# Lesson 3: GMM Estimation

# Sources

- Whited & Taylor (Summer School in Structural Estimation)
- Wooldridge (2001), *Econometric Analysis of Cross Section and Panel Data*
- Cochrane (2006), *Asset Pricing*

# Introduction

- **GMM** (Generalized Method of Moments) generalizes the method of moments.
- Formalized by Hansen (1982), it is widely used in economics and finance.
- It underlies methods like Simulated Method of Moments and Indirect Inference.
- **Key Benefit:** No need to fully specify the data's distribution.

# Method of Moments – Overview

- Estimate parameters by equating sample moments with population moments.
- Suppose we have $k$ unknown parameters $\theta_1, \theta_2, \ldots, \theta_k$.
- Data: A random variable $X$ with density $f_X(x; \theta_1, \ldots, \theta_k)$.

# Method of Moments – Moment Equations

- Assume the first $k$ moments satisfy:

$$\begin{aligned} \mu_1 &= E[X] = g_1(\theta_1, \ldots, \theta_k), \\ \mu_2 &= E[X^2] = g_2(\theta_1, \ldots, \theta_k), \\ &\vdots \\ \mu_k &= E[X^k] = g_k(\theta_1, \ldots, \theta_k). \end{aligned}$$

- Replace with sample moments:
    - $ \hat{\mu}*j = \frac{1}{n} \sum*{i=1}^n x_i^j $ for $j = 1, \ldots, k$.
- Solve:

  $\hat{\mu}_1 = g_1(\hat{\theta}_1, \ldots, \hat{\theta}_k)$, etc.

# Example: Normal Distribution

- For parameters $\mu$ and $\sigma$:
  - **Theoretical moments:** $E[X] = \mu$ and $E[X^2] = \mu^2 + \sigma^2$.
  - **Sample moments:**

$$\hat{\mu}_1 = \frac{1}{n} \sum_{i=1}^{n} x_i,$$

$$\hat{\mu}_2 = \frac{1}{n} \sum_{i=1}^{n} x_i^2.$$

- Then, $\hat{\mu} = \hat{\mu}_1$ and $\hat{\sigma}^2 = \hat{\mu}_2 - \hat{\mu}_1^2$.

# GMM: Overidentification

- When there are more moment conditions than parameters, the model is **overidentified**.
- **Advantages:**
  - Compute standard errors.
  - Test the validity of the moment restrictions.
- GMM optimally combines all available moments.

# GMM: Setup & Notation

- Let $w_i$ be an $(M \times 1)$ data vector for observation $i$.
- Parameter vector: $\theta$ (of size $(P \times 1)$).
- Define moment functions:
  $g(w_i, \theta) : \mathbb{R}^M \times \mathbb{R}^P \to \mathbb{R}^L$, with $L \geq P$.
- The key restriction is:
  $E[g(w_i, \theta_0)] = 0$, where $\theta_0$ is the true parameter.

# Moment Restrictions & Criterion

- In the sample, use:
  - $g_T(\theta) = \frac{1}{T}\sum_{i=1}^{T} g(w_i, \theta)$.
- **Objective:** Find $\hat{\theta}$ so that $g_T(\hat{\theta})$ is "close" to 0.
- Criterion function (when overidentified):

$$Q_T(\theta) = \left[g_T(\theta)\right]' \hat{W} \left[g_T(\theta)\right],$$

where $\hat{W}$ is a positive-definite weighting matrix.

## GMM Estimator & FOC

- **Estimator:**
  $\hat{\theta} = \arg\min_\theta Q_T(\theta).$
- **First-Order Condition (if $g$ is differentiable):**
  Define
  $D_T(\theta) = \frac{\partial g_T(\theta)}{\partial \theta}.$
  Then,
  $2\, D_T(\theta)' \hat{W}\, g_T(\theta) = 0.$

# Asymptotic Normality

Under standard regularity conditions:

$$\sqrt{T}\,(\hat{\theta} - \theta_0) \xrightarrow{d} N\Big(0,\ (D'WD)^{-1}\,D'W\,S\,WD\,(D'WD)^{-1}\Big),$$

where

- $D = E\left[\frac{\partial g(w_i,\theta_0)}{\partial \theta}\right]$,
- $S = \mathrm{Var}(g(w_i, \theta_0))$.

**Efficient GMM** uses $W = S^{-1}$, yielding:

$$\sqrt{T}\,(\hat{\theta}_{\text{eff}} - \theta_0) \sim N\Big(0,\ (D'S^{-1}D)^{-1}\Big).$$

# Standard Errors (Two-Step GMM)

1. **Step 1:**

   Estimate $\hat{\theta}_1$ with a simple weighting (e.g. $W = I$).

2. **Step 2:**

   Estimate $S$ from the residuals and re-estimate with $\hat{W} = \hat{S}^{-1}$:

$$\hat{\theta}_2 = \arg\min_{\theta} \Big[ g_T(\theta) \Big]' \hat{S}^{-1} \Big[ g_T(\theta) \Big].$$

- The asymptotic variance is estimated by:

$$\hat{V}(\hat{\theta}_2) = \frac{1}{T} \Big[ d' \hat{S}^{-1} d \Big]^{-1},$$

with $d = \frac{\partial g_T(\theta)}{\partial \theta} \Big|_{\theta = \hat{\theta}_2}$.

# Goodness of Fit

- The GMM criterion can test if the model is correctly specified.
- **J-Test:** Under the null,
  - $T\,Q_T(\hat{\theta}) \xrightarrow{d} \chi^2_{L-P'}$
    where $L$ is the number of moments and $P$ is the number of parameters.

# Example: OLS via GMM (Step 1)

Consider the linear regression:

$y_i = X_i\beta + \varepsilon_i.$

- **Moment Condition:**
  $E\big[X_i(y_i - X_i\beta)\big] = 0.$
- Form the moment vector:

$$g(w_i, \beta) = \begin{pmatrix} X_i'(y_i - X_i\beta) \\ y_i - X_i\beta \end{pmatrix}.$$

- **First-Step Estimator:**
  Use $W = I$:

$$\hat{\beta}_1 = \arg\min_{\beta} \left\{ \left[ \frac{1}{N} \sum_{i=1}^{N} g(w_i, \beta) \right]' \left[ \frac{1}{N} \sum_{i=1}^{N} g(w_i, \beta) \right] \right\}.$$

# Example: OLS via GMM (Step 2)

1. **Estimate Covariance:**

$$\hat{S} = \frac{1}{N} \sum_{i=1}^{N} g(w_i, \hat{\beta}_1) g(w_i, \hat{\beta}_1)'.$$

2. **Efficient Estimator:**

$$\hat{\beta}_2 = \arg\min_{\beta} \left\{ \left[ \frac{1}{N} \sum_{i=1}^{N} g(w_i, \beta) \right]' \hat{S}^{-1} \left[ \frac{1}{N} \sum_{i=1}^{N} g(w_i, \beta) \right] \right\}.$$

# GMM in Practice

- **Numerical Issues:** The estimated $\hat{S}$ may be nearly singular.
- **Solutions:**
    i. Use a one-step estimator.
    ii. Add a small amount of regularizing noise.
    iii. Use a generalized inverse (e.g. Moore-Penrose).

# Final Remarks

- **Regularity Assumptions:** Identification, smoothness, and finite moments are essential.
- **Choice of Instruments:** Overidentification allows testing instrument validity via the J-test.
- **Robustness:** GMM is robust to error distribution misspecification, but correct moment conditions are vital.