Линейная алгебра 1

Байрамкулов Аслан

Руководитель группы моделирования профиля
МТС Big Data

План:

- Вводная часть
- Роль линейной алгебры в DS
- Линейное пространство
- Скалярное произведение
- Векторные нормы
- Линейный оператор. Матрица
- Приложения

Вводная часть

О себе

Довелось поработать

- **в разных направлениях**: алготрейдинг, аналитика, A/B тестирование, classic ML
- **в разных компаниях**: WorldQuant, Yandex, X5 Retail Group. Чуть более 2 лет руковожу группой моделирования профиля клиента(ранее – экспериментальное направление) в МТС Big Data.

- **Сфера интересов**: прикладная статистика и A/B тестирование, механизмы принятия решений, теория игр, классический ML.
- Образование: ВМК МГУ (исследование операций)

Орг. моменты

План

- Линейная алгебра: 2 занятия
- Статистика: 6 занятий
- Оптимизация: 3 занятия
- (весной) A/B тесты: 2 занятия

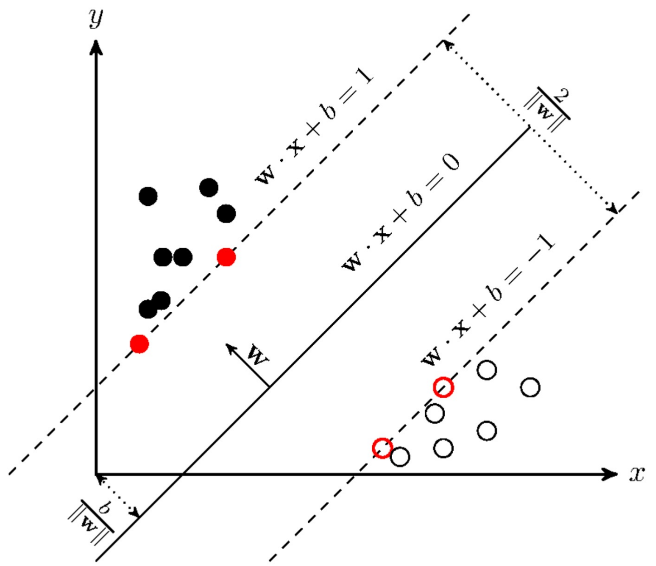
Орг. моменты

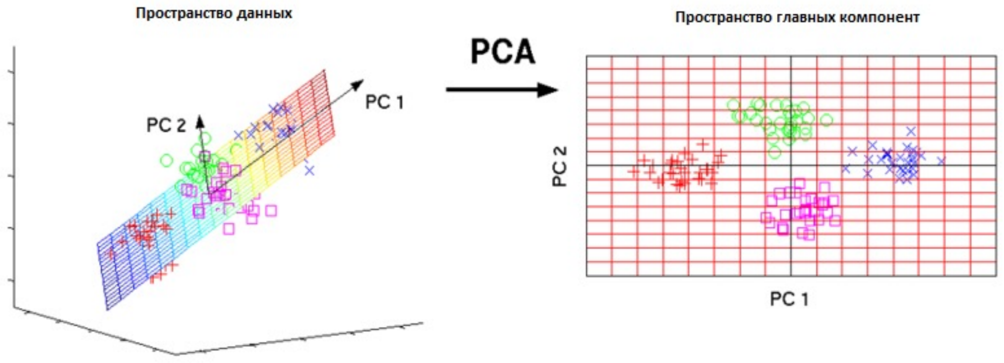
- После каждого занятия будет домашнее задание
- Крайний срок сдачи – 23:59 перед днём следующего занятия.
- В задачах, где ответ предполагает решение, нужно решение, иначе обнуление задачи.
- За каждую задачу оценка от 0 до 1, далее суммарная оценка переводится в 5-бальную систему.
- Для прохождения блока не менее 75% ДЗ должны быть сданы.

