Econometrics I - TA section

Eric Schulman

The University of Texas at Austin

November 10, 2020

Gauss-Markov (Theorem 4.4) "OLS" is "BLUE" i.e. best linear unbiased estimator

Gauss-Markov (Theorem 4.4) "OLS" is "BLUE" i.e. best linear unbiased estimator

- (1) unbiased - expected value is "true"

$$E(\hat{\beta}|X) = \beta$$

- (2) best - minimum variance

$$var(\tilde{\beta}|X) \geq var(\hat{\beta}|X)$$

Chapter 4

- Does it matter if there is 1 or 1000 obs for efficiency and unbiased-ness?

Chapter 4

- Does it matter if there is 1 or 1000 obs for efficiency and unbiased-ness?
  - No
- Trade-off between bias and variance i.e. is a constant effiecient and unbiased

- Does it matter if there is 1 or 1000 obs for efficiency and unbiased-ness?
    - No
- Trade-off between bias and variance i.e. is a constant effiecient and unbiased
    - Variance is 0, but super biased
    - There are estimators that have less variance than OLS, but these are biased – Ridge regression

4.1

For some integer k set $\mu_k = E(y^k)$

- Construct and estimator $\hat{\mu}_k$ for μ_k

4.1

For some integer k set $\mu_k = E(y^k)$

- Construct and estimator $\hat{\mu}_k$ for μ_k
  $\hat{\mu}_k = \frac{1}{n} \sum_{i=1}^{k} y_i^k$

4.1

For some integer k set $\mu_k = E(y^k)$

- Construct and estimator $\hat{\mu}_k$ for μ_k
  $\hat{\mu}_k = \dfrac{1}{n} \sum_{i=1}^{k} y_i^k$
- Show $\hat{\mu}_k$ is unbiased for μ_k

4.1

For some integer k set $\mu_k = E(y^k)$

- Construct and estimator $\hat{\mu}_k$ for μ_k
  $\hat{\mu}_k = \frac{1}{n} \sum_{i=1}^{k} y_i^k$
- Show $\hat{\mu}_k$ is unbiased for μ_k
  $E(\hat{\mu}_k) = E\left(\frac{1}{n} \sum_{i=1}^{n} y_i^k\right) = \frac{1}{n} \sum_{i=1}^{n} E\left(y_i^k\right) = E\left(y_i^k\right)$

4.1

For some integer k set $\mu_k = E(y^k)$

- Construct and estimator $\hat{\mu}_k$ for μ_k
  $\hat{\mu}_k = \frac{1}{n}\sum_{i=1}^{k} y_i^k$
- Show $\hat{\mu}_k$ is unbiased for μ_k
  $E(\hat{\mu}_k) = E\left(\frac{1}{n}\sum_{i=1}^{n} y_i^k\right) = \frac{1}{n}\sum_{i=1}^{n} E\left(y_i^k\right) = E\left(y_i^k\right)$

- Calculate the variance of $\hat{\mu}_k$ say $var(\hat{\mu}_k)$. What assumption is needed for $var(\hat{\mu}_k)$ to be finite?

- Calculate the variance of $\hat{\mu}_k$ say $var(\hat{\mu}_k)$. What assumption is needed for $var(\hat{\mu}_k)$ to be finite?
  $var(\hat{\mu}_k) = var\left(\frac{1}{n}\sum_{i=1}^{n} y_i^k\right) = \frac{1}{n^2}\sum_{i=1}^{n} var(y_i^k) = \frac{1}{n}var(y_i^k)$

- Calculate the variance of $\hat{\mu}_k$ say $var(\hat{\mu}_k)$. What assumption is needed for $var(\hat{\mu}_k)$ to be finite?
  $var(\hat{\mu}_k) = var\left(\frac{1}{n}\sum_{i=1}^{n} y_i^k\right) = \frac{1}{n^2}\sum_{i=1}^{n} var(y_i^k) = \frac{1}{n}var(y_i^k)$
  We need $var(y_i^k)$ to be finite

