Problem Set #1

Danny Edgel
Econ 715: Econometric Methods
Fall 2021

October 19, 2021

Question 1

(a) We can begin by simplifying the minimization problem:

$$\begin{aligned}
\hat{c} &= \underset{c\in\mathbb{R}}{\operatorname{argmin}}\ \mathbb{E}\left[\rho_\tau(Y-c)\right] \\
&= \underset{c\in\mathbb{R}}{\operatorname{argmin}}\ \mathbb{E}\left[\left[\tau - \mathbb{1}\left\{Y<c\right\}\right](Y-c)\right] \\
&= \underset{c\in\mathbb{R}}{\operatorname{argmin}}\ \tau\mathbb{E}\left[Y\right] - \mathbb{E}\left[Y\mathbb{1}\left\{Y<c\right\}\right] + c\mathbb{E}\left[\mathbb{1}\left\{Y<c\right\}\right] - \tau c
\end{aligned}$$

Where:

$$\begin{aligned}
\mathbb{E}\left[Y\right] &= \int_0^\infty f_{Y^*}(x)xdx \\
\mathbb{E}\left[Y\mathbb{1}\left\{Y<c\right\}\right] &= \begin{cases} \int_0^c f_{Y^*}(x)xdx, & c>0 \\ 0, & c\le 0\end{cases} \\
\mathbb{E}\left[\mathbb{1}\left\{Y<c\right\}\right] &= \begin{cases} \int_0^c f_{Y^*}(x)dx, & c>0 \\ 0, & c\le 0\end{cases}
\end{aligned}$$

First, consider the case when $Pr(Y\le 0)\le\tau$. Then, the expected value of the τth quantile of Y is 0. $\hat{c}=0$ is the unique solution in these cases, covering the $c\le 0$ case.

We can write the problem for $c\ge 0$ as:

$$\hat{c}_+ = \underset{c\ge 0}{\operatorname{argmin}}\ (\tau-1)\int_0^c f_{Y^*}(x)xdx + \tau\int_c^\infty f_{Y^*}(x)xdx + c\left[F_{Y^*}(c) - F_{Y^*}(0) - \tau\right]$$

Where $F_{Y^*}(\cdot)$ is the CDF of Y^*. Thus, we can solve for $\hat{c}_+$ using first order conditions:

$$\begin{aligned}
(\tau-1)cf_{Y^*}(c) - \tau cf_{Y^*}(c) + F_{Y^*}(c) + cf_{Y^*}(c) &= F_{Y^*}(0)+\tau \\
F_{Y^*}(c) &= F_{Y^*}(0)+\tau \\
\hat{c}_+ &= F_{Y^*}^{-1}\left(F_{Y^*}(0)+\tau\right)
\end{aligned}$$

Where, since f_{Y^*} is positive everywhere, F_{Y^*} is strictly increasing, ensuring a unique solution.

(b) To establish the uniform convergence of $\hat{Q}_n(\beta)$, we must show that the following conditions on $g(X_i, \beta) = \rho_\tau \left(Y_i - X_i'\beta \mathbb{1}\left\{X_i'\beta \geq 0\right\}\right)$ are met:

(i) At any $\beta \in B$, $g(X_i, \beta)$ is continuous in β at $\beta = \beta_0$ with probability one.

At any X_i, $X_i'\beta < 0$ or $X_i\beta \geq 0$. In the former case, g is clearly continuous in β at any β, since g is not a funciton of β. If $X_i\beta \geq 0$, then:

$$g(X_i, \beta) = \left(\tau - \mathbb{1}\left\{Y_i < X_i'\beta\right\}\right)\left(Y_i - X_i'\beta\right)$$

The function $Y_i - X_i'\beta$ is clearly continuous in β, but the existence of the indicator function that depends on β presents a potential discontinuity. However, this function is continuous, since

$$\lim_{Y_i - X_i'\beta \to 0^-} g(X_i, \beta) = \lim_{Y_i - X_i'\beta \to 0^+} g(X_i, \beta) = 0$$

Similarly, as $X_i'\beta \to 0^+$, the $X_i'\beta \geq 0$ case of $g(X_i, \beta)$ approaches the $X_i'\beta < 0$ case. Thus, $g(X_i, \beta)$ is continuous in β at $\beta = \beta_0$ with probability one.

