Spring 1997 | John Rust
Economics 551b | 37 Hillhouse, Rm. 27

PROBLEM SET 4: SOLUTIONS
Classical Methods(II)

QUESTION 1 (Proof) By Chebyshev's inequality, for scalar case

$$P\left(|X_N - X| > \varepsilon\right) \le \frac{E(X_N - X)^2}{\varepsilon^2} \to 0 \qquad \text{as } N \to \infty$$

or for vector case

$$P\left((X_N - X)'(X_N - X) > \varepsilon^2\right) \le \frac{E(X_N - X)'(X_N - X)}{\varepsilon^2} \to 0 \qquad \text{as } N \to \infty$$

QUESTION 2 (Proof)

$$\begin{aligned}
f(\omega) &= \frac{1}{2\pi}\sum_{h=-\infty}^{\infty}\gamma(h)e^{-i\omega h} \\
&= \frac{1}{2\pi}\sum_{h=-\infty}^{\infty}\gamma(h)\left(\cos(\omega h) - i\sin(\omega h)\right) \\
&= \frac{1}{2\pi}\gamma(0)\left(\cos(0) - i\sin(0)\right) \\
&\quad + \frac{1}{2\pi}\sum_{h=1}^{\infty}\gamma(h)\left(\cos(\omega h) + \cos(-\omega h) - i\sin(\omega h) - i\sin(-\omega h)\right) \\
&= \frac{1}{2\pi}\gamma(0)\left(1 - i0\right) \\
&\quad + \frac{1}{2\pi}\sum_{h=1}^{\infty}\gamma(h)\left(\cos(\omega h) + \cos(\omega h) - i\sin(\omega h) + i\sin(\omega h)\right) \\
&= \frac{1}{2\pi}\left[\gamma(0) + 2\sum_{h=1}^{\infty}\gamma(h)\cos(\omega h)\right].
\end{aligned}$$

QUESTION 3 Part 1: For the class of extremum estimator (or M-estimator) $\widehat{\theta}_N = \arg\max_{\theta\in\Theta} Q_N(\theta)$, following basic consistency theorem (see Newey-McFadden(1994) for example) is available.

Theorem. If

(a) (Unique Global Maximum) limit function $Q(\theta)$ achieves a unique global maximum at θ^*,

(b) (Compactness) parameter space Θ is a compact set,

(c) (Continuity) limit function $Q(\theta)$ is continuous,

(d) (Uniform Convergence)

$$\sup_{\theta\in\Theta}|Q_N(\theta) - Q(\theta)| \xrightarrow{p} 0,$$

then $\widehat{\theta}_N \xrightarrow{p} \theta^*$.

Using this theorem, consistency can be proved by showing four conditions above. In our case, the criterion function is

$$
\begin{aligned}
Q_N(\beta) &= -\frac{1}{N}\sum_{i=1}^{N}(y_i - \beta x_i)^4 \\
&= -\frac{1}{N}\sum_{i=1}^{N}\left(\varepsilon_i - (\beta-\beta^*)x_i\right)^4 \\
&= -\frac{1}{N}\sum_{i=1}^{N}(\varepsilon_i^4 - 4(\beta-\beta^*)\varepsilon_i^3 x_i + 6(\beta-\beta^*)^2\varepsilon_i^2 x_i^2 \\
&\quad -4(\beta-\beta^*)^3\varepsilon_i x_i^3 + (\beta-\beta^*)^4 x_i^4)
\end{aligned}
$$

By LLN, the limit function is

$$
\begin{aligned}
Q(\beta) &= -E\varepsilon_i^4 + 4(\beta-\beta^*)E\varepsilon_i^3 x_i - 6(\beta-\beta^*)^2 E\varepsilon_i^2 x_i^2 \\
&\quad +4(\beta-\beta^*)^3 E\varepsilon_i x_i^3 - (\beta-\beta^*)^4 E x_i^4
\end{aligned}
$$

Assuming independence of ε_i and x_i, and symmetric distribution of the error ($E\varepsilon_i^3 = 0$), we have

$$Q(\beta) = -E\varepsilon_i^4 - 6(\beta-\beta^*)^2 E\varepsilon_i^2 E x_i^2 - (\beta-\beta^*)^4 E x_i^4.$$

Since coefficients on the polynomials are all negative, the limit function is uniquely maximized at $\beta = \beta^*$(uniquely global maximum). Also the limit function is continuous (continuity). Convergence of criterion function is uniform over compact sets of β by virtue of the multiplicative way in which β enters (compactness and uniform convergence). Therefore $\widehat{\beta}_N \xrightarrow{p} \beta^*$.

