

```
class GDLinearRegression:
    def __init__(self, learning_rate=0.01, tolerance=1e-8):
        self.learning_rate = learning_rate
        self.tolerance = tolerance

    def fit(self, X, y):
        n_samples, n_features = X.shape
        self.bias, self.weights = 0, np.zeros(n_features)
        previous_db, previous_dw = 0, np.zeros(n_features)

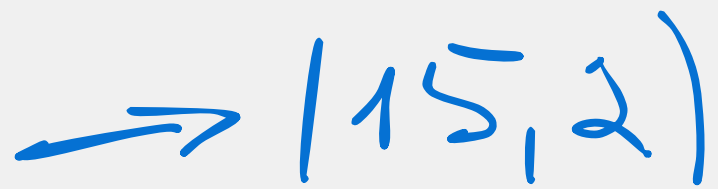
        while True:
            y_pred = X @ self.weights + self.bias
            db = 1 / n_samples * np.sum(y_pred - y)
            dw = 1 / n_samples * X.T @ (y_pred - y)
            self.bias -= self.learning_rate * db
            self.weights -= self.learning_rate * dw

            abs_db_reduction = np.abs(db - previous_db)
            abs_dw_reduction = np.abs(dw - previous_dw)

            if abs_db_reduction < self.tolerance:
                if abs_dw_reduction.all() < self.tolerance:
                    break

            previous_db = db
            previous_dw = dw

        def predict(self, X_test):
            return X_test @ self.weights + self.bias
```

  $(15, 2)$

$\cdot$  array([0., 0.]

# age	# experience	# income
25	1	30450
30	3	35670
47	2	31580
32	5	40130
43	10	47830
51	7	41630
28	5	41340
33	4	37650
37	5	40250
39	8	45150
29	1	27840
47	9	46110
54	5	36720
51	4	34800
44	12	51300
41	6	38900
58	17	63600
23	1	30870
44	9	44190
37	10	48700

```
df_path = "v"
income = pd.read_csv(df_path)
X1, y1 = income.iloc[:, :-1].values, income.iloc[:, -1].values
X1_scaled = scale(X1)
X1_train, X1_test, y1_train, y1_test = train_test_split(X1, y1, random_state=0)
X1_train_s, X1_test_s, y1_train, y1_test = train_test_split(X1_scaled, y1, random_state=0)
print(income)

correlation_matrix = income.corr()
correlation_matrix.style.background_gradient(cmap='coolwarm')

linear_regression = GDLinearRegression()
linear_regression.fit(X1_train_s, y1_train)
```

Градиентный спуск — это численный метод, задачей которого является нахождение локальных экстремумов функции.

