Econ 6200: Econometrics II
Prelim, May 15^{th}, 2023

This exam consists of 12 questions, not of equal length or difficulty, grouped into two exercises. Each question is worth 10 points. Remember to always explain your answer. Throughout, you may invoke theorems from class without proof. It is acceptable to invoke "existence of moments as needed."

Good luck!

1. ML and GMM Consider a model with correctly specified likelihood of individual observation $f(X;\theta)$.

1.1 Show that, if θ_0 is the true parameter value, then $\mathbb{E}\left(\frac{\partial \log f(X;\theta_0)}{\partial\theta}\right) = 0$.

1.2 Where did your proof use that the model is correctly specified?

1.3 Explain how the above result can be used to express the ML estimator as a GMM estimator.

1.4 Will there be an asymptotic efficiency gain from estimating the estimator from 1.3 by Two Stage GMM?

1.5 Define $\theta^* \equiv \arg\max_{\theta\in\Theta} \mathbb{E}(\log f(X,\theta))$ and assume it is unique. Give a set of conditions under which the maximum likelihood estimator $\hat{\theta}$ will be consistent for θ^*.

2. Binary Response and Identification Consider the model

$$Y = \mathbf{1}\{\alpha + \beta X + \varepsilon \geq 0\},$$

where we will generally assume that data are i.i.d. X is a scalar whose support includes $[0, 1]$. (Note that ε need not be independent of X.)

2.1 Argue that, if we assume that the distribution of ε is continuous with 0 on its support and $\text{med}(\varepsilon \mid X) = 0$, then (α, β) is identified, possibly up to some normalization that you provide.

2.2 Argue that, if we rather assume that $\mathbb{E}(\varepsilon \mid X) = 0$, then without further restrictions (α, β) is *not* identified.

3 This question is about extremum estimation with set-valued arg min. Thus, consider the following assumptions on a continuously differentiable, nonstochastic population criterion function $Q : \Theta \to \mathbb{R}$ and stochastic sample criterion function $Q_n : \Theta \to \mathbb{R}$. Recall also that point-set distance of point a from set A is $d(a, A) \equiv \min_{a' \in A} \|a - a'\|$.

1. $\Theta_0 \equiv \arg\min_{\theta\in\Theta} Q(\theta) \neq \emptyset$.
2. For each $\varepsilon > 0$, there exists $\delta > 0$ s.t.
$$\min_{\theta' \in \Theta : d(\theta', \Theta_0) \geq \varepsilon} Q(\theta') \geq \min_{\theta \in \Theta} Q(\theta) + \delta.$$
3. For each $\varepsilon > 0$, $\Pr(\sup_{\theta\in\Theta} |Q(\theta) - Q_n(\theta)| \geq \varepsilon) \to 0$.
4. Θ is a compact subset of $\mathbb{R}^k$, $k < \infty$.

Define $\hat{\theta} \equiv \arg\min_{\theta\in\Theta} Q_n(\theta)$. In case that $\hat{\theta}$ as defined is not unique, *any* selection is acceptable. You may also assume for simplicity that:

- $\min_{\theta\in\Theta} Q(\theta) = 0$,
- any relevant maxima and minima exist.

3.1 Explain how assumption 1 is weaker than a similar assumption from a result proved in class. Show that assumption 2 is not substantially weakened; that is, in the special case where the stronger version of assumption 1 holds, assumption 2 would specialize to an assumption made in class.

3.2 Prove: $d(\hat{\theta}, \Theta_0) \xrightarrow{p} 0$.

3.3 Can it in general be true that $\hat{\theta} \xrightarrow{p} \theta_0$ for *every* $\theta_0 \in \Theta_0$?

3.4 Does your finding from part 2 imply that $\hat{\theta} \xrightarrow{p} \theta_0$ for *some* $\theta_0 \in \Theta_0$?

3.5 Propose a nontrivial (i.e., not equal to Θ) *set-valued* estimator $\widehat{\Theta}$ with the feature that for all $\theta_0 \in \Theta_0$, $d(\theta_0, \widehat{\Theta}) \xrightarrow{p} 0$.

Hint: You may find helpful the following, admittedly contrived example.

- $\Theta = [-1, 1]$,
- $Q(\theta) = 0 \implies \Theta_0 = [-1, 1]$,
- $Q_n(\theta) = (-1)^n \theta / n \implies \hat{\theta} = (-1)^n$.