- Calculate the variance of $\hat{\mu}_k$ say $var(\hat{\mu}_k)$. What assumption is needed for $var(\hat{\mu}_k)$ to be finite?
  $var(\hat{\mu}_k) = var\left(\frac{1}{n}\sum_{i=1}^n y_i^k\right) = \frac{1}{n^2}\sum_{i=1}^n var(y_i^k) = \frac{1}{n}var(y_i^k)$
  We need $var(y_i^k)$ to be finite
- Propose an estimator of $var(\hat{\mu}_k)$

- Calculate the variance of $\hat{\mu}_k$ say $var(\hat{\mu}_k)$. What assumption is needed for $var(\hat{\mu}_k)$ to be finite?
  $var(\hat{\mu}_k) = var\left(\frac{1}{n}\sum_{i=1}^{n} y_i^k\right) = \frac{1}{n^2}\sum_{i=1}^{n} var(y_i^k) = \frac{1}{n}var(y_i^k)$
  We need $var(y_i^k)$ to be finite
- Propose an estimator of $var(\hat{\mu}_k)$
  $\frac{1}{n^2}\left(\sum_{i=1}^{n}\left(y_i^k - \hat{\mu}_k\right)^2\right)$

- Calculate the variance of $\hat{\mu}_k$ say $var(\hat{\mu}_k)$. What assumption is needed for $var(\hat{\mu}_k)$ to be finite?
  $var(\hat{\mu}_k) = var\left(\frac{1}{n}\sum_{i=1}^{n} y_i^k\right) = \frac{1}{n^2}\sum_{i=1}^{n} var(y_i^k) = \frac{1}{n} var(y_i^k)$
  We need $var(y_i^k)$ to be finite
- Propose an estimator of $var(\hat{\mu}_k)$
  $\frac{1}{n^2}\left(\sum_{i=1}^{n}\left(y_i^k - \hat{\mu}_k\right)^2\right)$

4.7

Let $\tilde{\beta}$ be the GLS estimator. Assume that $\Omega = c^2\Sigma$ with Σ known and c^2 unknown. Define the residual vector $\tilde{e} = y - X\tilde{\beta}$ and

$$\tilde{c}^2 = \frac{1}{n-k}\tilde{e}'\Sigma^{-1}\tilde{e}$$

4.7

Let $\tilde{\beta}$ be the GLS estimator. Assume that $\Omega = c^2\Sigma$ with Σ known and c^2 unknown. Define the residual vector $\tilde{e} = y - X\tilde{\beta}$ and

$$\tilde{c}^2 = \frac{1}{n-k}\tilde{e}'\Sigma^{-1}\tilde{e}$$

- Show (4.18) $E(\tilde{\beta}|X) = \beta$

4.7
Let $\tilde{\beta}$ be the GLS estimator. Assume that $\Omega = c^2\Sigma$ with Σ known and c^2 unknown. Define the residual vector $\tilde{e} = y - X\tilde{\beta}$ and

$$\tilde{c}^2 = \frac{1}{n-k}\tilde{e}'\Sigma^{-1}\tilde{e}$$

- Show (4.18) $E(\tilde{\beta}|X) = \beta$

$$\begin{aligned} E(\tilde{\beta}|X) &= E((X'\Sigma^{-1}X)^{-1}(X'\Sigma^{-1}y)|X) \\ &= E((X'\Sigma^{-1}X)^{-1}(X'\Sigma^{-1}(X\beta + e))|X) \\ &= \beta + E((X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}e|X) = \beta \end{aligned}$$

4.7

- Show (4.19) $var(\tilde{\beta}|X) = (X'\Omega^{-1}X)^{-1}$

4.7

- Show (4.19) $var(\tilde{\beta}|X) = (X'\Omega^{-1}X)^{-1}$

$$
\begin{aligned}
var(\tilde{\beta}|X) &= var((X'\Sigma^{-1}X)^{-1}(X'\Sigma^{-1}y)|X) \\
&= var((X'\Sigma^{-1}X)^{-1}(X'\Sigma^{-1}(X\beta + e))|X) \\
&= var((X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}e|X) \\
&= (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}\Omega(X'\Sigma^{-1})'((X'\Sigma^{-1}X)^{-1})' \\
&= (X'\Omega^{-1}X)^{-1}
\end{aligned}
$$

- Note

$$(A^{-1})' = (A')^{-1}$$

$$(AB)' = B'A'$$

4.23 Take the linear regression model with $E(y|X) = X\beta$. Define the ridge regression estimator $\hat{\beta} = (X'X + I_k\lambda)^{-1}X'y$ where λ is a fixed constant. Find $E(\hat{\beta}|X)$. Is $\hat{\beta}$ biased for β?