(ii) $|g(X_i, \beta)| \leq G(X_i)$ $\forall \beta \in B$ for some dominating function $G(X_i)$

$$\begin{aligned}
|g(X_i, \beta)| &\leq \left|Y_i - X_i'\beta \mathbb{1}\left\{X_i'\beta \geq 0\right\}\right| \leq |Y_i - X_i'\beta| \\
&= \left|(X_i'\beta_0 + \varepsilon_i)\mathbb{1}\left\{\varepsilon_i \geq -X_i'\beta_0\right\} - X_i'\beta\right| \\
&\leq |X_i'(\beta_0 - \beta) + \varepsilon_i| \leq |X_i'(\beta_0 - \beta)| + |\varepsilon_i| \\
&\leq |\varepsilon_i| + c||X_i|| = G(X_i)
\end{aligned}$$

For some $c \in \mathbb{R}_+$.

(iii) $\mathbb{E}\left[G(X_i)\right] < \infty$—This is true so long as $\mathbb{E}\left[|\varepsilon_i|\right] < \infty$ and $\mathbb{E}\left[||X_i||\right] < \infty$

(iv) B is compact—Since B is a subset of a compact space, B is compact.

The limiting criterion function is simply $Eg(X, \beta)$:

$$Q(\beta) = \mathbb{E}\left[\left(\tau - \mathbb{1}\left\{Y < X'\beta \mathbb{1}\left\{X'\beta \geq 0\right\}\right\}\right)\left(Y - X'\beta \mathbb{1}\left\{X'\beta\right\}\right)\right]$$

(c) The distribution of Y^* is given by $F_{\varepsilon|X}(Y - X'\beta_0)$. Thus, this problem is functionally identical to the problem in (a), where $c = X_i'\beta$ and $X_i'\beta \geq 0$ for $\tau \geq F(0)$ and $c = 0$ (i.e., $\beta = 0$) otherwise. Thus, $\hat{c} = F^{-1}(F(0) + \tau)$ is attained by $\mathbb{E}\left[Y_i - X_i'\beta\right] = 0$, which is true if $\beta = \beta_0$

(d) If $\mathbb{E}\left[XX'\mathbb{1}\left\{X'\beta_0 > 0\right\}\right]$ has full rank, then the solution to $\mathbb{E}\left[Y_i - X_i'\beta\right] = 0$ is unique, as well, ensuring that the criterion-minimizing value of β is unique, at β_0.[1]

[1] I understand that the last couple of answers are not sufficient. It is simply not feasible for me to rigorously complete this problem set under my time and priority constraints.

Question 2

(a) Letting $\widetilde{\theta}$ be some value between θ_0 and $\hat{\theta}$, the mean-value expansion of the first-order condition of the problem, at $\hat{\theta}$, is:

$$\begin{aligned}\frac{\partial \hat{Q}(\hat{\theta}_n)}{\partial \theta} &= \frac{1}{n}\sum_{i=1}^{n}\frac{\partial g(W_i,\hat{\theta}_n,\hat{\gamma}_n)}{\partial \theta} \\ &= \frac{1}{n}\sum_{i=1}^{n}\frac{\partial g(W_i,\theta_0,\hat{\gamma}_n)}{\partial \theta} + \frac{1}{n}\sum_{i=1}^{n}\frac{\partial^2 g(W_i,\widetilde{\theta}_n,\hat{\gamma}_n)}{\partial \theta \partial \theta'}(\hat{\theta}_n-\theta_0)\end{aligned}$$

Note that $\hat{\gamma}_n \to_p \gamma_0$, and since $\hat{\gamma}_n$ was acquired via a sample independent of $\{W_i\}$, $Cov(\hat{\gamma}_n, \hat{\theta}_n) = 0$. Then:

$$\sqrt{n}\frac{\partial \hat{Q}(\theta_0)}{\partial \theta} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\partial g(W_i,\theta_0,\hat{\gamma}_n)}{\partial \theta} \to_d \mathcal{N}(0,\Omega_0)$$

$$\text{Where } \Omega_0 = \mathbb{E}\left[\frac{\partial g(W_i,\theta_0,\gamma_0)}{\partial \theta}\frac{\partial g(W_i,\theta_0,\gamma_0)}{\partial \theta'}\right]$$

Denote $B_n = \frac{1}{n}\sum_{i=1}^{n}\frac{\partial^2 g(W_i,\widetilde{\theta}_n,\hat{\gamma}_n)}{\partial \theta \partial \theta'}$, where, since the conditions for ULLN are satisfied:

$$B_n \to_p B_0 = \frac{\partial^2 g(W_i,\theta_0,\gamma_0)}{\partial \theta \partial \theta'}$$

