

Линейная алгебра

Дима Трушин

Семинар 1

Немного общих слов

В линейной алгебре основные объекты изучения – векторные пространства, линейные отображения и билинейные формы на них. Существует два языка говорить об этих объектах:

1. Язык конкретных векторных пространств.
2. Язык абстрактных векторных пространств.

Давайте я пару слов скажу о каждой из составляющих, чтобы ориентироваться, что там есть и для чего все это нужно.

Конкретные векторные пространства Когда мы в школе изучали геометрию, то мы запомнили, что \mathbb{R} – это прямая. Когда надо было работать с плоскостью, мы на ней вводили координаты и потому каждая точка описывалась парой чисел (x, y) . Множество таких пар есть \mathbb{R}^2 – плоскость. В трех мерном пространстве все описывалось тремя координатами и мы получаем \mathbb{R}^3 . По аналогии получается пространство \mathbb{R}^n – n -мерное векторное пространство. Оно удобно тем, что с ним понятно как работать. Тут возникают такие вещи как: системы линейных уравнений, матрицы, определители, ранги матриц и прочие технические вещи. Эта часть линейной алгебры полезна, когда надо что-то посчитать руками и по-другому никак. Однако, основной недостаток в том, что на этом языке не все удобно делать и не все связи между объектами можно легко увидеть. Именно поэтому математики придумали абстрактный подход, а не для того, чтобы замучить людей до смерти.

Абстрактные векторные пространства Утиный принцип из программирования на самом деле был заимствован из математики, как и интерфейсы. Для того, чтобы было проще работать со сложными объектами, мы их скрываем за удобными лаконичными интерфейсами. У векторных пространств тоже есть свой интерфейс, он называется определением. Грубо говоря, к векторным пространствам относится все что угодно, у чего есть элементы при этом эти элементы можно складывать и умножать на числа.¹ На этом языке в линейной алгебре возникают не только векторные пространство, но и линейные отображения, операторы, билинейные формы, скалярные произведения, собственные значения и собственные векторы и куча других любопытных вещей. Понимание этих геометрических слов, помогает искать правильные подходы к решению задач. Однако, для начала я бы хотел поговорить о простых вещах. Потому начнем с систем линейных уравнений и матриц. Уже на эти темы полно серьезных задач.

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):

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

¹Еще требуются некоторые свойства, но это уже детали контракта интерфейса.

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.²

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

Прямой ход алгоритма Гаусса

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

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & | & b_1 \\ 0 & a_{22} & a_{23} & a_{24} & | & b_2 \\ a_{31} & a_{32} & a_{33} & a_{34} & | & b_3 \end{pmatrix} \quad \begin{array}{l} \text{3-я строка} - \frac{a_{32}}{a_{22}} \cdot \text{2-я строка} \end{array}$$

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & | & b_1 \\ 0 & a_{22} & a_{23} & a_{24} & | & b_2 \\ 0 & 0 & a_{33} & a_{34} & | & b_3 \end{pmatrix}$$

В результате данного хода какие-то коэффициенты, например a_{33} , могли занулиться, потому возможны следующие принципиально другие случаи⁴

$$\begin{pmatrix} \underline{a_{11}} & a_{12} & a_{13} & a_{14} & | & b_1 \\ 0 & \underline{a_{22}} & a_{23} & a_{24} & | & b_2 \\ 0 & 0 & 0 & \underline{a_{34}} & | & b_3 \end{pmatrix} \quad \begin{pmatrix} \underline{a_{11}} & a_{12} & a_{13} & a_{14} & | & b_1 \\ 0 & 0 & \underline{a_{23}} & a_{24} & | & b_2 \\ 0 & 0 & 0 & \underline{a_{34}} & | & b_3 \end{pmatrix} \quad \begin{pmatrix} \underline{a_{11}} & a_{12} & a_{13} & a_{14} & | & b_1 \\ 0 & \underline{a_{22}} & a_{23} & a_{24} & | & b_2 \\ 0 & 0 & 0 & 0 & | & \underline{b_3} \end{pmatrix} \quad \begin{pmatrix} \underline{a_{11}} & a_{12} & a_{13} & a_{14} & | & b_1 \\ 0 & \underline{a_{22}} & a_{23} & a_{24} & | & b_2 \\ 0 & 0 & 0 & 0 & | & 0 \end{pmatrix}$$

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.

