Deep Learning School

Ликбез по линейной алгебре

Векторы

- операции над векторами
- линейные подпространства и линейная оболочка
- линейная независимость и базис

Вектор

Определение. *Вектором* в n-мерном евклидовом пространстве $\mathbb{R}^n$ называется упорядоченный набор чисел $x = (x_1, x_2, \ldots, x_n)$ — собственно, элемент пространства $\mathbb{R}^n$.

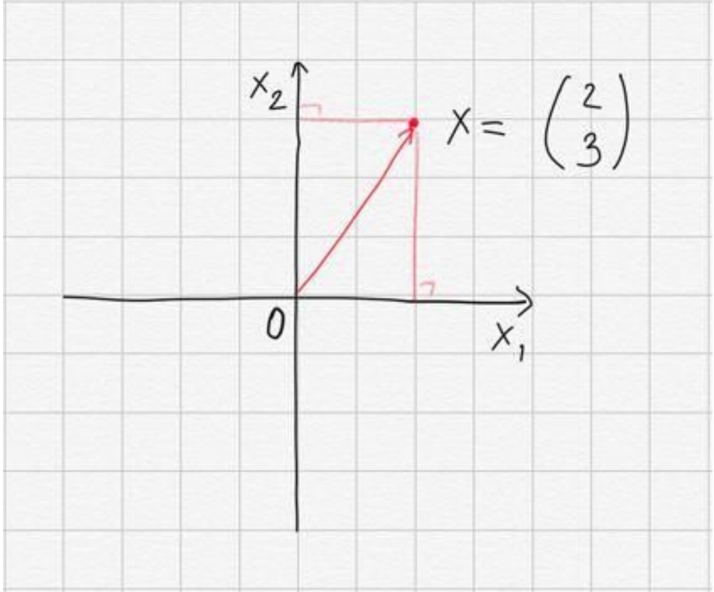
Вектор

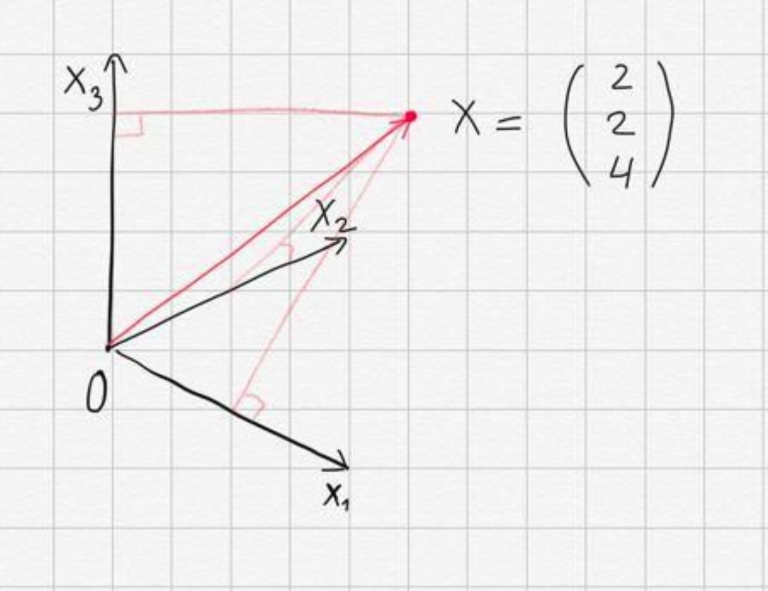
Определение. *Вектором* в n-мерном евклидовом пространстве $\mathbb{R}^n$ называется упорядоченный набор чисел $x = (x_1, x_2, \ldots, x_n)$ — собственно, элемент пространства $\mathbb{R}^n$.

Часто вектор удобнее записывать в столбец:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix}$$

Векторы на плоскости и в пространстве





Сложение и умножение на скаляр

Наблюдение. Векторы можно складывать и умножать на скаляр (число). Результат будет вектором, элементы которого суть результаты поэлементного применения операции.

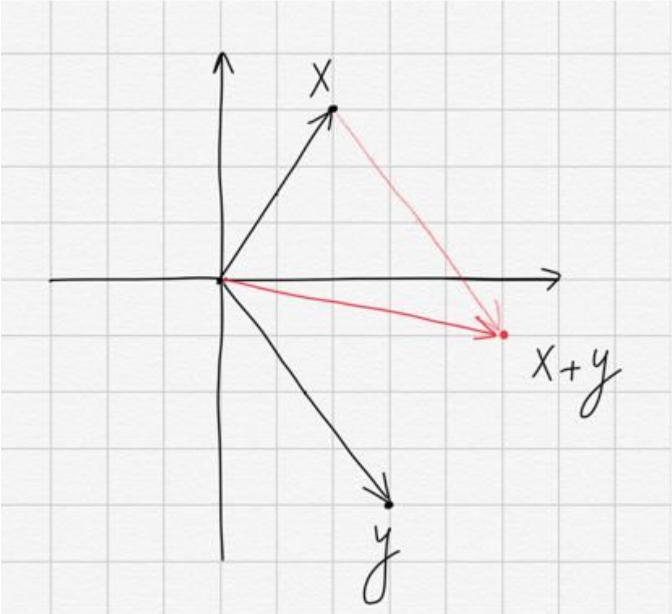
Сложение и умножение на скаляр

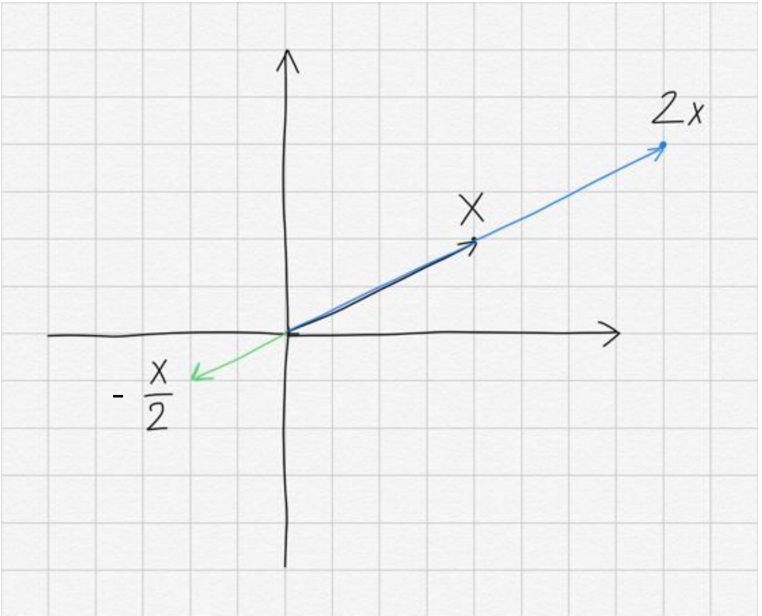
Наблюдение. Векторы можно складывать и умножать на скаляр (число). Результат будет вектором, элементы которого суть результаты поэлементного применения операции.

