☆ Partial Derivatives

let $f(x, y) = 2x^2 y + 3y^3 x^2 + 3y$

$\frac{\partial}{\partial x} f(x, y)$ = derivative (differentiation) with respect to x

$$= 4xy + 6y^3 x$$

$$\frac{\partial}{\partial y} f(x, y) = 2x^2 + 9y^2 x^2 + 3$$

$$f(x, y) = 2x^2 + 3y^3x^2 + 3y \quad \text{where } \boxed{y = \underbrace{5-2x}_{g(x)}}$$

$$\frac{\partial}{\partial x} f(x, y) = 2x^2 + 3\, g(x)^3 \cdot x^2$$

Age $\underbrace{\text{height} \quad \text{weight} \quad \text{Metabolism} \quad \boxed{\text{BMI}} \quad \text{Investment} \quad \text{medicksim}}$

$$BMI = \frac{\text{weight}}{(\text{height})^2}$$

in m

$\boxed{\text{Independent}}$

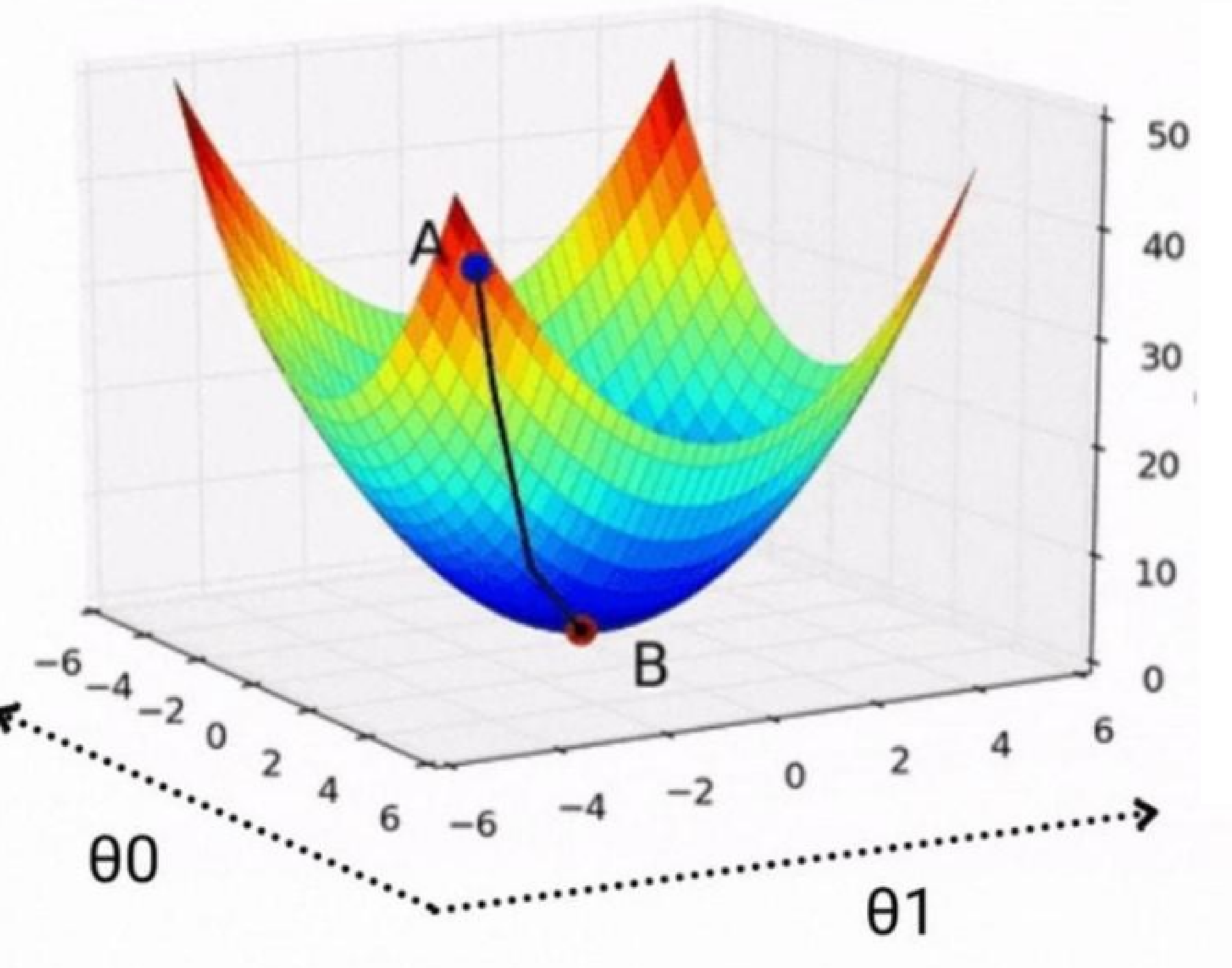
variables

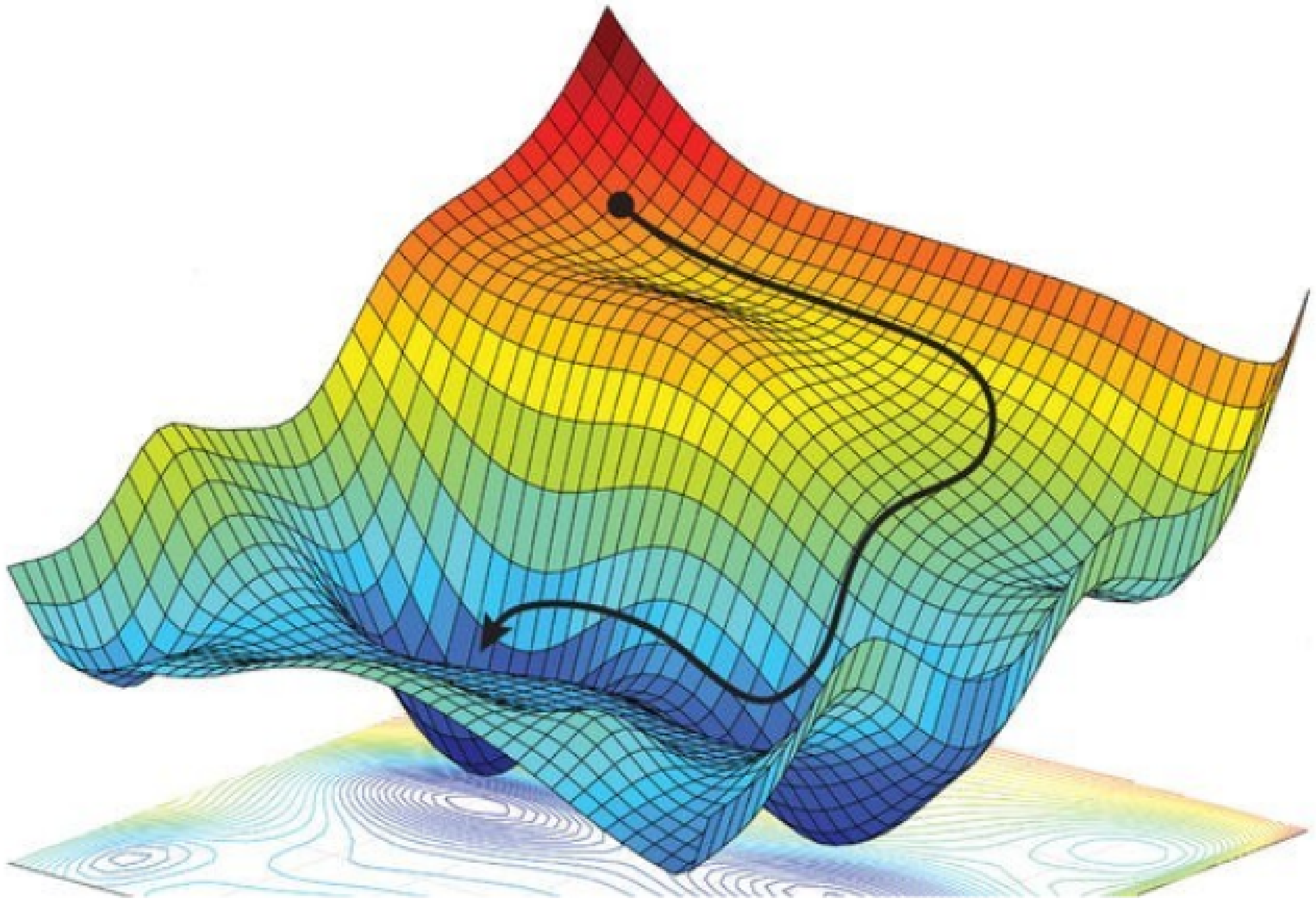
Loss function

$$\underset{w_1, w_2, \ldots w_n, w_0}{\arg\min} \quad L(w_1, w_2, \ldots, w_n, w_0)$$

$$\frac{\partial L}{\partial w_1}, \frac{\partial L}{\partial w_2}, \frac{\partial L}{\partial w_3} \ldots, \frac{\partial L}{\partial w_n}, \frac{\partial L}{\partial w_0}$$

$$\nabla L = \begin{bmatrix} \partial L/\partial w_0 \\ \partial L/\partial w_1 \\ \vdots \\ \partial L/\partial w_n \end{bmatrix} \longrightarrow \text{Gradient}$$


