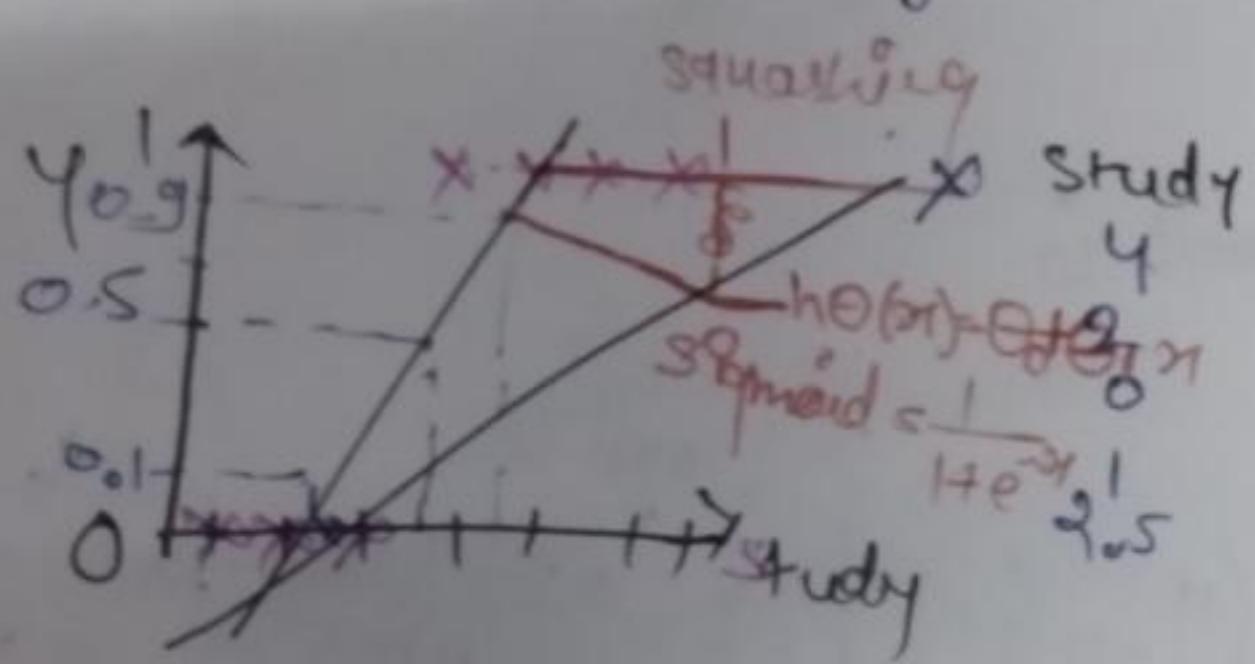


Logistic Regression - A statistical classification algorithm that predicts the probability of a binary outcome (0/1) (Yes/No) using a logistic sigmoid function.

→ output is always 0 and 1



when we try to solve this with linear regression.

(1) Outlier - so best fit line change.

(2) $z > 1$ or < -1 .
we cannot solve classification through linear.

Logic - Means a line gonna have 0 or 1 output shade.

$$h_{\theta}(x) = g(\theta_0 + \theta_1 x)$$

squeezing.
best fit line
 $z = \theta_0 + \theta_1 x$

$$g = \frac{1}{1 + e^{-z}}$$

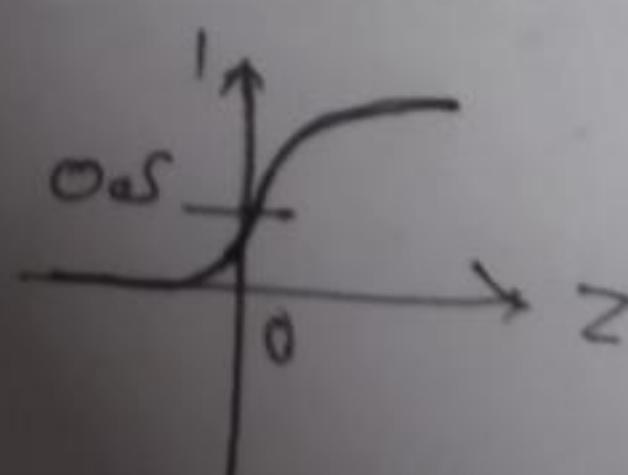
$$h_{\theta}(x) = \frac{1}{1 + e^{-z - \theta_0 + \theta_1 x}}$$

