1n}x_1x_n + \sum_{i=2}^n a_{ii}x_i^2 + 2\sum_{2\leq i<j} a_{ij}x_ix_j = a_{11}x_1^2 + 2x_1(a_{12}x_2 + \ldots + a_{1n}x_n) + R(x_2, \ldots, x_n)$, где $R(x_2, \ldots, x_n)$ – квадратичная форма с $n-1$ переменной.

Далее, $Q(y) = a_{11}(p_1y_1 + p_2y_2 + \ldots + p_ny_n)^2 + 2p_1y_1(a_{12}y_2 + \ldots + a_{1n}y_n) + R_1(y_2, \ldots, y_n)$, где члены, содержащие $y_2, \ldots, y_n$ оказались в $R_1(y_2, \ldots, y_n)$. Тогда, $Q(y) = a_{11}\left(p_1^2y_1^2 + 2p_1p_2y_1y_2 + 2p_1p_3y_1y_3 + \ldots + 2p_1p_ny_1y_n\right) + 2(p_1a_{12}y_1y_2 + p_1a_{13}y_1y_3 + \ldots + p_1a_{1n}y_1y_n) + R_2(y_2, \ldots, y_n)$, где $R_2(y_2, \ldots, y_n)$ содержит все компоненты с $y_2, \ldots, y_n$. Если обозначить $p_2 = -\frac{a_{12}}{a_{11}}, p_3 = -\frac{a_{13}}{a_{11}}, \ldots, p_n = \frac{-a_{1n}}{a_{11}}$, и $p_1 = \frac{1}{\sqrt{|a_{11}|}}$, то $Q(y) = \pm y_1^2 + R_1(y_2, \ldots, y_n)$, где знак коэффициента при y_1^2 – такой же, как и знак a_{11}.

Теперь рассмотрим случай $a_{11} = 0$, но хотя бы один из $a_{1i} \neq 0$. Без потери общности, предположим, что $a_{12} \neq 0$. Расммотрим изменение координат в форме $x_1 = y_1 + y_2, x_2 = y_1 - y_2$, и $x_i = y_i$ для $i \geq 3$. Мы снова вернулись к случаю, когда $a_{11} \neq 0$, поскольку $Q(y) = (y_1 + y_2)(y_1 - y_2) + \ldots = y_1^2 - y_2^2 + \ldots$ содержит y_1^2.

Кроме того, если $a_{1i} = 0$ для всех i, можно переходить к квадратичной форме с меньшим числом переменных.

Пара чисел (k, l), которые представляют собой количество положительных и отрицательных членов в каноническом представлении, соответственно, определена уникальным образом и называется сигнатурой квадратичной формы. Когда в канонической форме нет нулевых членов (т.е. $k + l = n$), и известна размерность nn векторного пространства, можно ограничиться знанием разности $k - l$, которая иногда также называется сигнатурой.[6] Например, сигнатура $(3, 1)$ в $\mathbb{R}^4$ также называется "сигнатурой 2". Больше не будем обращать внимания на различия между этими двумя формулировками.

Доказательство этой теоремы содержит процедуру, называемую выделением полных квадратов, которая может использоваться для поиска канонического базиса.

