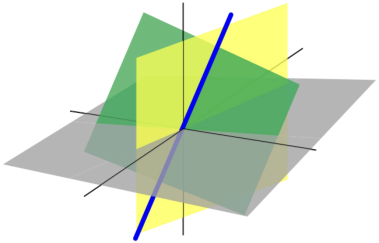
Ликбез по линейной алгебре

Векторы и матрицы



Векторы

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 $\mathbb{R}^n$.

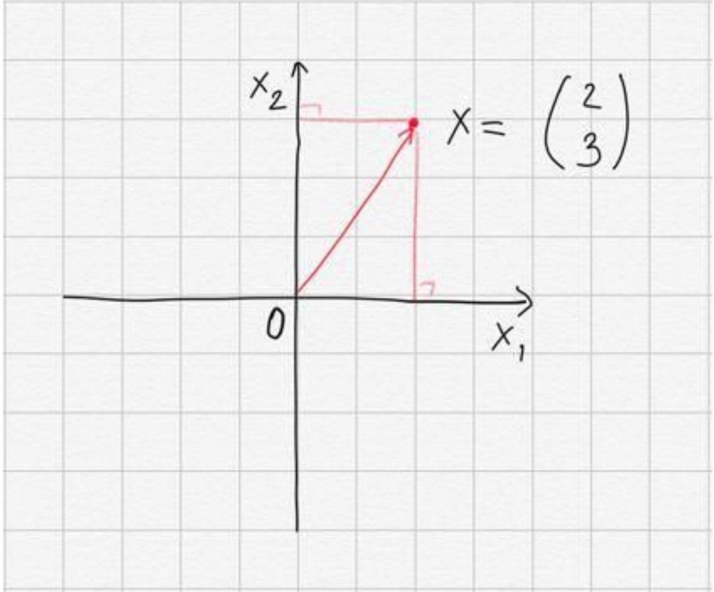
Вектор

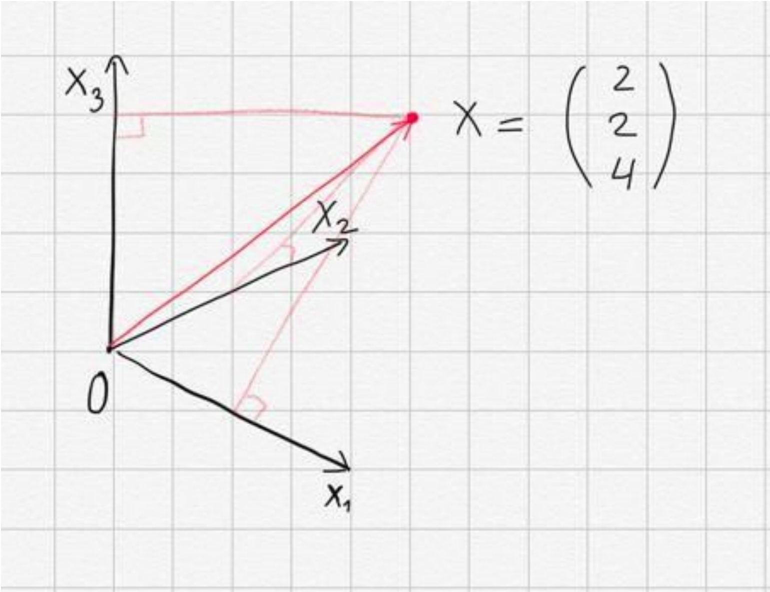
Определение. *Вектором* в n-мерном евклидовом пространстве $\mathbb{R}^n$ называется упорядоченный набор чисел $x = (x_1, x_2, \ldots, x_n)$ — собственно, элемент пространства $\mathbb{R}^n$.

Часто вектор удобнее записывать в столбец:

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

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.

