Multiple Linear Regression

In multiple linear Regression there are multiple input columns.

* Derivation

$$\hat{y} = \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \beta_3 x_{13} \dots \beta_m x_{1m}$$

$\vdots$

$$\hat{y}_n = \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \dots \beta_m x_{nm}.$$

where β_0 : is intercept and β are Coefficients

$$\begin{bmatrix} \hat{y}_1 \\ y_2 \\ \vdots \\ \hat{y}_n \end{bmatrix} = \begin{bmatrix} \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \dots + \beta_m x_{1m}. \\ \vdots \\ \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \dots + \beta_m x_{nm} \end{bmatrix}$$

$$\downarrow$$

$$\begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1m} \\ 1 \\ 1 \\ \vdots & x_{n1} & x_{n2} & \dots & x_{nm} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_m \end{bmatrix}$$

$(n \times (m+1))$ $\quad (m+1)(1)$

We can represent it as.

$$\boxed{\hat{Y} = XB} \quad - \quad (1)$$

Error function

$$E = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \vdots \\ Y_n \end{bmatrix} \qquad \hat{Y} = \begin{bmatrix} \hat{Y}_1 \\ \hat{Y}_2 \\ \vdots \\ \hat{Y}_n \end{bmatrix}$$ - matrix form

$$e = Y - \hat{Y} = \begin{bmatrix} Y_1 - \hat{Y}_1 \\ \vdots \\ Y_n - \hat{Y}_n \end{bmatrix}_{n \times 1}$$

$$e^T e = \begin{bmatrix} Y_1 - \hat{Y}_1, & Y_2 - \hat{Y}_2 & \cdots & Y_n - \hat{Y}_n \end{bmatrix}_{1 \times n} \begin{bmatrix} Y_1 - \hat{Y}_1 \\ Y_2 - \hat{Y}_2 \\ \vdots \\ Y_n - \hat{Y}_n \end{bmatrix}_{n \times 1}$$

$$e^T e = (Y_1 - \hat{Y}_1)^2 + (Y_2 - \hat{Y}_2)^2 + \cdots + (Y_n - \hat{Y}_n)^2$$

$$\boxed{e^T e = \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2}$$

$$\boxed{E = e^T e} \quad - \ (2)$$

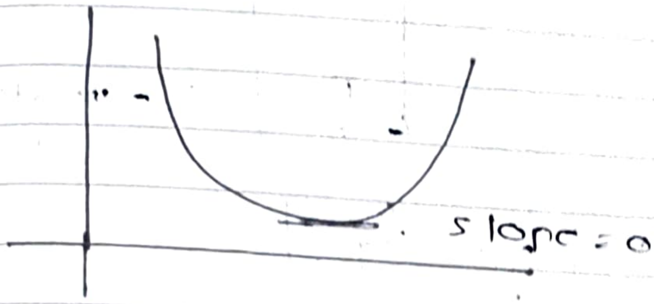
$$E = (Y - \hat{Y})^T (Y - \hat{Y})$$
$$= (Y^T - \hat{Y}^T)(Y - \hat{Y})$$

$$E = Y^T Y - \underbrace{Y^T \hat{Y} - \hat{Y}^T Y}_{} + \hat{Y}^T \hat{Y}$$

$$E = Y^T Y - \underline{2 Y^T \hat{Y}} + \hat{Y}^T \hat{Y}$$ by matrix Transform

$$E = Y^T Y - 2Y^T (XB) + (XB)^T (XB)$$ - from (1)

We have to find such value of β matrix for which E is min



$$\frac{dE}{d\beta} = 0$$

$$\frac{dE}{d\beta} = 0 = -2Y^T X + 2\beta^T X^T X = 0$$

matrix differentiation

$$\cancel{2}\beta^T X^T X = \cancel{2} Y^T X$$

$X^T X = A$ Symmetric

$$\boxed{\beta^T X^T X = Y^T X}$$

Multiply by $(X^T X)^{-1}$ on both side

$$\beta^T X^T X (X^T X)^{-1} = Y^T X (X^T X)^{-1}$$

$$\beta^T I = Y^T X (X^T X)^{-1}$$

$$\beta^T = Y^T X (X^T X)^{-1}$$

$$(\beta^T)^T = [\underbrace{Y^T X}_{A} \underbrace{(X^T X)^{-1}}_{B}]^T$$

$$\beta =$$

$$\beta = [(X^T X)^{-1}]^T (Y^T X)^T$$

$$\beta = [(X^T X)^{-1}]^T X^T Y$$

As $(X^T X^{-1})^T = X^T X^{-1}$ — a

$\therefore X^T X = A$ — consider

$$A A^{-1} = I$$

$$(AA^{-1})^T = I^T$$
$$(A^{-1})^T A^T = I$$
$$(A^{-1})^T A = I$$
$$(A^{-1})^T A A^{-1} = I A^{-1}$$
$$(A^{-1})^T I = A^{-1}$$
$$\boxed{(A^{-1})^T = A^{-1}}$$

$$\therefore \boxed{\beta = (X^T X)^{-1} X^T Y} \leftarrow (3)$$

For multiple linear Regression using OLS (ordinary least Square).