Пример. Найдите каноническую форму

$$Q(x_1, x_2, x_3) = x_1^2 - 4x_1x_2 + 2x_1x_3 + 6x_2^2 + 2x_3^2$$

Решение. $Q(x_1, x_2, x_3) = \left(x_1^2 - 4x_1x_2 + 4x_2^2\right) + 2x_1x_3 + 2x_2^2 + 2x_3^2 = (x_1 - 2x_2)^2 + 2x_1x_3 + 2x_2^2 + 2x_3^2$. Обозначим $y_1 = x_1 - 2x_2, y_2 = x_2$, и $y_3 = x_3$. Тогда, $x_1 = y_1 + 2y_2$ и получаем $Q(y_1, y_2, y_3) = y_1^2 + 2(y_1 + 2y_2)y_3 + 2y_2^2 + 2y_3^2 = y_1^2 + 2y_1y_3 + 4y_2y_3 + 2y_2^2 + 2y_3^2 = \left(y_1^2 + 2y_1y_3 + y_3^2\right) + 4y_2y_3 + 2y_2^2 + y_3^2 = (y_1 + y_3)^2 + 4y_2y_3 + 2y_2^2 + y_3^2$. Теперь обозначим $y_1 + y_3 = z_1, y_2 = z_2$ и $y_3 = z_3$. Получили $Q(z_1, z_2, z_3) = z_1^2 + 2z_2^2 + 4z_2z_3 + z_3^2 = z_1^2 + 2\left(z_2^2 + 2z_2z_3 + z_3^2\right) - z_3^2 = z_1^2 + 2(z_2 + z_3)^2 - z_3^2$. В итоге, обозначим $t_1 = z_1, t_2 = z_2 + z_3$ и $t_3 = z_3$. Теперь $Q(z_1, z_2, z_3) = t_1^2 + 2t_2^2 - t_3^2$. Поэтому, сигнатура этой квадратичной формы – это $(2, 1)$.

Теорема.25. Квадратичная форма $Q(x_1, x_2, \ldots, x_n)$ является положительно определённой (отрицательно определённой, положительно полуопределённой, отрицательно полуопределённой) тогда и только тогда, когда она имеет сигнатуру $(n, 0)$ (сигнатуру $(0, n), (k, 0)$ и $(0, k)$ соответственно, где $k < n$). Квадратичная

[6] Такое обозначение часто встречается в учебниках по физике.

форма $Q(x_1, x_2, \ldots, x_n)$ – неопределённая тогда и только тогда, когда она имеет сигнатуру (k, l), где $k > 0$ и $l > 0$.

Доказательство утверждения следует из определения.

6.4 Критерий Сильвестра

Иногда выделение полных квадратов – не самый удобный способ проверки знакоопределённости матрицы. Следующая теорема предлагает быстрый способ проверки, не требующий замены системы координат.

Теорема (критерий Сильвестра). Пусть $Q(x) = x^T Ax$ – квадратичная форма с n переменными, у которой $A = A^T$. Форма $Q(x)$ является положительно определённой тогда и только тогда, когда все её главные миноры (т.е. миноры, левый-верхний элемент которых – это левый-верхний элемент A) больше нуля, т.е.

$$\Delta_1 = a_{11} > 0$$

$$\Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0,$$

$$\Delta_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} > 0, \ldots$$

$$\Delta_n = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} > 0$$

Доказательство теоремы следует из следующей леммы.

Лемма. Пусть $x_1 = p_1 y_1 + p_2 y_2 + \ldots + p_n y_n$ и $x_i = y_i$ для $i \geq 2$ – замена системы координат, приводящая квадратичную форму к каноническому виду. Все главные миноры A размером не изменятся при $p_1 = 1$.

Следствие Пусть $Q(x) = x^T Ax$ – квадратичная форма с n переменным и пусть $\lambda_1 x_1^2 + \lambda_2 x_2^2 + \ldots + \lambda_n x_n^2$ – её диагональное представление, получаемое выделением полных квадратов $x_1 = p_1 y_1 + p_2 y_2 + \ldots + p_n y_n$, $x_i = y_i$ для $i \geq 2$ с $p_1 = 1$. Предположим, что $\Delta_i \neq 0$ для всех i. Тогда

$$\lambda_i = \frac{\Delta_i}{\Delta_{i-1}}$$

где $\Delta_0 = 1$.

Пример. Рассмотрим $Q(x_1, x_2, x_3) = x_1^2 - 4x_1 x_2 + 2x_1 x_3 + 6x_2^2 + 2x_3^2$.

$$A = \begin{pmatrix} 1 & -2 & 1 \\ -2 & 6 & 0 \\ 1 & 0 & 2 \end{pmatrix},$$

$$\Delta_1 = 1, \quad \Delta_2 = \begin{vmatrix} 1 & -2 \\ -2 & 6 \end{vmatrix} = 6 - 4 = 2,$$

$$\Delta_3 = \begin{vmatrix} 1 & -2 & 1 \\ -2 & 6 & 0 \\ 1 & 0 & 2 \end{vmatrix} = 12 - 6 - 8 = -2,$$

$$\lambda_1 = 1, \quad \lambda_2 = \frac{2}{1} = 2, \quad \lambda_3 = \frac{-2}{2} = -1,$$

в соответствии с канонической формой, полученной в предыдущем примере.

