Introductory Econometrics

Lecture 6: Gauss-Markov Theorem

Jun Ma
School of Economics
Renmin University of China

October 1, 2018
There are many alternatives 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

\[
c_1 = -\frac{1}{X_2 - X_1} \quad \text{and} \quad c_2 = \frac{1}{X_2 - X_1}.
\]
Unbiasedness of $\tilde{\beta}$

- If $Y_i = \alpha + \beta X_i + U_i$ and $E[U_i|X_1, \ldots, X_n] = 0$, then $\tilde{\beta}$ is unbiased:

\[
\tilde{\beta} = \frac{Y_2 - Y_1}{X_2 - X_1}
\]

\[
= \frac{(\alpha + \beta X_2 + U_2) - (\alpha + \beta X_1 + U_1)}{X_2 - X_1}
\]

\[
= \frac{\beta (X_2 - X_1)}{X_2 - X_1} + \frac{U_2 - U_1}{X_2 - X_1}
\]

\[
= \beta + \frac{U_2 - U_1}{X_2 - X_1}, \text{ and}
\]

\[
E[\tilde{\beta}|X_1, X_2] = \beta + E\left[\frac{U_2 - U_1}{X_2 - X_1} | X_1, X_2\right]
\]

\[
= \beta + \frac{E[U_2|X_1, X_2] - E[U_1|X_1, X_2]}{X_2 - X_1}
\]

\[
= \beta.
\]
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

1. \( Y_i = \alpha + \beta X_i + U_i. \)
2. \( E[U_i|X_1, \ldots, X_n] = 0. \)
3. \( E[U_i^2|X_1, \ldots, X_n] = \sigma^2 \) for all \( i = 1, \ldots, n \) (homoskedasticity).
4. For all \( i \neq j, \) \( E[U_i U_j|X_1, \ldots, X_n] = 0. \)

Then, conditionally on \( X \)'s, the OLS estimators are BLUE.
Gauss-Markov Theorem

- 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 $\beta$ 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 $\tilde{\beta} \neq \hat{\beta}$,
  $$\text{Var} [\tilde{\beta}] > \text{Var} [\hat{\beta}],$$
  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, $\text{Cov} [\tilde{\beta}, \hat{\beta}] = \text{Var} [\hat{\beta}]$.

3. Using the results of Step 2, we will show that conditionally on $X$’s, $\text{Var} [\tilde{\beta}] \geq \text{Var} [\hat{\beta}]$.

4. Lastly, we will show that $\text{Var} [\tilde{\beta}] = \text{Var} [\hat{\beta}]$ if and only if $\tilde{\beta} = \hat{\beta}$.
Proof: Step 1*

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

\[
\tilde{\beta} = \sum_{i=1}^{n} c_i (\alpha + \beta X_i + U_i)
\]

\[
= \alpha \sum_{i=1}^{n} c_i + \beta \sum_{i=1}^{n} c_i X_i + \sum_{i=1}^{n} c_i U_i.
\]

- Conditionally on $X$'s,

\[
E[\tilde{\beta}] = E\left(\alpha \sum_{i=1}^{n} c_i + \beta \sum_{i=1}^{n} c_i X_i + \sum_{i=1}^{n} c_i U_i\right)
\]

\[
= \alpha \sum_{i=1}^{n} c_i + \beta \sum_{i=1}^{n} c_i X_i + \sum_{i=1}^{n} c_i EU_i
\]

\[
= \alpha \sum_{i=1}^{n} c_i + \beta \sum_{i=1}^{n} c_i X_i.
\]
Proof: Step 1*

- 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 $\alpha$ and $\beta$, it follows now that

