BTVN

HongAnh

October, 2022

1 Prove that:

$$t_n = y(x,w) + \epsilon \Longleftrightarrow W = \left(X^t X\right)^{-1} X^T t$$

$$L = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{Y}_i - Y_i\right)^2$$

$$\hat{Y}_i = W_0 + W_1 X_1$$

$$w = \begin{bmatrix} w_0 \\ w_1 \end{bmatrix},\ x = \begin{bmatrix} 1\ x_1 \\ 1\ x_2 \\ ... \\ 1\ x_n \end{bmatrix},\ y = \begin{bmatrix} y_1 \\ y_2 \\ ... \\ y_n \end{bmatrix}$$

$$\hat{y} = \begin{bmatrix} \hat{y_1} \\ \hat{y_2} \\ ... \\ \hat{y_n} \end{bmatrix} = \begin{bmatrix} w_0 + w_1 x_1 \\ w_0 + w_1 x_2 \\ ... \\ w_0 + w_1 x_n \end{bmatrix} = x_w$$

$$L \ = \parallel \hat{y} - y \parallel_2^2 = \parallel X_w - y \parallel_2^2 = (X_w - y)^T (X_w - y)$$

$$\rightarrow \frac{L'}{w'} = 2X^T(X_w - y) = 0$$

$$\Leftrightarrow X^T X_w = X^T y$$

$$\Leftrightarrow w = (X^T X)^{-1} X^T y$$

2 Prove that $X^t X$ is invertable when X full rank

If X is full rank, X is linear independent.

$$\Rightarrow \vec{v}^T X^T X \vec{v} = \vec{v}^T \overrightarrow{0} = 0 \Rightarrow (X\vec{v})^T X \vec{v} = 0 \Rightarrow (X\vec{v}) \cdot (X\vec{v}) = 0 \Rightarrow X\vec{v} = \overrightarrow{0}$$

So we have: if $\vec{v} \in N\left(X^T X\right) \Rightarrow \vec{v} \in N(X)$

$$\Rightarrow \vec{v} can only be \overrightarrow{0} \Rightarrow N\left(X^T X\right) = N(X) = \{\overrightarrow{0}\}$$

$\Rightarrow$ $X^T X$ is linearly independent; and $X^T X$ is a square matrix $\Rightarrow$ $X^T X$ is invertible