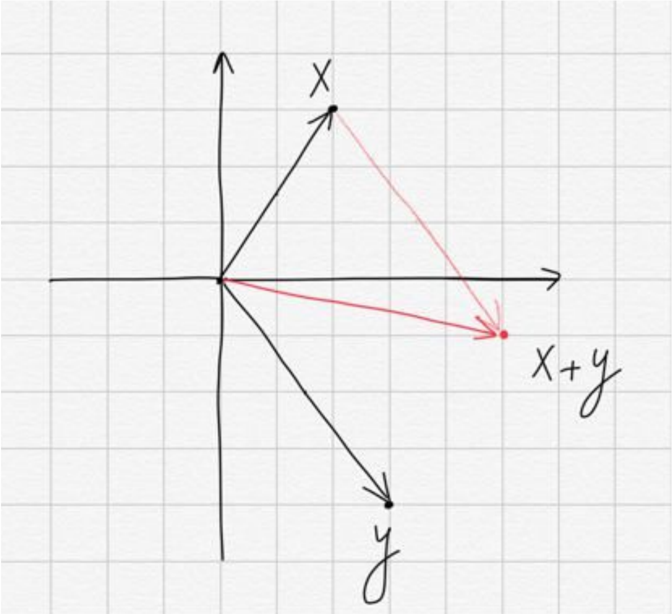
Сложение и умножение на скаляр

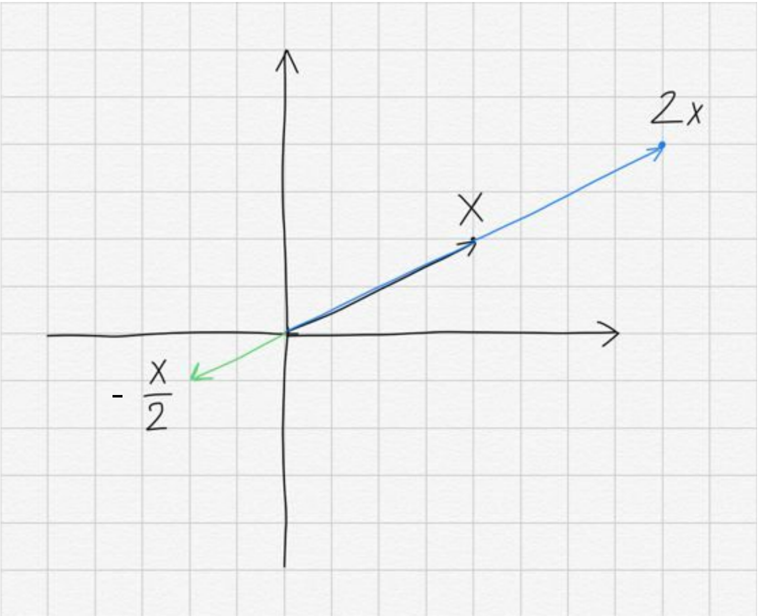
Наблюдение. Векторы можно складывать и умножать на скаляр (число). Результат будет вектором, элементы которого суть результаты поэлементного применения операции. Вектор — удобная форма представления различных математических объектов

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





Матрицы

Матрицы

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

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

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:

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

Умножение матрицы на вектор

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

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: $Ax = b$

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:

$$Xw = y$$

$$X = \begin{pmatrix} x_{11} & x_{12} & \ldots & x_{1n} \\ x_{21} & x_{22} & \ldots & x_{2n} \\ \multicolumn{4}{c}{\dotfill} \\ x_{m1} & x_{m2} & \ldots & x_{mn} \end{pmatrix}, \; w = \begin{pmatrix} w_1 \\ w_2 \\ \ldots \\ w_n \end{pmatrix}, \; y = \begin{pmatrix} y_1 \\ y_2 \\ \ldots \\ y_n \end{pmatrix}$$

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 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

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

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

Умножение матриц

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

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

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

Умножение матриц

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

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

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

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

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

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

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 $AB = BA$

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*.