Thus,

$$\sqrt{n}\frac{\partial \hat{Q}(\hat{\theta}_n)}{\partial \theta} = \sqrt{n}\frac{\partial \hat{Q}(\theta_0)}{\partial \theta} + \frac{1}{n}\sum_{i=1}^{n}\frac{\partial^2 g(W_i,\widetilde{\theta}_n,\hat{\gamma}_n)}{\partial \theta \partial \theta'}\sqrt{n}(\hat{\theta}_n-\theta_0) = 0$$

$$\sqrt{n}(\hat{\theta}_n-\theta_0) = -\hat{B}_n^{-1}\sqrt{n}\frac{\partial \hat{Q}(\theta_0)}{\partial \theta} \to_d \mathcal{N}\left(0, B_0^{-1}\Omega_0 B_0^{-1}\right)$$

Where B_0 and Ω_0 are known and given above.

(b) First, the additional conditions necessary to derive the asymptotic distribution of $\sqrt{n}(\hat{\theta}_n-\theta_0)$ are the necessary assumptions for ULLN, which are the assumptions given in (f) and (g) for g, but instead for m.

Since $\hat{\gamma}_n$ and $\hat{\theta}_n$ were retrieved from the same sample, we can no longer assume that their asymptotic covariance is zero and therefore must account for the asymptotic variance of $\hat{\theta}_n$ in the asumptotic variance of $\hat{\theta}_n$. Since m does not depend on θ, we can rewrite Σ_γ as $A_0^{-1}\Omega_0^\gamma A_0^{-1}$, where:

$$A_0 = \mathbb{E}\left[\frac{\partial^2 m(W_i,\gamma_0)}{\partial \gamma \partial \gamma'}\right], \qquad \Omega_0^\gamma = \mathbb{E}\left[\frac{\partial m(W_i,\gamma_0)}{\partial \gamma}\frac{\partial m(W_i,\gamma_0)}{\partial \gamma'}\right]$$

Now, the Taylor expansion from (a) becomes:

$$\frac{\partial \hat{Q}(\hat{\theta}_n)}{\partial \theta} = \frac{1}{n}\sum_{i=1}^{n}\frac{\partial g(W_i,\theta_0,\gamma_0)}{\partial \theta} + \frac{1}{n}\sum_{i=1}^{n}\begin{pmatrix}\frac{\partial^2 g(W_i,\widetilde{\theta}_n,\gamma_0)}{\partial \theta \partial \theta'} \\ \frac{\partial^2 g(W_i,\theta_0,\widetilde{\gamma}_n)}{\partial \theta \partial \gamma'}\end{pmatrix}'\begin{pmatrix}\hat{\theta}_n-\theta_0 \\ \hat{\gamma}_n-\gamma_0\end{pmatrix}$$

Thus, we can write:

$$\sqrt{n}\begin{pmatrix}\hat{\theta}_n-\theta_0\\ \hat{\gamma}_n-\gamma_0\end{pmatrix}=-C_n^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\partial g(W_i,\theta_0,\gamma_0)}{\partial\theta}$$

Where:

$$C_n=\frac{1}{n}\sum_{i=1}^{n}\begin{pmatrix}\frac{\partial^2 g(W_i,\widetilde{\theta}_n,\gamma_0)}{\partial\theta\partial\theta'}\\ \frac{\partial^2 g(W_i,\theta_0,\widetilde{\gamma}_n)}{\partial\theta\partial\gamma'}\end{pmatrix}'\to_p C_0=\mathbb{E}\left[\begin{pmatrix}\frac{\partial^2 g(W_i,\theta_0,\gamma_0)}{\partial\theta\partial\theta'}\\ \frac{\partial^2 g(W_i,\theta_0,\gamma_0)}{\partial\theta\partial\gamma'}\end{pmatrix}'\right]$$

Then, by the Central Limit Theorem,

$$\sqrt{n}\begin{pmatrix}\hat{\theta}_n-\theta_0\\ \hat{\gamma}_n-\gamma_0\end{pmatrix}\to_d\mathcal{N}\left(0,C_0^{-1}\begin{pmatrix}\Omega_0^{\theta}\\ \Omega_0^{\gamma}\end{pmatrix}(C_0^{-1})'\right)$$