

Семинар 1

Системы линейных уравнений

Наша задача научиться решать Системы Линейных Уравнений (СЛУ), то есть находить все их решения или доказывать, что решений нет. Общий вид СЛУ и ее однородная версия (ОСЛУ):

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ \dots \\ a_{m1}x_1 + \dots + a_{mn}x_n = b_m \end{cases} \quad \begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = 0 \\ \dots \\ a_{m1}x_1 + \dots + a_{mn}x_n = 0 \end{cases}$$

Коэффициенты

Где живут коэффициенты a_{ij} и b_j ? Варианты:

- Вещественные числа \mathbb{R}
- Комплексные числа \mathbb{C}
- Рациональные числа \mathbb{Q}

Для решения СЛУ **НЕ** имеет значения откуда берутся коэффициенты, так как решения будут лежать там же. Потому мы будем работать с числами из \mathbb{R} .

Матрицы связанные со СЛУ

Для каждой СЛУ введем следующие обозначения:

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad (A|b) = \left(\begin{array}{ccc|c} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{array} \right)$$

Названия:

- A – матрица системы
- b – вектор правой части
- $(A|b)$ – расширенная матрица системы
- x – вектор решений

Будем кратко записывать СЛУ и ее однородную версию так: $Ax = b$ и $Ax = 0$.

Количество решений

Случай одного уравнения и одной неизвестной

- $x = 0$ – одно решение
- $0x = 0$ – бесконечное число решений
- $0x = 1$ – нет решений

Элементарные преобразования

$$\begin{aligned}
 \text{I тип: } & \left(\begin{array}{ccc|c} a_{11} & \dots & a_{1n} & b_1 \\ a_{i1} & \dots & a_{in} & b_i \\ a_{j1} & \dots & a_{jn} & b_j \\ a_{m1} & \dots & a_{mn} & b_m \end{array} \right) \mapsto \left(\begin{array}{ccc|c} a_{11} & \dots & a_{1n} & b_1 \\ a_{i1} & \dots & a_{in} & b_i \\ a_{j1} + \lambda a_{i1} & \dots & a_{jn} + \lambda a_{in} & b_j + \lambda b_i \\ a_{m1} & \dots & a_{mn} & b_m \end{array} \right) \quad i \neq j \\
 \text{II тип: } & \left(\begin{array}{ccc|c} a_{11} & \dots & a_{1n} & b_1 \\ a_{i1} & \dots & a_{in} & b_i \\ a_{j1} & \dots & a_{jn} & b_j \\ a_{m1} & \dots & a_{mn} & b_m \end{array} \right) \mapsto \left(\begin{array}{ccc|c} a_{11} & \dots & a_{1n} & b_1 \\ a_{j1} & \dots & a_{jn} & b_j \\ a_{i1} & \dots & a_{in} & b_i \\ a_{m1} & \dots & a_{mn} & b_m \end{array} \right) \\
 \text{III тип: } & \left(\begin{array}{ccc|c} a_{11} & \dots & a_{1n} & b_1 \\ a_{i1} & \dots & a_{in} & b_i \\ a_{m1} & \dots & a_{mn} & b_m \end{array} \right) \mapsto \left(\begin{array}{ccc|c} a_{11} & \dots & a_{1n} & b_1 \\ \lambda a_{i1} & \dots & \lambda a_{in} & \lambda b_i \\ a_{m1} & \dots & a_{mn} & b_m \end{array} \right) \quad \lambda \neq 0
 \end{aligned}$$

Приведение к ступенчатому виду (алгоритм Гаусса)

Основной способ решения СЛУ – привести ее элементарными преобразованиями к простому виду, где множество решений очевидно.¹

Разберем типичный ход алгоритма Гаусса на примере 3 уравнений и 4 неизвестных.²

Прямой ход алгоритма Гаусса

$$\begin{aligned}
 & \left(\begin{array}{cccc|c} a_{11} & a_{12} & a_{13} & a_{14} & b_1 \\ a_{21} & a_{22} & a_{23} & a_{24} & b_2 \\ a_{31} & a_{32} & a_{33} & a_{34} & b_3 \end{array} \right) \quad \begin{array}{l} \text{2-я строка} - \frac{a_{21}}{a_{11}} \cdot \text{1-я строка} \\ \text{3-я строка} - \frac{a_{31}}{a_{11}} \cdot \text{1-я строка} \end{array} \\
 & \left(\begin{array}{cccc|c} a_{11} & a_{12} & a_{13} & a_{14} & b_1 \\ 0 & a_{22} & a_{23} & a_{24} & b_2 \\ a_{31} & a_{32} & a_{33} & a_{34} & b_3 \end{array} \right) \quad \begin{array}{l} \text{3-я строка} - \frac{a_{31}}{a_{11}} \cdot \text{1-я строка} \\ \text{3-я строка} - \frac{a_{32}}{a_{22}} \cdot \text{2-я строка} \end{array} \\
 & \left(\begin{array}{cccc|c} a_{11} & a_{12} & a_{13} & a_{14} & b_1 \\ 0 & a_{22} & a_{23} & a_{24} & b_2 \\ 0 & a_{32} & a_{33} & a_{34} & b_3 \end{array} \right) \quad \begin{array}{l} \text{3-я строка} - \frac{a_{32}}{a_{22}} \cdot \text{2-я строка} \\ \text{3-я строка} - \frac{a_{33}}{a_{23}} \cdot \text{2-я строка} \end{array} \\
 & \left(\begin{array}{cccc|c} a_{11} & a_{12} & a_{13} & a_{14} & b_1 \\ 0 & a_{22} & a_{23} & a_{24} & b_2 \\ 0 & 0 & a_{33} & a_{34} & b_3 \end{array} \right)
 \end{aligned}$$

В результате данного хода какие-то коэффициенты, например a_{33} , могли занулиться, потому возможны следующие принципиально другие случаи³

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & a_{13} & a_{14} & b_1 \\ 0 & a_{22} & a_{23} & a_{24} & b_2 \\ 0 & 0 & 0 & \underline{a_{34}} & b_3 \end{array} \right) \quad \left(\begin{array}{cccc|c} a_{11} & a_{12} & a_{13} & a_{14} & b_1 \\ 0 & 0 & a_{23} & a_{24} & b_2 \\ 0 & 0 & 0 & \underline{a_{34}} & b_3 \end{array} \right) \quad \left(\begin{array}{cccc|c} a_{11} & a_{12} & a_{13} & a_{14} & b_1 \\ 0 & a_{22} & a_{23} & a_{24} & b_2 \\ 0 & 0 & 0 & 0 & \underline{b_3} \end{array} \right) \quad \left(\begin{array}{cccc|c} a_{11} & a_{12} & a_{13} & a_{14} & b_1 \\ 0 & a_{22} & a_{23} & a_{24} & b_2 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

Подчеркнутые элементы считаются не равными нулю. В ступенчатом виде все переменные (и соответственно коэффициенты перед ними) делятся на главные и неглавные. Главные коэффициенты – это первые ненулевые коэффициенты в строке (подчеркнутые). Переменные при них называются главными, остальные ненулевые коэффициенты и переменные – неглавные или свободными.

¹ Данный метод является самым быстрым возможным как для написания программ, так и для ручного вычисления. При вычислениях руками, однако, полезно местами пользоваться «локальными оптимизациями», то есть, если вы видите, что какая-то хитрая комбинация строк сильно упростит вид системы, то сделайте ее.

² При переходе от одной матрицы к другой я новым коэффициентам даю старые имена, чтобы не захламлять текст новыми обозначениями.

³ Это не полный список всех случаев.

Обратный ход алгоритма Гаусса

Разберем типичный обратный ход алгоритма Гаусса. Подчеркнутые элементы считаются не равными нулю.

$$\begin{pmatrix} \underline{a_{11}} & a_{12} & a_{13} & a_{14} & | & b_1 \\ 0 & \underline{a_{22}} & a_{23} & a_{24} & | & b_2 \\ 0 & 0 & \underline{a_{33}} & a_{34} & | & b_3 \end{pmatrix} \quad \text{разделить } i\text{-ю строку на } a_{ii}$$

$$\begin{pmatrix} 1 & a_{12} & a_{13} & a_{14} & | & b_1 \\ 0 & 1 & a_{23} & a_{24} & | & b_2 \\ 0 & 0 & 1 & a_{34} & | & b_3 \end{pmatrix} \quad \text{2-я строка} - a_{23} \cdot \text{3-я строка}$$

$$\begin{pmatrix} 1 & a_{12} & a_{13} & a_{14} & | & b_1 \\ 0 & 1 & 0 & a_{24} & | & b_2 \\ 0 & 0 & 1 & a_{34} & | & b_3 \end{pmatrix} \quad \text{1-я строка} - a_{13} \cdot \text{3-я строка}$$

$$\begin{pmatrix} 1 & a_{12} & 0 & a_{14} & | & b_1 \\ 0 & 1 & 0 & a_{24} & | & b_2 \\ 0 & 0 & 1 & a_{34} & | & b_3 \end{pmatrix} \quad \text{1-я строка} - a_{12} \cdot \text{2-я строка}$$

$$\begin{pmatrix} 1 & 0 & 0 & a_{14} & | & b_1 \\ 0 & 1 & 0 & a_{24} & | & b_2 \\ 0 & 0 & 1 & a_{34} & | & b_3 \end{pmatrix}$$

В специальных случаях приведенных выше, получим

$$\begin{pmatrix} 1 & 0 & a_{13} & 0 & | & b_1 \\ 0 & 1 & a_{23} & 0 & | & b_2 \\ 0 & 0 & 0 & 1 & | & b_3 \end{pmatrix} \quad \begin{pmatrix} 1 & a_{12} & 0 & 0 & | & b_1 \\ 0 & 0 & 1 & 0 & | & b_2 \\ 0 & 0 & 0 & 1 & | & b_3 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & a_{13} & a_{14} & | & 0 \\ 0 & 1 & a_{23} & a_{24} & | & 0 \\ 0 & 0 & 0 & 0 & | & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & a_{13} & a_{14} & | & b_1 \\ 0 & 1 & a_{23} & a_{24} & | & b_2 \\ 0 & 0 & 0 & 0 & | & 0 \end{pmatrix}$$

Полученный в результате обратного хода вид расширенной матрицы называется улучшенным ступенчатым видом, т.е., это ступенчатый вид, где все коэффициенты при главных неизвестных – единицы, и все коэффициенты над ними равны нулю.

t.me/postypashki_old/1229

t.me/postypashki_old/1229

t.me/postypashki_old/1229

Получение решений

В системе ниже, выберем переменную x_4 как параметр

$$\begin{pmatrix} 1 & 0 & 0 & a_{14} & | & b_1 \\ 0 & 1 & 0 & a_{24} & | & b_2 \\ 0 & 0 & 1 & a_{34} & | & b_3 \end{pmatrix}$$

Тогда решения имеют вид⁴

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} - x_4 \begin{pmatrix} a_{14} \\ a_{24} \\ a_{34} \end{pmatrix}$$

Специальные случаи:

$$\begin{pmatrix} 1 & 0 & a_{13} & 0 & | & b_1 \\ 0 & 1 & a_{23} & 0 & | & b_2 \\ 0 & 0 & 0 & 1 & | & b_3 \end{pmatrix} \quad \text{Решения:} \quad \begin{pmatrix} x_1 \\ x_2 \\ x_4 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} - x_3 \begin{pmatrix} a_{13} \\ a_{23} \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & a_{12} & 0 & 0 & | & b_1 \\ 0 & 0 & 1 & 0 & | & b_2 \\ 0 & 0 & 0 & 1 & | & b_3 \end{pmatrix} \quad \text{Решения:} \quad \begin{pmatrix} x_1 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} - x_2 \begin{pmatrix} a_{12} \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & a_{13} & a_{14} & | & 0 \\ 0 & 1 & a_{23} & a_{24} & | & 0 \\ 0 & 0 & 0 & 0 & | & 1 \end{pmatrix} \quad \text{Решения:} \quad \text{Нет решений, т.к. последнее уравнение } 0 = 1$$

$$\begin{pmatrix} 1 & 0 & a_{13} & a_{14} & | & b_1 \\ 0 & 1 & a_{23} & a_{24} & | & b_2 \\ 0 & 0 & 0 & 0 & | & 0 \end{pmatrix} \quad \text{Решения:} \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} - x_3 \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix} - x_4 \begin{pmatrix} a_{14} \\ a_{24} \end{pmatrix}$$

⁴Операция умножения матрицы на число покомпонентная (умножаем каждый элемент на число). Сумма и разность двух матриц покомпонентная (складываем или вычитаем числа на одних и тех же позициях).

Вычислительная практика

Найти решения СЛУ соответствующих следующим расширенным матрицам:

$$\left(\begin{array}{ccc|c} 2 & 1 & 1 & 3 \\ 1 & 2 & 1 & 0 \\ 1 & 1 & 2 & 0 \end{array}\right) \quad \left(\begin{array}{cc|c} 9 & 3 & 5 \\ 7 & 5 & 3 \end{array}\right) \quad \left(\begin{array}{ccc|c} 1 & 1 & -1 & a \\ 1 & -1 & 1 & b \\ -1 & 1 & 1 & c \end{array}\right) \quad \left(\begin{array}{ccc|c} 1 & 1 & 2 & 0 \\ -1 & 0 & -1 & -1 \\ 1 & 1 & 2 & \lambda \end{array}\right) \quad \left(\begin{array}{ccc|c} \lambda & 1 & 1 & 1 \\ 1 & \lambda & 1 & 1 \\ 1 & 1 & \lambda & 1 \end{array}\right) \quad \left(\begin{array}{ccc|c} 1 & -1 & -2 & 0 \\ 0 & -1 & -1 & 0 \\ 1 & 2 & 1 & 1 \end{array}\right)$$

Замечание. Работая с целочисленными матрицами, старайтесь во время прямого хода алгоритма Гаусса не выходить за рамки целых чисел.

- Используйте элементарные преобразования I типа только с целым параметром.
- Полезно не злоупотреблять умножением на ненулевое целое, умножайте только на ± 1 . Иначе придется работать с большими числами.

На этапе обратного хода алгоритма Гаусса избавиться от деления уже не возможно.

Широкие системы Пусть вам дана однородная система из m уравнений с n неизвестными и $m < n$, тогда утверждается, что она обязательно имеет хотя бы одно ненулевое решение, то есть такое решение, где хотя бы одна переменная отличная от нуля. Действительно, так как у однородной системы всегда есть решение, то нам достаточно доказать, что у системы в улучшенном ступенчатом виде есть свободные переменные. Если это так, то придавая свободным переменным ненулевые значения, мы получим ненулевое решение. Чтобы понять, что есть свободные переменные, давайте оценим сверху количество главных. Главных переменных не больше чем строк, а строк строго меньше, чем всех переменных, вот и все.

Матрицы

Матрица – это прямоугольная таблица чисел

t.me/postypashki_old/1229

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}, \text{ где } a_{ij} \in \mathbb{R}$$

t.me/postypashki_old/1229

Множество всех матриц с m строками и n столбцами обозначается $M_{m,n}(\mathbb{R})$. Множество квадратных матриц размера n будем обозначать $M_n(\mathbb{R})$. Матрицы с одним столбцом или одной строкой называются векторами (вектор-столбцами и вектор-строками соответственно). Множество всех векторов с n координатами обозначается через \mathbb{R}^n . Мы по умолчанию считаем, что наши вектора – вектор-столбцы.⁵

Операции над матрицами

Сложение Пусть $A, B \in M_{m,n}(\mathbb{R})$. Тогда сумма $A + B$ определяется покомпонентно, т.е. $C = A + B$, то $c_{ij} = a_{ij} + b_{ij}$ или

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

Складывать можно только матрицы одинакового размера.

Умножение на скаляр Если $\lambda \in \mathbb{R}$ и $A \in M_{m,n}(\mathbb{R})$, то λA определяется так: $\lambda A = C$, где $c_{ij} = \lambda a_{ij}$ или

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

⁵Важно, directX и openGL используют вектор-строки! Потому часть инженерной литературы на английском связанной с трехмерной графикой оперирует со строками.

Умножение матриц Пусть $A \in M_{mn}(\mathbb{R})$ и $B \in M_{nk}(\mathbb{R})$, то произведение $AB \in M_{mk}(\mathbb{R})$ определяется так: $AB = C$, где $c_{ij} = \sum_{t=1}^n a_{it}b_{tj}$ или

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & \dots & b_{1k} \\ \vdots & \ddots & \vdots \\ b_{n1} & \dots & b_{nk} \end{pmatrix} = \begin{pmatrix} \sum_{t=1}^n a_{1t}b_{t1} & \dots & \sum_{t=1}^n a_{1t}b_{tk} \\ \vdots & \ddots & \vdots \\ \sum_{t=1}^n a_{mt}b_{t1} & \dots & \sum_{t=1}^n a_{mt}b_{tk} \end{pmatrix}$$

Умножение матриц полезно как-то визуализировать у себя в голове. Предположим мы хотим посчитать коэффициент c_{ij} . Тогда надо из матрицы A взять i -ю строку (она имеет длину n), а из матрицы B взять j -ый столбец (он тоже имеет длину n). Тогда их надо скалярно перемножить и результат подставить в c_{ij} , как показано ниже на картинке.

$$\begin{pmatrix} a_{11} & \dots & \dots & a_{1n} \\ \vdots & \ddots & & \vdots \\ \boxed{a_{i1} \quad \dots \quad a_{in}} \\ \vdots & & \ddots & \vdots \\ a_{m1} & \dots & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & \dots & \boxed{b_{1j}} & \dots & b_{1k} \\ \vdots & \ddots & \vdots & & \vdots \\ \vdots & & \vdots & & \vdots \\ \vdots & & \vdots & \ddots & \vdots \\ b_{n1} & \dots & \boxed{b_{nj}} & \dots & b_{nk} \end{pmatrix} = \begin{pmatrix} * & \dots & & \dots & * \\ \vdots & \ddots & & & \vdots \\ \vdots & & \boxed{\sum_{t=1}^n a_{it}b_{tj}} & & \vdots \\ \vdots & & \vdots & \ddots & \vdots \\ * & \dots & & \dots & * \end{pmatrix}$$

При этом умножение строки на столбец можно себе представлять так: мы приставляем строку к столбцу, потом поэлементно перемножаем их, а затем складываем все произведения вместе.

$$\begin{array}{c} \boxed{a_{i1} \quad \dots \quad a_{in}} \cdot \begin{array}{c} \boxed{b_{1j}} \\ \vdots \\ \boxed{b_{nj}} \end{array} \mapsto \begin{array}{c} \boxed{a_{i1} \quad \dots \quad a_{in}} \\ \boxed{b_{1j} \quad \dots \quad b_{nj}} \end{array} \mapsto \\ \mapsto \begin{array}{c} \boxed{a_{i1}} \\ \vdots \\ \boxed{b_{1j}} \end{array} + \dots + \begin{array}{c} \boxed{a_{in}} \\ \vdots \\ \boxed{b_{nj}} \end{array} = a_{i1}b_{1j} + \dots + a_{in}b_{nj} \end{array}$$

Транспонирование Пусть A – матрица вида

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \quad \text{или} \quad \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \quad \text{или} \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

Определим транспонированную матрицу $A^t = (a'_{ij})$ так: $a'_{ij} = a_{ji}$. Наглядно, транспонированная матрица для приведенных выше

$$\begin{pmatrix} a_{11} & \dots & a_{m1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \dots & a_{mn} \end{pmatrix} \quad \text{или} \quad \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \end{pmatrix} \quad \text{или} \quad (x_1 \quad x_2 \quad x_3)$$

След матрицы Пусть $A \in M_n(\mathbb{R})$, тогда определим след матрицы A , как сумму ее диагональных элементов: $\text{tr } A = \sum_{i=1}^n a_{ii}$. Давайте отметим следующие свойства следа:

1. Для любых матриц $A, B \in M_n(\mathbb{R})$ верно $\text{tr}(A + B) = \text{tr}(A) + \text{tr}(B)$.
2. Для любой матрицы $A \in M_n(\mathbb{R})$ и $\lambda \in \mathbb{R}$ выполнено $\text{tr}(\lambda A) = \lambda \text{tr}(A)$.
3. Для любых матриц $A \in M_{mn}(\mathbb{R})$ и $B \in M_{nm}(\mathbb{R})$ выполнено $\text{tr}(AB) = \text{tr}(BA)$.

Все эти свойства проверяются непосредственным вычислением по определению.

Специальные виды матриц

Ниже мы перечислим названия некоторых специальных классов матриц:

- $A = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix}$ – диагональная матрица. Все ненулевые элементы стоят на главной диагонали, то есть в позиции, где номер строки равен номеру столбца.
- $A = \begin{pmatrix} \lambda & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda \end{pmatrix}$ – скалярная матрица. Диагональная матрица с одинаковыми элементами на диагонали.

Замечание Пусть у нас задана система линейных уравнений с матрицей $A \in M_{m,n}(\mathbb{R})$ и правой частью $b \in \mathbb{R}^m$. Пусть $x = (x_1, \dots, x_n)$ – будет столбцом из переменных. Тогда систему $(A|b)$ можно задать в матричном виде $Ax = b$.

Свойства операций

Все три операции на матрицах обладают «естественными свойствами» и согласованы друг с другом. Вот перечень базовых свойств операций над матрицами:⁶

1. **Ассоциативность сложения** $(A + B) + C = A + (B + C)$ для любых $A, B, C \in M_{m,n}(\mathbb{R})$
2. **Существование нейтрального элемента для сложения** Существует единственная матрица $0 \in M_{m,n}(\mathbb{R})$ обладающая следующим свойством $A + 0 = 0 + A = A$ для всех $A \in M_{m,n}(\mathbb{R})$. Такая матрица целиком заполнена нулями.
3. **Коммутативность сложения** $A + B = B + A$ для любых $A, B \in M_{m,n}(\mathbb{R})$.
4. **Наличие обратного по сложению** Для любой матрицы $A \in M_{m,n}(\mathbb{R})$ существует матрица $-A$ такая, что $A + (-A) = (-A) + A = 0$. Такая матрица единственная и состоит из элементов $-a_{ij}$.
5. **Ассоциативность умножения** Для любых матриц $A \in M_{m,n}(\mathbb{R})$, $B \in M_{n,k}(\mathbb{R})$ и $C \in M_{k,t}(\mathbb{R})$ верно $(AB)C = A(BC)$.
6. **Существование нейтрального элемента для умножения** Для каждого k существует единственная матрица $E \in M_k(\mathbb{R})$ такая, что для любой $A \in M_{m,n}(\mathbb{R})$ верно $EA = AE = A$. У такой матрицы $E_{ii} = 1$, а $E_{ij} = 0$. Когда нет путаницы матрицу E обозначают через 1.
7. **Дистрибутивность умножения относительно сложения** Для любых матриц $A, B \in M_{m,n}(\mathbb{R})$ и $C \in M_{n,k}(\mathbb{R})$ верно $(A + B)C = AC + BC$. Аналогично, для любых $A \in M_{m,n}(\mathbb{R})$ и $B, C \in M_{n,k}(\mathbb{R})$ верно $A(B + C) = AB + AC$.
8. **Умножение на числа ассоциативно** Для любых $\lambda, \mu \in \mathbb{R}$ и любой матрицы $A \in M_{m,n}(\mathbb{R})$ верно $\lambda(\mu A) = (\lambda\mu)A$. Аналогично для любого $\lambda \in \mathbb{R}$ и любых $A \in M_{m,n}(\mathbb{R})$ и $B \in M_{n,k}(\mathbb{R})$ верно $\lambda(AB) = (\lambda A)B$.
9. **Умножение на числа дистрибутивно относительно сложения матриц и сложения чисел** Для любых $\lambda, \mu \in \mathbb{R}$ и $A \in M_{m,n}(\mathbb{R})$ верно $(\lambda + \mu)A = \lambda A + \mu A$. Аналогично, для любого $\lambda \in \mathbb{R}$ и $A, B \in M_{m,n}(\mathbb{R})$ верно $\lambda(A + B) = \lambda A + \lambda B$.
10. **Умножение на скаляр нетривиально** Если $1 \in \mathbb{R}$, то для любой матрицы $A \in M_{m,n}(\mathbb{R})$ верно $1A = A$.
11. **Умножение на скаляр согласовано с умножением матриц** Для любого $\lambda \in \mathbb{R}$ и любых $A \in M_{m,n}(\mathbb{R})$ и $B \in M_{n,k}(\mathbb{R})$ верно $\lambda(AB) = (\lambda A)B = A(\lambda B)$.

К этим свойствам надо относиться так. Доказывая что-то про матрицы, можно лезть внутрь определений операций над ними, а можно пользоваться свойствами операций. Так вот, список выше – это минимальный набор свойств операций, из которых можно вытащить базовую информацию про эти операции и при этом не лезть внутрь определений.

⁶Все эти свойства объединяет то, что они являются аксиомами в различных определениях для алгебраических структур. Позже мы столкнемся с такими структурами.

Нулевые строки и столбцы Пусть в матрице $A \in M_{m \times k}(\mathbb{R})$ i -я строка полностью состоит из нулей и нам дана матрица $B \in M_{k \times n}(\mathbb{R})$. Тогда в произведении AB i -я строка тоже будет нулевой. Изобразим это ниже графически

$$AB = \begin{pmatrix} * & * & \dots & * \\ * & * & \dots & * \\ 0 & 0 & \dots & 0 \\ * & * & \dots & * \end{pmatrix} \begin{pmatrix} * & * & \dots & * \\ * & * & \dots & * \\ * & * & \dots & * \\ * & * & \dots & * \end{pmatrix} = \begin{pmatrix} * & * & \dots & * \\ * & * & \dots & * \\ 0 & 0 & \dots & 0 \\ * & * & \dots & * \end{pmatrix}$$

Действительно, i -я строка произведения зависит от i -ой строки левого множителя (матрицы A) и всех столбцов B . Но умножая нулевую строку A на что угодно, получим нули в i -ой строке результата. Аналогичное утверждение верно для столбцов в матрице B , а именно. Пусть в матрице $B \in M_{k \times n}(\mathbb{R})$ i -ый столбец полностью состоит из нулей и нам дана матрица $A \in M_{m \times k}(\mathbb{R})$. Тогда в произведении AB i -ый столбец тоже будет нулевой.

$$AB = \begin{pmatrix} * & * & * & * \\ * & * & * & * \\ \vdots & \vdots & \vdots & \vdots \\ * & * & * & * \end{pmatrix} \begin{pmatrix} * & * & 0 & * \\ * & * & 0 & * \\ \vdots & \vdots & \vdots & \vdots \\ * & * & 0 & * \end{pmatrix} = \begin{pmatrix} * & * & 0 & * \\ * & * & 0 & * \\ \vdots & \vdots & \vdots & \vdots \\ * & * & 0 & * \end{pmatrix}$$

Нехорошие свойства операций

Матрицы как новые числа Рассмотрим множество квадратных матриц с введенными выше операциями: $(M_n(\mathbb{R}), +, -, \cdot, {}^t)$. Про это множество стоит думать как про новый вид чисел со своими операциями. Принципиальное отличие – нельзя делить на любую ненулевую матрицу, как это можно было делать с числами. У нас вообще пока нет понимания, что такое деление в матрицах. Однако, это не единственное отличие.

Аномалии матричных операций Матричные операции обладают несколькими аномалиями по сравнению со свойствами операций над обычными числами.

1. Существование вычитания следует из «хорошести» операции сложения. Она позволяет определить вычитание без проблем. Однако, операция умножения уже хуже, чем на обычных числах, потому не получится просто так определить на матрицах операцию деления. Про деление и как его можно было бы обобщать мы поговорим ниже.
2. Давайте обсудим порядок матриц в произведении. Тут может быть несколько проблем. Может так оказаться, что матрицы можно перемножить в одном порядке, но нельзя в другом. Например, $A \in M_{m \times n}(\mathbb{R})$, $B \in M_{n \times k}(\mathbb{R})$ и при этом $m \neq k$. В этом случае $AB \in M_{m \times k}(\mathbb{R})$, однако, произведение BA просто не определено из-за отсутствия условия согласованности матриц. Если мы рассмотрим матрицы, которые можно перемножить в обоих порядках, то есть $A \in M_{m \times n}(\mathbb{R})$, $B \in M_{n \times m}(\mathbb{R})$, то результаты будут $AB \in M_m(\mathbb{R})$ и $BA \in M_n(\mathbb{R})$. А значит при разных m и n это будут матрицы разного размера и не могут быть одинаковыми. Таким образом, чтобы AB могло совпасть с BA нам надо, чтобы матрицы A и B были квадратными. Но даже в этом случае умножение матриц НЕ коммутативно. Действительно, вот простой пример в малой размерности

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{но} \quad \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

3. В матрицах есть «делители нуля», т.е. существуют две ненулевые матрицы A и B такие, что $AB = 0$.⁷
Пример:

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

4. В матрицах есть «нильпотенты», то есть можно найти такую ненулевую матрицу A , что $A^n = 0$. Пример,

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

⁷На самом деле, это очень «хорошая» аномалия, так как она связана с тем, что ОСЛУ имеют решения. Действительно, вопрос решения ОСЛУ $Ax = 0$ – это в точности вопрос существования правых делителей нуля для A в множестве \mathbb{R}^n .

Блочное умножение матриц

Пусть даны две матрицы, которые разбиты на блоки как показано ниже:

$$\begin{matrix} & \begin{matrix} k & s \end{matrix} \\ \begin{matrix} m \\ n \end{matrix} & \begin{pmatrix} A & B \\ C & D \end{pmatrix} \end{matrix} \quad \begin{matrix} & \begin{matrix} u & v \end{matrix} \\ \begin{matrix} k \\ s \end{matrix} & \begin{pmatrix} X & Y \\ W & Z \end{pmatrix} \end{matrix}$$

Числа m, n, k, s, u, v – размеры соответствующих блоков. Наша цель понять, что эти матрицы можно перемножать блочно. А именно, увидеть, что результат умножения этих матриц имеет вид

$$\begin{matrix} & \begin{matrix} u & v \end{matrix} \\ \begin{matrix} m \\ n \end{matrix} & \begin{pmatrix} AX + BW & AY + BZ \\ CX + DW & CY + DZ \end{pmatrix} \end{matrix}$$

Делается это таким трюком. В начале заметим, что

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & B \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ C & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & D \end{pmatrix}$$

После чего методом «пристального взгляда» перемножаем матрицы с большим количеством нулей (попробуйте проделать это!).

На этот факт можно смотреть вот как. Матрица – это прямоугольная таблица заполненная числами. А можно составлять прямоугольные таблицы заполненные другими объектами, например матрицами. Тогда они складываются и перемножаются так же как и обычные матрицы из чисел. Единственное надо учесть, что в блочном умножении есть разница между $AX + BW$ и $XA + BW$, так как A, B, X и W не числа, а матрицы, то их нельзя переставлять местами, порядок теперь важен.

Вот полезный пример. Пусть дана матрица из $M_{n+1}(\mathbb{R})$ вида

$$\begin{pmatrix} A & v \\ 0 & \lambda \end{pmatrix}, \quad \text{где } A \in M_n(\mathbb{R}), \quad v \in \mathbb{R}^n, \quad \lambda \in \mathbb{R}$$

Тогда

$$\begin{pmatrix} A & v \\ 0 & \lambda \end{pmatrix} \begin{pmatrix} A & v \\ 0 & \lambda \end{pmatrix} = \begin{pmatrix} A^2 & Av + v\lambda \\ 0 & \lambda^2 \end{pmatrix} = \begin{pmatrix} A^2 & Av + \lambda v \\ 0 & \lambda^2 \end{pmatrix} = \begin{pmatrix} A^2 & (A + \lambda E)v \\ 0 & \lambda^2 \end{pmatrix}$$

Предпоследнее равенство верно, так как не важно с какой стороны умножать v на скаляр λ .

Задача. Пусть $A \in M_{m,n}(\mathbb{R})$ и $B \in M_{n,k}(\mathbb{R})$. Покажите, что

1. $(AB)^t = B^t A^t$.
2. Если A – блочная матрица следующего вида

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \quad \text{тогда } A^t = \begin{pmatrix} A_{11}^t & A_{21}^t \\ A_{12}^t & A_{22}^t \end{pmatrix}$$

Вот еще один полезный пример блочного умножения. Пусть $x_1, \dots, x_m \in \mathbb{R}^n$ и $y_1, \dots, y_m \in \mathbb{R}^n$ – столбцы. Составим из этих столбцов матрицы $X = (x_1 | \dots | x_m)$ и $Y = (y_1 | \dots | y_m)$.⁸ Заметим, что $X, Y \in M_{n,m}(\mathbb{R})$. Тогда

$$XY^t = (x_1 | \dots | x_m)(y_1 | \dots | y_m)^t = \sum_{i=1}^m x_i y_i^t$$

Элементарные преобразования

Тип I Пусть $S_{ij}(\lambda) \in M_n(\mathbb{R})$ – матрица, полученная из единичной вписыванием в ячейку i, j числа λ . Эта матрица имеет следующий вид:

$$S_{ij}(\lambda) = \begin{matrix} & & j & \\ i & \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \ddots & \lambda & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix} \end{matrix}$$

⁸Данная запись означает, что мы берем столбцы x_i и записываем их подряд в одну большую таблицу.

Тогда прямая проверка показывает, умножение $A \in M_{nm}(\mathbb{R})$ на $S_{ij}(\lambda)$ слева прибавляет j строку умноженную на λ к i строке матрицы A , а умножение $B \in M_{mn}(\mathbb{R})$ на $S_{ij}(\lambda)$ справа прибавляет i столбец умноженный на λ к j столбцу матрицы B .

Тип II Пусть $U_{ij} \in M_n(\mathbb{R})$ – матрица, полученная из единичной перестановкой i и j столбцов (или что то же самое – строк). Эта матрица имеет следующий вид

$$\begin{matrix} & i & & j \\ i & \begin{pmatrix} 1 & & & \\ & 0 & & 1 \\ & & \ddots & \\ & 1 & & 0 \end{pmatrix} & \\ j & & & 1 \end{matrix}$$

Тогда прямая проверка показывает, умножение $A \in M_{nm}(\mathbb{R})$ на U_{ij} слева переставляет i и j строки матрицы A , а умножение $B \in M_{mn}(\mathbb{R})$ на U_{ij} справа переставляет i и j столбцы матрицы B .

Тип III Пусть $D_i(\lambda) \in M_n(\mathbb{R})$ – матрица, полученная из единичной умножением i строки на $\lambda \in \mathbb{R}$ (или что то же самое – столбца). Эта матрица имеет следующий вид

$$\begin{matrix} & i \\ i & \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & \lambda & \\ & & & \ddots \\ & & & & 1 \end{pmatrix} \end{matrix}$$

Тогда прямая проверка показывает, умножение $A \in M_{nm}(\mathbb{R})$ на $D_i(\lambda)$ слева умножает i строку A на λ , а умножение $B \in M_{mn}(\mathbb{R})$ на $D_i(\lambda)$ справа умножает i столбец матрицы B на λ .

Внимание магия! Пусть есть матрицы U и V соответствующие элементарным преобразованиям и A – произвольная матрица. Тогда сделать преобразование U над строками, а потом V над столбцами это $(UA)V$, а сделать эти преобразования в обратном порядке (то есть сначала над столбцами, а потом над строками) это $U(AV)$. Так как умножение матриц ассоциативно (не важно как расставлять скобки), то это одно и то же. Значит действия над строками коммутируют с действиями над столбцами.

Умножение на специальные виды матриц Пусть $A, B \in M_n(\mathbb{R})$ имеют следующий вид

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1\ n-1} & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2\ n-1} & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n-1\ 1} & a_{n-1\ 2} & \dots & a_{n-1\ n-1} & a_{n-1\ n} \\ a_{n1} & a_{n2} & \dots & a_{nn-1} & a_{nn} \end{pmatrix} \quad B = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & \ddots & \\ & & & \ddots & 1 \\ & & & & 0 \end{pmatrix}$$

где все незаполненные клетки считаются равными нулю. Тогда прямая проверка показывает, что

$$AB = \begin{pmatrix} 0 & a_{11} & \dots & a_{1\ n-2} & a_{1\ n-1} \\ 0 & a_{21} & \dots & a_{2\ n-2} & a_{2\ n-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & a_{n-1\ 1} & \dots & a_{n-1\ n-2} & a_{n-1\ n-1} \\ 0 & a_{n1} & \dots & a_{nn-2} & a_{nn-1} \end{pmatrix} \quad \text{и} \quad BA = \begin{pmatrix} a_{21} & a_{22} & \dots & a_{2\ n-1} & a_{2n} \\ a_{31} & a_{32} & \dots & a_{3\ n-1} & a_{3n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn-1} & a_{nn} \\ 0 & 0 & \dots & 0 & 0 \end{pmatrix}$$

То есть умножение на B справа сдвигает все столбцы матрицы A вправо, а умножение на B слева сдвигает все строки матрицы A вверх. Если мы хотим переставлять столбцы и строки по циклу, то надо умножать на

матрицу B следующего вида

$$B = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & \ddots & \\ & & & \ddots & 1 \\ 1 & & & & 0 \end{pmatrix}$$

то есть добавить единичку слева снизу к предыдущей матрице. Проверьте, что умножение на нее действительно сдвигает столбцы и строки по циклу.

