ML | Cost function in Logistic Regression

Difficulty Level : Medium • Last Updated : 06 May, 2019

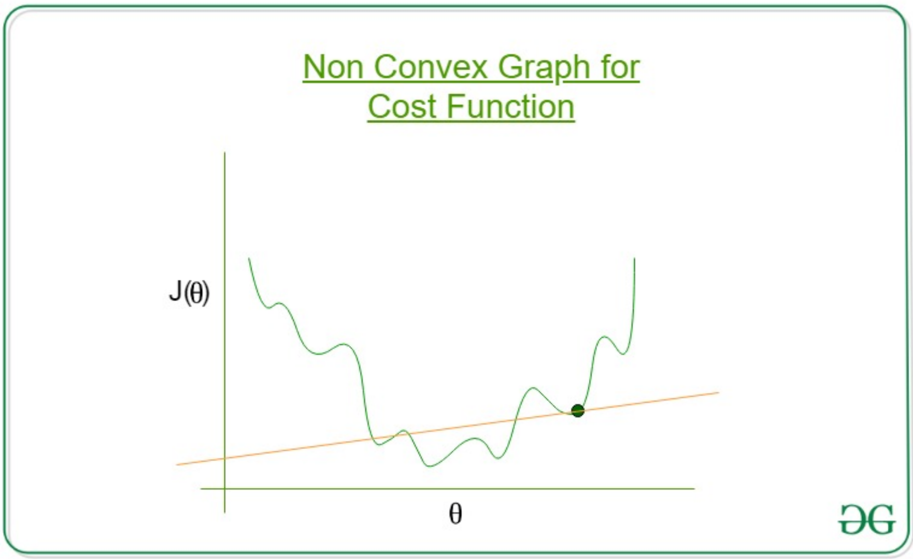
In the case of Linear Regression, the Cost function is –

$$J(\Theta) = \frac{1}{m} \sum_{i=1}^{m} \frac{1}{2} [h_\Theta(x^{(i)}) - y^{(i)}]^2$$

But for Logistic Regression,

$$h_\Theta(x) = g(\Theta^T x)$$

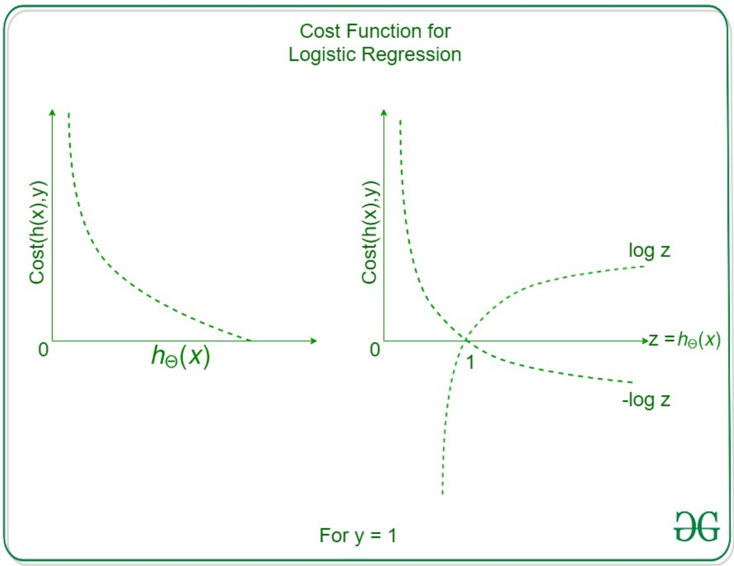
It will result in a non-convex cost function. But this results in cost function with local optima's which is a very big problem for Gradient Descent to compute the global optima.



So, for Logistic Regression the cost function is

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

If y = 1



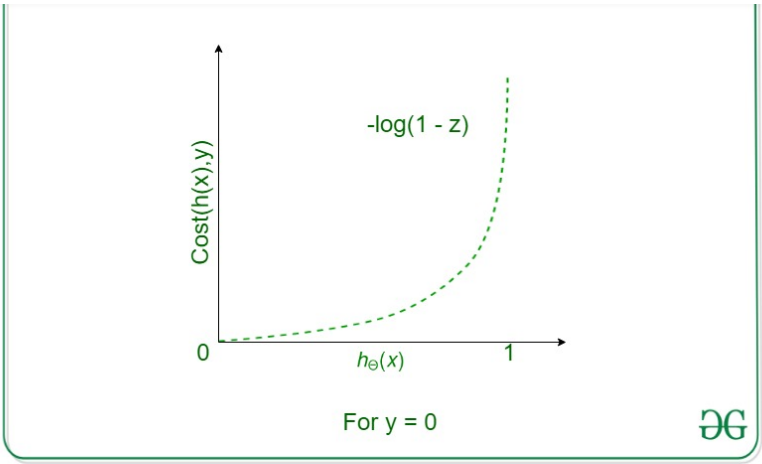
Cost = 0 if y = 1, $h_\theta(x)$ = 1

But as,

$h_\theta(x)$ -> 0

Cost -> Infinity

If y = 0



So,

$$\mathrm{Cost}(\mathrm{h}_\Theta(x), y) = \begin{cases} 0 & if \quad h_\Theta(x) = y \\ \infty & if \qquad y = 0 \qquad and \quad h_\Theta(x) \rightarrow 1 \\ \infty & if \qquad y = 1 \qquad and \quad h_\Theta(x) \rightarrow 0 \end{cases}$$

$$\mathrm{Cost}(\mathrm{h}_\Theta(x), y) = -ylog(h_\Theta(x)) - (1 - y)log(1 - h_\Theta(x))$$

$$\mathrm{J}(\Theta) = \frac{-1}{m} \sum_{i=1}^{m} Cost(h_\Theta(x), y)$$

To fit parameter **θ**, J(θ) has to be minimized and for that Gradient Descent is required.

Gradient Descent – Looks similar to that of Linear Regression but the difference lies in the hypothesis $h_\theta(x)$

$$\Theta_j := \Theta_j - \alpha \sum_{i=1}^{m} (h_\Theta(x^{(i)}) - y^{(i)}) x_j^{(i)}$$