LINEAR REGRESSION, MATRIX FORMULATION

THIS FORMULATION LEADS TO MULTILINEAR REGRESSION

GIVEN: $\{(x_i, y_i) \mid 1 \le i \le n\}$ SET OF DATA POINTS

GOAL: FIT THE DATA-POINTS TO A STRAIGHT LINE

DETAILS:

LET $y_i = \beta_0 + \beta_1 x_i + e_i \quad ; \quad 1 \le i \le n$

e_i = ERROR TERM

β_0, β_1 ARE DETERMINED FROM THE GIVEN ~~THAT~~ DATA,

SO THAT $\sum_{i=1}^{n} e_i^2 \triangleq E$ IS MINIMIZED.

LET $Y^T = [y_1 \; y_2 \; \ldots \; y_n]$ = RESPONSE VECTOR

$$X^T = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_1 & x_2 & x_3 & \cdots & x_n \end{bmatrix} = \text{DESIGN MATRIX}$$

$e^T = [e_1 \; e_2 \; \ldots \; e_n]$ = ERROR VECTOR

$\cancel{\beta^T = X\beta + e}$; $\beta^T = [\beta_0 \quad \beta_1]$ = VECTOR OF PARAMETERS

$Y = X\beta + e$

$$E = \sum_{i=1}^{n} e_i^2 = e^T e = (Y - X\beta)^T (Y - X\beta) = (Y^T - \beta^T X^T)(Y - X\beta)$$

$$= Y^T Y - Y^T X\beta - \beta^T X^T Y + \beta^T X^T X\beta$$

$$= Y^T Y - 2Y^T X\beta + \beta^T X^T X\beta$$

($Y^T X\beta$ IS SCALAR, AND $(Y^T X\beta)^T = \beta^T X^T Y$)

E IS A SCALAR. TAKE FIRST DERIVATIVE OF E WITH RESPECT TO β YIELDS

$$\frac{\partial E}{\partial \beta} = -2X^T y + 2X^T X\beta = 0$$

$$\boxed{\hat{\beta} = (X^T X)^{-1} X^T y} \qquad \text{IF } (X^T X)^{-1} \text{ EXISTS}$$

TO MAKE SURE THAT $\hat{\beta}$ ~~IS~~ INDEED PROVIDES A MINIMUM OF E, AND NOT A MAXIMUM, WE NEED TO TAKE THE SECOND DERIVATIVE, AND MAKE SURE THAT IT IS POSITIVE DEFINITE. THE 2X2 HESSIAN MATRIX OF SECOND DERIVATIVES IS:

$$\frac{\partial^2 E}{\partial \beta\, \partial \beta^T} = \frac{\partial}{\partial \beta^T}\left(-2X^T y + 2X^T X\beta\right) = 2X^T X$$

WHICH IS A POSITIVE DEFINITE MATRIX.

<u>CHECK:</u>

$$X^T X = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} = \begin{bmatrix} n & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \end{bmatrix}$$

DEFINE: $\bar{x} = \frac{1}{n}\sum_{i=1}^n x_i \quad ; \quad \bar{y} = \frac{1}{n}\sum_{i=1}^n y_i$

$$s_x^2 = \frac{1}{n}\sum_{i=1}^n (x_i - \bar{x})^2 \quad ; \quad s_y^2 = \frac{1}{n}\sum_{i=1}^n (y_i - \bar{y})^2$$

$$s_{xy} = \frac{1}{n}\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

$$s_x^2 = \frac{1}{n}\sum_{i=1}^n (x_i^2 - 2x_i\bar{x} + \bar{x}^2) = \frac{1}{n}\sum_{i=1}^n x_i^2 - \bar{x}^2$$

$$\sum_{i=1}^n x_i^2 = n\left(s_x^2 + \bar{x}^2\right)$$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{\Delta}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

$$\Delta = (ad - bc)$$

$\therefore$

$$X^T X = \begin{bmatrix} n & n\bar{x} \\ n\bar{x} & n(s_x^2 + \bar{x}^2) \end{bmatrix}$$

$$(X^T X)^{-1} = \frac{1}{n s_x^2}\begin{bmatrix} s_x^2 + \bar{x}^2 & -\bar{x} \\ -\bar{x} & 1 \end{bmatrix}$$

$$X^T Y = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} n\bar{y} \\ \sum_{i=1}^{n} x_i y_i \end{bmatrix}$$

$$\sum_{i=1}^{n} x_i y_i = n(s_{xy} + \bar{x}\bar{y})$$

$$X^T Y = n\begin{bmatrix} \bar{y} \\ s_{xy} + \bar{x}\bar{y} \end{bmatrix}$$

$$\hat{\beta} = (X^T X)^{-1}(X^T Y) = \frac{1}{s_x^2}\begin{bmatrix} s_x^2 + \bar{x}^2 & -\bar{x} \\ -\bar{x} & 1 \end{bmatrix}\begin{bmatrix} \bar{y} \\ s_{xy} + \bar{x}\bar{y} \end{bmatrix} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix}$$

$$\hat{\beta}_1 = \frac{s_{xy}}{s_x^2} \; ; \quad \hat{\beta}_0 = \frac{1}{s_x^2}\left[\bar{y}(s_x^2 + \bar{x}^2) - \bar{x}(s_{xy} + \bar{x}\bar{y})\right]$$

$$= \frac{1}{s_x^2}\left[\bar{y}\, s_x^2 - \bar{x}\, s_{xy}\right] = \bar{y} - \hat{\beta}_1 \bar{x} \qquad \square$$

$$\boxed{\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \; ; \quad \hat{\beta}_1 = \frac{s_{xy}}{s_x^2}} \; ; \quad \hat{\beta} = \begin{bmatrix} \hat{\beta}_0 & \hat{\beta}_1 \end{bmatrix}^T$$