

Linear Regression

- 1) It is a machine learning algorithm under supervised learning and regression tasks.
- 2) The hypothesis function is given by

$$h(x) = \sum_{i=0}^d \theta_i x_i = \theta^T x \quad (1)$$

- 3) Our aim is to minimize the cost function using the gradient descent algorithm, given as

$$J(\theta) = \frac{1}{2} \sum_{i=1}^n (h_{\theta}(x^{(i)}) - y^{(i)})^2 \quad (2)$$

- 4) We regularly update the parameters using the LMS update rule or Widrow-Hoff learning rule, α stands for learning rate.

$$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta) \quad (3)$$

$$\theta_j := \theta_j + \alpha (y^{(i)} - h_{\theta}(x^{(i)})) x_j^{(i)} \quad (4)$$