Derivation of the OLS Estimators

- The derivation of the OLS estimators is a standard calculus minimization problem of the sum of squared residuals. Given a set of data all the values like Y_i and X_i are fixed numbers. We have to find the values of $\hat{\beta}_0$ and $\hat{\beta}_1$ that make the SSR as small as possible. In this context $\hat{\beta}_0$ and $\hat{\beta}_1$ are the variables (arguments of the function) while the Ys and the Xs are constants.

- The sum of squared residuals is given by

$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)^2. \tag{1}$$

- To find the values of $\hat{\beta}_0$ and $\hat{\beta}_1$ that minimize this expression, we take partial derivatives of (1) with respect to $\hat{\beta}_0$ and $\hat{\beta}_1$ and set them equal to zero.

$$\frac{\delta}{\delta\hat{\beta}_0}\left[\sum (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)^2\right] = 0. \tag{2}$$

$$\frac{\delta}{\delta\hat{\beta}_1}\left[\sum (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)^2\right] = 0. \tag{3}$$

Evaluating these partial derivatives gives us

$$-2\sum(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i) = 0 \tag{4}$$

$$-2\sum(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)X_i = 0 \tag{5}$$

From both equations the sum of the terms in parenthesis has to be equal to zero. We can then rearrange the equations and write

$$\sum Y_i = n\hat{\beta}_0 + \hat{\beta}_1 \sum X_i \tag{6}$$

$$\sum Y_i X_i = \hat{\beta}_0 \sum X_i + \hat{\beta}_1 \sum X_i^2. \tag{7}$$

This pair of equations is a set of two simultaneous equations in which $\hat{\beta}_0$ and $\hat{\beta}_1$ are unknown and all the X_i and Y_i are known.

- We can solve this system of equations, for example by first solving (6) for

$$\hat{\beta}_0 = \left[\sum Y_i - \hat{\beta}_1 \sum X_i\right]/n, \tag{8}$$

and then substitute this expression in (7). That equation can then be solved to yield

$$\hat{\beta}_1 = \frac{n\sum X_i Y_i - \sum X_I \sum Y_i}{n\sum X_i^2 - (\sum X_i)^2}, \tag{9}$$

expression that can be arranged in many ways in order to make computations easy, for example

$$\hat{\beta}_1 = \frac{\sum(X_i - \overline{X})(Y_i - \overline{Y})}{\sum(X_i - \overline{X})^2}, \tag{10}$$

and also

$$\hat{\beta}_1 = \frac{\sum(X_i - \overline{X})Y_i}{\sum(X_i - \overline{X})^2}, \tag{11}$$

where this last expression is the same one we wrote in class. Now we can write

$$\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1\overline{X}, \tag{12}$$

which again is the expression we wrote in class.

- Notice that from the two equations that equalize the partial derivatives to zero (4) and (5) we conclude that $\sum e_i = 0$, and that e and X are uncorrelated.