Введение в анализ данных

Лекция 3

Линейная алгебра и анализ данных

Евгений Соколов

sokolov.evg@gmail.com

НИУ ВШЭ, 2016

Организационное

- Рассылка: hse-minor-datamining-2+subscribe@googlegroups.com
  - (написать пустое письмо)
- Почему у ИАД-6 и ИАД-8 по два семинара?
  - Потому что раз в две недели
- Скоро на всех экранах: https://www.coursera.org/specializations/machine-learning-data-analysis

Напоминание

- $\mathbb{X}$ — пространство объектов, $\mathbb{Y}$ — пространство ответов
- $x = (x^1, \dots, x^d)$ — признаковое описание
- $X = (x_i, y_i)_{i=1}^{\ell}$ — обучающая выборка
- $a(x)$ — алгоритм, модель
- $Q(a, X)$ — функционал качества алгоритма a на выборке X
- Обучение: $a(x) = \arg\max_{a \in \mathcal{A}} Q(a, X)$

Напоминание

- $\mathbb{X}$ — пространство объектов, $\mathbb{Y}$ — пространство ответов
- $x = (x^1, \ldots, x^d)$ — признаковое описание
- $X = (x_i, y_i)_{i=1}^{\ell}$ — обучающая выборка
- $a(x)$ — алгоритм, модель
- $Q(a, X)$ — функционал качества алгоритма a на выборке X
- Обучение: $a(x) = \arg\max_{a \in \mathcal{A}} Q(a, X)$

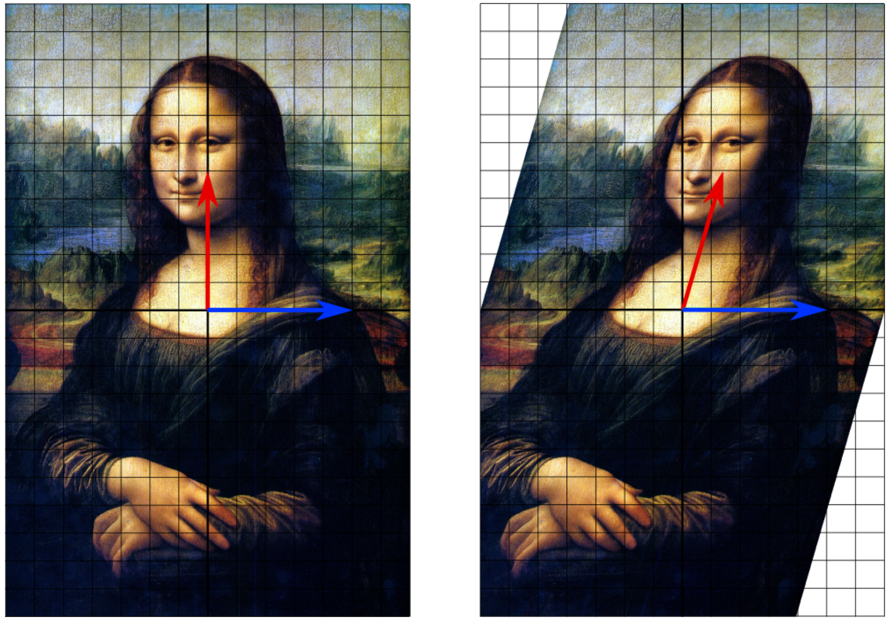
- (могло прийти в голову кому угодно 5 веков назад)

Неудобные вопросы

- Как удобно записать обучающую выборку? Какими особыми свойствами она может обладать?
- Каким должен быть $Q(a, X)$?
- Как найти $\arg\max_{a \in \mathcal{A}} Q(a, X)$?
- Как описать неопределенность в данных?
- Может ли алгоритм оценить истинную зависимость между признаками и ответом?