A
B
50
40
30
20
10
0
−6
−4
−2
0
2
4
6
−6
−4
−2
0
2
4
6
θ0
θ1



☆ Learning Rate

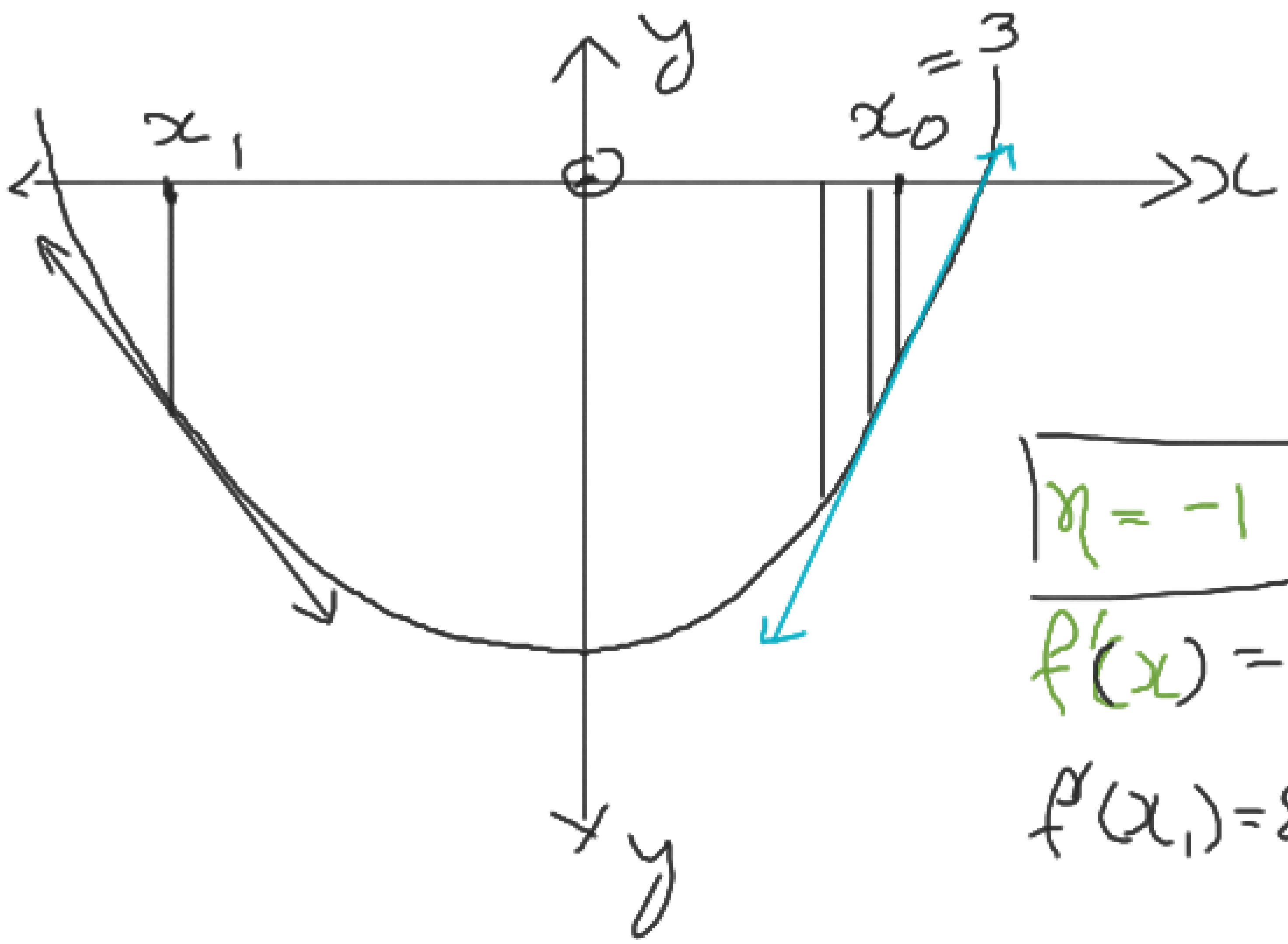
Suppose our $f(x) = x^2 - 30$

$\therefore \frac{\partial y}{\partial x} = 2x = f'(x)$

$f'(x) = 0 \Rightarrow 2x = 0 \Rightarrow x = 0$

$f''(x) = 2 \Rightarrow f''(x) > 0$

$\therefore x = 0$ is minima



$\boxed{\eta = -1}$

$f'(x) = 2(3) = 6$

$f'(x_1) = 2(-3) = -6$

Next guess $x_1 = x_0 + \eta f'(x_0)$

↓

Learning rate

(one of the Hyper Parameters)

