Toulouse School of Economics

M1 INTERMEDIATE ECONOMETRICS

Asymptotics for OLS

Koen Jochmans   François Poinas

2024 — 2025

Overview

This deck of slides goes through the asymptotic behavior of the ordinary least squares estimator.

The corresponding chapter in Hansen is 7.

Model and assumptions (H Assumption 7.2)

Parameter of interest is β in

$$Y = X'\beta + e, \qquad \mathbb{E}(Xe) = 0,$$

that is,

$$\beta = \mathbb{E}(XX')^{-1}\mathbb{E}(XY).$$

Assumptions:

1. Random sample: (Y_i, X_i) for $i = 1, \ldots, n$ on (Y, X).
2. Finite fourth-order moments: $\mathbb{E}(|Y|^4) < \infty$, and
3. $\mathbb{E}(\|X\|^4) < \infty$.
4. Full rank: $\mathbb{E}(XX')$ is positive definite.

Consistency (H Theorem 7.1)

The OLS estimator is

$$\hat{\beta} = \left(\frac{1}{n} \sum_{i=1}^{n} X_i X_i' \right)^{-1} \left(\frac{1}{n} \sum_{i=1}^{n} X_i Y_i \right) = \hat{\boldsymbol{Q}}_{XX}^{-1} \hat{\boldsymbol{Q}}_{XY}$$

while

$$\beta = \mathbb{E}(XX')^{-1} \mathbb{E}(XY) = \boldsymbol{Q}_{XX}^{-1} \boldsymbol{Q}_{XY}.$$

By the LLN, as $n \to \infty$,

$$\hat{\boldsymbol{Q}}_{XX} \underset{p}{\rightarrow} \boldsymbol{Q}_{XX}, \qquad \hat{\boldsymbol{Q}}_{XY} \underset{p}{\rightarrow} \boldsymbol{Q}_{XY}.$$

Because $\boldsymbol{Q}_{XX}$ is invertible, by the continuous-mapping theorem, as $n \to \infty$,

$$\hat{\boldsymbol{Q}}_{XX}^{-1} \hat{\boldsymbol{Q}}_{XY} \underset{p}{\rightarrow} \boldsymbol{Q}_{XX}^{-1} \boldsymbol{Q}_{XY}$$

that is,

$$\hat{\beta} \underset{p}{\rightarrow} \beta.$$

Limit distribution (H Theorem 7.3)

Note that

$$\begin{aligned}
\hat{\beta} &= \hat{\boldsymbol{Q}}_{XX}^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} X_i Y_i\right) \\
&= \hat{\boldsymbol{Q}}_{XX}^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} X_i [X_i'\beta + e_i]\right) \\
&= \hat{\boldsymbol{Q}}_{XX}^{-1}\left(\hat{\boldsymbol{Q}}_{XX}\beta + \frac{1}{n}\sum_{i=1}^{n} X_i e_i\right) \\
&= \beta + \hat{\boldsymbol{Q}}_{XX}^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} X_i e_i\right)
\end{aligned}$$

That is,

$$\sqrt{n}(\hat{\beta} - \beta) = \hat{\boldsymbol{Q}}_{XX}^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i e_i$$

Now,

$$\mathbb{E}(Xe) = 0, \qquad \mathbb{E}(\|Xe\|^2) \leq \mathbb{E}(\|X\|^4)^{1/2}\, \mathbb{E}(|e|^4)^{1/2} < \infty$$

so that, by the CLT,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i e_i \underset{d}{\rightarrow} N(0, \Omega), \qquad \Omega = \mathbb{E}(XX'e^2).$$

Therefore, by Slutzky's theorem,

$$\sqrt{n}(\hat{\beta} - \beta) \underset{d}{\rightarrow} N(0, \boldsymbol{V}_\beta), \qquad \boldsymbol{V}_\beta = \boldsymbol{Q}_{XX}^{-1} \Omega\, \boldsymbol{Q}_{XX}^{-1}$$

because $\hat{\boldsymbol{Q}}_{XX} \underset{p}{\rightarrow} \boldsymbol{Q}_{XX}$, as $n \rightarrow \infty$.

$\boldsymbol{V}_\beta$ is the asymptotic variance of $\hat{\beta}$.

If

$$\mathbb{E}(e^2|X = x) = \sigma^2$$

then

$$\Omega = \mathbb{E}(XX'e^2) = \mathbb{E}[XX'\mathbb{E}(e^2|X)] = \mathbb{E}(XX')\,\sigma^2 = \boldsymbol{Q}_{XX}\,\sigma^2.$$

Consequently, in this case,

$$\boldsymbol{V}_\beta = \sigma^2\,\boldsymbol{Q}_{XX}^{-1}.$$

Variance estimation (H Theorem 7.6)

An estimator of

$$\boldsymbol{V}_\beta = \boldsymbol{Q}_{XX}^{-1} \Omega \, \boldsymbol{Q}_{XX}^{-1}$$

is

$$\hat{\boldsymbol{V}}_\beta = \hat{\boldsymbol{Q}}_{XX}^{-1} \hat{\Omega} \, \hat{\boldsymbol{Q}}_{XX}^{-1}$$

for

$$\hat{\Omega} = \frac{1}{n} \sum_{i=1}^{n} X_i X_i' \hat{e}_i^2$$

where $\hat{e}_i = Y_i - X_i' \hat{\beta}$ are the least-squares residuals.

