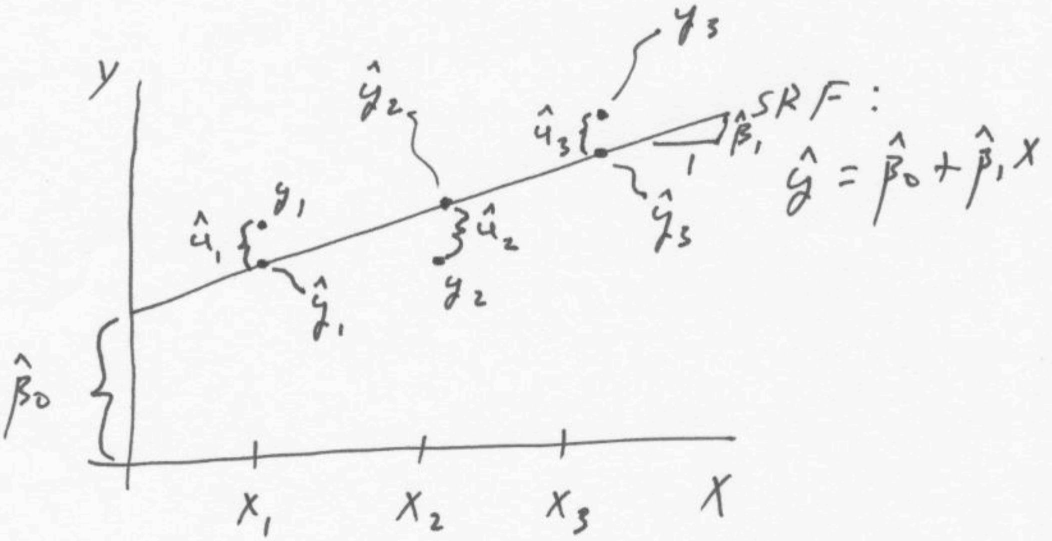
Lecture 5

Derivation of OLS Estimators $\hat{\beta}_0$ and $\hat{\beta}_1$

(Method of Least Squares)

Consider the data scatter



$$\hat{u}_1 = y_1 - \hat{y}_1 = y_1 - (\hat{\beta}_0 + \hat{\beta}_1 X_1)$$

$$\hat{u}_2 = y_2 - \hat{y}_2 = y_2 - (\hat{\beta}_0 + \hat{\beta}_1 X_2)$$

$$\hat{u}_3 = y_3 - \hat{y}_3 = y_3 - (\hat{\beta}_0 + \hat{\beta}_1 X_3)$$

Goal: we want to choose $\hat{\beta}_0$ and $\hat{\beta}_1$ to minimize the sum of squared residuals $\hat{u}_1^2 + \hat{u}_2^2 + \hat{u}_3^2$.

The more general problem is to minimize the sum of squared residuals for N observations

$$\min_{\hat{\beta}_0, \hat{\beta}_1} \sum_1^N \hat{u}_i^2 = \min_{\hat{\beta}_0, \hat{\beta}_1} \sum_1^N (y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)^2$$

$$= \min_{\hat{\beta}_0, \hat{\beta}_1} SSR \qquad (1)$$

where $SSR = \sum_1^N \hat{u}_i^2$ is the sum of squared residuals obtained by fitting a sample Regression Function (SRF) to the data (y_i, X_i), $i = 1, 2, \cdots, N$.

The problem in equation (1) is solved by using the Calculus. We need to solve the two first order conditions below simultaneously for $\hat{\beta}_0$ and $\hat{\beta}_1$.

$$\frac{d}{d\hat{\beta}_0} \sum_1^N (y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)^2 = 0$$

$$\sum_1^N 2(y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)(-1) = 0 \qquad (1)$$

$$\frac{d}{d\hat{\beta}_1} \sum_1^N (y_i - \beta_0 - \hat{\beta}_1 x_i)^2 = 0$$

$$\sum_1^N 2(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)(-x_i) = 0 \qquad (2)$$

From (1) we can get

$$\sum_1^N (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0$$

$$\Rightarrow \sum_1^N y_i - N\hat{\beta}_0 - \hat{\beta}_1 \sum_1^N x_i = 0$$

$$\Rightarrow \sum_1^N y_i - \hat{\beta}_1 \sum_1^N x_i = N\hat{\beta}_0$$

$$\Rightarrow \hat{\beta}_0 = \frac{\sum_1^N y_i}{N} - \hat{\beta}_1 \frac{\sum_1^N x_i}{N}$$

$$\Rightarrow \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \qquad (1')$$

Simplifying (2) we get

$$\sum_1^N (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) x_i = 0$$

$$\Rightarrow \sum_1^N x_i y_i - \hat{\beta}_0 \sum_1^N x_i - \hat{\beta}_1 \sum_1^N x_i^2 = 0 \qquad (2')$$

Now substituting $\hat{\beta}_0$ from (1') into (2') we get

$$\sum_1^N x_i y_i - (\bar{y} - \hat{\beta}_1 \bar{x}) \sum_1^N x_i - \hat{\beta}_1 \sum_1^N x_i^2 = 0$$

$$\Rightarrow \sum_1^N x_i y_i - \bar{y} \sum_1^N x_i + \hat{\beta}_1 \bar{x} \sum_1^N x_i - \hat{\beta}_1 \sum_1^N x_i^2 = 0$$

$$\Rightarrow \hat{\beta}_1 \left(\bar{x} \sum_1^N x_i - \sum_1^N x_i^2 \right) = \bar{y} \sum_1^N x_i - \sum_1^N x_i y_i$$

$$\Rightarrow \hat{\beta}_1 = \frac{\bar{y} \sum_1^N x_i - \sum_1^N x_i y_i}{\bar{x} \sum_1^N x_i - \sum_1^N x_i^2} = \frac{\sum_1^N x_i y_i - \bar{y} \sum_1^N x_i}{\sum_1^N x_i^2 - \bar{x} \sum_1^N x_i}$$

$$= \frac{\sum_1^N x_i y_i - N \bar{x} \bar{y}}{\sum_1^N (x_i - \bar{x})^2} = \frac{\sum_1^N (x_i - \bar{x})(y_i - \bar{y})}{\sum_1^N (x_i - \bar{x})^2}$$

$$= \frac{\sum_1^N (x_i - \bar{x}) y_i}{\sum_1^N (x_i - \bar{x})^2}$$

The last two lines just emphasize that there are several mathematically equivalent forms for the OLS estimator of β_1, the slope of the PRF.