Gradient descent for Logistic Regression

We need $\min_{\theta} J(\theta)$:

Repeat until convergence {

$$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j}(J(\theta)) \quad (\text{for } j = 0, 1, 2 \cdots, n)$$

} (simultaneously update all θ_j)

$$\frac{\partial}{\partial \theta_j} J(\theta) = \frac{\partial}{\partial \theta_j}\left(-\frac{1}{m}\sum_{i=1}^{m} y^{(i)} \log\left(h_\theta(x^{(i)})\right) - (1-y^{(i)}) \log\left(1-h_\theta(x^{(i)})\right)\right)$$

$$\boxed{h_\theta(x) = \sum_{j=0}^{n} \theta_j x_j}$$

$$= -\frac{1}{m}\left(\left(y^{(i)} \frac{1}{h_\theta(x^{(i)})} x_j^{(i)} - (1-y^{(i)}) \frac{1}{1-h_\theta(x^{(i)})} (-x_j^{(i)})\right)\right)$$

$$= \frac{1}{m}\sum_{i=1}^{m}\left(h_\theta(x^{(i)}) - y^{(i)}\right) x_j^{(i)}$$

$\therefore$ Repeat {

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

$$\theta_j := \theta_j - \frac{\alpha}{m}\sum_{i=1}^{m}\left(h_\theta(x^{(i)}) - y^{(i)}\right) x_j^{(i)}$$

$(\text{for } j = 0, 1, 2, \cdots n)$

} (simultaneous update all θ_j)

☆ Looks identical to linear regression even though hypothesis func^n $\Rightarrow$ $\underbrace{h_\theta(x) = \theta^T x}_{\text{Linear Regression}}$ & $\underbrace{h_\theta(x) = \frac{1}{1+e^{-\theta^T x}}}_{\text{logistic Regression}}$

$$\theta = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_n \end{bmatrix}_{n+1} \quad x = \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{bmatrix}_{n+1}$$