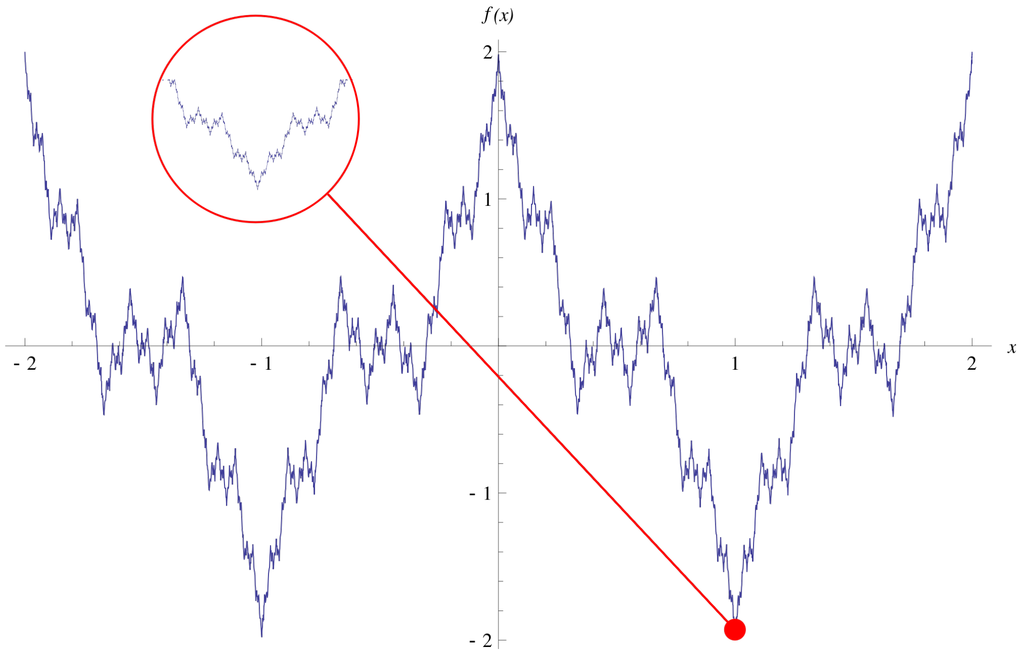
Линейная алгебра

- Обучающая выборка описывается **матрицей**
- Нет ли в ней избыточной информации? — линейная зависимость
- Системы линейных уравнений
- Линейные преобразования
- Матричные разложения



Математический анализ

- Как найти минимум функции? — производная и градиент
- А если функция сложная? — численная оптимизация
- Единственный ли минимум? — выпуклые функции

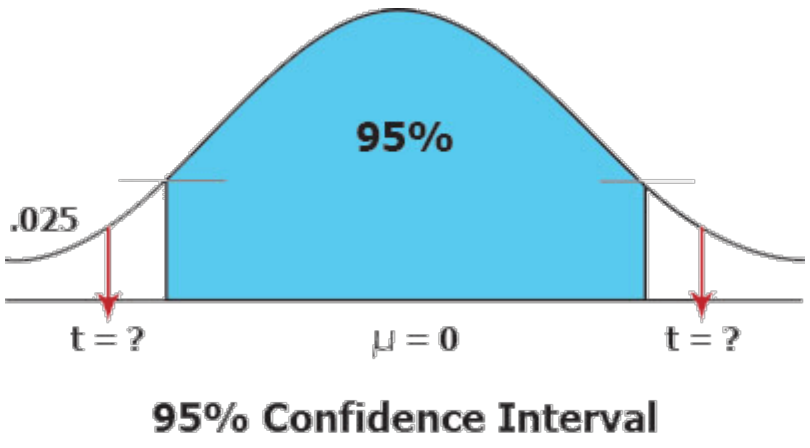


Теория вероятностей

- Как формально описать случайность?
- Какое распределение может быть у ответов или признаков?
- Как задавать вероятностные модели данных?
- Как описывать распределения? — среднее, медиана, моменты
- Как проверить связь двух случайных величин? — корреляция
- Что особенного в сумме большого числа факторов? — центральная предельная теорема

Математическая статистика

- Как оценить средний доход по небольшой выборке?
- Как оценить качество алгоритма по небольшой выборке?
- Как настроить алгоритм, чтобы он максимально точно описывал данные? — метод максимального правдоподобия
- Какова вероятность, что документ про математику, если в нем есть слово «градиент»?



Числа

Какие бывают числа?

- Натуральные $\mathbb{N}$: 1, 2, 3, …
- Целые $\mathbb{Z}$: 0, 1, -1, 2, -2, 3, -3, …
- Рациональные $\mathbb{Q}$: $\frac{m}{n}$, где $m \in \mathbb{Z}, n \in \mathbb{N}$
- И всё?

Какие бывают числа?

- Не все числа можно представить дробью
- L — длина окружности, D — диаметр окружности
- Уравнение: $L = xD$
- Чему равен x?

Какие бывают числа?

- Не все числа можно представить дробью
- L — длина окружности, D — диаметр окружности
- Уравнение: $L = xD$
- Чему равен x?
- $x = \pi$
- π — нельзя представить в виде дроби (иррациональное число)

Какие бывают числа?

- Вещественные числа $\mathbb{R}$: десятичные дроби (в том числе бесконечные)
- $\pi = 3.14159265359\ \ldots$
- Теперь всё?

Какие бывают числа?

- Уравнение: $x^2 = -1$
- Есть ли решение?

Какие бывают числа?

- Уравнение: $x^2 = -1$
- Есть ли решение?
- Среди вещественных чисел — нет
- Давайте дополним их решением уравнения!
- Мнимая единица: $i = \sqrt{-1}$
- Комплексные числа $\mathbb{C}$: $a + bi$, где $a \in \mathbb{R}$, $b \in \mathbb{R}$

Какие бывают числа?

- Может теперь всё?

Какие бывают числа?

- Может теперь всё?
- Да!
- Основная теорема алгебры: любой многочлен имеет хотя бы один комплексный корень

Векторы и матрицы

Вектор

- $x = (x^1, \dots, x^d)$ — признаковое описание
- $x^1, \dots, x^d$ — вещественные числа
- x — набор из d чисел — **вектор**

Вектор

- 5 — число
- (5, 3) — точка на плоскости
- (5, 3, 9) — точка в пространстве
- (5, 3, 9, 1) — точка в четырехмерном пространстве
- …

Векторное пространство

- Что мы будем называть вектором?
- Векторное пространство V — множество, состоящее из векторов
- Например, V — все наборы из d вещественных чисел

Векторное пространство

- Какие операции над векторами нам нужны?
- Простейшие: сложение и умножение на число
- Как их ввести? Какими свойствами они должны обладать?

