

$$\theta = \begin{bmatrix} \theta_0 \\ \vdots \\ \theta_n \end{bmatrix} \quad x_i^j = \begin{bmatrix} x_i^0 \\ x_i^1 \\ \vdots \\ x_i^n \end{bmatrix}$$

\mathbf{x}_i^j \leftarrow append

$$h = \theta^T x_i^j$$

$$= \sum \theta_i x_i^j$$

$$x_i^0 = 1$$

$$\theta = \begin{bmatrix} \theta_0 \\ \theta_1 \end{bmatrix}$$

$$x = \begin{bmatrix} x_0 \\ x_1 \end{bmatrix}$$

$$y = \theta^T x$$

$$= \begin{bmatrix} \theta_0 & \theta_1 \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \end{bmatrix}$$

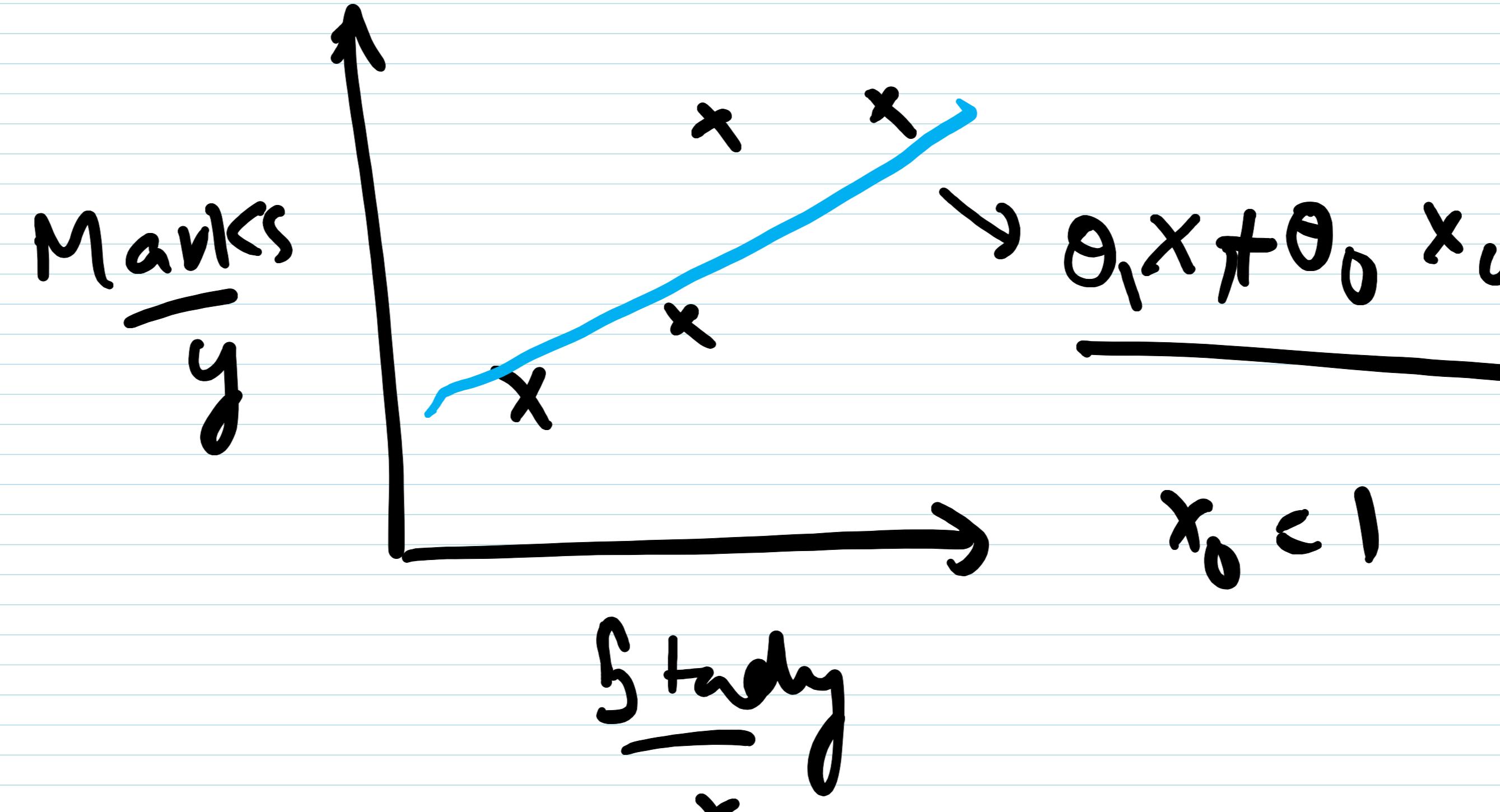
$$= \theta_0 x_0 + \theta_1 x_1$$

$$= \sum \theta_i x_i$$

$$\theta = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