4.23 Take the linear regression model with $E(y|X) = X\beta$. Define the ridge regression estimator $\hat{\beta} = (X'X + I_k\lambda)^{-1}X'y$ where λ is a fixed constant. Find $E(\hat{\beta}|X)$. Is $\hat{\beta}$ biased for β?

$$E(X'X + I_k\lambda)^{-1}X'y|X) = E((X'X)^{-1}X'y + I_k\lambda^{-1}X'y|X)$$

$$= E((X'X)^{-1}X'X\beta + I_k\lambda^{-1}X'X\beta|X) = \beta + \lambda I_k X'X\beta$$

Clearly it is biased.

Chapter 7

- Asymptotic analysis. Now n comes into play.
- Consistency - kind of like being unbiased.
- Asymptotic normality - Normal “approximates” the true distribution.
- Allows us to make statements about how likely our estimator is.

Consistency/LLN

- Consistency

$$\hat{\theta} \rightarrow \theta$$

- Weak LLN (1) x_i are i.i.d. (2) and $E|x_i| < \infty$. Then as $n \rightarrow \infty$

$$\frac{1}{n}\sum_{i=1}^{n} x_i \xrightarrow{p} \mu$$

  - As the sample gets large, variance gets small...
  - Then apply Markov's inequality

Asymptotic Normality/CLT Lindbergh-Levy CLT: (1) x_i i.i.d. and (2) $E(x_i^2) < \infty$ then

$$\sqrt{n}(\bar{x} - \mu) \xrightarrow{d} \mathcal{N}(0, \sigma^2)$$

Asymptotic Normality/CLT Lindbergh-Levy CLT: (1) x_i i.i.d. and (2) $E(x_i^2) < \infty$ then

$$\sqrt{n}(\bar{x} - \mu) \xrightarrow{d} \mathcal{N}(0, \sigma^2)$$

- Do we know μ? i.e. the population mean?

Asymptotic Normality/CLT Lindbergh-Levy CLT: (1) x_i i.i.d. and (2) $E(x_i^2) < \infty$ then

$$\sqrt{n}(\bar{x} - \mu) \xrightarrow{d} \mathcal{N}(0, \sigma^2)$$

- Do we know μ? i.e. the population mean?
- No! We usually need to choose it

Summarizing - "4 properties" you'll be asked about

- Does not depend on n
  - Unbiased
  - Efficient
- Depends on n
  - Consistent
  - Asymptotically normal

7.9 Take the model

$$y_i = x_i\beta + e_i$$

$$E(e_i|x_i) = 0$$

Consider two estimators

$$\hat{\beta} = \sum_{i=1}^{n} \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}$$

$$\tilde{\beta} = \frac{1}{n}\sum_{i=1}^{n} \frac{y_i}{x_i}$$

7.9

- Under stated assumptions, are both estimators consistent for β?

7.9

- Under stated assumptions, are both estimators consistent for β?
  First compute $E(\hat{\beta}|X) = E\left(\frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2} \mid X\right) =$
  $E\left(\frac{\sum_{i=1}^n x_i(\beta x_i + e_i)}{\sum_{i=1}^n x_i^2} \mid X\right) = E\left(\beta \frac{\sum_{i=1}^n x_i^2}{\sum_{i=1}^n x_i^2} \mid X\right) = \beta$
  This establishes unbiasedness i.e.
  $E(\hat{\beta}) = E(E(\hat{\beta}|X)) = E(\beta) = \beta$. Now can apply the weak LLN.

7.9

- Under stated assumptions, are both estimators consistent for β?