Пример.

$$3\begin{pmatrix}2\\-\frac{1}{2}\\4\end{pmatrix}+2\begin{pmatrix}-1\\0\\2\end{pmatrix}=\begin{pmatrix}4\\-\frac{3}{2}\\16\end{pmatrix}$$

Геометрия векторных операций





Линейные подпространства

- Векторное пространство $\mathbb{R}^n$ **замкнуто** относительно операций сложения и умножения на скаляр
- Обобщим это наблюдение

Линейные подпространства

- Векторное пространство $\mathbb{R}^n$ **замкнуто** относительно операций сложения и умножения на скаляр
- Обобщим это наблюдение

Определение. *Линейным (или векторным) подпространством* векторного пространства $\mathscr{L}$ называется множество векторов $\mathscr{M} \subset \mathscr{L}$, замкнутое относительно операций сложения и умножения на скаляр.

Линейная оболочка

Определение. *Линейной оболочкой* векторов $v_1, v_2, \ldots, v_n$ называется множество всех линейных комбинаций этих векторов с произвольными коэффициентами:

$$\mathscr{M} = \langle v_1, v_2, \ldots, v_n \rangle = \{\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n \mid \alpha_i \in \mathbb{R}\}$$

Линейная оболочка

Определение. *Линейной оболочкой* векторов $v_1, v_2, \ldots, v_n$ называется множество всех линейных комбинаций этих векторов с произвольными коэффициентами:

$$\mathscr{M} = \langle v_1, v_2, \ldots, v_n \rangle = \{\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n \mid \alpha_i \in \mathbb{R}\}$$

Утверждение. Линейная оболочка произвольного числа векторов является линейным подпространством в $\mathbb{R}^n$.

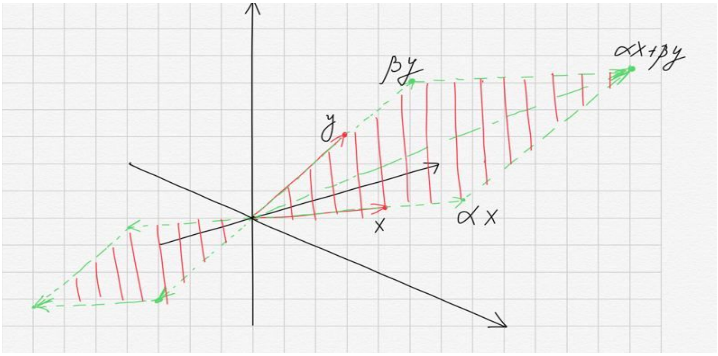
Линейная оболочка

Определение. *Линейной оболочкой* векторов $v_1, v_2, \ldots, v_n$ называется множество всех линейных комбинаций этих векторов с произвольными коэффициентами:

$$\mathscr{M} = \langle v_1, v_2, \ldots, v_n \rangle = \{\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n \mid \alpha_i \in \mathbb{R}\}$$

Утверждение. Линейная оболочка произвольного числа векторов является линейным подпространством в $\mathbb{R}^n$.

Пример. На данной картинке $<x, y>$ — плоскость



Линейная независимость

Определение. Векторы $v_1, v_2, \ldots, v_n$ называются *линейно независимыми*, если никакая нетривиальная линейная комбинация этих векторов не равна нуль-вектору. Иными словами, для любых $\alpha_i \in \mathbb{R}$, не все из которых нулевые, выполняется

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n \neq \overline{0}$$

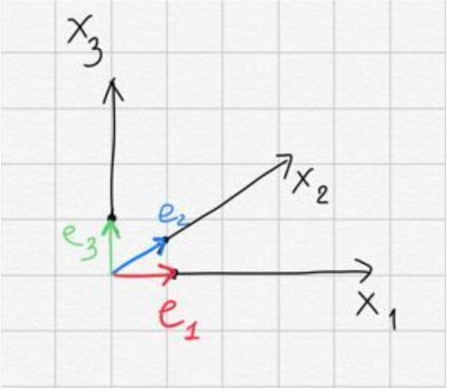
Линейная независимость

Определение. Векторы $v_1, v_2, \ldots, v_n$ называются *линейно независимыми*, если никакая нетривиальная линейная комбинация этих векторов не равна нуль-вектору. Иными словами, для любых $\alpha_i \in \mathbb{R}$, не все из которых нулевые, выполняется

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n \neq \overline{0}$$

Пример.

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$



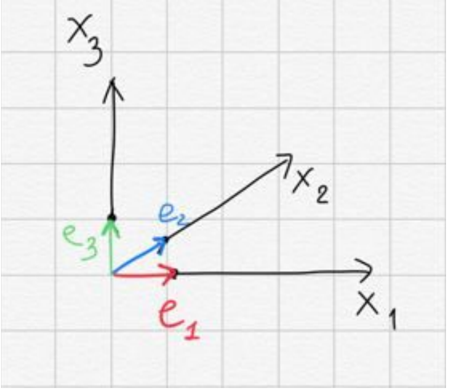
Линейная независимость

Определение. Векторы $v_1, v_2, \ldots, v_n$ называются *линейно независимыми*, если никакая нетривиальная линейная комбинация этих векторов не равна нуль-вектору. Иными словами, для любых $\alpha_i \in \mathbb{R}$, не все из которых нулевые, выполняется

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n \neq \overline{0}$$

Пример.

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$



Доказательство ЛНЗ. $\alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3 = 0 \;\Rightarrow\; \alpha_i = 0$ $\square$

Базис

Определение. Пусть $\mathscr{M}$ — линейное подпространство.
Базисом в $\mathscr{M}$ называется минимальная система векторов $v_1, v_2, \ldots, v_n$, для которой $\mathscr{M} = \langle v_1, v_2, \ldots, v_n \rangle$.

Базис

Определение. Пусть $\mathscr{M}$ — линейное подпространство.
Базисом в $\mathscr{M}$ называется минимальная система векторов $v_1, v_2, \ldots, v_n$, для которой $\mathscr{M} = \langle v_1, v_2, \ldots, v_n \rangle$.

Пример. $\mathscr{M} = \mathbb{R}^3$

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \text{ — базис}$$

Базис

Определение. Пусть $\mathscr{M}$ — линейное подпространство.
Базисом в $\mathscr{M}$ называется минимальная система векторов $v_1, v_2, \ldots, v_n$,
для которой $\mathscr{M} = \langle v_1, v_2, \ldots, v_n \rangle$.

Свойства базиса

- Базис является ЛНЗ системой
- Векторы из $\mathscr{M}$ выражается через базис единственным способом
- Любую ЛНЗ систему можно дополнить до базиса
- В любой системе образующих можно выбрать базис
- Любые два базиса равномощны

Базис

Определение. Пусть $\mathscr{M}$ — линейное подпространство.
Базисом в $\mathscr{M}$ называется минимальная система векторов $v_1, v_2, \ldots, v_n$, для которой $\mathscr{M} = \langle v_1, v_2, \ldots, v_n \rangle$.

