Derivation of OLS Estimators in a Simple Regression

1 A Simple Regression Model with Both Intercept and Slope

Consider the model

$$y_t = \beta_1 + \beta_2 x_t + e_t. \tag{1}$$

The sum of errors squared

$$S \equiv \sum_{t=1}^{T} (y_t - \beta_1 - \beta_2 x_t)^2. \tag{2}$$

The first-order derivatives are

$$\begin{aligned}
\frac{\partial S}{\partial \beta_1} &= \sum 2(y_t - \beta_1 - \beta_2 x_t)(-1) = 2T\beta_1 + 2(\sum x_t)\beta_2 - 2\sum y_t, \\
\frac{\partial S}{\partial \beta_2} &= \sum 2(y_t - \beta_1 - \beta_2 x_t)(-x_t) = 2(\sum x_t)\beta_1 + 2(\sum x_t^2)\beta_2 - 2\sum x_t y_t.
\end{aligned}$$

The first-order conditions (FOCs) are

$$2Tb_1 + 2(\sum x_t)b_2 - 2\sum y_t = 0, \tag{3}$$

$$2(\sum x_t)b_1 + 2(\sum x_t^2)b_2 - 2\sum x_t y_t = 0. \tag{4}$$

From equation (3) one can obtain

$$Tb_1 + (\sum x_t)b_2 - \sum y_t = 0, \tag{5}$$

which leads to

$$\Longrightarrow \quad \begin{aligned}
Tb_1 &= \sum y_t - (\sum x_t)b_2 \\
b_1 &= \frac{\sum y_t}{T} - \frac{\sum x_t}{T} b_2 \\
&\text{(Note that } b_1 = \bar{y} - \bar{x} b_2\text{)}
\end{aligned} \tag{6}$$

From equation (4) one can obtain

$$(\sum x_t)b_1 + (\sum x_t^2)b_2 - \sum x_t y_t = 0. \tag{7}$$

Plug (6) into (7):

$$\begin{aligned}
& (\sum x_t)(\frac{\sum y_t}{T} - \frac{\sum x_t}{T} b_2) + (\sum x_t^2)b_2 - \sum x_t y_t = 0 \\
\Longrightarrow \quad & (\sum x_t)(\frac{\sum y_t}{T}) - (\sum x_t)(\frac{\sum x_t}{T})b_2 + (\sum x_t^2)b_2 - \sum x_t y_t = 0 \\
\Longrightarrow \quad & (\sum x_t)(\sum y_t) - (\sum x_t)^2 b_2 + T(\sum x_t^2)b_2 - T\sum x_t y_t = 0 \text{ (Multiplying both sides by } T) \\
\Longrightarrow \quad & \left[T(\sum x_t^2) - (\sum x_t)^2\right] b_2 = T\sum x_t y_t - (\sum x_t)(\sum y_t) \\
\Longrightarrow \quad & b_2 = \frac{T\sum x_t y_t - (\sum x_t)(\sum y_t)}{\left[T(\sum x_t^2) - (\sum x_t)^2\right]}. \qquad (8)
\end{aligned}$$

Combine equations (6) and (8), we have the estimators

$$b_2 = \frac{T\sum x_t y_t - (\sum x_t)(\sum y_t)}{\left[T(\sum x_t^2) - (\sum x_t)^2\right]}, \tag{9}$$

$$b_1 = \bar{y} - \bar{x} b_2. \tag{10}$$

Note that b_2 has several equivalent but different forms, including

$$b_2 = \frac{\sum (x_t - \bar{x}) y_t}{\sum (x_t - \bar{x})^2} \tag{11}$$

and

$$b_2 = \frac{\sum (x_t - \bar{x})(y_t - \bar{y})}{\sum (x_t - \bar{x})^2}. \tag{12}$$

You can prove that equations (8), (11), and (12) are equivalent.

2 A Simple Regression Model with Slope Only

If there is no intercept, i.e., $\beta_1 = 0$, the model becomes

$$y_t = \beta_2 x_t + e_t. \tag{13}$$

The sum of errors squared

$$S \equiv \sum_{t=1}^{T} (y_t - \beta_2 x_t)^2. \tag{14}$$

The first-order derivative is

$$\frac{\partial S}{\partial \beta_2} = \sum 2(y_t - \beta_2 x_t)(-x_t) = 2\sum y_t x_t - 2\beta_2 \sum x_t^2. \tag{15}$$

The FOC is

$$2\sum y_t x_t - 2b_2 \sum x_t^2 = 0, \tag{16}$$

which leads to

$$b_2 = \frac{\sum x_t y_t}{\sum x_t^2}. \tag{17}$$

3 A Simple Regression Model with Intercept Only

If there is no slope (no x variable), i.e., $\beta_2 = 0$, the model becomes

$$y_t = \beta_1 + e_t. \tag{18}$$

The sum of errors squared

$$S \equiv \sum_{t=1}^{T} (y_t - \beta_1)^2. \tag{19}$$

The first-order derivative is

$$\frac{\partial S}{\partial \beta_1} = \sum 2(y_t - \beta_1)(-1) = (-2)\sum (y_t - \beta_1). \tag{20}$$

The FOC is

$$(-2)\sum (y_t - b_1) = 0, \tag{21}$$

which leads to

$$b_1 = \frac{\sum y_t}{T} = \bar{y}. \tag{22}$$