Коммутирующие матрицы

В общем виде матрицы не коммутируют друг с другом. Потому интересной является такая задача: для конкретной матрицы $A \in M_n(\mathbb{R})$ найти все такие матрицы $X \in M_n(\mathbb{R})$, которые коммутируют с A , то есть $AX = XA$. Давайте рассмотрим парочку конкретных матриц.

Диагональные матрицы Пусть матрица $A \in M_n(\mathbb{R})$ имеет вид

$$A = \begin{pmatrix} \lambda_1 & & & & \\ & \lambda_2 & & & \\ & & \lambda_3 & & \\ & & & \ddots & \\ & & & & \lambda_n \end{pmatrix}$$

где $\lambda_i \neq \lambda_j$ при $i \neq j$. Найдем $\{X \in M_n(\mathbb{R}) \mid AX = XA\}$. Ответ на эту задачу достаточно простой. Оказывается в качестве X подходят диагональные матрицы и только они. Из определения умножения для диагональных матриц мы видим, что все диагональные точно коммутируют с A . Надо лишь показать, почему других матриц нет. Действительно, пусть X – произвольная матрица с неопределенными коэффициентами. Рассмотрим условие $AX = XA$. В произвольной ячейке с индексами i, j в левой матрице стоит элемент $\lambda_i x_{ij}$, а в правой матрице $\lambda_j x_{ij}$. Если $i = j$, то это один и тот же элемент и их равенство не дает никакого условия. Однако, для $i \neq j$ мы получаем, что условие $\lambda_i x_{ij} = \lambda_j x_{ij}$ влечет $x_{ij} = 0$ (тут мы пользуемся тем, что $\lambda_i \neq \lambda_j$). А это и означает, что все внедиагональные элементы матрицы X равны нулю, а значит она диагональна.

Матрица сдвига Пусть матрица $A \in M_n(\mathbb{R})$ имеет вид

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

Найдем $\{X \in M_n(\mathbb{R}) \mid AX = XA\}$. Давайте для определенности рассмотрим задачу в случае $n = 4$, так будет проще уловить наши вычисления, то есть пусть

$$A = \begin{pmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & 0 & 1 \\ & & & 0 \end{pmatrix} \quad \text{и} \quad X = \begin{pmatrix} x_{11} & x_{12} & x_{13} & x_{14} \\ x_{21} & x_{22} & x_{23} & x_{24} \\ x_{31} & x_{32} & x_{33} & x_{34} \\ x_{41} & x_{42} & x_{43} & x_{44} \end{pmatrix}$$

Тогда

$$AX = \begin{pmatrix} x_{21} & x_{22} & x_{23} & x_{24} \\ x_{31} & x_{32} & x_{33} & x_{34} \\ x_{41} & x_{42} & x_{43} & x_{44} \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{и} \quad XA = \begin{pmatrix} 0 & x_{11} & x_{12} & x_{13} \\ 0 & x_{21} & x_{22} & x_{23} \\ 0 & x_{31} & x_{32} & x_{33} \\ 0 & x_{41} & x_{42} & x_{43} \end{pmatrix}$$

Так как матрицы AX и XA равны, то у них равны все элементы. Давайте посмотрим на первый столбец для начала сверху вниз. Тогда $x_{21} = (AX)_{11} = (XA)_{11} = 0$. Значит

$$X = \begin{pmatrix} x_{11} & x_{12} & x_{13} & x_{14} \\ 0 & x_{22} & x_{23} & x_{24} \\ x_{31} & x_{32} & x_{33} & x_{34} \\ x_{41} & x_{42} & x_{43} & x_{44} \end{pmatrix}, \quad AX = \begin{pmatrix} 0 & x_{22} & x_{23} & x_{24} \\ x_{31} & x_{32} & x_{33} & x_{34} \\ x_{41} & x_{42} & x_{43} & x_{44} \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{и} \quad XA = \begin{pmatrix} 0 & x_{11} & x_{12} & x_{13} \\ 0 & 0 & x_{22} & x_{23} \\ 0 & x_{31} & x_{32} & x_{33} \\ 0 & x_{41} & x_{42} & x_{43} \end{pmatrix}$$

Аналогично Тогда $x_{31} = (AX)_{21} = (XA)_{21} = 0$, то есть имеем

$$X = \begin{pmatrix} x_{11} & x_{12} & x_{13} & x_{14} \\ 0 & x_{22} & x_{23} & x_{24} \\ 0 & x_{32} & x_{33} & x_{34} \\ x_{41} & x_{42} & x_{43} & x_{44} \end{pmatrix}, \quad AX = \begin{pmatrix} 0 & x_{22} & x_{23} & x_{24} \\ 0 & x_{32} & x_{33} & x_{34} \\ x_{41} & x_{42} & x_{43} & x_{44} \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{и} \quad XA = \begin{pmatrix} 0 & x_{11} & x_{12} & x_{13} \\ 0 & 0 & x_{22} & x_{23} \\ 0 & 0 & x_{32} & x_{33} \\ 0 & x_{41} & x_{42} & x_{43} \end{pmatrix}$$

Потом видим, что $x_{41} = (AX)_{31} = (XA)_{31} = 0$ и значит

$$X = \begin{pmatrix} x_{11} & x_{12} & x_{13} & x_{14} \\ 0 & x_{22} & x_{23} & x_{24} \\ 0 & x_{32} & x_{33} & x_{34} \\ 0 & x_{42} & x_{43} & x_{44} \end{pmatrix}, \quad AX = \begin{pmatrix} 0 & x_{22} & x_{23} & x_{24} \\ 0 & x_{32} & x_{33} & x_{34} \\ 0 & x_{42} & x_{43} & x_{44} \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{и} \quad XA = \begin{pmatrix} 0 & x_{11} & x_{12} & x_{13} \\ 0 & 0 & x_{22} & x_{23} \\ 0 & 0 & x_{32} & x_{33} \\ 0 & 0 & x_{42} & x_{43} \end{pmatrix}$$

Условие $(AX)_{41} = (XA)_{41}$ тривиальное и не дает никаких ограничений. Теперь сравним элементы матриц AX и XA во втором столбце начиная с диагонали вниз. Получим $x_{32} = (AX)_{22} = (XA)_{22} = 0$ и здесь мы пользуемся полученным знанием о том, что элемент $(XA)_{22}$ нулевой, получаем

$$X = \begin{pmatrix} x_{11} & x_{12} & x_{13} & x_{14} \\ 0 & x_{22} & x_{23} & x_{24} \\ 0 & 0 & x_{33} & x_{34} \\ 0 & x_{42} & x_{43} & x_{44} \end{pmatrix}, \quad AX = \begin{pmatrix} 0 & x_{22} & x_{23} & x_{24} \\ 0 & 0 & x_{33} & x_{34} \\ 0 & x_{42} & x_{43} & x_{44} \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{и} \quad XA = \begin{pmatrix} 0 & x_{11} & x_{12} & x_{13} \\ 0 & 0 & x_{22} & x_{23} \\ 0 & 0 & 0 & x_{33} \\ 0 & 0 & x_{42} & x_{43} \end{pmatrix}$$

Продолжая в том же духе и переходя к следующим столбцам с диагонали и ниже, мы видим, что у нас будет следующая ситуация

$$X = \begin{pmatrix} x_{11} & x_{12} & x_{13} & x_{14} \\ 0 & x_{22} & x_{23} & x_{24} \\ 0 & 0 & x_{33} & x_{34} \\ 0 & 0 & 0 & x_{44} \end{pmatrix}, \quad AX = \begin{pmatrix} 0 & x_{22} & x_{23} & x_{24} \\ 0 & 0 & x_{33} & x_{34} \\ 0 & 0 & 0 & x_{44} \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{и} \quad XA = \begin{pmatrix} 0 & x_{11} & x_{12} & x_{13} \\ 0 & 0 & x_{22} & x_{23} \\ 0 & 0 & 0 & x_{33} \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Теперь нам надо рассмотреть все ячейки в матрицах AX и XA выше главного диагонали. Давайте начнем с элемента 1, 2 и будем идти по диагонали вправо вниз. Видим, что $x_{22} = (AX)_{12} = (XA)_{12} = x_{11}$. Аналогично $x_{33} = (AX)_{23} = (XA)_{23} = x_{22}$ и $x_{44} = (AX)_{34} = (XA)_{34} = x_{33}$. То есть мы видим

$$X = \begin{pmatrix} a & x_{12} & x_{13} & x_{14} \\ 0 & a & x_{23} & x_{24} \\ 0 & 0 & a & x_{34} \\ 0 & 0 & 0 & a \end{pmatrix}, \quad AX = \begin{pmatrix} 0 & a & x_{23} & x_{24} \\ 0 & 0 & a & x_{34} \\ 0 & 0 & 0 & a \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{и} \quad XA = \begin{pmatrix} 0 & a & x_{12} & x_{13} \\ 0 & 0 & a & x_{23} \\ 0 & 0 & 0 & a \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Продолжая аналогично рассматривать элементы выше главного диагонали на диагоналях параллельных главной, мы видим, что

$$X = \begin{pmatrix} a & b & c & d \\ 0 & a & b & c \\ 0 & 0 & a & b \\ 0 & 0 & 0 & a \end{pmatrix}, \quad AX = \begin{pmatrix} 0 & a & b & c \\ 0 & 0 & a & b \\ 0 & 0 & 0 & a \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{и} \quad XA = \begin{pmatrix} 0 & a & b & c \\ 0 & 0 & a & b \\ 0 & 0 & 0 & a \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Таким образом в общем случае множество матриц X коммутирующих с A имеет следующий вид

$$\begin{pmatrix} a_1 & a_2 & \ddots & a_{n-1} & a_n \\ & a_1 & a_2 & \ddots & a_{n-1} \\ & & a_1 & \ddots & \ddots \\ & & & \ddots & a_2 \\ & & & & a_1 \end{pmatrix}$$

То есть это верхнетреугольная матрица у которой на диагоналях главное параллельной стоят одинаковые числа.

Квадратные матрицы

Заметим, что множество квадратных матриц $M_n(\mathbb{R})$ замкнуто относительно сложения, умножения, содержит 0 и 1 (в смысле содержит единичную матрицу E , которая является аналогом числа 1 для чисел). На это множество можно смотреть как на обобщение обычных чисел. Однако, надо не забывать, что произведение ненулевых матриц может стать нулем или даже степень ненулевой матрицы может оказаться нулем.

Что значит деление в числах? Предположим, что у нас есть два числа $a, b \in \mathbb{R}$. Тогда деление $a/b = a \cdot b^{-1}$ – это просто умножение на обратный элемент, а обратный элемент b^{-1} определяется свойством $bb^{-1} = 1$. Данное наблюдение дает ключ к распространению деления и обращения на случай матриц.

Обратимые матрицы Чтобы определить обратную матрицу нам нужен аналог единицы из чисел. В качестве такого аналога выступает единичная матрица E . Для матрицы $A \in M_n(\mathbb{R})$ матрица C называется обратной, если выполняется равенство $AC = E = CA$. Такая матрица C не обязательно существует, но если существует, то единственна. Тогда обратная матрица обозначается A^{-1} . Еще стоит обратить внимание на то, что умножение матриц не коммутативно (имеет значение в каком порядке умножаются матрицы). Так как деление – это умножение на обратную матрицу, то у нас возникает аж два деления: деление справа и деление слева. Чтобы не было путаницы операцию деления вообще не воодят на матрицах, а пользуются умножением на обратную.

Утверждение. Пусть $A \in M_n(\mathbb{R})$ – произвольная квадратная матрица. Тогда следующие условия эквивалентны:

1. Система линейных уравнений $Ax = 0$ имеет только нулевое решение $x \in \mathbb{R}^n$.
2. Система линейных уравнений $A^t y = 0$ имеет только нулевое решение $y \in \mathbb{R}^n$.
3. Матрица A представляется в виде $A = U_1 \cdot \dots \cdot U_k$, где U_i – матрицы элементарных преобразований.
4. Матрица A обратима.
5. Матрица A обратима слева, т.е. существует $L \in M_n(\mathbb{R})$ такая, что $LA = E$.
6. Матрица A обратима справа, т.е. существует $R \in M_n(\mathbb{R})$ такая, что $AR = E$.

Нахождение обратной матрицы методом Гаусса

Дано Матрица $A \in M_n(\mathbb{R})$.

Задача Понять обратима ли матрица A и если она обратима, то найти ее обратную A^{-1} .

Алгоритм

1. Нам надо по сути решить систему $AX = E$, где E – единичная матрица. Потому составим расширенную матрицу системы $(A|E)$.
2. Приведем эту матрицу к улучшенному ступенчатому виду.
3. В результате возможны 2 случая:
 - (a) После приведения получили матрицу $(E|B)$. Тогда A обратима и $A^{-1} = B$.
 - (b) После приведения получили матрицу $(D|B)$ и у матрицы D есть свободные переменные (или что то же самое нулевая строка). Тогда матрица A не обратима.

Заметим, что если в процессе алгоритма, мы слева от черты в расширенной матрице нашли свободную переменную, то на этом можно остановиться – матрица A необратима.

Обратимость прямоугольных матриц

Утверждение. Пусть $A \in M_{m,n}(\mathbb{R})$ и $B \in M_{n,m}(\mathbb{R})$ такие, что $AB = E_m$ и $BA = E_n$ (здесь E_n – единичная матрица размера n). Тогда $m = n$.

Доказательство. Все доказательство основано на догадливости: надо рассмотреть равенство $\text{tr}(AB) = \text{tr}(BA)$. Откуда получаем, $m = \text{tr}(AB) = \text{tr}(BA) = n$, все. \square

Односторонняя обратимость Пусть $A \in M_{m,n}(\mathbb{R})$ и при этом $m < n$. Давайте покажем следующие вещи:

1. Если матрица A содержит нулевой столбец, то она не может быть обратима слева, а если содержит нулевую строку, то она не может быть обратима справа.
2. Матрица A обратима справа (слева) тогда и только тогда, когда обратима справа (слева) матрица CAD , где $C \in M_m(\mathbb{R})$ и $D \in M_n(\mathbb{R})$ – обратимые матрицы.
3. Не существует матрицы $B \in M_{n,m}(\mathbb{R})$ такой, что $BA = E \in M_n(\mathbb{R})$. То есть прямоугольная матрица никогда не обратима с короткой стороны.
4. Найдется матрица $B \in M_{n,m}(\mathbb{R})$ такая, что $AB = E \in M_m(\mathbb{R})$, тогда и только тогда, когда количество главных переменных A равно m . Таким образом прямоугольная матрица может быть обратима только с длинной стороны и критерий обратимости – максимально возможное количество ступенек в улучшенном ступенчатом виде.⁹
5. В случае существования правой обратной для матрицы $A \in M_{m,n}(\mathbb{R})$ количество правых обратных бесконечно.

(1) Пусть у матрицы A есть нулевой столбец, тогда для любой матрицы B , в матрице BA тот же самый столбец будет нулевым. А значит BA никогда не может быть единичной. Аналогично делается и для нулевой строки.

(2) Докажем для обратимости справа, для обратимости слева делается аналогично. Пусть матрица A обратима справа, тогда найдется $B \in M_{n,m}(\mathbb{R})$ такая, что $AB = E \in M_m(\mathbb{R})$. Тогда $E = AB = ADD^{-1}B$. Домножим это равенство на C слева и на C^{-1} справа, получим $E = CADD^{-1}BC$. То есть $D^{-1}BC$ – правый обратный для CAD . В обратную сторону делается аналогично ибо, если $A' = CAD$, то $A = C^{-1}A'D^{-1}$.

(3) Покажем, что матрица A не обратима с короткой стороны. Приведем матрицу A^t к ступенчатому виду Q^t . Так как $n > m$, то это означает, что Q^t имеет нулевую строку в конце (матрица высокая). Это значит, что в матрице A после выполнения элементарных преобразований столбцов, мы получим матрицу Q с нулевым столбцом. В частности, по пункту (1), матрица Q не обратима слева. Так как элементарные преобразования столбцов означают домножение на обратимую матрицу справа, это значит, что найдется обратимая матрица $U \in M_n(\mathbb{R})$ такая, что $AU = Q$. Тогда по пункту (2) следует, что и матрица A не обратима слева.

(4) Пусть у матрицы A в ступенчатом виде Q главных переменных меньше m (количества строк). Тогда в матрице Q есть нулевая строка. То есть матрица Q не обратима справа. Значит для некоторой обратимой матрицы $U \in M_m(\mathbb{R})$ выполнено $UA = Q$. Теперь из пункта (2) следует, что A необратима Справа.

Теперь рассмотрим случай, когда количество главных переменных равно количеству строк. В этом случае матрицу A можно привести к виду $(E \mid 0)$ с помощью преобразований строк и столбцов. То есть для обратимых матриц $C \in M_m(\mathbb{R})$ и $D \in M_n(\mathbb{R})$ выполнено

$$CAD = (E \mid 0)$$

С другой стороны, мы видим, что матрица $R = \begin{pmatrix} E \\ 0 \end{pmatrix}$ является правой обратной к $(E \mid 0)$. Опять ссылка на пункт (2) завершает доказательство обратимости A справа. При этом заметим, что правая обратная к A имеет вид $DRC \in M_{n,m}(\mathbb{R})$.

(5) Для этого достаточно заметить, что в случае выше мы могли бы взять матрицу $R' = \begin{pmatrix} E \\ P \end{pmatrix}$, где P – произвольная матрица нужных размеров. Тогда формула $DR'C \in M_{n,m}(\mathbb{R})$ дает нам бесконечное множество правых обратных.

Примеры Очень полезно привести пару примеров односторонних обратимых матриц. Самый простой случай $A = (1, 0)$. Тогда понятно, что у нее нет левой обратной (мы это уже доказали, но тем не менее, я повторю аргумент явно), действительно, в этом случае у BA второй столбец всегда нулевой. С другой стороны матрицы вида $\begin{pmatrix} 1 \\ a \end{pmatrix}$ являются всевозможными обратными для A справа.

⁹Теперь то мы знаем, что он корректно определен.

Семинар 1

Задачи:

1. Найти общее решение и одно частное решение системы линейных уравнений, используя метод Гаусса:

$$\begin{cases} -3x_1 + 6x_2 + 2x_3 - 5x_4 = -6, \\ 2x_1 - 4x_2 + 2x_4 = 4, \\ 3x_1 - 6x_2 - x_3 + 4x_4 = 6 \end{cases}$$

2. Исследовать систему и найти общее решение в зависимости от значения параметра λ :

$$\begin{cases} (1 + \lambda)x_1 + x_2 + x_3 = 1, \\ x_1 + (1 + \lambda)x_2 + x_3 = \lambda, \\ x_1 + x_2 + (1 + \lambda)x_3 = \lambda^2 \end{cases}$$

3. Пусть матрица $A \in M_{56}(\mathbb{R})$ имеет вид

$$\begin{pmatrix} 1 & x & 1 & 1 & x & 1 \\ x & 1 & x & x & 1 & x \\ x & 1 & 1 & 1 & 1 & x \\ 1 & x & 1 & 1 & x & 1 \\ 1 & 1 & x & x & 1 & 1 \end{pmatrix}$$

Для системы $Ay = 0$, где $y \in \mathbb{R}^6$, найти количество главных переменных при любом значении $x \in \mathbb{R}$.

4. Пусть матрица $J(\lambda) \in M_n(\mathbb{R})$ имеет следующий вид¹

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

Доказать, что для любого k верна формула

$$J(\lambda)^k = \begin{pmatrix} \lambda^k & C_k^1 \lambda^{k-1} & C_k^2 \lambda^{k-2} & \dots & C_k^{n-1} \lambda^{k-n+1} \\ 0 & \lambda^k & C_k^1 \lambda^{k-1} & \dots & C_k^{n-2} \lambda^{k-n+2} \\ 0 & 0 & \lambda^k & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & C_k^1 \lambda^{k-1} \\ 0 & 0 & \dots & 0 & \lambda^k \end{pmatrix}$$

где $C_k^m = \frac{k!}{n!(k-n)!}$, а $n! = 1 \cdot 2 \cdot \dots \cdot n$.

5. Вычислить для любого n :

$$\left(\begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix} \right)^n$$

6. Пусть $A \in M_n(\mathbb{R})$ такая, что $A^m = 0$ для некоторого m . Показать, что $E + A$ и $E - A$ обратимы, где $E \in M_n(\mathbb{R})$ – единичная матрица.

¹Такая матрица называется Жордановой клеткой.

Семинар 2

Классификационный результат

Утверждение. Пусть $A, B \in M_{m,n}(\mathbb{R})$ и пусть $E_A, E_B \subseteq \mathbb{R}^n$ – множества решений систем $Ax = 0$ и $Bx = 0$, соответственно. Тогда следующее эквивалентно:

1. $E_A = E_B$, т.е. системы имеют одно и то же множество решений.
2. A приводится к B элементарными преобразованиями.
3. Существует обратимая $C \in M_m(\mathbb{R})$ такая, что $B = CA$.
4. Матрица улучшенного ступенчатого вида для A совпадает с матрицей улучшенного ступенчатого вида для B .

Подстановка матриц в многочлен

Пусть $p(x) = a_0 + a_1x + \dots + a_nx^n$ – многочлен с вещественными коэффициентами, а $A \in M_n(\mathbb{R})$. Тогда можно определить $p(A) = a_0E + a_1A^1 + \dots + a_nA^n$, где E – единичная матрица. Множество всех многочленов с вещественными коэффициентами я буду обозначать $\mathbb{R}[x]$.

Свойства подстановки в многочлен

1. Если $A \in M_n(\mathbb{R})$ и $f \in \mathbb{R}[x]$ – многочлен, то $f(C^{-1}AC) = C^{-1}f(A)C$ для любой обратимой $C \in M_n(\mathbb{R})$.
2. Если $A \in M_n(\mathbb{R})$ и $f, g \in \mathbb{R}[x]$ – многочлены, то матрицы $f(A)$ и $g(A)$ коммутируют между собой.

Спектр матрицы

Пусть $A \in M_n(\mathbb{R})$ определим спектр матрицы A следующим образом:

$$\text{спес } A = \{\lambda \in \mathbb{C} \mid A - \lambda E \text{ не обратима}\}$$

Это определение спектра принадлежит функциональному анализу. На самом деле этот спектр совпадает со множеством собственных значений матрицы, о которых речь пойдет чуть позже. Обратите внимание, что хотя матрица имеет вещественные коэффициенты, спектр мы рассматриваем комплексный. Это все заговор среди математиков, которые любят комплексные числа больше вещественных за их более приличное математическое поведение. Еще можно ввести спектр с конкретным видом чисел, например вещественный спектр

$$\text{спес}_{\mathbb{R}} A = \{\lambda \in \mathbb{R} \mid A - \lambda E \text{ не обратима}\}$$

Аналогично спектр можно определить для любого множества чисел (рациональных, например или еще каких).

Примеры

1. Пусть $A \in M_n(\mathbb{R})$ – диагональная матрица

$$A = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

Тогда $\text{спес } A = \{\lambda_1, \dots, \lambda_n\}$.

2. Пусть $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in M_2(\mathbb{R})$. Можно показать, что $\text{спес } A = \{i, -i\}$.

Минимальный многочлен Пусть $A \in M_n(\mathbb{R})$ – некоторая матрица. Рассмотрим множество всех ненулевых многочленов аннулирующих A . Формально мы смотрим на множество

$$M = \{f \in \mathbb{R}[x] \mid f(A) = 0, f \neq 0\}$$

Пусть $f_{\min} \in M$ – многочлен самой маленькой степени и со старшим коэффициентом 1. Тогда он называется минимальным многочленом матрицы A . Обратите внимание, что минимальный многочлен зависит от того, с какими коэффициентами мы его рассматриваем. Комплексный и вещественный минимальный многочлен могут быть разными.

Утверждение. Пусть $A \in M_n(\mathbb{R})$, тогда верны следующие утверждения:

1. Минимальный многочлен f_{\min} существует и единственный, при этом $\deg f_{\min} \leq n$.
2. Минимальный многочлен делит любой другой многочлен аннулирующий A .
3. $\lambda \in \text{спес } A$ тогда и только тогда, когда $f_{\min}(\lambda) = 0$.
4. Для любого аннулирующего многочлена $g \in \mathbb{R}[x]$, то есть $g(A) = 0$, верно $\text{спес } A \subseteq \text{корни } g$.¹
5. Верно равенство $\text{спес } A = \text{корни } f_{\min}$.

Поиск минимального многочлена Пусть задана матрица $A \in M_n(\mathbb{R})$. Тогда мы знаем, что найдется многочлен $f \in \mathbb{R}[x]$ такой, что $f(A) = 0$. Кроме того, я сообщил, что $\deg f \leq n$. Давайте обсудим, как найти подобный многочлен. Будем искать его с неопределенными коэффициентами $f(x) = a_0 + a_1x + \dots + a_nx^n$. Подставим в многочлен матрицу A и приравняем результат к нулю.

$$f(A) = a_0E + a_1A + \dots + a_nA^n = 0$$

Тогда то, что написано, является системой из n^2 уравнений, а именно

t.me/postypashki_old/1229

$$\{1 \leq i, j \leq n\} E_{ij}a_0 + A_{ij}a_1 + \dots + (A^n)_{ij}a_n = 0$$

t.me/postypashki_old/1229

Здесь через B_{ij} обозначены коэффициенты матрицы B , например, E_{ij} – это ij -ый коэффициент единичной матрицы, а $(A^n)_{ij}$ – ij -ый коэффициент матрицы A^n .

Теперь нас интересует ненулевое решение этой системы, у которого как можно больше нулей справа. Давайте поясню. Такое решение отвечает аннулирующему многочлену. Мы хотим выбрать такой многочлен как можно меньшей степени. То есть мы хотим по возможности занулить a_n , потом a_{n-1} , потом a_{n-2} и так далее, пока находится ненулевое решение. Предположим, что мы привели систему к ступенчатому виду и a_k – самая левая свободная переменная. Я утверждаю, что k и будет степенью минимального многочлена, а чтобы его найти надо положить $a_k = 1$, и все остальные свободные переменные равными нулю.

Действительно, если мы сделали, как описано, то все главные переменные правее a_k тоже равны нулю, ибо они зависят от свободных переменных, стоящих правее, а они в нашем случае нулевые. То есть a_k будет старший ненулевой коэффициент в искомом многочлене, а значит k будет его степенью. Почему нельзя найти меньше. Чтобы найти меньше надо занулить еще и a_k . То есть все свободные переменные в этом случае будут нулевыми, а тогда и все главные будут нулевыми, а это даст нулевое решение, что противоречит нашим намерениям найти ненулевой многочлен.

Вычленение из какого-то аннулирующего Предположим, что вы угадали какой-нибудь аннулирующий многочлен для вашей матрицы $A \in M_n(\mathbb{R})$, а именно, нашли какой-то $f \in \mathbb{R}[x]$ такой, что $f(A) = 0$. Тогда можно попытаться найти минимальный многочлен среди делителей многочлена f . Эта процедура требует уметь искать эти самые делители. Но в некоторых ситуациях эта процедура тоже бывает полезна. Например, в случае большой блочной матрицы A бывает проще найти аннулирующий многочлен.

¹Аналогичное выполнено для вещественного спектра, а именно, вещественный спектр матрицы A лежит среди вещественных корней g .

Замечание о спектре Можно показать, что любой вещественный многочлен $f \in \mathbb{R}[x]$ единственным образом разваливается в произведение

$$f(x) = (x - \lambda_1) \dots (x - \lambda_k) q_1(x) \dots q_r(x)$$

где числа $\lambda_i \in \mathbb{R}$ могут повторяться, а $q_i(x)$ – многочлены второй степени с отрицательным дискриминантом (то есть без вещественных корней).

Пусть теперь f_{\min} – минимальный многочлен некоторой матрицы A . Разложим его подобным образом. Тогда мы видим из предыдущего утверждения, что $\text{spes}_{\mathbb{R}} A$ помнит информацию только о первой половине сомножителей и теряет информацию о квадратичных многочленах. Однако, если бы мы рассмотрели f_{\min} как многочлен с комплексными коэффициентами, то мы бы могли доразложить все $q_i(x)$ на линейные множители и $\text{spes}_{\mathbb{C}} A$ помнит информацию о всех сомножителях f_{\min} . Еще надо понимать, что каждое $x - \lambda$ может несколько раз участвовать в разложении f_{\min} , но спектр не помнит это количество, он лишь знает был ли там данный $x - \lambda$ или нет.

Замечание об арифметических свойствах матриц Если вы работаете с матрицами, то готовьтесь к тому, чтобы думать про них как про более сложную версию чисел. А значит, вы будете писать с ними различного рода алгебраические выражения. Например, для какой-нибудь матрицы $A \in M_n(\mathbb{R})$ можно написать $A^3 + 2A - 3E$. И предположим вы хотите упростить это выражение как-нибудь, не зная как именно выглядит ваша матрица A . Единственное, что вам поможет в этом случае – зануляющий многочлен. Пусть, например, $f(x) = x^2 - 3$ зануляет A . Это значит, что $A^2 = 3E$. Тогда выражение выше можно упростить так

$$A^3 + 2A - 3E = 3A + 2A - 3E = 5A - 3E$$

Роль минимального многочлена заключается в том, что это «самый лучший» многочлен, который помнит как можно больше соотношений на матрицу A , чтобы можно было упрощать выражения. Более того, минимальный многочлен автоматически говорит, когда можно делить на выражение от матрицы, а когда нет. Например, на $A - E$ поделить можно, так как 1 не является корнем f , с другой стороны на матрицы $A \pm \sqrt{3}E$ делить нельзя.

Обратимость и минимальный многочлен Обратимость матрицы по определению равносильна тому, что в ее спектре нет нуля, а это то же самое, что у минимального многочлена свободный член отличен от нуля. В этом случае мы можем явно выразить обратную матрицу через исходную. Действительно, пусть $f_{\min} = a_0 + a_1x + \dots + a_mx^m$ для некоторой матрицы $A \in M_n(\mathbb{R})$. Тогда

$$a_0E + a_1A + \dots + a_mA^m = 0 \Rightarrow A(a_1E + \dots + a_mA^{m-1}) = -a_0E \Rightarrow A\left(-\frac{a_1}{a_0}E - \dots - \frac{a_m}{a_0}A^{m-1}\right) = E$$

То есть по определению

$$A^{-1} = -\frac{a_1}{a_0}E - \dots - \frac{a_m}{a_0}A^{m-1}$$

Обратите внимание, что данная формула работает при условии, что $a_0 \neq 0$. Эта процедура похожа на процедуру избавления от иррациональности в знаменателе дробей или избавления от мнимой части в знаменателе в комплексных дробях. Это неспроста, это в точности тот же самый метод.

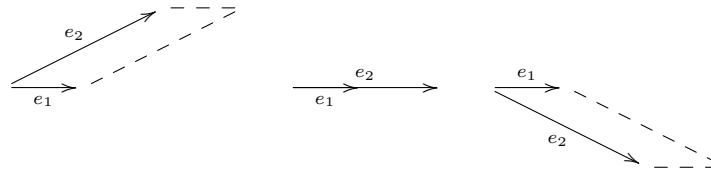
Определитель (немного философии)

Сейчас я хочу обсудить «ориентированный объем» на прямой, плоскости и в пространстве.

Прямая На прямой мы можем выбрать «положительное» направление. Обычно на рисунке выбирают с лева на право. Тогда длина вектора, который смотрит с лева на право, считается положительной, а с права на лево – отрицательной.

Плоскость Здесь объем будет задаваться парой векторов, то есть некоторой квадратной матрицей размера 2, где вектора – это ее столбцы. Основная идея такая: пусть мы хотим посчитать площадь между двумя

векторами на плоскости, точнее площадь параллелограмма натянутого на вектора e_1 и e_2 как на первом рисунке ниже.

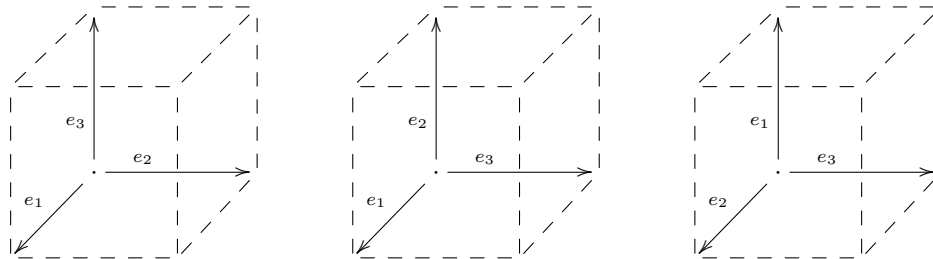


Давайте двигать вектор e_2 к вектору e_1 . Тогда площадь будет уменьшаться и когда вектора лягут на одну прямую, она будет равна нулю. Однако, если мы продолжим двигать вектор e_2 , то площадь между векторами опять начнет расти и картинка в конце концов станет симметрична исходной, а полученный параллелограмм равен изначальному. Однако, эта ситуация отличается от предыдущей и вот как можно понять чем. Предположим, что между векторами была натянута хорошо сжимаемая ткань, одна сторона которой красная, другая зеленая. Тогда в самом начале на нас смотрит красная сторона этой ткани, но как только e_2 прошел через e_1 на нас уже смотрит зеленая сторона. Мы бы хотели научиться отличать эти две ситуации с помощью знака, если на нас смотрит красная сторона – знак положительный, если зеленая – отрицательный.

Еще один способ думать про эту ситуацию. Представим, что плоскость – это наш стол, а параллелограмм вырезан из бумаги. Мы можем положить параллелограмм на стол двумя способами: лицевой стороной вверх или же вниз. В первом случае мы считаем площадь положительной, а во втором – отрицательной. Возможность определить лицевую сторону связана с тем, что мы знаем, где у стола верх, а где низ. Это возможно, потому что наша плоскость лежит в трех мерном пространстве и мы можем глядеть на нее извне. Однако, если бы мы жили на плоскости и у нас не было бы возможности выглянуть за ее пределы, то единственный способ установить «какой стороной вверх лежит параллелограмм» был бы с помощью порядка векторов.

Еще одно важное замечание. Если мы берем два одинаковых параллелограмма на нашем столе, которые лежат лицевой стороной вверх, то мы можем передвинуть один в другой, не отрывая его от стола. А вот если один из параллелограммов имеет положительный объем, а другой отрицательный, то нельзя перевести один в другой, не отрывая от стола. То есть, если вы живете на плоскости, то вам не получится переместить положительный параллелограмм в отрицательный, не сломав или не разобрав его.

Пространство В пространстве дело с ориентацией обстоит абсолютно аналогично. Мы хотим уже считать объемы параллелепипедов натянутых на три вектора. И мы так же хотим, чтобы эти объемы показывали «с какой стороны» мы смотрим на параллелепипед.



Здесь знак объема определяется по порядку векторов.² На рисунке объемы первого и третьего положительные, а у второго отрицательный.³ Если вы сделаете модельки этих кубиков из подписанных спичек, то третий кубик – это первый, но лежащий на другой грани. А вот второй кубик получить из первого вращениями не получится. Надо будет его разобрать и присобачить ребра по-другому.

Как и в случае с плоскостью, если бы мы могли выйти за пределы нашего трехмерного пространства, то у нас появилась бы лицевая и тыльная сторона, как у стола. И тогда первый и третий кубики лежали бы лицевой стороной вверх, а второй – вниз. Мы, конечно же, так сделать не сможем и никогда в жизни не увидим подобное, но думать про такое положение вещей по аналогии с плоскостью можем и эта интуиция бывает полезна.

²В общем случае можно определить понятие знака для любого упорядочивания векторов, так что этот знак будет вести себя хорошим образом. Однако, мы не будем углубляться в эту ненужную нам теорию.

³Чтобы понять, почему именно так, надо развить теорию знака перестановки. Однако, сейчас это не важно, а ниже, я расскажу метод, который не потребует этого знания.

Определитель

Существует несколько способов определить определитель.⁴ Вот классическое определение через перестановки. Пусть $A \in M_n(\mathbb{R})$ положим

$$\det A = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}$$

Именно эту формулу любят составители учебников по линейной алгебре и именно ее никто не использует, когда начинают решать задачи. А если учесть, что в этой формуле полно всяких непонятных значков, смысл которых еще надо объяснить, то это по сути означает, что данная формула, мягко говоря, не такая уж полезная и нужная.