Свойства базиса

- Базис является ЛНЗ системой
- Векторы из $\mathscr{M}$ выражается через базис единственным способом
- Любую ЛНЗ систему можно дополнить до базиса
- В любой системе образующих можно выбрать базис
- Любые два базиса равномощны

Последнее свойство свидетельствует о корректности определения *размерности линейного пространства* как размера базиса в этом линейном пространстве.

Базис

Теорема. $n+1$ векторов в n-мерном пространстве всегда линейно зависимы.

Доказательство.

Базис

Теорема. $n+1$ векторов в n-мерном пространстве всегда линейно зависимы.

Доказательство. От противного.

Базис

Теорема. $n{+}1$ векторов в n-мерном пространстве всегда линейно зависимы.

Доказательство. От противного. Они линейно независимы

Базис

Теорема. $n+1$ векторов в n-мерном пространстве всегда линейно зависимы.

Доказательство. От противного. Они линейно независимы $\rightarrow$ можно дополнить до базиса

Базис

Теорема. $n+1$ векторов в n-мерном пространстве всегда линейно зависимы.

Доказательство. От противного. Они линейно независимы → можно дополнить до базиса → противоречие с тем, что любые два базиса равномощны.

$\square$

Задачи

Задача 1. Доказать, что следующие вектора ЛНЗ:

$$a = \begin{pmatrix} 3 \\ -1 \\ 0 \\ 1 \end{pmatrix}, b = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 0 \end{pmatrix}, c = \begin{pmatrix} 2 \\ 0 \\ 1 \\ 2 \end{pmatrix}, d = \begin{pmatrix} 1 \\ 0 \\ -1 \\ -1 \end{pmatrix}$$

Задача 2. Найти базис в пространстве

$$\mathscr{M} = \{(x, y, z, t) : x + 2y - 3z + t = 0\}$$

Задача 3. Выразить вектор $\begin{pmatrix} 3 \\ 2 \\ 2 \\ -1 \end{pmatrix}$ через базисные вектора, найденные в задаче 2

Задачи

Задача 1. Доказать, что следующие вектора ЛНЗ:

$$a = \begin{pmatrix} 3 \\ -1 \\ 0 \\ 1 \end{pmatrix}, b = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 0 \end{pmatrix},$$

$$c = \begin{pmatrix} 2 \\ 0 \\ 1 \\ 2 \end{pmatrix}, d = \begin{pmatrix} 1 \\ 0 \\ -1 \\ -1 \end{pmatrix}$$

Задачи

Задача 2

Найти базис

в пространстве

$$\mathscr{M} = \{(x, y, z, t) :$$
$$x + 2y - 3z + t = 0\}$$

Задачи

Задача 3

Выразить вектор $\begin{pmatrix} 3 \\ 2 \\ 2 \\ -1 \end{pmatrix}$

через базисные вектора,

найденные в задаче 2

Резюме

- Вектор — удобная форма представления различных математических объектов
- Линейное пространство — множество векторов, замкнутое относительно сложения и умножения на скаляр
- Базис — универсальный способ описания линейных подпространств в $\mathbb{R}^n$
- Размер базиса, или размерность пространства, является важной характеристикой ЛП

Матрицы

- определение
- (не)вырожденность и ранг
- умножение матрицы на вектор и матричный вид СЛУ
- пример: линейная регрессия

Матрицы

Определение. *Матрицей размера* $m \times n$ называется прямоугольная таблица с числами из m строк и n столбцов:

$$\begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \dots & \dots & \dots & \dots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{pmatrix}$$

Матрицы

Определение. *Матрицей размера* $m \times n$ называется прямоугольная таблица с числами из m строк и n столбцов:

$$\begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \dots & \dots & \dots & \dots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{pmatrix}$$

Удобно представлять матрицу как совокупность из n векторов-столбцов, записанных в строчку:

$$(x_{*1}, x_{*2}, \dots, x_{*n})$$

Вырожденная матрица

Определение. Квадратная матрица называется *(не)вырожденной*, если её строки линейно (не)зависимы.

Пример.

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

Вырожденная матрица

Определение. Квадратная матрица называется *(не)вырожденной*, если её строки линейно (не)зависимы.

Пример.

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

Утверждение. Строки квадратной матрицы ЛНЗ тогда и только тогда, когда её столбцы ЛНЗ.

Случай произвольных матриц

Определение. *Строчным рангом* матрицы A называется размер наибольшего подмножества линейно независимых строк A. Аналогично определяется *столбцовый ранг*.

Случай произвольных матриц

Определение. *Строчным рангом* матрицы A называется размер наибольшего подмножества линейно независимых строк A. Аналогично определяется *столбцовый ранг*.

Пример.

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

строчный и столбцовый ранг равны 3

Случай произвольных матриц

Определение. *Строчным рангом* матрицы A называется размер наибольшего подмножества линейно независимых строк A. Аналогично определяется *столбцовый ранг*.

Пример.

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

строчный и столбцовый ранг равны 3

Утверждение. Строчный и столбцовый ранг равны.

Умножение матрицы на вектор

$$\begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \hdotsfor{4} \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{pmatrix}$$

Умножение матрицы на вектор

$$\begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \hdotsfor{4} \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{pmatrix} \times \begin{pmatrix} w_1 \\ w_2 \\ \dots \\ w_k \end{pmatrix} =$$

Умножение матрицы на вектор

$$\begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \dots & \dots & \dots & \dots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{pmatrix} \times \begin{pmatrix} w_1 \\ w_2 \\ \dots \\ w_k \end{pmatrix} = \begin{pmatrix} \langle x_{1*}, w \rangle \\ \langle x_{2*}, w \rangle \\ \dots \\ \langle x_{m*}, w \rangle \end{pmatrix}$$

где

$$\langle x, y \rangle = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n$$

Пример: система линейных уравнений

- Решаем систему

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{cases}$$

- Запись системы в матричном виде: $Ax = b$

Система линейных уравнений

Теорема. Пусть A — невырожденная квадратная матрица.
Тогда система линейных уравнений $Ax = b$ имеет единственное решение при любом значении b.

