Details of the proof of consistency for the GMM estimator for β^o (Theorem 8.2)

Step 0. Key objects and the tools to apply

- Introduce the key objects.
  - β is a generic label for the unknown parameter. β^o is assumed to be the parameter of interest (see Assumption 8.3). $\widehat{\beta}$ is a GMM estimator. Θ is the parameter space.
  - Define the sample GMM objective function as

$$\widehat{Q}(\beta) = -\widehat{m}(\beta)' \widehat{W}^{-1} \widehat{m}(\beta)$$

    and the population GMM objective function as

$$Q(\beta) = -m(\beta)' W^{-1} m(\beta).$$

  - W as seen in $Q(\beta)$ and Assumption 8.4 has to be positive definite.
- Understand the key tools.
  - (One version of Cauchy-Schwarz) For $a \in \mathbb{R}^p$, $b \in \mathbb{R}^p$, and $p \times p$ matrix Ω, we have $|a'\Omega b| \leq \|a\| \|b\| \|\Omega\|$.
  - (Weierstrass Theorem, statement from Wikipedia) If K is a compact set and $f : K \to \mathbb{R}$ is a continuous function, then f is bounded and there exist $a_1, a_2 \in K$ such that

$$f(a_1) = \sup_{x \in K} f(x), \quad f(a_2) = \inf_{x \in K} f(x).$$

  - (Lemma 8.2) The uniform law of large numbers for ergodic stationary processes
  - (Extremum Estimator Lemma, Lemma 8.3) Suppose that (i) Θ is a compact subset of $\mathbb{R}^K$. (ii) $\widehat{Q}(\beta)$ is continuous in β for any data. (iii) $\widehat{Q}(\beta)$ is a measurable function of the data for all $\beta \in \Theta$. If there is a function $Q(\beta)$ such that (C1) $Q(\beta)$ is uniquely maximized on Θ at $\beta^o \in \Theta$, (C2) $\widehat{Q}(\beta)$ converges uniformly in probability to $Q(\beta)$. Then $\widehat{\beta} \xrightarrow{p} \beta^o$.
- The goal is to apply the Extremum Estimator Lemma.
  - (C2) Show that the largest possible absolute discrepancy (or difference) between the two objective functions disappears in the limit, regardless of the value of β:

$$\sup_{\beta \in \Theta} \left| \widehat{Q}(\beta) - Q(\beta) \right| \xrightarrow{p} 0.$$

    Another way to say this is that you have to show the uniform convergence of $\widehat{Q}(\beta)$ to $Q(\beta)$.
  - (C1) Show that β^o uniquely maximizes $Q(\beta)$.

Step 1. Show conditions (i) to (iii) of the Extremum Estimator Lemma hold for the GMM case.

- Assumption 8.1 satisfies (i).
- Assumption 8.2(a), along with the fact that $\widehat{Q}(\beta)$ is quadratic, ensures that (ii) and (iii) are satisfied.

Step 2. Decompose into components which are easy to analyze. Do some algebra (add and subtract tricks) on $\widehat{Q}(\beta) - Q(\beta)$, apply the triangle inequality repeatedly, and show that

$$\begin{aligned}
\left|\widehat{Q}(\beta) - Q(\beta)\right| \quad \leq \quad & \underbrace{\left|\left[\widehat{m}(\beta) - m(\beta)\right]' \widehat{W}^{-1} \left[\widehat{m}(\beta) - m(\beta)\right]\right|}_{(a)} \\
& +2 \underbrace{\left|m(\beta)' \widehat{W}^{-1} \left[\widehat{m}(\beta) - m(\beta)\right]\right|}_{(b)} \\
& + \underbrace{\left|m(\beta)' \left(\widehat{W}^{-1} - W^{-1}\right) m(\beta)\right|}_{(c)}
\end{aligned}$$

Step 3. Apply a version of Cauchy-Schwarz inequality, the Weierstrass theorem, and the uniform LLN.