Тем не менее, давайте я объясню, что тут написано. По-хорошему, я должен объяснить, что такое перестановки, их знак и как они тут используются. Но вместо этого я пойду коротким путем и объясню по рабоче-крестьянски.

Внутри стоят произведения вида $a_{1\sigma(1)} \cdots a_{n\sigma(n)}$, где первый индекс – это индекс строки и эти индексы идут по-порядку. Это значит, что мы из каждой строки выбираем по элементу. Хитрая запись для второго индекса означает, что мы выбираем элементы так, чтобы из каждого столбца выбран в точности один элемент. Например, если мы в первой строке выбрали элемент из 3-го столбца, то далее из 3-го столбца выбирать уже нельзя. То есть такое слагаемое получено так: мы выбираем n элементов из матрицы A так, чтобы в каждой строке и каждом столбце встретился только один элемент. Все эти элементы перемножаем, добавляем знак ± 1 и получаем одно слагаемое. А потом мы все подобные слагаемые складываем между собой. Всего оказывается целых $n!$ слагаемых. А это означает, что их сильно дофига и считать по этой формуле практически невозможно.

Как это определение связано с ориентированным объемом. Думать про это надо так: пусть $A = (A_1 | \dots | A_n) \in M_n(\mathbb{R})$ составлена из столбцов $A_i \in \mathbb{R}^n$. Тогда $\det A$ – это ориентированный объем n -мерного параллелепипеда натянутого на вектора A_1, \dots, A_n .

Примеры

1. Если $A \in M_1(\mathbb{R}) = \mathbb{R}$, то $\det A = A$.
2. Если $A \in M_2(\mathbb{R})$ имеет вид $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, то $\det A = ad - bc$. Графически: главная диагональ минус побочная.
3. Если $A \in M_3(\mathbb{R})$ имеет вид $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$, то определитель получается из 6 слагаемых три из них с $+$ три с $-$. Графически слагаемые можно изобразить так:

$$\det A = + \left(\begin{array}{|c|} \hline \diagdown \\ \hline \end{array} \right) + \left(\begin{array}{|c|} \hline \diagup \\ \hline \end{array} \right) + \left(\begin{array}{|c|} \hline \diagdown \\ \hline \end{array} \right) - \left(\begin{array}{|c|} \hline \diagup \\ \hline \end{array} \right) - \left(\begin{array}{|c|} \hline \diagdown \\ \hline \end{array} \right) - \left(\begin{array}{|c|} \hline \diagup \\ \hline \end{array} \right)$$

Точная формула⁵

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

Свойства определителя

Считать определитель по явной формуле весьма проблематично. Слишком уж много слагаемых. Потому по определению его можно вычислить лишь для очень специальных матриц. Для произвольных матриц используются некоторые полезные свойства, с помощью которых их определители сводятся к определителям специальных матриц.

Пусть $A \in M_n(\mathbb{R})$ – матрица, тогда на нее можно смотреть как на набор из n столбцов $A = (A_1 | \dots | A_n)$. Тогда определитель $\det(A)$ можно рассматривать, как функцию от столбцов матрицы A , то есть $\det(A) = \det(A_1 | \dots | A_n)$. Думаю таким образом, мы можем сформулировать следующие свойства:

1. $\det(A_1 | \dots | A_i + A'_i | \dots | A_n) = \det(A_1 | \dots | A_i | \dots | A_n) + \det(A_1 | \dots | A'_i | \dots | A_n)$. Например,

$$\det \begin{pmatrix} 1 & 3 \\ 2 & 7 \end{pmatrix} = \det \left(\begin{pmatrix} 1 \\ 2 \end{pmatrix} \middle| \begin{pmatrix} 3 \\ 7 \end{pmatrix} \right) = \det \left(\begin{pmatrix} 1 \\ 2 \end{pmatrix} \middle| \begin{pmatrix} 1 \\ 3 \end{pmatrix} + \begin{pmatrix} 2 \\ 4 \end{pmatrix} \right) = \det \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix} + \det \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$$

⁴Как бы глупо это ни звучало.

⁵Начиная с этого момента можно забыть честное определение. Оно почти никогда не нужно для использования, нам лишь пригодятся его свойства. Определитель никогда не считают по определению. Это слишком долго.

2. $\det(A_1 | \dots | A_i | \dots | A_j | \dots | A_n) = -\det(A_1 | \dots | A_j | \dots | A_i | \dots | A_n)$. То есть если поменять местами два столбца, то определитель изменит знак, например,

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

3. $\det(A_1 | \dots | A' | \dots | A' | \dots | A_n) = 0$, то есть если у матрицы есть два одинаковых столбца, то определитель автоматически равен нулю, например

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

4. $\det(A_1 | \dots | \lambda A_i | \dots | A_n) = \lambda \det(A_1 | \dots | A_i | \dots | A_n)$. То есть, если один столбец умножить на одно и то же число, то весь определитель умножится на это число, например,

$$\det \begin{pmatrix} 1 & 3 \\ 2 & 9 \end{pmatrix} = 3 \det \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix}$$

5. $\det(A_1 | \dots | A_i | \dots | A_j | \dots | A_n) = \det(A_1 | \dots | A_i | \dots | A_j + \lambda A_i | \dots | A_n)$. То есть, если к одному столбцу матрицы прибавить другой умноженный на коэффициент, то определитель не изменится, например

$$\det \begin{pmatrix} 1 & 3 \\ 2 & 7 \end{pmatrix} = \det \left(\begin{pmatrix} 1 \\ 2 \end{pmatrix} \middle| \begin{pmatrix} 3 \\ 7 \end{pmatrix} \right) = \det \left(\begin{pmatrix} 1 \\ 2 \end{pmatrix} \middle| \begin{pmatrix} 3 \\ 7 \end{pmatrix} + 2 \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right) = \det \left(\begin{pmatrix} 1 \\ 2 \end{pmatrix} \middle| \begin{pmatrix} 5 \\ 11 \end{pmatrix} \right) = \det \begin{pmatrix} 1 & 5 \\ 2 & 11 \end{pmatrix}$$

6. $\det A = \det A^t$. То есть определитель матрицы равен определителю транспонированной матрицы. А значит, все свойства сформулированные выше для столбцов автоматически верны и для строк.

7. Определитель треугольной матрицы

$$\det \begin{pmatrix} \lambda_1 & * & \dots & * \\ & \lambda_2 & \dots & * \\ & & \ddots & \vdots \\ & & & \lambda_n \end{pmatrix} = \det \begin{pmatrix} \lambda_1 & & & \\ * & \lambda_2 & & \\ \vdots & \vdots & \ddots & \\ * & * & \dots & \lambda_n \end{pmatrix} = \lambda_1 \lambda_2 \dots \lambda_n$$

То есть у треугольной матрицы определитель равен произведению ее диагональных элементов. В частности $\det E = 1$ и $\det(\lambda E) = \lambda^n$.

Вычисление определителя с помощью элементарных преобразований

Из сформулированных свойств выше следует, что определитель можно считать так: надо привести матрицу A элементарными преобразованиями к треугольному виду и по пути запоминать некоторые коэффициенты, после чего надо перемножить эти коэффициенты с определителем треугольной матрицы. Давайте продемонстрируем на примере, будем приводить матрицу к ступенчатому виду элементарными преобразованиями строк

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

Связь определителя с произведением

Для определителя верны следующие формулы

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

3. Пусть $A \in M_n(\mathbb{R})$, $C \in M_m(\mathbb{R})$ и $B \in M_{nm}(\mathbb{R})$, тогда

$$\det \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = \det A \det C$$

Оказывается, что определитель является единственной функцией $\phi: M_n(\mathbb{R}) \rightarrow \mathbb{R}$ такой, что

$$1. \phi(AB) = \phi(A)\phi(B)$$

$$2. \phi \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & \lambda \end{pmatrix} = \lambda$$

Миноры и алгебраические дополнения

Пусть $A \in M_n(\mathbb{R})$ с элементами a_{ij} . Рассмотрим матрицу $D_{ij} \in M_{n-1}(\mathbb{R})$ полученную из A вычеркиванием i -ой строки и j -го столбца. Определитель матрицы D_{ij} обозначается M_{ij} и называется *минором* матрицы A или ij -минором для определенности. Число $(-1)^{i+j} M_{ij}$ называется алгебраическим дополнением элемента a_{ij} и обозначается A_{ij} .

Покажем как это все выглядит на картинках. Если мы представим матрицу A в виде

$$A = \begin{pmatrix} X_{ij} & \begin{matrix} * \\ \vdots \\ * \end{matrix} & Y_{ij} \\ * & \dots & a_{ij} & \dots & * \\ Z_{ij} & \begin{matrix} \vdots \\ * \end{matrix} & W_{ij} \end{pmatrix}$$

Тогда

$$D_{ij} = \begin{pmatrix} X_{ij} & Y_{ij} \\ Z_{ij} & W_{ij} \end{pmatrix}, \quad M_{ij} = \det \begin{pmatrix} X_{ij} & Y_{ij} \\ Z_{ij} & W_{ij} \end{pmatrix} \quad \text{и} \quad A_{ij} = (-1)^{i+j} \det \begin{pmatrix} X_{ij} & Y_{ij} \\ Z_{ij} & W_{ij} \end{pmatrix}$$

Разложение определителя по строке (столбцу)

Разложение по столбцу Для матрицы $A \in M_n(\mathbb{R})$ и любого k с условием $1 \leq k \leq n$ верна формула $\det A = \sum_{i=1}^n a_{ik} A_{ik}$. Здесь A_{ij} – алгебраическое дополнение a_{ij} . Например, разложим по второму столбцу

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

Для разложения по строкам верны ровно те же самые формулы. Их можно получить просто перейдя к транспонированной матрице

Явная формула обратной матрицы

Пусть $A \in M_n(\mathbb{R})$. Оказывается, что матрица обратима тогда и только тогда, когда $\det A \neq 0$. В этом случае мы можем написать явные формулы для обратной матрицы через определитель. Пусть A имеет вид

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

Составим матрицу A^+ из алгебраических дополнений, т.е.

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

Сопряженная матрица \hat{A} для A определяется как $(A^+)^t$, т.е.

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

Тогда $A^{-1} = \frac{1}{\det A} \hat{A}$.

Пример Пусть $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, тогда $A^+ = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$, $\hat{A} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$, $A^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

Формулы Крамера

Матрица $A \in M_n(\mathbb{R})$ называется *невыврожденной*, если $\det A \neq 0$. В противном случае она называется вырожденной. Если матрица A не вырожденная, то из явных формул для обратной матрицы следует, что существует A^{-1} . В частности система вида $Ax = b$ имеет единственное решение для любой правой части $b \in \mathbb{R}^n$.

Пусть $A \in M_n(\mathbb{R})$ – невырожденная квадратная матрица, $b \in \mathbb{R}^n$ – произвольный столбец и $Ax = b$ – система линейных уравнений с n неизвестными.

Пусть $A = (A_1 | \dots | A_n)$, где $A_i \in \mathbb{R}^n$ – столбцы матрицы A . Определим матрицы $B_i = (A_1 | \dots | A_{i-1} | b | A_{i+1} | \dots | A_n)$, т.е. B_i получается из A , если в A заменить i -ый столбец A_i на вектор b из правой части системы.

Так как A – невырождена, то система $Ax = b$ всегда имеет единственное решение. Это решение вычисляется по формулам $x_i = \frac{\det B_i}{\det A}$. Эти формулы и называются формулами Крамера.⁶

Пример Пусть $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ и $b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$. Решим систему $Ax = b$. Тогда $\det A = ad - bc$, $\det B_1 = \det \begin{pmatrix} b_1 & b \\ b_2 & d \end{pmatrix} = b_1d - b_2b$, $\det B_2 = \det \begin{pmatrix} a & b_1 \\ c & b_2 \end{pmatrix} = ab_2 - cb_1$. Тогда $x_1 = \frac{b_1d - b_2b}{ad - bc}$ и $x_2 = \frac{ab_2 - cb_1}{ad - bc}$.

Характеристический многочлен

Пусть $A \in M_n(\mathbb{R})$, тогда выражение

$$\chi_A(\lambda) = \det(\lambda E - A) = (-1)^n \det(A - \lambda E)$$

является многочленом от λ степени n и называется характеристическим многочленом для A . Главная польза от характеристического многочлена – его корни это и есть спектр матрицы.

Утверждение. Пусть $A \in M_n(\mathbb{R})$ и $\chi_A(\lambda)$ – характеристический многочлен. Тогда

1. Легко посчитать следующие коэффициенты⁷

$$\chi_A(\lambda) = \lambda^n - \operatorname{tr}(A)\lambda^{n-1} + \dots + (-1)^n \det(A)$$

2. Для произвольного числа λ верно, что $\lambda \in \operatorname{спекс} A$ тогда и только тогда, когда $\chi_A(\lambda) = 0$.

3. Многочлен $\chi_A(\lambda)$ зануляющий для A , то есть $\chi_A(A) = 0$. Это называется теорема Гамильтона-Кэли.

Явные формулы для коэффициентов характеристического многочлена Вначале давайте введем некоторые обозначения. Пусть $A \in M_n(\mathbb{R})$ – некоторая матрица. Рассмотрим произвольное k элементное подмножество в множестве чисел от 1 до n заданное в виде i_1, \dots, i_k ⁸. Вычеркнем из матрицы A столбцы и строки с этими номерами и обозначим полученную матрицу через R_{i_1, \dots, i_k} . Графически эта процедура

⁶У этих формул нет особого практического смысла, в основном только теоретическое применение, за исключением малого размера.

⁷На самом деле есть формулы для всех коэффициентов, но в них нет смысла для нас.

⁸Здесь предполагается, что $i_1 < \dots < i_k$.

выглядит так:

$$\begin{matrix} & & i_1 & \dots & i_k & & \\ & & \boxed{\begin{matrix} a_{1i_1} \\ \vdots \\ a_{1i_k} \end{matrix}} & \dots & \boxed{\begin{matrix} a_{1i_k} \\ \vdots \\ a_{1i_1} \end{matrix}} & & \\ i_1 & \left(\begin{array}{cc|cc|cc} R_{11} & & & & & R_{1k+1} \\ \hline a_{i_11} & \dots & & & & \dots & a_{i_1n} \\ \hline \vdots & & & & & & \vdots \\ \hline a_{i_k1} & \dots & & & & \dots & a_{i_kn} \\ \hline R_{k+11} & & \vdots & \dots & \vdots & & R_{k+1k+1} \\ & & a_{ni_1} & & a_{ni_k} & & \end{array} \right) & \mapsto & R_{i_1, \dots, i_k} = \begin{pmatrix} R_{11} & \dots & R_{1k+1} \\ \vdots & \ddots & \vdots \\ R_{k+11} & \dots & R_{k+1k+1} \end{pmatrix} \in M_{n-k}(\mathbb{R})
 \end{matrix}$$

Пользуясь этими обозначениями покажем следующее.

Утверждение. Пусть $A \in M_n(\mathbb{R})$ и его характеристический многочлен имеет вид

$$\chi_A(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_1\lambda + a_0$$

Тогда, в обозначениях выше, для коэффициентов a_k верна следующая формула⁹

$$a_k = (-1)^{n-k} \left(\sum_{i_1 < \dots < i_k} \det R_{i_1, \dots, i_k} \right)$$

⁹Заметим, что эта формула также имеет смысл при $k = 0$ и при $k = n$. Если $k = 0$, то множество индексов пусто \emptyset и $R_\emptyset = A$, потому формула превращается в равенство $a_0 = (-1)^n \det A$. При условии $k = n$, мы вычеркиваем все строки из матрицы и в этом случае $R_{i_1, \dots, i_n} \in M_0(\mathbb{R})$. Такого объекта не существует, но мы можем для удобства считать, что в этом случае формула означает $\det R_{i_1, \dots, i_n} = 1$.

Семинар 2 дополнительно

Принцип «продолжения по непрерывности»

Задача. Рассмотрим квадратную матрицу из $M_{n+m}(\mathbb{R})$ вида $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$, где $A \in M_n(\mathbb{R})$, $B \in M_{n,m}(\mathbb{R})$, $C \in M_{m,n}(\mathbb{R})$, $D \in M_m(\mathbb{R})$. Покажите следующее:

1. Если A обратима, то

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(A) \det(D - CA^{-1}B)$$

2. Если $m = n$ и $AC = CA$, то¹

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(AD - CB)$$

Решение. (1) Идея в том, чтобы применить блочные элементарные преобразования к строкам матрицы. А именно, мы берем первую строку матрицы $(A|B)$ и умножаем ее слева на «коэффициент» CA^{-1} , получим $(C|CA^{-1}B)$. После этого вычитаем эту строку из второй строки $(C|D)$ и получаем $(0|D - CA^{-1}B)$. На одной картинке мы пределали следующее

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \xrightarrow{\text{вычитаем из II-ой строки I-ю умноженную на } CA^{-1} \text{ слева}} \begin{pmatrix} A & B \\ 0 & D - CA^{-1}B \end{pmatrix}$$

По аналогии с элементарными преобразованиями первого типа над строками такая процедура не должна поменять определитель. Давайте строго поймем, почему определитель не меняется. Для этого заметим, что такая процедура эквивалентна умножению слева на блочную матрицу, а именно

$$\begin{pmatrix} E & 0 \\ -CA^{-1} & E \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A & B \\ 0 & D - CA^{-1}B \end{pmatrix}$$

Самая левая матрица – нижнетреугольная с единицами на диагонали, а ее определитель равен 1. Потому определитель исходной матрицы равен определителю матрицы слева. А ее определитель равен требуемому, так как тут есть угол нулей.

(2) **Шаг 1** Давайте в начале рассмотрим случай A обратима. В этом случае, по первому пункту мы получаем

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det A \det(D - CA^{-1}B) = \det(AD - ACA^{-1}B) = \det(AD - CB)$$

В последнем равенстве мы воспользовались тем, что матрицы A и C коммутируют.

Шаг 2 А здесь я вам продемонстрирую идею, на которой держится решение многих задач. Например, если вы можете доказать что-то только для обратимых матриц, то этим способом очень легко свести доказательство необратимого случая к обратимому.

Давайте рассмотрим матрицу $A_\lambda = A - \lambda E$ и поймем при каких λ такая матрица может быть обратима, а при каких нет. Я утверждаю, что только для конечного числа λ такая матрица не обратима. Действительно, мы знаем, что матрица обратима тогда и только тогда, когда ее определитель не ноль. Потому A_λ необратима тогда и только тогда, когда $\det A_\lambda = 0$. Давайте внимательно посмотрим на $\det(A - \lambda E)$.

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

Я утверждаю, что полученное выражение будет многочленом от λ .² Действительно, если посчитать по формуле для определителя, то мы будем перемножать какие-то элементы из матрицы $A - \lambda E$ и потом складывать все вместе. Все элементы являются либо константами либо линейными многочленами от λ . А значит их произведение будет многочленом степени не более n от λ , ну и сумма в том числе. Легко видеть, что старший член этого многочлена может взяться только с диагонали и равен $(-\lambda)^n$.

¹Обратите внимание, матрица A может не быть обратимой в этом случае.

²На самом деле даже многочленом степени n , но это сейчас не важно.

Так как $\det(A_\lambda)$ является ненулевым многочленом от λ , то он имеет только конечное количество корней.³ А значит матрица A_λ необратима только для конечного числа λ .

Теперь рассмотрим нашу задачу для матрицы A_λ вместо матрицы A . То есть нам надо посчитать

$$\det \begin{pmatrix} A_\lambda & B \\ C & D \end{pmatrix}$$

Матрицы A и C коммутировали. Так как $A_\lambda = A - \lambda E$, то и A_λ и C коммутируют. А значит для всех $\lambda \in \mathbb{R}$ кроме конечного числа мы можем воспользоваться предыдущим шагом и написать

$$\det \begin{pmatrix} A_\lambda & B \\ C & D \end{pmatrix} = \det(A_\lambda D - CB) \quad \text{для всех } \lambda \in \mathbb{R} \text{ кроме конечного числа}$$

Теперь обратим внимание, что левая часть

$$\det \begin{pmatrix} A_\lambda & B \\ C & D \end{pmatrix}$$

является каким-то многочленом $f(\lambda)$ и аналогично правая часть

$$\det(A_\lambda D - CB)$$

тоже является каким-то многочленом $g(\lambda)$. И мы только что доказали, что $f(\lambda) = g(\lambda)$ для всех $\lambda \in \mathbb{R}$ кроме конечного числа. То есть многочлен $f - g$ имеет бесконечное число корней. А такое бывает лишь когда $f = g$, а значит $f(\lambda) = g(\lambda)$ вообще для любого числа $\lambda \in \mathbb{R}$. Подставим значение $\lambda = 0$ и получим желаемое. \square

На последний шаг можно смотреть так. Мы знаем результат для каждой матрицы A_λ в силу обратимости таких матриц при достаточно малом λ . А теперь в финальном равенстве переходим к пределу при $\lambda \rightarrow 0$. Это тоже корректное доказательство, но приведенное выше не использует анализ в явном виде. Кроме того, при правильных словах оно годится даже для экзотических полей вроде конечных (если вы понимаете о чем я говорю).

³Не более чем n – его степень.

Семинар 2

Задачи:

1. Найти определитель следующей матрицы

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

2. Посчитайте характеристический многочлен матрицы

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

3. Найдите определители следующих матриц

$$(a) \begin{pmatrix} -t & & & & a_1 \\ a_2 & -t & & & \\ & & \ddots & & \\ & & & a_{n-1} & -t \\ & & & a_n & -t \end{pmatrix} \quad \text{и} \quad (b) \begin{pmatrix} \lambda & 1 & \dots & 1 & \lambda \\ 1 & \lambda & \dots & \lambda & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & \dots & \lambda & 1 \\ 1 & 1 & \dots & 1 & \lambda \end{pmatrix}$$

4. Найдите обратную матрицу

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

5. Пусть $X = (X_1 \mid \dots \mid X_n) \in M_n(\mathbb{R})$ и $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. Найти $\det(\lambda_1 X_1 X_1^t + \dots + \lambda_n X_n X_n^t)$.

6. Пусть $A \in M_n(\mathbb{R})$ – произвольная матрица. Построим из нее матрицу $B \in M_n(\mathbb{R})$ следующим образом: сдвинем все столбцы матрицы A по циклу на два вправо и результат прибавим к A . Выразите определитель B через определитель A .

Семинар 3

Конкретные векторные пространства

Основной объект изучения – пространство столбцов \mathbb{R}^n . Любая матрица $A \in M_{m,n}(\mathbb{R})$ определяет отображение $\mathbb{R}^n \xrightarrow{A} \mathbb{R}^m$ заданное $x \mapsto Ax$. В каком-то смысле это единственный пример, а потому – самый важный.

Однако есть и другие, внешне не похожие, примеры:

1. Матрицы $M_n(\mathbb{R})$. В качестве отображений нужно рассматривать $M_n(\mathbb{R}) \rightarrow M_m(\mathbb{R})$ заданное по правилу $X \mapsto \sum_i P_i X Q_i$, где $P_i, Q_i^t \in M_{m,n}(\mathbb{R})$.
2. Решения СЛУ $\{y \in \mathbb{R}^n \mid Ay = 0\}$ для некоторой $A \in M_{m,n}(\mathbb{R})$. Хороший вопрос: какие тут должны быть отображения между решениями?
3. Еще более интересный вопрос, а какие должны быть отображения между $M_n(\mathbb{R})$ и $\{y \in \mathbb{R}^m \mid Ay = 0\}$?

Именно поэтому нужна общая теория, которая бы помогла нам одним языком описать все эти случаи и придумать правильные определения, когда они не очевидны.

Абстрактные векторные пространства

В определении векторного пространства надо зафиксировать откуда берутся коэффициенты. Вариантов несколько: вещественные числа \mathbb{R} , комплексные числа \mathbb{C} , рациональные числа \mathbb{Q} .¹ Для простоты, все определения будем формулировать с вещественными числами.²

Определение. Векторное пространство над \mathbb{R} это следующие данные:

1. множество V .³
2. операция сложения векторов, т.е. отображение $+: V \times V \rightarrow V$ вида $(v_1, v_2) \mapsto v_1 + v_2$.
3. операция умножения векторов на число, т.е. отображение $\cdot: \mathbb{R} \times V \rightarrow V$ вида $(r, v) \mapsto rv$.

И эти данные удовлетворяют следующим аксиомам:

1. Для любых $v, u, w \in V$ верно $(v + u) + w = v + (u + w)$.
2. Существует вектор $0 \in V$ такой, что для любого вектора $v \in V$ имеем $0 + v = v + 0 = v$.
3. Для любого вектора $v \in V$ существует вектор $-v$ такой, что $v + (-v) = (-v) + v = 0$.
4. Для любых векторов $v, u \in V$ верно $v + u = u + v$.
5. Для любых $r \in \mathbb{R}$ и $v, u \in V$ верно $r(v + u) = rv + ru$.
6. Для любых $r_1, r_2 \in \mathbb{R}$ и $v \in V$ верно $(r_1 + r_2)v = r_1v + r_2v$.
7. Для любых $r_1, r_2 \in \mathbb{R}$ и $v \in V$ верно $(r_1 r_2)v = r_1(r_2 v)$.
8. Для любого $v \in V$ верно $1v = v$.

Обычно элементы V называют *векторами*, а элементы \mathbb{R} называют *скалярами*. Даже в абстрактном случае, полезно думать геометрически, представляя себе в первую очередь \mathbb{R}^n как главные пример.

¹На самом деле годится все что угодно, если в этом «чем угодно» можно складывать, вычитать, умножать и делить. Такие объекты называются *полями*.

²Определение векторного пространства может показаться жутко сложным и формальным. Не стоит бросаться учить наизусть все, что в нем находится. Главное понимать как с ним работать. Действительно, никто из нас не знает строгого определения \mathbb{R} , но это не мешает нам с ним работать!

³Это будет как раз множество векторов.

Замечания Полезно понимать, что указанный выше набор аксиом – это минимальный набор. Но из него можно вывести простыми но нудными манипуляциями еще полезные свойства, которые надо иметь перед глазами. В дальнейшем окажется, что все разумные свойства, к которым вы привыкли в \mathbb{R}^n выполняются в произвольном векторном пространстве. Потому интуитивно надо все время думать про \mathbb{R}^n , эта интуиция вас не подведет. А вот пример полезных свойств:

1. Нулевой вектор всегда единственный.
2. Для любого вектора $v \in V$ противоположный вектор $-v$ всегда единственный и совпадает с $-1 \cdot v$.
3. Если умножить число 0 на любой вектор, то получится нулевой вектор, то есть $0 \cdot v = 0$ (тут слева ноль – это число, а справа – вектор).
4. Аналогично, если умножить нулевой вектор на что угодно, то получится нулевой вектор, то есть $\lambda \cdot 0 = 0$ (тут с обеих сторон ноль означает нулевой вектор).

Определение. Пусть V – векторное пространство над \mathbb{R} , тогда подмножество $U \subseteq V$ называется подпространством, если

1. $0 \in U$
2. Если $u, v \in U$, то и $v + u \in U$.
3. Если $r \in \mathbb{R}$ и $v \in U$, то и $rv \in U$.

Замечания

- Отметим, что всякое подпространство само является векторным пространством.
- Обратите внимание, что первое свойство равносильно тому, что подпространство U обязательно не пусто. Действительно, если есть 0 вектор, то подпространство не пусто. Если же U не пусто, то там есть какой-то вектор v , тогда мы v можем умножить на число 0 и получим нулевой вектор внутри U по третьему свойству.

Примеры Пространства:

1. Пусть $V = \mathbb{R}^n$ с обычными операциями сложения и умножения на скаляр.
2. Пусть $V = M_{m,n}(\mathbb{R})$ с обычными операциями сложения и умножения на скаляр.
3. Пусть $V = \{0\} = \mathbb{R}^0$.
4. Пусть $\mathbb{R}[x] = \{a_0 + a_1x + \dots + a_nx^n \mid a_i \in \mathbb{R}, n \in \mathbb{N}\}$ – множество многочленов от переменной x с вещественными коэффициентами с обычными операциями сложения и умножения на число.
5. Пусть $C[0, 1] = \{f: [0, 1] \rightarrow \mathbb{R} \mid f \text{ непрерывна}\}$ – множество непрерывных функций на отрезке $[0, 1]$, тогда оно является векторным пространством над \mathbb{R} .

Подпространства (а значит тоже пространства):

1. $\{y \in \mathbb{R}^n \mid Ay = 0\}$ для некоторой $A \in M_{m,n}(\mathbb{R})$ с обычными операциями сложения и умножения на скаляр.
2. $\{X \in M_n(\mathbb{R}) \mid AXB = 0\}$ для некоторых $A \in M_{m,n}(\mathbb{R})$ и $B \in M_{n,k}(\mathbb{R})$ с обычными операциями сложения и умножения на скаляр.

Линейные комбинации

Пусть V – векторное пространство над \mathbb{R} . Главные вопросы: как надо про него думать? Ответ прост: это куча из элементов и эти элементы можно складывать и умножать на числа. То есть, мы всегда можем вытащить элементы⁴ $v_1, \dots, v_n \in V$ и начать их складывать с коэффициентами, получив некий новый вектор $r_1 v_1 + \dots + r_n v_n \in V$, где $r_i \in \mathbb{R}$. Поэтому все, что можно сказать про векторное пространство, обязательно формулируется в терминах таких выражений. Поэтому предлагается изучать подобные выражения.

Пусть $v_i \in V$ и $r_i \in \mathbb{R}$ как выше, тогда выражение $r_1 v_1 + \dots + r_n v_n$ называется *линейной комбинацией* векторов v_1, \dots, v_n . Линейная комбинация называется *тривиальной*, если все $r_i = 0$, в противном случае *нетривиальной*. Вектора v_1, \dots, v_n называются *линейно зависимыми*, если существует нетривиальная линейная комбинация этих векторов равная нулю (нулевому вектору), т.е. v_1, \dots, v_n – линейно зависимы, если существуют $r_1, \dots, r_n \in \mathbb{R}$ так, что хотя бы один из r_i не равен нулю и $r_1 v_1 + \dots + r_n v_n = 0$. В противном случае вектора называются *линейно независимыми*.

Пусть $v_i \in V$ – произвольный набор векторов, тогда набор v_1, \dots, v_n называется порождающим V (или просто порождающим), если любой вектор из V представляется в виде их линейной комбинации, то есть для любого $v \in V$ найдутся числа $r_1, \dots, r_n \in \mathbb{R}$ (не обязательно единственный) такой, что $v = r_1 v_1 + \dots + r_n v_n$. Обратите внимание, что если набор v_1, \dots, v_n является порождающим для V , то V – минимальное подпространство внутри V , содержащее все эти векторы. Любое другое подпространство хотя бы один из них не содержит.

Примеры

1. Рассмотрим $V = \mathbb{R}^3$ и пусть

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad v_4 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad v_5 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix},$$

Тогда вектора v_1, v_2, v_3 линейно независимы. Вектора v_1, v_2, v_4 тоже линейно независимы. Но вот вектора v_1, v_2, v_3, v_4 уже зависимы, так как $v_1 + v_2 + v_3 - v_4 = 0$. Также зависимы вектора v_1, v_2 и v_5 , ибо $v_1 + v_2 - v_5 = 0$.

2. Один вектор $v \in V$ линейно зависим тогда и только тогда, когда он равен нулю.
3. Два вектора $v, u \in V$ линейно зависимы тогда и только тогда, когда они пропорциональны, т.е. либо $v = \lambda u$ либо $u = \lambda v$.

Базис

Утверждение. Пусть V – векторное пространство над \mathbb{R} и пусть $v_1, \dots, v_n \in V$. Следующие утверждения эквивалентны:

1. Вектора v_1, \dots, v_n линейно независимы и набор v_1, \dots, v_n является порождающим.
2. Вектора v_1, \dots, v_n – максимальный линейно независимый набор, то есть для любого $u \in V$ вектора v_1, \dots, v_n, u уже линейно зависимы.
3. Вектора v_1, \dots, v_n – минимальный порождающий набор, то есть для любого i набор $\{v_1, \dots, v_n\} \setminus \{v_i\}$ уже не порождающий.

Доказательство. (1) \Rightarrow (2). Возьмем любой $u \in V$, тогда $u = \lambda_1 v_1 + \dots + \lambda_n v_n$, тогда $\lambda_1 v_1 + \dots + \lambda_n v_n - u = 0$ – линейная зависимость между векторами.

(2) \Rightarrow (1). Пусть $u \in V$, тогда существует какая-то линейная зависимость $\lambda_1 v_1 + \dots + \lambda_n v_n + \mu u = 0$. Если $\mu = 0$, то вектора v_i линейно зависимы, но это не так. Значит $\mu \neq 0$. Тогда на него можно поделить и получим $u = -\frac{\lambda_1}{\mu} v_1 - \dots - \frac{\lambda_n}{\mu} v_n$.

(1) \Rightarrow (3). Так как вектора v_1, \dots, v_n линейно независимы, то ни один из v_i не выражается через оставшиеся (иначе бы мы нашли нетривиальную линейную комбинацию). А значит, выкинув любой из v_i мы обязательно получим систему, линейные комбинации которой не содержат v_i , то есть не порождающую систему.

⁴называемые векторами

(3) \Rightarrow (1). Надо доказать линейную независимость векторов. Если бы это было не так, то нашлась бы нетривиальная линейная комбинация этих векторов и как в доказательстве (2) \Rightarrow (1) мы бы могли выразить какой-то v_i через оставшиеся. Давайте предположим для простоты, что это v_n . Но тогда v_1, \dots, v_{n-1} была бы меньшая система порождающих. Действительно, любой вектор v расписывается как $r_1 v_1 + \dots + r_{n-1} v_{n-1} + r_n v_n$. Но теперь мы можем выразить v_n через v_1, \dots, v_{n-1} . Подставляя это выражение, получаем разложение для v через v_1, \dots, v_{n-1} . Получено противоречие с минимальностью системы векторов v_1, \dots, v_n . Значит она была линейно независима. \square

Если в векторном пространстве V существует система векторов v_1, \dots, v_n обладающая одним из свойств выше, то мы будем называть такую систему векторов *базисом* пространства V .

Описание всех векторных пространств с базисами

Пусть V – векторное пространство над \mathbb{R} и пусть у нас есть базис $v_1, \dots, v_n \in V$. Тогда любой вектор $u \in V$ единственным образом представляется в виде $u = a_1 v_1 + \dots + a_n v_n$. Надо лишь объяснить единственность. Если $u = b_1 v_1 + \dots + b_n v_n$, то $(a_1 - b_1)v_1 + \dots + (a_n - b_n)v_n = 0$ – линейная комбинация равна нулю. Но так как v_i линейно независимы, это может быть лишь тривиальная линейная комбинация, т.е. все $a_i - b_i = 0$. Таким образом получаем биекцию между множеством векторов V и пространством столбцов \mathbb{R}^n , а именно каждому вектору v сопоставим столбец его координат. Тогда обратное отображение действует по правилу $u = \sum_{i=1}^n a_i v_i \mapsto \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$. Кроме того обратите, что это сопоставление переводит сумму векторов в сумму столбцов,

а также вектор умноженный на число соответствует столбцу умноженному на число. То есть эта биекция согласована со структурой векторного пространства. По сути это означает, что эти два пространства устроены одинаково с точки зрения теории векторных пространств. Любое свойство V , которое можно вытащить через «интерфейс» векторного пространства будет выполнено тогда и только тогда, когда то же самое свойство будет выполнено в \mathbb{R}^n .

Теперь важное замечание. Базис существует всегда! Только не всегда он состоит из конечного числа векторов. Я не хочу обсуждать бесконечные базисы. Однако, важно понимать следующее.

Утверждение. Пусть V – векторное пространство. Любые два базиса V имеют одинаковое число элементов.⁵

Если V – векторное пространство, то число элементов в базисе называется *размерностью* векторного пространства V и обозначается $\dim V$. Если конечного базиса нет, будем писать $\dim V = \infty$.⁶

Примеры Для простых пространств размерность в точности равна числу коэффициентов, которые необходимы для задания векторов: $\dim \mathbb{R}^n = n$ или $\dim M_{m,n}(\mathbb{R}) = mn$. Чуть позже мы увидим, что для $V = \{y \in \mathbb{R}^n \mid Ay = 0\}$, где $A \in M_{m,n}(\mathbb{R})$, $\dim V$ равна числу свободных переменных СЛУ $Ay = 0$. То есть это число характеризует на сколько много есть решений у системы.

Утверждение. Пусть V – векторное пространство и пусть $U \subseteq V$ – подпространство.

1. $\dim U \leq \dim V$. В частности, если V обладает конечным базисом, то и U обязательно обладает конечным базисом.
2. $\dim U = \dim V$ тогда и только тогда, когда $U = V$.

