07-Simple Linear Regression

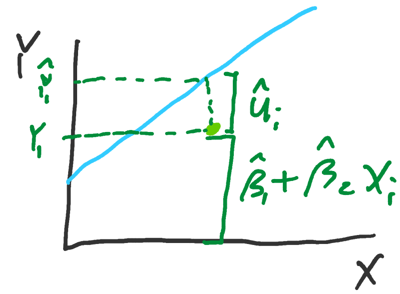
Monday, October 28, 2019 11:58 AM

Preliminaries

$$Y_i = \beta_1 + \beta_2 X_i + u_i$$

Decompose Y_i:

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + \hat{u}_i$$



$$\hat{u}_i = Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i$$

$$RSS = \sum_{i=1}^{n} u_i^2 = \sum (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)^2$$

Objective Function

$$\min_{\hat{\beta}_1, \hat{\beta}_2} \sum (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)^2$$

$$RSS = \sum (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)(Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)$$

$$= \sum \big(Y_i^2 - Y_i \hat{\beta}_1 - Y_i \hat{\beta}_2 X_i - \hat{\beta}_1 Y_i + \hat{\beta}_1^2 + \hat{\beta}_1 \hat{\beta}_2 X_i - \hat{\beta}_2 X_i Y_i + \hat{\beta}_2 X_i \hat{\beta}_1 + \hat{\beta}_2^2 X_i^2\big)$$

$$= \sum \big(Y_i^2 - 2 Y_i \hat{\beta}_1 - 2 Y_i \hat{\beta}_2 X_i + \hat{\beta}_1^2 + 2 \hat{\beta}_1 \hat{\beta}_2 X_i + \hat{\beta}_2^2 X_i^2\big)$$

$$+\hat{\beta}_1^2 + 2\hat{\beta}_1\hat{\beta}_2 X_i + \hat{\beta}_2^2 X_i^2 \Big)$$

FOCs

$$(1) \quad \frac{\partial RSS}{\partial \hat{\beta}_1} = \sum\left(-2Y_i + 2\hat{\beta}_1 + 2\hat{\beta}_2 X_i\right) = 0$$

$$(2) \quad \frac{\partial RSS}{\partial \hat{\beta}_2} = \sum\left(-2Y_iX_i + 2\hat{\beta}_1 X_i + 2\hat{\beta}_2 X_i^2\right) = 0$$

Intercept $(\hat{\beta}_1)$

$$(1) \Rightarrow 2\sum \hat{\beta}_1 = 2\sum Y_i - 2\hat{\beta}_2 \sum X_i$$

$$\sum \hat{\beta}_1 = \sum Y_i - \hat{\beta}_2 \sum X_i$$

Notice that $\sum Y_i = \frac{n}{n}\sum Y_i = n\frac{1}{n}\sum Y_i$
$= n\bar{Y}$

$$n\hat{\beta}_1 = n\bar{Y} - \hat{\beta}_2 n\bar{X}$$

$$\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{X}$$ OLS intercept

Slope $(\hat{\beta}_2)$

$$(2) \Rightarrow \quad -\sum Y_iX_i + \hat{\beta}_1 \sum X_i + \hat{\beta}_2 \sum X_i^2 = 0$$

$(2) \Rightarrow \quad -\sum Y_i X_i + \hat{\beta}_1 \sum X_i + \hat{\beta}_2 \sum X_i^2 = 0$

Plug in $\hat{\beta}_1$:

$$-\sum Y_i X_i + (\bar{Y} - \hat{\beta}_2 \bar{X}) \sum X_i + \hat{\beta}_2 \sum X_i^2 = 0$$

$$-\sum Y_i X_i + \bar{Y} \sum X_i - \hat{\beta}_2 \bar{X} \sum X_i + \hat{\beta}_2 \sum X_i^2 = 0$$

Collect $\hat{\beta}_2$ terms:

$$\hat{\beta}_2 \left(\sum X_i^2 - \bar{X} \sum X_i \right) = \sum Y_i X_i - \bar{Y} \sum X_i$$

$$\hat{\beta}_2 = \frac{\sum (Y_i X_i) - \bar{Y} \sum X_i}{\sum X_i^2 - \bar{X} \sum X_i}$$

Preliminary slope estimator

Simplify $\hat{\beta}_2$

Notice that

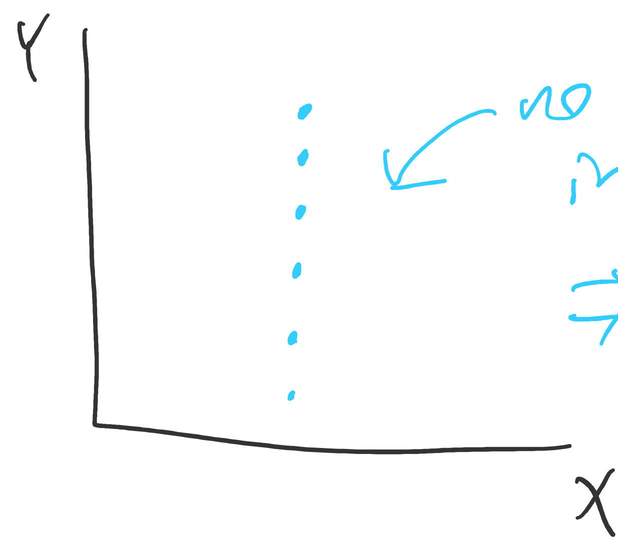
$$\begin{aligned} \sum (X_i - \bar{X}) &= \sum X_i - \sum \bar{X} \\ &= \sum X_i - n\bar{X} \\ &= \sum X_i - n \frac{1}{n} \sum X_i \\ &= \sum X_i - \sum X_i \end{aligned}$$

$$\hat{\beta}_2 = \frac{\sum Y_i X_i - \bar{Y}\sum X_i \overbrace{- \bar{X}\sum (Y_i - \bar{Y})}^{=0}}{\sum X_i^2 - \bar{X}\sum X_i - \bar{X}\sum (X_i - \bar{X})}$$

$$= \frac{\sum Y_i X_i - \sum \bar{Y} X_i - \sum \bar{X} Y_i + \sum \bar{Y}\bar{X}}{\sum X_i^2 - \bar{X}\sum X_i - \sum \bar{X} X_i + \sum \bar{X}^2}$$

$$\hat{\beta}_2 = \frac{\sum (Y_i - \bar{Y})(X_i - \bar{X})}{\sum (X_i - \bar{X})(X_i - \bar{X})}$$

OLS Slope!



no variation in X

$\Rightarrow$ OLS doesn't work

Example

i	X_i	Y_i	$\bar{X}$	$\bar{Y}$	$X_i - X$	$Y_i - \bar{Y}$
1	8	1	5	3	$8-5=3$	$1-3=-2$
2	4	3	5	3	$4-5=-1$	$3-3=0$
3	6	2	5	3	$6-5=1$	$2-3=-1$
4	2	6	5	3	$2-5=-3$	$6-3=3$

Use the OLS formulas