MSE

Приближенный численный метод

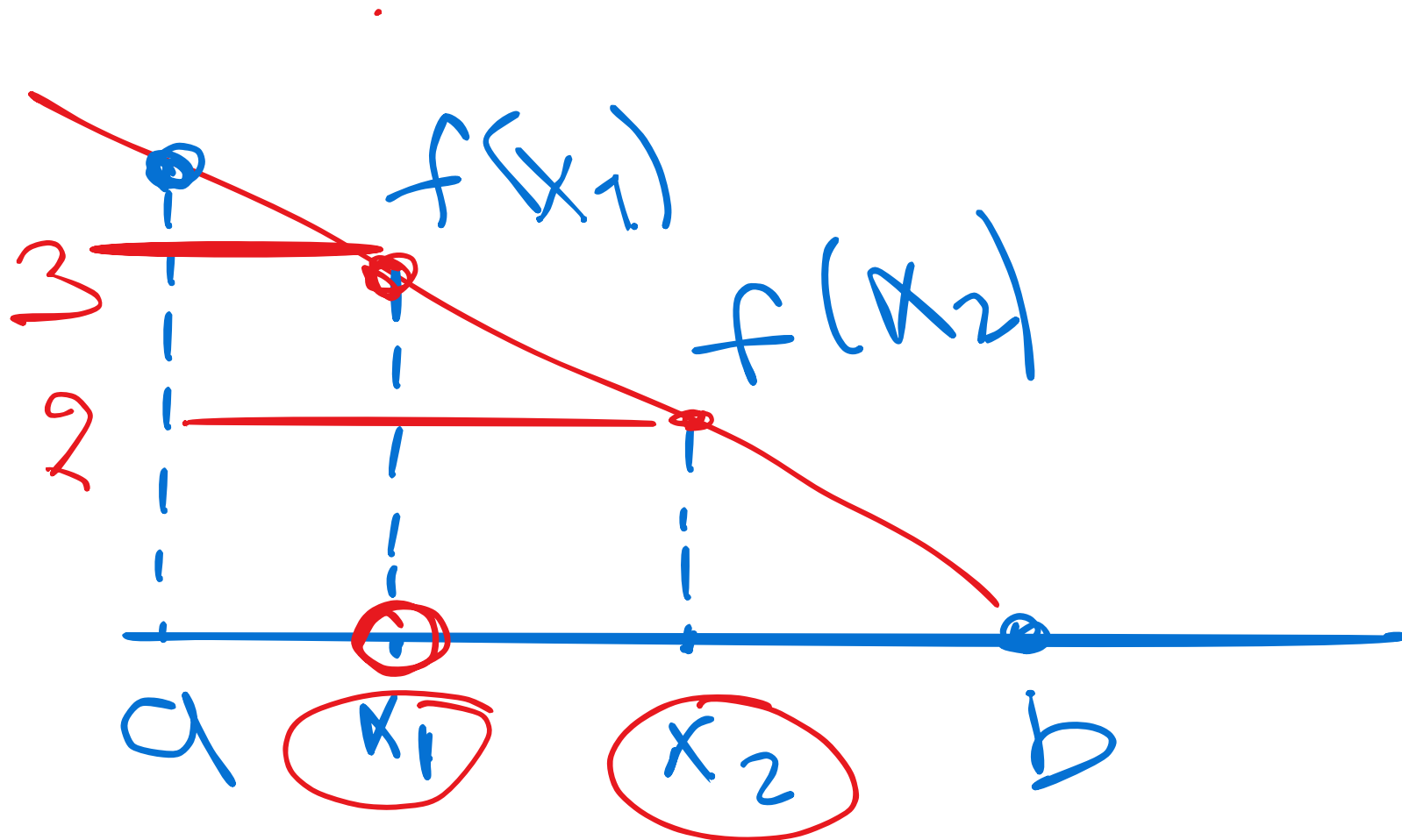
Минимизируемый функционал является гладким и выпуклым, а это значит, что можно эффективно искать точку его минимума с помощью итеративных градиентных методов.

цикл

численный метод это набор техник и подходов для приближённого решения математических задач на компьютере.

## Монотонность функции

Пусть некоторая функция  $f(x)$  задана на промежутке  $[a, b]$ , тогда рассмотрим несколько вариаций её поведения на этом промежутке:

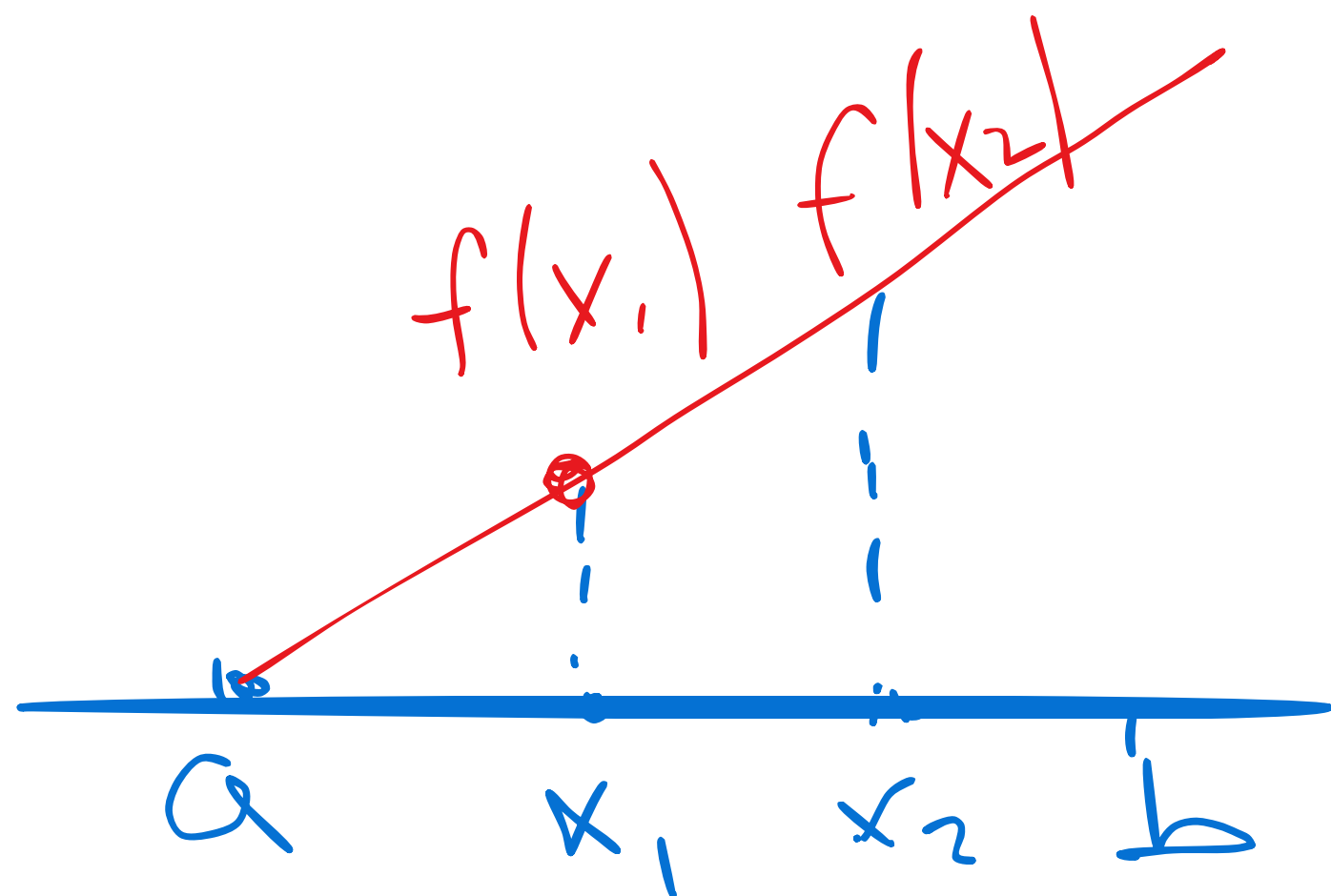


$f(x)$  убывает (строго) на промежутке  $[a, b]$   
если  $\forall x_1, x_2 \in [a, b] \wedge x_1 < x_2$   
 $\Rightarrow f(x_1) > f(x_2)$

Функция ↗ на промежутке  $[a; b]$

$$\neg \forall x_1, x_2 \in [a, b] \wedge x_1 < x_2$$

$$f(x_1) < f(x_2)$$



Logical operation	Logical symbol	Macro
There is	$\exists$	(isE)
There is exactly	$\exists!$	(isE!)
if, then	$\supset$	ifthen
inclus. or	$\vee$	vr*
exclus. or xor	$\oplus$	vr**
and	$\wedge$	a*d
logical. equiv.	$\equiv$	lgeq
material. equiv.	$\Leftrightarrow$	mteq
Because	$\therefore$	b/c*
For all	$\forall$	fral*
proport. to	$\propto$	prpto*
intersection	$\cap$	nxn
union	$\cup$	uni*
tilde	$\sim$	tld*
asymptot. equal	$\approx$	aeql*
is true	$\models$	isT*
is not true	$\not\models$	isnT*
necessarily	$\Box$	ncsry
element of	$\in$	lmnt
inferred from	$\vdash$	infr*