Размерность – это величина, показывающая на сколько векторное пространство большое и характеризует «количество степеней свободы» в пространстве. Кроме того, это понятие согласовано с нашей интуицией: прямая \mathbb{R}^1 имеет размерность 1, плоскость \mathbb{R}^2 – размерность 2, а пространство \mathbb{R}^3 – размерность 3.

Смысл базиса

Напомню, что «по-простому» векторное пространство – это все что угодно, где элементы можно складывать и умножать на числа. Формально надо проверить еще какие-то аксиомы, но все, что возникает в реальной жизни, обязательно будет удовлетворять им. Главные примеры – \mathbb{R}^n и $M_{m,n}(\mathbb{R})$.

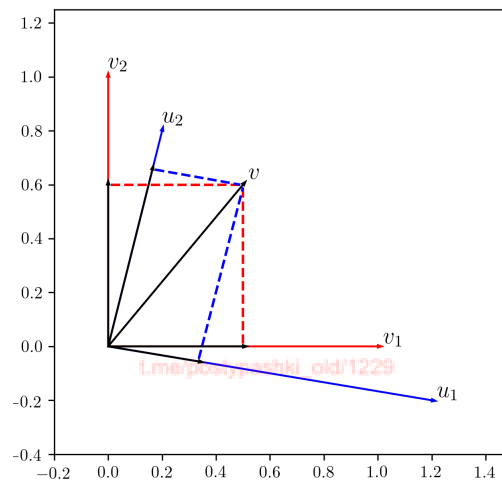
⁵Это утверждение верно и для конечных и для бесконечных базисов, но для работы с бесконечными базисами надо знать как сравнивать бесконечные множества.

⁶На самом деле, теория множеств позволяет различать какие-то из бесконечных множеств, но мы этого делать не будем.

Пусть V – векторное пространство. Напомню, что базис – это набор векторов v_1, \dots, v_n который линейно независим и через них все выражается. То есть уравнение $x_1 v_1 + \dots + x_n v_n = 0$ имеет только нулевое решение $x_i = 0$ и любой вектор $v \in V$ представляется (по безысходности единственным образом) в виде $v = \lambda_1 v_1 + \dots + \lambda_n v_n$. Смысл базиса вот в чем: если вы его выбрали, то вы можете отождествить V с \mathbb{R}^n следующим образом: если $v = \lambda_1 v_1 + \dots + \lambda_n v_n$, то ему соответствует единственный столбец $\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$. Таким образом, про любое векторное пространство, когда это необходимо, можно думать как просто про \mathbb{R}^n . Но надо помнить, что (1) часто необязательно выбирать базис и без выбора базиса задачи могут проще решаться и (2) базис можно выбрать не единственным образом, и если его выбрать по-другому вычисления могут стать либо проще либо сложнее (как повезет).

Смена базиса

На рисунке ниже изображена плоскость с двумя базисами: красный v_1, v_2 и синий u_1, u_2 . При этом вектор v можно разложить как по одному, так и по другому базису. В зависимости от этого у него будут разные координаты.



Давайте теперь поговорим, как это устроено в общем случае. Пусть у нас есть векторное пространство V . Можно считать, что $V = \mathbb{R}^n$ для удобства. Пусть у нас в V есть два базиса: e_1, \dots, e_n и f_1, \dots, f_n .⁷ И пусть у нас есть вектор $v \in V$. Тогда он раскладывается по обоим базисам:

$$v = x_1 e_1 + \dots + x_n e_n = \begin{pmatrix} e_1 & \dots & e_n \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{и} \quad v = y_1 f_1 + \dots + y_n f_n = \begin{pmatrix} f_1 & \dots & f_n \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

Вопрос: а как связаны координаты x_i и координаты y_i ? Вот на этот вопрос мы и попытаемся ответить. Для начала надо знать, как связаны базисы e_i и f_i . По определению базиса для e_i каждый вектор f_i представляется в виде:

$$f_i = c_{1i} e_1 + \dots + c_{ni} e_n = \begin{pmatrix} e_1 & \dots & e_n \end{pmatrix} \begin{pmatrix} c_{1i} \\ \vdots \\ c_{ni} \end{pmatrix}$$

Если положить $C = (c_{ij}) \in M_n(\mathbb{R})$, то предыдущий набор равенств можно записать кратко в виде:

$$\begin{pmatrix} f_1 & \dots & f_n \end{pmatrix} = \begin{pmatrix} e_1 & \dots & e_n \end{pmatrix} \begin{pmatrix} c_{11} & \dots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \dots & c_{nn} \end{pmatrix} = \begin{pmatrix} e_1 & \dots & e_n \end{pmatrix} C$$

⁷Напомню, что базисы обязательно имеют одинаковый размер.

Такая матрица называется *матрицей перехода* от e_i к f_i . Теперь запишем наш вектор v так

$$v = (f_1 \quad \dots \quad f_n) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = (e_1 \quad \dots \quad e_n) C \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \quad \text{и} \quad v = (e_1 \quad \dots \quad e_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Так как разложение по базису e_i однозначно (по определению базиса), то получаем связь на координаты

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = C^{-1} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Запоминать это правило надо так: если от базиса e к базису f мы перешли с помощью умножения справа на матрицу C , то на координатах у нас отображение в обратную сторону с помощью умножения на матрицу C слева (то есть тоже с другой стороны). Еще полезно держать перед глазами вот эту таблицу.

базис	новый \xleftarrow{C} старый
координаты	новые \xrightarrow{C} старые

Смена координат в \mathbb{R}^n

В случае, когда мы работаем в \mathbb{R}^n вот как можно думать про равенство

$$(f_1 \quad \dots \quad f_n) = (e_1 \quad \dots \quad e_n) C$$

В этом случае каждый вектор f_i – это вектор столбец высоты n . Потому левая часть равенства (когда там записаны n столбцов) представляет из себя матрицу n на n . То есть $(f_1 \dots f_n) \in M_n(\mathbb{R})$. Аналогично можно думать, что $(e_1 \dots e_n) \in M_n(\mathbb{R})$. Таким образом можно найти матрицу перехода: $C = (e_1 \dots e_n)^{-1} (f_1 \dots f_n)$. Тут есть важный частный случай, предположим, что e_i – стандартный базис, т.е. e_i имеет 1 на i -ом месте и нули в остальных местах. Тогда матрица перехода $C = (f_1 \dots f_n)$. То есть C составлена из координат f_i в стандартном базисе.

Линейные оболочки

Пусть V – векторное пространство. Для простоты можно думать, что это \mathbb{R}^n . И пусть у нас задан произвольный набор векторов $v_1, \dots, v_k \in V$.⁸ Понятно, что конечный набор не образует подпространство, но можно рассмотреть наименьшее подпространство, содержащее данные вектора. Это подпространство обозначается $\langle v_1, \dots, v_k \rangle$ и состоит из всех линейных комбинаций v_i , т.е.

$$\langle v_1, \dots, v_k \rangle = \{ \lambda_1 v_1 + \dots + \lambda_k v_k \in V \mid \lambda_i \in \mathbb{R} \}$$

Заметим, что $\dim \langle v_1, \dots, v_k \rangle \leq k$ и равенство достигается тогда и только тогда, когда v_i линейно независимы. Кроме того, по определению v_1, \dots, v_k являются порождающими для линейной оболочки $\langle v_1, \dots, v_k \rangle$.

Пусть вектора v_1, \dots, v_k линейно зависимы. Выделим среди них наибольшее линейно независимое подмножество. После перенумерации векторов, можно считать, что это v_1, \dots, v_m , где $m \leq k$. Тогда каждый вектор v_j при $j > m$ будет линейно выражаться через первые m векторов. Последнее означает, что $\langle v_1, \dots, v_k \rangle = \langle v_1, \dots, v_m \rangle$, но теперь вектора справа линейно независимы.

Выделение базиса из системы векторов

Дано Пусть $v_1, \dots, v_m \in \mathbb{R}^n$ – вектора и $V = \langle v_1, \dots, v_m \rangle$ – их линейная оболочка.

Задача Среди векторов v_1, \dots, v_m найти базис пространства V и разложить оставшиеся вектора по этому базису.

⁸Вектора могут быть любые, могут быть линейно зависимыми или независимыми, могут быть хоть все нулевыми или просто одинаковыми.

Алгоритм

1. Запишем вектора v_1, \dots, v_m по столбцам в матрицу $A \in M_{nm}(\mathbb{R})$. Например, при $n = 3, m = 5$

$$A = \begin{pmatrix} v_{11} & v_{21} & v_{31} & v_{41} & v_{51} \\ v_{12} & v_{22} & v_{32} & v_{42} & v_{52} \\ v_{13} & v_{23} & v_{33} & v_{43} & v_{53} \end{pmatrix}$$

2. Приведем матрицу A элементарными преобразованиями строк к улучшенному ступенчатому виду. Например

$$A' = \begin{pmatrix} 1 & 0 & a_{31} & 0 & a_{51} \\ 0 & 1 & a_{32} & 0 & a_{52} \\ 0 & 0 & 0 & 1 & a_{53} \end{pmatrix}$$

3. Пусть k_1, \dots, k_r – номера главных позиций в матрице A' . Тогда вектора v_{k_1}, \dots, v_{k_r} образуют базис V . Например, в примере выше это вектора v_1, v_2 и v_4 .
4. Пусть v_i – вектор соответствует неглавной позиции в A' . Тогда в i -ом столбце A' записаны координаты разложения v_i через найденный базис выше. Например, в примере выше $v_3 = a_{31}v_1 + a_{32}v_2$ и $v_5 = a_{51}v_1 + a_{52}v_2 + a_{53}v_4$.

Пример Пусть

$$v_1 = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, v_3 = \begin{pmatrix} 5 \\ 12 \\ 7 \end{pmatrix}, v_4 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, v_5 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \in \mathbb{R}^3$$

Тогда

$$\begin{pmatrix} 1 & 1 & 5 & 1 & -1 \\ 3 & 2 & 12 & 1 & 1 \\ 2 & 1 & 7 & 1 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 1 & 5 & 1 & -1 \\ 0 & 0 & 0 & -1 & 2 \\ 2 & 1 & 7 & 1 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 1 & 5 & 1 & -1 \\ 0 & 0 & 0 & -1 & 2 \\ 1 & 0 & 2 & 0 & 1 \end{pmatrix} \mapsto$$

$$\mapsto \begin{pmatrix} 1 & 0 & 2 & 0 & 1 \\ 1 & 1 & 5 & 1 & -1 \\ 0 & 0 & 0 & 1 & -2 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 2 & 0 & 1 \\ 0 & 1 & 3 & 1 & -2 \\ 0 & 0 & 0 & 1 & -2 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 2 & 0 & 1 \\ 0 & 1 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 & -2 \end{pmatrix}$$

Тогда v_1, v_2 и v_4 – базис линейной оболочки. $v_3 = 2v_1 + 3v_2$ и $v_5 = v_1 - 2v_4$.

Нахождение какого-то базиса линейной оболочки

Дано Пусть $v_1, \dots, v_m \in \mathbb{R}^n$ – вектора и $V = \langle v_1, \dots, v_m \rangle$ – их линейная оболочка.

Задача Найти какой-нибудь базис подпространства V .

Алгоритм

1. Уложить все вектора v_i в строки матрицы $A \in M_{mn}(\mathbb{R})$.
2. Элементарными преобразованиями строк привести матрицу к ступенчатому виду.
3. Ненулевые строки полученной матрицы будут искомым базисом.

Дополнение линейно независимой системы до базиса всего пространства стандартными векторами

Дано Пусть $v_1, \dots, v_m \in \mathbb{R}^n$ – линейно независимая система векторов, $V = \langle v_1, \dots, v_m \rangle$ – их линейная оболочка и e_i – стандартные базисные векторы, т.е. на i -ом месте стоит 1, а в остальных 0.

Задача Найти такие вектора $e_{k_1}, \dots, e_{k_{n-m}}$, что система $v_1, \dots, v_m, e_{k_1}, \dots, e_{k_{n-m}}$ является базисом \mathbb{R}^n .

Алгоритм

1. Уложить вектора v_i в строки матрицы $A \in M_{m \times n}(\mathbb{R})$.
2. Привести матрицу A к ступенчатому виду.
3. Пусть k_1, \dots, k_{n-m} – номера неглавных столбцов. Тогда $e_1, \dots, e_{k_{n-m}}$ – искомое множество.

Подпространства в \mathbb{R}^n

Мы хотим понять как устроены все возможные подпространства в \mathbb{R}^n . Для начала надо понять, а как вообще задавать подпространства в \mathbb{R}^n . Существует два способа:

1. С помощью образующих векторов: $U \subseteq \mathbb{R}^n$ задано в виде $U = \langle v_1, \dots, v_k \rangle$, где $v_1, \dots, v_k \in \mathbb{R}^n$ – некоторые вектора. В этом случае часто бывает полезно, чтобы вектора v_i были линейно независимыми.
2. С помощью СЛУ: $U \subseteq \mathbb{R}^n$ задано в виде $U = \{y \in \mathbb{R}^n \mid Ay = 0\}$, где $A \in M_{m \times n}(\mathbb{R})$ – некоторая матрица. В этом случае часто бывает полезно, чтобы строки матрицы A были линейно независимыми.

Любое пространство можно задать любым из этих двух способов, а значит, если пространство задано одним из этих способов, его можно задать и другим.

Фундаментальная система решений (ФСР)

Так как \mathbb{R}^n обладает конечным базисом, то и $\{y \in \mathbb{R}^n \mid Ay = 0\}$, где $A \in M_{m \times n}(\mathbb{R})$, тоже обладает конечным базисом, причем количество базисных элементов не превосходит n . Любой базис такого пространства называется *фундаментальной системой решений*. Наша задача научиться находить его.

Дано Матрица $A \in M_{m \times n}(\mathbb{R})$.

Задача Найти базис пространства $U = \{y \in \mathbb{R}^n \mid Ay = 0\}$.

Алгоритм

1. Приведем матрицу A к улучшенному ступенчатому виду. Пусть например она имеет вид

$$\begin{pmatrix} 1 & a_{12} & 0 & a_{14} & 0 & a_{16} \\ 0 & 0 & 1 & a_{24} & 0 & a_{26} \\ 0 & 0 & 0 & 0 & 1 & a_{36} \end{pmatrix}$$

2. Теперь $\dim U$ равна количеству свободных переменных. ФСР строится так: для каждой свободной переменной будет свой базисный вектор. Такую свободную переменную полагаем 1, а остальные свободные переменные 0. После чего рассчитываем значения главных переменных. В примере выше, свободные переменные x_2 , x_4 и x_6 . Тогда ФСР

$$v_2 = \begin{pmatrix} -a_{12} \\ \underline{1} \\ 0 \\ \underline{0} \\ 0 \\ \underline{0} \end{pmatrix}, \quad v_4 = \begin{pmatrix} -a_{14} \\ \underline{0} \\ -a_{24} \\ \underline{1} \\ 0 \\ \underline{0} \end{pmatrix}, \quad v_6 = \begin{pmatrix} -a_{16} \\ \underline{0} \\ -a_{26} \\ \underline{0} \\ -a_{36} \\ \underline{1} \end{pmatrix},$$

В векторах выше подчеркнуты позиции свободных переменных, которые мы задаем сами.

Ранг системы векторов

Пусть V – некоторое векторное пространство. Системой векторов называется последовательность (v_1, \dots, v_k) из векторов V , в которой векторы v_i могут повторяться.⁹

По определению рангом системы (v_1, \dots, v_k) называется максимальное количество линейно независимых векторов в этой системе. Ранг такой системы будет обозначаться $\text{rk}(v_1, \dots, v_k)$.

Утверждение. Если (v_1, \dots, v_k) – некоторая система векторов в векторном пространстве V , то $\text{rk}(v_1, \dots, v_k) = \dim\langle v_1, \dots, v_k \rangle$.

Матричный ранг

Пусть $A \in M_{m,n}(\mathbb{R})$ – некоторая матрица. Сейчас я определю пять разных определений ранга матрицы. Все эти ранги между собой совпадают и полученная величина будет просто называться рангом матрицы A и обозначаться $\text{rk } A$.

Определение. Пусть $A_1, \dots, A_n \in \mathbb{R}^m$ – столбцы матрицы A , то есть $A = (A_1 | \dots | A_n)$. Тогда столбцовым рангом матрицы A называется ранг системы (A_1, \dots, A_n) , то есть $\text{rk}_{\text{столб}} A = \text{rk}(A_1, \dots, A_n)$.

Определение. Пусть $A_1, \dots, A_m \in \mathbb{R}^n$ – строки матрицы A , то есть $A^t = (A_1 | \dots | A_m)$. Тогда строковым рангом матрицы A называется ранг системы (A_1, \dots, A_m) , то есть $\text{rk}_{\text{стр}} A = \text{rk}(A_1, \dots, A_m)$.

Определение. Факториальным рангом матрицы A называется следующее число

$$\min\{k \mid A = BC, \text{ где } B \in M_{m,k}(\mathbb{R}), C \in M_{k,n}(\mathbb{R})\}$$

то есть это минимальное число k такое, что матрица A представима в виде произведения матриц BC , где общая размерность для B и C , по которой они перемножаются, есть k .

Определение. Тензорным рангом матрицы A называется следующее число

$$\min\{k \mid A = x_1 y_1^t + \dots + x_k y_k^t, \text{ где } x_i \in \mathbb{R}^m, y_i \in \mathbb{R}^n\}$$

то есть это минимальное число k такое, что матрица A представима в виде суммы k «тощих» матриц вида xy^t , где $x \in \mathbb{R}^m$ и $y \in \mathbb{R}^n$.

Если я в матрице A выделю какой-нибудь набор из k строк и одновременно набор из k столбцов, а потом возьму матрицу составленную из элементов на пересечении этих строк и столбцов, то я получу квадратную матрицу размера k . Такие матрицы мы будем называть квадратными подматрицами матрицы A .

Определение. Минорным рангом матрицы A называется размер наибольшей невырожденной квадратной подматрицы.¹⁰

Главное для нас следующее утверждение.

Утверждение. Для любой матрицы $A \in M_{m,n}(\mathbb{R})$ все пять видов ранга совпадают и не превосходят $\min(m, n)$.

Примеры

1. В начале заметим, что матрица имеет ранг 0 тогда и только тогда, когда $A = 0$.
2. Ранг матрицы A равен единице тогда и только тогда, когда она не нулевая и все столбцы пропорциональны одному общему столбцу (или что эквивалентно, все строки пропорциональны одной общей строке). Если воспользоваться определением факториального ранга, то мы видим, что тогда матрица A имеет вид $A = xy^t$, где $x \in \mathbb{R}^m$ и $y \in \mathbb{R}^n$ – ненулевые вектора.

⁹В подобной ситуации повторяющиеся векторы различаются по индексу – «ключу».

¹⁰На самом деле можно дать более сильное определение, а именно, минорный ранг – это размер любой максимальной невырожденной подматрицы. То есть мы берем какую-то квадратную подматрицу, которая невырождена, а любая большая подматрица уже вырождена. Оказывается, что все максимальные невырожденные подматрицы имеют одинаковый размер и он называется минорным рангом.

Свойства ранга

Прежде всего надо запомнить как ранг связан с матричными операциями.

Утверждение. Пусть $A, B \in M_{m,n}(\mathbb{R})$, тогда

$$|\operatorname{rk} A - \operatorname{rk} B| \leq \operatorname{rk}(A + B) \leq \operatorname{rk} A + \operatorname{rk} B$$

Надо понимать, что, во-первых, все эти эффекты можно увидеть на диагональных матрицах; во-вторых, все границы неравенств достигаются. Смысл этого утверждения вот в чем: если вы шевелите матрицу A с помощью матрицы B , то ранг A может измениться не более чем на ранг B в любую сторону. Теперь посмотрим на матрицы: $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ и $C = -A$. Тогда $\operatorname{rk}(A + B) = \operatorname{rk} A + \operatorname{rk} B$ и $\operatorname{rk}(A + C) = \operatorname{rk} A - \operatorname{rk} C$.

Утверждение. Пусть $A \in M_{m,n}(\mathbb{R})$ и $B \in M_{n,k}(\mathbb{R})$, тогда

$$\operatorname{rk} A + \operatorname{rk} B - n \leq \operatorname{rk}(AB) \leq \min(\operatorname{rk} A, \operatorname{rk} B)$$

Как и в предыдущем случае, все обе границы неравенства достигаются и все можно пронаблюдать на диагональных матрицах. Пусть $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ и $B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. Тогда $\operatorname{rk}(AA) = \operatorname{rk} A$ и $\operatorname{rk}(AB) = \operatorname{rk} A + \operatorname{rk} B - 2$.

Утверждение. Пусть $A \in M_n(\mathbb{R})$ – квадратная матрица. Тогда $\operatorname{rk} A = n$ тогда и только тогда, когда A невырождена, т.е. $\det A \neq 0$.

Таким образом на ранг можно смотреть как на степень невырожденности матрицы A . Самый высокий ранг у невырожденных матриц, самый маленький у нулевой, но есть еще и промежуточные состояния.

Утверждение. Если матрица $A \in M_{m,n}(\mathbb{R})$ находится в ступенчатом виде и имеет k ступенек, то ее ранг равен k .

Это утверждение вместе со следующим дают эффективный способ считать ранг.

Утверждение. Для любой матрицы $A \in M_{m,n}(\mathbb{R})$ и любых невырожденных матриц $C \in M_m(\mathbb{R})$ и $D \in M_n(\mathbb{R})$ верно: $\operatorname{rk} A = \operatorname{rk}(CA) = \operatorname{rk}(AD)$.¹¹

В частности ранг матрицы не меняется при элементарных преобразованиях столбцов и строк. Обычно этим пользуются для нахождения ранга. Более того, если $A \in M_{m,n}(\mathbb{R})$ имеет ранг r , то элементарными преобразованиями строк и столбцов она приводится к виду

$$A \mapsto \begin{pmatrix} E & 0 \\ 0 & 0 \end{pmatrix}, \text{ где } E \in M_r(\mathbb{R}) \text{ – единичная матрица}$$

Следствием данного замечания является следующее.

Утверждение. Для любых матриц $A \in M_{m,n}(\mathbb{R})$ и $B \in M_{s,t}(\mathbb{R})$ имеем

$$\operatorname{rk} \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} = \operatorname{rk} A + \operatorname{rk} B$$

¹¹В частности ранг не меняется при элементарных преобразованиях строк и столбцов.

Семинар 3

Задачи:

1. Даны векторы

$$a_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, a_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, a_3 = \begin{pmatrix} 3 \\ 8 \\ 11 \end{pmatrix}, a_4 = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}, a_5 = \begin{pmatrix} 1 \\ 3 \\ 3 \end{pmatrix}$$

Среди этих векторов найти базис их линейной оболочки и выразить все оставшиеся вектора через базисные.

2. Найдите базис векторного пространства $U = \{y \in \mathbb{R}^5 \mid Ay = 0\}$, где

$$A = \begin{pmatrix} 2 & 1 & 3 & 0 & 7 \\ 1 & 2 & 0 & 1 & 4 \\ 0 & 1 & -1 & 1 & 0 \end{pmatrix}$$

3. Определите можно ли из системы векторов

$$v_1 = \begin{pmatrix} -4 \\ -5 \\ 2 \\ 3 \end{pmatrix}, v_2 = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 1 \end{pmatrix}, v_3 = \begin{pmatrix} 5 \\ 5 \\ 3 \\ -2 \end{pmatrix}, v_4 = \begin{pmatrix} 4 \\ 5 \\ -2 \\ -3 \end{pmatrix}$$

выбрать ФСР для системы линейных уравнений

$$\begin{cases} x_1 - x_3 + 2x_4 & = 0 \\ x_2 - 2x_3 + 3x_4 & = 0 \\ 2x_1 - 2x_2 + 2x_3 - 2x_4 & = 0 \end{cases}$$

4. Являются ли функции $1, \sin(x), \sin^2(x), \dots, \sin^n(x)$ линейно зависимыми?
5. Пусть $\mathbb{R}[x]_n$ – множество всех многочленов с вещественными коэффициентами степени не больше n . Показать, что системы

$$\{1, x, x^2, \dots, x^n\} \text{ и } \{1, x - a, (x - a)^2, \dots, (x - a)^n\}, \text{ где } a \in \mathbb{R}$$

являются базисами в $\mathbb{R}[x]_n$ и найти матрицы перехода от первого базиса ко второму и от второго к первому.

6. Найти ранг следующей матрицы при различных значениях параметра λ :

$$\begin{pmatrix} 7 - \lambda & -12 & 6 \\ 10 & -19 - \lambda & 10 \\ 12 & -24 & 13 - \lambda \end{pmatrix}$$

7. Пусть A и B – квадратные матрицы одного размера. Доказать, что

$$\text{rk} \begin{pmatrix} A & AB \\ B & B + B^2 \end{pmatrix} = \text{rk } A + \text{rk } B$$

Семинар 4

Линейные отображения

Сами по себе векторные пространства не интересны. Нам бы хотелось уметь их сравнивать между собой. Для этого нам нужны линейные отображения. Кроме того, многие вопросы, возникающие в линейной алгебре формулируются именно в терминах линейных отображений.

Определение. Пусть V и U – векторные пространства над \mathbb{R} . Отображение $\phi: V \rightarrow U$ называется линейным, если

1. $\phi(v_1 + v_2) = \phi(v_1) + \phi(v_2)$ для любых $v_1, v_2 \in V$.
2. $\phi(rv) = r\phi(v)$ для любых $r \in \mathbb{R}$ и $v \in V$.

Множество всех линейных отображений из V в U обозначается $\text{Hom}(V, U)$. Если надо подчеркнуть какие скаляры имеются в виду, можно написать $\text{Hom}_{\mathbb{R}}(V, U)$.

Примеры

1. Любое линейное отображение $\phi: \mathbb{R}^n \rightarrow \mathbb{R}^m$ задается в виде $\phi(x) = Ax$, где $A \in M_{m \times n}(\mathbb{R})$. Таким образом множество $\text{Hom}_{\mathbb{R}}(V, U)$ отождествляется с множеством матриц $M_{m \times n}(\mathbb{R})$.
2. Любое линейное отображение $\phi: M_n(\mathbb{R}) \rightarrow M_m(\mathbb{R})$ задается единственным образом в виде $\phi(X) = \sum_{i,k=1}^m \sum_{j,l=1}^n a_{ijkl} E_{ij} X E_{lk}$, где E_{ij} – матричная единица, т.е. матрица такая, что на i -ой строке j -ом столбце стоит 1, а все остальные элементы равны 0.

Определение. Если V – векторное пространство и $\phi: V \rightarrow V$ – линейное отображение, то ϕ называется *линейным оператором*.¹

Правильно думать про линейные операторы как про «линейные деформации пространства V ». Например, в \mathbb{R}^n мы можем делать растяжения вдоль координатных осей (на самом деле растяжения вдоль любых прямых годятся). Или можем делать повороты вокруг каких-то прямых. Можно «наклонить» одну координатную ось, зеркальная симметрия, симметрия относительно прямой, плоскости, проекция вектора на прямую, плоскость и еще куча других преобразований описывается линейными операторами.

Определение. Пусть V и U – векторные пространства и $\phi: V \rightarrow U$ – линейное отображение. Мы говорим, что оно является *изоморфизмом* (а пространства V и U *изоморфными*), если существует $\psi: U \rightarrow V$ – линейное отображение такое, что $\phi\psi = \text{Id}$ и $\psi\phi = \text{Id}$.²

Про изоморфные пространства надо думать как про одинаковые. На множество V можно смотреть как на «имена векторов», соответственно, U – множество «новых имен», а ϕ – это переименование наших векторов. А раз это всего лишь переименование, то от него ничего не должно зависеть. Потому изучать V – это все равно, что изучать U .

Важный вопрос: а как задавать линейные отображения и операторы? Оказывается для этого достаточно знать куда отправляется базис.

Утверждение. Пусть e_1, \dots, e_n – некоторый базис векторного пространства V и u_1, \dots, u_n – произвольный набор векторов другого пространства U . Тогда существует единственное линейное отображение $\phi: V \rightarrow U$ такое, что $\phi(e_i) = u_i$.

Доказательство. Действительно, пусть $v = x_1 e_1 + \dots + x_n e_n$ – произвольный вектор из V . Тогда, если ϕ существует, то он должен действовать по правилу

$$\phi(v) = \phi(x_1 e_1 + \dots + x_n e_n) = x_1 \phi(e_1) + \dots + x_n \phi(e_n) = x_1 u_1 + \dots + x_n u_n$$

С другой стороны, легко видеть, что данное равенство однозначно задает линейное отображение. \square

¹Технически линейный оператор ничем не отличается от линейного отображения, просто мы требуем чтобы мы действовали на одном пространстве, а не между двумя. Однако, это порождает огромную разницу в поведении этих объектов и чтобы не путать их между собой люди даже специально ввели для линейных отображений на одном пространстве отдельное название – «оператор».

²Другими словами ϕ обратимо, где обратный $\phi^{-1} = \psi$.

Примеры

- В частности этот критерий позволяет отвечать на вопросы следующего вида: существует ли отображение $\phi: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, со следующим свойством

$$\phi \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad \phi \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \quad \phi \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

В данном случае векторы

$$v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

являются базисом, а

$$v_3 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{2}(v_1 + v_2)$$

По утверждению, векторы v_1 и v_2 можно отправить куда угодно и тогда найдется единственное $\phi: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ со свойствами

$$\phi \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad \phi \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

Теперь осталось лишь проверить, удовлетворяет ли наше ϕ последнему свойству. С одной стороны мы хотим, чтобы

$$\phi \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

С другой стороны, как мы выяснили $v_3 = \frac{1}{2}(v_1 + v_2)$. Значит

$$\phi(v_3) = \frac{1}{2}(\phi(v_1) + \phi(v_2)) = \frac{1}{2} \left(\begin{pmatrix} -1 \\ 1 \end{pmatrix} + \begin{pmatrix} 2 \\ 0 \end{pmatrix} \right) = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Не сходится. Значит, не существует. Если бы сошлось, то существовал бы. Отметим, что наивный подход заключается в том, чтобы задать отображение ϕ в виде $x \mapsto Ax$, где $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.³ Тогда условия на ϕ можно переписать как систему линейных уравнений на a, b, c, d . Три вектора, по две координаты, будет всего 6 условий и 4 неизвестные. Это намного неприятнее, чем предложенный выше метод.

- А что если нам даны векторы v_1, \dots, v_k в \mathbb{R}^n и векторы u_1, \dots, u_k в \mathbb{R}^m и нас спрашивают существует ли линейное отображение $\phi: \mathbb{R}^n \rightarrow \mathbb{R}^m$ такое, что $\phi(v_i) = u_i$. Выше я рассказывал, как решать задачу, если $\langle v_1, \dots, v_k \rangle = \mathbb{R}^n$. На самом деле был изложен способ понять, существует ли линейное отображение $\psi: \langle v_1, \dots, v_k \rangle \rightarrow \mathbb{R}^m$ такое, что $\psi(v_i) = u_i$. Если такое отображение не существует на подпространстве $\langle v_1, \dots, v_k \rangle$, то очевидно, что оно не существует на всем пространстве \mathbb{R}^n . То есть в негативном случае задача решается проще. Однако, если же отображение $\psi: \langle v_1, \dots, v_k \rangle \rightarrow \mathbb{R}^m$ найдется. То искомое ϕ можно построить так. Пусть v_1, \dots, v_s – базис в $\langle v_1, \dots, v_k \rangle$. Так как это линейно независимое множество в \mathbb{R}^n его можно дополнить до базиса в \mathbb{R}^n . То есть мы можем найти векторы $w_{s+1}, \dots, w_n \in \mathbb{R}^n$ такие, что $v_1, \dots, v_s, w_{s+1}, \dots, w_n$ являются базисом \mathbb{R}^n . Отображение ψ отображает v_1, \dots, v_s в u_1, \dots, u_s . Отправим векторы w_{s+1}, \dots, w_n в 0. Это нам даст отображение уже на всем пространстве \mathbb{R}^n , которое продолжает желаемое отображение и обладает свойством, что v_i идут в u_i . Тут не важно, во что отправить w_j . Как мы видим таких отображений будет много.
- Другой разумный пример использования утверждения. Если пространство V имеет размерность не меньше, чем пространство U , то всегда можно найти сюръективное линейное отображение $\phi: V \rightarrow U$. Действительно, пусть e_1, \dots, e_n – базис V и f_1, \dots, f_m – базис U и $n \geq m$. Тогда существует единственное линейное отображение со свойством $e_1 \mapsto f_1, \dots, e_m \mapsto f_m, e_{m+1} \mapsto 0, \dots, e_n \mapsto 0$. Как легко видеть такое отображение получается сюръективным.

³В следующем разделе как раз показано, что все линейные отображения между \mathbb{R}^n и \mathbb{R}^m задаются умножением на матрицу.

Линейные отображения между \mathbb{R}^n и \mathbb{R}^m

В случае $V = \mathbb{R}^n$ и $U = \mathbb{R}^m$ мы можем полностью описать линейные отображения в терминах матриц. Действительно, пусть (в обозначениях предыдущего утверждения) $\phi(e_i) = u_i = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix}$, где e_i – стандартный базисный вектор \mathbb{R}^n . Тогда

$$\phi \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \phi(x_1 e_1 + \dots + x_n e_n) = x_1 u_1 + \dots + x_n u_n = x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Пристально взглядевшись в то, что мы только что сделали, можно получить следующее.

Утверждение. Отображение $M_{m,n}(\mathbb{R}) \rightarrow \text{Hom}(\mathbb{R}^n, \mathbb{R}^m)$, которое каждой матрице A ставит в соответствие линейное отображение ϕ_A , действующее $\phi_A(x) = Ax$, где $x \in \mathbb{R}^n$, является изоморфизмом векторных пространств, т.е. это правило биективно и $\phi_{A+B} = \phi_A + \phi_B$ и $\phi_{\lambda A} = \lambda \phi_A$.

Заметим, что под действием биекции из упражнения выше операция композиции линейных отображений соответствует операции умножения матриц: если $A \in M_{m,n}(\mathbb{R})$ и $B \in M_{n,k}(\mathbb{R})$, то они соответствуют $\phi_A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ и $\phi_B: \mathbb{R}^k \rightarrow \mathbb{R}^n$. Тогда $\phi_A \phi_B: \mathbb{R}^k \rightarrow \mathbb{R}^m$ совпадает с ϕ_{AB} . Таким образом, как только в пространствах V и U выбраны базисы, нет разницы между изучением линейных отображений и матриц.

Удобный формализм

Матрица линейного отображения Пусть у нас есть линейное отображение $\phi: V \rightarrow U$ и пусть e_1, \dots, e_n – некоторый базис V и f_1, \dots, f_m – некоторый базис U . Тогда каждый вектор $\phi(e_i)$ является линейной комбинацией векторов f_i , т.е. $\phi(e_i) = a_{1i} f_1 + \dots + a_{mi} f_m$. Это можно записать в матричном виде так

$$(\phi(e_1) \quad \dots \quad \phi(e_n)) = (f_1 \quad \dots \quad f_m) \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

или еще короче

$$\phi(e_1 \quad \dots \quad e_n) = (f_1 \quad \dots \quad f_m) A$$

Здесь $\phi(e_1, \dots, e_n)$ имеется в виду покомпонентное умножение вектора из e_i на ϕ слева. Это одна из форм блочного умножения матриц. Матрица A в этом случае называется матрицей линейного отображения ϕ в базисах e_i и f_i .

Действие линейного отображения в координатах Пусть теперь $v \in V$ – некоторый вектор, который раскладывается по базису $v = x_1 e_1 + \dots + x_n e_n = (e_1, \dots, e_n)x$, где $x \in \mathbb{R}^n$. Тогда

$$\phi(v) = \phi(e_1, \dots, e_n)x = (f_1, \dots, f_m)Ax$$

То есть вектор $\phi(v)$ раскладывается по базису f_i с координатами Ax . Значит в координатах, наше линейное отображение задается по правилу $x \mapsto Ax$. На этот факт можно смотреть так. Если есть отображение $\phi: V \rightarrow U$, то после выбора базиса в V оно превращается в \mathbb{R}^n , после выбора базиса в U оно превращается в \mathbb{R}^m , а ϕ должен превратиться в отображение умножения на некоторую матрицу слева. Так вот матрица линейного оператора для ϕ – это в точности та самая матрица, в которую превратился ϕ после выбора базиса.

Смена базиса и линейные отображения

Линейные отображения – это отображения прежде всего и потому они ничего не знают про выбор базиса. С другой стороны, такие отображения задаются разными матрицами в разных базисах. Тут есть пара вещей которые надо понимать: (1) как меняется матрица линейного отображения и (2) смена базиса позволяет упростить вид матрицы.