Роль линейной алгебры в DS

Роль линейной алгебры в анализе данных

- Любая работа с векторами и матрицами
- МНК в линейной регрессии
- Матричные разложения (SVD)
- Методы понижения размерности (PCA)

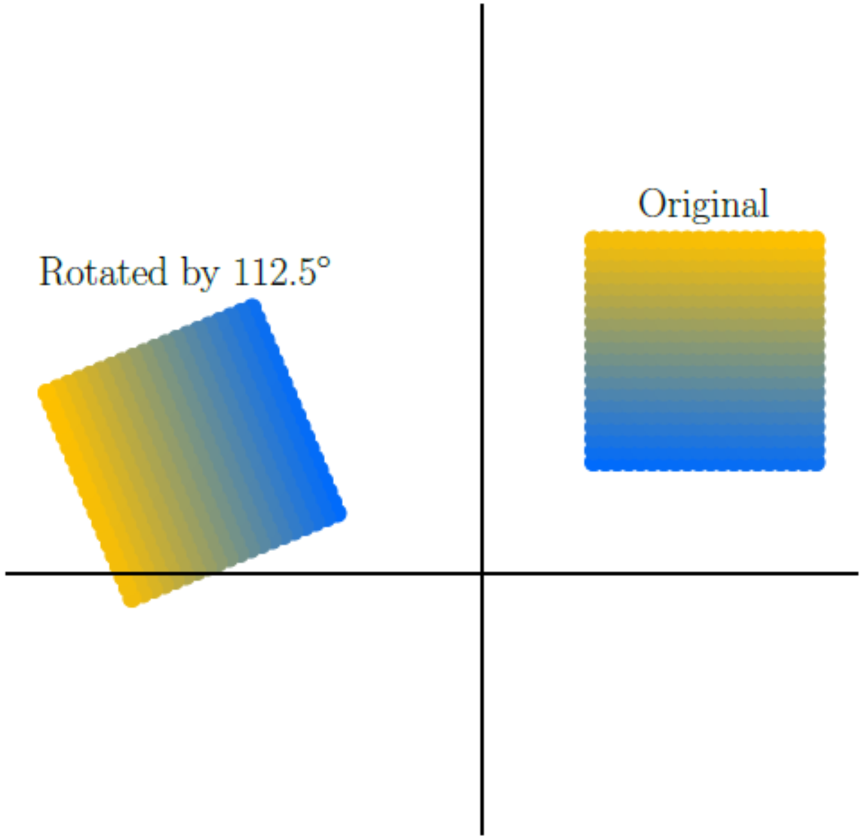




Матричный поворот

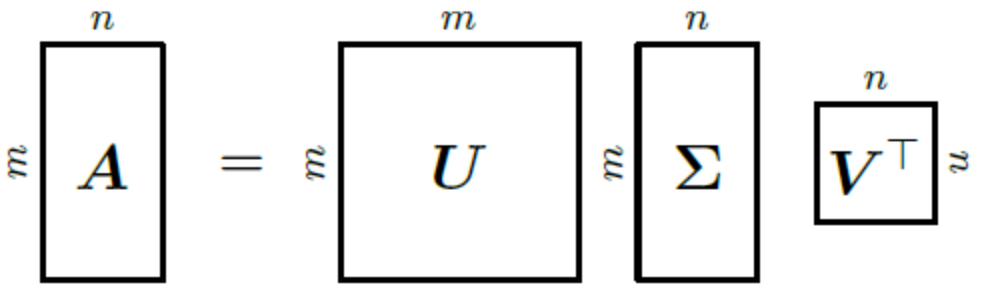
- Пусть у нас есть набор изображений для обучения нейронной сети. Дополним этот набор поворотами изображений.
- Поворот изображения – это частный случай умножения матрицы векторов координат на матрицу трансформации.
- Матрица трансформации для поворота координат выглядит так:

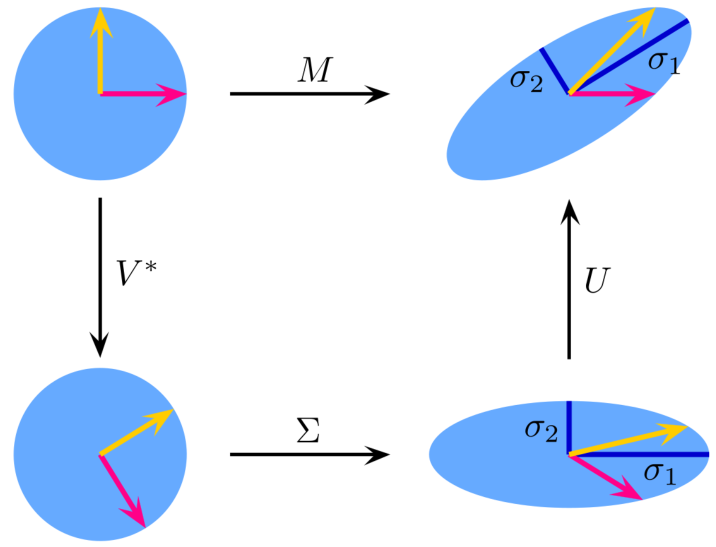
$$T_\theta = \begin{Vmatrix} cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{Vmatrix}$$



SVD разложение

- SVD разложение широко используется в рекомендательных системах.
- Например, если строки матрицы соответствуют читателям, столбцы – книгам, а сама матрица содержит оценки книг пользователями, то такое разложение матрицы выделит "типичных читателей" и "типичные книги".
- Каждого реального читателя и каждую реальную книгу можно представить линейной комбинацией "типичных", после чего мы сможем рассчитать ожидаемую оценку любой книги любым читателем.



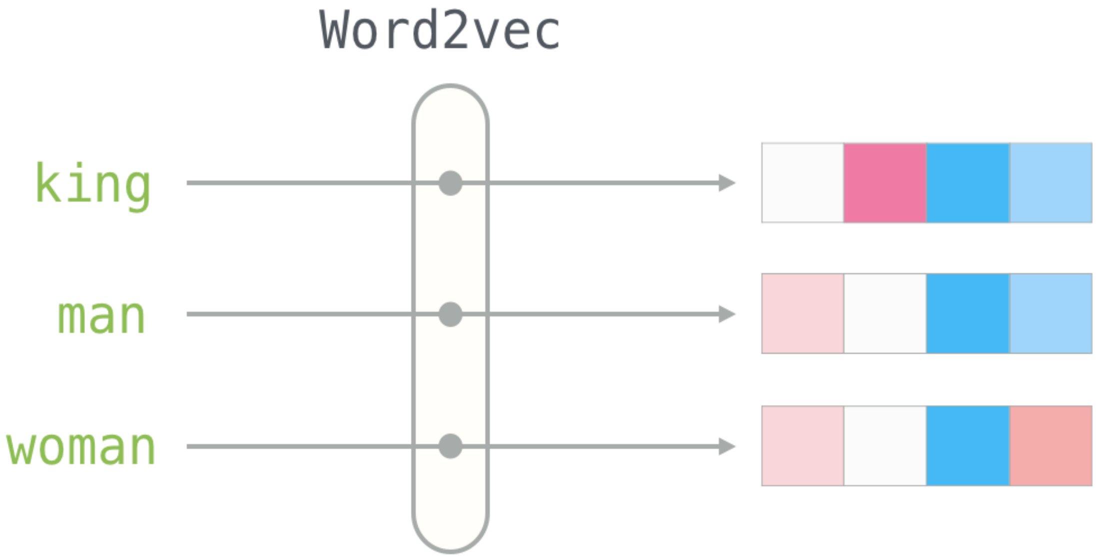


$$M = U \cdot \Sigma \cdot V^*$$

Линейное пространство

Линейное пространство

- Удобно представлять исследуемые объекты набором чисел.
- Пример: модель обработки текстов word2vec:



1. Vec('Paris') – Vec('France') + Vec('Italy') ~ Vec('Rome')

2. Vec('king') – Vec('man') + Vec('woman') ~ Vec('queen')

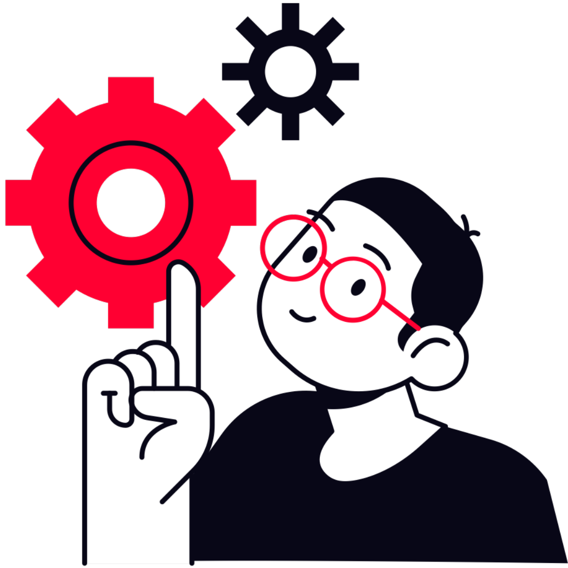
Линейное пространство (строгое определение)

Линейное пространство – совокупность (V, P, +, *), где:

1. V – непустое множество элементов (множество векторов)
2. P – поле, на элементы из которого можно умножать векторы
3. + – правило сложения векторов из V. Должны выполняться основные аксиомы сложения
4. * – правило умножения векторов из V на числа из P. Должны выполняться основные свойства умножения

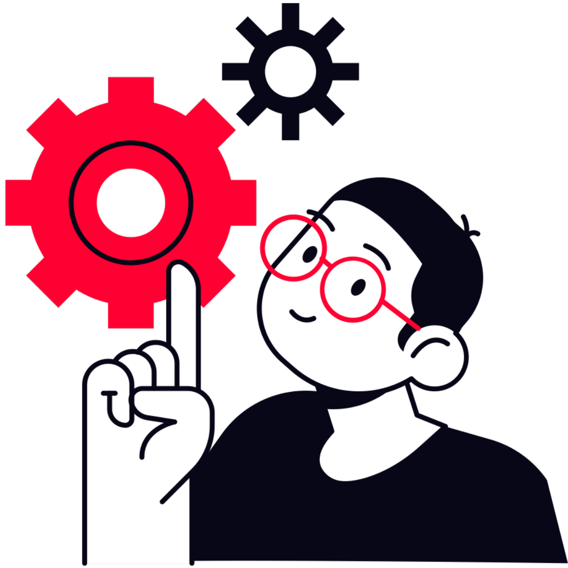
Линейное пространство
Поле. Свойства сложения

1. $a + b = b + a$ – коммутативность
2. $(a + b) + c = a + (b + c)$ – ассоциативность
3. Существует единственный элемент 0: $a + 0 = a$ (существование нейтрального элемент относительно сложения)
4. Существует –a: $a + (-a) = 0$ (существование обратного элемента относительно сложения)



Линейное пространство
Поле. Свойства умножения

1. a * b = b * a - коммутативность
2. (a * b) * c = a * (b * c) – ассоциативность
3. Существует единственный элемент 1: a * 1 = a (существование нейтрального элемента относительно умножения)
4. a != 0 => существует b: a * b = 1 (существование обратного элемента относительно умножения для ненулевого элемента)
5. a * (b + c) = a * b + a * c – дистрибутивность умножения относительно сложения



Линейное пространство
Описание

- Вектор – упорядоченный набор элементов поля
- Пример: $\mathbb{R}^3 = \{(x, y, z) : x, y, z \in \mathbb{R}\}$

Сложение:

(1, 2, 3) + (4, -5, 3) = (1 + 4, 2 - 5, 3 + 3) = (5, -3, 6)

(1, 2, 3) + (1, 4, -5, 3) = ???

