Derivation of Gradient Descent

we know that

cost function $\longrightarrow$ mean squared error

$$J(\omega,b) = \frac{1}{2m}\sum_{i=1}^{m}\left(f_{\omega,b}(x^{(i)}) - (y^{(i)})\right)^2$$

where $f_{\omega,b}(x) = \omega x + b$

$m \longrightarrow$ no. of tuples

$$\omega = \omega - \alpha \frac{\partial J}{\partial \omega}(\omega,b)$$

$$b = b - \alpha \frac{\partial J}{\partial b}(\omega,b)$$

This process will be repeated upto the point of convergence

Date :

$$\frac{\partial}{\partial w} J(w,b) = \frac{\partial}{\partial w} \frac{1}{2m} \sum_{i=1}^{m} \left(f_{wb}(x_i^i) - (y_i^i) \right)^2$$

$$= \frac{\partial}{\partial w} \frac{1}{2m} \sum_{1}^{n} (w x_i + b - y_i)^2$$

$$= \frac{1}{\cancel{2}m} \sum_{1}^{m} (w x_i + b - y_i)^{\cancel{2}} \times \cancel{2} x_i$$

$$= \frac{1}{m} \sum_{1}^{m} (w x_i + b - y_i) \cdot x_i$$

From chain rule,

$$\boxed{\frac{\partial}{\partial w} J(w,b) = \frac{1}{m} \sum_{1}^{m} (f(x_i) - y_i) \cdot x_i}$$

$$\frac{\partial}{\partial b} J(w,b) = \frac{\partial}{\partial b} \sum_{1}^{m} (f(x_i) - y_i)^2 \times \frac{1}{2m}$$

$$= \frac{\partial}{\partial b} \sum_{1}^{m} (w x_i + b - y_i)^2 \times \frac{1}{2m}$$

$$= \frac{1}{\cancel{2}m} \sum_{1}^{m} (w x_i + b - y_i) \times \cancel{2}$$

$$\boxed{\frac{\partial}{\partial b} J(w,b) = \frac{1}{m} \sum_{1}^{m} (w x_i + b - y_i)}$$

$$w = w - \alpha \frac{\left(\sum_{1}^{m} (f(x_i) - y_i) x_i \right)}{m}$$

~~b = ew~~

$$b = b - \alpha \frac{\left(\sum_{1}^{n} (f(x_i) - y_i) \right)}{m}$$

→ Gradiant descent for linear Regression