Part 2: First order condition satisfied by the estimator is

$$\nabla_\beta Q_N(\widehat{\beta}_N) = -\frac{1}{N}\sum_{i=1}^{N} 4(y_i - \widehat{\beta}_N x_i)^3 x_i = 0$$

Taylor series expansion around β^* yields

$$0 = -\frac{1}{N}\sum_{i=1}^{N} 4(y_i - \beta^* x_i)^3 x_i - \left[\frac{1}{N}\sum_{i=1}^{N} 12(y_i - \overline{\beta} x_i)^2 x_i^2\right](\widehat{\beta}_N - \beta^*)$$

where $\overline{\beta}$ is between β^* and $\widehat{\beta}_N$. Rearranging yields

$$\sqrt{N}(\widehat{\beta}_N - \beta^*) = \left[\frac{1}{N}\sum_{i=1}^{N} 12(y_i - \overline{\beta} x_i)^2 x_i^2\right]^{-1} \frac{1}{\sqrt{N}}\sum_{i=1}^{N} 4(y_i - \beta^* x_i)^3 x_i.$$

For the first (denominator) term, $\widehat{\beta}_N \xrightarrow{p} \beta^*$ implies $\overline{\beta} \xrightarrow{p} \beta^*$, and by LLN uniform in β converging to β^*, we have

$$\frac{1}{N}\sum_{i=1}^{N} 12(y_i - \overline{\beta} x_i)^2 x_i^2 \xrightarrow{p} 12 E\varepsilon_i^2 E x_i^2.$$

For the second (numerator) term, by CLT,

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}4\varepsilon_i^3 x_i \xrightarrow{d} N\left[0,16E\varepsilon_i^6 Ex_i^2\right].$$

Finally by Slutsky theorem,

$$\sqrt{N}(\widehat{\beta}_N-\beta^*)\xrightarrow{d} N\left[0,\frac{16E\varepsilon_i^6 Ex_i^2}{\left(12E\varepsilon_i^2 Ex_i^2\right)^2}\right]=N\left[0,\frac{1}{9}\frac{E\varepsilon_i^6 Ex_i^2}{\left(E\varepsilon_i^2 Ex_i^2\right)^2}\right].$$

Part 3: From (2) and $x_i=1$ and $E\varepsilon_i^6=15\left(\sigma^2\right)^3$(this can be obtained by taking 6th derivative of normal MGF), asymptotic variance of this M-estimator is

$$V(\widehat{\beta})=\frac{1}{9}\frac{E\varepsilon_i^6 E1^2}{\left(E\varepsilon_i^2 E1^2\right)^2}=\frac{1}{9}\frac{15\left(\sigma^2\right)^3}{\left(\sigma^2\right)^2}=\frac{5}{3}\sigma^2.$$

Asymptotic variance of OLS estimator is

$$V(\widehat{\beta}_{OLS})=\frac{\sigma^2}{Ex_i^2}=\sigma^2<\frac{5}{3}\sigma^2=V(\widehat{\beta}).$$

QUESTION 4 (1) GMM estimator is in the class of M-estimators with criterion function

$$Q_N(\theta)=-\widehat{g}_N(\theta)'W_N\widehat{g}_N(\theta)$$

and limit function

$$Q(\theta)=-g_0(\theta)'Wg_0(\theta)$$

where

$$g_0(\theta)=E[g(x,\theta)].$$

In order to apply the basic consistency theorem for M-estimator in Question 3, we need to check each conditions by providing appropriate assumptions.

(Unique Global Maximum) By assuming $Wg_0(\theta)=0$ only if $\theta=\theta^*$, we have $W^{\frac{1}{2}\prime}g_0(\theta)\neq 0$ only if $\theta\neq\theta^*$. Therefore, $Q(\theta)=-\left[W^{\frac{1}{2}\prime}g_0(\theta)\right]'\left[W^{\frac{1}{2}\prime}g_0(\theta)\right]<Q(\theta^*)=0$ for $\theta\neq\theta^*$. We now utilize following matrix ULLN.

