Normal eq. for Logistic Regression :

$$\Theta = (X^T X)^{-1} X^T Y \quad \text{(i.e "Slope")}$$

$\boxed{\Theta \approx \text{"Slope"}}$

Θ is the vector of parmeters

X is data (with featuress of all data points)

Y is output/lable

hypothesis eq :

$$h(\Theta) = \Theta_0 x_0 + \Theta_1 x_1 + \Theta_2 x_2 + \Theta_3 x_3 + \cdots + \Theta_m x_m \quad (1)$$

$$h(\Theta) = \Theta^T . x \Leftarrow h_\theta(x)$$

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}, \quad \Theta = \begin{bmatrix} \Theta_0 \\ \vdots \\ \Theta_m \end{bmatrix}, \quad \Theta^T = [\Theta_0, \cdots \Theta_m]$$

Cost function :-

$$J(\Theta) = \frac{1}{2m} (h_\theta(x) - y)^T (h_\theta(x) - y_0)$$

$\boxed{\text{sigma form}}$

$$J(\theta) = \frac{1}{2m} \sum_{i=1}^{n} (h_\theta(x^i) - y^{(i)})^2$$

{ x_j^i : i^{th} data ; n : no. of training data ; y^i : i^{th} expected value }

$$J(\Theta) = \begin{bmatrix} h_\theta(x^0) \\ \vdots \\ h_\theta(x^n) \end{bmatrix} - \begin{bmatrix} y_0 \\ y^1 \\ \vdots \\ y^m \end{bmatrix} = \begin{bmatrix} \Theta^T(x^0) \\ \vdots \\ \Theta^T(x^n) \end{bmatrix} - y$$

Using (1)

$$J(\theta) = \begin{bmatrix} \Theta_0 (x_0^0) + \Theta_1 (x_1^0) + \cdots \Theta_n (x_n^0) \\ \vdots \\ \Theta_0 (x^n) + \cdots + \Theta_m (x_m^n) \end{bmatrix} - y$$