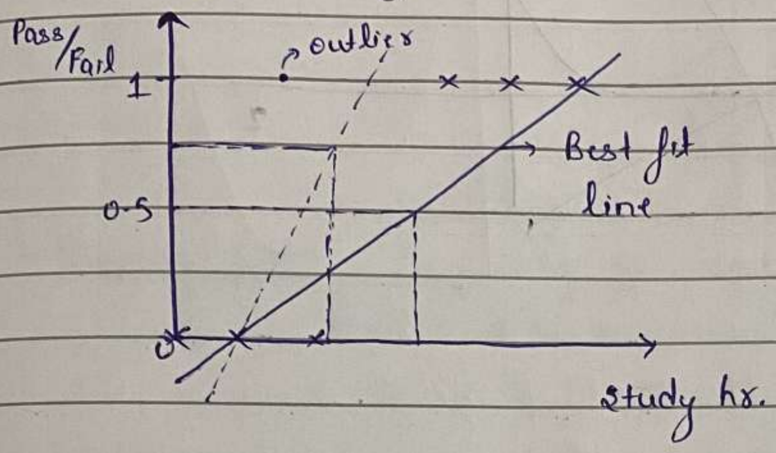
# Logistic Regression

→ Solving classification problem using linear regression




Therefore, we cannot use linear regression to solve classification problem.

- Outliers
- Sometimes predicted value can >1 or <0.

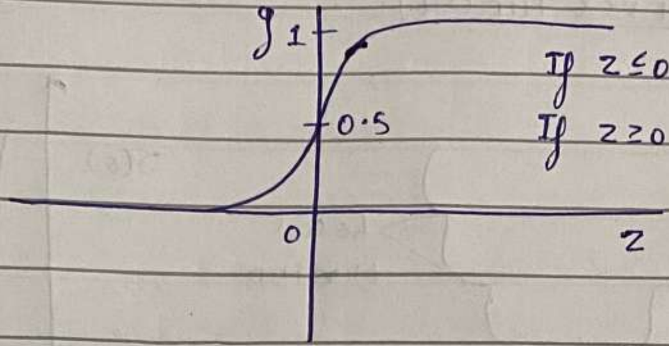
→ Logistic Regression is a supervised machine learning algorithm ~~by fin~~ which is used to solve classification problem by finding a best fit line or best separation line (equation) and applying sigmoid function.

$h_\theta(x) = \theta_0 + \theta_1 x$ } → Linear Regression hypothesis

$$\left.\begin{aligned} Z &= \theta_0 + \theta_1 x \\ g &= \frac{1}{1+e^{-Z}} \\ h_\theta(x) &= \frac{1}{1+e^{-(\theta_0+\theta_1 x)}} \end{aligned}\right\} \rightarrow \text{Logistic Regression Hypothesis}$$

→ **Sigmoid Function Graph :**

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




If $z \le 0 \Rightarrow g(z) \le 0.5$

If $z \ge 0 \Rightarrow g(z) \ge 0.5$

**Training Set** $\{(x_1,y_1), (x_2,y_2), ------, (x_n, y_n)\}$

$x \in \{x_1, x_2, ----, x_n\}$

$y \in \{0, 1\}$

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

$$= \frac{1}{1+e^{-\theta^T x}}$$

* **Cost Function :**

→ Linear Regression :

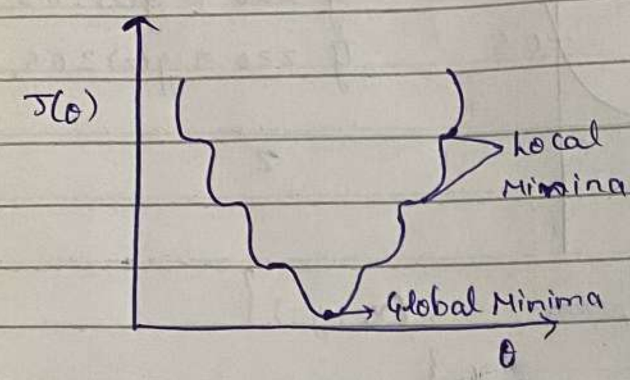
$$J(\theta_0,\theta_1) = \frac{1}{2m}\sum_{i=1}^{m}\left(h_\theta(x)^{(i)} - y^{(i)}\right)^2 \Rightarrow h_\theta(x)^{(i)} = \theta_0 + \theta_1 x$$