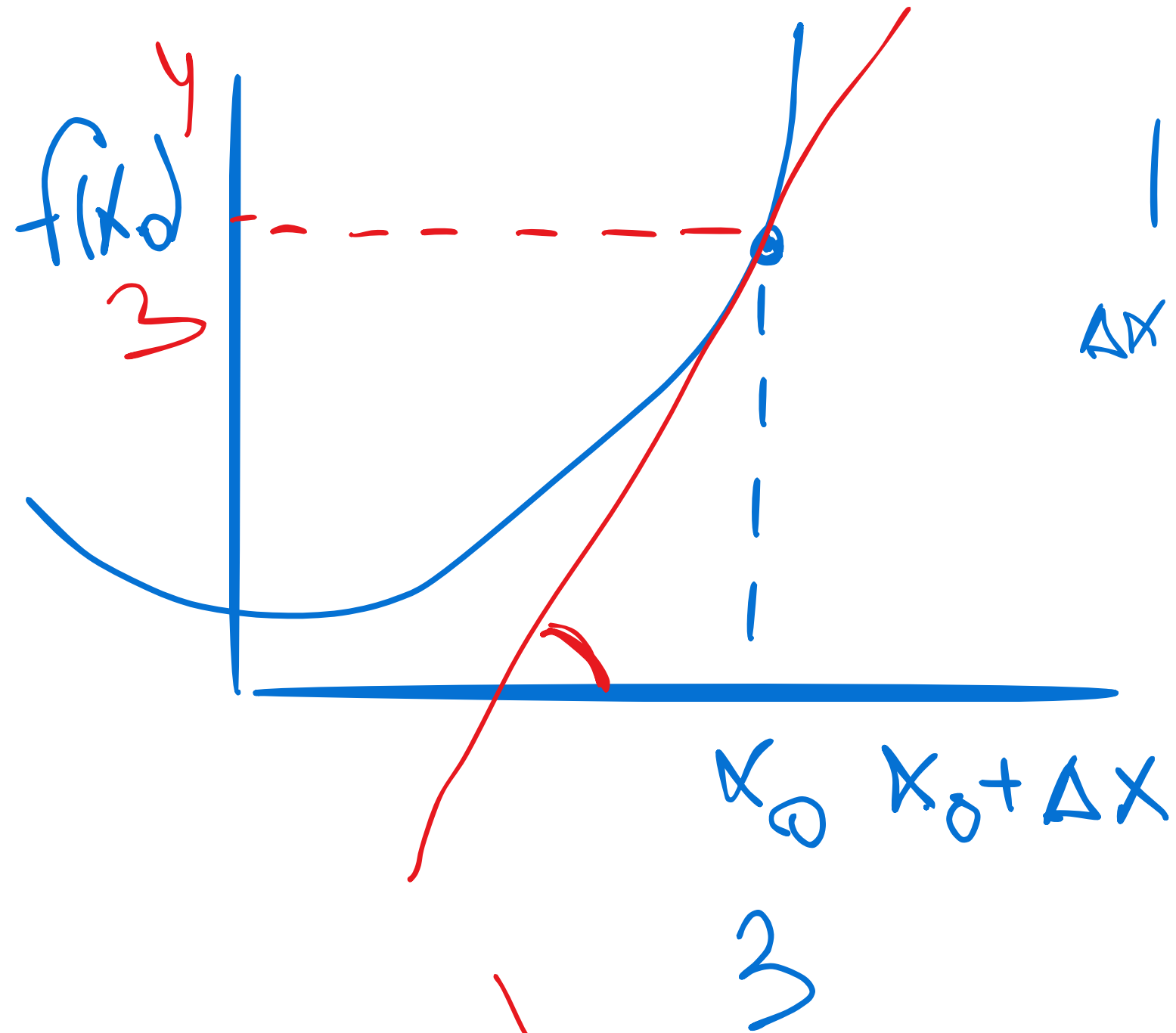
- $\in$  - знак принадлежности (принадлежит);
- $\notin$  - не принадлежит;
- $\subset$  - знак включения, подмножества;
- $\Rightarrow$  - знак следования;
- $\Leftrightarrow$  - знак равносильности;
- $\{ \}$  - знак множества;
- $\{$  - знак системы;
- $[$  - знак совокупности;
- $\cup$  - знак объединения;
- $\cap$  - знак пересечения;
- $\emptyset$  - пустое множество;
- $\forall$  - квантор всеобщности (для любого, каждого);
- $\exists$  - квантор существования (найдется, существует);
- $A$  - аксиома;
- $Th$  - теорема;
- $lh$  - лемма;
- $cl$  - следствие;
- $\blacksquare$  - что и требовалось доказать;
- $\circ$  - определение;
- $\wp$  - пример;
- $\wp$  - утверждение;
- $\sqsupset$  - пусть;
- $\textcircled{z}$  - замечание;
- $\lrcorner$  - если;
- $\textcircled{q}$  - рассмотрим;
- $\vdots$  - деление без остатка;

