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

$$= \frac{1}{\cancel{2}m}\cancel{2}\sum_{i=1}^{m}\left(h_\theta(x^{(i)}) - y^{(i)}\right)\frac{\partial}{\partial \theta_j}\sum_{i=1}^{n}\left(h_\theta(x^{(i)}) - \dots y^{(i)}\right)$$

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

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

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

Gradient descent for $n \geq 1$:

Repeat until convergence {

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

} (for every $j = 0, 1, 2, \dots, n$)

[simultaneous update]

Feature Scaling $\Rightarrow$ Idea: Make sure features are on a similar scale.

Eg) x_1 = size (0-2000 feet^2)

x_2 = # bedrooms (1-5)