Two Routes to the Asymptotics of OLS
and Their Extension to GMM

1. Setup

Consider the linear model

$$y = X\beta + u,$$

with X $(n \times k)$ full column rank and

$$E(u|X) = 0, \qquad Var(u|X) = \sigma^2 I.$$

Assume

$$\frac{1}{n}X'X \to \Sigma, \quad \Sigma \text{ finite and positive definite.}$$

OLS:

$$\hat{\beta} = (X'X)^{-1}X'y.$$

We derive consistency and asymptotic normality in two ways.

---

2. Preliminaries: Quadratic Mean

Definition

$Z_n \to c$ in quadratic mean (or L^2) if

$$E\|Z_n - c\|^2 \to 0.$$

Lemma ($L^2 \Rightarrow p$)

If $Z_n \to c$ in quadratic mean, then $Z_n \xrightarrow{p} c$.

Proof For $\varepsilon > 0$,

$$\Pr(\|Z_n - c\| > \varepsilon) \le \frac{E\|Z_n - c\|^2}{\varepsilon^2} \to 0.$$

□

Bias–Variance Identity

$$E\|\hat{\theta} - \theta\|^2 = \operatorname{tr}\left(Var(\hat{\theta})\right) + \|\operatorname{Bias}(\hat{\theta})\|^2.$$

Thus bias $\to 0$ and variance $\to 0$ imply quadratic mean convergence.

---

3. Method A: Quadratic Mean Route

Bias

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

Variance

$$Var(\hat{\beta}|X) = \sigma^2(X'X)^{-1} = \frac{\sigma^2}{n}\left(\frac{1}{n}X'X\right)^{-1}.$$

Since $\frac{1}{n}X'X \to \Sigma$,

$$Var(\hat{\beta}) \to \frac{\sigma^2}{n}\Sigma^{-1} \to 0.$$

Consistency

Bias $= 0$ and variance $\to 0$ imply

$$E\|\hat{\beta} - \beta\|^2 \to 0 \quad \Rightarrow \quad \hat{\beta} \xrightarrow{p} \beta.$$

Asymptotic Normality

$$\sqrt{n}(\hat{\beta} - \beta) = \left(\frac{1}{n}X'X\right)^{-1} \frac{1}{\sqrt{n}}X'u.$$

By the multivariate CLT,

$$\frac{1}{\sqrt{n}}X'u \xrightarrow{d} N(0, \sigma^2\Sigma),$$

so

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{d} N(0, \sigma^2\Sigma^{-1}).$$

---

4. Method B: LLN + Continuous Mapping + Slutsky

Rewrite:

$$\hat{\beta} - \beta = \left(\frac{1}{n}X'X\right)^{-1}\left(\frac{1}{n}X'u\right).$$

Consistency

LLN gives:

$$\frac{1}{n}X'X \xrightarrow{p} \Sigma, \qquad \frac{1}{n}X'u \xrightarrow{p} 0.$$

By continuity of inversion and Slutsky:

$$\hat{\beta} \xrightarrow{p} \beta.$$

Asymptotic Normality

Using

$$\sqrt{n}(\hat{\beta} - \beta) = \left(\frac{1}{n}X'X\right)^{-1} \frac{1}{\sqrt{n}}X'u,$$

combine CLT + Slutsky to obtain

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{d} N(0, \sigma^2\Sigma^{-1}).$$

---

5. Conceptual Difference

Method A: Variance collapse ⇒ convergence.

Method B: Sample moments converge ⇒ estimator converges.

Method B generalizes to nonlinear estimators; Method A relies on closed-form variance expressions.

---

6. Plug-In, Method of Moments, and GMM

Plug-In Principle

If $\theta_0 = T(F)$, estimate

$$\hat{\theta} = T(F_n).$$

No optimization is required.

Method of Moments (Exactly Identified)

If

$$E[g(W_i, \theta_0)] = 0, \quad g : \mathbb{R}^k \to \mathbb{R}^k,$$

solve

$$\bar{g}_n(\hat{\theta}) = 0.$$

OLS is exactly identified MoM:

$$\frac{1}{n}X'(y - X\hat{\beta}) = 0.$$

GMM (Overidentified)

If g is $m \times 1$ with $m > k$, solve

$$\hat{\theta} = \arg\min_{\theta} \bar{g}_n(\theta)' W_n \bar{g}_n(\theta).$$

Now the estimator is defined by optimization, not substitution.

---

7. Extension of Method B to GMM

Assume:

- $E[g(W_i, \theta_0)] = 0$ uniquely identifies θ_0.
- Uniform LLN:
$$\sup_\theta \|\bar{g}_n(\theta) - g(\theta)\| \xrightarrow{p} 0.$$
- $W_n \xrightarrow{p} W$ positive definite.

Then by the argmin theorem:
$$\hat{\theta}_n \xrightarrow{p} \theta_0.$$

For asymptotic normality, linearize:

$$\sqrt{n}(\hat{\theta} - \theta_0) = -(G'WG)^{-1}G'W\sqrt{n}\bar{g}_n(\theta_0) + o_p(1),$$

where
$$G = \frac{\partial}{\partial \theta'} E[g(W_i, \theta)]\Big|_{\theta_0}, \quad \Omega = Var(g(W_i, \theta_0)).$$

By CLT,
$$\sqrt{n}\bar{g}_n(\theta_0) \xrightarrow{d} N(0, \Omega),$$

so
$$\sqrt{n}(\hat{\theta} - \theta_0) \xrightarrow{d} N\big(0, (G'WG)^{-1}G'W\Omega WG(G'WG)^{-1}\big).$$

---

8. Takeaway

OLS illustrates two asymptotic philosophies:

Variance collapse vs. *Moment convergence.*

Modern econometrics adopts the second because it extends seamlessly to GMM, M-estimation, and nonlinear models.