Econ 7010 - Assignment 1

Ryan T. Godwin

1. Given the population model:

$$\boldsymbol{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon},$$

derive the OLS estimator for $\boldsymbol{\beta}$. Which assumptions do you need?

Answer.

The OLS estimator is defined as the vector of estimates $\boldsymbol{b}$ which minimizes the sum of squared residuals $\boldsymbol{e}'\boldsymbol{e}$, where $\boldsymbol{y} = X\boldsymbol{b} + \boldsymbol{e}$.

The optimization problem can be stated as:

$$\min_{\boldsymbol{b}} \boldsymbol{e}'\boldsymbol{e}$$

Substituting $\boldsymbol{y} - X\boldsymbol{b}$ into $\boldsymbol{e}$, we get[1]:

$$\boldsymbol{e}'\boldsymbol{e} = (\boldsymbol{y} - X\boldsymbol{b})'(\boldsymbol{y} - X\boldsymbol{b}) = \boldsymbol{y}'\boldsymbol{y} - 2\left(\boldsymbol{y}'X\right)\boldsymbol{b} + \boldsymbol{b}'\left(X'X\right)\boldsymbol{b}$$

Taking the derivative of $\boldsymbol{e}'\boldsymbol{e}$ with respect to the vector $\boldsymbol{b}$, and setting it equal to zero, we get:

$$\frac{\partial \boldsymbol{e}'\boldsymbol{e}}{\boldsymbol{b}} = \boldsymbol{0} - 2\left(\boldsymbol{y}'X\right)' + 2\left(X'X\right)\boldsymbol{b} = 0$$

or

$$X'X\boldsymbol{b} = X'\boldsymbol{y}.$$

Using assumption A.2 (full rank) so that $(X'X)^{-1}$ exists, we can solve for $\boldsymbol{b}$:

$$\boldsymbol{b} = \left(X'X\right)^{-1} X'\boldsymbol{y}$$

The full rank assumption also ensures that we have minimized (not maximized) $\boldsymbol{e}'\boldsymbol{e}$.

2. Suppose the population model is:

$$y_i = \beta_1 + \beta_2 x_i + \epsilon_i$$

The $\boldsymbol{y}$ and $\boldsymbol{x}$ variables are:

$$\boldsymbol{y} = \begin{bmatrix} -1 \\ 2 \\ 5 \end{bmatrix}; \quad \boldsymbol{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

[1] See pg. 9 of Topic 1

Calculate the OLS estimators for β_1 and β_2.

Answer.

The full X matrix is:

$$X = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}$$

The $(X'X)^{-1}$ matrix is then:

$$(X'X)^{-1} = \left(\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \right)^{-1} = \begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix}^{-1} = \frac{1}{6} \begin{bmatrix} 14 & -6 \\ -6 & 3 \end{bmatrix} = \begin{bmatrix} \frac{7}{3} & -1 \\ -1 & \frac{1}{2} \end{bmatrix}$$

The $(X'\boldsymbol{y})$ matrix is:

$$(X'\boldsymbol{y}) = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 6 \\ 18 \end{bmatrix}$$

Finally, the vector of OLS estimates is:

$$\boldsymbol{b} = \begin{bmatrix} \frac{7}{3} & -1 \\ -1 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 6 \\ 18 \end{bmatrix} = \begin{bmatrix} -4 \\ 3 \end{bmatrix}$$

That is, the estimated intercept is $b_1 = -4$ and the estimated slope is $b_1 = 3$.

3. Suppose again that we have the population model:

$$\boldsymbol{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon},$$

but where the X matrix contains only a column of 1s. In this case, prove that $b = \bar{y}$.

Answer.

Again, the OLS estimator is:

$$\boldsymbol{b} = (X'X)^{-1} X'\boldsymbol{y}$$

If the model contains only an intercept, then the X matrix is only a column of 1s. The OLS estimator then becomes:

$$\boldsymbol{b} = \left(\begin{bmatrix} 1 & 1 & \dots & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & 1 & \dots & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \frac{1}{n} \sum_{i=1}^{n} y_i = \bar{y}$$

4. Prove that the OLS residuals sum to zero.

The OLS residuals only sum to zero (in general) if the model includes an intercept. From the "OLS Normal Equations" (the first order condition used to solve for the OLS estimator) we have:

$$X'\boldsymbol{e} = \boldsymbol{0}$$

With the intercept in the model, we have:

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ x_{11} & x_{21} & \dots & x_{n1} \\ \vdots & \vdots & & \vdots \\ x_{1k} & x_{2k} & \dots & x_{nk} \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n e_i \\ \sum_{i=1}^n x_{i1} e_i \\ \vdots \\ \sum_{i=1}^n x_{ik} e_i \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

From the first element of the above vectors, we see that $\sum_{i=1}^n e_i = 0$.

