Lecture 5: Gauss-Markov Theorem

Economics 326 — Introduction to Econometrics II

Vadim Marmer, UBC

There are many alternative estimators

- The OLS estimator is not the only estimator we can construct. There are alternative estimators with some desirable properties.
- Example: Using only the first two observations, suppose that $X_2 \neq X_1$.

$$\tilde{\beta} = \frac{Y_2 - Y_1}{X_2 - X_1}.$$

- $\tilde{\beta}$ is **linear**:

$$\tilde{\beta} = c_1 Y_1 + c_2 Y_2,$$

  where

$$c_1 = -\frac{1}{X_2 - X_1} \text{ and } c_2 = \frac{1}{X_2 - X_1}.$$

Unbiasedness of $\tilde{\beta}$

- If $Y_i = \alpha + \beta X_i + U_i$ and $E(U_i | X_1, \dots, X_n) = 0$, then $\tilde{\beta}$ is **unbiased**:

$$\begin{aligned}
\tilde{\beta} &= \frac{Y_2 - Y_1}{X_2 - X_1} \\
&= \frac{(\alpha + \beta X_2 + U_2) - (\alpha + \beta X_1 + U_1)}{X_2 - X_1} \\
&= \frac{\beta(X_2 - X_1)}{X_2 - X_1} + \frac{U_2 - U_1}{X_2 - X_1} \\
&= \beta + \frac{U_2 - U_1}{X_2 - X_1}, \text{ and}
\end{aligned}$$

$$\begin{aligned}
E(\tilde{\beta} | X_1, X_2) &= \beta + E\left(\left. \frac{U_2 - U_1}{X_2 - X_1} \right| X_1, X_2 \right) \\
&= \beta + \frac{E(U_2 | X_1, X_2) - E(U_1 | X_1, X_2)}{X_2 - X_1} \\
&= \beta.
\end{aligned}$$

An optimality criterion

- Among all **linear** and **unbiased** estimators, an estimator with the **smallest variance** is called the **Best Linear Unbiased Estimator** (**BLUE**).
- Note that the statement is **conditional** on X's:
  - The estimators are **unbiased** conditionally on X's.
  - The **variance** is conditional on X's.

Gauss-Markov Theorem

Suppose that

- $Y_i = \alpha + \beta X_i + U_i$.
- $E(U_i | X_1, \dots, X_n) = 0$.
- $E(U_i^2 | X_1, \dots, X_n) = \sigma^2$ for all $i = 1, \dots, n$ (homoskedasticity).
- For all $i \neq j$, $E(U_i U_j | X_1, \dots, X_n) = 0$.

Then, conditionally on X's, the OLS estimators are **BLUE**.

Gauss-Markov Theorem (setup)

- We already know that the OLS estimator $\hat{\beta}$ is linear and unbiased (conditionally on X's).
- Let $\tilde{\beta}$ be any other estimator of β such that
  - $\tilde{\beta}$ is linear:
    $$\tilde{\beta} = \sum_{i=1}^{n} c_i Y_i,$$
    where c's depend only on X's.
  - $\tilde{\beta}$ is unbiased:
    $$E\tilde{\beta} = \beta,$$
    where expectation is conditional on X's.
- We need to show that for **any** such $\tilde{\beta} \neq \hat{\beta}$,
  $$Var(\tilde{\beta}) > Var(\hat{\beta}),$$
  where the variance is conditional on X's.

An outline of the proof

1. First, we are going to show that the c's in $\tilde{\beta} = \sum_{i=1}^{n} c_i Y_i$ satisfy $\sum_{i=1}^{n} c_i = 0$ and $\sum_{i=1}^{n} c_i X_i = 1$.
2. Using the results of Step 1, we will show that conditionally on X's, $Cov(\tilde{\beta}, \hat{\beta}) = Var(\hat{\beta})$.
3. Using the results of Step 2, we will show that conditionally on X's, $Var(\tilde{\beta}) \geq Var(\hat{\beta})$.
4. Lastly, we will show that $Var(\tilde{\beta}) = Var(\hat{\beta})$ if and only if $\tilde{\beta} = \hat{\beta}$.

Proof: Step 1

- Since $\tilde{\beta} = \sum_{i=1}^{n} c_i Y_i$,

$$\begin{aligned}\tilde{\beta} &= \sum_{i=1}^{n} c_i(\alpha + \beta X_i + U_i) \\ &= \alpha \sum_{i=1}^{n} c_i + \beta \sum_{i=1}^{n} c_i X_i + \sum_{i=1}^{n} c_i U_i.\end{aligned}$$

- Conditionally on X's,

$$\begin{aligned} E\tilde{\beta} &= E\left(\alpha \sum_{i=1}^{n} c_i + \beta \sum_{i=1}^{n} c_i X_i + \sum_{i=1}^{n} c_i U_i\right) \\ &= \alpha \sum_{i=1}^{n} c_i + \beta \sum_{i=1}^{n} c_i X_i + \sum_{i=1}^{n} c_i E U_i \\ &= \alpha \sum_{i=1}^{n} c_i + \beta \sum_{i=1}^{n} c_i X_i. \end{aligned}$$

Proof: Step 1 (continued)