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta f}{\Delta x} = \lim_{x \rightarrow 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} = f'(x_0)$$

$\Delta f < 0$  when  $f'(x_0) < 0$ , where  $f'(x_0) > 0$

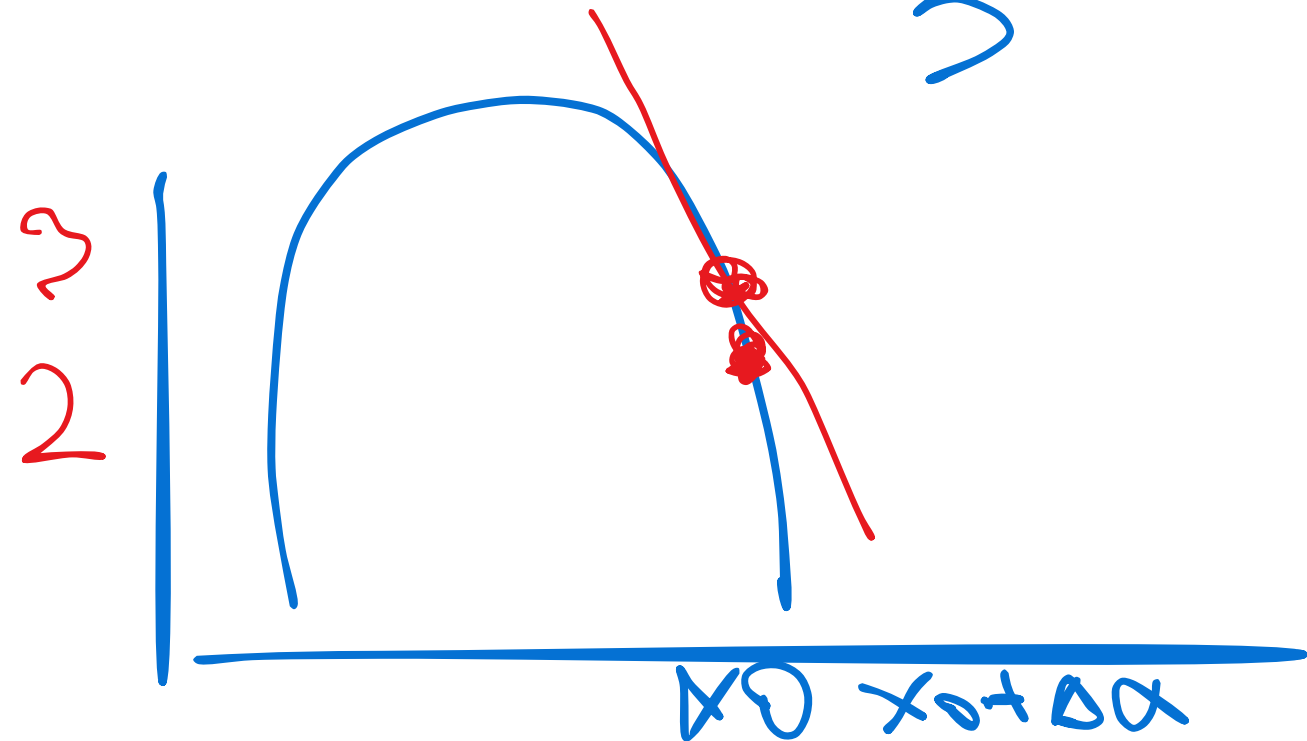
$$f'(x) > 0 \nearrow \quad f'(x) < 0 \searrow$$