Lemma.(Newey-McFadden p.2129) If
(a)(IID data) $\{x_i\}$ are i.i.d.,
(b)(Compactness) parameter space Θ is a compact set,
(c)(Continuity) matrix function $g(x,\theta)$ is continuous with probability 1,
(d)(Dominating Function) there exists dominating function $d(x)$ such that $\|g(x,\theta)\|\leq d(x)$.

then (i) $E[g(x,\theta)]$ is continuous
and (ii)

$$\sup_{\theta\in\Theta}\left\|\frac{1}{N}\sum_{i=1}^{N}g(x_i,\theta)-E[g(x,\theta)]\right\|\xrightarrow{p}0.$$

Therefore by assuming condition (a)-(d), we have (compactness), (continuity) (from result(i)) and (uniform convergence)(see below) for basic consistency theorem.

(Uniform convergence)

$$\begin{aligned}
& |Q_N(\theta) - Q(\theta)| \\
= & \left|\widehat{g}_N(\theta)' W_N \widehat{g}_N(\theta) - g_0(\theta)' W g_0(\theta)\right| \\
\leq & \left|\left[\widehat{g}_N(\theta) - g_0(\theta)\right]' W_N \left[\widehat{g}_N(\theta) - g_0(\theta)\right]\right| \\
& + \left|2 g_0(\theta)' W_N \left[\widehat{g}_N(\theta) - g_0(\theta)\right]\right| \\
& + \left|g_0(\theta)' \left(W_N - W\right) g_0(\theta)\right| \qquad \text{(by triangle inequality)} \\
\leq & \left\|\widehat{g}_N(\theta) - g_0(\theta)\right\|^2 \left\|W_N\right\| \\
& + 2\left\|g_0(\theta)\right\| \left\|\widehat{g}_N(\theta) - g_0(\theta)\right\| \left\|W_N\right\| \\
& + \left\|g_0(\theta)\right\|^2 \left\|W_N - W\right\| \qquad \text{(by Cauchy-Schwartz inequality)}
\end{aligned}$$

By taking supremum, result (ii) and W $=\text{plim}_{N\to\infty} W_N$ imply the RHS converges to zero. Therefore

$$\sup_{\theta\in\Theta} |Q_N(\theta) - Q(\theta)| \xrightarrow{p} 0.$$

Part 2: First order condition satisfied by the GMM estimator is

$$\nabla_\theta Q_N(\widehat{\theta}_N) = -\left[\nabla_\theta \widehat{g}_N(\widehat{\theta}_N)\right]' W_N \widehat{g}_N(\widehat{\theta}_N) = -\widehat{G}_N(\widehat{\theta}_N)' W_N \widehat{g}_N(\widehat{\theta}_N) = 0$$

Taylor series expansion of $\widehat{g}_N(\widehat{\theta}_N)$ around θ^* yields

$$\widehat{g}_N(\widehat{\theta}_N) = \widehat{g}_N(\theta^*) - \widehat{G}_N(\overline{\theta})(\widehat{\theta}_N - \theta^*)$$

where $\overline{\theta}$ is between θ^* and $\widehat{\theta}_N$.. Substitution to FOC implies

$$0 = -\widehat{G}_N(\widehat{\theta}_N)' W_N \left[\widehat{g}_N(\theta^*) - \widehat{G}_N(\overline{\theta})(\widehat{\theta}_N - \theta^*)\right]$$

Rearranging yields

$$\sqrt{N}(\widehat{\theta}_N - \theta^*) = \left[\widehat{G}_N(\widehat{\theta}_N)' W_N \widehat{G}_N(\overline{\theta})\right]^{-1} \widehat{G}_N(\widehat{\theta}_N)' W_N \sqrt{N} \widehat{g}_N(\theta^*).$$

For the first term, $\widehat{\theta}_N \xrightarrow{p} \theta^*$ implies $\overline{\theta} \xrightarrow{p} \theta^*$, and by LLN and continuous mapping theorem, we have

$$\left[\widehat{G}_N(\widehat{\theta}_N)' W_N \widehat{G}_N(\overline{\theta})\right]^{-1} \widehat{G}_N(\widehat{\theta}_N)' W_N \xrightarrow{p} (G'WG)^{-1} G'W,$$

where $G = E[\nabla_\theta g(x, \theta^*)]$. For the second term, by CLT,

$$\sqrt{N}\widehat{g}_N(\theta^*) \xrightarrow{d} N\left[0, \Omega\right].$$

where $\Omega = E[g(x,\theta^*) g(x,\theta^*)']$. Finally by Slutsky theorem,

$$\sqrt{N}\left(\widehat{\theta}_N - \theta^*\right) \xrightarrow{d} N\left[0, (G'WG)^{-1} G'W\Omega WG(G'WG)^{-1}\right].$$

Part 3: Choose W_N such that $W = \Omega^{-1} = E[g(x,\theta^*)g(x,\theta^*)']^{-1}$. For example, $W_N = [\frac{1}{N}\sum_{i=1}^N g(x_i,\widehat{\theta})g(x_i,\widehat{\theta})']^{-1}$ where $\widehat{\theta}$ is the 1st stage GMM estimator with $W_N = I$.