Начнем с первого вопроса. Тут есть две ситуации: $\phi: V \rightarrow U$ и $\phi: V \rightarrow V$, т.е. случай общего линейного отображения и случай линейного оператора. Главная разница в том, что в первом случае мы можем менять одновременно два базиса и в области определения ϕ и в области куда ϕ бьет. Во втором случае, базисы меняются одновременно.

Утверждение. Пусть e_1, \dots, e_n и e'_1, \dots, e'_n – два базиса V , также f_1, \dots, f_m и f'_1, \dots, f'_m – два базиса U . Пусть

$$(e'_1, \dots, e'_n) = (e_1, \dots, e_n)C \text{ и } (f'_1, \dots, f'_m) = (f_1, \dots, f_m)D$$

где $C \in M_n(\mathbb{R})$ и $D \in M_m(\mathbb{R})$ – матрицы перехода. Если ϕ задается матрицей A в базисах e_i и f_i , то в базисах e'_i и f'_i он задается матрицей $D^{-1}AC$.

Доказательство. Для доказательства воспользуемся замечанием из предыдущего раздела. Нам известно, что $\phi(e_1, \dots, e_n) = (f_1, \dots, f_m)A$, а надо найти матрицу A' такую, что $\phi(e'_1, \dots, e'_n) = (f'_1, \dots, f'_m)A'$. Давайте посчитаем:

$$\phi(e'_1 \dots e'_n) = \phi(e_1 \dots e_n)C = (f_1 \dots f_m)AC = (f'_1 \dots f'_m)D^{-1}AC$$

Значит $A' = D^{-1}AC$, что и требовалось. \square

Следствие. Если $\phi: V \rightarrow V$ в базисе e_1, \dots, e_n записывается матрицей A , то в базисе e'_1, \dots, e'_n заданном $(e'_1, \dots, e'_n) = (e_1, \dots, e_n)C$, ϕ записывается матрицей $C^{-1}AC$.

Смена базиса в координатах

Пусть теперь $V = \mathbb{R}^n$ и $U = \mathbb{R}^m$, также e_1, \dots, e_n обозначает стандартный базис в \mathbb{R}^n и f_1, \dots, f_m – стандартный базис в \mathbb{R}^m . Пусть e'_1, \dots, e'_n – другой базис \mathbb{R}^n . Это вектор столбцы, из которых я могу соорудить матрицу $C \in M_n(\mathbb{R})$, поставив e'_i подряд в качестве столбцов.⁴ Аналогично, если f'_1, \dots, f'_m – другой базис из \mathbb{R}^m я могу составить из них матрицу $D \in M_m(\mathbb{R})$. Обе матрицы C и D невырождены.

Любой вектор $v \in \mathbb{R}^n$ можно записать как

$$v = x_1 e_1 + \dots + x_n e_n = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \text{ в этом случае мы говорим, что задали его в координатах } x_i$$

С другой стороны, мы можем записать v так

$$v = y_1 e'_1 + \dots + y_n e'_n = C \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \text{ в этом случае мы говорим, что задали его в координатах } y_i$$

Аналогично в пространстве \mathbb{R}^m любой вектор u может быть записан в двух системах координат:

$$u = w_1 f_1 + \dots + w_m f_m = \begin{pmatrix} w_1 \\ \vdots \\ w_m \end{pmatrix} \text{ или } u = z_1 f'_1 + \dots + z_m f'_m = D \begin{pmatrix} z_1 \\ \vdots \\ z_m \end{pmatrix}$$

Пусть теперь наше отображение $\phi: \mathbb{R}^n \rightarrow \mathbb{R}^m$ задано матрицей A , то есть вектор в координатах x_i переходит в вектор в координатах w_i по правилу

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} w_1 \\ \vdots \\ w_m \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \text{ или кратко } x \mapsto w = Ax$$

Мы хотим переписать ϕ в координатах y_i и z_i , то есть записать отображение ϕ в виде $y \mapsto z = A'y$. Для этого надо пройти по следующей диаграмме

$$\begin{array}{ccc} x = Cy & \xrightarrow{\quad} & w = Ax = ACy \\ \uparrow & & \downarrow \\ y & \xrightarrow{\quad} & z = D^{-1}w = D^{-1}ACy \end{array}$$

Стартуем с координат y (левый нижний угол). По ним сначала рассчитываем координаты x (вверх по диаграмме). Потом действуем отображением ϕ с помощью матрицы A и получаем вектор $\phi(v)$ в координатах w (вправо по стрелке). Потом пересчитываем координаты w в координаты z (вниз по диаграмме). В результате получаем, что $y \mapsto z = D^{-1}ACy$, т.е. $A' = D^{-1}AC$.

⁴В этом случае мы также имеем $(e'_1, \dots, e'_n) = (e_1, \dots, e_n)C$. Это лишь другой способ описать ту же конструкцию, что и в предыдущем пункте. В столбцах матрицы C стоят координаты векторов e'_i относительно стандартного базиса e_i .

Образ и ядро отображения

Если $\phi: V \rightarrow U$ – линейное отображение (как и выше $V = \mathbb{R}^n$ и $U = \mathbb{R}^m$), то с ним можно связать два подпространства. Первое из них – *ядро* ϕ , а именно: $\ker \phi = \{v \in V \mid \phi(v) = 0\}$.⁵ Второе – *образ* ϕ : $\operatorname{Im} \phi = \phi(V) \subseteq U$, то есть все, что можно получить из V , применив к нему ϕ .

Связь со СЛУ Пусть ϕ задается матрицей $A \in M_{m,n}(\mathbb{R})$, то есть наше отображение имеет вид $\phi: \mathbb{R}^n \rightarrow \mathbb{R}^m$ по правилу $x \mapsto y = Ax$, здесь $x \in \mathbb{R}^n$ и $y \in \mathbb{R}^m$.

- **Ядро** – это пространство решений однородной системы линейных уравнений $\{y \in \mathbb{R}^n \mid Ay = 0\}$.
- **Образ.** Введем следующие обозначения для столбцов матрицы A : $A = (A_1 \mid \dots \mid A_n)$. Тогда по определению в образе ϕ лежат все возможные векторы вида Ax . Давайте распишем это так:

$$\operatorname{Im} \phi = \{Ax \mid x \in \mathbb{R}^n\} = \{x_1 A_1 + \dots + x_n A_n \mid x_i \in \mathbb{R}\} = \langle A_1, \dots, A_n \rangle$$

То есть образ – это линейная оболочка столбцов матрицы A . Если e_1, \dots, e_n – это стандартный базис \mathbb{R}^n , то есть все координаты e_i кроме i -ой равны нулю, а i -я равна единице, тогда i -ый столбец матрицы A – это образ вектора e_i .

- **Прообраз вектор.** Пусть мы зафиксировали вектор $b \in \mathbb{R}^m$ и хотим найти все векторы $x \in \mathbb{R}^n$ такие, что они переходят в b под действием ϕ . Тогда это означает, что нам надо решить уравнение $Ax = b$, то есть решение неоднородной системы означает, что мы ищем прообраз к некоторому вектору.
- **Связь между ОСЛУ и СЛУ.** Пусть x_0 – произвольное решение для $Ax = b$ и $\ker \phi = \{y \in \mathbb{R}^n \mid Ay = 0\}$ – решения однородной системы. Тогда все решения системы $Ax = b$ имеют вид $x_0 + z$, где $z \in \ker \phi$. То есть прообраз любого вектора b является сдвигом ядра отображения ϕ . Однако, обратите внимание, прообраз вектора b может быть пуст, а ядро всегда не пусто, в нем как минимум всегда найдется нулевой вектор. Таким образом ядро отвечает за единственность решения, если оно есть.

Полезно понимать, что для любого b найдется прообраз относительно ϕ , если в системе $Ax = 0$ (или $Ax = b$) количество главных переменных равно количеству строк матрицы A , то есть m . В терминах ранга это означает, что $\operatorname{rk} A = m$.

Свойства ядра и образа

Утверждение. Пусть V и U – векторные пространства и $\varphi: V \rightarrow U$ – линейное отображение. Тогда

1. φ сюръективно тогда и только тогда, когда $\operatorname{Im} \varphi = U$.
2. φ инъективно тогда и только тогда, когда $\ker \varphi = 0$.
3. $\dim \ker \varphi + \dim \operatorname{Im} \varphi = \dim V$.

Доказательство. (1) Это просто переформулировка сюръективности на другом языке.

(2) Так как $\ker \varphi = \varphi^{-1}(0)$ и прообраз всегда содержит 0, то из инъективности вытекает, что $\ker \varphi = 0$. Наоборот, пусть $\varphi(v) = \varphi(v')$, тогда $\varphi(v) - \varphi(v') = 0$. А значит, $\varphi(v - v') = 0$. То есть $v - v'$ лежит в ядре, а значит равен 0, что и требовалось.

(3) Этот пункт я пояснять не буду.

□

Еще полезно понимать, что если в пространствах V и U задать пару подпространств $V' \subseteq V$ и $U' \subseteq U$ такую, что $\dim V' + \dim U' = \dim V$, то найдется (и не одно) линейное отображение $\phi: V \rightarrow U$ такое, что $\ker \phi = V'$, а $\operatorname{Im} \phi = U'$.

Найти линейное отображение с заданными ядром и образом

Дано Пространства $U \subseteq \mathbb{R}^n$ и $W \subseteq \mathbb{R}^m$ такие, что $\dim U + \dim W = n$.

Задача Найти матрицу линейного отображения $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^m$ такого, что $U = \ker \varphi$ и $W = \operatorname{Im} \varphi$.

⁵В англоязычной технической литературе ядро еще называют nullspace, что можно перевести как нулевой пространство.

Алгоритм

1. Задать подпространство W с помощью базиса. Пусть b_1, \dots, b_k – базис W . Определим матрицу $B = (b_1 | \dots | b_k)$.
2. Задать подпространство U системой с линейно независимыми строками $U = \{y \in \mathbb{R}^n \mid Ay = 0\}$.
3. В силу условия $\dim U + \dim W = n$ матрица A будет иметь столько же строк, сколько столбцов в матрице B . В этом случае искомое линейное отображение задается матрицей BA .

Линейные операторы

Напоминание В этом разделе я наконец-то вам начну рассказывать о самых важных объектах в линейной алгебре – линейных операторах. Пусть V – векторное пространство, тогда линейным оператором на V называется линейное отображение $\varphi: V \rightarrow V$, то есть такое отображение, что $\varphi(v_1 + v_2) = \varphi(v_1) + \varphi(v_2)$ и $\varphi(\lambda v) = \lambda \varphi(v)$. Так как линейный оператор – это частный случай линейного отображения, то для него применимо все, о чем мы уже говорили в случае отображений. Про линейный оператор надо думать как про линейную деформацию пространства V .

Примеры

1. $\text{Id}: V \rightarrow V, v \mapsto v$. Тожественный линейный оператор, ничего не деформирует.
2. $0: V \rightarrow V, v \mapsto 0$. Нулевой линейный оператор, который все отправляет в ноль.
3. Растяжения вдоль осей: $D: \mathbb{R}^n \rightarrow \mathbb{R}^n$, заданный по правилу $x \mapsto Dx$, где

$$D = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix}$$

Тогда отображение D растягивает i -ю координату в d_i раз. Если $d_i > 1$, то это растяжение, если $0 < d_i < 1$, то это сжатие, если $-1 < d_i < 0$, то это сжатие и отражение вдоль оси, если $d_i < -1$, то это растяжение и отражение вдоль оси.

4. Поворот на плоскости: $\rho_\alpha: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, где вектор x поворачивается на угол α против часовой стрелки. Такое отображение в матричном виде задается так⁶

$$\rho_\alpha \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

5. Поворот в пространстве вокруг оси OX : $\rho_{1,\alpha}: \mathbb{R}^3 \rightarrow \mathbb{R}^3$, где вектор x поворачивается вокруг оси OX на угол α против часовой стрелки, если смотреть со стороны вектора $e_1 = (1, 0, 0)^t$ на начало координат. В матричном виде эта штука имеет вид

$$\rho_{1,\alpha} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

Теперь давайте объясним некоторую специфику.

⁶Строго говоря я еще не рассказывал про то, что такое движение, но этот и следующий пример можно понять и без общей науки, которая будет чуть позже.

Матрица линейного оператора Пусть в векторном пространстве V задан некоторый базис e_1, \dots, e_n и пусть $\varphi: V \rightarrow V$ – линейный оператор. Так как у оператора пространство из которого он бьет и то в которое он бьет совпадают, то мы фиксируем всего лишь один базис (пространство-то у нас одно). Тогда по определению матрица линейного оператора φ – это такая матрица $A_\varphi \in M_n(\mathbb{R})$, что выполнено $\varphi e = e A_\varphi$, где $e = (e_1, \dots, e_n)$.

Пусть теперь у нас задан другой базис e'_1, \dots, e'_n в пространстве V с матрицей перехода $C \in M_n(\mathbb{R})$, то есть $(e'_1, \dots, e'_n) = (e_1, \dots, e_n)C$. Пусть так же $e' = (e'_1, \dots, e'_n)$. Тогда матрица φ в базисе e' пусть будет A'_φ , то есть $\varphi e' = e' A'_\varphi$. В этом случае связь между матрицами следующая $A'_\varphi = C^{-1} A_\varphi C$. То есть матрица A_φ сопряжена матрице A'_φ .⁷

Замечания

- Отметим, что матрица линейного оператора обязательно квадратная. Таким образом, изучение линейного отображения – это изучение прямоугольной матрицы, а изучение линейного оператора – это всегда изучение только квадратной матрицы.
- Если линейное отображение $\psi: V \rightarrow U$ бьет между двумя разными пространствами одинаковой размерности, то ему тоже соответствует квадратная матрица. Но принципиальная разница с линейным оператором заключается в том, что для линейного отображения мы можем независимо менять базисы в V и U , что соответствует замене $A'_\psi = C^{-1} A_\psi D$, а для линейного оператора, так как пространство одно и то же, базисы меняются одновременно, что соответствует $A'_\varphi = C^{-1} A_\varphi C$.
- Так как линейные операторы – это линейные отображения, то задавать их можно так же как и линейные отображения, например: либо с помощью образа базисных векторов, либо с помощью матрицы в фиксированном базисе.

Например, можно фиксировать векторы $e_1, \dots, e_n \in V$, которые являются базисом и зафиксировать любой набор векторов $u_1, \dots, u_n \in V$, тогда существует единственный линейный оператор $\varphi: V \rightarrow V$ такой, что $\varphi(e_i) = u_i$. Это самый простой и эффективный способ строить новые линейные операторы на пространстве. Но может быть не всегда самый удобный, чтобы потом с такими операторами работать.

Характеристики операторов

След оператора Пусть $A \in M_n(\mathbb{R})$ – квадратная матрица. Тогда *след матрицы* A – это сумма ее диагональных элементов, т.е. $\text{tr } A = a_{11} + \dots + a_{nn} = \sum_{i=1}^n a_{ii}$. Заметим важное свойство следа: $\text{tr}(AB) = \text{tr}(BA)$ (это непосредственная проверка влоб). В частности $\text{tr}(C^{-1}AC) = \text{tr}(A)$ для любых $A, C \in M_n(\mathbb{R})$ с условием, что C обратима.

Пусть теперь $\phi: V \rightarrow V$ – некоторый линейный оператор. Тогда в некотором базисе e_1, \dots, e_n он задается матрицей A . Определим *след линейного оператора* ϕ как след этой матрицы A . Это определение не зависит от базиса. Действительно, в другом базисе оператор ϕ задается матрицей $A' = C^{-1}AC$, тогда $\text{tr}(A') = \text{tr}(C^{-1}AC) = \text{tr}(A)$. След оператора ϕ также обозначается через $\text{tr } \phi$. Важно понимать, что след – это характеристика линейного оператора, а не его матрицы, т.е. эта штука не зависит от матрицы, которой задается оператор. Однако, мы не можем определить эту характеристику не пользуясь базисом. Более того, в принципе невозможно определить след без базиса!

Определитель оператора Пусть $\phi: V \rightarrow V$ – линейный оператор. Тогда в некотором базисе он задается матрицей A . Положим $\det \phi = \det A$. Надо лишь проверить, что это определение не зависит от выбора базиса. Действительно, в другом базисе ϕ задается $C^{-1}AC$, а значит $\det(C^{-1}AC) = \det A$. Величина $\det \phi$ называется *определителем линейного оператора*. Как и в случае следа, определитель линейного оператора не зависит от базиса, но его нельзя определить не пользуясь базисом.

Многочлены от операторов Пусть $\phi: V \rightarrow V$ – линейный оператор. Тогда определен линейный оператор $\lambda\phi$ для любого $\lambda \in \mathbb{R}$. Уточним на всякий случай, что $(\lambda\phi)(v) = \lambda\phi(v)$ по определению. Более того, определены все натуральные степени оператора ϕ , как композиция, т.е. $\phi^n(v) = \phi(\dots\phi(v)\dots)$ – где композиция берется

⁷Напомню, что квадратные матрицы B и D называются сопряженными, если найдется обратимая матрица C такая, что $D = C^{-1}BC$.

n раз, например, $\phi^3(v) = \phi(\phi(\phi(v)))$.⁸ Кроме того, линейные операторы можно складывать. Напомню, что $(\phi + \psi)(v) = \phi(v) + \psi(v)$ по определению.

Из выше сказанного следует, что можно составлять выражения вида $\sqrt{\pi}\phi - 2\phi^4$ и их результат будет линейный оператор на V . Более того, можно определить ϕ^0 как тождественный оператор Id , т.е. $\text{Id}(v) = v$.⁹ Тогда можно писать выражения вроде $2/3 + \phi^2$, где имеется в виду $2/3 \text{Id} + \phi^2$.

Таким образом для любого многочлена $p(t) = a_0 + a_1 t + \dots + a_n t^n$, где $a_i \in \mathbb{R}$, и любого линейного оператора $\phi: V \rightarrow V$ определен линейный оператор $p(\phi): V \rightarrow V$. Когда мы перейдем к базисам, оператор ϕ будет соответствовать матрице A . В этом случае $p(\phi)$ соответствует матрице $p(A)$.¹⁰

Основной бонус от подстановки матриц и линейных операторов в многочлены состоит вот в чем. Пусть многочлен $p(t)$ раскладывается в произведение $(t - \lambda_1) \dots (t - \lambda_n)$. Тогда оператор $p(\phi)$ раскладывается в композицию операторов $(\phi - \lambda_1 \text{Id}) \dots (\phi - \lambda_n \text{Id})$. Скоро (очень очень скоро) будет видно зачем все это нужно.

Характеристический многочлен Пусть $\phi: V \rightarrow V$ – некоторый линейный оператор. Опять же для удобства, можно считать, что после выбора базиса $V = \mathbb{R}^n$ и ϕ соответствует некоторой матрице $A \in M_n(\mathbb{R})$. Тогда выражение $\det(\lambda \text{Id} - \phi) = \det(\lambda E - A)$ является многочленом от λ степени n . Действительно, указанный определитель с точностью до $(-1)^n$ будет равен

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

С усилием вспоминая явную формулу для определителя через перестановки, понимаем, что получается многочлен от λ . Еще чуть внимательнее присмотревшись к нему, можно заметить, что

$$\det(A - \lambda E) = \det(A) + \dots + (-1)^{n-1} \text{tr } A \lambda^{n-1} + (-1)^n \lambda^n$$

или

$$\det(\lambda E - A) = \lambda^n - \text{tr}(A) \lambda^{n-1} + \dots + (-1)^n \det(A)$$

Данный многочлен обозначим через $\chi_\phi(\lambda)$ или $\chi_A(\lambda)$ и будем называть *характеристическим многочленом* оператора ϕ или соответствующей матрицы A (в зависимости от того, о чем идет речь).¹¹

Утверждение (Теорема Гамильтона-Кэли). Пусть $\phi: V \rightarrow V$ – линейный оператор и $\chi(t)$ – его характеристический многочлен. Тогда $\chi(\phi) = 0$.¹²

Минимальный многочлен Если $\phi: V \rightarrow V$ – линейный оператор на n мерном пространстве, то как мы видели выше, он зануляется своим характеристическим многочленом, т.е. существуют многочлены $p(t)$ такие, что $p(\phi) = 0$. На самом деле, факт существования таких многочленов доказывается проще, чем теорема Гамильтона-Кэли. Наш ϕ соответствует матрице A . Но тогда $E, A, A^2, \dots, A^{n^2}$ не могут быть линейно независимы. А значит $a_0 + a_1 A + \dots + a_{n^2} A^{n^2} = 0$, то есть A зануляется многочленом $p(t) = a_0 + a_1 t + \dots + a_{n^2} t^{n^2}$. Суть теоремы Гамильтона-Кэли в том, чтобы понизить степень многочлена с n^2 до n .

Мы скажем, что ненулевой многочлен $p(t)$ является минимальным для ϕ , если $p(\phi) = 0$ и степень многочлена p является минимально возможной. Минимальный многочлен делит любой многочлен зануляющий ϕ , так как остаток тоже должен занулять ϕ и степени меньше. Потому, если у минимальный многочлен нормировать так, чтобы его старший коэффициент был единицей, то он определен однозначно.

Утверждение. Пусть $\phi: V \rightarrow V$ – линейный оператор. Тогда

⁸На самом деле умные люди вообще не пишут скобки, ибо они только загромождают обозначения. Ведь куда приятнее смотреть на $\phi\phi\phi v$ вместо $\phi(\phi(\phi(v)))$. Но еще приятнее смотреть на $\phi^3 v$.

⁹В любом базисе тождественный оператор соответствует единичной матрице.

¹⁰Предлагаю вам самим восстановить детали того, что такое подставить матрицу в многочлен по аналогии с операторами.

¹¹Надо отметить, что часто характеристическим многочленом называют $\det(A - \lambda E)$, потому что так удобнее считать руками на бумажке. Наш многочлен от этого отличается на $(-1)^n$. Для многих вопросов это не принципиально.

¹²Есть два доказательства этого факта: (1) кустарное, методами линейной алгебры и (2) концептуальное методами коммутативной алгебры. Первое доказательство использует теорему о Жордановой нормальной форме (по сути классификацию всех линейных операторов) и очень геморройное. Второе доказательство в одну строчку – формулы Крамера для модулей над коммутативными кольцами. Его беда в том, что надо объяснить все дурацкие слова в доказательстве. Это не сложно, но требует кучу времени и усилий, чтобы их осознать.

1. Минимальный многочлен $p(t)$ для ϕ со старшим коэффициентом определен однозначно.
2. Многочлен $p(t)$ делит любой многочлен аннулирующий ϕ .
3. Существует такое число m , что характеристический многочлен $\chi_\phi(t)$ делит $p^m(t)$.

Собственные значения и вектора оператора

Пусть $\phi: V \rightarrow V$ – линейный оператор на пространстве V . Будем говорить, что вектор $v \in V$ является *собственным*, если $\phi v = \lambda v$.¹³ То есть на собственный вектор оператор ϕ действует растяжением. Если $\phi v = \lambda v$ для $v \neq 0$, число λ называется *собственным значением* оператора ϕ . При фиксированном $\lambda \in \mathbb{R}$ множество всех собственных векторов с собственным значением λ , т.е. $\{v \in V \mid \phi v = \lambda v\}$, будем обозначать через V_λ . Все V_λ обязательно будут подпространствами.¹⁴

Если мы выберем базис в пространстве V , то оно превратится в \mathbb{R}^n . Наш оператор ϕ будет задаваться матрицей $A \in M_n(\mathbb{R})$. В этом случае, собственный вектор задается уравнением $Ax = \lambda x$, где $x \in \mathbb{R}^n$. Беда в том, что мы пока заранее не знаем, какие λ нам подходят. Чтобы это выяснить нужно переписать уравнение так: $(A - \lambda E)x = 0$. Оно имеет решение тогда и только тогда, когда $A - \lambda E$ – вырожденная матрица. Это, в свою очередь, происходит тогда и только тогда, когда $\det(A - \lambda E) = 0$. Напомним, что характеристический многочлен ϕ (он же характеристический для A) это $\chi(\lambda) = \det(\lambda E - A)$, т.е. получаем следующее.

Утверждение. Пусть $\phi: V \rightarrow V$ – некоторый линейный оператор с матрицей $A \in M_n(\mathbb{R})$ в некотором базисе. Тогда

1. Все собственные значения оператора ϕ – это в точности корни характеристического многочлена $\chi(\lambda) = \det(\lambda E - A)$.
2. Если λ – НЕ корень характеристического многочлена, то $V_\lambda = 0$.
3. Если λ – корень характеристического многочлена, то V_λ – ненулевое подпространство V . Кроме того, $\dim V_\lambda$ не превосходит кратности корня λ у характеристического многочлена.¹⁵

Привет от комплексных чисел

Заметим, что собственные значения являются корнями многочлена. С действительными числами есть беда: многочлены могут вообще не иметь корней. Например: пусть $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, тогда $\chi_A(\lambda) = \lambda^2 + 1$. У этого многочлена нет вещественных корней. Потому нет собственных значений, а значит и ненулевых собственных векторов. На самом деле, для любого многочлена можно подобрать матрицу так, что он будет ее характеристическим многочленом. Так что это не случайное явление.

Так как собственные значения и вектора хотелось бы иметь, то нам придется в этом вопросе переходить к комплексным числам. и вместо пространства \mathbb{R}^n рассматривать \mathbb{C}^n . Тогда над комплексными числами каждый многочлен имеет ровно столько корней (с учетом кратности), какова его степень. Это первое место в линейной алгебре, где появляется разница в том, какие коэффициенты использовать.

Кто такие комплексные числа

По простому, мы хотим построить множество «чисел», которые бы содержали вещественные числа и на них были определены все нужные операции: сложения, вычитания, умножения и деления на любое ненулевое число. Есть несколько конструкций, я рассмотрю две.

Классическая конструкция Рассмотрим множество картинок вида $a + bi$, где $a, b \in \mathbb{R}$, а i – просто символ. Как множество $\mathbb{C} = \{a + bi \mid a, b \in \mathbb{R}\}$. Теперь на \mathbb{C} определим следующие операции:

1. Сложение: $(a + bi) + (c + di) = (a + c) + (b + d)i$

¹³Нулевой вектор является собственным для любого $\lambda \in \mathbb{R}$. Обратите внимание, что в некоторых учебниках собственные вектора обязательно считаются ненулевыми, но это идейно не правильно.

¹⁴Заметим, что $V_\lambda = \ker(\phi - \lambda \text{Id})$.

¹⁵Для многочлена $p(t)$ число λ является корнем тогда и только тогда, когда $p(t) = (t - \lambda)q(t)$. Если λ корень для $q(t)$, мы можем еще раз вынести множитель $t - \lambda$ и так далее. В итоге, можно записать $p(t) = (t - \lambda)^k h(t)$, где $h(\lambda) \neq 0$. Такое число k называется кратностью корня λ .

2. Вычитание: $(a + bi) - (c + di) = (a - c) + (b - d)i$
3. Умножение: $(a + bi)(c + di) = (ac - bd) + (ad + bc)i$.
4. Сопряжение: $\overline{a + bi} = a - bi$.

В этом случае нулем будет число вида $0 + 0i$, единицей $1 + 0i$. Если $z = a + bi$, то число $z\bar{z} = a^2 + b^2$ является неотрицательным вещественным числом. Модуль комплексного числа z – это $|z| = \sqrt{z\bar{z}}$. Обратный к числу z имеет вид $\frac{\bar{z}}{|z|^2}$.

Числа вида $a + 0i$ можно отождествить с вещественными числами $a \in \mathbb{R}$. Таким образом $\mathbb{R} \subseteq \mathbb{C}$. Более того, операции определены так, что это вложение с ними согласовано. Обратим внимание на новое число $i = 0 + 1i$. По определению $i^2 = -1$. На самом деле верно следующее.

Утверждение. Для любого многочлена $p(t) = a_0 + a_1t + \dots + a_{n-1}t^{n-1} + t^n$, где $a_i \in \mathbb{C}$ существует ровно n комплексных корней с учетом кратности.

Матричная конструкция Рассмотрим матрицы вида

$$T = \left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \in M_2(\mathbb{R}) \right\}$$

Заметим, что если сложить или перемножить любые две матрицы из T , то получим матрицу из T . Более того, все матрицы из T кроме нулевой обратимы. Множество T можно отождествить с \mathbb{C} , построенным выше, следующим образом: $a + bi \mapsto \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$. То есть T и \mathbb{C} – это одно и то же. Обратим внимание, что на этом языке сопряжение – это транспонирование, а определитель равен квадрату модуля комплексного числа.¹⁶

Собственный базис

Утверждение. Пусть $\phi: V \rightarrow V$ – линейный оператор и пусть $\lambda_1, \dots, \lambda_k$ – его разные собственные значения (тут не важно из \mathbb{R} или \mathbb{C}) и $v_1, \dots, v_k \in V$ – соответствующие им ненулевые собственные вектора. Тогда v_1, \dots, v_k линейно независимы.

Доказательство. Предположим противное, что $a_1v_1 + \dots + a_kv_k = 0$. Мы можем считать, что все a_i не равны нулю. Это можно записать так

$$a_1v_1 + \dots + a_kv_k = \begin{pmatrix} a_1v_1 & \dots & a_kv_k \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = 0$$

Применим к этой линейной комбинации ϕ , получим новую линейную комбинацию

$$a_1\lambda_1v_1 + \dots + a_k\lambda_kv_k = \begin{pmatrix} a_1v_1 & \dots & a_kv_k \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_k \end{pmatrix} = 0$$

Продолжим применять ϕ суммарно $k - 1$ раз. В результате имеем

$$\begin{pmatrix} a_1v_1 & \dots & a_kv_k \end{pmatrix} \begin{pmatrix} 1 & \lambda_1 & \dots & \lambda_1^{k-1} \\ 1 & \lambda_2 & \dots & \lambda_2^{k-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_k & \dots & \lambda_k^{k-1} \end{pmatrix} = \begin{pmatrix} 0 & \dots & 0 \end{pmatrix}$$

Но определитель матрицы выше есть определитель вандермонда. Значит, матрица обратима и на нее можно поделить. Значит, все вектора $a_iv_i = 0$. Так как по предположению $v_i \neq 0$ это означает, что $a_i = 0$, противоречие. \square

¹⁶Первая конструкция обычно рассказывается в школе и потому более привычная. Вторая хороша тем, что нам не надо проверять, что операции ведут себя хорошо, все следует из знаний о матрицах. Плюс это дает некий мостик в правильную линейную алгебру над вещественными числами.

Утверждение. Пусть $\phi: V \rightarrow V$ – оператор на n -мерном пространстве (не важно комплексном или вещественном), при этом его характеристический многочлен имеет n различных корней $\lambda_1, \dots, \lambda_n$. Тогда соответствующие ненулевые собственные вектора v_1, \dots, v_n образуют базис V и в этом базисе матрица ϕ диагональная с числами λ_i на диагонали.

Доказательство. Действительно, для каждого такого λ_i обязательно найдется ненулевой собственный вектор. Из предыдущего утверждения все такие собственные вектора линейно независимы, а значит образуют базис. По определению в этом базисе $\phi v_i = \lambda_i v_i$, т.е. $\phi(v_1, \dots, v_n) = (v_1, \dots, v_n) \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix}$. \square

На это утверждение можно смотреть так: если есть квадратная матрица $A \in M_n(\mathbb{R})$ (или $A \in M_n(\mathbb{C})$) такая, что $\det(A - \lambda E)$ имеет n различных корней, то существует такая невырожденная матрица $C \in M_n(\mathbb{R})$ (соответственно из $M_n(\mathbb{C})$), что $C^{-1}AC$ является диагональной и на диагонали стоят корни многочлена $\det(A - \lambda E)$. Комплексный случай хорош лишь тем, что корни обязательно существуют у многочлена, надо лишь чтобы они были различными. В вещественном случае существование корней не гарантировано.

Важно понимать, что если матрицу взялась «из жизни» или из «непрерывных случайных данных», то с вероятностью один, характеристический многочлен такой матрицы будет иметь n различных комплексных корней. То есть над комплексными числами любая случайная матрица с вероятностью один превращается в диагональную в некотором базисе.

Поиск собственных значений и векторов

Следующий алгоритм годится как для комплексных так и для вещественных матриц. Разница лишь в том, что в вещественном случае у нас вообще говоря будет меньше собственных значений. Для определенности алгоритм рассказывается для комплексных матриц.

Дано Матрица $A \in M_n(\mathbb{C})$.

Задача Найти все собственные значения λ_i для A и для каждого λ_i найти базис пространства V_{λ_i} .

Алгоритм

1. Посчитать характеристический многочлен $(-1)^n \chi_A(\lambda) = \det(A - \lambda E)$.
2. Найти корни многочлена $\chi(\lambda)$. Корни $\{\lambda_1, \dots, \lambda_k\}$ будут собственными значениями A .
3. Для каждого λ_i найти ФСР системы $(A - \lambda_i E)x = 0$. Тогда ФСР будет базисом V_{λ_i} .

Отметим, что общее количество собственных векторов для всех собственных значений λ_i не превосходит n – размерности матрицы, так как $\dim V_{\lambda_i}$ не превосходит кратности корня λ_i , а сумма кратностей всех корней в точности равна степени многочлена $\chi(\lambda)$, которая есть n – размер матрицы A .

Если количество собственных векторов оказалось равно n , то матрица A приводится в диагональный вид. Пусть v_{i1}, \dots, v_{in_i} – собственные вектора с собственным значением λ_i , при этом n_i будет кратность собственного значения λ_i . Пусть C – матрица составленная из векторов v_{ij} . Пусть D – диагональная матрица с диагональю $(\lambda_1, \dots, \lambda_1, \lambda_2, \dots, \lambda_2, \dots, \lambda_k, \dots, \lambda_k)$, где каждое λ_i повторяется n_i раз. Тогда $C^{-1}AC = D$.

Проверка на диагонализуемость

Геометрическая проверка на диагонализуемость

Дано Матрица $A \in M_n(\mathbb{R})$, задающая линейный оператор $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^n$.

Задача Выяснить существует ли базис, в котором φ задается диагональной матрицей и если задается, то какой именно. На матричном языке: существует ли невырожденная матрица $C \in M_n(\mathbb{R})$ такая, что $C^{-1}AC$ является диагональной и найти эту диагональную матрицу.

Алгоритм

1. Найдем характеристический многочлен $\chi(t)$ для φ , он же для A по формуле $(-1)^n \chi(t) = \det(A - tE)$.
2. Проверим, раскладывается ли $\chi(t)$ на линейные множители, то есть представляется ли он в виде $\chi(t) = (t - \lambda_1)^{d_1} \dots (t - \lambda_k)^{d_k}$. Если не представляется, то φ (или что то же самое A) не диагонализуется.
3. Если $\chi(t) = (t - \lambda_1)^{d_1} \dots (t - \lambda_k)^{d_k}$. Найдем для каждого λ_i базис V_{λ_i} как ФСР системы $(A - \lambda_i E)x = 0$. Если для хотя бы одного i количество элементов в ФСР меньше соответствующей кратности корня d_i , то φ не диагонализуется.
4. Если для каждого i мы получили, что размер ФСР совпадает с кратностью корня, то есть $\dim V_{\lambda_i} = d_i$. То φ диагонализуется и диагональная матрица $C^{-1}AC$ на диагонали содержит числа λ_i в количестве d_i штук.

Заметим, что если задача изначально дана для комплексной матрицы $A \in M_n(\mathbb{C})$, которая задает в этом случае оператор $\varphi: \mathbb{C}^n \rightarrow \mathbb{C}^n$, то первый шаг алгоритма выполнен автоматически, а именно, над комплексными числами любой многочлен разлагается на линейные множители. Потому над комплексными числами вопрос о диагонализуемости – это лишь проверка всех равенств $\dim V_{\lambda_i} = d_i$.

Обратите внимание, что если $A \in M_n(\mathbb{R})$ – матрица диагонализуемого оператора, то $\text{rk } A$ равен количеству ненулевых диагональных элементов в ее диагональном виде. То есть количеству ненулевых собственных значений с учетом их кратности в характеристическом многочлене.

Алгебраическая проверка на диагонализуемость

Дано Матрица $A \in M_n(\mathbb{R})$, задающая линейный оператор $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^n$.

Задача Выяснить существует ли базис, в котором φ задается диагональной матрицей. На матричном языке: существует ли невырожденная матрица $C \in M_n(\mathbb{R})$ такая, что $C^{-1}AC$ является диагональной и найти эту диагональную матрицу.

