② Vectorization of gradient descent

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

(for all $j = 0, 1, \dots, n$)

$$\theta_0 := \theta_0 - \alpha \frac{1}{m}\sum_{i=1}^{m}\left(h_\theta(x^{(i)}) - y^{(i)}\right) x_0^{(i)}$$

$$\theta_1 := \theta_1 - \alpha \frac{1}{m}\sum_{i=1}^{m}\left(h_\theta(x^{(i)}) - y^{(i)}\right) x_1^{(i)}$$

$$\theta_2 := \theta_2 - \alpha \frac{1}{m}\sum_{i=1}^{m}\left(h_\theta(x^{(i)}) - y^{(i)}\right) x_2^{(i)}$$

(for $j = 0, 1, 2$)

Vectorized Implementation:

$$\underbrace{\theta}_{\mathbb{R}^{n+1}} := \underbrace{\theta}_{\mathbb{R}^{n+1}} - \overbrace{\alpha}^{\mathbb{R}} \overbrace{\delta}^{\mathbb{R}^{n+1}}$$

where, $\delta = \frac{1}{m}\sum_{i=1}^{m}\left(\overbrace{h_\theta(x^{(i)}) - y^{(i)}}^{\mathbb{R}}\right)\overbrace{x^{(i)}}^{\mathbb{R}^{n+1}}$

$$\delta = \begin{bmatrix} \delta_0 \\ \delta_1 \\ \delta_2 \end{bmatrix}$$

$$\delta_0 = \frac{1}{m}\sum_{i=1}^{m} h_\theta(x^{(i)}) - y^{(i)}) x_0^{(i)}$$

$$(h_\theta(x^{(1)}) - y^{(1)})x^{(1)} + (h_\theta(x^{(2)}) - y^{(2)})x^{(2)} + (h_\theta(x^{(3)}) - y^{(3)})x^{(3)}$$

$$X = \begin{bmatrix} x_0^{(i)} \\ x_1^{(i)} \\ x_2^{(i)} \end{bmatrix}$$