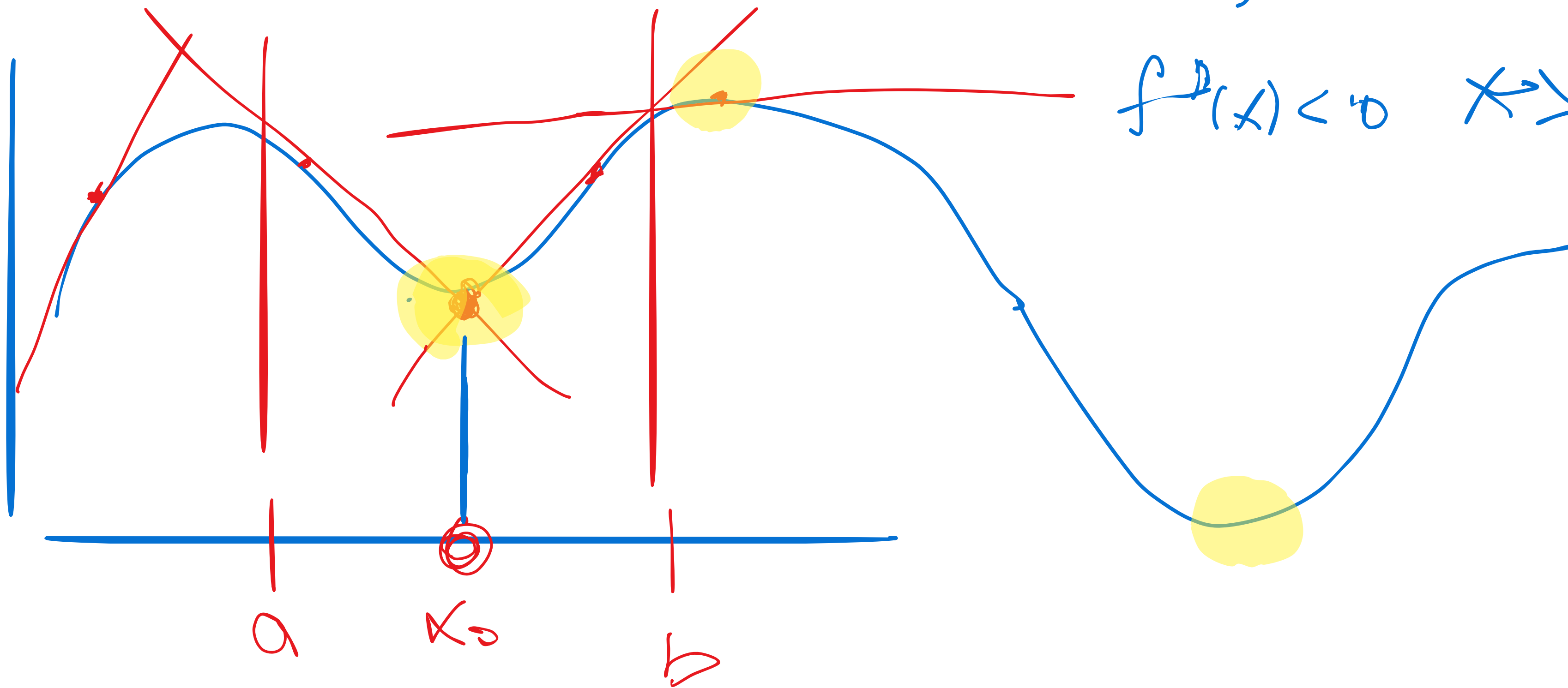


$$\lim_{\Delta x \rightarrow 0} \frac{\Delta f}{\Delta x} = \lim_{x \rightarrow 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}$$