$\boxed{\eta = -0.01}$

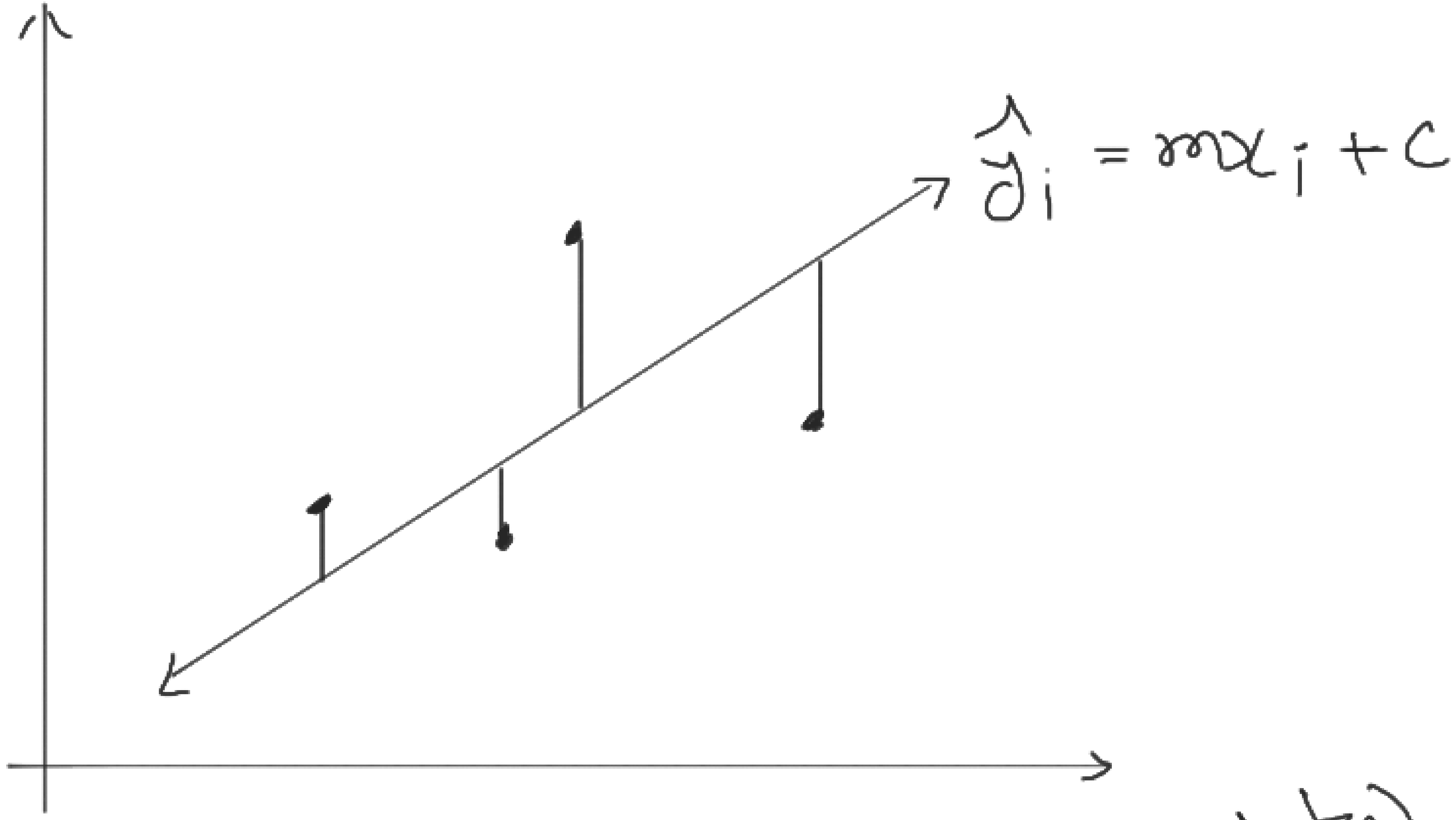
$x_i \rightarrow$	3	2.94
$x_{i+1} \rightarrow$	2.94	2.88

☆ Linear Regression

$\hat{y}_i = m x_i + C$

$\hookrightarrow$ calculated

$y_i \longrightarrow$ actual



$$\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = J(m, c)$$

$$\Rightarrow \quad J(m, c) = \sum_{i=1}^{n} \frac{(y_i - (mx_i + c))^2}{N} \quad \text{(Normalizing the data)}$$

$$\frac{\partial J}{\partial m} = \frac{-2}{N} \sum_{i=1}^{n} x_i (y_i - (mx_i + c))$$

$$\frac{\partial J}{\partial c} = \frac{-2}{N} \sum_{i=1}^{n} (y_i - (mx_i + c))$$

To normalize them (as we used to do in case of z-score)

$$\frac{\partial J}{\partial m} = \frac{-2}{N} \sum_{i=1}^{n} x_i(y_i - (mx_i + c)) \quad \& \quad \frac{\partial J}{\partial c} = \frac{-2}{N} \sum_{i=1}^{n} (y_i - (mx_i + c))$$