7.9

- Under stated assumptions, are both estimators consistent for β?
  $E(\tilde{\beta}|X) = E\left(\frac{1}{n}\sum_{i=1}^{n}\frac{y_i}{x_i} \mid X\right) =$
  $E\left(\frac{1}{n}\sum_{i=1}^{n}\frac{x_i\beta + e_i}{x_i} \mid X\right) = \beta$
  Similarly, can apply the weak LLN.

7.9

- Under stated assumptions, are both estimators consistent for β?
  $E(\tilde{\beta}|X) = E\left(\frac{1}{n}\sum_{i=1}^{n}\frac{y_i}{x_i} \mid X\right) =$
  $E\left(\frac{1}{n}\sum_{i=1}^{n}\frac{x_i\beta + e_i}{x_i} \mid X\right) = \beta$
  Similarly, can apply the weak LLN.
- Are there conditions under which either estimator is efficient?
  Yes Guass-Markov, (1) is the OLS estimator

7.14
Take the model

$$y_i = x_{1i}\beta_1 + x_{2i}\beta_2 + e_i$$

$$E(x_i e_i) = 0$$

Where β_1 and β_2 are scalars. Define $\theta = \beta_1\beta_2$

7.14

- What is the appropriate estimator $\hat{\theta}$ for θ

7.14

- What is the appropriate estimator $\hat{\theta}$ for θ $\begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = (X'X)^{-1}X'y$

  Then
  $\hat{\beta}_1\hat{\beta}_2 = \hat{\theta}$. Can show consistency using Slutsky's theorem i.e. $\hat{\beta}_1$ and $\hat{\beta}_2$ converge to a constant in probability. Their product also converges.

7.14

- What is the appropriate estimator $\hat{\theta}$ for θ $\begin{bmatrix}\hat{\beta}_1 \\ \hat{\beta}_2\end{bmatrix} = (X'X)^{-1}X'y$

  Then
  $\hat{\beta}_1\hat{\beta}_2 = \hat{\theta}$. Can show consistency using Slutsky's theorem i.e. $\hat{\beta}_1$ and $\hat{\beta}_2$ converge to a constant in probability. Their product also converges.
- Note, convergence in probability is stronger than convergence in distribution. As a result, the conditions for Slutsky's theorem (as written in the textbook) apply.

7.14

- Find the asymptotic distribution of $\hat{\theta}$ under standard regularity conditions.

7.14

- Find the asymptotic distribution of $\hat{\theta}$ under standard regularity conditions.
  $\begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} \to \mathcal{N}(0, \Sigma)$
  Need to compute a change of variable for $\hat{\theta} = \hat{\beta}_1 \hat{\beta}_2$. Similar to 4.8 on previous homework.

7.14

- Find the asymptotic distribution of $\hat{\theta}$ under standard regularity conditions.
  $\begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} \to \mathcal{N}(0, \Sigma)$
  Need to compute a change of variable for $\hat{\theta} = \hat{\beta}_1 \hat{\beta}_2$. Similar to 4.8 on previous homework.
- Show how to calculate an asymptotic 95% confidence interval for θ

7.14

- Find the asymptotic distribution of $\hat{\theta}$ under standard regularity conditions.
  $\begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} \to \mathcal{N}(0, \Sigma)$
  Need to compute a change of variable for $\hat{\theta} = \hat{\beta}_1 \hat{\beta}_2$. Similar to 4.8 on previous homework.
- Show how to calculate an asymptotic 95% confidence interval for θ
  Need to solve for θ where $Pr(|\hat{\beta}_1 \hat{\beta}_2| \leq \theta) = .95$

7.14

- Find the asymptotic distribution of $\hat{\theta}$ under standard regularity conditions.
  $\begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} \rightarrow \mathcal{N}(0, \Sigma)$
  Need to compute a change of variable for $\hat{\theta} = \hat{\beta}_1 \hat{\beta}_2$. Similar to 4.8 on previous homework.
- Show how to calculate an asymptotic 95% confidence interval for θ
  Need to solve for θ where $Pr(|\hat{\beta}_1 \hat{\beta}_2| \leq \theta) = .95$