$$x = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$$

$$\underline{\underline{\theta^T x}} = 1.4 + 2.5 + 3.6 = \text{Number}$$

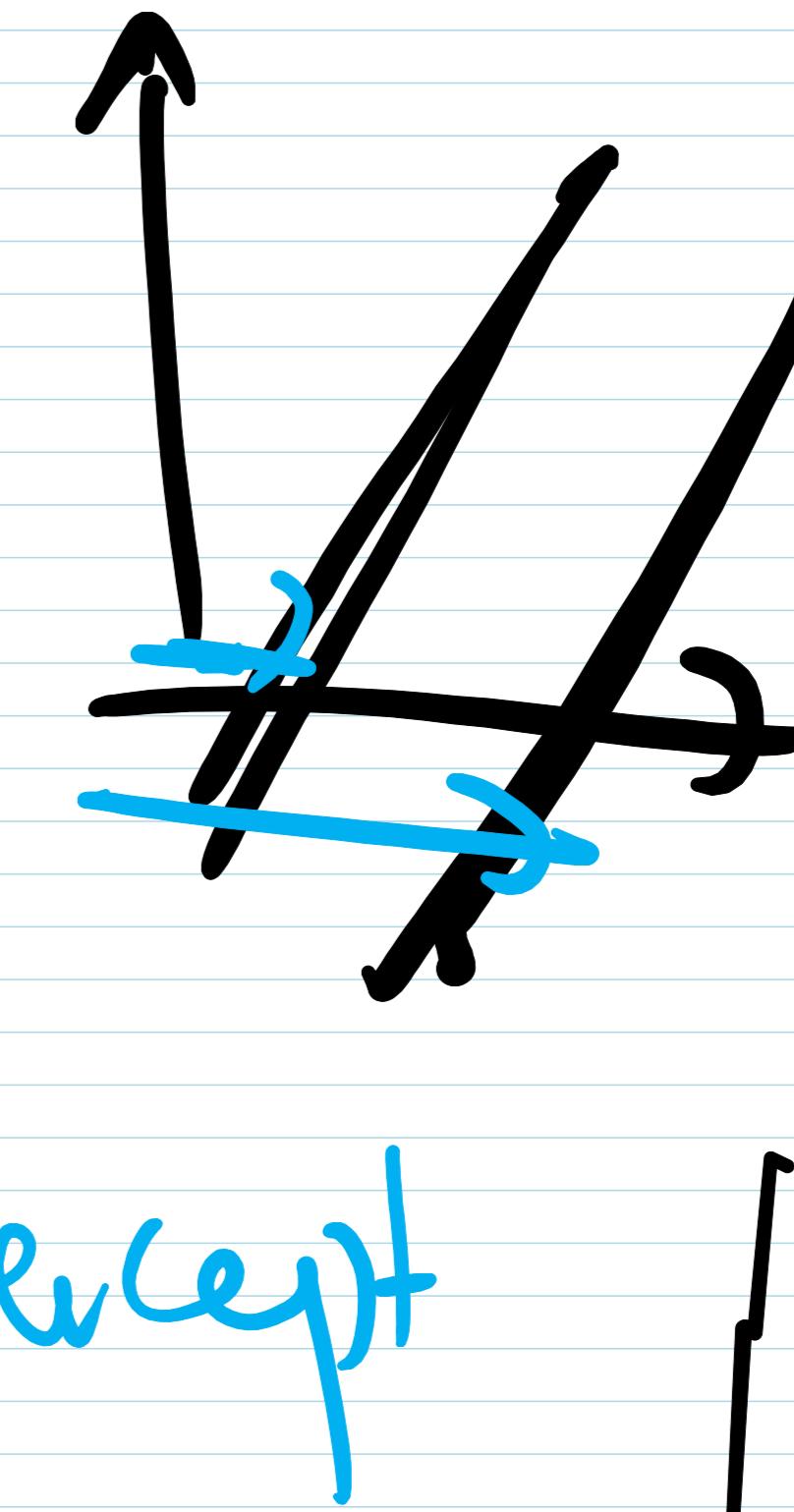


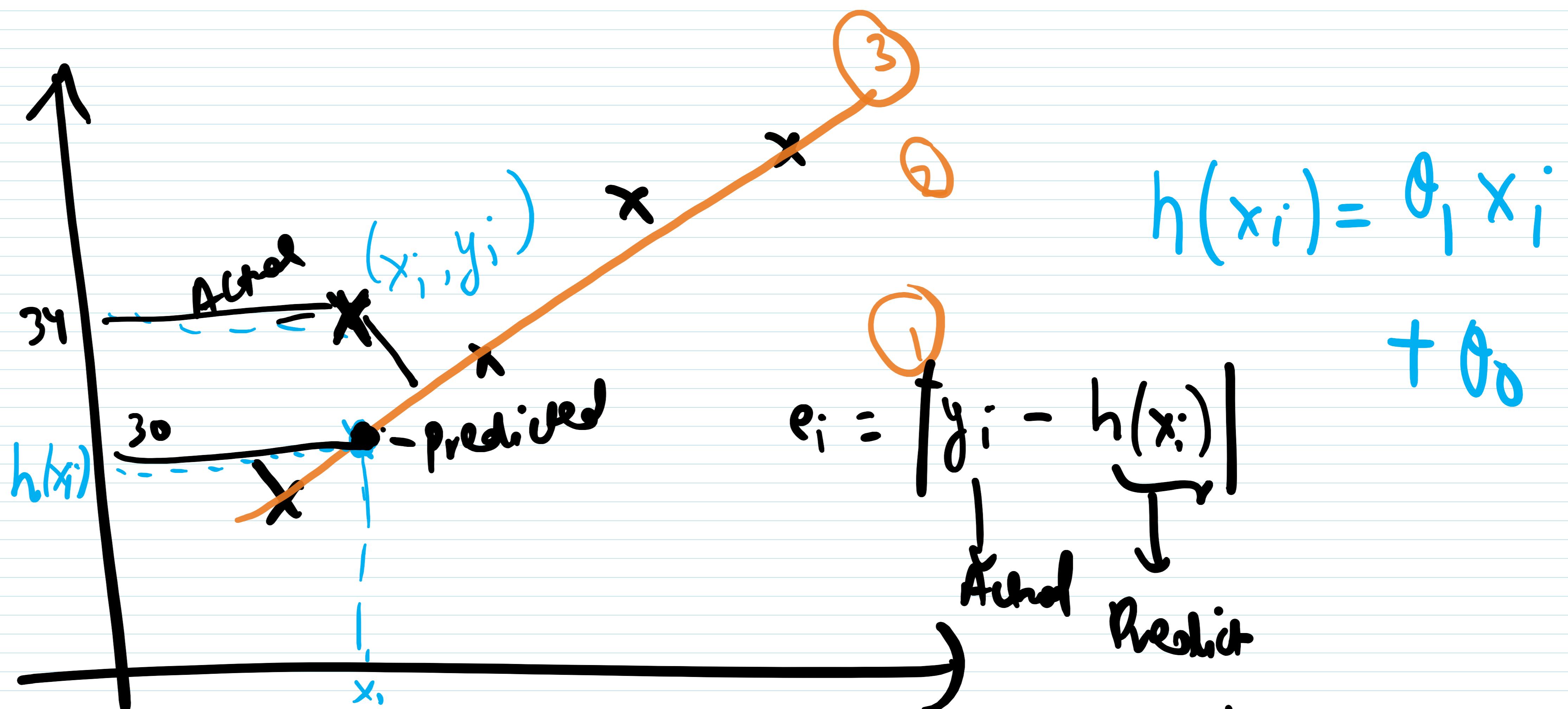
$$y = \theta_1 x_1 + \theta_0 x_0$$

$$\theta = \begin{bmatrix}] \\] \end{bmatrix} \quad x = \begin{bmatrix}] \\] \end{bmatrix}$$