Умножение:

(1, 2, 3) * 5 = (1 * 5, 2 * 5, 3 * 5) = (5, 10, 15)

Линейное пространство
Примеры

- $\mathbb{R}$, $\mathbb{C}$ - линейные пространства с векторами размерности 1.
- $\mathbb{R}^n$ - вещественное линейное пространство - линейное пространство векторов фиксированной длины (размерности), у которых все координаты - вещественные числа.
- $\mathbb{C}^n$ - комплексное линейное пространство - линейное пространство векторов фиксированной длины (размерности), у которых все координаты - комплексные числа.
- $\mathbb{P}^n$ - пространство многочленов степени не больше n - линейное пространство, в котором выражению a0 + a1 *t1 + · · · + an * tn соответствует вектор (a0, a1, . . . , an)

Линейное пространство. Линейная зависимость

- **Линейная комбинация** – сумма векторов, умноженных на числовые коэффициенты:

$$2 * (4, 1, 1) – 3 * (0, 1, 0) = (8, -1, 2)$$

Система векторов x1, … , xn называется

- **Линейно зависимой**, если существует такой набор чисел a1, …, an, не равных 0 одновременно: a1 * x1 + … + an * xn = 0
- **Линейно независимой**, если равенство 0 достигается ⇔ все коэффиценты равны 0.

Линейное пространство. Базис

- **Базис** – минимальная по включению система линейно независимых векторов линейного пространства, через которые выражаются все остальные векторы пространства.

  - Естественный базис в R3: e1 = (1, 0, 0), e2 = (0, 1, 0), e3= (0, 0, 1).
  - Число векторов в любом базисе одинаково и совпадает с размерностью пространства.
  - Любой вектор пространства единственным образом представим в виде линейной комбинации векторов базиса.
  - Координаты вектора – коэффициенты в его разложении по базису:
    - (2, -3, 7) = 2 * (1, 0, 0) – 3 * (0, 1, 0) + 7 * (0, 0, 1)

Линейное пространство. Проверка линейной зависимости

$$a_1 = [6,\ 12,\ 0],\ a_2 = [0,\ 17,\ 0],\ a_3 = [1,\ 0,\ 1]$$

Максимально зануляем координаты векторов, чтобы в каждом векторе появилась уникальная (не встречающаяся в других векторах) позиция с ненулевым элементом

$$\begin{bmatrix} 6 & 12 & 0 \\ 0 & 17 & 0 \\ 1 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Линейное пространство. Ранг системы векторов

Рангом системы векторов называется максимальное число её линейно независимых элементов

$$\begin{bmatrix} 6 & 12 & 6 \\ 0 & 17 & 0 \\ 1 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Третий вектор занулился, первые 2 линейно независимы, поэтому ранг системы равен 2.

Скалярное произведение

Скалярное произведение. Определение

Скалярным произведением называется правило, по которому любой паре элементов линейного пространства ставится в соответствие число со следующим набором свойств:

- Симметричность: $< a, b >=< b, a >$
- Линейность: $< \alpha a + \beta b, c >= \alpha < a, c > +\beta < b, c >$
- Неотрицательность: $< a, a >\geq 0, < a, a >= 0 \iff a = 0$

Скалярное произведение. Ортонормированный базис

Система (базис) линейного пространства называется **ортогональной**, если для любых различных её векторов их скалярное произведение равно 0 (и ортонормированной, если скалярное произведение любого вектора на себя равно 1).

Скалярное произведение в вещественном n-мерном пространстве (с ОНБ):