5. Prove that the fitted regression "line" passes through the sample mean of the data.

   Answer.

   See pg.11 in Topic 1.

6. Prove that the sample mean of the OLS predicted values ($\hat{\boldsymbol{y}}$) equals the sample mean of the actual $\boldsymbol{y}$ values.

   Answer.

   See pg.11 in Topic 1.

7. Prove that $\boldsymbol{y}' M^0 \boldsymbol{y} = \hat{\boldsymbol{y}}' M^0 \hat{\boldsymbol{y}} + \boldsymbol{e}'\boldsymbol{e}$.

   Answer.

   See pg.13 in Topic 1.

8. Suppose that we have our usual linear model:

$$\boldsymbol{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon},$$

   but that we partition the X matrix and $\boldsymbol{\beta}$ vector and write the model as:

$$\boldsymbol{y} = X_1\boldsymbol{\beta}_1 + X_2\boldsymbol{\beta}_2 + \boldsymbol{\epsilon}.$$

   All of the usual assumptions are satisfied, except that $\mathrm{E}[\boldsymbol{\epsilon}] = X_1\boldsymbol{\gamma}$. That is, the mean vector for the disturbances is a linear combination of a subset of the regressors. Let $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$ be the OLS estimators for $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$. Obtain the expressions for $\mathrm{E}[\boldsymbol{\beta}_1]$ and $\mathrm{E}[\boldsymbol{\beta}_1]$, and interpret your results.

   Answer.

   The OLS estimator for $\boldsymbol{\beta}_1$ is:

$$\boldsymbol{b}_1 = \left(X_1' M_2 X\right)^{-1} X_1' M_2 \boldsymbol{y},$$

   where $M_2 = I - X_2\left(X_2'X_2\right)^{-1} X_2'$. Taking the expectation of $\boldsymbol{b}_1$, and substituting in the population model for $\boldsymbol{y}$, we have:

$$\mathrm{E}\left[\boldsymbol{b}_1\right] = \mathrm{E}\left[\left(X_1' M_2 X_1\right)^{-1} X_1' M_2 \left(X_1\boldsymbol{\beta}_1 + X_2\boldsymbol{\beta}_2 + \epsilon\right)\right]$$

   Using A.5, and expanding the expression:

$$\mathrm{E}\left[\boldsymbol{b}_1\right] = \left(X_1' M_2 X_1\right)^{-1} X_1' M_2 X_1 \boldsymbol{\beta}_1 + \left(X_1' M_2 X_1\right)^{-1} X_1' M_2 X_2 \boldsymbol{\beta}_2 + \left(X_1' M_2 X_1\right)^{-1} X_1' M_2 \mathrm{E}\left[\epsilon\right]$$

Using $\mathrm{E}\left[\epsilon\right] = X_1 \gamma$, $M_2 X_2 = 0$, and simplifying, we get:

$$\mathrm{E}\left[\boldsymbol{b}_1\right] = \boldsymbol{\beta}_1 + 0 + \gamma$$

We will get a similar result for $\mathrm{E}\left[\boldsymbol{\beta}_2\right]$, except that the third term will cancel since $M_1 X_1 = 0$:

$$\mathrm{E}\left[\boldsymbol{b}_2\right] = 0 + \boldsymbol{\beta}_2 + \left(X_2' X_2\right)^{-1} X_2' M_1 X_1 \gamma = \boldsymbol{\beta}_2$$

A critical assumption for the unbiasedness of the OLS estimator is that the regressors are unrelated to the error term (assumption A.3). The results above show that if some of the regressors (X_1 for example) are related to the error term, then only those corresponding OLS estimators will be biased ($\boldsymbol{\beta}_1$). $\boldsymbol{\beta}_2$ remains unbiased as long as X_2 is unrelated to ϵ.