Econometric Theory 1: Homework 04

b06303009@ntu.edu.tw

December 30, 2021

Problem 01. Consider the model

$$y_i = \mathbf{x}_i'\beta + \varepsilon_i$$

with instruments $\mathbf{z}_i$. The GMM estimator is of the form

$$\hat{\beta}(\hat{\mathbf{W}}) = (\mathbf{S}_{\mathbf{zx}}'\hat{\mathbf{W}}\mathbf{S}_{\mathbf{zx}})^{-1}\mathbf{S}_{\mathbf{zx}}'\hat{\mathbf{W}}\mathbf{s}_{\mathbf{z}y}$$

where $\mathbf{S}_{\mathbf{zx}} = \frac{1}{n}\sum_{i=1}^n \mathbf{z}_i\mathbf{x}_i'$ and $\mathbf{s}_{\mathbf{z}y} = \frac{1}{n}\sum_{i=1}^n \mathbf{z}_i y_i$. Note that we emphasize that the estimator is a function of our choice of weight by denoting it as $\hat{\beta}(\hat{\mathbf{W}})$. By central limit theorem (with assumptions on the cross-moments of ε and $\mathbf{x}_i$), we have

$$\sqrt{n}(\hat{\beta}(\hat{\mathbf{W}}) - \beta_0) \xrightarrow{d} \mathcal{N}\Big(0, \operatorname{Avar}(\hat{\beta}(\hat{\mathbf{W}}))\Big)$$

where

$$\operatorname{Avar}(\hat{\beta}(\hat{\mathbf{W}})) = (\boldsymbol{\Sigma}_{\mathbf{zx}}'\mathbf{W}\boldsymbol{\Sigma}_{\mathbf{zx}})^{-1}\boldsymbol{\Sigma}_{\mathbf{zx}}'\mathbf{W}\boldsymbol{\Omega}\mathbf{W}\boldsymbol{\Sigma}_{\mathbf{zx}}(\boldsymbol{\Sigma}_{\mathbf{zx}}'\mathbf{W}\boldsymbol{\Sigma}_{\mathbf{zx}})^{-1}$$

with $\hat{\mathbf{W}} \xrightarrow{p} \mathbf{W}$, $\mathbf{S}_{\mathbf{zx}} \xrightarrow{p} \boldsymbol{\Sigma}_{\mathbf{zx}}$, and $\boldsymbol{\Omega} = \mathbb{E}\big[\varepsilon^2\mathbf{z}_i\mathbf{z}_i'\big]$. We want to show that in GMM, the most efficient (in the sense the covariance matrix is the smallest) choice of weight matrix is $\hat{\mathbf{W}} \coloneqq \boldsymbol{\Omega}^{-1}$.

Consider another choice of weight matrix $\mathbf{Q}$. We want to show that $\operatorname{Avar}(\hat{\beta}(\mathbf{Q})) - \operatorname{Avar}(\hat{\beta}(\boldsymbol{\Omega}^{-1}))$ is semi-positive definite. Consider the following:

$$\begin{aligned}
&(\boldsymbol{\Sigma}_{\mathbf{zx}}'\mathbf{Q}\boldsymbol{\Sigma}_{\mathbf{zx}})^{-1}\boldsymbol{\Sigma}_{\mathbf{zx}}'\mathbf{Q}\boldsymbol{\Omega}\mathbf{Q}\boldsymbol{\Sigma}_{\mathbf{zx}}(\boldsymbol{\Sigma}_{\mathbf{zx}}'\mathbf{Q}\boldsymbol{\Sigma}_{\mathbf{zx}})^{-1} - (\boldsymbol{\Sigma}_{\mathbf{zx}}'\boldsymbol{\Omega}^{-1}\boldsymbol{\Sigma}_{\mathbf{zx}})^{-1} \\
&= \bigcirc\Big[\boldsymbol{\Sigma}_{\mathbf{zx}}'\mathbf{Q}\boldsymbol{\Omega}\mathbf{Q}\boldsymbol{\Sigma}_{\mathbf{zx}} - (\boldsymbol{\Sigma}_{\mathbf{zx}}'\mathbf{Q}\boldsymbol{\Sigma}_{\mathbf{zx}})(\boldsymbol{\Sigma}_{\mathbf{zx}}'\boldsymbol{\Omega}^{-1}\boldsymbol{\Sigma}_{\mathbf{zx}})^{-1}(\boldsymbol{\Sigma}_{\mathbf{zx}}'\mathbf{Q}\boldsymbol{\Sigma}_{\mathbf{zx}})\Big]\bigcirc \\
&= \bigcirc\Big[\boldsymbol{\Omega} - \boldsymbol{\Sigma}_{\mathbf{zx}}(\boldsymbol{\Sigma}_{\mathbf{zx}}'\boldsymbol{\Omega}^{-1}\boldsymbol{\Sigma}_{\mathbf{zx}})^{-1}\boldsymbol{\Sigma}_{\mathbf{zx}}'\Big]\bigcirc \\
&= \bigcirc\Big[I - (L')^{-1}\boldsymbol{\Sigma}_{\mathbf{zx}}(\boldsymbol{\Sigma}_{\mathbf{zx}}'\boldsymbol{\Omega}^{-1}\boldsymbol{\Sigma}_{\mathbf{zx}})^{-1}\boldsymbol{\Sigma}_{\mathbf{zx}}'L^{-1}\Big]\bigcirc \\
&= A[I - B]A'
\end{aligned}$$

where the first equality is obtain by factoring out $(\boldsymbol{\Sigma}_{\mathbf{zx}}'\mathbf{Q}\boldsymbol{\Sigma}_{\mathbf{zx}})^{-1}$ on both sides; the second equality is obtain by factoring out $\boldsymbol{\Sigma}_{\mathbf{zx}}'\mathbf{Q}$ in the front and $\mathbf{Q}\boldsymbol{\Sigma}_{\mathbf{zx}}$ at the back; the penultimate equality is obtain by decomposing $\boldsymbol{\Omega}$ as $L'L$ (since $\boldsymbol{\Omega}$ is positive definite, there exists a decomposition where L is invertible); the ultimate equality is to redefine the matrices in simpler notation.

Notice that B is symmetric and idempotent. Thus, $I - B$ is also idempotent. We have

$$A(I-B)A' = A(I-B)(I-B)A' \geq 0 \qquad \text{(positive definite)}$$

Therefore, $\hat{\mathbf{W}} = \boldsymbol{\Omega}^{-1}$ is the most efficient choice of weights. ■

Problem 02. We want to show that

$$\mathbb{E}\left[\ddot{\ell}(X;\theta) \,\middle|\, \theta\right] = -\,\mathbb{E}\left[\dot{\ell}(X;\theta)^2 \,\middle|\, \theta\right]$$

where ℓ denotes the log-likelihood function and dot(s) denote partial derivative with respect to θ.

Consider the following:

$$\begin{aligned}
\mathbb{E}\left[\ddot{\ell} \,\middle|\, \theta\right] = \int_A \ddot{\ell} f \, dx &= \int_A \left(-f^{-2}\dot{f}\dot{f} + f^{-1}\ddot{f}\right) f \, dx && \text{(def. of derivative)} \\
&= -\int_A \dot{f}\dot{f}f^{-2}f + \ddot{f} \, dx \\
&= -\int_A \dot{\ell}^2 f \, dx + \frac{\partial^2}{\partial\theta^2}\int_A f \, dx && \text{(def. of } \ell \text{ and Leibniz rule)} \\
&= -\,\mathbb{E}\left[\dot{\ell}^2 \,\middle|\, \theta\right] + \cancelto{0}{\frac{\partial^2}{\partial\theta^2}\int_A f \, dx} && (f \text{ integrates to } 1) \\
&= -\,\mathbb{E}\left[\dot{\ell}^2 \,\middle|\, \theta\right]
\end{aligned}$$

where A is the support of X and f is the PDF of X. (PDF of X exists since we assume the likelihood function exists.) ■