

Notes on UFLDL Tutorial from Stanford University

June 29, 2015

1 Linear regression

1. Linear function: $h_{\theta}(x) = \sum_j \theta_j x_j$
2. The cost function: $J(\theta) = \frac{1}{2} \sum_i (h_{\theta}(x^{(i)}) - y^{(i)})^2$
3. Gradient descent is used to minimize the cost function
4. The gradient of the cost function $\nabla_{\theta} J(\theta)$ is the differential of function $J(\theta)$ and it is a vector that points in the direction of steepest increase as a function of θ .

$$\frac{\partial J(\theta)}{\partial \theta_j} = \sum_i x_j^{(i)} (h_{\theta}(x^{(i)}) - y^{(i)})$$