$$y = \underline{\underline{m}} x + \underline{\underline{c}}$$

slope intercept





$$\text{Total } e = \sum_i |y_i - h(x_i)|$$

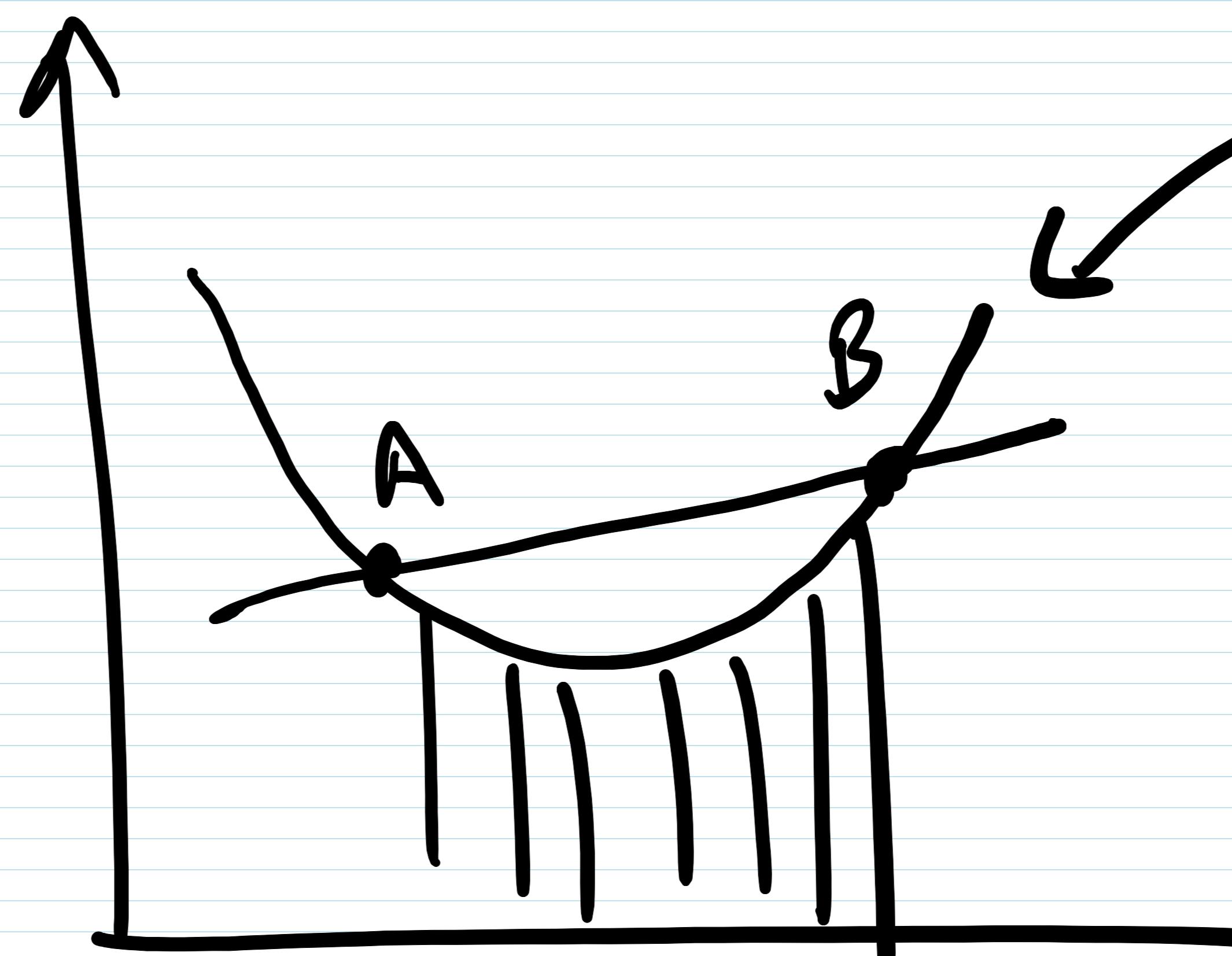
Not
Differ

Loss
Squared = $\frac{1}{2} \sum_i (y_i - h(x_i))^2$