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

$$\sum_{i=1}^{n} c_i X_i = 1.$$
Proof: Step 2*

We have

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

Conditionally on \( X \)'s,

\[
\text{Cov} [\tilde{\beta}, \hat{\beta}] = E \left[ (\tilde{\beta} - \beta)(\hat{\beta} - \beta) \right] \\
= E \left[ \left( \sum_{i=1}^{n} c_i U_i \right) \left( \sum_{i=1}^{n} w_i U_i \right) \right] \\
= \sum_{i=1}^{n} c_i w_i E \left[ U_i^2 \right] + \sum_{i=1}^{n} \sum_{j \neq i} c_i w_j E \left[ U_i U_j \right].
\]
Proof: Step 2*

\[
\text{Cov} \left[ \tilde{\beta}, \hat{\beta} \right] = \sum_{i=1}^{n} c_i w_i E \left[ U_i^2 \right] + \sum_{i=1}^{n} \sum_{j \neq i} c_i w_j E \left[ U_i U_j \right].
\]

- Since \( E \left[ U_i^2 \right] = \sigma^2 \) for all \( i \)'s:
  \[
  \sum_{i=1}^{n} c_i w_i E \left[ U_i^2 \right] = \sigma^2 \sum_{i=1}^{n} c_i w_i.
  \]

- Since \( E \left[ U_i U_j \right] = 0 \) for all \( i \neq j \),
  \[
  \sum_{i=1}^{n} \sum_{j \neq i} c_i w_j E \left[ U_i U_j \right] = 0.
  \]

- Thus,
  \[
  \text{Cov} \left[ \tilde{\beta}, \hat{\beta} \right] = \sigma^2 \sum_{i=1}^{n} c_i w_i.
  \]
Proof: Step 2*

Conditionally on $X$’s:

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

\[
\text{Cov} \left[ \tilde{\beta}, \hat{\beta} \right] = \sigma^2 \sum_{i=1}^{n} c_i \frac{X_i - \bar{X}}{\sum_{j=1}^{n} (X_j - \bar{X})^2} \\
= \frac{\sigma^2}{\sum_{j=1}^{n} (X_j - \bar{X})^2} \sum_{i=1}^{n} c_i (X_i - \bar{X}) \\
= \frac{\sigma^2}{\sum_{j=1}^{n} (X_j - \bar{X})^2} \left( \sum_{i=1}^{n} c_i X_i - \bar{X} \sum_{i=1}^{n} c_i \right) \\
= \frac{\sigma^2}{\sum_{j=1}^{n} (X_j - \bar{X})^2} (1 + \bar{X} \cdot 0) \\
= \text{Var} \left[ \hat{\beta} \right].
\]
Proof: Step 3*

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

\[
\text{Cov} [\tilde{\beta}, \hat{\beta}] = \text{Var} [\hat{\beta}].
\]

- Let's consider \( \text{Var} [\tilde{\beta} - \hat{\beta}] \):

\[
\begin{align*}
\text{Var} [\tilde{\beta} - \hat{\beta}] &= \text{Var} [\tilde{\beta}] + \text{Var} [\hat{\beta}] - 2\text{Cov} [\tilde{\beta}, \hat{\beta}] \\
&= \text{Var} [\tilde{\beta}] + \text{Var} [\hat{\beta}] - 2\text{Var} [\hat{\beta}] \\
&= \text{Var} [\tilde{\beta}] - \text{Var} [\hat{\beta}]
\end{align*}
\]

- But since \( \text{Var} [\tilde{\beta} - \hat{\beta}] \geq 0 \),

\[
\text{Var} [\tilde{\beta}] - \text{Var} [\hat{\beta}] \geq 0
\]

or

\[
\text{Var} [\tilde{\beta}] \geq \text{Var} [\hat{\beta}].
\]
Proof: Step 4 (Uniqueness)*

Suppose that \( \text{Var} [\tilde{\beta}] = \text{Var} [\hat{\beta}] \).

- Then, \( \text{Var} [\tilde{\beta} - \hat{\beta}] = \text{Var} [\tilde{\beta}] - \text{Var} [\hat{\beta}] = 0 \).

- Thus, \( \tilde{\beta} - \hat{\beta} \) is not random or \( \tilde{\beta} - \hat{\beta} = \text{constant} \).

- This constant also has to be zero because
  \[
  E [\tilde{\beta}] = E [\hat{\beta}] + \text{constant} \\
  = \beta + \text{constant},
  \]
  and in order for \( \tilde{\beta} \) to be unbiased
  \[\text{constant}=0 \text{ or } \tilde{\beta} = \hat{\beta}.\]