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:

$$Ax = b$$

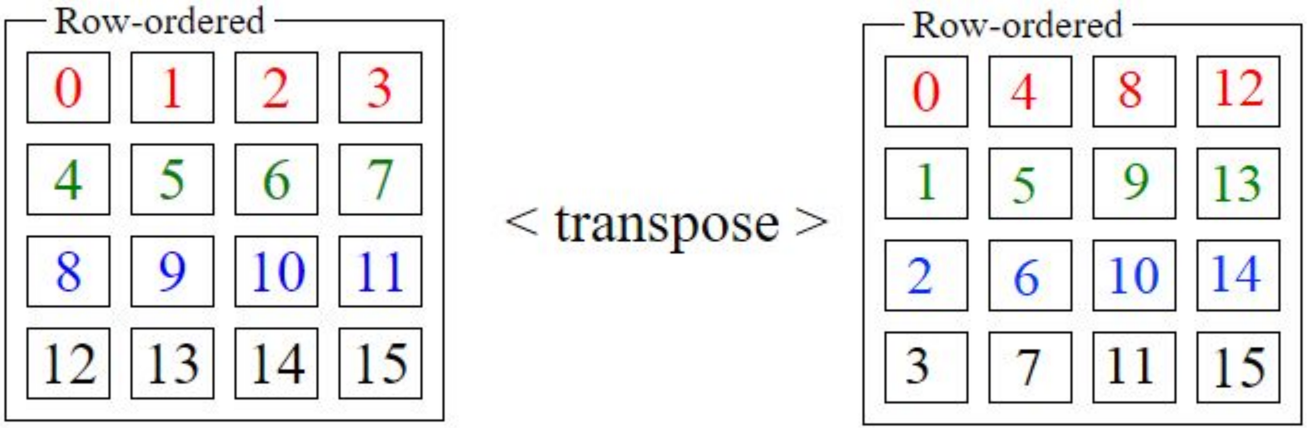
- Если существует A^{-1}, то у системы есть единственное решение:

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

Транспонированная матрица

Транспонирование — операция отражения матрицы относительно главной диагонали

Пишут: $A^\top$



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:

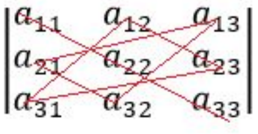
$$\langle x, y\rangle = x^\top y$$

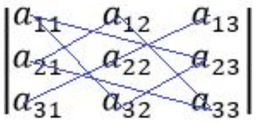
Определитель матрицы

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

- $\begin{vmatrix} a \end{vmatrix} = a$
- $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$
- $\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \cdot a_{22} \cdot a_{33} + a_{12} \cdot a_{23} \cdot a_{31} + a_{13} \cdot a_{21} \cdot a_{32} -$

$$- a_{13} \cdot a_{22} \cdot a_{31} - a_{11} \cdot a_{23} \cdot a_{32} - a_{12} \cdot a_{21} \cdot a_{33}$$





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

Резюме

С матрицами можно делать следующее:

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

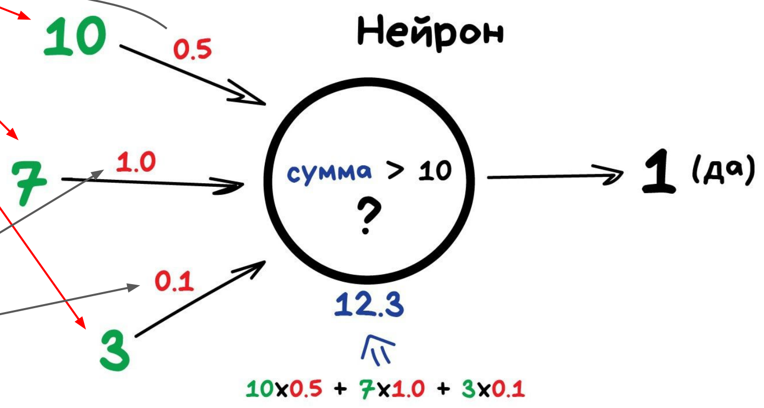
- Все эти операции так или иначе необходимы для теоретического понимания матричного исчисления
- Большая часть операций так или иначе используется в нейросетях

Пример: матрица в нейронной сети

Опишем нейронную сеть в терминах матриц!