Hypothesis for logistic

$$h_{\theta}(x) = \frac{1}{1 + e^{-(\theta_0 + \theta_1 x)}}$$

Sigmoid $\rightarrow g = \frac{1}{1 + e^{-z}}$

$$\begin{array}{ll} z < 0 & z \geq 0 \\ g(z) < 0.5 & g(z) \geq 0.5 \end{array}$$



train set - $\{(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \dots, (x^m, y^m)\}$

$$x \in \begin{bmatrix} x_0 \\ x_1 \\ x_2 \end{bmatrix} \quad y \in \{0, 1\}$$

$$h_{\theta}(x) = \frac{1}{1+e^{-(\theta_0 + \theta_1 x_1)}} \quad z = \theta_0 + \theta_1 x_1 \\ = \theta^T x.$$

$$h_{\theta}(x) = \frac{1}{1 + e^{-(\theta^T x)}} \rightarrow \text{Hypothesis}$$

Linear Regression

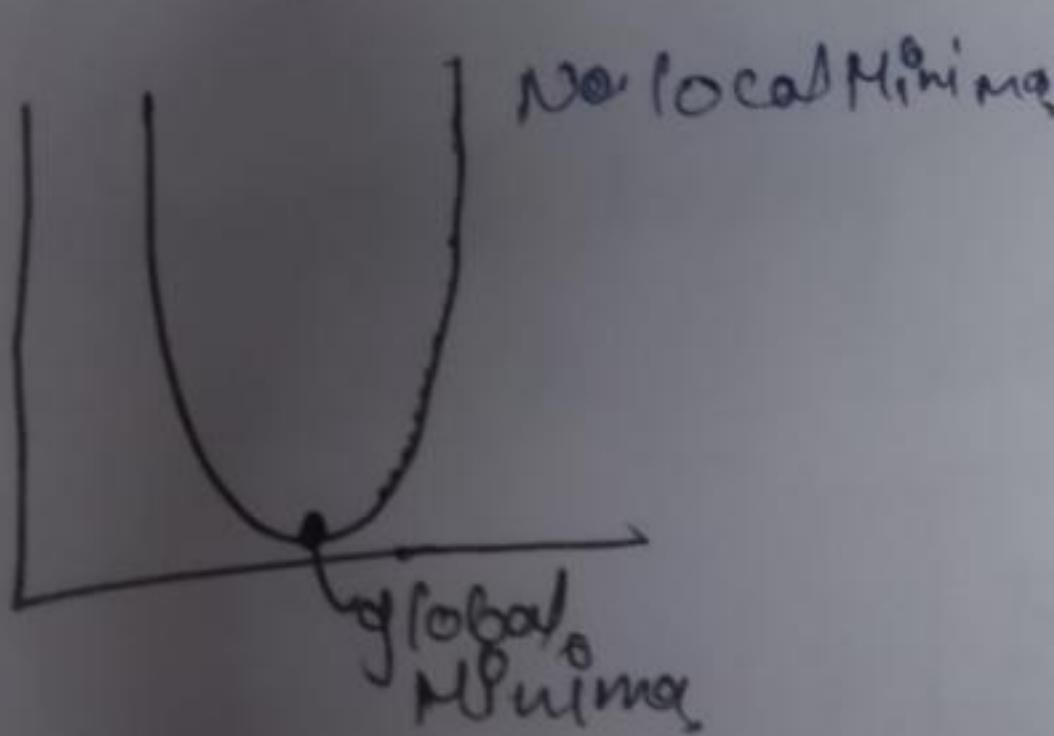
$$\text{Linear Regression} \quad J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^m (h_\theta(x^{(i)}) - y^{(i)})^2$$