Алгоритм

1. Найдем характеристический многочлен $\chi(t)$ для φ , он же для A по формуле $(-1)^n \chi(t) = \det(A - tE)$.
2. Проверим, раскладывается ли $\chi(t)$ на линейные множители, то есть представляется ли он в виде $\chi(t) = (t - \lambda_1)^{d_1} \dots (t - \lambda_k)^{d_k}$. Если не представляется, то φ (или что то же самое A) не диагонализуется.
3. Избавимся от кратных корней в хар многочлене по правилу $h = \frac{\chi}{(\chi', \chi)}$, где в знаменателе стоит НОД хар многочлена и производной.
4. Проверим зануляет ли h оператор, то есть выполнено ли $h(A) = 0$. Если выполнено, то оператор диагонализуем, если нет, то не диагонализуем.

Признак диагонализуемости Есть очень удобный признак диагонализуемости, который часто помогает. Если у вас есть линейный оператор $\varphi: V \rightarrow V$, то обычно в начале фиксируется какой-то базис и после этого V превращается в \mathbb{R}^n , а оператор φ в оператор умножения на матрицу $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^n$ задан $x \mapsto Ax$, где $A \in M_n(\mathbb{R})$. В итоге нам надо понять, можно ли выбрать другой базис для φ , чтобы матрицу A заменить на диагональную.

Утверждение. Пусть $A \in M_n(\mathbb{R})$ и $g \in \mathbb{R}[x]$ – зануляющий многочлен для A , то есть $g(A) = 0$. Если он раскладывается на линейные множители $g(x) = (x - \lambda_1) \dots (x - \lambda_k)$ и все λ_i различны, то существует обратимая матрица $C \in M_n(\mathbb{R})$ такая, что $C^{-1}AC$ будет диагональной и на диагонали будут числа из множества $\{\lambda_1, \dots, \lambda_k\}$ (но может быть не все из них и часть может повторяться).

Обратите внимание, что здесь очень важно, чтобы корни λ_i были различными! Если это условие не выполнено, то утверждение не верно. Например, возьмем матрицу $J_n(\lambda)$. У нее минимальный многочлен будет $(x - \lambda)^n$, но про нее можно доказать, что нельзя сопряжением ее сделать диагональной.

Кроме того, в силу теоремы Гамильтона-Кэли, вам достаточно проверить, что характеристический или минимальный многочлены раскладываются на линейные множители и не имеют кратных корней.

Квадратные корни из единицы Это соображение полезно, если вы хотите решать различные матричные уравнения. Например, пусть мы хотим найти все возможные $A \in M_n(\mathbb{R})$ такие, что $A^2 = E$, то есть хотим найти все квадратные корни из единицы в матрицах. Тогда это означает, что $g(x) = x^2 - 1$ зануляет матрицу A . При этом $g(x) = (x - 1)(x + 1)$. Это значит, что найдется обратимая матрица C такая, что $C^{-1}AC$ будет диагональной с числами 1 и -1 на диагонали. Кроме того, если у диагональной матрицы 1 и -1 не идут подряд, то мы ее можем сопрячь некоторой матрицей так, что 1 и -1 пойдут подряд. То есть мы показали, что если A является решением уравнения $A^2 = E$, то найдется такая невырожденная матрица C , что

$$C^{-1}AC = \begin{pmatrix} E & 0 \\ 0 & -E \end{pmatrix}, \text{ следовательно } A = C \begin{pmatrix} E & 0 \\ 0 & -E \end{pmatrix} C^{-1}$$

Здесь подразумевается, что блоки с единичной и минус единичной матрицей могут быть пустыми (то есть только одни единицы или минус единицы допустимы).

С другой стороны. Легко видеть, что матрицы полученного вида являются решениями данного уравнения

$$A^2 = C \begin{pmatrix} E & 0 \\ 0 & -E \end{pmatrix} C^{-1} C \begin{pmatrix} E & 0 \\ 0 & -E \end{pmatrix} C^{-1} = C \begin{pmatrix} E & 0 \\ 0 & -E \end{pmatrix}^2 C^{-1} = C E C^{-1} = E$$

Жорданова нормальная форма (ЖНФ)

Самый главный вопрос о линейных операторах: на сколько хорошим можно выбрать базис, чтобы максимально упростить матрицу оператора в этом базисе? В случае «общего положения» как в предыдущем параграфе мы можем диагонализировать матрицу. И это самый популярны в приложениях случай. Но есть и плохие матрицы, которые нельзя диагонализировать. В общем случае ответ будет чуть-чуть сложнее.

Для начала несколько определений. *Жорданова клетка* $J_n(\lambda)$ размера n с числом $\lambda \in \mathbb{C}$ – это матрица вида¹⁷

$$J_n(\lambda) = \begin{pmatrix} \lambda & 1 & \dots & 0 \\ 0 & \lambda & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda \end{pmatrix} \in M_n(\mathbb{C})$$

Будем говорить, что матрица $A \in M_n(\mathbb{C})$ имеет *Жорданову нормальную форму*, если она имеет следующий блочный вид

$$A = \begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A_k \end{pmatrix} \text{ где } A_i \in M_{n_i}(\mathbb{R}) \text{ имеет вид } \begin{pmatrix} J_{r_1}(\lambda_i) & 0 & \dots & 0 \\ 0 & J_{r_2}(\lambda_i) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_{r_s}(\lambda_i) \end{pmatrix}$$

Следующая теорема – это полная классификация линейных операторов на векторном пространстве.

Утверждение (Теорема о Жордановой нормальной форме). Пусть V – комплексное векторное пространство и $\phi: V \rightarrow V$ – произвольный линейный оператор. Пусть $\lambda_1, \dots, \lambda_k$ – корни его характеристического многочлена с кратностями n_1, \dots, n_k . Тогда, существует базис V такой, что матрица ϕ имеет следующий вид:

$$A = \begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A_k \end{pmatrix}$$

где $A_i \in M_{n_i}(\mathbb{C})$ (размер равен кратности собственного значения). А каждая A_i имеет вид

$$\begin{pmatrix} J_{r_{i1}}(\lambda_i) & 0 & \dots & 0 \\ 0 & J_{r_{i2}}(\lambda_i) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_{r_{is_i}}(\lambda_i) \end{pmatrix}$$

¹⁷Можно определить Жорданову клетку и для действительных чисел и для рациональных и вообще каких угодно, но я буду тут обсуждать только комплексный случай.

где λ_i – соответствующее собственное значение. При этом числа r_{i1}, \dots, r_{is_i} определены однозначно и могут отличаться только порядком.¹⁸

Замечания Пусть $A \in M_n(\mathbb{C})$, тогда она обязательно приводится в ЖНФ. Если λ – это собственное значение для A , тогда оно является корнем характеристического многочлена и пусть его кратность будет d .¹⁹

1. Число d – это суммарный размер жордановых клеток в ЖНФ для A с числом λ на диагонали.
2. Число $\dim V_\lambda$ – это количество жордановых клеток в ЖНФ для A с числом λ на диагонали.

¹⁸Существуют алгоритмы нахождения базиса, в котором матрица имеет Жорданову нормальную форму, но мы их изучать не будем.

¹⁹На самом деле можно дать полный список числовых инвариантов, которые характеризуют ЖНФ для A , но это выходит за рамки нашего обсуждения.

Семинар 4 дополнительно

Рекуррентные соотношения

Вместо того, чтобы тут развивать супер общую теорию, я все проиллюстрирую на конкретном примере – последовательность Фибоначчи. Что это такое? Это последовательность чисел $a_n \in \mathbb{R}$, где $n \in \mathbb{Z}_{\geq 0}$ удовлетворяющая следующим условиям

$$\begin{cases} a_n = a_{n-1} + a_{n-2} \\ a_0 = 1 \\ a_1 = 1 \end{cases}$$

Если мы захотим по этим правилам посчитать n -ый член последовательности, то нам понадобится $O(n)$ операций, то есть последовательно посчитать все n членов последовательности, чтобы добраться до a_n . Однако, можно несколько схитрить и сделать это быстрее за $O(\log n)$ с помощью матричных операций. Для этого введем вектор

$$x_n = \begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix} \text{ при этом мы знаем, что } x_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Заметим, что x_n выражается через x_{n-1} следующим образом

$$\begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_{n-1} \\ a_{n-2} \end{pmatrix} \text{ то есть } x_n = Ax_{n-1} \text{ где } A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

Но тогда $x_n = Ax_{n-1} = A^2x_{n-2} = \dots = A^{n-1}x_1$. А значит для нахождения x_n нам лишь надо возвести матрицу A в степень n . Для этого подходит хорошо известный алгоритм быстрого возведения в степень для чисел, который слово в слово работает для матриц.¹ Давайте заведем две квадратные матрицы $X, Y \in M_2(\mathbb{R})$ и число $m \in \mathbb{Z}_{\geq 0}$. В самом начале положим $X = E$, $Y = A$ и $m = n$. Будем поддерживать следующий инвариант $XY^m = A^n$. Алгоритм остановим тогда, когда $m = 0$, тогда X будет нашим ответом. Шаги алгоритма следующие. Если m четно, то $XY^{2m'} = X(Y^2)^{m'}$ поделим m на 2, а Y возведем в квадрат. Если m нечетно, то $XY^{2m'+1} = (XY)Y^{2m'}$ уменьшим m' на единицу и умножим X на Y .

На самом деле, можно проверить, что $A = CDC^{-1}$, где

$$D = \begin{pmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{pmatrix}, \quad C = \begin{pmatrix} \frac{2}{\sqrt{5}-1} & -\frac{2}{\sqrt{5}+1} \\ 1 & 1 \end{pmatrix} \text{ и } C^{-1} = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{5-\sqrt{5}}{10} \\ -\frac{1}{\sqrt{5}} & \frac{5+\sqrt{5}}{10} \end{pmatrix}$$

Для простоты обозначений, будем считать $D = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$, тогда

$$A^n = (CDC^{-1})^n = CD^nC^{-1} = C \begin{pmatrix} \lambda^n & 0 \\ 0 & \mu^n \end{pmatrix} C^{-1}$$

То есть, зная как получить из матрицы A супер крутое разложение, мы можем еще сильнее упростить вычисления степени матрицы и свести его к вычислению степеней чисел и произведения трех матриц.

¹То есть от степени у нас будет логарифмическая сложность, но константа будет включать в себя сложность перемножения матриц, которая зависит от размера матрицы, равному длине рекурренты в нашем случае.

Семинар 4

Общая информация:

- Квадратные матрицы $A, B \in M_n(\mathbb{R})$ называются сопряженными, если найдется невырожденная матрица $C \in M_n(\mathbb{R})$ такая, что $B = C^{-1}AC$.

Задачи:

1. Какие из следующих матриц сопряжены? Если они сопряжены, то укажите с помощью какой матрицы:
 - $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ и $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, • $\begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}$ и $\begin{pmatrix} 3 & -1 \\ 2 & 1 \end{pmatrix}$, • $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ и $\begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}$.

2. В пространстве \mathbb{R}^3 заданы следующие векторы

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, v_3 = \begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix}, u_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, u_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, u_3 = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}$$

Найдите матрицу A линейного оператора $\phi: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ по правилу $x \mapsto Ax$, такого, что $Av_i = u_i$ для всех $1 \leq i \leq 3$.

3. Пусть $\phi: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ – линейное отображение, заданное в стандартном базисе матрицей $A = \begin{pmatrix} 1 & 2 & 0 \\ -1 & 0 & 2 \end{pmatrix}$. Пусть

$$f_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, f_2 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, f_3 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \text{ вектора в } \mathbb{R}^3, \quad g_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, g_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ вектора в } \mathbb{R}^2$$

Найти матрицу отображения ϕ в базисах f_1, f_2, f_3 и g_1, g_2 .

4. Найти собственные значения и собственные векторы линейных операторов, заданных в некотором базисе матрицами: (a) $\begin{pmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{pmatrix}$, (b) $\begin{pmatrix} 4 & -5 & 2 \\ 5 & -7 & 3 \\ 6 & -9 & 4 \end{pmatrix}$. Можно ли эти матрицы диагонализировать в каком-нибудь базисе?

5. Пусть $\mathbb{R}[x]_{\leq n}$ – пространство многочленов степени не выше n . Рассмотрим на нем линейное отображение по правилу $f \mapsto (x+1)f'(x) - 2f(x) + \frac{1}{x} \int_0^x f(t) dt$. Найдите матрицу этого линейного отображения в базисе $1, x, x^2, \dots, x^n$.

Семинар 5

Билинейные формы

Пусть V – векторное пространство, можно думать для простоты, что $V = \mathbb{R}^n$. Тогда *билинейная форма* на V – это отображение $\beta: V \times V \rightarrow \mathbb{R}$ такое,¹ что

1. $\beta(v_1 + v_2, u) = \beta(v_1, u) + \beta(v_2, u)$ для всех $v_1, v_2, u \in V$.
2. $\beta(\lambda v, u) = \lambda \beta(v, u)$ для всех $v, u \in V$ и $\lambda \in \mathbb{R}$.
3. $\beta(v, u_1 + u_2) = \beta(v, u_1) + \beta(v, u_2)$ для всех $v, u_1, u_2 \in V$.
4. $\beta(v, \lambda u) = \lambda \beta(v, u)$ для всех $v, u \in V$ и $\lambda \in \mathbb{R}$.

Думать про эту процедуру надо так: у нас даны два вектора v и u из V , мы их «перемножаем» и получаем число $\beta(v, u) \in \mathbb{R}$. Самый важный пример билинейной формы – стандартное скалярное произведение: пусть даны два вектора $x, y \in \mathbb{R}^n$, зададим тогда $\beta(x, y) = x^t y = \sum_{i=1}^n x_i y_i$.

Основной план взаимодействия с билинейными формами такой. Среди всех билинейных форм мы выделим «хорошие» и назовем их скалярными произведениями. Эти товарищи будут иметь хороший геометрический смысл, с помощью которого мы определим *движения* в векторных пространствах. Но нашей конечной целью будет изучение самих движений с помощью скалярных произведений.

Как задавать билинейные формы

Пусть V – векторное пространство с базисом e_1, \dots, e_n и β – билинейная форма на V . Тогда определим числа $b_{ij} = \beta(e_i, e_j)$ – произведения базисных векторов, и составим из них матрицу $B \in M_n(\mathbb{R})$. Тогда для любых векторов $v = x_1 e_1 + \dots + x_n e_n$ и $u = y_1 e_1 + \dots + y_n e_n$ имеем

$$\beta(v, u) = \sum_{ij} x_i y_j \beta(e_i, e_j) = (x_1 \dots x_n) B \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

В частности, если $V = \mathbb{R}^n$ и $x, y \in \mathbb{R}^n$, а e_i – стандартный базис. То получаем $\beta(x, y) = x^t B y$.

Таким образом и линейные операторы и билинейные формы задаются матрицами. Основная разница между ними – как эта самая матрица меняется при замене базиса. Для операторов ответы мы знаем, для билинейных форм мы сейчас займемся данным вопросом.

Смена базиса

Пусть в векторном пространстве V заданы два базиса e_1, \dots, e_n и f_1, \dots, f_n с матрицей перехода $C \in M_n(\mathbb{R})$, т.е. $(f_1, \dots, f_n) = (e_1, \dots, e_n)C$. Пусть нам даны два вектора v и u в V . Тогда их можно разложить по базисным векторам следующим образом

$$\begin{aligned} v = x_1 e_1 + \dots + x_n e_n &= (e_1 \dots e_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, & u = y_1 e_1 + \dots + y_n e_n &= (e_1 \dots e_n) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \\ v = x'_1 f_1 + \dots + x'_n f_n &= (f_1 \dots f_n) \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}, & u = y'_1 f_1 + \dots + y'_n f_n &= (f_1 \dots f_n) \begin{pmatrix} y'_1 \\ \vdots \\ y'_n \end{pmatrix} \end{aligned}$$

Благодаря матрице перехода C , мы знаем, что

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = C \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} = C x' \text{ и } y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = C \begin{pmatrix} y'_1 \\ \vdots \\ y'_n \end{pmatrix} = C y'$$

¹На самом деле можно рассматривать отображения $\beta: V \times U \rightarrow \mathbb{R}$, то есть можно перемножать вектора из разных пространств, но мы этого делать не будем.

Тогда в базисе e_i форма записывается в виде $\beta(v, u) = x^t B y$, а в базисе f_i в виде $\beta(v, u) = (x')^t B' y'$. Но это одно и то же число посчитанное в разных базисах. Значит

$$(x')^t B' y' = x^t B y = (C x')^t B C y' = (x')^t C^T B C y'$$

для всех $x', y' \in \mathbb{R}^n$. Значит $B' = C^t B C$.²

Симметричность и кососимметричность

Форма $\beta: V \times V \rightarrow \mathbb{R}$ называется *симметричной*, если $\beta(v, u) = \beta(u, v)$ для всех $v, u \in V$. Она называется *кососимметричной*, если $\beta(v, u) = -\beta(u, v)$.

Если в координатах $\beta(x, y) = x^t B y$, то $\beta(y, x) = y^t B x$. Так как выражение $y^t B x$ является числом, то оно не меняется при транспонировании, то есть $y^t B x = (y^t B x)^t = x^t B^t y$. Значит симметричность означает $x^t B y = x^t B^t y$ для всех $x, y \in \mathbb{R}^n$. А это равносильно тому, что $B = B^t$. Такая матрица B называется *симметричной*. Аналогично, форма кососимметрична, тогда и только тогда, когда $B^t = -B$. В этом случае матрица B называется *кососимметричной*.

Характеристики билинейных форм

Как и выше $\beta: V \times V \rightarrow \mathbb{R}$ – билинейная форма. И пусть в некотором базисе она задана в виде $\beta(x, y) = x^t B y$ для некоторой матрицы $B \in M_n(\mathbb{R})$. Посмотрим какие характеристики матрицы B не зависят от выбора базиса.

1. Ранг матрицы B не меняется при замене $B \mapsto C^t B C$, где $C \in M_n(\mathbb{R})$ – невырожденная матрица.
2. Знак определителя B не меняется при замене $B \mapsto C^t B C$, где $C \in M_n(\mathbb{R})$ – невырожденная матрица. Но сам определитель меняется на $\det(C)^2$. Потому можно лишь говорить о ситуации определитель меньше нуля, больше нуля или равен нулю.
3. Обратим внимание, что невырожденность матрицы B не меняется при замене $B \mapsto C^t B C$, где $C \in M_n(\mathbb{R})$ – невырожденная матрица.
4. След матрицы B вообще говоря может стать каким угодно при замене $B \mapsto C^t B C$. Потому он не несет никакой информации.
5. Симметричность и кососимметричность матрицы B не зависят от замены $B \mapsto C^t B C$.

Ядра и ортогональные дополнения

Как и выше $\beta: V \times V \rightarrow \mathbb{R}$ – билинейная форма. Множества ${}^{\perp}V = \{v \in V \mid \beta(v, V) = 0\}$ и $V^{\perp} = \{v \in V \mid \beta(V, v) = 0\}$ называются *левым* и *правым ядрами* формы β . Эти подмножества являются подпространствами в V . Если в координатах форма задана $\beta(x, y) = x^t B y$, то ${}^{\perp}\mathbb{R}^n = \{x \in \mathbb{R}^n \mid x^t B = 0\}$ и $(\mathbb{R}^n)^{\perp} = \{y \in \mathbb{R}^n \mid B y = 0\}$. В частности отсюда видно, что ядра имеют одинаковую размерность равную $n - \text{rk } B$. Если форма симметричная или кососимметричная, то нет разницы между правыми и левыми ядрами. Ядра – это неинтересная часть пространства, которая «ортогональна» всему относительно этой формы.

Более обще, пусть $U \subseteq V$ – подпространство в V . Тогда его *левым ортогональным дополнением* является подпространство ${}^{\perp}U = \{v \in V \mid \beta(v, U) = 0\}$. Аналогично, *правое ортогональное дополнение* это $U^{\perp} = \{v \in V \mid \beta(U, v) = 0\}$. Если в координатах форма задана $\beta(x, y) = x^t B y$ и $U = \langle u_1, \dots, u_k \rangle$. Пусть D – матрица составленная из столбцов u_i . Тогда ${}^{\perp}U = \{x \in \mathbb{R}^n \mid x^t B D = 0\}$ и $U^{\perp} = \{y \in \mathbb{R}^n \mid D^t B y = 0\}$. Обычно ортогональные дополнения и ядра интересны в случае симметрических или кососимметрических форм, так как в этом случае левые и правые ортогональные дополнения равны между собой.

Двойственность для подпространств

Напомню, что для двух подпространств $U, W \subseteq V$ определены их сумма $U + W = \{u + w \mid u \in U, w \in W\}$ и пересечение $U \cap W$. Сумма – это наименьшее подпространство, которое содержит U и W одновременно

²Напомним, что матрица A линейного оператора $\phi: V \rightarrow V$ меняется по правилу $A' = C^{-1} A C$.

(объединение не является подпространством вообще говоря). А пересечение – это наибольшее подпространство, которое лежит и в том и в том. Про них надо думать, как бы как про НОК и НОД подпространств, соответственно.

В общем случае нет хорошей связи между подпространством и его ортогональным (левым или правым) дополнением. Например, если билинейная форма нулевая, то есть $\beta: V \times V \rightarrow \mathbb{R}$ все переводит в ноль (это соответствует нулевой матрице в любом базисе), то любые два подпространства ортогональны, и ортогональным дополнением к чему угодно будет все пространство V . Однако для хороших билинейных форм можно доказать теорему о двойственности на подпространствах.

Утверждение. Пусть $\beta: V \times V \rightarrow \mathbb{R}$ – невырожденная билинейная форма (то есть ее матрица не вырождена). Тогда:

1. Для любого подпространства $W \subseteq V$ выполнено

$$\dim W^\perp + \dim W = \dim V$$

2. Для любого подпространства $W \subseteq V$ выполнено ${}^\perp(W^\perp) = W$.

3. Для любых подпространств $W \subseteq E \subseteq V$ верно, что $W^\perp \supseteq E^\perp$. Причем $W = E$ тогда и только тогда, когда $W^\perp = E^\perp$.

4. Для любых подпространств $W, E \subseteq V$ выполнено равенство

$$(W + E)^\perp = W^\perp \cap E^\perp$$

5. Для любых подпространств $W, E \subseteq V$ выполнено равенство

$$(W \cap E)^\perp = W^\perp + E^\perp$$

Аналогично выполнены все свойства для подпространств $W \subseteq U$ и их левых ортогональных дополнений ${}^\perp W$.

Таким образом, мы как бы переворачиваем подпространства вверх ногами под действием операции взятия ортогонального дополнения. Операции взятия левого и правого ортогонального дополнения становятся взаимнообратными. При этом эти операции большие подпространства переводят в маленькие, обращают вложения и меняют НОК и НОД подпространств местами.

Симметричные формы

Утверждение. Пусть $\beta: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ – симметрическая билинейная форма заданная $\beta(x, y) = x^t B y$. Тогда существует такой базис, что матрица B диагональная и на диагонали стоят либо 1, либо -1 , либо 0, т.е. блочно имеет вид $B' = \begin{pmatrix} E & & \\ & -E & \\ & & 0 \end{pmatrix}$. При этом количество единиц и минус единиц на диагонали не зависит от базиса.

На это утверждение еще можно смотреть так. Для любой симметрической матрицы $B \in M_n(\mathbb{R})$ можно найти такую невырожденную матрицу $C \in M_n(\mathbb{R})$, что матрица $C^t B C$ имеет описанный диагональный вид.

Суммарно количество единиц и минус единиц дает ранг матрицы B , то есть ранг билинейной формы. Тот факт, что количество единиц и минус единиц является инвариантом формы надо понимать так: у нас ранг как бы складывается из положительной и отрицательной части и размеры этих частей определены однозначно.

Количество единиц $\#1$ в таком виде называется положительным индексом инерции формы, количество минус единиц $\#-1$ – отрицательным индексом, а количество нулей $\#0$ – нулевым индексом. Вместе набор чисел $(\#1, \#-1, \#0)$ называется сигнатурой формы.

Определение сигнатуры формы

Для определения сигнатуры билинейной формы можно воспользоваться разными методами. Самый простой – Симметричный Гаусс. Мы диагонализуем форму и считаем количество положительных, отрицательных и нулевых элементов на диагонали. Этот метод работает всегда.

Симметрический Гаусс Теперь, когда мы знаем, что симметрические билинейные формы диагонализуются в каком-то базисе, хорошо было бы иметь какой-нибудь (ну хотя бы плохонький) алгоритм, приводящий форму к диагональному виду, если она задана в каком-то случайном базисе. Пусть, скажем, нам задана билинейная форма $\beta: F^n \times F^n \rightarrow F$ по правилу $(x, y) \mapsto x^t B y$, где $B \in M_n(F)$ – некоторая симметричная матрица. Тогда в новом базисе матрица будет иметь вид $C^t B C$, где $C \in M_n(F)$ – некоторая невырожденная матрица.³ Любая невырожденная матрица C раскладывается в произведение элементарных матриц. С другой стороны, если C – матрица элементарного преобразования, то $B \mapsto C^t B C$ – это выполнение одного и того же преобразования и над строками и над столбцами (не важно в каком порядке, так как произведение матриц ассоциативно). То есть у нас есть следующий запас операций:

- Прибавляем i -ю строку умноженную на λ к j -ой строке, потом прибавляем i -ый столбец умноженный на λ к j -ому столбцу.
- Меняем местами i и j строки, после чего меняем местами i и j столбцы.
- Умножаем на ненулевое λ i -ю строку, потом умножаем на λ i -ый столбец.

Таким образом предыдущая теорема гласит, что выполняя подобные симметричные элементарные преобразования над симметрической матрицей, мы обязательно приведем ее к диагональному виду. Если при этом надо восстановить матрицу C , то рассматриваем $(B|E)$ и делаем симметричные элементарные преобразования над ней в том смысле, что преобразования над строками выполняются над всей матрицей, а преобразования над столбцами только над частью, где лежит B . Тогда матрица приведется к виду $(B'|C^t)$.

Метод Якоби Также для определения сигнатуры формы используется метод Якоби. Этот метод работает почти всегда и я поясню, что это значит и что делать, когда он не работает. Но прежде всего я хочу обратить внимание, что у него есть ограничения на входные данные. Матрица B обязательно должна быть невырожденной. Это в частности означает, что метод работает только для форм у которых в сигнатуре только единицы и минус единицы и совсем нет нулей.

Пусть $B \in M_n(\mathbb{R})$ – симметричная невырожденная матрица и $\beta(x, y) = x^t B y$. Выделим в матрице B верхние левые блоки:

$$B = \begin{pmatrix} \begin{pmatrix} b_{11} & & \\ & \ddots & \\ & & B_k \end{pmatrix} & & \\ & & \ddots \end{pmatrix}$$

То есть B_k – подматрица состоящая из первых k строк и столбцов. Теперь определим числа $\Delta_k = \det(B_k)$, которые называются угловыми минорами. Если так получилось, что все числа Δ_k НЕ равны нулю⁴, то мы строим последовательность

$$\Delta_1, \frac{\Delta_2}{\Delta_1}, \frac{\Delta_3}{\Delta_2}, \dots, \frac{\Delta_n}{\Delta_{n-1}}$$

Тогда положительный индекс инерции для B равен количеству положительных чисел в этой последовательности, а отрицательный индекс инерции равен количеству отрицательных чисел в этой последовательности.

Организация вычислений в методе Якоби Считать в матрице B сразу все угловые миноры – накладная задача. Оказывается это можно сделать за один проход гаусса. Давайте применять к матрице B элементарные преобразования I типа, когда прибавляем с коэффициентом более высокие строки к более низким. Такие

³На самом деле C – матрица перехода из старого в новый базис.

⁴Матрицу B с таким условием можно разложить в виде $B = LU$, где L – нижнетреугольная матрица, U – верхнетреугольная матрица с 1 на диагонали. Это называется LU разложением. Для этого надо применить Гаусса к B вычитая из более верхних строчек более низкие. Тогда мы приведем B к верхнетреугольному виду. Отсюда можно вытащить LU разложение стандартным рассуждением.

преобразования не меняют значения угловых миноров. Значит, если мы смогли привести матрицу к верхнетреугольному виду

$$\begin{pmatrix} d_1 & * & \dots & * & * \\ & d_2 & \dots & * & * \\ & & \ddots & \vdots & \vdots \\ & & & d_{n-1} & * \\ & & & & d_n \end{pmatrix}$$

То тогда определитель Δ_k матрицы B будет совпадать с угловым определителем k на k полученной матрицы, то есть $\Delta_k = d_1 \dots d_k$. А значит сигнатура билинейной формы совпадает со знаками последовательности диагональных элементов d_1, \dots, d_n . Что нам может помешать привести матрицу к верхнетреугольному виду? Единственная проблема – мы встретили ноль на диагонали и под ним есть не нули, то есть встретилась ситуация вида

$$\begin{pmatrix} * & * & * & * & * \\ & * & * & * & * \\ & & 0 & * & * \\ & & * & * & * \\ & & * & * & * \end{pmatrix}$$

Но это означает, что угловой минор (на примере 3 на 3) вырожден (то есть $\Delta_3 = 0$), а значит не выполняются условия для применения метода Якоби. Таким образом мы либо находим, что какой-то Δ_i равен нулю, либо приводим матрицу B к верхнетреугольному виду и определяем сигнатуру матрицы B .

Что делать, если встретились нули По-хорошему надо пользоваться симметричным Гауссом. Но если все таки хочется воспользоваться именно методом Якоби, то надо чуть-чуть пошевелить матрицу B правильным образом. Мы можем сгенерировать случайную матрицу C . Она с вероятностью один будет невырожденной. Потом надо рассмотреть матрицу $B' = C^t B C$ и применить метод Якоби к матрице B' вместо B . Сделаю одно замечание по организации вычислений. В этом методе надо генерировать случайную матрицу C и НЕ проверять ее на невырожденность. Вместо этого, надо сразу применить метод Якоби к матрице B' . Если все Δ_k оказались не нулевыми, то нам повезло и метод и так сработал (матрица C в этом случае автоматически окажется невырожденной). А если не повезло, то нам все равно надо будет генерировать новую матрицу C и не важно какой она была.

Продвинутый метод определения сигнатуры Пусть $B \in M_n(\mathbb{R})$ – симметричная матрица и $\beta(x, y) = x^t B y$. Тогда найдем спектр матрицы B с кратностями. В случае симметрической матрицы окажется, что спектр будет обязательно вещественным. Тогда количество положительных чисел в спектре с учетом кратности равно положительному индексу инерции, количество отрицательных чисел в спектре с кратностью равно отрицательному индексу инерции, а количество нулей – нулевому индексу. Надо понимать, что сам спектр не является корректно определенной величиной для билинейной формы, он может измениться кардинальной при смене базиса, но знаки собственных значений, оказывается, не изменятся.

Последнее свойство дает способ оценить количество собственных значений у матрицы на отрезке. Например у матрицы $A - \lambda E$ спектр состоит из $\lambda_i - \lambda$, где λ_i – собственные значения A . Тогда количество положительных собственных значений у $A - \lambda E$ можно определить симметричным Гауссом, и оно совпадает с количеством собственных значений A строго больших λ . Подбирая λ и считая количество положительных или отрицательных корней, мы можем определить количество собственных значений на заданном отрезке.

LU разложение Пусть нам дана матрица B и мы хотим представить ее в виде $B = LU$, где L – нижнетреугольная матрица с единицами на диагонали, а U – верхнетреугольная матрица. Давайте для простоты будем считать, что матрица U , а значит и матрица B не вырожденная. Пусть B_k – угловая подматрица B размера k на k . То есть подматрица стоящая на первых k строках и столбцах. Аналогично определим L_k и U_k . В силу того, что матрица L нижнетреугольная, а U верхнетреугольная можно заметить, что $B_k = L_k U_k$.⁵

Предположим, B невырожденная матрица и для нее существует разложение LU , где L нижнетреугольная с единицами на диагонали, а U верхнетреугольная. Тогда в частности U тоже не вырождена и все ее диагональные коэффициенты не равны нулю. В частности матрицы L_k и U_k невырождены. А значит и матрица

⁵Например примените блочные формулы.

$B_k = L_k U_k$ тоже не вырождена. То есть существование LU разложения влечет, что все угловые подматрицы B_k невырождены.

Наоборот, если все матрицы B_k невырождены, то мы можем найти LU разложение следующим образом. Давайте будем применять к матрице B элементарные преобразования I типа, где мы будем прибавлять более высокие строки к более низким с каким-то коэффициентом. Тогда применение таких преобразований равносильно умножению на матрицу C , которая будет нижнетреугольной с единицами на диагонали. Если в результате таких действий, мы сможем привести матрицу B к верхнетреугольному виду U , то это означает, что $CB = U$. А значит $B = C^{-1}U$. Но тогда матрица $L = C^{-1}$ тоже нижнетреугольная с единицами на диагонали и разложение получено. Потому нам лишь надо показать, что B приводится к верхнетреугольному виду указанными преобразованиями и можно пользоваться, что все матрицы B_k не вырождены. Обратите внимание что прибавление более верхней строки к более нижней не меняет определители матриц B_k . Так что после указанных преобразований угловые подматрицы будут оставаться невырожденными. Теперь будем идти по матрице B сверху вниз и занулять более нижние элементы с помощью диагональных элементов. Мы хотим реализовать следующую последовательность действий:

$$\begin{pmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{pmatrix} \mapsto \begin{pmatrix} * & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \end{pmatrix} \mapsto \begin{pmatrix} * & * & * & * \\ 0 & * & * & * \\ 0 & 0 & * & * \\ 0 & 0 & * & * \end{pmatrix} \mapsto \begin{pmatrix} * & * & * & * \\ 0 & * & * & * \\ 0 & 0 & * & * \\ 0 & 0 & 0 & * \end{pmatrix}$$

Когда у нас может возникнуть проблема? Если мы встретим ситуацию такую

$$\begin{pmatrix} * & * & * & * & * \\ & * & * & * & * \\ & & 0 & * & * \\ & & * & * & * \\ & & * & * & * \end{pmatrix}$$

Когда диагональный элемент ноль, а под ним есть ненулевые элементы. Но в этом случае угловая подматрица (в данном случае подматрица 3 на 3) будет верхнетреугольной и вырожденной, что противоречит тому, что все угловые подматрицы остаются невырожденными. Значит такой ситуации встретиться не может и мы всегда можем привести матрицу B к верхнетреугольному виду.

Квадратичные формы

Если нам дана какая-то билинейная форма $\beta: V \times V \rightarrow \mathbb{R}$, то отображение $Q: V \rightarrow \mathbb{R}$ вида $Q(x) = \beta(x, x)$ называется квадратичной формой. Если векторное пространство $V = \mathbb{R}^n$, то билинейная форма превращается в $\beta(x, y) = x^t B y$, а соответствующая квадратичная форма в $Q(x) = x^t B x$. Если расписать явно последнее выражение, то мы получим

$$Q(x) = \beta(x, x) = x^t B x = \sum_{ij} b_{ij} x_i x_j = \sum_i b_{ii} x_i^2 + \sum_{i < j} (b_{ij} + b_{ji}) x_i x_j$$

Обратите внимание, что в отличие от билинейной формы, квадратичная форма не однозначно задается матрицей B . Действительно,

$$Q(x_1, x_2) = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 2x_1 x_2$$

За счет этого эффекта, при переходе к квадратичным формам от билинейных, мы теряем часть информации. Однако, квадратичная форма однозначно задается симметрической матрицей B , то есть матрицей B с условием $B^t = B$. В примере выше – это последний случай.

Для полноты картины добавлю, что в случае симметричной матрицы B или что то же самое симметричной билинейной формы β , мы можем вернуться от квадратичной формы к билинейной с помощью так называемой поляризационной формулы, а именно

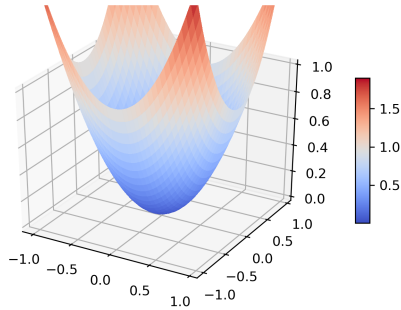
$$\beta(x, y) = \frac{Q(x+y) - Q(x) - Q(y)}{2}$$

Идейно это означает, что изучать симметричные билинейные формы – это то же самое, что изучать квадратичные формы. Но у квадратичных форм есть красивый геометрический смысл. Его мы и обсудим далее.

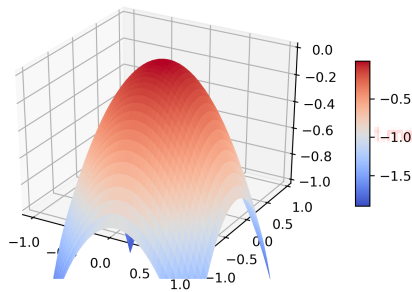
Графики квадратичных форм

Пусть $V = \mathbb{R}^2$. Тогда квадратичная форма $Q(x, y)$ задает функцию от двух переменных, а именно $z = Q(x, y)$. Давайте нарисуем ее графики в некоторых частных случаях.