Обратный ход алгоритма Гаусса

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

$$\begin{pmatrix} \underline{a_{11}} & a_{12} & a_{13} & a_{14} & | & b_1 \\ 0 & \underline{a_{22}} & a_{23} & a_{24} & | & b_2 \\ 0 & 0 & \underline{a_{33}} & a_{34} & | & b_3 \end{pmatrix} \quad \begin{array}{l} \text{разделить } i\text{-ю строку на } a_{ii} \\ \text{2-я строка} - a_{23} \cdot \text{3-я строка} \end{array}$$

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

$$\begin{pmatrix} 1 & a_{12} & 0 & a_{14} & | & b_1 \\ 0 & 1 & 0 & a_{24} & | & b_2 \\ 0 & 0 & 1 & a_{34} & | & b_3 \end{pmatrix}$$

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

²Данный 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.

³При переходе от одной матрицы к другой я новым коэффициентам даю старые имена, чтобы не захламлять текст новыми обозначениями.

⁴Это не полный список всех случаев.

В специальных случаях приведенных выше, получим

$$\left(\begin{array}{cccc|c} 1 & 0 & a_{13} & 0 & b_1 \\ 0 & 1 & a_{23} & 0 & b_2 \\ 0 & 0 & 0 & 1 & b_3 \end{array}\right) \quad \left(\begin{array}{cccc|c} 1 & a_{12} & 0 & 0 & b_1 \\ 0 & 0 & 1 & 0 & b_2 \\ 0 & 0 & 0 & 1 & b_3 \end{array}\right) \quad \left(\begin{array}{cccc|c} 1 & 0 & a_{13} & a_{14} & 0 \\ 0 & 1 & a_{23} & a_{24} & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right) \quad \left(\begin{array}{cccc|c} 1 & 0 & a_{13} & a_{14} & b_1 \\ 0 & 1 & a_{23} & a_{24} & b_2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

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.

Получение решений

В системе ниже, выберем переменную x_4 как параметр

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & a_{14} & b_1 \\ 0 & 1 & 0 & a_{24} & b_2 \\ 0 & 0 & 1 & a_{34} & b_3 \end{array}\right)$$

Тогда решения имеют вид⁵

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

Специальные случаи:

$$\begin{array}{ll} \left(\begin{array}{cccc|c} 1 & 0 & a_{13} & 0 & b_1 \\ 0 & 1 & a_{23} & 0 & b_2 \\ 0 & 0 & 0 & 1 & b_3 \end{array}\right) & \text{Решения: } \begin{pmatrix} x_1 \\ x_2 \\ x_4 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} - x_3 \begin{pmatrix} a_{13} \\ a_{23} \\ 0 \end{pmatrix} \\ \left(\begin{array}{cccc|c} 1 & a_{12} & 0 & 0 & b_1 \\ 0 & 0 & 1 & 0 & b_2 \\ 0 & 0 & 0 & 1 & b_3 \end{array}\right) & \text{Решения: } \begin{pmatrix} x_1 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} - x_2 \begin{pmatrix} a_{12} \\ 0 \\ 0 \end{pmatrix} \\ \left(\begin{array}{cccc|c} 1 & 0 & a_{13} & a_{14} & 0 \\ 0 & 1 & a_{23} & a_{24} & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right) & \text{Решения: Нет решений, т.к. последнее уравнение } 0 = 1 \\ \left(\begin{array}{cccc|c} 1 & 0 & a_{13} & a_{14} & b_1 \\ 0 & 1 & a_{23} & a_{24} & b_2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) & \text{Решения: } \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} - x_3 \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix} - x_4 \begin{pmatrix} a_{14} \\ a_{24} \end{pmatrix} \end{array}$$

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.

⁵Операция 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).

Матрицы

Матрица – это прямоугольная таблица чисел

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

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.⁶

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

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

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

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

На умножение матриц можно смотреть следующим образом. Чтобы получить коэффициент c_{ij} 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} .