- From the **linearity** we have that, conditionally on X's,

$$E\tilde{\beta} = \alpha \sum_{i=1}^{n} c_i + \beta \sum_{i=1}^{n} c_i X_i.$$

- From the **unbiasedness** we have that conditionally on X's,

$$\beta = E\tilde{\beta} = \alpha \sum_{i=1}^{n} c_i + \beta \sum_{i=1}^{n} c_i X_i.$$

- Since this has to be true for **any** α and β, it follows now that

$$\sum_{i=1}^{n} c_i = 0, \quad \sum_{i=1}^{n} c_i X_i = 1.$$

Proof: Step 2

- We have

$$\tilde{\beta} = \beta + \sum_{i=1}^{n} c_i U_i, \text{ with } \sum_{i=1}^{n} c_i = 0, \sum_{i=1}^{n} c_i X_i = 1.$$

$$\hat{\beta} = \beta + \sum_{i=1}^{n} w_i U_i, \text{ with } w_i = \frac{X_i - \bar{X}}{\sum_{j=1}^{n} (X_j - \bar{X})^2}.$$

- Conditionally on X's,

$$\begin{aligned} Cov(\tilde{\beta}, \hat{\beta}) &= E[(\tilde{\beta} - \beta)(\hat{\beta} - \beta)] \\ &= E\left[\left(\sum_{i=1}^{n} c_i U_i\right)\left(\sum_{i=1}^{n} w_i U_i\right)\right] \\ &= \sum_{i=1}^{n} c_i w_i E(U_i^2) + \sum_{i=1}^{n} \sum_{j \neq i} c_i w_j E(U_i U_j). \end{aligned}$$

Proof: Step 2 (continued)

$$Cov(\tilde{\beta}, \hat{\beta}) = \sum_{i=1}^{n} c_i w_i E(U_i^2) + \sum_{i=1}^{n} \sum_{j \neq i} c_i w_j E(U_i U_j).$$

- Since $E(U_i^2) = \sigma^2$ for all i's:

$$\sum_{i=1}^{n} c_i w_i E(U_i^2) = \sigma^2 \sum_{i=1}^{n} c_i w_i.$$

- Since $E(U_iU_j) = 0$ for all $i \neq j$,

$$\sum_{i=1}^{n} \sum_{j \neq i} c_i w_j E(U_i U_j) = 0.$$

- Thus,

$$Cov(\tilde{\beta}, \hat{\beta}) = \sigma^2 \sum_{i=1}^{n} c_i w_i.$$

Proof: Step 2 (continued)

Conditionally on X's:

$$Cov(\tilde{\beta}, \hat{\beta}) = \sigma^2 \sum_{i=1}^{n} c_i w_i \text{ and } w_i = \frac{X_i - \bar{X}}{\sum_{j=1}^{n}(X_j - \bar{X})^2}.$$

$$\begin{aligned}
Cov(\tilde{\beta}, \hat{\beta}) &= \sigma^2 \sum_{i=1}^{n} c_i \frac{X_i - \bar{X}}{\sum_{j=1}^{n}(X_j - \bar{X})^2} \\
&= \frac{\sigma^2}{\sum_{j=1}^{n}(X_j - \bar{X})^2} \sum_{i=1}^{n} c_i (X_i - \bar{X}) \\
&= \frac{\sigma^2}{\sum_{j=1}^{n}(X_j - \bar{X})^2} \left(\sum_{i=1}^{n} c_i X_i - \bar{X} \sum_{i=1}^{n} c_i \right) \\
&= \frac{\sigma^2}{\sum_{j=1}^{n}(X_j - \bar{X})^2} (1 + \bar{X} \cdot 0) \\
&= Var(\hat{\beta}).
\end{aligned}$$

Proof: Step 3

- We know now that for any **linear** and **unbiased** $\tilde{\beta}$,

$$Cov(\tilde{\beta}, \hat{\beta}) = Var(\hat{\beta}).$$

- Let's consider $Var(\tilde{\beta} - \hat{\beta})$:

$$\begin{aligned}
Var(\tilde{\beta} - \hat{\beta}) &= Var(\tilde{\beta}) + Var(\hat{\beta}) - 2Cov(\tilde{\beta}, \hat{\beta}) \\
&= Var(\tilde{\beta}) + Var(\hat{\beta}) - 2Var(\hat{\beta}) \\
&= Var(\tilde{\beta}) - Var(\hat{\beta}).
\end{aligned}$$

- But since $Var(\tilde{\beta} - \hat{\beta}) \geq 0$,

$$Var(\tilde{\beta}) - Var(\hat{\beta}) \geq 0$$

  or

$$Var(\tilde{\beta}) \geq Var(\hat{\beta}).$$

Proof: Step 4 (Uniqueness)

Suppose that $Var(\tilde{\beta}) = Var(\hat{\beta})$.

- Then,

$$Var(\tilde{\beta} - \hat{\beta}) = Var(\tilde{\beta}) - Var(\hat{\beta}) = 0.$$

- Thus, $\tilde{\beta} - \hat{\beta}$ is not random or

$$\tilde{\beta} - \hat{\beta} = \text{constant}.$$

- This constant also has to be zero because

$$
\begin{aligned}
E\tilde{\beta} &= E\hat{\beta} + \text{constant} \\
&= \beta + \text{constant},
\end{aligned}
$$

  and in order for $\tilde{\beta}$ to be **unbiased**

$$
\text{constant} = 0 \text{ or } \tilde{\beta} = \hat{\beta}.
$$