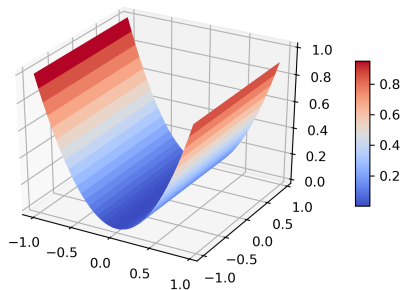
1. $z = x^2 + y^2$. Начало координат – точка минимума. Матричная запись $z = Q(x, y) = \begin{pmatrix} x & y \end{pmatrix}^t \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$



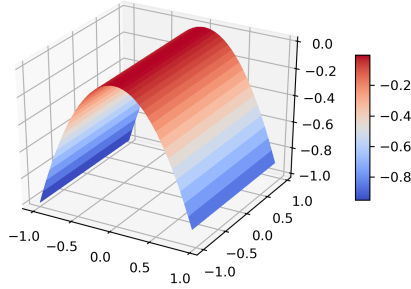
2. $z = -x^2 - y^2$. Начало координат – точка максимума. Матричная запись $z = Q(x, y) = \begin{pmatrix} x & y \end{pmatrix}^t \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$



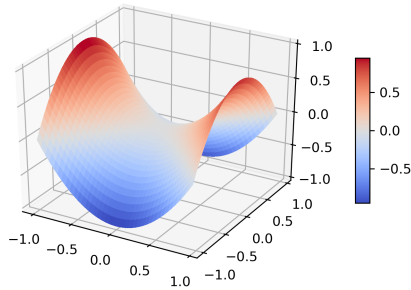
3. $z = x^2$. Минимум достигается на прямой $x = 0$. Матричная запись $z = Q(x, y) = \begin{pmatrix} x & y \end{pmatrix}^t \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$



4. $z = -x^2$. Максимум достигается на прямой $x = 0$. Матричная запись $z = Q(x, y) = \begin{pmatrix} x & y \end{pmatrix}^t \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$



5. $z = x^2 - y^2$. Начало координат – седловая точка. Матричная запись $z = Q(x, y) = \begin{pmatrix} x \\ y \end{pmatrix}^t \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$



Обратите внимание, что поведение графика зависит от знаков чисел на диагонали матрицы два на два. В общем случае поведение графика зависит от сигнатуры формы.

Классификация билинейных и квадратичных форм Ниже я все определения отразил в единой таблице. В ней подразумевается, что билинейная форма задана на пространстве размерности n .

Термин	Обозначения	Условие	Индексы
Положительная	$\beta > 0$ или $Q > 0$	$\forall x \neq 0 \Rightarrow Q(x) > 0$	$\#1 = n$
Отрицательная	$\beta < 0$ или $Q < 0$	$\forall x \neq 0 \Rightarrow Q(x) < 0$	$\# - 1 = n$
Неотрицательная	$\beta \geq 0$ или $Q \geq 0$	$\forall x \Rightarrow Q(x) \geq 0$	$\# - 1 = 0$
Неположительная	$\beta \leq 0$ или $Q \leq 0$	$\forall x \Rightarrow Q(x) \leq 0$	$\#1 = 0$
Неопределенная		$\exists x, y \Rightarrow Q(x) > 0$ и $Q(y) < 0$	$\#1 > 0$ и $\# - 1 > 0$

Скалярные произведения

Билинейная форма $\beta: V \times V \rightarrow \mathbb{R}$ называется *скалярным произведением*, если она

1. симметрична $\beta(v, u) = \beta(u, v)$.
2. *положительно определена*, т.е. для любого ненулевого вектора $v \in V$ имеем $\beta(v, v) > 0$.

В этом случае пишут (v, u) вместо $\beta(v, u)$. Самый важный пример – стандартное скалярное произведение: $(x, y) = x^t y$, где $x, y \in \mathbb{R}^n$. Векторное пространство, в котором зафиксировано какое-либо скалярное произведение называется *Евклидовым пространством*.

По определению скалярного произведения у него в сигнатуре присутствуют только единицы, а минус единиц и нулей нет. В частности это означает, что матрица скалярного произведения всегда невырождена. Кроме того это еще означает, что для любого скалярного произведения существует такой базис, что в нем матрица B становится единичной матрицей. По-другому, на этот факт можно смотреть так: какие-бы два евклидовых пространства одинаковой размерности вы ни взяли бы, они оказываются одинаковыми (формально изоморфными). И наоборот, если вы хотите задать скалярное произведение, то достаточно взять

какой-то базис и объявить его ортонормированным. Тогда существует единственное скалярное произведение с таким свойством, действительно, потому что это означает, что в этом базисе вы выбрали в качестве матрицы билинейной формы матрицу $B = E$.

Экзотические скалярные произведения

Ради интереса вот два примера любопытных скалярных произведений в неожиданных ситуациях.

1. Пусть в качестве векторного пространства у нас будет пространство матриц $V = M_{m,n}(\mathbb{R})$ не обязательно квадратных. Пусть $A, B \in M_{m,n}(\mathbb{R})$, зададим скалярное произведение следующим образом $(A, B) = \text{tr}(A^t B)$.⁶ Можно руками проверить, что $(A, B) = (B, A)$ и что $(A, A) = \sum_{ij} a_{ij}^2 > 0$, если $A \neq 0$. А потому эта штука удовлетворяет свойствам скалярного произведения. Значит можно в пространстве матриц мерить длины матриц и углы между ними. Длина матрицы в этом случае будет $|A|_F = \sqrt{\text{tr}(A^t A)}$ и называется «нормой Фробениуса» матрицы A .
2. Пусть теперь в качестве векторного пространства у нас будет множество всех непрерывных функций на отрезке $[0, 1]$, то есть $V = C[0, 1] = \{f: [0, 1] \rightarrow \mathbb{R} \mid f \text{ непрерывна}\}$. Тогда для двух функций $f, g \in C[0, 1]$ определим скалярное произведение следующим образом $(f, g) = \int_0^1 f(x)g(x) dx$. Получается, что теперь можно мерить длины функций и углы между функциями. Например, длина функции f будет $\sqrt{\int_0^1 f(x)^2 dx}$.

Углы и расстояния

Пусть V – евклидово пространство. Тогда *длина* вектора v это $|v| = \sqrt{(v, v)}$. Если $v, u \in V$ – два вектора, то определим *угол* $\alpha_{v,u}$ между этими векторами из равенства $\cos \alpha_{v,u} = \frac{(v,u)}{|v||u|}$.

Два вектора v и u называются *ортгоналными*, если $(v, u) = 0$, т.е. угол между векторами 90° . Базис e_1, \dots, e_n называется *ортгоналным*, если любая пара векторов из базиса ортгонална, т.е. $(e_i, e_j) = 0$ при $i \neq j$. Базис называется *ортонормированным*, если он ортгонален и все вектора имеют длину 1, т.е. $(e_i, e_j) = 0$ при $i \neq j$ и $(e_i, e_i) = 1$. По определению матрицы билинейной формы $b_{ij} = (e_i, e_j)$, а значит в ортонормированном базисе скалярное произведение имеет вид $(x, y) = x^t y$.

Для любого множества векторов $S \subseteq \mathbb{R}^n$ мы можем определить ортгоналное дополнение

$$S^\perp = \{v \in V \mid (v, s) = 0 \text{ для любого } s \in S\}$$

То есть это все возможные вектора, которые ортгогалны всем векторам из S . Обратите внимание, что S может быть любым множеством, например, может состоять из одного вектора, но при этом S^\perp всегда будет векторным подпространством в \mathbb{R}^n .

Расстоянием между двумя векторами v и u пространства \mathbb{R}^n называется $\rho(v, u) = |v - u|$. Если мы хотим найти расстояние между двумя подмножествами, например, $X, Y \subseteq V$, то по определению расстояние между ними – это наименьшее расстояние между всеми парами точек из них, то есть

$$\rho(X, Y) = \inf_{\substack{x \in X \\ y \in Y}} \rho(x, y)$$

В частности можно говорить о расстоянии от вектора до подпространства. Напомню, что если $U \subseteq \mathbb{R}^n$ – некоторое подпространство и U^\perp – его ортгоналное дополнение, то любой вектор однозначно раскладывается в сумму вектора из U и вектора из U^\perp , то есть любой $v \in V$ имеет вид $v = u + w$, где $u \in U$ и $w \in U^\perp$. Тогда вектор u называется ортгогалной проекцией v на U , а вектор w называется ортгогалной составляющей v относительно U .

Утверждение. Пусть $V = \mathbb{R}^n$ – евклидово пространство, $U \subseteq V$ – подпространство, $v \in V$ – некоторый вектор. Тогда расстояние от v до U равно длине ортгогалной составляющей v относительно U .

Ортгогаланизация Грама-Шмидта

Дано Множество векторов $v_1, \dots, v_k \in \mathbb{R}^n$.

⁶Обратите внимание, что матрица $A^t B$ будет квадратной размера n на n , а потому для нее корректно определено понятие следа.

Задача Найти множество u_1, \dots, u_s такое, что u_i попарно ортогональны и $\langle v_1, \dots, v_k \rangle = \langle u_1, \dots, u_s \rangle$.

Алгоритм

1. Берем первый ненулевой вектор среди v_i . Пусть это будет v_1 . Тогда полагаем $u_1 = v_1$.
2. Рассмотрим $v_2 - \frac{(v_2, u_1)}{(u_1, u_1)} u_1$. Если этот вектор не ноль, то обозначим его за u_2 . Если ноль, то выкинем v_2 и перенумеруем вектора так, что v_3 теперь будет вектором v_2 . Повторяем этот шаг до тех пор, пока не найдем u_2 или пока не закончатся вектора v_i .
3. Рассмотрим $v_3 - \frac{(v_3, u_1)}{(u_1, u_1)} u_1 - \frac{(v_3, u_2)}{(u_2, u_2)} u_2$. Если он не ноль, то обозначим его за u_3 . Иначе как и в предыдущем пункте переходим к следующему вектору и повторяем этот шаг.
4. Для поиска u_i надо рассмотреть вектор $v_i - \frac{(v_i, u_1)}{(u_1, u_1)} u_1 - \dots - \frac{(v_i, u_{i-1})}{(u_{i-1}, u_{i-1})} u_{i-1}$. Аналогично предыдущему пункту, если этот вектор не ноль, то это u_i . Если ноль, то рассматриваем следующий v_{i+1} вместо него и повторяем этот шаг.

Пример Пусть у нас заданы векторы

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 3 \\ 3 \\ 1 \\ 1 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 2 \\ 2 \\ 1 \\ 1 \end{pmatrix} \text{ и } v_4 = \begin{pmatrix} 2 \\ 0 \\ 1 \\ -1 \end{pmatrix} \in \mathbb{R}^4$$

Первый вектор не ноль, значит $u_1 = v_1$. Теперь рассмотрим

$$v_2 - \frac{(v_2, u_1)}{(u_1, u_1)} u_1 = \begin{pmatrix} 3 \\ 3 \\ 1 \\ 1 \end{pmatrix} - \frac{3+3+1+1}{1+1+1+1} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}$$

t.me/postypashki_old/1229

t.me/postypashki_old/1229

t.me/postypashki_old/1229

Значит $u_2 = v_2 - 2u_1$. Теперь рассмотрим

$$v_3 - \frac{(v_3, u_1)}{(u_1, u_1)} u_1 - \frac{(v_3, u_2)}{(u_2, u_2)} u_2 = \begin{pmatrix} 2 \\ 2 \\ 1 \\ 1 \end{pmatrix} - \frac{2+2+1+1}{1+1+1+1} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} - \frac{2+2-1-1}{1+1+1+1} \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix} = 0$$

Значит забываем про v_3 и переходим к следующему вектору.

$$v_4 - \frac{(v_4, u_1)}{(u_1, u_1)} u_1 - \frac{(v_4, u_2)}{(u_2, u_2)} u_2 = \begin{pmatrix} 2 \\ 0 \\ 1 \\ -1 \end{pmatrix} - \frac{2+1-1}{1+1+1+1} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} - \frac{2-1+1}{1+1+1+1} \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$$

Таким образом ответ

$$u_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad u_2 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix} \text{ и } u_3 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$$

Ортогональные матрицы

Утверждение. Для матрицы $A \in M_n(\mathbb{R})$ следующие условия эквивалентны:

1. $A^t A = E$.
2. $AA^t = E$.
3. $A^t = A^{-1}$.

Матрица обладающая одним из этих эквивалентных условий называется *ортогональной*. Таким образом в ортонормированном базисе движение задается ортогональной матрицей.

QR-разложение Пусть у нас дан набор векторов $v_1, \dots, v_k \in \mathbb{R}^n$ и задано стандартное скалярное произведение $(x, y) = x^t y$. Давайте для начала для простоты будем считать, что все векторы линейно независимы. Проведем для них процесс ортогонализации Грама-Шмидта и получим векторы v'_1, \dots, v'_k . Если в формулах для Грама-Шмидта перенести все старые векторы влево, а новые вправо, то получится следующий набор равенств:

$$(v_1 \quad \dots \quad v_k) = (v'_1 \quad \dots \quad v'_k) \begin{pmatrix} 1 & \frac{(v_2, v'_1)}{(v'_1, v'_1)} & \frac{(v_3, v'_1)}{(v'_1, v'_1)} & \dots & \frac{(v_k, v'_1)}{(v'_1, v'_1)} \\ & 1 & \frac{(v_3, v'_2)}{(v'_2, v'_2)} & \dots & \frac{(v_k, v'_2)}{(v'_2, v'_2)} \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}$$

При этом векторы v'_1, \dots, v'_k являются ортогональными. В силу предположенной линейной независимости исходных векторов новые векторы тоже будут линейно независимыми, а в частности не нулевыми. Если теперь поделить каждый из них на его длину, то получится

$$(v_1 \quad \dots \quad v_k) = \begin{pmatrix} \frac{v'_1}{|v'_1|} & \dots & \frac{v'_k}{|v'_k|} \end{pmatrix} \begin{pmatrix} |v'_1| & \frac{(v_2, v'_1)}{|v'_1|} & \frac{(v_3, v'_1)}{|v'_1|} & \dots & \frac{(v_k, v'_1)}{|v'_1|} \\ & |v'_2| & \frac{(v_3, v'_2)}{|v'_2|} & \dots & \frac{(v_k, v'_2)}{|v'_2|} \\ & & |v'_3| & & \\ & & & \ddots & \\ & & & & |v'_k| \end{pmatrix}$$

Пусть теперь $A = (v_1 | \dots | v_k) \in M_{n \times k}(\mathbb{R})$ – матрица составленная из исходных векторов, а $Q' = \left(\frac{v'_1}{|v'_1|} | \dots | \frac{v'_k}{|v'_k|} \right)$, а $R' \in M_k(\mathbb{R})$ – матрица из коэффициентов справа в верхнем равенстве. Тогда верхнее равенство записывается так $A = Q' R'$. Мы можем добавить столбцы в матрицу Q' в конце так, чтобы получилась ортогональная матрица Q , тогда $A = Q(R' | 0)$. Таким образом A представлена в виде произведения ортогональной матрицы Q и верхнетреугольной матрицы $R = (R' | 0)$. Такое разложение называется QR разложением.

Если же у исходной матрицы A столбцы оказались линейно зависимыми, то в результате нашего алгоритма, Q' может содержать нулевые столбцы, но все они будут ортогональны. В этом случае поступают так. Для каждого нулевого столбца Q' зануляют соответствующую строку в матрице R' . После этого все равно, что стоит в соответствующем столбце Q' . В этом случае в место нулевых столбцов вставляют недостающие векторы до ортонормированного базиса. А последний шаг перехода от Q' и R' к Q и R абсолютно такой же.

Матрица Грама

Давайте поговорим о еще одном объекте, который возникает в связи с конечной системой векторов в Евклидовом пространстве. Таким объектом является матрица Грама. Она в частности используется для определения объемов.

Определение. Пусть V – евклидово пространство и $v_1, \dots, v_k \in V$ – произвольный набор векторов (k НЕ обязательно равно размерности пространства). Тогда матрица

$$G(v_1, \dots, v_k) = \begin{pmatrix} (v_1, v_1) & \dots & (v_1, v_k) \\ \vdots & \ddots & \vdots \\ (v_k, v_1) & \dots & (v_k, v_k) \end{pmatrix} \in M_k(\mathbb{R})$$

называется матрицей Грама системы векторов v_1, \dots, v_k .⁷

Если e_1, \dots, e_n – некоторый базис пространства V , то $B = G(e_1, \dots, e_n)$ – матрица скалярного произведения заданная в базисе e_1, \dots, e_n . Таким образом матрица Грама – это некоторое обобщение матрицы билинейной формы.

Теперь вспомним, что у любой билинейной формы есть операторная запись. Давайте введем следующее обозначение: для произвольных векторов $w, u \in V$ положим $w \cdot u = (w, u)$. Тогда для набора $v = (v_1, \dots, v_k)$

⁷Обратите внимание, тут важен порядок векторов. То есть формально матрица Грама зависит от набора $(v_1, \dots, v_k) \in V^k$.

выполнено

$$G(v_1, \dots, v_k) = \begin{pmatrix} v_1 \\ \vdots \\ v_k \end{pmatrix} \cdot (v_1 \ \dots \ v_k) = v^t \cdot v$$

Пусть теперь $C \in M_{k \times r}(\mathbb{R})$ – некоторая матрица. Тогда из набора (v_1, \dots, v_k) можно построить новый набор $(u_1, \dots, u_r) = (v_1, \dots, v_k)C$ или кратко $u = vC$. Тогда

$$G(u) = G(vC) = (vC)^t \cdot vC = C^t v^t \cdot vC = C^t G(v)C$$

То есть $G((v_1, \dots, v_k)C) = C^t G(v_1, \dots, v_k)C$.

Утверждение. Пусть $v_1, \dots, v_k \in V$ – произвольный набор векторов в евклидовом пространстве. Тогда

1. Пусть $\alpha_1, \dots, \alpha_k \in \mathbb{R}$, введем обозначение $\alpha = (\alpha_1, \dots, \alpha_k)^t$. Тогда следующие условия эквивалентны:

(a) $\alpha_1 v_1 + \dots + \alpha_k v_k = 0$.

(b) $G(v_1, \dots, v_k)\alpha = 0$.

(c) $\alpha^t G(v_1, \dots, v_k)\alpha = 0$.

2. $\text{rk } G(v_1, \dots, v_k) = \dim \langle v_1, \dots, v_k \rangle$.

3. $\det G(v_1, \dots, v_k) \geq 0$. При этом равенство достигается тогда и только тогда, когда векторы линейно зависимы.

4. Если $C \in M_k(\mathbb{R})$ является матрицей элементарного преобразования I или II типа, то

$$\det G(v_1, \dots, v_k) = \det G((v_1, \dots, v_k)C)$$

Заметим, что из третьего пункта следует вот какое наблюдение. Если v_1, \dots, v_k – линейно независимый набор векторов, из которого процессом ортогонализации Грама-Шмидта мы получили набор u_1, \dots, u_k , то $\det G(u_1, \dots, u_k) = \det G(v_1, \dots, v_k)$.

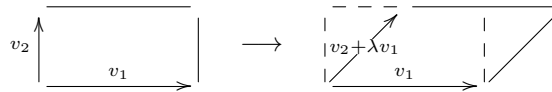
Объемы

Теперь самое время прикоснуться к объемам. Надо сказать, что существует общая теория вычисления объемов в евклидовом пространстве V . Она позволяет посчитать «объем» любого подмножества в V . Данная конструкция ведет к понятию меры Лебега и далее к интегралу Лебега. Мы, конечно же, не будем развивать подобную теорию в такой общности, а всего лишь ограничимся вычислением объемов для простых и естественных с точки зрения линейной алгебры фигур – многомерных параллелепипедов.

Определение. Пусть V – евклидово пространство и $v_1, \dots, v_k \in V$ – набор векторов, тогда k -мерным параллелепипедом натянутым на v_1, \dots, v_k называется следующее множество

$$\Pi(v_1, \dots, v_k) = \left\{ \sum_{i=1}^k x_i v_i \mid 0 \leq x_i \leq 1 \right\}$$

VIP пример Я хочу разобрать один важный пример. Пусть $V = \mathbb{R}^2$ – плоскость со стандартным скалярным произведением $(x, y) = x^t y$ и $v_1, v_2 \in V$ – два ортогональных вектора. Если я заменю v_2 на $v_2 + \lambda v_1$, то геометрически я наклоню мой параллелепипед вдоль направления v_1 как на рисунке ниже:



На рисунке мы видим, что параллелепипед справа отличается от параллелепипеда слева перестановкой треугольника отмеченного пунктиром. А значит их площади одинаковые. Давайте посмотрим на матрицы Грама двух наборов векторов:

$$G(v_1, v_2 + \lambda v_1) = \begin{pmatrix} 1 & 0 \\ \lambda & 1 \end{pmatrix} G(v_1, v_2) \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix} \quad \text{при этом} \quad G(v_1, v_2) = \begin{pmatrix} |v_1|^2 & 0 \\ 0 & |v_2|^2 \end{pmatrix}$$

То есть $\det G(v_1, v_2 + \lambda v_1) = \det G(v_1, v_2) = |v_1|^2 |v_2|^2 = S_{\Pi(v_1, v_2)}^2$. То есть определитель матрицы Грамма дает нам квадрат площади параллелограмма, натянутого на векторы v_1, v_2 . Этот пример подсказывает, как надо определять объемы в общем случае.

Определение. Пусть $v_1, \dots, v_k \in V$ – произвольный набор векторов в евклидовом пространстве. Тогда определим k -мерный объем параллелепипеда $\Pi(v_1, \dots, v_k)$ по следующей формуле:

$$\text{Vol}_k(\Pi(v_1, \dots, v_k)) = \sqrt{\det G(v_1, \dots, v_k)}$$

Давайте теперь покажем, что объем k -мерного параллелепипеда можно считать через площадь ($k - 1$ -мерный объем) основания на высоту.

Утверждение. Пусть $v_1, \dots, v_k \in V$ – произвольный набор векторов в евклидовом пространстве. Тогда

$$\text{Vol}_k(\Pi(v_1, \dots, v_k)) = \text{Vol}_{k-1}(\Pi(v_1, \dots, v_{k-1})) \rho(v_k, \langle v_1, \dots, v_{k-1} \rangle)$$

Последняя формула позволяет нам вычислять расстояние от вектора до подпространства с помощью объемов. А именно, если $v \in V$ и $L \subseteq V$ – подпространство с базисом e_1, \dots, e_k , то

$$\rho(v, L) = \sqrt{\frac{\det G(e_1, \dots, e_k, v)}{\det G(e_1, \dots, e_k)}}$$

Пример Давайте рассмотрим векторное пространство \mathbb{R}^n со стандартным скалярным произведением $(x, y) = x^t y$. И пусть даны векторы $v_1, \dots, v_n \in \mathbb{R}^n$. Сложим эти векторы в матрицу $A = (v_1 | \dots | v_n)$. Тогда

$$\text{Vol}_n(\Pi(v_1, \dots, v_n)) = \sqrt{\det(A^t A)} = |\det A|$$

Таким образом в ортонормированном базисе теория неориентированного объема превращается в теорию вычисления модуля определителя.

Утверждение. Пусть V – евклидово пространство размерности n , $v_1, \dots, v_n \in V$ – некоторый набор векторов и $\varphi: V \rightarrow V$ – линейный оператор. Тогда

$$\text{Vol}_n(\varphi(\Pi(v_1, \dots, v_n))) = |\det \varphi| \text{Vol}_n(\Pi(v_1, \dots, v_n))$$

Утверждение. Пусть $v_1, \dots, v_k \in V$ – набор векторов в евклидовом пространстве и $C \in M_k(\mathbb{R})$ – произвольная матрица. Тогда

$$\text{Vol}_k \Pi((v_1, \dots, v_k)C) = |\det C| \text{Vol}_k \Pi(v_1, \dots, v_k)$$

Таким образом, при линейной замене образующих параллелепипеда мы получаем другой параллелепипед, объем которого меняется на модуль определителя матрицы замены. Мы хотим избавиться от модуля в этой формуле, чтобы объем мог быть положительным и отрицательным.

Ориентированный объем

Конструкция ориентированного объема Пусть V – евклидово пространство размерности n . Рассмотрим все возможные наборы из n векторов – V^n . Множество V^n разбивается на две части: когда набор (v_1, \dots, v_n) линейно зависим и когда он линейно независим. В первом случае $\text{Vol}_n \Pi(v_1, \dots, v_n) = 0$, а во втором $\text{Vol}_n \Pi(v_1, \dots, v_n) \neq 0$. Мы хотим поделить все линейно независимые наборы (то есть базисы) на два класса: для одного класса объемы будут положительные, а для другого – отрицательные.

Определение. Пусть (v_1, \dots, v_n) и (u_1, \dots, u_n) – два базиса пространства V . Тогда существует единственная матрица $C \in M_n(\mathbb{R})$ такая, что $(v_1, \dots, v_n) = (u_1, \dots, u_n)C$.⁸ Будем говорить, что (v_1, \dots, v_n) эквивалентно (u_1, \dots, u_n) и писать $(v_1, \dots, v_n) \sim (u_1, \dots, u_n)$, если $\det C > 0$.

Утверждение. Пусть V – евклидово пространство размерности n . Тогда

1. Отношение, введенное на базисах, является отношением эквивалентности на множестве всех упорядоченных базисов

$$G_n(V) = \{(v_1, \dots, v_n) \mid v_i \in V, v_i \text{ линейно независимы} \}$$

⁸Потому что любой вектор однозначно раскладывается по базису.

2. Множество упорядоченных базисов $G_n(V)$ разбивается на два класса эквивалентности.

Таким образом все базисы у нас поделились на две группы. Какую-то из этих групп нам надо назвать положительной, другую отрицательной. Какую выбрать – это наша свобода. После подобного выбора все наборы в V^n делятся на три группы: (1) линейно зависмые, у них объем ноль, (2) положительные, натянутые на них параллелепипеды имеют положительный объем, (3) отрицательные, натянутые на них параллелепипеды имеют отрицательный объем. Положительные базисы будем еще называть положительно ориентированными, а отрицательные – отрицательно ориентированными. Если зафиксированы положительные и отрицательные базисы, будем говорить, что на V зафиксирована ориентация.

Определение. Пусть V – евклидово пространство с фиксированной ориентацией. Тогда определим ориентированный n -мерный объем следующим образом. Пусть $(v_1, \dots, v_n) \in V$, тогда

$$\text{Vol}_n^{\text{or}} \Pi(v_1, \dots, v_n) = \begin{cases} 0, & v_1, \dots, v_n \text{ линейно зависмы} \\ \text{Vol}_n \Pi(v_1, \dots, v_n), & (v_1, \dots, v_n) \text{ положительно ориентирован} \\ -\text{Vol}_n \Pi(v_1, \dots, v_n), & (v_1, \dots, v_n) \text{ отрицательно ориентирован} \end{cases}$$

Утверждение. Пусть V – евклидово пространство размерности n с фиксированной ориентацией, $(v_1, \dots, v_n) \in V^n$ и $C \in M_n(\mathbb{R})$. Тогда

$$\text{Vol}_n^{\text{or}} \Pi((v_1, \dots, v_n)C) = \det C \text{Vol}_n^{\text{or}} \Pi(v_1, \dots, v_n)$$

Пример Пусть $V = \mathbb{R}^n$ со стандартным скалярным произведением $(x, y) = x^t y$. И пусть ориентация зафиксирована так, что стандартный базис является положительным. Возьмем $v_1, \dots, v_n \in \mathbb{R}^n$ произвольный набор векторов. Образует матрицу $A = (v_1 | \dots | v_n) \in M_n(\mathbb{R})$. Тогда

- $G(v_1, \dots, v_n) = A^t A$.
- $\det G(v_1, \dots, v_n) = \det A^2$.
- $\text{Vol}_n \Pi(v_1, \dots, v_n) = |\det A|$.
- $\text{Vol}_n^{\text{or}} \Pi(v_1, \dots, v_n) = \det A$.

Заметьте, что ориентация набора (как и знак соответствующего объема) меняется при перестановке векторов в наборе на знак совершенной перестановки. Другая причина знака объема – смена направления вектора, то есть когда вектор v_i в наборе меняется на вектор $-v_i$.

Утверждение. Пусть V – ориентированное евклидово пространство размерности n , $v_1, \dots, v_n \in V$ – некоторый набор векторов и $\varphi: V \rightarrow V$ – линейный оператор. Тогда

$$\text{Vol}_n^{\text{or}}(\varphi(\Pi(v_1, \dots, v_n))) = \det \varphi \text{Vol}_n^{\text{or}}(\Pi(v_1, \dots, v_n))$$

Семинар 5

Общая информация:

- Напомню, что стандартным скалярным произведением на \mathbb{R}^n называется $(x, y) = x^t y$.
- Через $\mathbb{R}[x]_{\leq n}$ обозначается пространство многочленов степени не более n , то есть $\mathbb{R}[x]_{\leq n} = \{a_0 + a_1 x + \dots + a_n x^n \mid a_i \in \mathbb{R}\}$.

Задачи:

1. В пространстве \mathbb{R}^4 задана билинейная форма

$$\beta(x, y) = 2x_2 y_1 + x_4 y_4$$

По ней построили квадратичную форму $Q: \mathbb{R}^4 \rightarrow \mathbb{R}$. После этого Q ограничили на подпространство $V = \{x \in \mathbb{R}^4 \mid x_1 - 3x_2 - 3x_3 + x_4 = 0\}$. Найдите сигнатуру Q и сигнатуру ограничения Q на V .

2. Рассмотрим евклидово пространство $\mathbb{R}[x]_{\leq 3}$ со скалярным произведением $(f, g) = \int_{-1}^1 f(x)g(x) dx$. Методом Грама-Шмидта ортогонализуйте базис $1, x, x^2, x^3$.
3. Найти длины сторон, внутренние углы и площадь треугольника ABC в пространстве \mathbb{R}^5 со стандартным скалярным произведением, где

$$A = \begin{pmatrix} 2 \\ 4 \\ 2 \\ 4 \\ 2 \end{pmatrix}, \quad B = \begin{pmatrix} 6 \\ 4 \\ 4 \\ 4 \\ 6 \end{pmatrix}, \quad C = \begin{pmatrix} 5 \\ 7 \\ 5 \\ 7 \\ 2 \end{pmatrix}$$

4. Пусть $U \subseteq \mathbb{R}^4$ – векторное подпространство заданное следующим образом $U = \langle v_1, v_2, v_3, v_4 \rangle$, где

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 4 \\ 3 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 5 \\ 5 \\ 8 \end{pmatrix}, \quad v_3 = \begin{pmatrix} -2 \\ 6 \\ -2 \\ 24 \end{pmatrix}, \quad v_4 = \begin{pmatrix} 2 \\ -4 \\ 3 \\ -19 \end{pmatrix}$$

Задайте это подпространство в виде $U = \{y \in \mathbb{R}^4 \mid Ay = 0\}$ для некоторой матрицы $A \in M_{m4}(\mathbb{R})$.

5. Существует ли скалярное произведение на пространстве матриц $n \times n$ ($n > 1$), относительно которого матрица из всех единиц была бы ортогональна любой верхнетреугольной матрице? (Указание: билинейные формы можно задавать в разных базисах.)

Семинар 6

Линейные классификаторы

Начнем с простого примера. Предположим, что мы хотим разделить пространство \mathbb{R}^n на две части с помощью подпространства. Оказывается такое подпространство должно быть размерности $n - 1$. При наличии скалярного произведения такое подпространство очень легко задается с помощью одного единственного вектора ему ортогонального. Однако, подпространства всегда проходят через начало координат и потому это не лучшая конструкция. Тем не менее, смещение подпространства в сторону от начала координат делается смещением начала отсчета в векторном пространстве, это приводит к понятию линейного многообразия (просто сдвинутое подпространство). Давайте я сначала проговорю случай с чистым подпространством без сдвига, а потом общий случай, который и дает так называемый линейный классификатор.

Пример Пусть $V = \mathbb{R}^n$ и задано стандартное скалярное произведение $(x, y) = x^t y$. Пусть U – подпространство заданное системой следующего вида: $U = \{y \in \mathbb{R}^n \mid (a_1, \dots, a_n)y = 0\}$ (мы считаем, что хотя бы одно из чисел $a_i \neq 0$). Если положим $v = (a_1, \dots, a_n)^t$. То $U = \langle v \rangle^\perp$ по определению. То есть оно задается в виде $\{y \in \mathbb{R}^n \mid (y, v) = 0\}$ для некоторого фиксированного ненулевого вектора v . Заметим, что это подпространство делит все пространство на два класса:

1. Положительные векторы $\{y \in \mathbb{R}^n \mid (y, v) > 0\}$.
2. Отрицательные векторы $\{y \in \mathbb{R}^n \mid (y, v) < 0\}$.

Между этими классами как раз и расположено подпространство $\langle v \rangle^\perp$.

Пусть $w \in V$ – произвольный вектор. Давайте поймем как найти расстояние от w до подпространства U . Вспомним, что $w = \alpha v + u$, где $u \in U$, а вектор αv будет ортогональной составляющей v относительно U . Значит его длина и будет расстоянием. Надо лишь найти неизвестное α . Для этого умножим равенство $w = \alpha v + u$ скалярно на v и получим $(w, v) = \alpha(v, v)$. Откуда $\alpha = \frac{(w, v)}{(v, v)}$. Значит

$$\rho(w, U) = \left| \frac{(w, v)}{(v, v)} v \right| = \frac{|(w, v)|}{|v|} = |(w, v/|v|)|$$

Если вектор v имел единичную длину, то формула упрощается так

$$\rho(w, \langle v \rangle^\perp) = |(w, v)| = |w^t v|¹$$

Линейные многообразия По-простому, подпространства – это «линейные поверхности» проходящие через начало координат. А если мы их хотим сдвинуть, то такие штуки называются линейными многообразиями. А именно, пусть $V = \mathbb{R}^n$, $v \in V$ – некоторый вектор и $U \subseteq V$ – подпространство. Тогда множество $v + U$ называется линейным многообразием. В этом случае подпространство U однозначно определено и называется направляющим подпространством, а вектор v вообще говоря можно заменить на любой вектор вида $v + u$, где $u \in U$.

Линейные многообразия в \mathbb{R}^n Любое линейное многообразие в \mathbb{R}^n можно задать двумя способами:

1. $L = v + \langle v_1, \dots, v_k \rangle$, где v_1, \dots, v_k можно выбрать линейно независимыми.
2. $L = \{y \in \mathbb{R}^n \mid Ay = b\}$, где $A \in M_{m \times n}(\mathbb{R})$, $b \in \mathbb{R}^m$. Кроме того, строки матрицы A можно выбрать линейно независимыми.

Таким образом линейные многообразия возникают, когда мы решаем неоднородные линейные системы, а по своей сути они всего лишь сдвиги подпространств на какой-то вектор.

¹Последнее равенство в силу того, что у нас стандартное скалярное произведение.

Примеры

1. Прямая. Пусть $p, v \in V$ – некоторые векторы, причем $v \neq 0$. Тогда $p + \langle v \rangle$ – это сдвиг одномерного подпространства. Такое подпространство называется прямой. Явно на такой прямой лежат вектора вида $p + \lambda v$, где $\lambda \in \mathbb{R}$.
2. Гиперповерхность. Пусть $p \in V$ и $U \subseteq V$ подпространство размерности на 1 меньше, чем V . Тогда $p + U$ называется гиперповерхностью. Как было показано выше, $U = \langle v \rangle^\perp$ для некоторого ненулевого вектора v . В этом случае $p + U = \{y \in V \mid (y, v) = (p, v)\}$. Кроме этого мы можем разделить все пространство на два класса:
 - (a) Положительные векторы $\{y \in V \mid (y, v) > (p, v)\}$.
 - (b) Отрицательные векторы $\{y \in V \mid (y, v) < (p, v)\}$.

Гиперповерхность является простейшим примером линейного классификатора. Если точки находятся по одну сторону от нее, то они относятся к одному классу, а по другую – к другому.

Пусть $w \in V$ – произвольный вектор. Как посчитать расстояние от w до $p + U$. Для этого надо сдвинуть начало координат в точку p и посчитать расстояние до подпространства U . Из вычислений ранее получаем

$$\rho(w, p + U) = \rho(w - p, U) = \frac{|(w - p, v)|}{|v|}$$

Сопряженный оператор

Пусть V – евклидово пространство и пусть $\phi: V \rightarrow V$ – линейный оператор. Тогда *сопряженный* к нему линейный оператор ϕ^* – это такой оператор, что $(\phi(v), u) = (v, \phi^*(u))$ для всех $v, u \in V$. Понять почему такой оператор вообще существует и единственный помогает вычисление в координатах.