$$< a, b >= a_1 b_1 + \cdots + a_n b_n$$

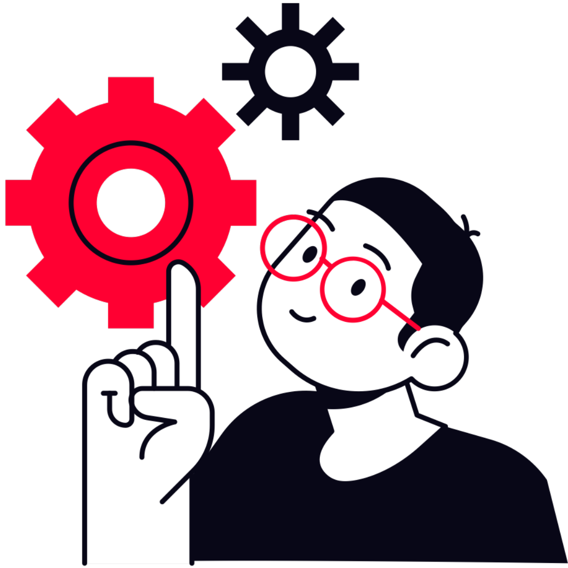
Скалярное произведение. Примеры

$$< (1, 4, 6), (3, -8, 5) >= 3 - 32 + 30 = 1$$

$$< (-15, 70, 0, 0), (0, 0, 1, 2) >= 0 + 0 + 0 + 0 = 0$$

$$< (1, 0, 0), (0, 1, 0) >= 0$$

$$< (1, 0, 0), (1, 0, 0) >= 1$$



Векторные нормы

Норма

Норма – обобщение длины (модуля) с множества чисел на линейное пространство.

$$||a|| \geq 0, \;\; ||a|| = 0 \iff a = 0$$

$$||\alpha a|| = |\alpha|||a||, \;\; \alpha \in \mathbb{C}$$

$$||a + b|| \leq ||a|| + ||b||$$

В пространстве со скалярным произведением норма вводится как корень из скалярного квадрата вектора (евклидова норма):

$$||a|| = \sqrt{< a, a >}$$

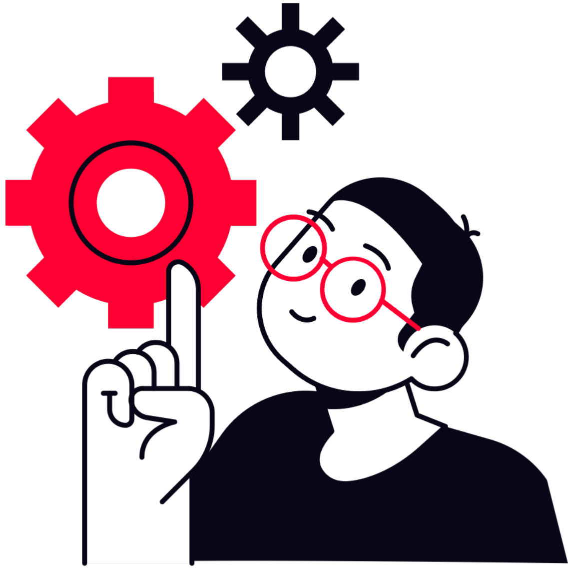
$$||(1, 2, 3, 4)|| = \sqrt{< (1, 2, 3, 4), (1, 2, 3, 4) >} = \sqrt{1 + 4 + 9 + 16} = \sqrt{30}$$

Норма

Расстояние между двумя векторами можно определить как норму разности между ними.

$$d(x, y) = \|x - y\|$$

Нормы могут быть разными!



Норма

Евклидова норма– это частный случай важного класса p-норм

$$\|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2} \qquad \|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}$$

2 важных частных случая p-норм:

- **Норма Чебышёва** $\|x\|_\infty = \max_i |x_i|$
- **Манхэттенское расстояние** $\|x\|_1 = \sum_i |x_i|$

Скалярное произведение. Угол между векторами

Косинусом угла в прямоугольном треугольнике называется отношение прилежащего катета к гипотенузе. Косинус равен 1, если угол равен 0, то есть векторы коллинеарны (направление совпадает) и 0, если они ортогональны (перпендикулярны друг другу).

$$cos(a, b) = \frac{(a, b)}{||a||\ ||b||}$$

$$cos((17, 6, 2), (-1, 2, 2)) = \frac{-17 + 12 + 4}{\sqrt{289 + 36 + 4}\sqrt{1 + 4 + 4}} = \frac{-1}{3\sqrt{329}}$$

Линейный оператор

Линейный оператор. Определение

Пусть задано отображение, которое любому вектору ставит в соответствие вектор, возможно, в другом пространстве, со следующими свойствами:

$$A(x + y) = Ax + Ay$$
$$A(\alpha x) = \alpha Ax$$

Такое отображение называется **линейным оператором**.