Заметим, что каждая положительно определённая квадратичная форма $Q(x_1, x_2, \ldots, x_n)$ превращается в отрицательно определённую при умножении на -1 (доказательство: если $Q(x_1, x_2, \ldots, x_n) > 0$ для $(x_1, x_2, \ldots, x_n) \neq 0$, то $-Q(x_1, x_2, \ldots, x_n) < 0$ при $(x_1, x_2, \ldots, x_n) \neq 0$). Учитывая, что матрица квадратичной формы при этом тоже умножается на -1,

$$\Delta_n = \begin{vmatrix} -a_{11} & -a_{12} & \ldots & -a_{1n} \\ -a_{21} & -a_{22} & \ldots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \ldots & -a_{nn} \end{vmatrix} = (-1)^{n^2} \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix},$$

получаем, что $\Delta_n(-Q) = (-1)^{n^2}\Delta_n(Q) = (-1)^n\Delta_n(Q)$, поскольку n^2 – чётное число тогда и только тогда, когда n чётное. Таким образом, получили формулировку теоремы для отрицательно-определённых квадратичных форм.

Теорема. Пусть $Q(x) = x^T Ax$ – квадратичная форма с n переменными, где $A = A^T$. Форма $Q(x)$ является отрицательно-определённой тогда и только тогда, когда главные миноры размера $k \times k$ матрицы A образуют знакочередующуюся последовательность, начинающуюся с отрицательного числа, т.е. $\Delta_1 < 0, \Delta_2 > 0, \Delta_3 < 0, \ldots$ и т. д.

Критерий Сильвестра – удобный инструмент для анализа квадратичных форм, но не всегда позволяющий получить однозначный ответ, поскольку какие-то из главных миноров могут оказаться равными нулю. В этом случае необходимо делать процедуру выделения полных квадратов.

Пример. Пусть $Q(x_1, x_2) = x_1 x_2$. Критерий Сильвестра не позволяет понять знакоопределённость $Q(\cdot)$, поскольку $\Delta_1 = 0$. Однако, линейное преобразование $x_1 = y_1 - y_2$ и $x_2 = y_1 + y_2$ приводит $Q(y_1, y_2) = (y_1 - y_2)(y_1 + y_2) = y_1^2 - y_2^2$, откуда ясно, что квадратичная форма $Q(x_1, x_2)$ является неопределённой.

6.5 Упражнения

1. Используйте процедуру выделения полных квадратов для того, чтобы привести квадратичную форму к каноническому виду.

a) $-4x_1^2 + 8x_1x_2 + 10x_1x_3 - 2x_2^2 - 10x_2x_3 - 5x_3^2$

b) $-x_1^2 - 6x_1x_2 - 2x_1x_3 + 4x_2^2 + 8x_2x_3 + 2x_3^2$

c) $-5x_1^2 - 4x_1x_2 + 10x_1x_3 - 2x_2^2 + 10x_2x_3 - 5x_3^2$

d) $x_1^2 + 4x_1x_2 + 10x_1x_3 + 2x_2^2 + 8x_2x_3 + 5x_3^2$

e) $2x_1^2 - 4x_1x_2 + 8x_1x_3 - 4x_1x_4 - x_2^2 + 4x_2x_3 + 4x_2x_4 - x_3^2 - 2x_3x_4$

f) $-x_1^2 - 6x_1x_2 - 6x_1x_3 - 8x_1x_4 + 5x_2^2 + 6x_2x_4 + 3x_3^2 + 6x_3x_4 + 4x_4^2$

g) $x_1^2 - 4x_1x_2 - 2x_1x_3 + 5x_2^2 - 4x_2x_3 + 4x_2x_4 - x_3^2 - 4x_3x_4 + 4x_4^2$

h) $2x_1^2 + 6x_1x_2 - 2x_1x_3 - 8x_1x_4 - 3x_2^2 - 6x_2x_3 + x_3^2 + 8x_3x_4 + 5x_4^2$