Пример: линейная регрессия

- Есть m объектов (квартир)
- Объект описывается n признаками (площадь, этаж, количество комнат, ...)

$$x_{k*} = (x_{k1}, x_{k2}, \ldots, x_{kn})$$

- Необходимо предсказать целевую переменную y_k (стоимость квартиры)
- Ищем закономерность в *линейном виде*:

$$y_k = w_1 x_{k1} + w_2 x_{k2} + \ldots + w_n x_{kn}$$

Линейная регрессия в матричном виде

- Ищем закономерность в *линейном виде*:

$$y_k = w_1 x_{k1} + w_2 x_{k2} + \ldots + w_n x_{kn}$$

- В матричном виде уравнение записывается так:

$$Xw = y$$

$$X = \begin{pmatrix} x_{11} & x_{12} & \ldots & x_{1n} \\ x_{21} & x_{22} & \ldots & x_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ x_{m1} & x_{m2} & \ldots & x_{mn} \end{pmatrix}, \; w = \begin{pmatrix} w_1 \\ w_2 \\ \ldots \\ w_n \end{pmatrix}, \; y = \begin{pmatrix} y_1 \\ y_2 \\ \ldots \\ y_n \end{pmatrix}$$

Линейная регрессия в матричном виде

- Ищем закономерность в *линейном виде*:

$$y_k = w_1 x_{k1} + w_2 x_{k2} + \ldots + w_n x_{kn}$$

- В матричном виде уравнение записывается так:

$$Xw = y$$

$$X = \begin{pmatrix} x_{11} & x_{12} & \ldots & x_{1n} \\ x_{21} & x_{22} & \ldots & x_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ x_{m1} & x_{m2} & \ldots & x_{mn} \end{pmatrix}, \; w = \begin{pmatrix} w_1 \\ w_2 \\ \ldots \\ w_n \end{pmatrix}, \; y = \begin{pmatrix} y_1 \\ y_2 \\ \ldots \\ y_n \end{pmatrix}$$

- Если $m = n$, то решение (скорее всего) единственное
- Если $m > n$, то решения (скорее всего) нет
- Если $m < n$, то решений (скорее всего) бесконечно много

Линейная регрессия в матричном виде

Если объектов меньше, чем признаков,
то линейная регрессия будет работать плохо!

Операции над матрицами

- сложение матриц
- умножение матриц
- транспонирование и обратная матрица
- определитель матрицы

Сложение и вычитание матриц

- Сложение и вычитание происходит поэлементно
- Можно применять только к матрицам одинакового размера

Пример.

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

Умножение матриц

- Даны матрицы A размера $m \times k$ и B размера $k \times n$
- Хотим научиться вычислять матричное произведение AB

Умножение матриц

- Даны матрицы A размера $m \times k$ и B размера $k \times n$
- Хотим научиться вычислять матричное произведение AB

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

Умножение матриц

- Даны матрицы A размера $m \times k$ и B размера $k \times n$
- Хотим научиться вычислять матричное произведение AB

$$\begin{pmatrix} 2 & 3 & -1 \\ 3 & -2 & 0 \end{pmatrix} \times \begin{pmatrix} -2 & 2 & 0 \\ 2 & -2 & 1 \\ -1 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 3 & -2 & 1 \\ -10 & 10 & -2 \end{pmatrix}$$

Умножение матриц

- Даны матрицы A размера $m \times k$ и B размера $k \times n$
- Хотим научиться вычислять матричное произведение AB

$$\begin{pmatrix} \boxed{2 \quad 3 \quad -1} \\ 3 \quad -2 \quad 0 \end{pmatrix} \times \begin{pmatrix} -2 & 2 & 0 \\ 2 & -2 & 1 \\ -1 & 0 & 2 \end{pmatrix} = \begin{pmatrix} \boxed{3} & -2 & 1 \\ -10 & 10 & -2 \end{pmatrix}$$

$$2 \cdot (-2) + 3 \cdot 2 + (-1) \cdot (-1) = 3$$

Умножение матриц

- Даны матрицы A размера $m \times k$ и B размера $k \times n$
- Хотим научиться вычислять матричное произведение AB

$$\begin{pmatrix} \boxed{2 \quad 3 \quad -1} \\ 3 \quad -2 \quad 0 \end{pmatrix} \times \begin{pmatrix} -2 & \boxed{\begin{matrix} 2 \\ -2 \\ 0 \end{matrix}} & \begin{matrix} 0 \\ 1 \\ 2 \end{matrix} \end{pmatrix} = \begin{pmatrix} 3 & \boxed{-2} & 1 \\ -10 & 10 & -2 \end{pmatrix}$$

$$2 \cdot 2 + 3 \cdot (-2) + (-1) \cdot 0 = -2$$

Умножение матриц

- Даны матрицы A размера $m \times k$ и B размера $k \times n$
- Хотим научиться вычислять матричное произведение AB

$$\begin{pmatrix} \boxed{2 \quad 3 \quad -1} \\ 3 \quad -2 \quad 0 \end{pmatrix} \times \begin{pmatrix} -2 & 2 & 0 \\ 2 & -2 & 1 \\ -1 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 3 & -2 & \boxed{1} \\ -10 & 10 & -2 \end{pmatrix}$$

$$2 \cdot 0 + 3 \cdot 1 + (-1) \cdot 2 = 1$$

Умножение матриц

- Даны матрицы A размера $m \times k$ и B размера $k \times n$
- Хотим научиться вычислять матричное произведение AB

$$\begin{pmatrix} 2 & 3 & -1 \\ \boxed{3 \quad -2 \quad 0} \end{pmatrix} \times \begin{pmatrix} \boxed{-2} & 2 & 0 \\ \boxed{2} & -2 & 1 \\ \boxed{-1} & 0 & 2 \end{pmatrix} = \begin{pmatrix} 3 & -2 & 1 \\ \boxed{-10} & 10 & -2 \end{pmatrix}$$

$$3 \cdot (-2) + (-2) \cdot 2 + 0 \cdot (-1) = -10$$

Умножение матриц

- Даны матрицы A размера $m \times k$ и B размера $k \times n$
- Хотим научиться вычислять матричное произведение AB

$$\begin{pmatrix} 2 & 3 & -1 \\ \boxed{3 \quad -2 \quad 0} \end{pmatrix} \times \begin{pmatrix} -2 & 2 & 0 \\ 2 & -2 & 1 \\ -1 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 3 & -2 & 1 \\ -10 & \boxed{10} & -2 \end{pmatrix}$$

$$3 \cdot 2 + (-2) \cdot -2 + 0 \cdot 0 = 10$$

Умножение матриц

- Даны матрицы A размера $m \times k$ и B размера $k \times n$
- Хотим научиться вычислять матричное произведение AB

$$\begin{pmatrix} 2 & 3 & -1 \\ \boxed{3 \quad -2 \quad 0} \end{pmatrix} \times \begin{pmatrix} -2 & 2 & \boxed{0} \\ 2 & -2 & \boxed{1} \\ -1 & 0 & \boxed{2} \end{pmatrix} = \begin{pmatrix} 3 & -2 & 1 \\ -10 & 10 & \boxed{-2} \end{pmatrix}$$

$$3 \cdot 0 + (-2) \cdot 1 + 0 \cdot 2 = -2$$

Произведение матриц

- Частный случай — произведение матрицы на вектор
- Произведение матриц встречается тогда, когда совокупность векторов умножается на матрицу (например, при подаче в нейронную сеть батча данных)

Свойства произведения матриц

- Ассоциативность: $A(BC) = (AB)C$
- Дистрибутивность: $A(B{+}C) = AB + AC$
- Существование нейтрального элемента E_n (единичная матрица):

Свойства произведения матриц

- Ассоциативность: $A(BC) = (AB)C$
- Дистрибутивность: $A(B+C) = AB + AC$
- Существование нейтрального элемента E_n (единичная матрица):