→ Logistic Regression :

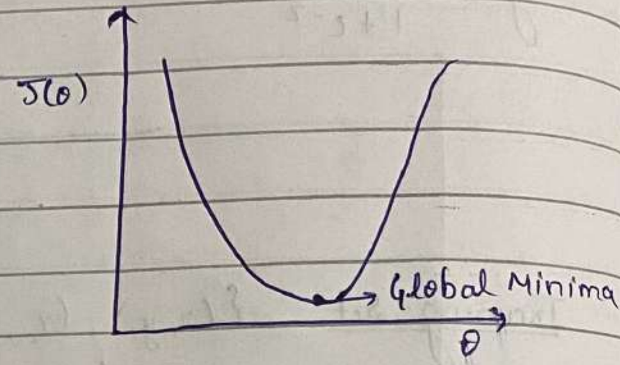
$$J(\theta_0,\theta_1) = \frac{1}{2m}\sum_{i=1}^{m}\left(h_\theta(x)^{(i)} - y^{(i)}\right)^2$$

$$\Rightarrow h_\theta(x) = \frac{1}{1+e^{-\theta^T x}}$$ → Non convex function

## Non convex Function



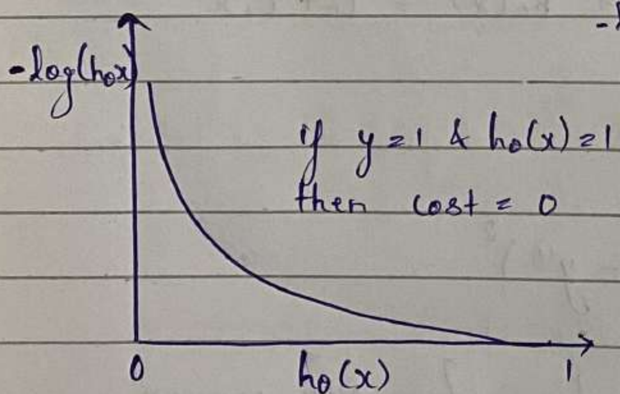

## Convex Function

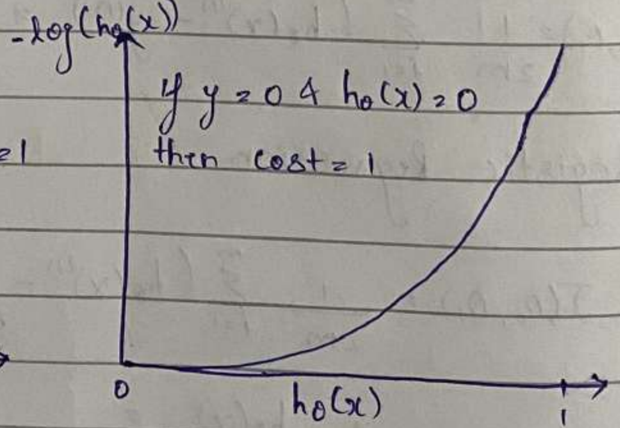



Note: It is very difficult to find global minima in a non-convex function. Therefore, in logistic regression we use another cost function which is a convex function.

$$(h_\theta(x)^{(i)} - y^{(i)})^2 = cost(h_\theta(x), y)$$

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



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

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

$$\boxed{J(\theta_0, \theta_1) = \frac{-1}{2m} \sum_{i=1}^{m} \left( y^i \log(h_\theta(x)^i) + (1-y^i)\log(1-h_\theta(x)^i) \right)}$$

To find the global minima, we use repeat convergence Theorem :

$$\left\{ \quad \theta_j := \theta_j - \alpha \frac{\partial J(\theta_0, \theta_1)}{\partial \theta_j} \quad \right\}$$