$$\theta^T = [\theta_0 \ \theta_1 \ \cdots \ \theta_n]_{1 \times (n+1)}$$

row matrix vector

$$\theta^T x = [\theta_0 \ \theta_1 \ \cdots \ \theta_n] \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{bmatrix}$$

$$= \theta_0 x_0 + \theta_1 x_1 + \cdots + \theta_n x_n$$

→ a.k.a multivariate linear regression.

$\therefore$, $h_\theta(x) = \sum_{i=0}^{n} \theta_i x_i = \theta^T x$

where, $x_0 = 1$

$$J(\theta_0, \theta_1, \cdots \theta_n) = \underset{[\text{cost func}^n]}{J(\theta)} = \frac{1}{2m} \sum_{i=1}^{n} (h_\theta(x^{(i)}) - y^{(i)})^2$$

$$= \frac{1}{2m} \sum_{i=1}^{n} \left(\sum_{j=0}^{n} \theta_j x_j^{(i)} - y^{(i)} \right)^2$$

where, $x_0 = 1$

$$= \frac{1}{2m} \sum_{i=1}^{n} (\theta^T x^{(i)} - y^{(i)})^2$$

Gradient descent :

Repeat until convergence {

$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta)$ (for every $j = 0, 1, \cdots, n$)

[simuntaneous update]

}