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

Свойства произведения матриц

- Ассоциативность: $A(BC) = (AB)C$
- Дистрибутивность: $A(B+C) = AB + AC$
- Существование нейтрального элемента E_n (единичная матрица):

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

$$AE = EA = A$$

Свойства произведения матриц

- Ассоциативность: $A(BC) = (AB)C$
- Дистрибутивность: $A(B+C) = AB + AC$
- Существование нейтрального элемента E_n (единичная матрица):

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

$AE = EA = A$

- Отсутствие коммутативности: не всегда $AB = BA$

Свойства произведения матриц

- Ассоциативность: $A(BC) = (AB)C$
- Дистрибутивность: $A(B+C) = AB + AC$
- Существование нейтрального элемента E_n (единичная матрица):

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

  $AE = EA = A$

- Отсутствие коммутативности: не всегда $AB = BA$
- Для квадратных матриц: если A вырождена, то AB также вырождена

Обратная матрица

Определение. Пусть A — **квадратная** матрица. Если существует такая матрица A^{-1}, что $AA^{-1} = A^{-1}A = E$, то A^{-1} называется *обратной матрицей* к A. Матрица A в таком случае называется *обратимой*.

Обратная матрица

Определение. Пусть A — **квадратная** матрица. Если существует такая матрица A^{-1}, что $AA^{-1} = A^{-1}A = E$, то A^{-1} называется *обратной матрицей* к A. Матрица A в таком случае называется *обратимой*.

Утверждение. Пусть A — квадратная матрица. Если строки (или столбцы) A линейно независимы (т.е. A невырожденная), то обратная матрица существует и единственна.

Решаем систему линейных уравнений

- Есть СЛУ, записанная в матричном виде:

$$Ax = b$$

- Если существует A^{-1}, то у системы есть единственное решение:

Решаем систему линейных уравнений

- Есть СЛУ, записанная в матричном виде:

$$Ax = b$$

- Если существует A^{-1}, то у системы есть единственное решение:

$$x = A^{-1}b$$

Транспонированная матрица

Транспонирование — операция отражения матрицы относительно главной диагонали
Пишут: $A^\top$

Вектор-столбец при транспонировании переходит в вектор-строку. Поэтому скалярное произведение можно записать так:

$$\langle x, y \rangle = x^\top y$$

Определитель матрицы

Определитель квадратной матрицы — это её числовая характеристика, которая определяется рекурсивно:

- $\begin{vmatrix} a \end{vmatrix} = a$
- $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$

Определитель матрицы

Определитель квадратной матрицы — это её числовая характеристика, которая определяется рекурсивно:

$$|A| = a_{11}|A_{11}| - a_{12}|A_{12}| + \ldots + (-1)^n|A_{1n}|$$

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

$$A_{ij} = \begin{pmatrix} a_{11} & \ldots & a_{1,j-1} & a_{1,j+1} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{i-1,1} & \ldots & a_{i-1,j-1} & a_{i-1,j+1} & \ldots & a_{i-1,n} \\ a_{i+1,1} & \ldots & a_{i+1,j-1} & a_{i+1,j+1} & \ldots & a_{i+1,n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{n,1} & \ldots & a_{n,j-1} & a_{n,j+1} & \ldots & a_{n,n} \end{pmatrix}$$

Свойства определителя

- $|AB| = |A||B|$
- $|A| = 0$ тогда и только тогда, когда A вырожденная

Вычисление обратной матрицы

- Цель: хотим научиться вычислять обратную матрицу

Алгоритм (метод Крамера). Дана квадратная матрица A.

1. Вычислим $|A|$
2. Построим матрицу миноров:

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

3. Обратная матрица вычисляется по формуле

$$A^{-1} = \frac{(A')^\top}{|A|}$$

Задачи

Задача 1
Вычислить определитель
матрицы

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

Задачи

Задача 2
Найти обратную
к матрице

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

Резюме

С матрицами можно делать следующее:

- Складывать, вычитать
- Умножать
- Находить обратную
- Транспонировать
- Считать определитель

- Все эти операции так или иначе необходимы для теоретического понимания матричного исчисления
- Большая часть операций так или иначе используется в нейросетях

Матрица и линейный оператор

- матрица и линейный оператор
- геометрический смысл линейного преобразования
- линейный оператор в нейронной сети

Матрица и линейный оператор

- Пусть A — некоторая матрица размера $m \times n$
- Рассмотрим оператор $f : \mathbb{R}^n \to \mathbb{R}^m$, действующий по формуле $f(x) = Ax$
- Отображение f является линейным, то есть

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$$

Матрица и линейный оператор

- Пусть A — некоторая матрица размера $m \times n$
- Рассмотрим оператор $f : \mathbb{R}^n \to \mathbb{R}^m$, действующий по формуле $f(x) = Ax$
- Отображение f является линейным, то есть

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$$

$$f(\alpha x + \beta y) = A(\alpha x + \beta y) =$$

Матрица и линейный оператор

- Пусть A — некоторая матрица размера $m \times n$
- Рассмотрим оператор $f : \mathbb{R}^n \to \mathbb{R}^m$, действующий по формуле $f(x) = Ax$
- Отображение f является линейным, то есть

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$$

$$f(\alpha x + \beta y) = A(\alpha x + \beta y) =$$

$$= \begin{pmatrix} \langle a_{1*}, (\alpha x + \beta y) \rangle \\ \langle a_{2*}, (\alpha x + \beta y) \rangle \\ \dots \\ \langle a_{m*}, (\alpha x + \beta y) \rangle \end{pmatrix}$$

Матрица и линейный оператор

- Пусть A — некоторая матрица размера $m \times n$
- Рассмотрим оператор $f : \mathbb{R}^n \to \mathbb{R}^m$, действующий по формуле $f(x) = Ax$
- Отображение f является линейным, то есть

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$$

$$f(\alpha x + \beta y) = A(\alpha x + \beta y) =$$

$$= \begin{pmatrix} \langle a_{1*}, (\alpha x + \beta y) \rangle \\ \langle a_{2*}, (\alpha x + \beta y) \rangle \\ \dots \\ \langle a_{m*}, (\alpha x + \beta y) \rangle \end{pmatrix} = \alpha \begin{pmatrix} \langle a_{1*}, x \rangle \\ \langle a_{2*}, x \rangle \\ \dots \\ \langle a_{m*}, x \rangle \end{pmatrix} + \beta \begin{pmatrix} \langle a_{1*}, y \rangle \\ \langle a_{2*}, y \rangle \\ \dots \\ \langle a_{m*}, y \rangle \end{pmatrix}$$

Геометрия линейного преобразования

- Рассмотрим линейный оператор $f : \mathbb{R}^n \to \mathbb{R}^m$ при $m = n = 2$
- f задаётся матрицей A и задаёт некоторое преобразование плоскости