$$h\theta(x) = \theta^T x$$

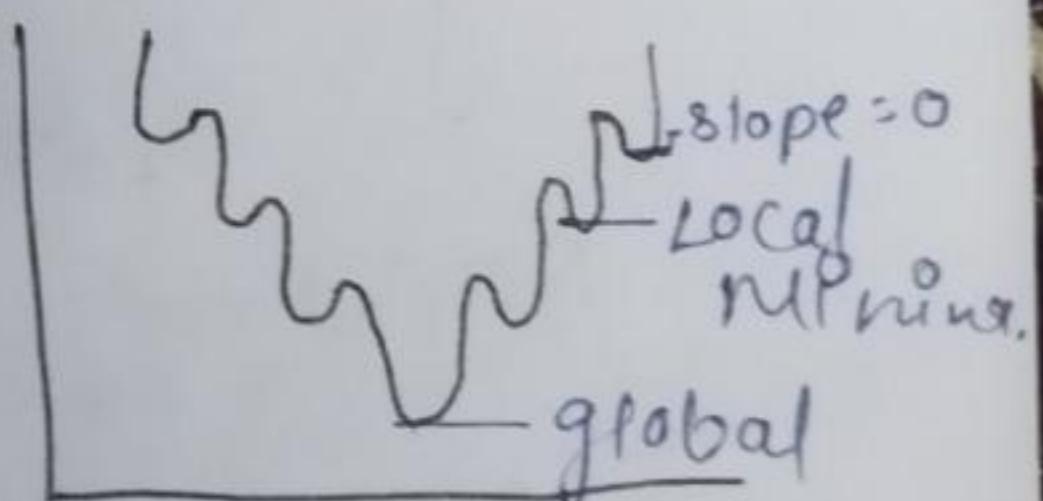
$$= \theta_0 + \theta_1 x$$

$$h_0(x) = \frac{1}{1 + e^{-(\text{const} x)}}$$

Non Convex function



Non Convex



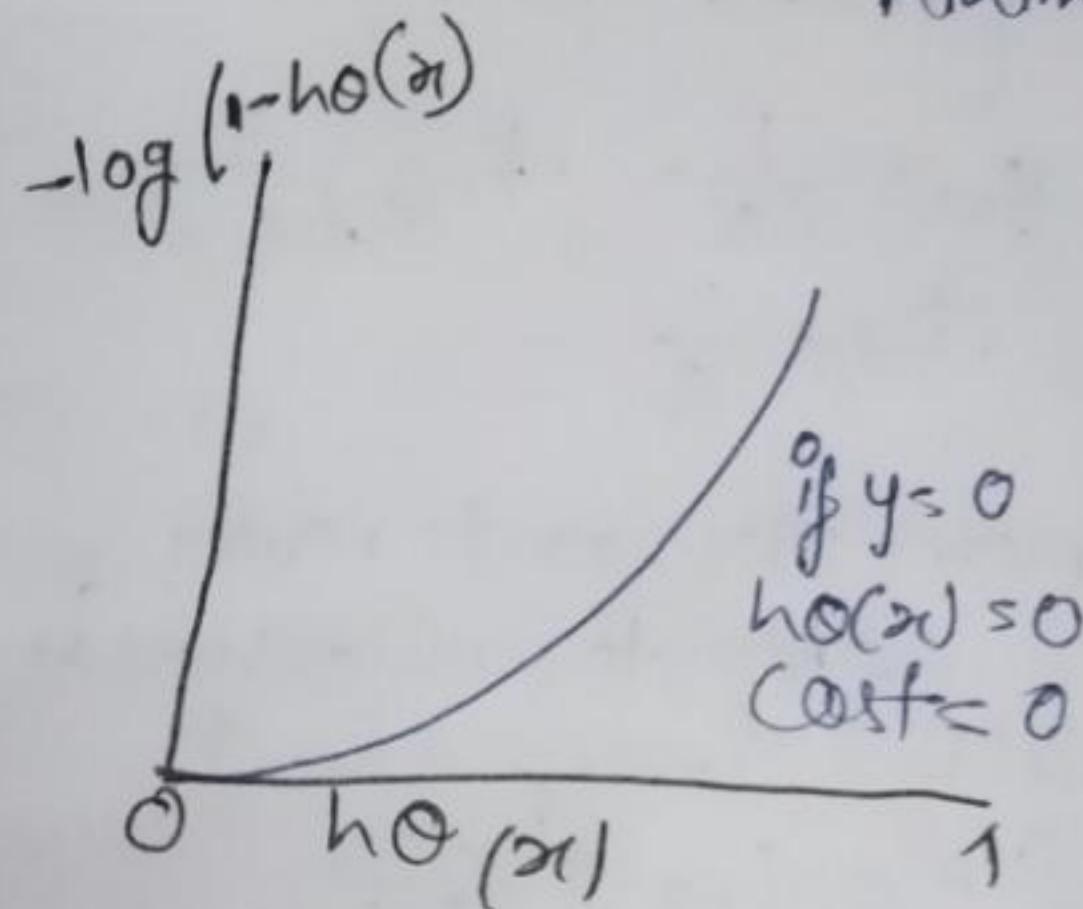
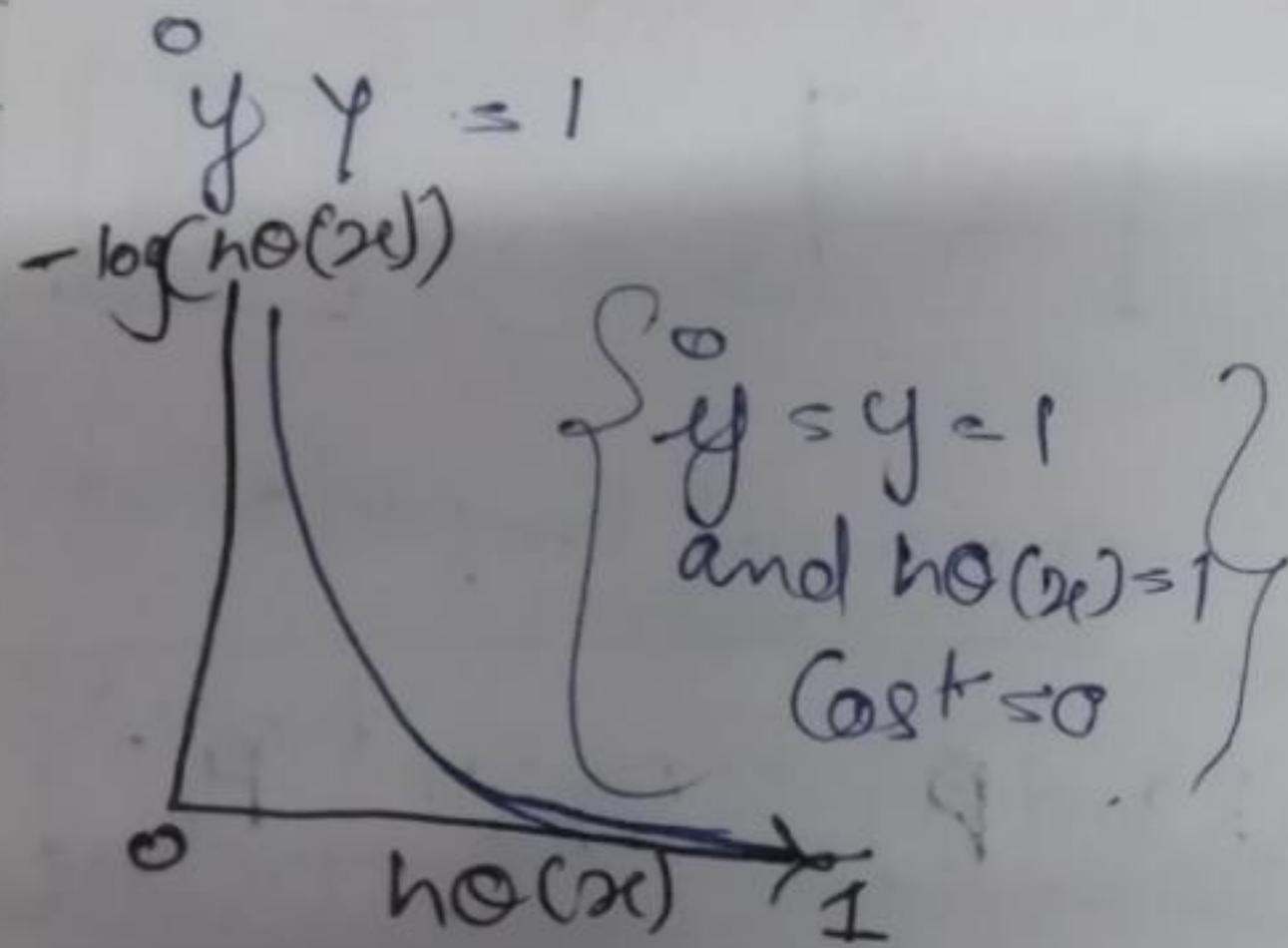
$$\text{Logistic Regression} \quad J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^m (\hat{y}_i - y_i)^2$$

$$h_0(x) = \frac{1}{1+e^z} \text{ (sigmoid function)} \quad \text{Cost}(h_0(x), y)$$

$$\text{Cost}(h_\theta(x), y) = \begin{cases} -\log(h_\theta(x)) & \text{if } y=1 \\ -\log(1-h_\theta(x)) & \text{if } y=0 \end{cases}$$

— Cost function

when we use this representation so, we will get convex curve or single global minima.



$$\text{Cost}(h_\theta(x), y) = -y \log(h_\theta(x)) - (1-y) \log(1-h_\theta(x))$$

if $y=1$
 $= -\log(h_\theta(x))$

if $y=0$
 $= -\log(1-h_\theta(x))$

$$J(\theta_0, \theta_1) = \frac{1}{m} \sum_{i=1}^m g^{(i)} \log h_\theta(x^{(i)}) + (1-g^{(i)}) \log(1-h_\theta(x^{(i)}))$$

↓ Repeat
Convergence Repeat

$$\theta_j \approx \theta_j - \frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1)$$

This convergence theorem yeh stayega curve me global minima kb pochhege.