$$\frac{4 - 3}{\Delta x} > 0$$



$$\frac{2 - 3}{\Delta x} < 0$$



$$f'(x) > 0 \quad x < x_0$$

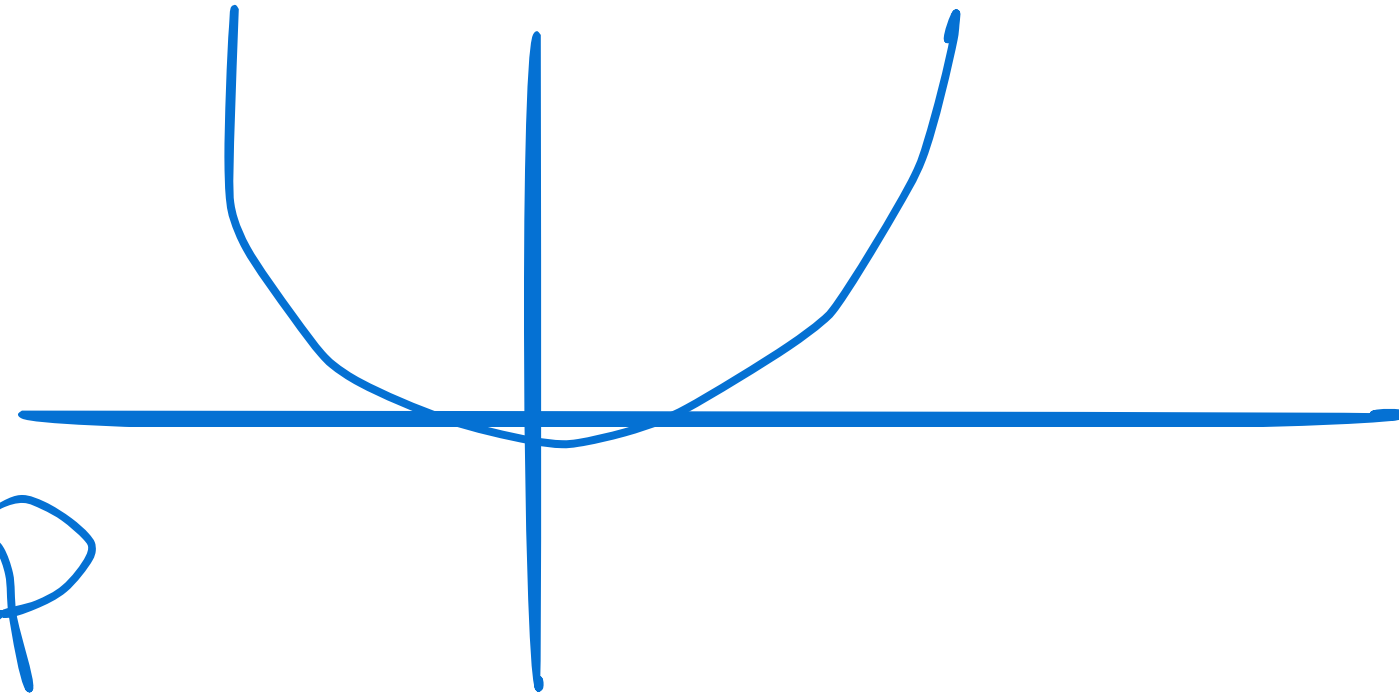
$$f'(x) < 0 \quad x > x_0$$

Теория гашения. УСЗ.



Функция – это соответствие между двумя множествами, при котором каждому элементу одного множества соответствует единственный элемент второго множества

$$y = f(x^2)$$



оооо

$D(f)$

$E(f)$

$f: X \rightarrow Y$



~~\_\_\_\_\_~~

Стандартизация приводит все признаки к одному масштабу, что ускоряет сходимость и делает обучение более стабильным.