В начале возьмем ортонормированный базис. В этом случае $V = \mathbb{R}^n$, $(x, y) = x^t y$, а $\phi(x) = Ax$, а $\phi^*(x) = Bx$. Тогда условие $(Ax, y) = (x, By)$ означает $x^t A^t y = x^t B y$. То есть $B = A^t$. То есть матрица для ϕ^* это A^t . Но это только в ортонормированном базисе.

В случае произвольного базиса скалярное произведение задается $(x, y) = x^t B y$, где B – симметричная невырожденная положительно определенная матрица. Тогда если $\phi x = Ax$ и $\phi^* x = A'x$, то условие $(Ax, y) = (x, A'y)$ расписывается так: $(Ax)^t B y = x^t B A' y$. То есть $x^t A^t B y = x^t B A' y$ для всех $x, y \in \mathbb{R}^n$. Последнее значит, что $A^t B = B A'$. Значит $A' = B^{-1} A^t B$ – это формула связывает матрицу ϕ и ϕ^* в произвольных базисах.

Замечание

- Обратите внимание, что определение ϕ^* очень формальное. Совершенно неясно, что это такое геометрически. Более того, никто не знает и не узнает, что это такое геометрически. Тут ситуация как с матричным умножением, есть какие-то частные случаи, когда мы можем как-то понять, что значит ϕ^* , но в общем случае – только считать по формуле.
- Технически про сопряженный оператор надо думать так, если вы встретили внутри скалярного произведения оператор ϕ , то его можно «перебросить» в другую сторону скалярного произведения, но при этом сам оператор заменится на ϕ^* .
- Почему-то оказывается, что это понятие очень хорошо описывает важные ситуации, когда действие оператора согласовано со скалярным произведением. Мы поговорим о таких ситуациях ниже.

Проекторы и ортопроекторы

Пусть V – некоторое векторное пространство и $U, W \subseteq V$ – некоторые подпространства. Будем говорить, что V раскладывается в прямую сумму этих подпространств, если $U \cap W = 0$ и $V = U + W$, т.е. любой вектор $v \in V$ представляется в виде $v = u + w$, где $u \in U$ и $w \in W$ (то есть $U + W = \{u + w \mid u \in U, w \in W\}$). Думать про это надо так, U и W – это «непересекающиеся» подпространства и V является наименьшим пространством их содержащим. Такое разложение всегда получается так: берем какой-нибудь базис e_1, \dots, e_n пространства V , делим его на две части e_1, \dots, e_k и e_{k+1}, \dots, e_n и полагаем $U = \langle e_1, \dots, e_k \rangle$ и $W = \langle e_{k+1}, \dots, e_n \rangle$. Если пространство V является прямой суммой подпространств U и W , то мы будем обозначать это дело следующим

образом $V = U \oplus W$. В этом случае любой вектор v единственным образом раскладывается в виде $v = u + w$, где $u \in U$ и $w \in W$. Еще в этом случае $\dim U + \dim W = \dim V$.

Утверждение. Пусть V – евклидово пространство и $U \subseteq V$ – произвольное подпространство. Тогда $V = U \oplus U^\perp$.

Таким образом в евклидовом пространстве V при фиксированном подпространстве $U \subseteq V$, любой вектор $v \in V$ единственным образом раскладывается в сумму $v = \text{pr}_U v + \text{ort}_U v$, где $\text{pr}_U v \in U$ и $\text{ort}_U v \in U^\perp$.

Определение. Если V – евклидово пространство, $U \subseteq V$ – произвольное подпространство и $v \in V$, то

- Вектор $\text{pr}_U v$ называется ортогональной проекцией v на U .
- Вектор $\text{ort}_U v$ называется ортогональной составляющей v относительно U .

Обратите внимание, что ортогональная проекция v на U – это проекция v на U вдоль U^\perp , а ортогональная составляющая – проекция v на U^\perp вдоль U .

Утверждение. Пусть V – евклидово пространство и $P: V \rightarrow V$ – некоторый оператор. Тогда

1. Оператор P является проектором тогда и только тогда, когда $P^2 = P$.
2. Оператор P является ортопроектором тогда и только тогда, когда $P^2 = P$ и $P^* = P$.

Обратите внимание, что выше мы имеем два описания проекторов (и ортопроекторов). Одно геометрическое, а второе чисто алгебраическое. Давайте поймем, как восстановить подпространства U (на которое мы проектируем) и W (вдоль которого мы проектируем) в случае, когда все что нам известно это лишь равенство $P^2 = P$.

Положим в этом случае $U = \text{Im } P$ и $W = \ker P$. Тогда

1. Для любого $u \in U$ верно $Pu = u$.
2. $U \cap W = 0$.
3. Любой вектор $v \in V$ есть сумма $v = u + w$ для некоторых $u \in U$ и $w \in W$.

Действительно, если $u \in U$, это значит, что $u = Pv$. Тогда $Pu = P^2v = Pv = u$. Если вектор $v \in U \cap W$, то с одной стороны $v = Pv$ из предыдущего, с другой стороны $Pv = 0$ по определению W . Для любого вектора $v \in V$ верно $v = Pv + (E - P)v$. Тогда вектор $Pv \in U$, а вектор $(E - P)v \in W$, так как $P(E - P)v = (P - P^2)v = 0$.

Если же дополнительно выполнено $P^* = P$, то $U = \text{Im } P$ и $W = \ker P$ оказываются ортогональными друг другу. Действительно, если $z \in U$, то $z = Pv$ и если $w \in W$, то

$$(z, w) = (Pv, w) = (v, P^*w) = (v, Pw) = (v, 0) = 0$$

То есть любой вектор из U ортогонален любому вектору из W . Можно показать, что верно и в обратную сторону, если $U \perp W$, то это влечет условие $P^* = P$.

Давайте заметим еще одну вещь. Если $P^2 = P$, это означает, что многочлен $f(t) = t^2 - t$ является аннулирующим для P . То есть спектр P лежит среди корней этого многочлена, а именно, среди $\{0, 1\}$. Кроме того, так как этот многочлен не имеет кратных корней, то оператор P диагонализуем и диагональный вид есть

$$A_P = \begin{pmatrix} E & 0 \\ 0 & 0 \end{pmatrix}$$

Где размер блока E равен размерности образа и рангу оператора r . Пусть f_1, \dots, f_n – базис, в котором P диагонализуется, тогда $U = \langle f_1, \dots, f_r \rangle$ и $W = \langle f_{r+1}, \dots, f_n \rangle$. Еще полезно заметить, что след для проектора будет равен размерности U . Это следует из вычисления следа по матрице A_P .

Формула БАБА

Давайте я в начале разберу задачу нахождения проекции вектора на подпространство вдоль другого подпространства (здесь нам не нужно никакое скалярное произведение). Пусть V – некоторое векторное пространство и $V = U \oplus W$. Тогда на пространстве V задан оператор проекции $P: V \rightarrow V$ такой, что $\ker P = W$ и $P|_U = \text{Id}$, то есть, если $v \in V$ раскладывается в сумму $v = u + w$, где $u \in U$ и $w \in W$, то $Pv = u$ – оператор вычисления проекции на U вдоль W .

Теперь мы хотим научиться эффективно считать P . Для этого предположим $V = \mathbb{R}^n$, $U = \langle u_1, \dots, u_k \rangle$, $W = \{y \in \mathbb{R}^n \mid Ay = 0\}$, где $A \in M_{s \times n}(\mathbb{R})$. В этом случае $P: \mathbb{R}^n \rightarrow \mathbb{R}^n$ задается некоторой матрицей. Наша задача – найти эту матрицу.

Предположим для простоты, что векторы u_1, \dots, u_k образуют базис U , а строки матрицы A линейно независимы. Определим матрицу $B = (u_1 | \dots | u_k) \in M_{n \times k}(\mathbb{R})$. Тогда утверждаются следующие вещи:

1. Количество столбцов B совпадает с количеством строк A , то есть $k = s$.
2. Матрица AB обратима.
3. Оператор проекции задается формулой $P = B(AB)^{-1}A$. Мнемоническое правило «БАБА».

Доказательство. Матрица A задает линейное отображение $A: \mathbb{R}^n \rightarrow \mathbb{R}^s$ такое, что $\ker A = W$ и $\text{Im } A = \mathbb{R}^s$ (так как строки матрицы A линейно независимы, то $\text{rk } A = s$, но $\text{rk } A = \dim \text{Im } A$). Матрица B задает отображение $B: \mathbb{R}^k \rightarrow \mathbb{R}^n$ такое, что $\text{Im } B = U$ и $\ker B = 0$ (так как столбцы B линейно независимы).

(1) Мы знаем, что

$$\begin{array}{lcl} \dim U + \dim W = n & \text{то есть} & k + \dim W = n \\ \dim \ker A + \dim \text{Im } A = n & & \dim W + s = n \end{array} \quad \text{откуда} \quad s = k$$

(2) Теперь рассмотрим отображение $AB: \mathbb{R}^k \rightarrow \mathbb{R}^k$. Заметим, что $\text{Im } B \cap \ker A = U \cap W = 0$. Значит $\ker AB = 0$, то есть AB – обратимый оператор.

(3) Теперь выведем формулу для P . Пусть $v = u + w$, где $v \in \mathbb{R}^n$ – произвольный вектор, $u \in U$ и $w \in W$ – его единственное разложение по прямой сумме подпространств. Тогда $Av = Au + Aw = Au$. С другой стороны, так как $u \in U$, мы имеем $u = Bx$ для некоторого $x \in \mathbb{R}^k$. Тогда $Av = ABx$. Так как AB обратимая квадратная матрица, имеем $x = (AB)^{-1}Av$. Значит $u = Bx = B(AB)^{-1}Av$, что и требовалось. \square

Обратите внимание, что проектор P на U вдоль W зависит от двух подпространств, а не только от U . Если вы измените одно из них, то проектор изменится.

Формула Атата

Теперь я хочу разобрать случай проектора на подпространство вдоль его ортогонального дополнения. Такой проектор называется ортопроектором. Пусть $V = \mathbb{R}^n$ со стандартным скалярным произведением $(x, y) = x^t y$ и пусть подпространство $U \subseteq V$ задано своим базисом $U = \langle u_1, \dots, u_k \rangle$. Составим матрицу $A = (u_1 | \dots | u_k) \in M_{n \times k}(\mathbb{R})$. Тогда $U^\perp = \{y \in \mathbb{R}^n \mid A^t y = 0\}$. Пусть теперь $v \in V$ – произвольный вектор и $v = \text{pr}_U v + \text{ort}_U v$. Тогда формула «БАБА» превращается в $\text{pr}_U v = A(A^t A)^{-1} A^t v$. Мнемоническое правило для запоминания: в евклидовом пространстве БАБА – это Атата.

Обратите внимание, что проектор P всегда зависит от двух подпространств: то, на которое проектируем U , и то, вдоль которого проектируем W . Но в случае ортогонального проектирования $W = U^\perp$, потому ортопроектор P реально зависит только от одного подпространства.

Метод наименьших квадратов

Пусть мы хотим решить систему $Ax = b$, где $A \in M_{m \times n}(\mathbb{R})$, $b \in \mathbb{R}^m$ и $x \in \mathbb{R}^n$ – столбец неизвестных. И предположим, что система не имеет решений, но от этого наше желание ее решить не становится слабее. Давайте обсудим, как удовлетворить наши желания в подобной ситуации и когда такие ситуации обычно встречаются.

Введем на пространстве \mathbb{R}^m стандартное скалярное произведение $(x, y) = x^t y$. Тогда, на процесс решения системы можно смотреть так: мы подбираем $x \in \mathbb{R}^n$ так, чтоб $|Ax - b| = 0$. Если решить систему невозможно,

то этот подход подсказывает, как надо поступить. Надо пытаться минимизировать расстояние между Ax и b . То есть решить задачу

$$|Ax - b| \rightarrow \min \\ x \in \mathbb{R}^n$$

Теперь давайте поймем, как надо решать такую задачу. Пусть матрица A имеет вид $A = (A_1 | \dots | A_n)$, где $A_i \in \mathbb{R}^m$ – ее столбцы. Тогда система $Ax = b$ означает, $x_1 A_1 + \dots + x_n A_n = b$. То есть система разрешима тогда и только тогда, когда $b \in \langle A_1, \dots, A_n \rangle$. Значит наша задача минимизировать расстояние между b и $\langle A_1, \dots, A_n \rangle$. Мы можем разложить вектор b на проекцию и ортогональную составляющую относительно $\langle A_1, \dots, A_n \rangle$. Обычная теорема пифагора говорит, что минимум расстояния достигается на $b_0 = \text{pr}_{\langle A \rangle} b$. В этом случае вместо исходной системы $Ax = b$ мы должны решить систему $Ax = b_0$. И если x_0 – ее решение, то $|Ax_0 - b|$ как раз и будет минимальным.

Давайте теперь предположим, что столбцы матрицы A линейно независимы. Тогда по формуле «Атата» мы знаем, что $b_0 = A(A^t A)^{-1} A^t b$. Кроме этого должно выполняться $b_0 = Ax_0$. Так как столбцы A линейно независимы, такое x_0 должно быть единственным. Но мы видим, что в качестве x_0 подходит $x_0 = (A^t A)^{-1} A^t b$.

Как организовывать вычисления Формула $x_0 = (A^t A)^{-1} A^t b$ выглядит очень приятной, но реально по ней считать – то еще удовольствие. Если вы хотите считать руками, то куда проще рассмотреть систему $A^t A x = A^t b$ и уже решать ее методом Гаусса.

Если же вы хотите считать проекцию по формуле $b_0 = A(A^t A)^{-1} A^t b$, то не надо считать матрицу $A(A^t A)^{-1} A^t$. Она будет большой и страшной. Вместо этого лучше применять части этой формулы к b справа налево, а именно считаем такую последовательность векторов

$$b \mapsto A^t b \mapsto (A^t A)^{-1} A^t b \mapsto A(A^t A)^{-1} A^t b$$

При этом второй переход лучше организовать как решение системы $A^t A x = A^t b$.

Движения и ортогональные матрицы

Так как углы и расстояния выражаются через скалярное произведение и наоборот, мы получаем следующее:

Утверждение. Пусть теперь $\phi: V \rightarrow V$ – линейный оператор в евклидовом пространстве. Следующие утверждения эквивалентны:

1. ϕ сохраняет скалярное произведение, т.е. $(\phi(v), \phi(u)) = (v, u)$ для любых $v, u \in V$.
2. ϕ сохраняет длины и углы, т.е. $|\phi(v)| = |v|$ и $\alpha_{\phi(v), \phi(u)} = \alpha_{v, u}$ для всех $v, u \in V$.
3. ϕ сохраняет длины, т.е. $|\phi(v)| = |v|$ для всех $v \in V$.
4. $\phi^* = \phi^{-1}$
5. Для любого ортонормированного базиса e_1, \dots, e_n его образ $\phi(e_1), \dots, \phi(e_n)$ – тоже ортонормированный базис.
6. Для любого ортонормированного базиса матрица ϕ в этом базисе будет ортогональной.

Линейные операторы, обладающие одним из эквивалентных свойств выше, называются *движениями*. Теперь обсудим эти свойства.

Если первое выполняется, то сохраняется все, что можно посчитать через скалярные произведения. Например это длины и углы. Наоборот, из-за формулы $(v, u) = |v||u| \cos(\angle v, u)$ следует, что сохранение углов и длин сохраняет скалярное произведение. Ясно, что второе влечет третье условие. Но если выполнено третье условие, то у любого треугольника сохраняются длины сторон, значит сохраняются любые треугольники, а значит в них сохраняются углы.²

Пусть в V выбрали ортонормированный базис. Это значит, что V можно отождествить с \mathbb{R}^n и при этом скалярное произведение превращается в стандартное $(x, y) = x^t y$. Пусть отображение $\phi: \mathbb{R}^n \rightarrow \mathbb{R}^n$ задано матрицей $A \in M_n(\mathbb{R})$. Тогда условие движения записывается так $(Ax, Ay) = (x, y)$. То есть $x^t A^t A y = x^t y$ для любых $x, y \in \mathbb{R}^n$. То есть $A^t A = E$, то есть A должна быть ортогональной матрицей.

Из условия три следует, что ϕ в ноль переводит только ноль, а значит он обратим. Потому условие $(\phi v, \phi u) = (v, u)$ можно переписать как $(\phi v, u) = (v, \phi^{-1} u)$ (для этого подставить $\phi^{-1} u$ вместо u).

²Для любителей формальности можно сослаться на поляризационную формулу.

Утверждение. Пусть $C \in M_n(\mathbb{R})$ – ортогональная матрица и пусть $\lambda \in \mathbb{C}$ – ее собственное значение. Тогда

1. $\bar{\lambda}$ тоже является собственным значением для C .
2. $|\lambda| = 1$.
3. Собственные векторы для разных собственных значений ортогональны.

Примеры

1. Пусть $V = \mathbb{R}^2$ со стандартным скалярным произведением. Тогда любое движение это:
 - (a) центральная симметрия относительно начала координат $C = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$.
 - (b) симметрия относительно какой-то прямой $C = D^t \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} D$, где D – матрица поворота (см. далее).
 - (c) поворот на некоторый угол, $C = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$ – матрица поворота.
2. Пусть $V = \mathbb{R}^3$ со стандартным скалярным произведением и $C \in M_3(\mathbb{R})$ – ортогональная матрица. Тогда $\chi_C(t)$ – многочлен степени 3. Любой многочлен нечетной степени имеет хотя бы один вещественный корень.³ А значит это ± 1 . То есть соответствующий собственный вектор v либо неподвижен, либо отражается в $-v$ под действием C . Кроме того, ортогональное дополнение $\langle v \rangle^\perp$ является двумерной плоскостью, на которой C действует одним из трех способов описанных в предыдущем пункте. Короче говоря, если задано движение в трехмерном пространстве, то в каком-то ортонормированном базисе оно имеет один из следующих видов:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi & -\sin \varphi \\ 0 & \sin \varphi & \cos \varphi \end{pmatrix} \quad \text{или} \quad \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos \varphi & -\sin \varphi \\ 0 & \sin \varphi & \cos \varphi \end{pmatrix}$$

Первое из них является поворотом вокруг некоторой оси, а второе является поворотом вокруг оси и отражением вдоль оси.

Утверждение. Пусть V евклидово пространство и $\phi: V \rightarrow V$ – некоторый оператор. Тогда эквивалентно

1. ϕ является движением (ортогональный оператор).
2. В некотором ортонормированном базисе матрица оператора ϕ имеет вид:

$$A_\phi = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_r \end{pmatrix}, \quad \text{где } A_i \text{ либо } 1, \text{ либо } -1, \text{ либо } \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$$

Для ортогональной матрицы $\det C = \pm 1$ (примените \det к равенству $C^t C = E$). Если $\det C = 1$, движение называется *собственным* и если $\det C = -1$, то *несобственным*.

Если e_1, \dots, e_n и f_1, \dots, f_n – ортонормированные базисы пространства V и пусть $(f_1, \dots, f_n) = (e_1, \dots, e_n)C$, где C – матрица перехода. Тогда C является ортогональной матрицей. Это вторая ситуация, когда появляются ортогональные матрицы.

Самосопряженные операторы

Пусть V – евклидово пространство и пусть $\phi: V \rightarrow V$ – линейный оператор. Тогда оператор называется *самосопряженным*, если $\phi^* = \phi$. Если в ортонормированном базисе матрица ϕ есть $A \in M_n(\mathbb{R})$, то условие самосопряженности это равенство $A^t = A$. Это определение сугубо формальное, из него совершенно не понятно, как этот оператор действует в пространстве.

Утверждение. Пусть $\phi: V \rightarrow V$ – самосопряженный оператор в евклидовом пространстве. Тогда

³Потому что такое многочлен устроен $\chi(t) = t^n(1 + o(1))$ при $t \rightarrow \pm\infty$. То есть на плюс бесконечности многочлен уходит в плюс бесконечность, а на минус бесконечности – в минус бесконечность. То есть по непрерывности он где-то должен был пересечь горизонтальную ось координат. А эта точка и есть корень.

1. Все его собственные значения вещественны.
2. Собственные вектора с разными собственными значениями ортогональны друг другу.
3. Существует ортонормированный базис пространства V состоящий из собственных векторов ϕ .
4. В некотором ортонормированном базисе матрица ϕ имеет диагональный вид, с вещественными числами на диагонали.

Таким образом самосопряженный оператор – это растяжение во взаимно перпендикулярных направлениях. Об этом говорит последний пункт предыдущего утверждения. Это дает полное геометрическое описание самосопряженных операторов. Теперь переформулируем это утверждение на языке матриц.

Утверждение. Пусть $A \in M_n(\mathbb{R})$ – симметрическая матрица. Тогда

1. Все собственные значения A вещественные.
2. Все собственные вектора с разными собственными значениями ортогональны.
3. Существует ортогональная матрица $C \in M_n(\mathbb{R})$ такая, что $C^{-1}AC$ является диагональной вещественной матрицей.⁴

Алгоритм разложения симметрических матриц

Дано Матрица $A \in M_n(\mathbb{R})$ такая, что $A^t = A$.

Задача Найти разложение $A = C\Lambda C^t$, где $C \in M_n(\mathbb{R})$ – ортогональная матрица, $\Lambda \in M_n(\mathbb{R})$ – диагональная матрица.

Алгоритм

1. Найти собственные значения матрицы A .
 - (a) Составить характеристический многочлен $\chi(\lambda) = \det(A - \lambda E)$.
 - (b) Найти корни $\chi(\lambda)$ с учетом кратностей: $\{(\lambda_1, n_1), \dots, (\lambda_k, n_k)\}$, где λ_i – корни, n_i – кратности.
2. Для каждого λ_i найти ортонормированный базис в пространстве собственных векторов отвечающему λ_i .
 - (a) Найти ФСР системы $(A - \lambda_i E)x = 0$. Пусть это будет $v_1^i, \dots, v_{n_i}^i$. Обратите внимание, что их количество будет в точности равно кратности n_i .
 - (b) Ортогонализировать $v_1^i, \dots, v_{n_i}^i$ методом Грама-Шмидта. Обратите внимание, после ортогонализации останется ровно n_i векторов.
 - (c) Сделать каждый вектор длинны один: $v_j^i \mapsto \frac{v_j^i}{|v_j^i|}$.
3. Матрица Λ будет диагональной с числами $\lambda_1, \dots, \lambda_1, \lambda_2, \dots, \lambda_2, \dots, \lambda_k, \dots, \lambda_k$ на диагонали, где каждое λ_i повторяется n_i раз. Обратите внимание, всего получится n чисел.
4. Матрица C будет составлена из столбцов $v_1^1, \dots, v_{n_1}^1, v_1^2, \dots, v_{n_2}^2, \dots, v_1^k, \dots, v_{n_k}^k$. Обратите внимание, порядок собственных векторов соответствует порядку собственных значений в матрице Λ .

⁴Обратите внимание, что тут нет разницы между $C^{-1}AC$ и C^tAC , так как C ортогональная.

Сингулярное разложение (SVD)

Это утверждение я в начале сформулирую на матричном языке.

Утверждение. Пусть дана матрица $A \in M_{m \times n}(\mathbb{R})$. Тогда

1. Существует $U \in M_m(\mathbb{R})$ такая, что $U^t U = E$.
2. Существует $V \in M_n(\mathbb{R})$ такая, что $V^t V = E$.
3. Существует последовательность вещественных чисел $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$.

такие, что $A = U \Sigma V^t$, где

$$\Sigma = \begin{pmatrix} \overbrace{\begin{pmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & \ddots \\ & & & & 0 \end{pmatrix}}^n \end{pmatrix}^m$$

При этом последовательность чисел $\sigma_1, \sigma_2, \dots, \sigma_r$ определена однозначно.

Пусть столбцы матрицы U – это вектора u_i , а столбцы матрицы V – это вектора v_i . Тогда утверждение означает, что

$$A = \sigma_1 u_1 v_1^t + \dots + \sigma_s u_s v_s^t$$

То есть мы представили матрицу A в виде «ортогональной» суммы матриц ранга один, в том смысле, что все u_i ортогональны друг другу и все v_i ортогональны друг другу.

Геометрически сингулярное разложение означает следующее.

Утверждение. Пусть V и U – евклидовы или эрмитовы пространства и $\phi: V \rightarrow U$ – линейное отображение. Тогда существует ортонормированный базис e_1, \dots, e_n в V , ортонормированный базис f_1, \dots, f_m в U и последовательность вещественных чисел $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ такие, что матрица ϕ имеет вид

$$\phi(e_1, \dots, e_n) = (f_1, \dots, f_m) \begin{pmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & \ddots \\ & & & & 0 \end{pmatrix}$$

При этом числа $\sigma_1, \dots, \sigma_r$ определены однозначно и называются сингулярными значениями отображения ϕ .

Компактное сингулярное разложение Пусть $A \in M_{m \times n}(\mathbb{R})$, $U \in M_{m \times s}(\mathbb{R})$, $V \in M_{n \times s}(\mathbb{R})$ и $\Sigma \in M_s(\mathbb{R})$ – диагональная матрица с числами $\sigma_1 \geq \dots \geq \sigma_s > 0$ на диагонали. Предположим, что столбцы матриц U и V ортонормированны (то есть все между собой ортогональны и длины один). Тогда равенство вида $A = U \Sigma V^t$ называется компактным сингулярным разложением.

Если нам известно сингулярное разложение $A = U \Sigma V^t$, то компактное из него делается так: 1) составим матрицу U' , состоящую из первых s столбцов матрицы U , 2) составим матрицу V' , состоящую из первых s столбцов матрицы V , 3) определим матрицу Σ' как квадратную s на s матрицу с диагональю из матрицы Σ . Тогда $A = U' \Sigma' (V')^t$ будет компактным разложением.

Замечание Философский смысл этого разложения следующий. Пусть наша матрица – это прямоугольная черно-белая картинка, где числа – интенсивности черного цвета. На вектора v_i и u_i надо смотреть как на «ортогональные» компоненты «базовых» цветовых интенсивностей. А λ_i – это мощности этих самых сигналов. Потому, если λ_i достаточно малы, то наш глаз не способен различить соответствующие сигналы. Потому, если мы выкинем их из нашей матрицы, то на глаз, матрица A не будет отличаться от полученной.

Обычно в реальной жизни выходит, что достаточно только первых штук пять слагаемых. Тогда $A' = \lambda_1 v_1 u_1^t + \dots + \lambda_5 v_5 u_5^t$ будет на глаз не отличима от A . В чем же польза от такого? На хранение матрицы A

нам потребуется mn чисел. Для хранения матрицы A' нам надо 5 чисел λ_i и еще 5 пар векторов v_i и u_i , на хранение каждого из которых надо m и n чисел соответственно. Итого затраты $5 + 5m + 5n = 5(m + n + 1)$. Это дает огромный выигрыш в количестве хранимой информации и является основой для многих алгоритмов архивации с потерей данных вроде JPG.

Задача о низкоранговом приближении

Теперь я хочу пару слов сказать о том, в каком смысле описанные выше процедуры являются оптимальными или попросту говоря «самыми лучшими». Пусть $A \in M_{m,n}(\mathbb{R})$. Зададим на пространствах матриц скалярное произведение по формуле $(A, B) = \text{tr}(A^t B)$. Длина относительно заданного скалярного произведения называется нормой фробениуса и выражается следующим образом:

$$\|A\|_F = \sqrt{\sum_{ij} a_{ij}^2}$$

Если матрица A имеет вид $A = (A_1 | \dots | A_n)$, тогда $\|A\|_F^2 = \sum_{i=1}^n |A_i|^2$, где $|A_i|$ – длина относительно стандартного скалярного произведения для столбца A_i .

Теперь наша задача – заменить матрицу A на матрицу B ранга не выше k , причем мы хотим выбрать B ближайшей в смысле нормы фробениуса. То есть мы зафиксируем матрицу A и число k и будем решать задачу

$$\begin{cases} \|A - B\|_F \rightarrow \min_B \\ \text{rk } B \leq k \end{cases}$$

Важно понимать, что множество матриц ранга не выше k не образуют линейное подпространство в пространстве матриц. А значит, тут не получится решить эту задачу просто применением ортогональных проекторов. Кроме того, задача может иметь не единственное решение, в некоторых ситуациях ближайших матриц может оказаться бесконечное число.

Обратите внимание, что если $k \geq \text{rk } A$, то ответом будет сама матрица A . А если $k < \text{rk } A$, то оказывается, что SVD дает нужный ответ к данной задаче. Нужно найти для матрицы A сингулярное разложение. После чего, выбрать в качестве нужной матрицы матрицу

$$B_k = \sigma_1 u_1 v_1^t + \dots + \sigma_k u_k v_k^t$$

Алгоритм нахождения компактного сингулярного разложения

Дано Матрица $A \in M_{m,n}(\mathbb{R})$.⁵

Задача Найти разложение $A = U \Sigma V^t$, где $U \in M_{m,s}(\mathbb{R})$, $V \in M_{n,s}(\mathbb{R})$ – матрицы с ортонормированными столбцами, $\Sigma \in M_s(\mathbb{R})$ – диагональная матрица с элементами $\sigma_1 \geq \dots \geq \sigma_s > 0$ на диагонали.

Алгоритм

1. Составим матрицу $S = AA^t \in M_m(\mathbb{R})$. Тогда $S = U \Sigma^2 U^t$.
2. Так как $S^t = S$. То с помощью алгоритма для симметрических матриц найдем ее разложение $S = CDC^t$.⁶ Причем, обязательно получится, что диагональная матрица $D = \text{diag}(\lambda_1, \dots, \lambda_m)$ состоит из неотрицательных элементов и мы можем выбрать порядок так, чтобы $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m \geq 0$.
3. Пусть $C = (C_1 | \dots | C_m)$, тогда положим $U = (C_1 | \dots | C_s) \in M_{m,s}(\mathbb{R})$. А матрица $\Sigma \in M_s(\mathbb{R})$ будет диагональной с числами $\sigma_i = \sqrt{\lambda_i}$ на диагонали, то есть $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_s)$.
4. Теперь надо найти V из условия $A = U \Sigma V^t$.⁷ Положим $U = (u_1 | \dots | u_s)$ и $V = (v_1 | \dots | v_s)$. Тогда $A^t U \Sigma^{-t} = V$, то есть $v_i = \frac{1}{\sigma_i} A^t u_i$ при $1 \leq i \leq s$.

⁵Этот алгоритм рекомендуется применять при $m \leq n$, в противном случае, применить его к матрице A^t , а потом транспонировать полученное разложение.

⁶Здесь D будет диагональной матрицей, а C ортогональной.

⁷Обратите внимание, что Σ квадратная и обратимая матрица.

Алгоритм нахождения сингулярного разложения

Дано Матрица $A \in M_{m,n}(\mathbb{R})$.

Задача Найти разложение $A = U\Sigma V^t$, где $U \in M_m(\mathbb{R})$ ортогональная, $V \in M_n(\mathbb{R})$ ортогональная, $\Sigma \in M_{m,n}(\mathbb{R})$ содержит на диагонали элементы $\sigma_1 \geq \dots \geq \sigma_s > 0$, а все остальные нули.

Алгоритм

1. Составим матрицу $S = AA^t \in M_m(\mathbb{R})$. Тогда $S = U\Sigma\Sigma^tU^t$.
2. Так как $S^t = S$. То с помощью алгоритма для симметрических матриц найдем ее разложение $S = CDC^t$. Причем, обязательно получится, что диагональная матрица $D = \text{diag}(\lambda_1, \dots, \lambda_m)$ состоит из неотрицательных элементов и мы можем выбрать порядок так, чтобы $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m \geq 0$.
3. Тогда $U = C$, а $\Sigma\Sigma^t = D$. То есть $\sigma_i^2 = \lambda_i$. Так как $\sigma_i \geq 0$, то они находятся как $\sigma_i = \sqrt{\lambda_i}$.
4. Теперь надо найти V из условия $A = U\Sigma V^t$.⁸ Пусть $\sigma_1 \geq \dots \geq \sigma_s > 0$. Положим $U = (u_1 | \dots | u_m)$ и $V = (v_1 | \dots | v_n)$. Тогда $A^tU = V\Lambda^t$, то есть $v_i = \frac{1}{\sigma_i}A^tu_i$ при $1 \leq i \leq s$. Так мы находим первые s столбцов матрицы V .
5. Найдем ФСР для $\{y \in \mathbb{R}^n \mid Ay = 0\}$ и ортонормировать его (ортогонализуем Грамом-Шмидтом, а потом нормируем). Полученные векторы и будут оставшиеся v_{s+1}, \dots, v_n .

Замечания

1. Надо заметить, что нельзя попытаться составить матрицу A^tA и из нее найти матрицу V . Так как матрицы V и U определены не однозначно и зависят друг от друга. Если вы нашли какую-то матрицу U , то к ней подойдет не любая найденная матрица V , а только та, что является решением $A = U\Sigma V^t$.
2. Приведенным выше алгоритмом имеет смысл пользоваться, если у матрицы A количество строк меньше, чем количество столбцов. Если же столбцов меньше, чем строк, то надо найти сингулярное разложение для $A^t = U\Sigma V^t$. Тогда $A = V\Sigma^tU^t$ будет искомым сингулярным разложением для A .

Пример Пусть у нас есть матрица

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

Найдем ее сингулярное разложение. В начале рассмотрим

$$AA^t = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

Теперь надо найти хар многочлен, это будет

$$\chi_{AA^t}(t) = \det(tE - AA^t) = t^2 - 4t + 3$$

У многочлена два корня $\lambda_1 = 3$ и $\lambda_2 = 1$. Откуда получаем, что

$$\Sigma = \begin{pmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

Теперь найдем матрицу $U = (u_1 | u_2)$. Вектор u_1 найдем как собственный для λ_1 и нормируем его длину, а вектор u_2 найдем как собственный для λ_2 и нормируем его длину.

Для λ_1 надо решить систему $(AA^t - 3E)x = 0$, то есть систему с матрицей

$$AA^t - 3E = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & -1 \end{pmatrix}$$

ФСР такой системы состоит из вектора $(1, 1)^t$. Его длина $\sqrt{2}$. Потому $u_1 = (1/\sqrt{2}, 1/\sqrt{2})^t$.

⁸Обратите внимание Σ не обязательно квадратная и тем более не обязательно обратимая.

Аналогично для λ_2 надо решить систему $(AA^t - E)x = 0$, то есть систему с матрицей

$$AA^t - E = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

ФСР такой системы состоит из вектора $(-1, 1)^t$. Его длина $\sqrt{2}$. Потому $u_2 = (-1/\sqrt{2}, 1/\sqrt{2})^t$. Значит

$$U = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

Теперь найдем $V = (v_1 | v_2 | v_3)$. Первые два вектора находятся по формулам

$$v_1 = \frac{1}{\sigma_1} A^t u_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{pmatrix} \quad v_2 = \frac{1}{\sigma_2} A^t u_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{pmatrix}$$

Осталось найти последний вектор в матрице V . Для этого надо решить систему $Ax = 0$ и нормировать единственное решение этой системы. Решаем

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

ФСР такой системы будет $w = (-1, 1, 1)^t$. Его длина будет $\sqrt{3}$. Значит $v_3 = (-1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$. А значит матрица V будет иметь вид

$$V = \begin{pmatrix} \frac{2}{\sqrt{6}} & 0 & \frac{-1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{pmatrix}$$

И в итоге

t.me/postypashki_old/1229

t.me/postypashki_old/1229

$$\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{2}{\sqrt{6}} & 0 & \frac{-1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{pmatrix}^t$$

Или без транспонирования

$$\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix}$$

Семинар 6

Задачи:

1. «Решите» систему методом наименьших квадратов

$$\left(\begin{array}{ccc|c} 2 & 7 & -3 & 6 \\ 1 & -1 & 3 & 3 \\ 2 & 1 & -1 & 0 \\ 2 & -2 & 3 & -9 \end{array} \right)$$

2. Диагонализовать следующую симметричную матрицу в ортонормированном базисе (то есть получить разложение $A = CDC^t$, где C – ортогональная матрица, а D диагональная).

$$\begin{pmatrix} 5 & -1 & -1 \\ -1 & 5 & -1 \\ -1 & -1 & 5 \end{pmatrix}$$

3. Пусть задана матрица

$$A = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 3 & -1 \\ 1 & -1 & 3 \end{pmatrix}$$

Найдите какую-нибудь симметричную матрицу B такую, что $B^2 = A$.

4. Найти сингулярное разложение следующих матриц

$$(a) \quad \begin{pmatrix} 5 & 1 & 8 \\ 7 & 5 & 4 \end{pmatrix}, \quad (b) \quad \begin{pmatrix} 1 & 1 & 2 & 2 \\ 2 & 2 & 1 & 1 \end{pmatrix}$$