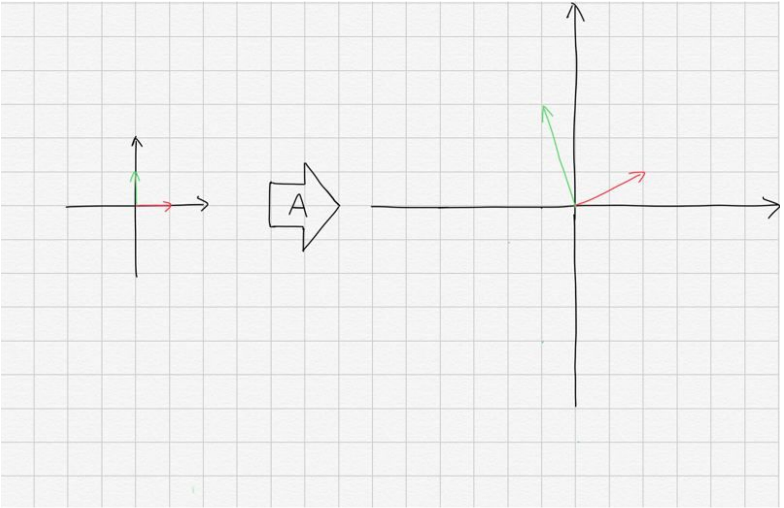
Геометрия линейного оператора

Пример 1.

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

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \to \begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} \to \begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \end{pmatrix}$$



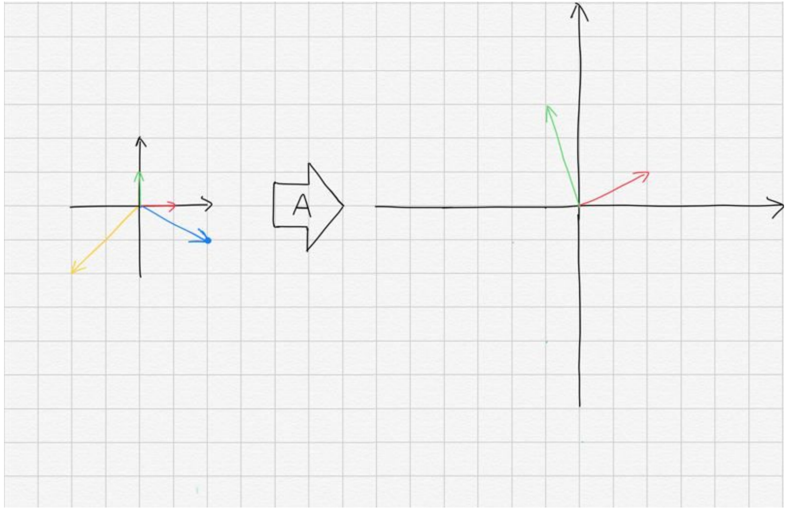
Геометрия линейного оператора

Пример 1.

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

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \to \begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} \to \begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \end{pmatrix}$$



$$\begin{pmatrix} 2 \\ -1 \end{pmatrix} \to$$

$$\begin{pmatrix} -2 \\ -2 \end{pmatrix} \to$$

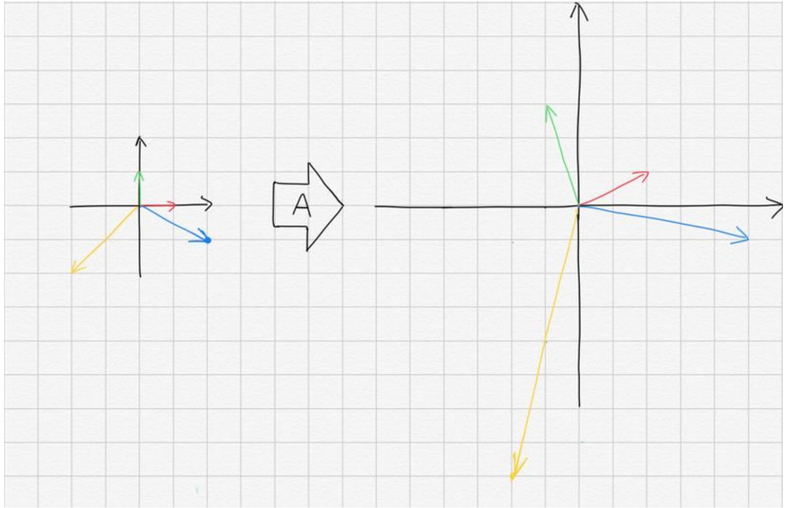
Геометрия линейного оператора

Пример 1.

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

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \to \begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} \to \begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \end{pmatrix}$$



$$\begin{pmatrix} 2 \\ -1 \end{pmatrix} \to \begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 5 \\ -1 \end{pmatrix}$$

$$\begin{pmatrix} -2 \\ -2 \end{pmatrix} \to \begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix} \cdot \begin{pmatrix} -2 \\ -2 \end{pmatrix} = \begin{pmatrix} -2 \\ -8 \end{pmatrix}$$

Геометрия линейного оператора

Пример 2. Что если матрица A вырожденная?

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

Геометрия линейного оператора

Пример 2. Что если матрица A вырожденная?

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

Все Ax укладываются на прямую $-2x_1 = x_2$

Геометрия линейного оператора

Пример 2. Что если матрица A вырожденная?

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

Все Ax укладываются на прямую $-2x_1 = x_2$

Вывод. Вырожденный линейный оператор отображает пространство в пространство меньшей размерности

Матрица поворота

Утверждение. Преобразованию поворота на угол α (против часовой стрелки) соответствует матрица

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

Матрица поворота

Утверждение

Преобразованию

поворота на угол α

(против часовой стрелки)

соответствует матрица

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

Матрица поворота

Утверждение. Преобразованию поворота на угол α (против часовой стрелки) соответствует матрица

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

Доказательство.

Композиция операторов

- Даны два оператора, имеющие матрицы A и B
- Необходимо найти матрицу композиции двух операторов

Утверждение. Матрица композиции линейных операторов равна произведению матриц этих линейных операторов

Композиция операторов

Утверждение. Матрица композиции линейных операторов равна произведению матриц этих линейных операторов

Доказательство.

