MULTIPLE LINEAR REGRESSION

SLR: $y = mx + b$ → single input column
2D data

nD:

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta x_2 + \dots \beta_n x_n$$

How would we write this for a whole dataset in a matrix form.
↳ rows → n, cols ⇒ m

$$\hat{Y} = \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_n \end{bmatrix} = \begin{bmatrix} \beta_0 + \beta_1 X_{11} + \beta_2 X_{12} \dots \beta_m X_{1m} \\ \beta_0 + \beta_1 X_{21} + \beta_2 X_{22} \dots \beta_m X_{2m} \\ \vdots \\ \beta_n + \beta_1 X_{n1} + \beta_2 X_{n2} \dots \beta_n X_{nm} \end{bmatrix}$$

⇒ Further Decomposing

$$\underbrace{\begin{bmatrix} 1 & X_{11} & X_{12} & \dots & X_{1m} \\ 1 & X_{21} & X_{22} & \dots & X_{2m} \\ \vdots & \vdots & & & \\ 1 & X_{n1} & X_{n2} & \dots & X_{nm} \end{bmatrix}}_{X} \underbrace{\begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_n \end{bmatrix}}_{\beta}$$

$$\boxed{\hat{Y} = X\beta}$$

Got one main Eq. down

This is from where we separate df into x, y at start.

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \qquad e = y - \hat{y} \Rightarrow \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} - \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_n \end{bmatrix}$$

$$e = \begin{bmatrix} y_1 - \hat{y}_1 \\ y_2 - \hat{y}_2 \\ \vdots \\ y_n - \hat{y}_n \end{bmatrix}$$

In Simple LR.:

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

should be equal

In Multiple Linear R

$$\cancel{E}^T \quad E = e^T e.$$

How ?? :

$$\left[(y_1 - \hat{y}_1) \quad (y_2 - \hat{y}_2) \quad \cdots \quad (y_n - \hat{y}_n) \right] \begin{bmatrix} (y_1 - \hat{y}_1) \\ (y_2 - \hat{y}_2) \\ \vdots \\ (y_n - \hat{y}_n) \end{bmatrix}$$

$$(1 \times n)(n \times 1)$$

(1×1) resultant shape

now?

$$(y_1 - \hat{y}_1)^2 + (y_2 - \hat{y}_2)^2 \cdots + (y_n - \hat{y}_n)^2$$

and it can be re-written as

$$\Rightarrow \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$ and ~~it~~ Now: $E = e^T e = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$

Now that we know that this can be loss function let's furthere dive in

$E = e^T e$

$= \cancel{(y - \hat{y})^T (y - \hat{y})}$

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

$= (\underline{Y^T} - \underline{\hat{Y}^T})(Y - \hat{Y}) \quad \therefore (A - B)^T = A^T - B^T$

$\Rightarrow ((\cancel{Y})^T Y^T - (X\beta)^T)(Y - (X\beta))$

re-writting

$= (Y^T - (X\beta)^T)(Y - (X\beta))$

$= Y^T Y - \underbrace{Y^T X\beta - (X\beta)^T Y}_{2} + X\beta (X\beta^T)$

These two have to be proved

equal (go 2 pages left)

$$\boxed{E = Y^T Y - 2Y^T X\beta + \beta^T X^T X\beta}$$

Loss function

Now to find minimum value we take its differental and equal it ~~to~~ 0.

$$\frac{dE}{d\beta} = \frac{d}{d\beta}\left[\cancel{Y^T Y}^{=0} - 2Y^T X\beta + \beta^T X^T X\beta\right] = 0.$$

$$\Rightarrow \left[0 - 2Y^T X^T + \frac{d}{d\beta}(\beta^T X^T X\beta)\right] = 0$$

$$\Rightarrow -2Y^T X^T + 2X^T X\beta = 0$$

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

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

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

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