Модель нейрона

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

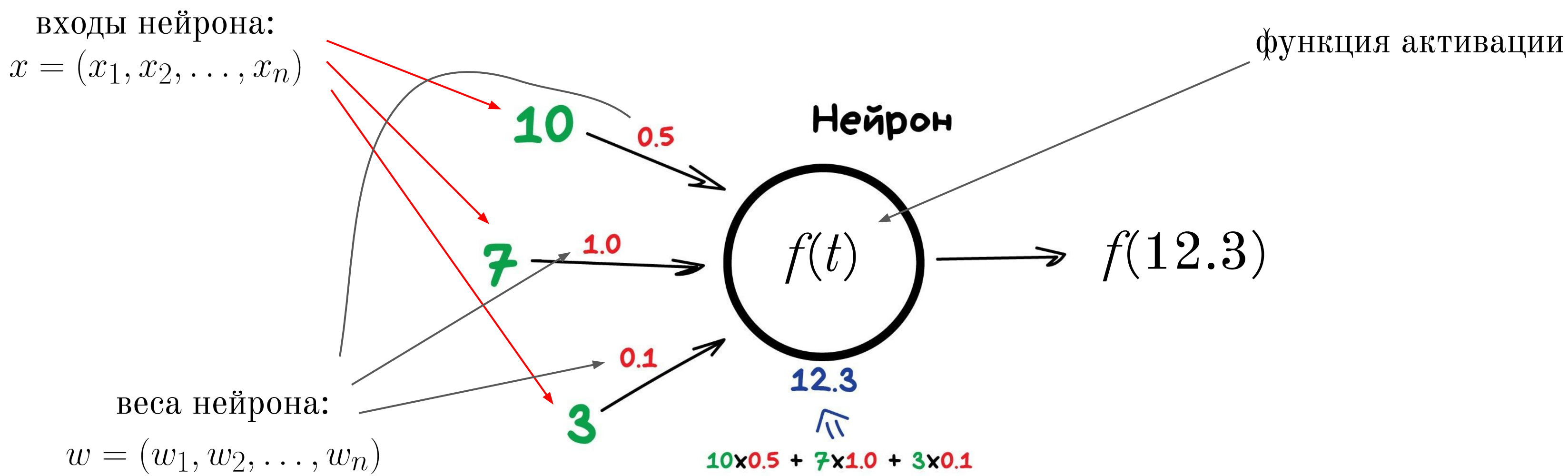


веса нейрона:
$w = (w_1, w_2, \ldots, w_n)$

скалярное произведение векторов x, w:

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

Модель нейрона

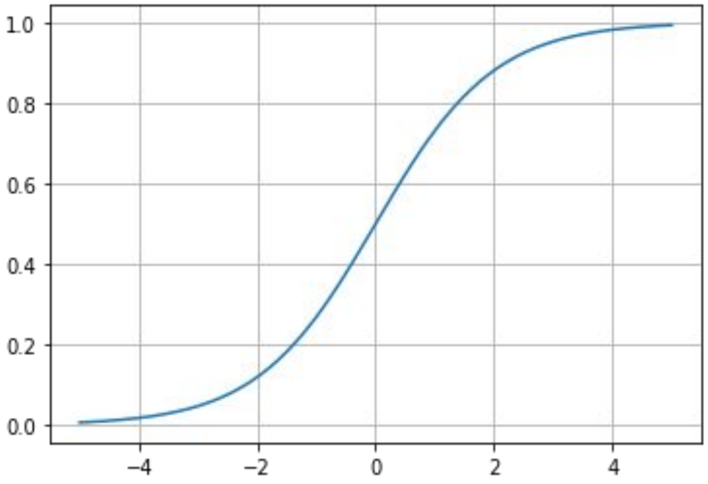


скалярное произведение векторов x, w:

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

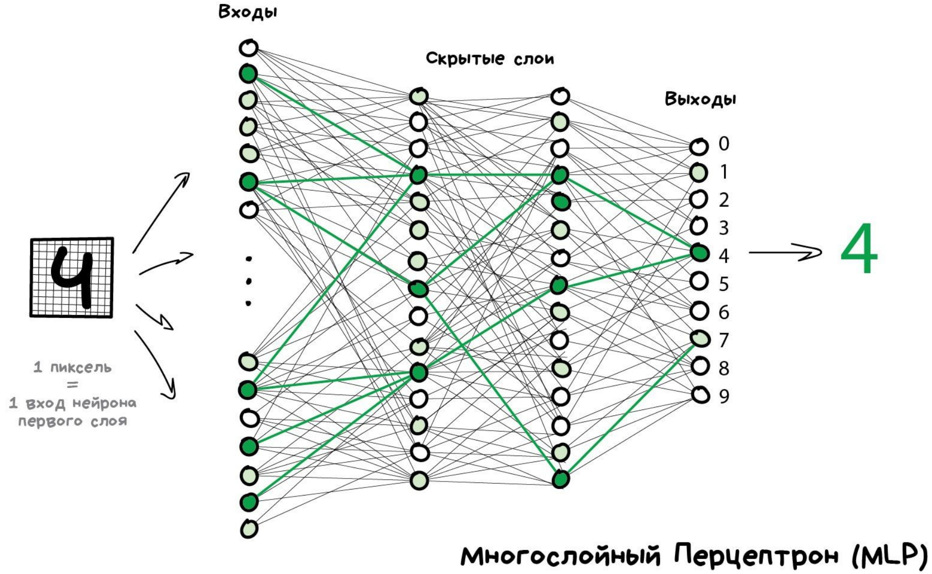
Функция сигмоиды

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

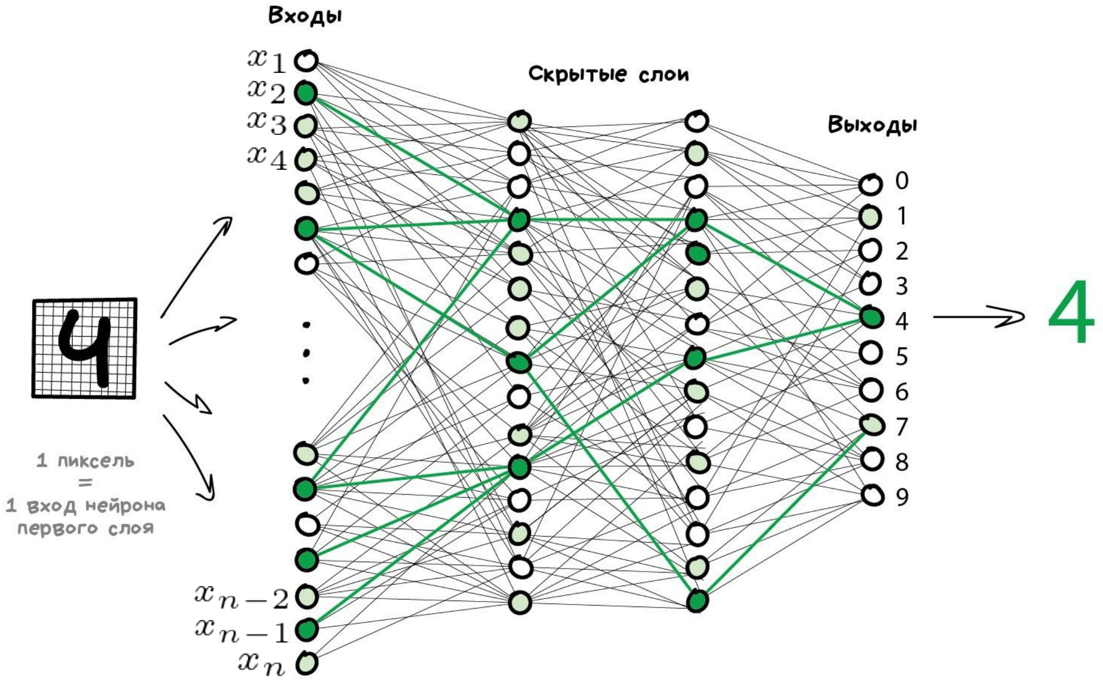


Многослойный перцептрон

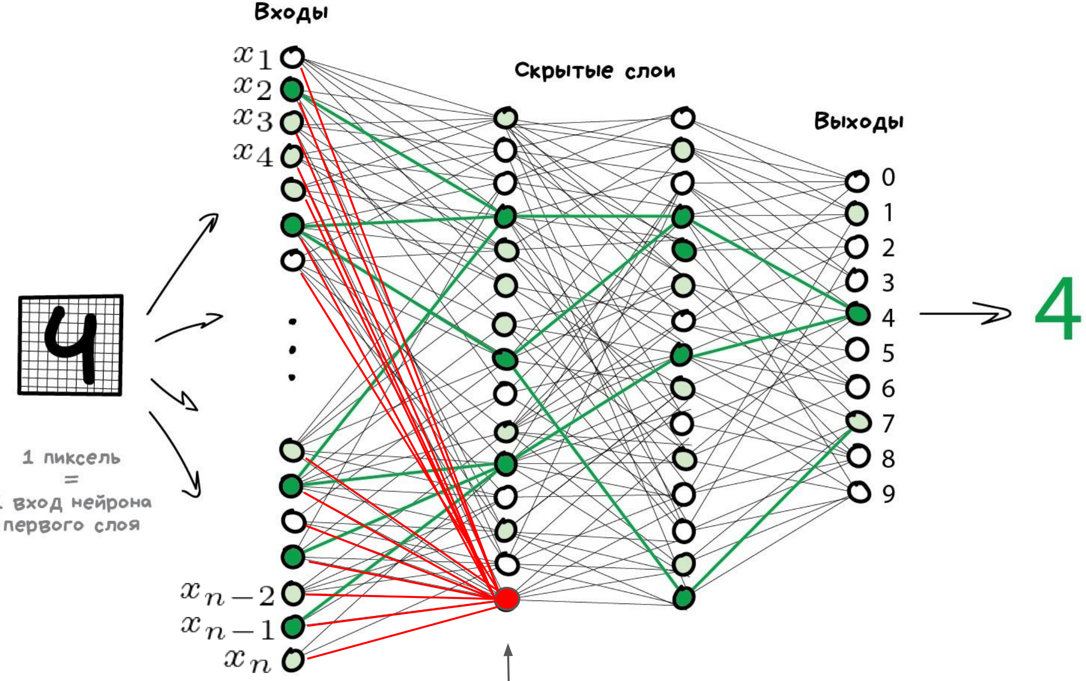
- Многослойный перцептрон — простейшая архитектура нейронной сети
