Patrik Guggenberger 4/2/2009
Economics 203C (Spring, 2009)

Problem Set I (due 4/14/9)

1. Examine the finite-sample properties of two GMM estimators via a Monte Carlo simulation. Consider the linear model $y_i = x_i\beta + \varepsilon_i \in \mathbb{R}$ for $i = 1, 2, ..., n$, where the scalar β equals 0. Suppose the $z_i \in \mathbb{R}^K$ are *iid* $N(0, \Lambda)$–vectors of instrumental variables, where $\Lambda \in \mathbb{R}^{K\times K}$ is the positive definite matrix given by DD' for a nonsingular matrix $D \in \mathbb{R}^{K\times K}$ of your choice (but don't pick the identity matrix). Assume that $x_i = z_i'\pi + u_i$, where (ε_i, u_i) are *iid* $N(0, \Omega)$, where Ω is a 2×2-matrix with diagonal and off diagonal elements 1 and ρ, respectively. Let $\pi = (\eta, \eta, \ldots, \eta)'$. Set $n = 100$ and simulate $R = 1000$ data samples for all 8 parameter combinations of $K = (1; 10)$, $\eta = (.05; 1)$ and $\rho = (0; .5)$.

The definition of the GMM stochastic criterion function depends on the weighting matrix A_n. The first estimator sets $A_n := I_K$. The second estimator (called "two-stage least squares", TSLS) picks A_n such that

$$\begin{aligned} A_n'A_n &= \left(\frac{1}{n}\sum_{i=1}^{n} Z_i Z_i'\right)^{-1}, \text{ and is given by} \\ \widehat{\beta} &= (X'P_Z X)^{-1} X'P_Z Y, \text{ where} \\ P_Z &= Z(Z'Z)^{-1}Z', \end{aligned}$$

and X, Y are the n-vectors with components x_i and y_i, and Z is the $n\times K$-matrix with rows z_i'.

In each of the 8 cases and each sample, calculate the two estimators. In each of the 8 cases, report summary statistics (bias, standard deviation, and mean squared error) of the estimators. Interpret the parameters η and ρ and discuss your results. Are your finite sample findings consistent with the asymptotic efficiency results?

2. Review of some mathematical notions that came up in class:

(i) Prove, starting from the definition of convergence in probability:

(a) $o_p(1) + o_p(1) = o_p(1)$, $o_p(1)o_p(1) = o_p(1)$.

(b) Assume the function g is continuous at a. Show that $g(a + o_p(1)) = g(a) + o_p(1)$.

(ii) Prove that for real-valued sequences (a_n) and (b_n), $n \in \mathbb{N}$, we have

$$\text{limsup}(a_n + b_n) \leq \text{limsup}(a_n) + \text{limsup}(b_n).$$

Do we even have equality?

3. Minimum distance (MD) estimation of a parameter vector θ_0 is another example of extremum estimation.

Let $\widehat{\pi}_n$ be an estimator of a k-vector parameter π_0. Suppose π_0 is known to be a function of $\theta_0 \in \mathbb{R}^d$, where $\pi_0 = g(\theta_0)$ for a certain known function $g(\cdot)$.[1]

[1] Clarification: The estimator $\widehat{\pi}_n$ is not defined as $g(\widehat{\theta})$. The estimator $\widehat{\pi}_n$ is observed and the estimator $\widehat{\theta}$ is defined as the θ that minimizes the distance between $\widehat{\pi}_n$ and $g(\theta)$.

Let A_n be a $k \times k$ random weight matrix. Define the MD estimator $\widehat{\theta}_n$ as the minimizer of

$$Q_n(\theta) = ||A_n(\widehat{\pi}_n - g(\theta))||^2/2$$

over $\theta \in \Theta$ where Θ is the parameter space. Under appropriate assumptions prove consistency of the MD estimator. [Assume 1) $A_n \to_p A$ and $\widehat{\pi}_n \to_p \pi_0$ where A is nonsingular 2) there exists a unique value $\theta_0 \in \Theta$ such that $\pi_0 = g(\theta_0)$ and maybe more.] What is Q in this example?

4. Prove that Assumption ID is implied by ID1. Provide examples that show that any two of the three conditions in ID1 are not enough to imply ID.