Аксиомы

1. $\mathbf{x}+\mathbf{y}=\mathbf{y}+\mathbf{x}$, для любых $\mathbf{x},\mathbf{y}\in V$ (*коммутативность сложения*);
2. $\mathbf{x}+(\mathbf{y}+\mathbf{z})=(\mathbf{x}+\mathbf{y})+\mathbf{z}$, для любых $\mathbf{x},\mathbf{y},\mathbf{z}\in V$ (*ассоциативность сложения*);
3. существует такой элемент $\mathbf{0}\in V$, что $\mathbf{x}+\mathbf{0}=\mathbf{x}$ для любого $\mathbf{x}\in V$ (*существование нейтрального элемента относительно сложения*), называемый **нулевым вектором** или просто **нулём** пространства V;
4. для любого $\mathbf{x}\in V$ существует такой элемент $-\mathbf{x}\in V$, что $\mathbf{x}+(-\mathbf{x})=\mathbf{0}$, называемый вектором, **противоположным** вектору $\mathbf{x}$;
5. $\alpha(\beta\mathbf{x})=(\alpha\beta)\mathbf{x}$ (*ассоциативность умножения на скаляр*);
6. $1\cdot\mathbf{x}=\mathbf{x}$ (*унитарность: умножение на нейтральный (по умножению) элемент поля F сохраняет вектор*).
7. $(\alpha+\beta)\mathbf{x}=\alpha\mathbf{x}+\beta\mathbf{x}$ (*дистрибутивность умножения на вектор относительно сложения скаляров*);
8. $\alpha(\mathbf{x}+\mathbf{y})=\alpha\mathbf{x}+\alpha\mathbf{y}$(*дистрибутивность умножения на скаляр относительно сложения векторов*).

Векторное пространство

Множество V называется векторным пространством, если:

- На нем заданы операции $+$ (сложение) и $*$ (умножение на число)
- Оно замкнуто относительно этих операций
- Для этих операций выполнены 8 аксиом

Евклидово пространство

- Пространство наборов из d вещественных чисел — евклидово пространство $\mathbb{R}^d$
- Бывают пространства с более сложными элементами: многочленами, уравнениями, функциями

Евклидово пространство

- Сложение и умножение на число — покоординатно
- Сложение
  - $a = (a_1, \ldots, a_d)$
  - $b = (b_1, \ldots, b_d)$
  - $a + b = (a_1 + b_1, \ldots, a_d + b_d)$
- Умножение на число:
  - $a = (a_1, \ldots, a_d)$
  - $\beta \in \mathbb{R}$
  - $\beta a = (\beta a_1, \ldots, \beta a_d)$

Матрицы

- Вектор описывает один объект
- А если объектов несколько?

Матрицы

36,18	2	1	2	3	59090	1
46,47671233	0	1	4	3	14773	1
45,13424658	0	1	3	3	19376	2
25,88767123	1	1	4	3	16098	0
25,70410959	1	1	3	3	20338	0
33,03	0	1	3	1	501667	2
46,44931507	3	1	2	1	26100	0
51,24383562	0	0	4	2	20727	0
46,8739726	0	1	1	3	27861	0
39,8630137	3	1	4	2	33495	1
37,09	0	1	4	3	55825	1
38,14	3	1	3	2	60000	1
45,46849315	3	1	1	1	40000	1
42,99726027	3	1	4	2	40343	0
29,98082192	3	0	4	2	27583,78	2
46,20547945	3	1	2	2	45385	1

Объект

Признак

Матрицы

- Матрица — таблица с числами
- Пример:

$$A = \begin{pmatrix} 1 & 2 & 5 & 1 \\ 5 & 3 & 9 & 0 \\ 0 & 7 & 1 & 4 \end{pmatrix}$$

- Два индекса: строка и столбец
- $a_{11} = 1$
- $a_{23} = 9$

Матрицы

- Матрица — таблица с числами
- Пример:

$$A = \begin{pmatrix} 1 & 2 & 5 & 1 \\ 5 & 3 & 9 & 0 \\ 0 & 7 & 1 & 4 \end{pmatrix}$$

- Два индекса: строка и столбец
- $a_{11} = 1$
- $a_{23} = 9$
- Пространство матриц 3 на 4: $\mathbb{R}^{3\times 4}$(тоже векторное!)

Матрицы

- Выборка объектов описывается матрицей «объекты-признаки»
- По строкам — объекты
- По столбцам — признаки

$$X = \begin{pmatrix} 1 & 1000 & 5 & 3 & 4 \\ 9 & 9000 & 10 & 5 & 7.5 \\ 5 & 5000 & 1 & 3 & 2 \end{pmatrix}$$

Линейная независимость

Объекты-признаки

$$X = \begin{pmatrix} 1 & 1000 & 5 & 3 & 4 \\ 9 & 9000 & 10 & 5 & 7.5 \\ 5 & 5000 & 1 & 3 & 2 \end{pmatrix}$$

- Задача предсказания прибыли магазина в следующем месяце
- Рассмотрим в качестве векторов столбцы матрицы (признаки)

Подозрительные зависимости

$$X = \begin{pmatrix} 1 & 1000 & 5 & 3 & 4 \\ 9 & 9000 & 10 & 5 & 7.5 \\ 5 & 5000 & 1 & 3 & 2 \end{pmatrix}$$

- Первый и второй признаки: $x_2 = 1000x_1$
- Первый — общий вес товаров в тоннах, второй — в килограммах