This is $\hat{\boldsymbol{V}}_\beta^{\mathrm{HC0}}$ in Hansen.

Consider the case where $k = 1$ for simplicity, then

$$
\begin{aligned}
\hat{\Omega} &= \frac{1}{n}\sum_{i=1}^{n} X_i^2 \hat{e}_i^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 \left(Y_i - X_i\hat{\beta}\right)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n} X_i^2 \left(e_i - X_i(\hat{\beta} - \beta)\right)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n} X_i^2 e_i^2 - \frac{2}{n}\sum_{i=1}^{n} X_i^2 \left(X_i e_i(\hat{\beta} - \beta)\right) + \frac{1}{n}\sum_{i=1}^{n} X_i^2 \left(X_i(\hat{\beta} - \beta)\right)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n} X_i^2 e_i^2 - 2\left(\frac{1}{n}\sum_{i=1}^{n} X_i^3 e_i\right)(\hat{\beta} - \beta) + \left(\frac{1}{n}\sum_{i=1}^{n} X_i^4\right)(\hat{\beta} - \beta)^2 \\
&= (\Omega + o_p(1)) + O_p(1)\, o_p(1) + O_p(1)\, o_p(1) \\
&\underset{p}{\to} \Omega
\end{aligned}
$$

as $n \to \infty$

We have

$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 e_i^2 \underset{p}{\to} \mathbb{E}(X^2 e^2) = \Omega.$$

We also know that

$$\hat{\beta} - \beta = o_p(1).$$

and that

$$\frac{1}{n}\sum_{i=1}^{n} X_i^4 \underset{p}{\to} \mathbb{E}(X^4) < \infty,$$

and

$$\frac{1}{n}\sum_{i=1}^{n} X_i^3 e_i \underset{p}{\to} \mathbb{E}(X^3 e) < \infty$$

because

$$\mathbb{E}(|X^3 e|) \leq \mathbb{E}(|X^2||Xe|) \leq \mathbb{E}(|X^2|^2)^{1/2}\,\mathbb{E}(|Xe|^2)^{1/2} = \mathbb{E}(X^4)^{1/2}\,\Omega^{1/2} < \infty.$$

Suppose we are interested in the vector of linear contrasts

$$\theta = \boldsymbol{R}'\beta$$

for some matrix $k \times q$ (non-random) matrix $\boldsymbol{R}$.

An estimator of θ is

$$\hat{\theta} = \boldsymbol{R}'\hat{\beta}.$$

Consistency is immediate.

By linearity of the transformation,

$$\sqrt{n}(\hat{\theta} - \theta) \underset{d}{\rightarrow} N(0, \boldsymbol{V}_\theta), \qquad \boldsymbol{V}_\theta = \boldsymbol{R}'\boldsymbol{V}_\beta\boldsymbol{R},$$

as $n \to \infty$.

Now consider nonlinear transformations

$$\theta = r(\beta)$$

for $r : \mathbb{R}^k \to \mathbb{R}^q$.

Then

$$\hat{\theta} = r(\hat{\beta}) \underset{p}{\rightarrow} \theta$$

as $n \to \infty$ provided that r is continuous at β.

Further, if r is continuously differentiable with Jacobian $\boldsymbol{R}' = \boldsymbol{R}'(\beta)$, then

$$\sqrt{n}(\hat{\theta} - \theta) \underset{d}{\rightarrow} N(0, \boldsymbol{V}_\theta), \qquad \boldsymbol{V}_\theta = \boldsymbol{R}'\boldsymbol{V}_\beta \boldsymbol{R},$$

as $n \to \infty$.

Standardized statistics: t-statistic and Wald statistic (H7.12 and H.16)

Let $\hat{\boldsymbol{R}} = \boldsymbol{R}(\hat{\beta})$.

Easy to see that

$$\hat{\boldsymbol{V}}_\theta = \hat{\boldsymbol{R}}' \hat{\boldsymbol{V}}_\beta \hat{\boldsymbol{R}} \underset{p}{\rightarrow} \boldsymbol{V}_\theta$$

as $n \to \infty$.

Then, provided that $\hat{\boldsymbol{V}}_\theta$ is invertible (wpa1),

$$\sqrt{n}\, \hat{\boldsymbol{V}}_\theta^{-1/2} (\hat{\theta} - \theta) \underset{d}{\rightarrow} N(0, \boldsymbol{I}_q). \tag{1}$$

The diagonal entries of the matrix

$$\hat{\boldsymbol{V}}_\theta^{1/2} / \sqrt{n}$$

are the standard errors of the entries of $\hat{\theta}$. They are a measure of precision.

We will often work with

$$n\,(\hat{\theta}-\theta)'\hat{\boldsymbol{V}}_{\theta}^{-1}(\hat{\theta}-\theta) \underset{d}{\rightarrow} \chi_q^2,$$

which is the inner product of (1).

This is known as a Wald statistic.

When $r:\mathbb{R}^k \rightarrow \mathbb{R}$ then θ is a scalar and the left-hand side of (1) is known as a t-statistic.