Name: Dhanraj Bhosale
ASU ID: 1225506620 (dbhosal1)

Assignment - Linear Regression model parameter Learning - Gradient Derivation

Q.1 Cost function -

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

$$h_\theta(x) = \theta^T x = \theta_0 + \theta_1 x$$ m - training set

Gradient of $J(\theta)$,

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

~~Let~~

let $\theta_0 + \theta_1 x^{(i)} - y^{(i)} = f(\theta_0,\theta_1)^{(i)}$ — (1)

$$= \frac{\partial}{\partial\theta_0}\frac{1}{2m}\sum_{i=1}^{m}\left(f(\theta_0,\theta_1)^{(i)}\right)^2$$

$$= \cancel{2}\times\frac{1}{\cancel{2}m}\sum_{i=1}^{m} f(\theta_0,\theta_1)\cdot\frac{\partial}{\partial\theta_j} f(\theta_0,\theta_1)$$ — chain rule — (2)

$$f(\theta_0,\theta_1) = \theta_0 + \theta_1 x^{(i)} - y^{(i)}$$ — from (1)

$$\therefore \frac{\partial}{\partial\theta_j} f(\theta_0,\theta_1) = \frac{\partial}{\partial\theta_j}\left(\theta_0+\theta_1 x^{(i)} - y^{(i)}\right)$$

for $j=0$,

$$\therefore \frac{\partial}{\partial\theta_0} f(\theta_0,\theta_1) = \frac{\partial}{\partial\theta_0}\left(\theta_0+\theta_1 x^{(i)} - y^{(i)}\right) = 1$$ — (3)

for $j=1$

$$\therefore \frac{\partial}{\partial\theta_1} f(\theta_0,\theta_1) = \frac{\partial}{\partial\theta_1}\left(\theta_0+\theta_1 x^{(i)} - y^{(i)}\right) = x_i$$ — (4)

Now substituting ③ & ④ in ①.

$$\boxed{\begin{aligned} \frac{\partial}{\partial \theta_0} J(\theta_0,\theta_1) &= \frac{1}{m}\sum_{i=1}^{m}\left(h_\theta(x_i)-y_i\right), \\ \frac{\partial}{\partial \theta_1} J(\theta_0,\theta_1) &= \frac{1}{m}\sum_{i=1}^{m}\left(h_\theta(x_i)-y_i\right)x_i. \end{aligned}}$$