Подозрительные зависимости

$$X = \begin{pmatrix} 1 & 1000 & 5 & 3 & 4 \\ 9 & 9000 & 10 & 5 & 7.5 \\ 5 & 5000 & 1 & 3 & 2 \end{pmatrix}$$

- $x_5 = 0.5x_3 + 0.5x_4$
- Пятый — средняя прибыль за последние два месяца
- Третий и четвертый — прибыль в прошлом и позапрошлом месяце

Линейная зависимость

— один из векторов равен сумме с весами остальных векторов

- Это плохо:
  - Избыточная информация
  - Лишние затраты на хранение данных
  - Вредит некоторым методам машинного обучения

Линейная зависимость

- Пусть дан набор векторов $x_1, \dots, x_n$
- Они линейно зависимы, если
  - существуют такие числа $\beta_1, \dots, \beta_n$,
  - хотя бы одно из которых не равно нулю,
  - что сумма векторов с такими коэффициентами равна нулю

$$\beta_1 x_1 + \cdots + \beta_n x_n = 0$$

Линейная зависимость

- Пусть дан набор векторов $x_1, \ldots, x_n$
- Они линейно зависимы, если
  - существуют такие числа $\beta_1, \ldots, \beta_n$,
  - хотя бы одно из которых не равно нулю,
  - что сумма векторов с такими коэффициентами равна нулю

$$\beta_1 x_1 + \cdots + \beta_n x_n = 0$$

Вектор из всех нулей

Линейная зависимость

- Аналогично: линейно зависимы, если один из векторов линейно выражается через остальные
$$x_i = \beta_1 x_1 + \ldots + \beta_{i-1} x_{i-1} + \beta_{i+1} x_{i+1} + \cdots + \beta_n x_n$$

Линейная независимость

- Пусть дан набор векторов $x_1, \dots, x_n$
- Они линейно независимы, если уравнение

$$\beta_1 x_1 + \cdots + \beta_n x_n = 0$$

выполнено только при $\beta_1 = \beta_2 = \cdots = \beta_n = 0$

Размерность векторного пространства

- Векторное пространство V
- Размерность — максимальное число линейно независимых векторов в нем
- Обозначение: $\dim \mathrm{V}$

Размерность векторного пространства

- Для евклидова пространства: $\dim \mathbb{R}^d = d$
- Пример системы линейно независимых векторов:

$$e_1 = (1, 0, \dots, 0)$$
$$e_2 = (0, 1, \dots, 0)$$
$$\dots$$
$$e_d = (0, 0, \dots, 1)$$

- Любой вектор из $\mathbb{R}^d$ можно записать как

$$x = a_1 e_1 + \cdots + a_d e_d$$

Пример

- $x_1 = (2, 1, 2)$
- $x_2 = (4, 2, 4)$
- $x_3 = (1, 2, 3)$
- $2x_1 - x_2 + 0x_3 = 0$
- Линейно зависимы

Пример

- $x_1 = (2, 1, 2)$
- $x_2 = (5, 2, 4)$
- $x_3 = (1, 2, 3)$
- Линейно независимы

Как это можно было понять?

- Вариант 1 — метод Гаусса
- Вариант 2 — вычислить ранг

Ранг матрицы

- Ранг системы строк матрицы A — максимальное число линейно независимых строк среди них
- Ранг системы столбцов матрицы A — максимальное число линейно независимых столбцов среди них
- **Важный результат:** ранг по столбцам и по строкам всегда совпадает
- Обозначение: $\mathrm{rg}\, A$

Ранг матрицы

- Характеристика количества информации в матрице

- $A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 0.5 & 1 & 0.5 \end{pmatrix}$

- $\mathrm{rg}\, A = 1$

- Можно заменить на один столбец: $A = \begin{pmatrix} 2 \\ 4 \\ 1 \end{pmatrix}$

Ранг матрицы

- Задача кредитного скоринга
- 100,000 клиентов, 1000 признаков
- Нужно хранить 10^8 чисел!
- Вычисляем ранг: $\mathrm{rg}\, A = 10$
- Очень много линейно зависимых столбцов: избыточность в базах данных, таблицы-дубликаты и т.д.
- Достаточно хранить 10^6 чисел
- Уменьшили необходимый объем памяти в 100 раз!

Линейные преобразования

Векторы и матрицы

- Вектор размера d — тоже матрица
- Вектор-строка: $w = (w_1, \dots, w_d) \in \mathbb{R}^{1 \times d}$
- Вектор-столбец: $w = \begin{pmatrix} w_1 \\ \cdots \\ w_d \end{pmatrix} \in \mathbb{R}^{d \times 1}$

Линейная модель

- $a(x) = w_1 x^1 + \dots + w_d x^d$
- Как применить модель к целой выборке?

$$\begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1d} \\ x_{21} & x_{22} & \cdots & x_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ x_{\ell 1} & x_{\ell 2} & \cdots & x_{\ell d} \end{pmatrix}$$

$$\begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_d \end{pmatrix}$$

$$\begin{pmatrix} \sum_{i=1}^{d} w_i x_{1i} \\ \sum_{i=1}^{d} w_i x_{2i} \\ \vdots \\ \sum_{i=1}^{d} w_i x_{\ell i} \end{pmatrix}$$