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:⁷

1. **Ассоциативность сложения** $(A + B) + C = A + (B + C)$ для любых $A, B, C \in M_{m,n}(\mathbb{R})$

⁶Важно, directX и OpenGL используют вектор-строки! Потому часть инженерной литературы на английском связанной с трехмерной графикой оперирует со строками.

⁷Все эти свойства объединяет то, что они являются аксиомами в различных определениях для алгебраических структур. Позже мы столкнемся с такими структурами.

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.

Нехорошие свойства 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 определить на матрицах операцию деления.
2. Умножение матриц НЕ коммутативно. Действительно

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

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

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

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 .

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

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

Задача. Пусть $A \in M_{mn}(\mathbb{R})$ и $B \in M_{nk}(\mathbb{R})$. Покажите, что

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

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

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 $X, Y \in M_{n \times m}(\mathbb{R})$. 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:

$$i \quad \begin{matrix} & & j & \\ \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \ddots & \lambda & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix} \end{matrix}$$

Тогда прямая проверка показывает, умножение $A \in M_{n \times m}(\mathbb{R})$ на $S_{ij}(\lambda)$ слева прибавляет j строку умноженную на λ к i строке матрицы A , а умножение $B \in M_{m \times n}(\mathbb{R})$ 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 $A \in M_{n \times m}(\mathbb{R})$ на U_{ij} слева переставляет i и j строки матрицы A , а умножение $B \in M_{m \times n}(\mathbb{R})$ 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

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

Тогда прямая проверка показывает, умножение $A \in M_{n \times m}(\mathbb{R})$ на $D_i(\lambda)$ слева умножает i строку A на λ , а умножение $B \in M_{m \times n}(\mathbb{R})$ 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.

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

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.

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 $AR = E$.

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 .

Линейная зависимость в \mathbb{R}^m

Пусть нам даны векторы $v_1, \dots, v_n \in \mathbb{R}^m$. Будем говорить, что они линейно зависимы, если найдется ненулевой набор чисел $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ (то есть хотя бы одно из $\lambda_i \neq 0$), что $\lambda_1 v_1 + \dots + \lambda_n v_n = 0$. В противном случае они называются линейно независимыми. То есть, в случае линейно независимых v_1, \dots, v_n , какие бы числа $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ мы ни взяли, если хотя бы одно из чисел λ_i не ноль, то и сумма $\lambda_1 v_1 + \dots + \lambda_n v_n$ будет не ноль. Выражения вида $\lambda_1 v_1 + \dots + \lambda_n v_n$ называется линейной комбинацией векторов v_1, \dots, v_n .

Примеры

1. Рассмотрим \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 $v \in \mathbb{R}^n$ линейно зависим тогда и только тогда, когда он равен нулю.
3. Два вектора $v, u \in \mathbb{R}^n$ линейно зависимы тогда и только тогда, когда они пропорциональны, т.е. либо $v = \lambda u$ либо $u = \lambda v$.
4. Пусть $A \in M_{m,n}(\mathbb{R})$ имеет вид $A = (A_1 | \dots | A_n)$. Тогда система вида $Ax = 0$ переписывается так $x_1 A_1 + \dots + x_n A_n = 0$. То есть решение однородной СЛУ – это выяснение какие есть линейные комбинации между столбцами матрицы A . Если же нам дана система $Ax = b$, то она переписывается в виде $x_1 A_1 + \dots + x_n A_n = b$. То есть решения СЛУ – это выяснение как можно разложить вектор b по столбцам матрицы A .

Подстановка матриц в многочлен

Пусть $p(x) = a_0 + a_1 x + \dots + a_n x^n$ – многочлен с вещественными коэффициентами, а $A \in M_n(\mathbb{R})$. Тогда можно определить $p(A) = a_0 E + a_1 A^1 + \dots + a_n A^n$, где E – единичная матрица.

Задача. Докажите, что для любой матрицы $A \in M_n(\mathbb{R})$ найдется многочлен $p(x)$ степени $n^2 + 1$ такой, что $p(A) = 0$.⁹

Подстановка матрицы в многочлены помогает построить из исходной матрицы другие с заданным свойством, а многочлен зануляющий нашу матрицу может стать неплохой исходной точкой для подобных манипуляций.

⁹На самом деле можно показать, что найдется многочлен степени n .

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.