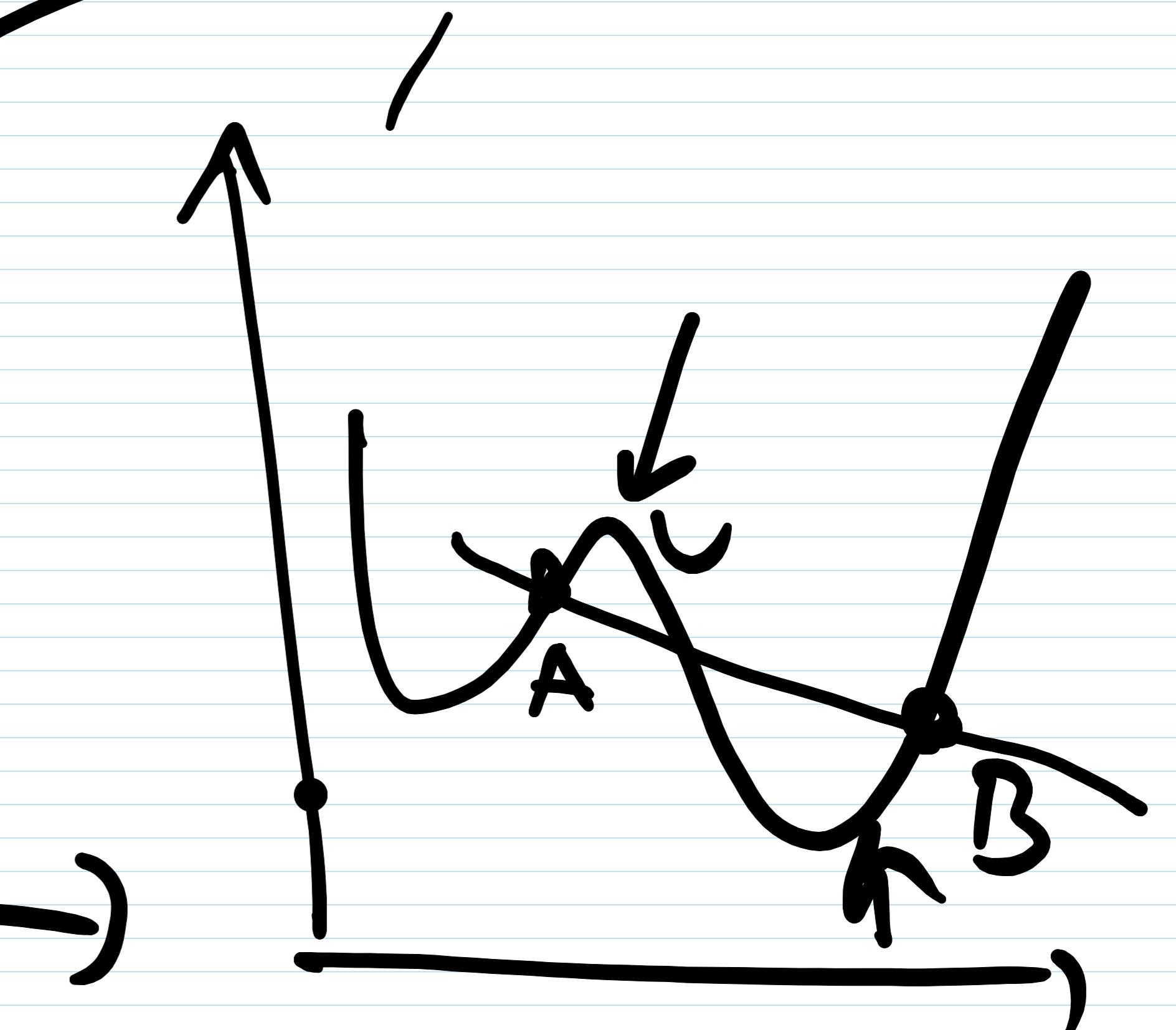
Differentiable

Convex Function

$$\frac{1}{2} (y - h(x_i))^2$$

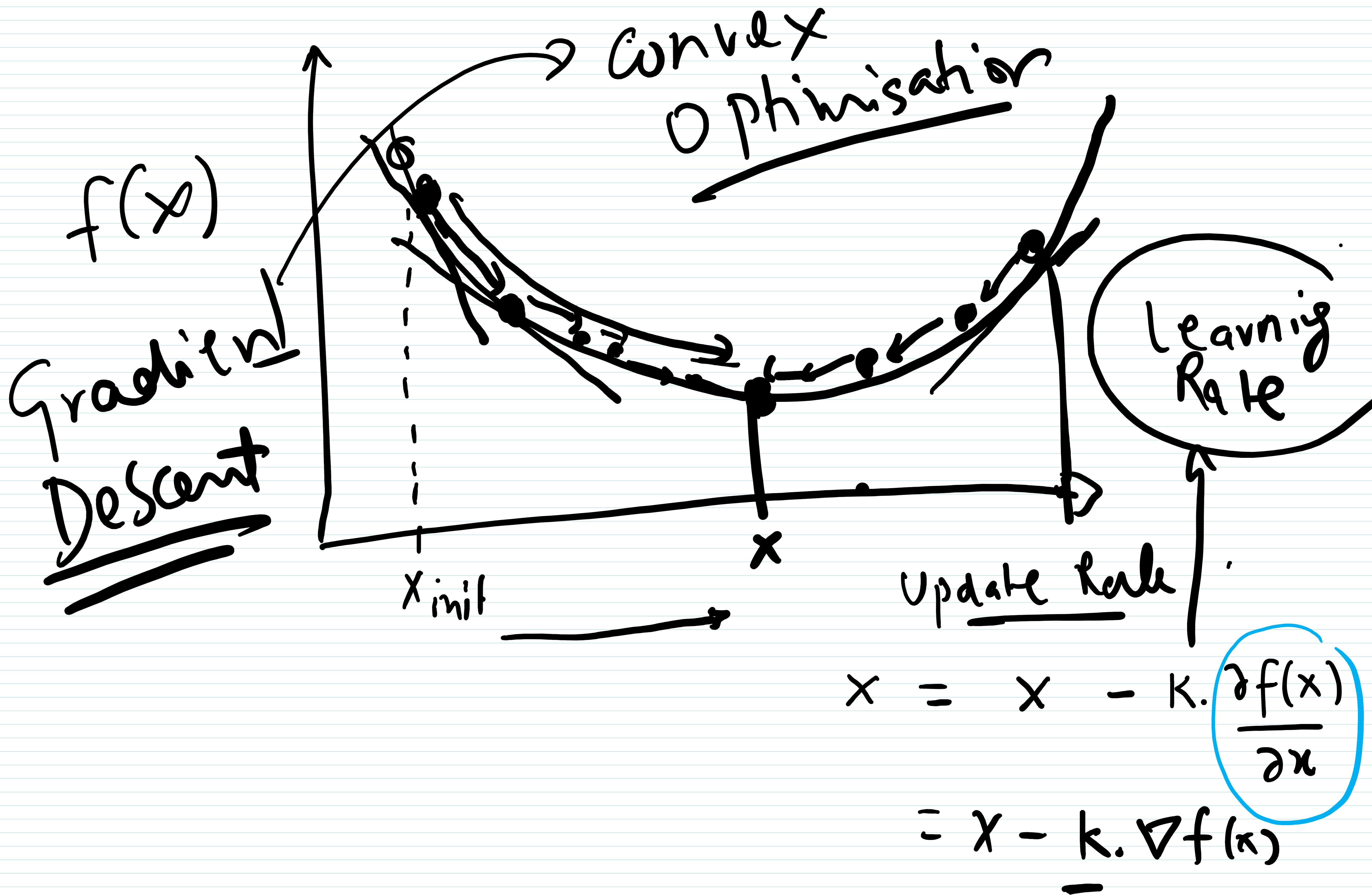


V convex

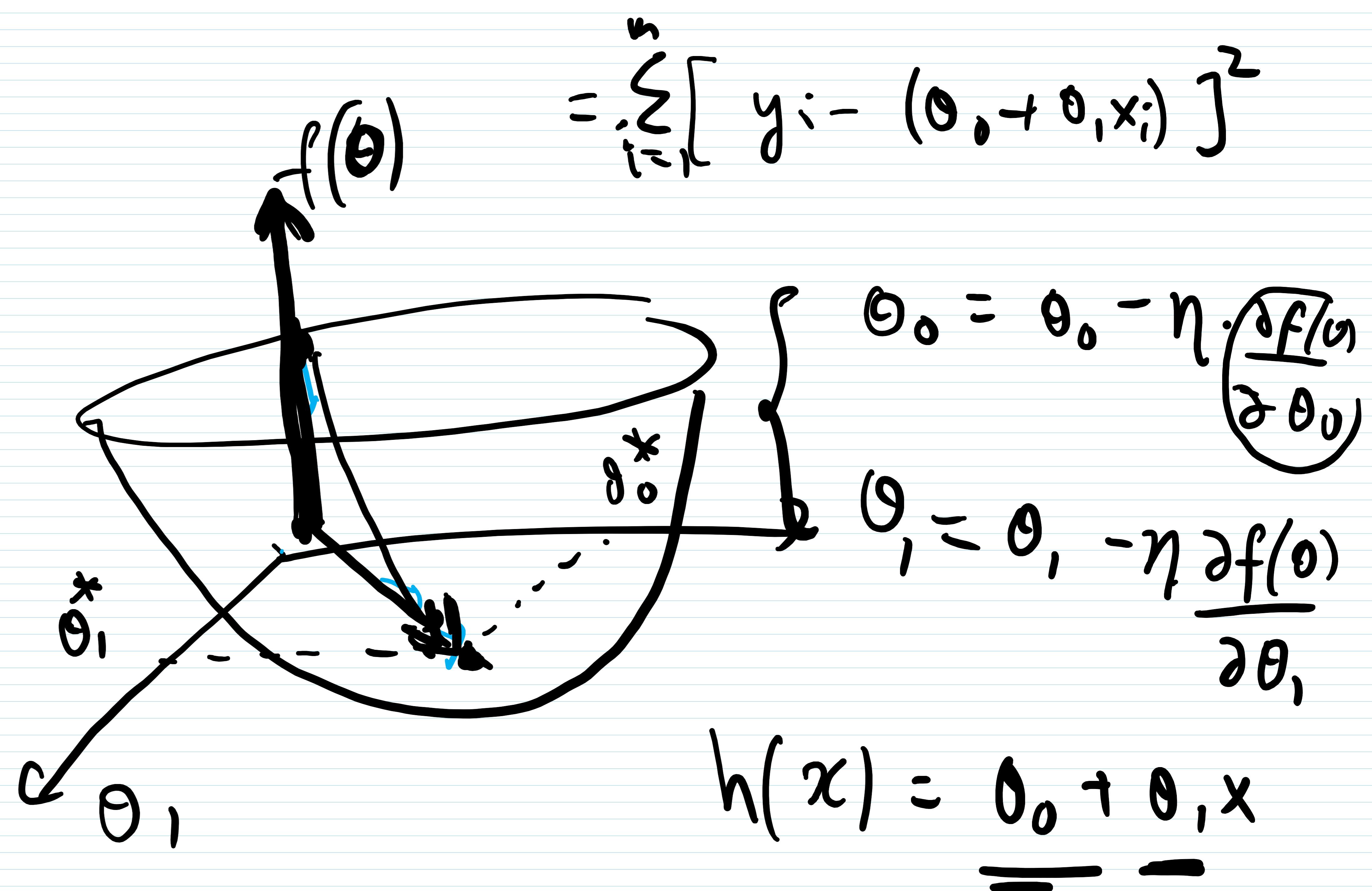
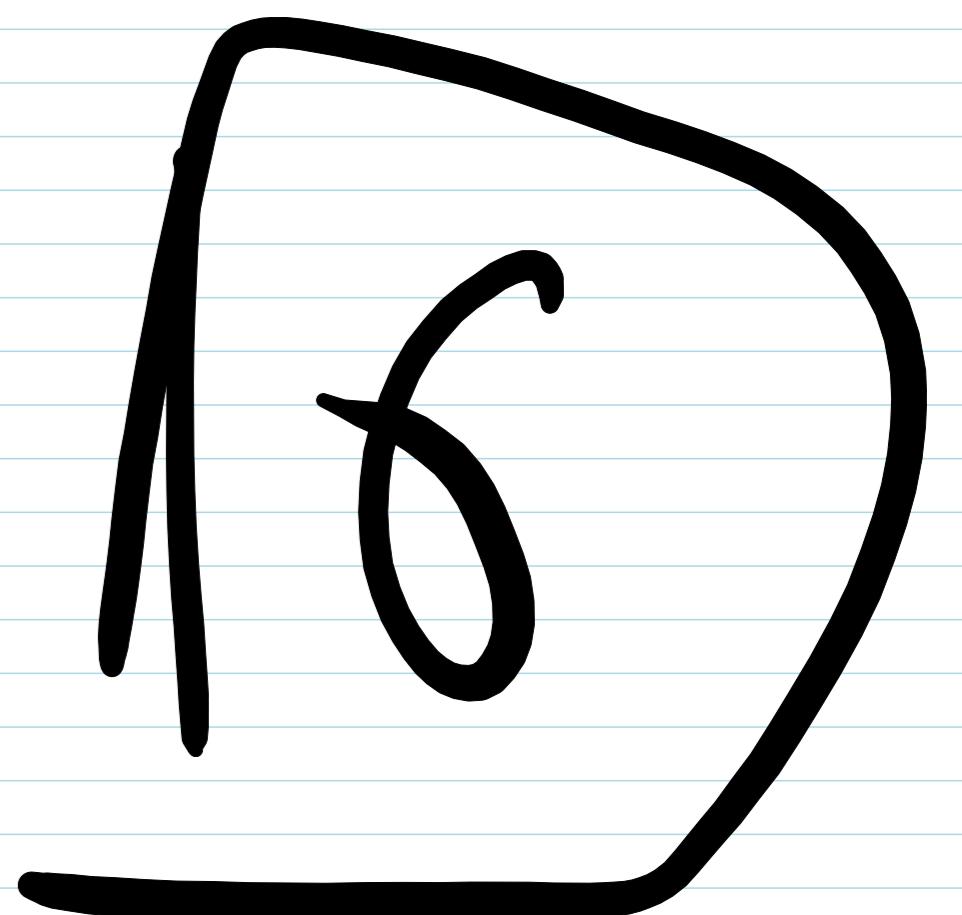


X convex

optimise



$$\text{reshape} \left(\begin{pmatrix} 8 & -1 \end{pmatrix} \right) = \begin{pmatrix} 1, 2, 3, 4 \\ 5, 6, 7, 8 \end{pmatrix}$$



$$\sum_i \frac{1}{2} (y_i - (\theta_0 + \theta_1 x_i))^2$$

$$\frac{\partial f(\theta)}{\partial \theta_0} =$$

$$= \sum_i \cancel{x_i} \left(y_i - (\underline{\theta_0} + \underline{\theta_1} \underline{x_i}) \right)$$

$$\Rightarrow h(x_i) - y_i \quad \leftarrow \text{grad } \theta_0$$

$$\frac{\partial f(\theta)}{\partial \theta_1} =$$

$$= \left(y_i - (\underline{\theta_0} + \underline{\theta_1} \underline{x_i}) \right) \cdot \cancel{x_i}$$

$$\Rightarrow (h(x_i) - y_i) x_i \quad \leftarrow \text{grad } \theta_1$$