Линейный оператор. Примеры

- Тождественный (не производящий изменений) оператор
- Нулевой оператор

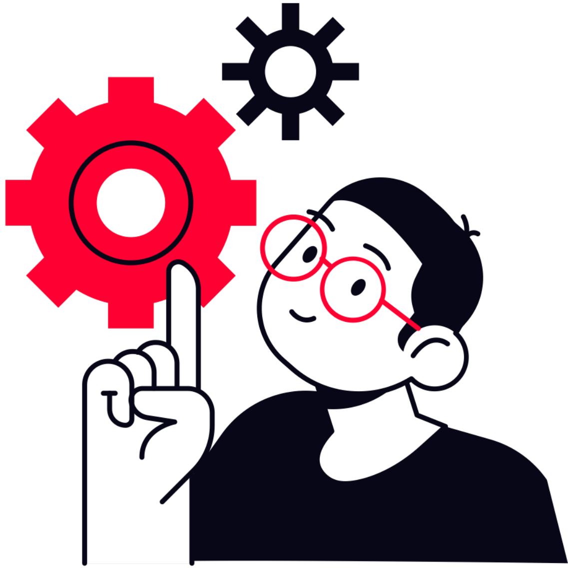
$$\Theta x = \theta$$

- Оператор проецирования

$$A \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

- Оператор дифференцирования

$$A \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix} = \{(1 + 2t + t^3)' = 3t^2 + 2\} = \begin{bmatrix} 2 \\ 0 \\ 3 \\ 0 \end{bmatrix}$$



Линейный оператор. Матрица

$$x = 2e_1 - 3e_2 + 7e_3$$

$$Ax = 2A\begin{bmatrix}1\\0\\0\end{bmatrix} - 3A\begin{bmatrix}0\\1\\0\end{bmatrix} + 7A\begin{bmatrix}0\\0\\1\end{bmatrix}$$

$$A\begin{bmatrix}1\\0\\0\end{bmatrix} = \begin{bmatrix}2\\1\\1\end{bmatrix}, A\begin{bmatrix}0\\1\\0\end{bmatrix} = \begin{bmatrix}-2\\0\\0\end{bmatrix}, A\begin{bmatrix}0\\0\\1\end{bmatrix} = \begin{bmatrix}0\\4\\0\end{bmatrix}$$

$$Ax = 2\begin{bmatrix}2\\1\\1\end{bmatrix} - 3\begin{bmatrix}-2\\0\\0\end{bmatrix} + 7\begin{bmatrix}0\\4\\0\end{bmatrix} = \begin{bmatrix}2*2-3*(-2)+7*0\\2*1-3*0+7*4\\2*1-3*0+7*0\end{bmatrix} = \begin{bmatrix}10\\30\\2\end{bmatrix}$$

Линейный оператор. Матрица

Если записать действие оператора A на базисные элементы по столбцам, то получим матрицу (набор элементов поля, упорядоченный по строкам и столбцам):

$$A = \begin{bmatrix} 2 & -2 & 0 \\ 1 & 0 & 4 \\ 1 & 0 & 0 \end{bmatrix}, \quad x = 2e_1 - 3e_2 + 7e_3 = \begin{bmatrix} 2 \\ -3 \\ 7 \end{bmatrix}$$

$$Ax = \begin{bmatrix} 2*2 - 3*(-2) + 7*0 \\ 2*1 - 3*0 + 7*4 \\ 2*1 - 3*0 + 7*0 \end{bmatrix} = \begin{bmatrix} <(2,-2,0), x> \\ <(1,0,4), x> \\ <(1,0,0), x> \end{bmatrix}$$

Между прямоугольными матрицами размера NxM и линейными операторами, действующими из N-мерного в M-мерное пространство существует взаимно-однозначное соответствие.