2. Используйте критерий Сильвестра чтобы проверить определённость квадратических форм.

a) $5x_1^2 - 6x_1x_2 - 6x_1x_3 + 3x_2^2 + 6x_2x_3 + 3x_3^2$

b) $3x_1^2 - 6x_1x_2 + 8x_1x_3 + x_2^2 - 6x_2x_3 + 4x_3^2$

c) $-3x_1^2 - 6x_1x_2 + 6x_1x_3 + 6x_2x_3 + x_3^2$

d) $-x_1^2 + 8x_1x_2 - 8x_1x_3 - 4x_2^2 + 8x_2x_3 - x_3^2$

e) $-x_1^2 - 2x_1x_2 - 8x_1x_3 - 6x_1x_4 - 2x_2^2 + 2x_2x_3 + 4x_3^2 + 6x_3x_4 - 3x_4^2$

f) $-2x_1^2 - 4x_1x_2 - 4x_1x_3 + 2x_2^2 + 4x_2x_3 - 8x_2x_4 + x_3^2 - 10x_3x_4 + x_4^2$

g) $-x_1^2 - 2x_1x_3 - 4x_2^2 - 8x_2x_4 - 3x_3^2 - 4x_3x_4 - 9x_4^2$

h) $-4x_2^2 + 16x_2x_4 - 3x_3^2 - 20x_4^2$

6.6 Евклидово пространство

Евклидово пространство – это линейное пространство с операцией скалярного произведения векторов.

Определение. Билинейная форма на векторном пространстве V – это отображение $b : V \times V \to \mathbb{R}$ со следующими свойствами:

1. $\forall x_1, x_2, y \in V \quad b\,(x_1 + x_2, y) = b\,(x_1, y) + b\,(x_2, y)$,

2. $\forall x, y_1, y_2 \in V \quad b\,(x, y_1 + y_2) = b\,(x, y_1) + b\,(x, y_2)$,

3. $\forall x, y \in V, \lambda \in \mathbb{R} b(\lambda x, y) = b(x, \lambda y) = \lambda b(x, y)$.

Другими словами, $b(x, y)$ – это функция, линейная по обоим своим аргументам. Билинейная форма $b(x, y)$ называется симметричной, если

4. $\forall x, y \in V \quad b(x, y) = b(y, x)$.

Билинейная форма $b(x, y)$ найзывается положительно определённой, если

5. $\forall x \in V \quad b(x, x) \geq 0$ и $b(x, x) = 0$ только при $x = 0$.

Аналогично тому, что обсуждалось в предыдущем разделе, каждая билинейная форма в конечномерном пространстве допускает представление в виде матричного произведения:

$$b(x, y) = x^T By,$$

где $x, y \in \mathbb{R}^n$, и B – $n \times n$ матрица, которая называется матрицей билинейной формы. Очевидно, что билинейная форма симметрична тогда и только тогда, когда её матрица симметрична. Изменение базисаа билинейной формы происходит аналогично тому, что мы видели для квадратичных форм. Если $x = Cx'$ то $b(x, y) = x^T By = (x')^T C^T BCx'$, т.е., матрица билинейной формы также преобразуется согласно правилу $B \mapsto C^T BC$.

Каждая симметричная билинейная форма $b(x, y)$ определяет квадратичную форму с той же матрицей выражением $Q(x) = b(x, x)$. С другой стороны, если $Q(x) = x^T Ax$ – квадратичная форма, то можно построить билинейную форму с той же матрицей, заданную выражением $b(x, y) = x^T Ay$. Это можно сделать даже не используя матрицы, поскольку $b(x+y, x+y) = b(x, x)+b(x, y)+b(y, x)+b(y, y) = Q(x)+Q(y)+2b(x, y) = Q(x+y)$ и

$$b(x, y) = \frac{1}{2}\big(Q(x+y) - Q(x) - Q(y)\big).$$

Поэтому каждая вадратичная форма определяет симметричную билинейную и наоборот, симметричная билинейная форма определяет квадратичную.