Asymptotic variance with this weight matrix reduces to $(G'\Omega G)^{-1}$ and this achieves minimum variance bound. This can be show as follows.

$$
\begin{aligned}
& (G'WG)^{-1}G'W\Omega WG(G'WG)^{-1} - (G'\Omega G)^{-1} \\
= \; & (G'WG)^{-1}G'W\left[\Omega - G(G'\Omega G)^{-1}G'\right]WG(G'WG)^{-1} \\
= \; & (G'WG)^{-1}G'W\Omega^{\frac{1}{2}}\left[I - \Omega^{-\frac{1}{2}}G(G'\Omega G)^{-1}G'\Omega^{-\frac{1}{2}}\right]\Omega^{\frac{1}{2}}WG(G'WG)^{-1} \\
\geq \; & 0
\end{aligned}
$$

since $I - \Omega^{-\frac{1}{2}}G(G'\Omega G)^{-1}G'\Omega^{-\frac{1}{2}}$ is an idempotent matrix.

QUESTION 5 First order condition satisfied by the MLE is

$$\nabla_\theta l_N(\widehat{\theta}_N) = \frac{1}{N}\sum_{i=1}^N \nabla_\theta \log f(x_i|\widehat{\theta}_N) = 0$$

Taylor series expansion around $\theta_N = \theta^* + \frac{\delta}{N}$ yields

$$0 = \frac{1}{N}\sum_{i=1}^N \nabla_\theta \log f(x_i|\theta_N) + \frac{1}{N}\sum_{i=1}^N \nabla_{\theta\theta} \log f(x_i|\overline{\theta})(\widehat{\theta}_N - \theta_N)$$

where $\overline{\theta}$ is between θ_N and $\widehat{\theta}_N$. Rearranging yields

$$\sqrt{N}(\widehat{\theta}_N - \theta_N) = \left[\frac{1}{N}\sum_{i=1}^N \nabla_{\theta\theta}\log f(x_i|\overline{\theta})\right]^{-1} \frac{1}{\sqrt{N}}\sum_{i=1}^N \nabla_\theta \log f(x_i|\theta_N).$$

For the first term, $\theta^* + \frac{\delta}{N} \to \theta^*$ implies $\widehat{\theta}_N \xrightarrow{p} \theta^*$ and $\overline{\theta} \xrightarrow{p} \theta^*$, and by LLN and continuous mapping theorem, we have

$$\frac{1}{N}\sum_{i=1}^N \nabla_{\theta\theta}\log f(x_i|\overline{\theta}) \xrightarrow{p} E\left[\nabla_{\theta\theta}\log f(x_i|\theta^*)\right] = \mathcal{I}(\theta^*).$$

For the second term, by CLT,

$$\frac{1}{\sqrt{N}}\sum_{i=1}^N \nabla_\theta \log f(x_i|\theta_N) \xrightarrow{d} N\left[0, \mathcal{I}(\theta^*)\right].$$

Finally by Slutsky theorem,

$$\sqrt{N}\left(\widehat{\theta}_N - \theta_N\right) \xrightarrow{d} N\left[0, \mathcal{I}^{-1}(\theta^*)\right].$$

Therefore MLE is regular.

To establish this result, we need (i) $\theta^* \in$int(Θ) so that $\theta_N \in \Theta$, (ii) $\{x_i\}$ are independent, and (iii) Lindeberg condition

For all $\varepsilon > 0$,

$$\lim_{N \to \infty} S_N^{-2} \sum_{i=1}^{N} \int_{|y_{iN}| > \varepsilon S_N} y_{iN}^2 f(x_i|\theta_N) dx_i = 0$$

where

$$\begin{aligned} y_{iN} &= \nabla_\theta \log f(x_i|\theta_N) \\ S_N^2 &= \sum_{i=1}^{N} Var(y_{iN}) \end{aligned}$$

in order to apply CLT for triangular array $\nabla_\theta \log f(x_i|\theta_N)$.