- Каждый слой нейронов связан со всем нейронами с предыдущего слоя
- Выходные нейроны соответствуют классам изображений



Многослойный перцептрон



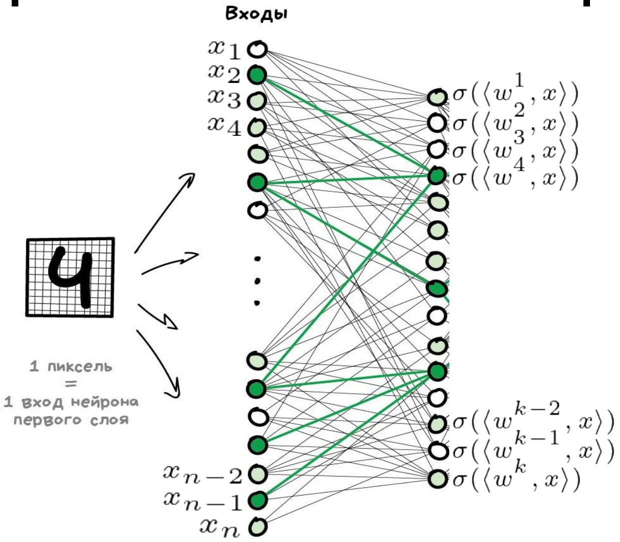
Многослойный перцептрон



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

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

Q&A

Спасибо за внимание!