- Use a particular version of the Cauchy-Schwarz inequality and the definition of a supremum (informally think of this as a maximum) to show that:

$$\begin{aligned}
(a) \quad & \leq \quad \left(\sup_{\beta\in\Theta} \|\widehat{m}(\beta) - m(\beta)\|\right)^2 \left\|\widehat{W}^{-1}\right\| \\
(b) \quad & \leq \quad \left(\sup_{\beta\in\Theta} \|m(\beta)\|\right) \left(\sup_{\beta\in\Theta} \|\widehat{m}(\beta) - m(\beta)\|\right) \left\|\widehat{W}^{-1}\right\| \\
(c) \quad & \leq \quad \left(\sup_{\beta\in\Theta} \|m(\beta)\|\right)^2 \left\|\widehat{W}^{-1} - W^{-1}\right\|
\end{aligned}$$

- All the suprema in the previous expressions exist (note these are not guaranteed to exist).
  - $\sup_{\beta\in\Theta} \|m(\beta)\|$ exists because of Assumpton 8.1, 8.2(d), and the Weierstrass theorem.
  - $\sup_{\beta\in\Theta} \|\widehat{m}(\beta) - m(\beta)\|$ exists because of Assumption 8.2(c).
- All the terms involving $\widehat{W}$ and W should be finite and nonsingular, which follows from Assumption 8.4 and Definition 8.1.

- In the end, we have

$$\begin{aligned}\sup_{\beta\in\Theta}\left|\widehat{Q}(\beta)-Q(\beta)\right| &\leq \left(\sup_{\beta\in\Theta}\|\widehat{m}(\beta)-m(\beta)\|\right)^2\left\|\widehat{W}^{-1}\right\| \\ &+\left(\sup_{\beta\in\Theta}\|m(\beta)\|\right)\left(\sup_{\beta\in\Theta}\|\widehat{m}(\beta)-m(\beta)\|\right)\left\|\widehat{W}^{-1}\right\| \\ &+\left(\sup_{\beta\in\Theta}\|m(\beta)\|\right)^2\left\|\widehat{W}^{-1}-W^{-1}\right\|.\end{aligned}$$

  All the three terms on the right hand side exist and are bounded. Assumptions 8.2(b), 8.2(c), and 8.4, along with the continuous mapping theorem, ensure that

$$\sup_{\beta\in\Theta}\left|\widehat{Q}(\beta)-Q(\beta)\right| \xrightarrow{p} 0.$$

  As a result, (C2) holds.

Step 4. Show that β^o uniquely maximizes $Q(\beta)$. Let $\widetilde{\beta}$ be another maximizer of $Q(\beta)$, so that $\widetilde{\beta}\neq\beta^o$. Let $\mathbf{P}$ be some nonsingular matrix such that $W=\mathbf{P}\mathbf{P}'$. This $\mathbf{P}$ exists because of Assumption 8.4 where nonsingularity must be strengthened to positive definiteness. Thus,

$$Q\left(\widetilde{\beta}\right)=-m\left(\widetilde{\beta}\right)'W^{-1}m\left(\widetilde{\beta}\right)=-m\left(\widetilde{\beta}\right)'\left(\mathbf{P}^{-1}\right)'\mathbf{P}^{-1}m\left(\widetilde{\beta}\right)=-\left\|\mathbf{P}^{-1}m\left(\widetilde{\beta}\right)\right\|^2$$

Since there exists a unique $\beta^o\in\Theta$ such that $m(\beta^o)=0$ by Assumption 8.3, then $Q\left(\widetilde{\beta}\right)<0=Q(\beta^o)$.

As a result, (C1) holds.

Step 5. Show that $\widehat{\beta}\xrightarrow{p}\beta^o$. Because (i), (ii), (iii) hold by Step 1, (C2) holds by Step 3, and (C1) holds by Step 4, we must have $\widehat{\beta}\xrightarrow{p}\beta^o$ by the Extremum Estimator Lemma.