Econ 4042 - Assignment 2

Ryan T. Godwin

Due: Oct. 4th. Worth 3% of your assignment grade. Question 8 will not be marked.

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:

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

2. Suppose the population model is:

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

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

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

Calculate the OLS estimators for β_1 and β_2.

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 $b = \bar{y}$.

4. Prove that the OLS residuals sum to zero.

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

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

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

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.