- Рассмотрим образ вектора x
- Под действием первого оператора он переходит в вектор Ax
- Вектор Ax под действием второго оператора переходит в $B(Ax)$
- Под действием композиции x переходит в $(BA)x$

Композиция операторов

Утверждение. Матрица композиции линейных операторов равна произведению матриц этих линейных операторов

Пример. Композиция двух поворотов с центром в нуле — поворот с центром в нуле

$$\begin{pmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{pmatrix} \cdot \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} =$$

$$= \begin{pmatrix} \cos\beta\cos\alpha - \sin\beta\sin\alpha & -\cos\beta\sin\alpha - \sin\beta\cos\alpha \\ \sin\beta\cos\alpha + \cos\beta\sin\alpha & -\sin\beta\sin\alpha + \cos\beta\cos\alpha \end{pmatrix} =$$

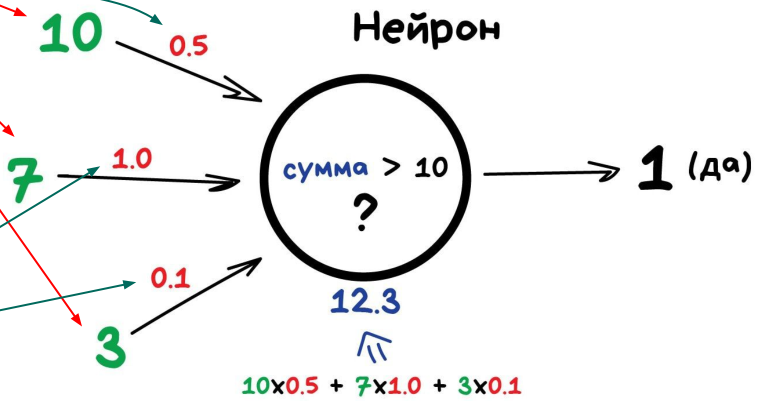
$$= \begin{pmatrix} \cos(\alpha+\beta) & -\sin(\alpha+\beta) \\ \sin(\alpha+\beta) & \cos(\alpha+\beta) \end{pmatrix}$$

Пример: матрица в нейронной сети

Опишем нейронную сеть в терминах матриц!

Модель нейрона

входы нейрона:
$x = (x_1, x_2, \dots, x_n)$



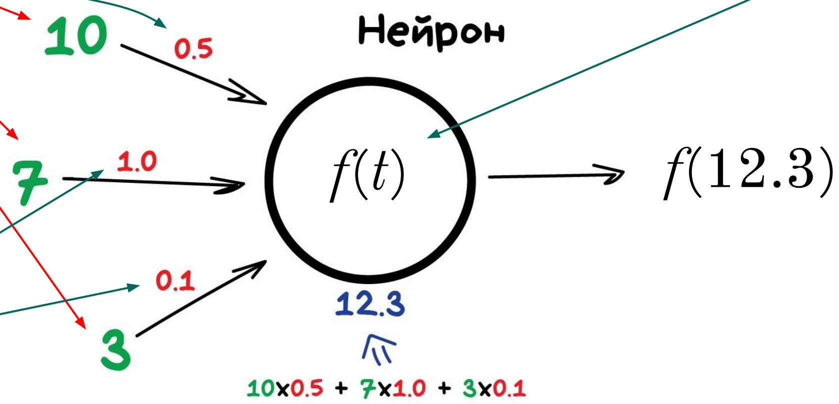
веса нейрона:
$w = (w_1, w_2, \dots, w_n)$

скалярное произведение векторов x, w:
$$w_1 \cdot x_1 + w_2 \cdot x_2 + \ldots + w_n \cdot x_n = \langle w, x \rangle$$

Модель нейрона

входы нейрона:
$x = (x_1, x_2, \ldots, x_n)$

функция активации

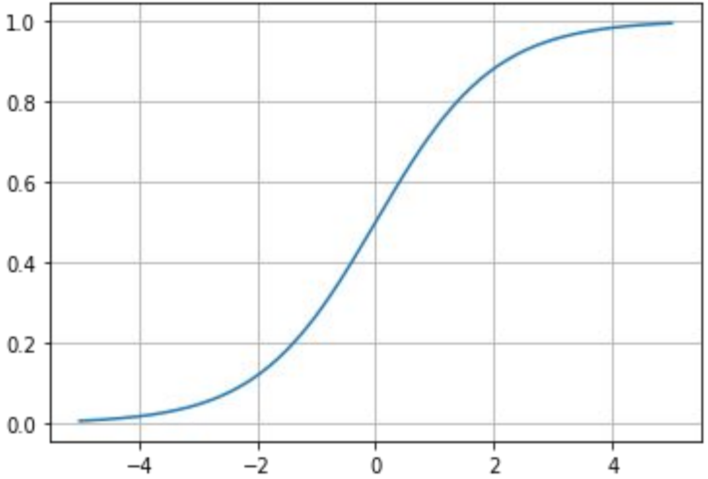


веса нейрона:
$w = (w_1, w_2, \ldots, w_n)$

скалярное произведение векторов x, w:
$$w_1 \cdot x_1 + w_2 \cdot x_2 + \ldots + w_n \cdot x_n = \langle w, x \rangle$$

