Date / /

Page No.

Normal equation derivation:

The best fit line is given by: $\theta_0 x_0 + \theta_1 x_1 + \dots \theta_n x_n$

$\Rightarrow \theta^T x$

The motive in linear regission is to minimize the cost function:

$$J(\theta) = \frac{1}{n} \sum_{i=1}^{n} (\theta^T x_i - y_i)^2$$

Since we cannot ~~sett~~ minimize n, we try to minimize,

$(\theta^T x_i - y_i)^2$

which can be written as

$(X\theta - y)^T (X\theta - y) = J(\theta)$

Since, θ is the independent variable we find minima of $J(\theta)$ wrt θ by equating derivate of cost function equal to 0.

$$\Rightarrow \frac{\delta J_\theta}{\delta \theta} = \frac{\delta}{\delta \theta} \left[(X\theta - y)^T (X\theta - y) \right] = 0$$

$\Rightarrow 2X^T X\theta - 2X^T y = 0$

$\Rightarrow 2X^T X\theta = 2X^T y$

$\Rightarrow (X^T X)^{-1} (X^T X)\theta = (X^T X)^{-1} . (X^T y)$

$\Rightarrow \theta = (X^T X)^{-1} . (X^T y)$