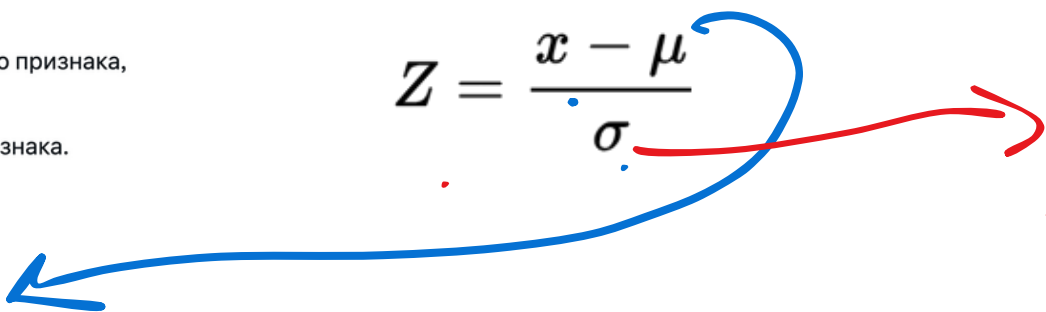
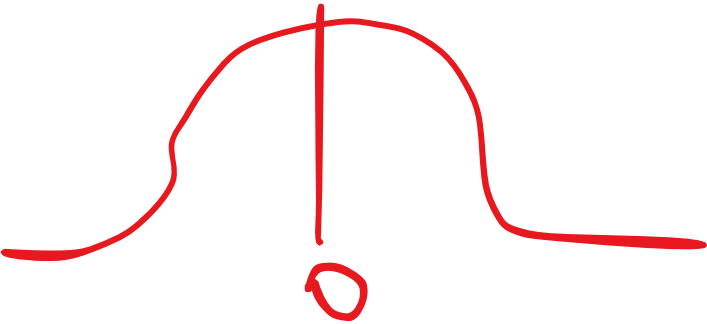
$$X_{\text{scaled},ij} = \frac{X_{ij} - \mu_j}{\sigma_j},$$

где:

- $X_{ij}$  — значение ( $i$ )-го примера для ( $j$ )-го признака,
- $\mu_j$  — среднее значение ( $j$ )-го признака,
- $\sigma_j$  — стандартное отклонение ( $j$ )-го признака.

Функция scale стандартизирует данные, приводя их к виду с нулевым средним и единичной дисперсией (z-score стандартизация):

$$Z = \frac{x - \mu}{\sigma}$$



$\mu = \sum_{i=1}^3 x_i = \frac{25 + 30 + 47}{3} = \frac{102}{3} = 34$

$\sigma = \sqrt{D}$

$D = \frac{\sum (x_i - \mu)^2}{N}$

# age	# experience	# income
25	1	30450
30	3	35670
47	2	31580
32	5	40130
43	10	47830
51	7	41630
28	5	41340
33	4	37650
37	5	40250
39	8	45150
29	1	27840
47	9	46110
54	5	36720
51	4	34800
44	12	51300
41	6	38900
58	17	63600
23	1	30870
44	9	44190
37	10	48700

$$25, 30, 47$$

$$\sigma^2 = \frac{\sum (x_i - \mu)^2}{N} = \frac{\overset{81}{(25-34)^2} + \overset{266}{(30-34)^2} + \overset{169}{(47-34)^2}}{3} =$$

$$\frac{266}{3} \approx 88.6667$$

array([-0.95578964, -0.4247954, 1.38058503])

$$\sigma = \sqrt{\sigma^2} = \sqrt{88.6667} \approx 9.416$$

$$Z = \frac{25 - 34}{9.416} \approx \underline{-0.956} \quad Z = \frac{47 - 34}{9.416} \approx \underline{1.380}$$

$$Z = \frac{30 - 34}{9.416} \approx \underline{-0.425}$$