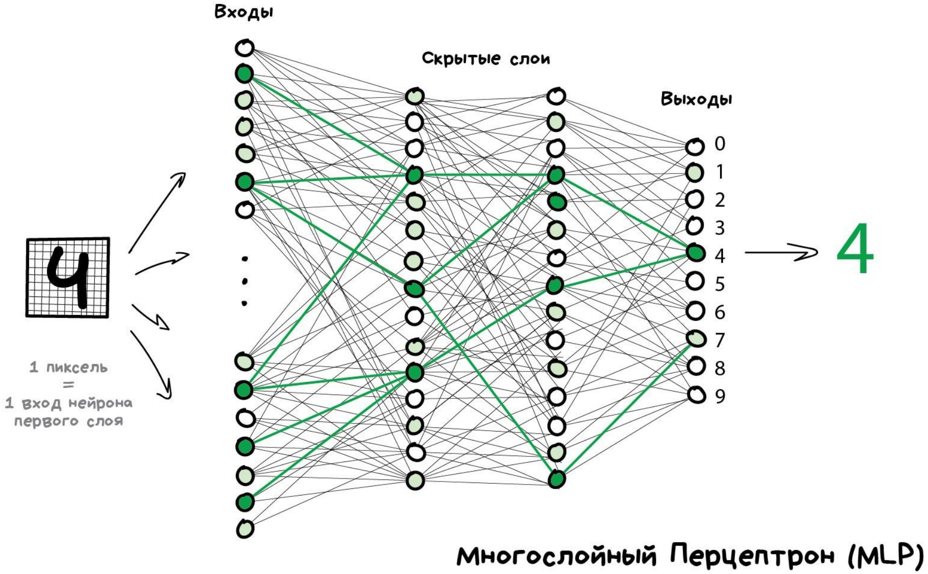
Функция сигмоиды

$$\sigma(t) = \frac{1}{1 + e^{-t}}$$

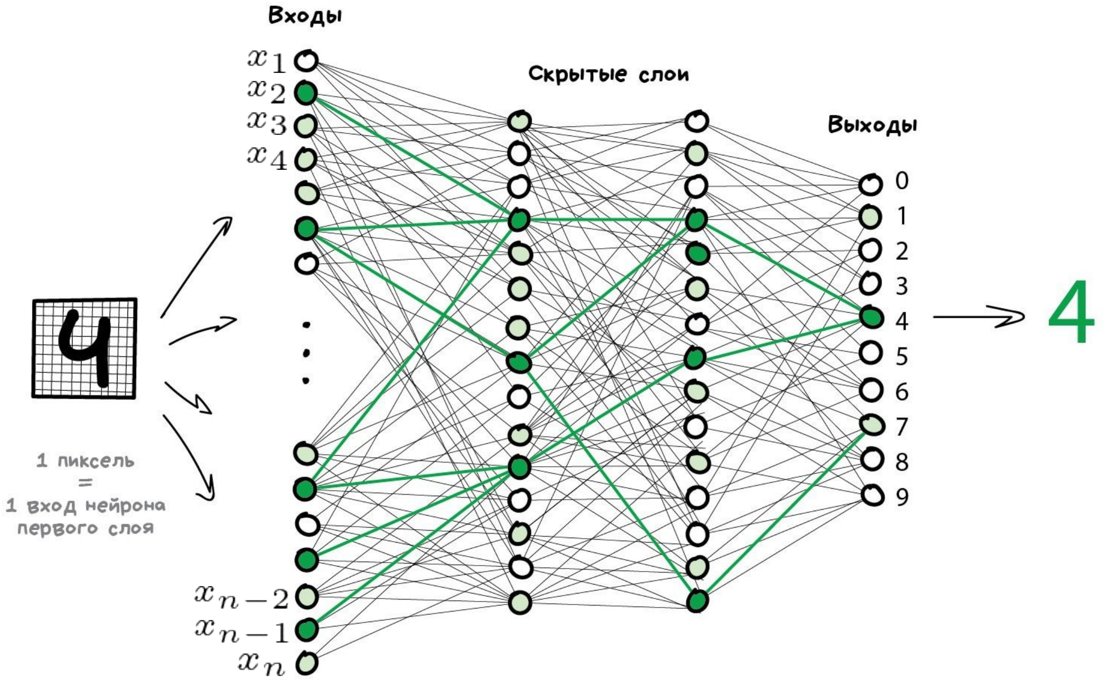


Многослойный перцептрон

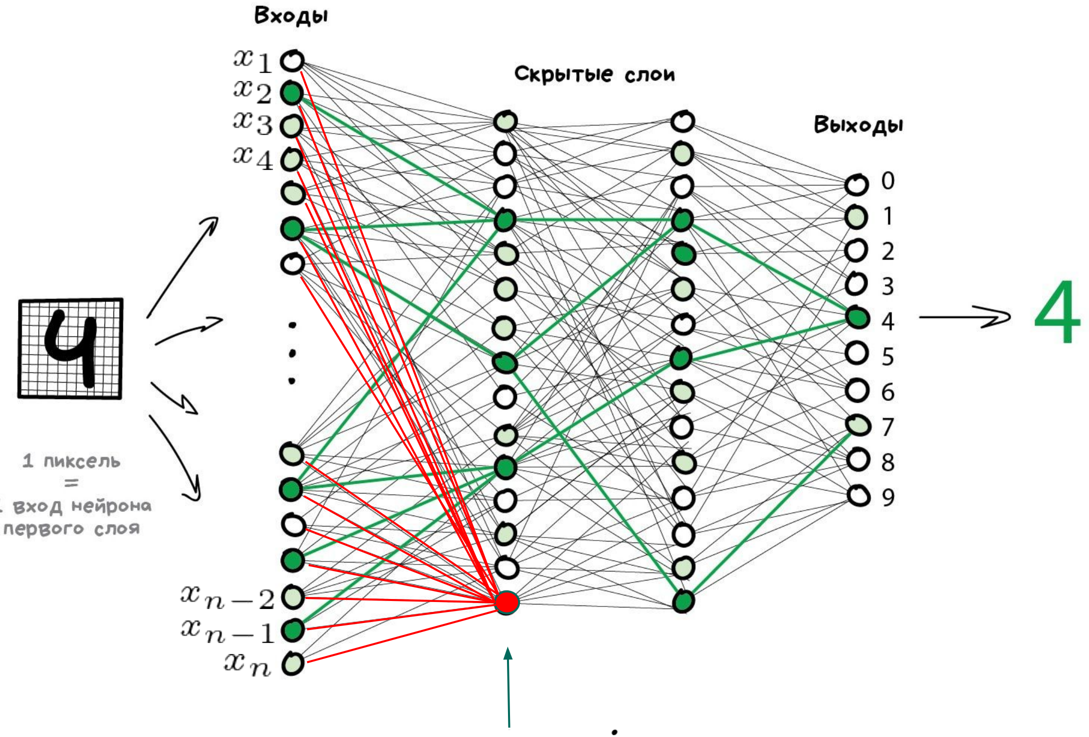
- Многослойный перцептрон — простейшая архитектура нейронной сети
- Каждый слой нейронов связан со всем нейронами с предыдущего слоя
- Выходные нейроны соответствуют классам изображений



Многослойный перцептрон



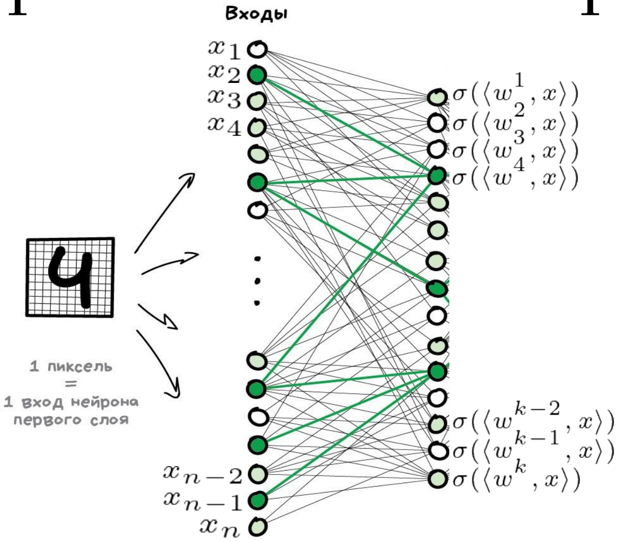
Многослойный перцептрон



$$y_i = \sigma(\langle w^i, x \rangle)$$

w^i — вектор весов i-ого нейрона

Преобразование вектора в перцептроне



$$= \sigma \begin{pmatrix} \langle w^1, x\rangle \\ \langle w^2, x\rangle \\ \dots \\ \langle w^k, x\rangle \end{pmatrix} = \sigma(Wx)$$

линейное преобразование вектора x

$$W = \begin{pmatrix} w_1^1 & w_2^1 & \dots & w_n^1 \\ w_1^2 & w_2^2 & \dots & w_n^2 \\ \vdots & \ddots & \ddots & \vdots \\ w_1^k & w_2^k & \dots & w_n^k \end{pmatrix} = \begin{pmatrix} w^1 \\ w^2 \\ \dots \\ w^k \end{pmatrix} \qquad x = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}$$

Преобразование вектора в перцептроне

- Полносвязный слой нейронной сети выполняет линейный оператор
- Функция активации создаёт нелинейность: без неё нейронная сеть была бы просто линейным алгоритмом

Резюме

- Матрица соответствует линейному оператору
- Умножение матриц соответствует композиции линейных операторов
- Линейные операторы имеют естественную геометрическую интерпретацию
- Нейронные сети удобно описывать в терминах матриц и линейных операторов

The End