Умножение

- Мы еще не вводили умножение матрицы на вектор
- Определим его именно так
- Только для матрицы $\ell \times d$ и вектора $d \times 1$

$$Xw = \begin{pmatrix} \sum_{i=1}^{d} w_i x_{1i} \\ \sum_{i=1}^{d} w_i x_{2i} \\ \vdots \\ \sum_{i=1}^{d} w_i x_{\ell i} \end{pmatrix}$$

Линейные преобразования

- Умножая матрицу $A \in \mathbb{R}^{m\times n}$ на вектор $x \in \mathbb{R}^{n\times 1}$, получаем вектор $z \in \mathbb{R}^{m\times 1}$
- Матрица задает функцию из $\mathbb{R}^{n\times 1}$ в $\mathbb{R}^{m\times 1}$
- Эта функция — линейная:
  - $A(x_1 + x_2) = Ax_1 + Ax_2$
  - $A(\alpha x) = \alpha Ax$
- Любая линейная функция описывается некоторой матрицей

Линейные преобразования

- Функции можно применять последовательно: $g(f(x))$
- В том числе линейные: $A(Bx)$
- Композиция линейных функций — тоже линейная функция:
  - $A\big(B(x_1 + x_2)\big) = A(Bx_1) + A(Bx_2)$
  - $A\big(B(\alpha x)\big) = \alpha A(Bx)$
- А какая у нее матрица?
- Зададим матричное умножение так, чтобы оно соответствовало композиции линейных преобразований

Матричное умножение

- Только для матриц $A \in \mathbb{R}^{m \times k}$ и $A \in \mathbb{R}^{k \times n}$
- Результат: $AB = C \in \mathbb{R}^{m \times n}$
- Правило:

$$c_{ij} = \sum_{p=1}^{k} a_{ip} b_{pj}$$

Пример

$$\begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 10 & 0 \end{pmatrix} * \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} \quad & \quad & \quad \\ & & \\ & & \end{pmatrix}$$

Пример

$$\begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 10 & 0 \end{pmatrix} * \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} \boxed{1} & & \\ & & \\ & & \end{pmatrix}$$

Пример

$$\begin{pmatrix} \boxed{1 \quad 2} \\ 0 \quad 1 \\ 10 \quad 0 \end{pmatrix} * \begin{pmatrix} 1 & \boxed{0} & 0 \\ 0 & \boxed{0} & 2 \end{pmatrix} = \begin{pmatrix} 1 & \boxed{0} & \\ & & \\ & & \end{pmatrix}$$

Пример

$$\begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 10 & 0 \end{pmatrix} * \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & \boxed{4} \\ & & \\ & & \end{pmatrix}$$

Пример

$$\begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 10 & 0 \end{pmatrix} * \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 4 \\ \boxed{0} & & \\ & & \end{pmatrix}$$

Сложение

- Можем складывать только матрицы одинакового размера
- Поэлементно:
- $C = AB$
- $c_{ij} = a_{ij} + b_{ij}$

Умножение на число

- Поэлементно:
- $C = \alpha B$
- $c_{ij} = \alpha b_{ij}$

Транспонирование

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

Транспонирование

- $A = \begin{pmatrix} 1 & 4 & 1 \\ 2 & 3 & 3 \\ 2 & 1 & 0 \\ 5 & 1 & 1 \end{pmatrix}$
- Главная диагональ — a_{ii}

Транспонирование

- Транспонирование — поворот относительно главной диагонали
- A размера $m \times n$ превращается в A^T размера $n \times m$
- $a_{ij}^T = a_{ji}$
- Строки становятся столбцами

Транспонирование

$$A = \begin{pmatrix} 1 & 4 & 1 \\ 2 & 3 & 3 \\ 2 & 1 & 0 \\ 5 & 1 & 1 \end{pmatrix} \Longrightarrow A^T = \begin{pmatrix} 1 & 2 & 2 & 5 \\ 4 & 3 & 1 & 1 \\ 1 & 3 & 0 & 1 \end{pmatrix}$$

Системы линейных уравнений

Система линейных уравнений

- Линейная модель должна давать верные ответы на обучении

$$\begin{cases} w_1 x_{11} + w_2 x_{12} + \dots + w_d x_{1d} = y_1 \\ \dots \\ w_1 x_{\ell 1} + w_2 x_{\ell 2} + \dots + w_d x_{\ell d} = y_\ell \end{cases}$$

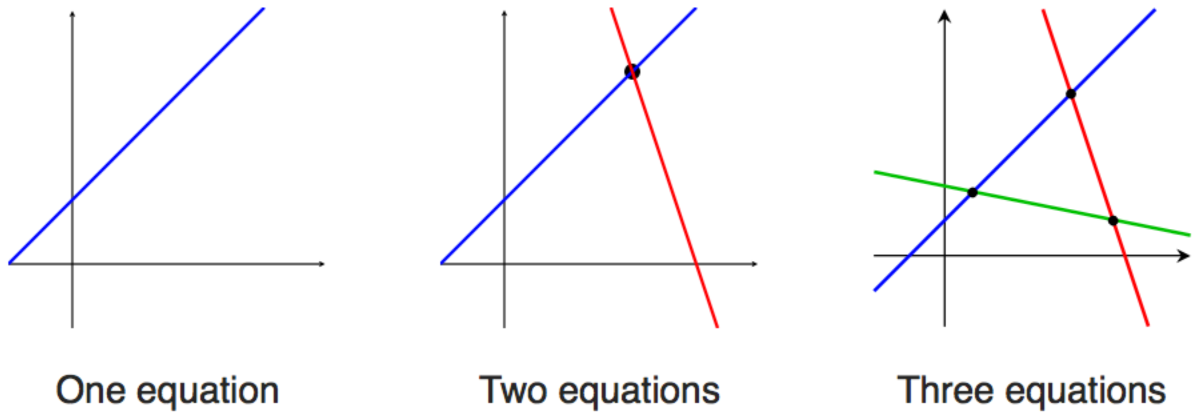
Система линейных уравнений

- В матричном виде: $Xw = y$

Система линейных уравнений

- Три уравнения:

$$y = x + 1$$
$$y = -3x + 9$$
$$y = -1x + 2$$



Три случая

- Бесконечно много решений
- Одно решение
- Ни одного решения

Как определить тип?

- Сравниваем ранг матрицы X и ранг $(X \mid y)$

$$\begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1d} \\ x_{21} & x_{22} & \cdots & x_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ x_{\ell 1} & x_{\ell 2} & \cdots & x_{\ell d} \end{pmatrix} \qquad \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1d} & y_1 \\ x_{21} & x_{22} & \cdots & x_{2d} & y_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_{\ell 1} & x_{\ell 2} & \cdots & x_{\ell d} & y_\ell \end{pmatrix}$$

- Если равны, то правая часть выражается через признаки
- Если ранги равны, то есть хотя бы одно решение
- Если ранг второй матрицы больше, то решений нет

Сколько решений у системы?

- Нужно сравнить ранг X и число неизвестных (т.е. d)
- Если $\mathrm{rg}\, X = d$, то решение одно
- Если $\mathrm{rg}\, X < d$, то решений бесконечно много (уравнений меньше, чем неизвестных)

- Пример: $x_1 + 5x_2 = 10$
- Решение: $x_1 = 10 - 5x_2$ для любого $x_2 \in \mathbb{R}$

Как решать?

- Численное решение систем линейных уравнений — на следующей лекции

Обратная матрица

- A^{-1}— обратная к A
- $AA^{-1} = A^{-1}A = I$
- I — единичная матрица
- Только для квадратных матриц

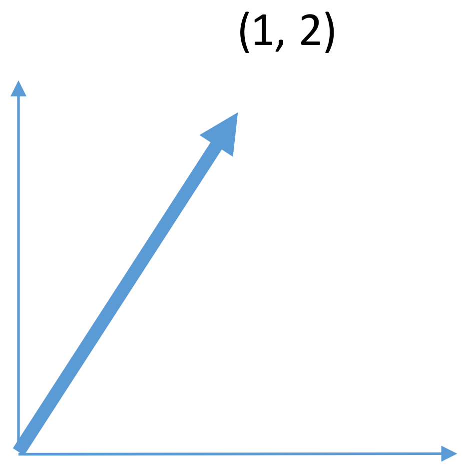
- Существует тогда и только тогда, когда $\det A \neq 0$
- Можно найти с помощью SciPy

Квадратные матрицы

- $Ax = b$
- Решение: $x = A^{-1}b$
- Решение единственно тогда и только тогда, когда существует A^{-1}

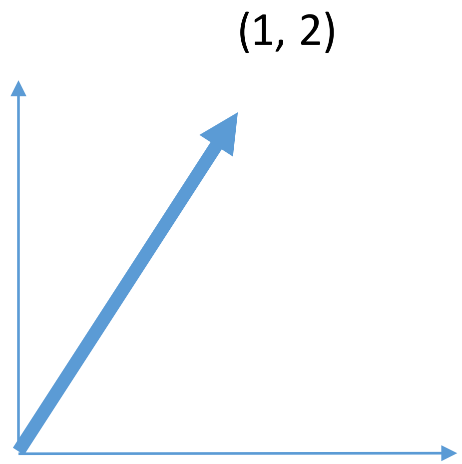
Операции в векторных пространствах

- Вектор — точка и стрелка, идущая к ней из нуля

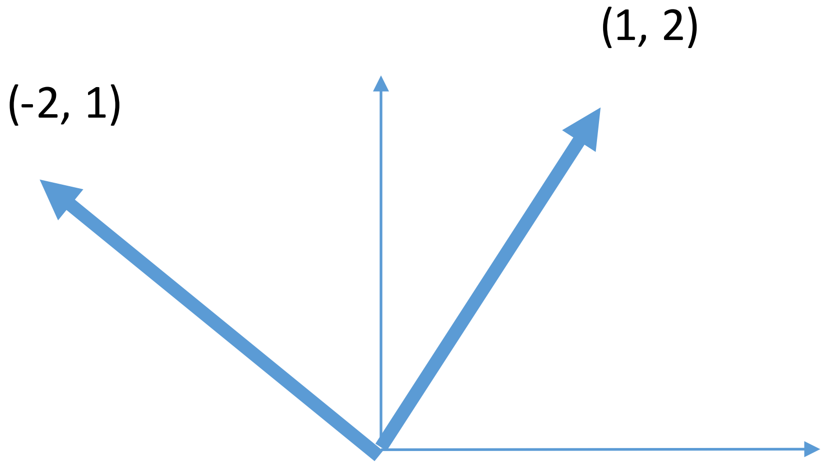


- Длина вектора:

$$\sqrt{1^2 + 2^2} = \sqrt{5}$$

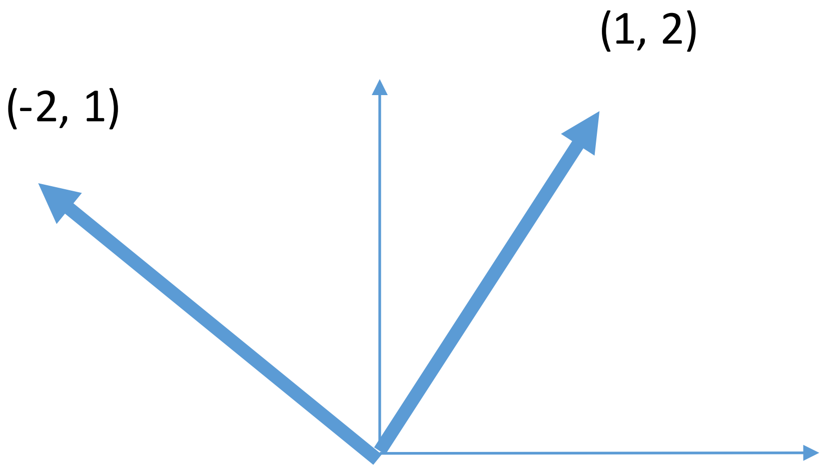


- Можем измерить угол с помощью транспортира: 90 градусов



- Можем измерить расстояние между точками:

$$\sqrt{\left(1-(-2)\right)^2+(2-1)^2}=\sqrt{10}$$



Норма

- Обобщение понятия длины вектора
- Функция $\|x\|$ от вектора
- Если $\|x\| = 0$, то $x = 0$
- $\|x + y\| \leq \|x\| + \|y\|$
- $\|\alpha x\| = |\alpha| \, \|x\|$

- Векторное пространство с нормой — нормированное

Примеры норм

- Евклидова норма:

$$\|x\|_2 = \sqrt{\sum_{i=1}^{d} x_i^2}$$

- Манхэттенская норма:

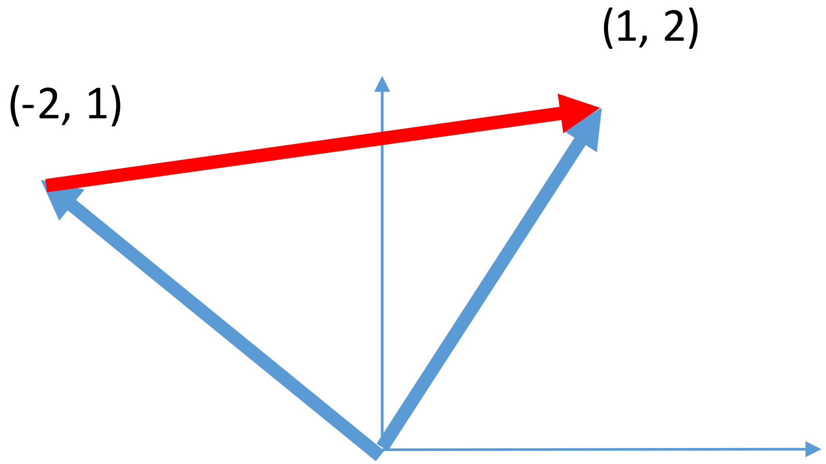
$$\|x\|_1 = \sum_{i=1}^{d} |x_i|$$

Метрика

- Обобщение понятия расстояния
- $\rho(x, y) = \|x - y\|$

- Соответствует геометрическим представлениям

- Векторное пространство с метрикой — метрическое



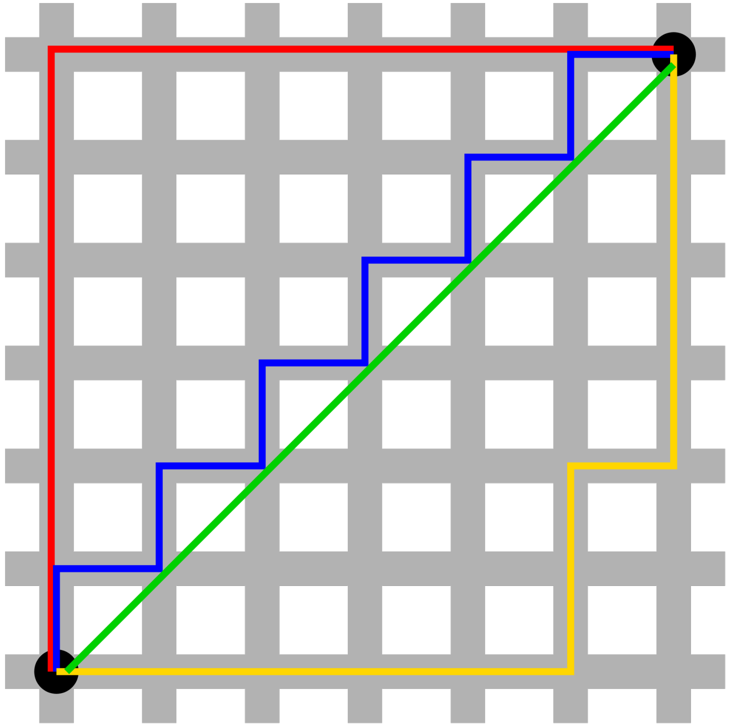
Примеры метрик

- Евклидова метрика:

$$\rho_2(x,z) = \sqrt{\sum_{i=1}^{d} (x_i - z_i)^2}$$

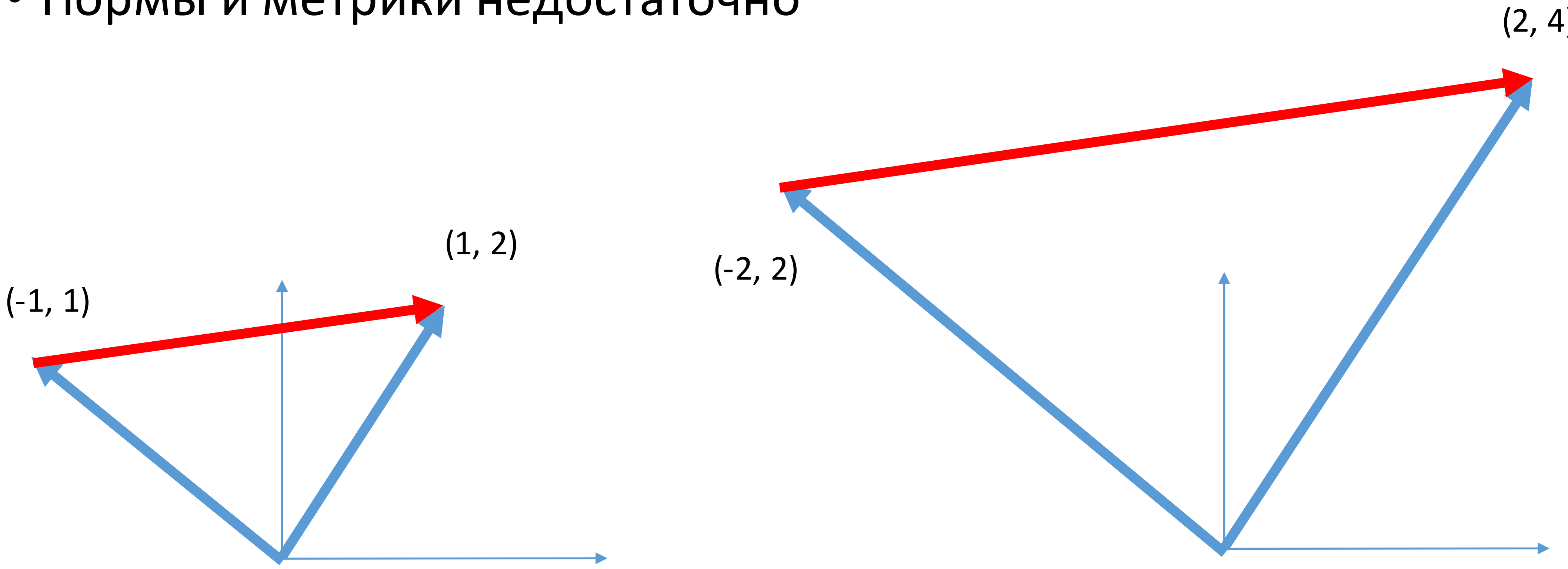
- Манхэттенская метрика:

$$\rho_1(x,z) = \sum_{i=1}^{d} |x_i - z_i|$$



Как искать углы?

- Нормы и метрики недостаточно



Скалярное произведение

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i$$

Скалярное произведение

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i$$

- Норма: $\|x\|_2 = \sqrt{\langle x, x \rangle}$

Скалярное произведение

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i$$

- Норма: $\|x\|_2 = \sqrt{\langle x, x \rangle}$
- Расстояние: $\rho_2(x, z) = \|x - z\| = \sqrt{\langle x - z, x - z \rangle}$

Скалярное произведение

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i$$

- Норма: $\|x\|_2 = \sqrt{\langle x, x \rangle}$
- Расстояние: $\rho_2(x, z) = \|x - z\| = \sqrt{\langle x - z, x - z \rangle}$
- Угол?

Скалярное произведение

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i$$

- Важное соотношение: $\langle x, y \rangle = \|x\| \|y\| \cos \angle(x, y)$

Скалярное произведение

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i$$

- Важное соотношение: $\langle x, y \rangle = \|x\| \|y\| \boxed{\cos \angle(x, y)}$

Косинус угла

Скалярное произведение

- Косинус угла: $\cos \angle(x, y) = \frac{\langle x, y \rangle}{\|x\| \|y\|}$

Скалярное произведение

- Косинус угла: $\cos\angle(x, y) = \frac{\langle x, y\rangle}{\|x\|\|y\|}$
- Мера сонаправленности векторов
- Для параллельных векторов $\cos\angle(x, y) = 1$
- Для перпендикулярных векторов $\cos\angle(x, y) = 0$

Резюме

- Виды чисел
- Векторы и матрицы — способ описания данных
- Линейная независимость и ранг — измерение количества информации в данных
- Матричное умножение — для удобства работы с линейными моделями
- Норма, метрика, угол — измерение сходства между объектами

На следующей лекции

- Математический анализ
- Как настроить линейную модель под данные?
- Как грамотно подобрать функционал качества?