$$= 2 \cdot \frac{1}{2} (h_\theta(x) - y) \cdot \frac{\partial}{\partial \theta_j} (h_\theta(x) - y)$$

$$= (h_\theta(x) - y) \cdot \frac{\partial}{\partial \theta_j} (\theta_0 x_0 + \theta_1 x_1 + \dots + \theta_n x_n - y)$$

for $j = 1 \Rightarrow (h_\theta(x) - y) \cdot \frac{\partial}{\partial \theta_1} (\theta_0 x_0 + \theta_1 x_1 + \dots + \theta_n x_n)$

$$= (h_\theta(x) - y) \cdot x_1$$

$$\therefore = (h_\theta(x) - y) \cdot \frac{\partial}{\partial \theta_j} (\theta_0 x_0 + \theta_1 x_1 + \dots + \theta_j x_j + \dots + \theta_n x_n)$$

$$= (h_\theta(x) - y) \cdot x_j$$

$$\theta_j := \theta_j - \alpha (h_\theta(x) - (y)) \cdot x_j$$

(only one training example)

Repeat until convergence

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

for ~~all~~ n training example

(for $j = 0, 1, \dots n$) a.k.a Batch gradient descent