Матрица. Линейное пространство матриц

Матрицы можно рассматривать как многомерные векторы (векторы векторов) . Они так же образуют линейное пространство с поэлементным сложением при согласованности размерностей и поэлементным умножением на число:

$$17 * \begin{pmatrix} 2 & 1 \\ 3 & 5 \end{pmatrix} = \begin{pmatrix} 17 * 2 & 17 * 1 \\ 17 * 3 & 17 * 5 \end{pmatrix} = \begin{pmatrix} 34 & 17 \\ 51 & 85 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 1 \\ 3 & 5 \end{pmatrix} + \begin{pmatrix} -3 & 0 \\ 13 & 24 \end{pmatrix} = \begin{pmatrix} 2-3 & 1+0 \\ 3+13 & 5+24 \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 16 & 29 \end{pmatrix}$$

Матрица. Произведение матриц

Произведение матриц определено, если число столбцов 1-й матрицы совпадает с числом строк 2-й матрицы. При этом условии произведением является матрица, у которой на пересечении i-й строки и j-го столбца матрицы – результата будет находиться сумма

$$C = AB \Rightarrow c_{i,j} = \sum_{\ell=1}^{n} a_{i,\ell} b_{\ell,j}$$

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

$$= \begin{pmatrix} <(2,1,6),(2,3,0)> & <(2,1,6),(1,5,4)> \\ <(3,5,-5),(2,3,0)> & <(3,5,-5),(1,5,4)> \end{pmatrix}$$

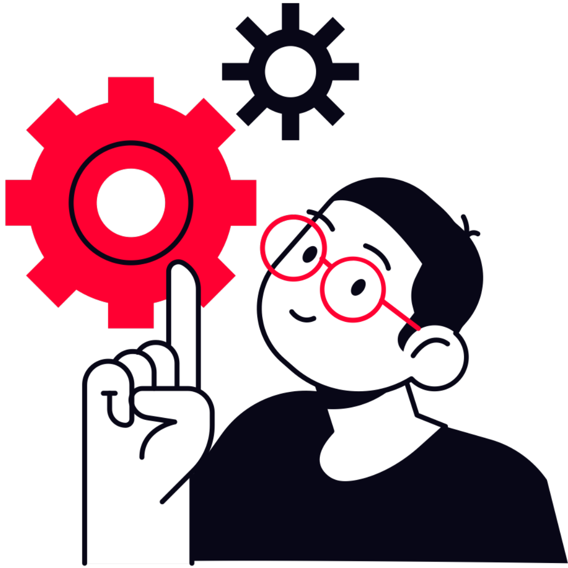
$$= \begin{pmatrix} 2*2+1*3+6*0 & 2*1+1*5+6*4 \\ 3*2+5*3-5*0 & 3*1+5*5-5*4 \end{pmatrix} = \begin{pmatrix} 13 & 30 \\ 16 & 8 \end{pmatrix}$$

Матрица. Транспонирование

Транспонированием матрицы называется операция, после которой строки матрицы становятся её столбцами и наоборот (с сохранением номеров), то есть для любых i, j:

$$\hat{a}_{i,j} = a_{j,i}$$

$$(AB)^T = B^T A^T$$



Приложения

Приложения

Найти стоимости товаров, если известна суммарная стоимость товаров и их количество в чеках

$$\begin{pmatrix} 2 & 1 \\ 3 & 5 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 10 \\ 16 \end{pmatrix}$$

$$A = \begin{pmatrix} 2 & 1 \\ 3 & 5 \end{pmatrix}, \; x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \; y = \begin{pmatrix} 10 \\ 16 \end{pmatrix}$$

$$Ax = y$$

Приложения

Вспомогательную роль линейная алгебра играет во многих областях:

- Эконометрика (метод наименьших квадратов)
- Машинное обучение (классический ML, нейронные сети)
- Методы оптимизации (задачи линейного программирования)
- И другие…

Что мы узнали

- Что из себя представляет линейное пространство
- Как работать с векторами и матрицами
- Рассмотрели скалярное произведение и нормы векторов
- Разобрали случаи линейной зависимости векторов
- Затронули понятие ранга системы векторов (далее ранг матрицы)

Полезные ссылки

- https://www.cis.upenn.edu/~jean/math-deep.pdf
- https://mml-book.github.io/
- https://youtube.com/playlist?list=PLZHQObOWTQDPD3MizzM2xVFitgF8hE_ab&si=PCfYtyqnveHtF2My