Patrik Guggenberger 4/10/2009
Economics 203C (Spring, 2009)

Problem Set II (due 4/21/9)

1) Prove the following statement: *Suppose* (i) $\widehat{\beta}_n \to_p \beta_0 \in R^s$, (ii) $\sup_{\gamma\in\Gamma}\sup_{\beta\in B(\beta_0,\varepsilon)} |L_n(\gamma,\beta) - L(\gamma,\beta)| \to_p 0$ *for some* $\varepsilon > 0$, *and* (iii) $L(\gamma,\beta)$ *is continuous in* β *at* β_0 *uniformly over* $\gamma \in \Gamma$ (*i.e.,* $\lim_{\beta\to\beta_0}\sup_{\gamma\in\Gamma} |L(\gamma,\beta) - L(\gamma,\beta_0)| = 0$.) *Then,*

$$\sup_{\gamma\in\Gamma} |L_n(\gamma,\widehat{\beta}_n) - L(\gamma,\beta_0)| \xrightarrow{p} 0. \tag{1}$$

Note that condition (iii) holds if Γ is compact and $L(\gamma,\beta)$ is continuous in (γ,β) on $\Gamma \times B(\beta_0,\varepsilon)$.

2) (a) Covariance estimation: Suppose $Y_t = X_t'\beta + U_t$ for $t = 1,...,T$, $\{(X_t,U_t) : t \geq 1\}$ are *iid* with $E(U_t|X_t) = 0$ a.s. and $E(U_t^2|X_t) < \infty$ a.s.. The least squares estimator $\widehat{\beta}_{LS}$ satisfies

$$\sqrt{T}(\widehat{\beta}_{LS} - \beta) \to_d Z \sim N(0,\Omega) \text{ as } T \to \infty, \tag{2}$$

where $\Omega := (EX_tX_t')^{-1}EU_t^2X_tX_t'(EX_tX_t')^{-1}$. Show that the so-called Eicker-White estimator

$$\widehat{\Omega} := (\frac{1}{T}\sum_{t=1}^{T} X_tX_t')^{-1}\frac{1}{T}\sum_{t=1}^{T}\widehat{U}_t^2X_tX_t'(\frac{1}{T}\sum_{t=1}^{T} X_tX_t')^{-1}, \tag{3}$$

where $\widehat{U}_t := Y_t - X_t'\widehat{\beta}_{LS}$, satisfies $\widehat{\Omega} \to_p \Omega$ as $T \to \infty$. State clearly in each step of your argument which rule you use and which assumptions are needed.

(b) In the linear regression model of part a) with $\sigma^2 = E(U_t^2|X_t) < \infty$ a.s. (conditional homoskedasticity), design a Monte Carlo study in which you compare the precision of the Eicker-White estimator with the one of the covariance matrix estimator that assumes conditional homoskedasticity. For simplicity take $EX_tX_t' = I_k$. Report finite sample bias, variance, and mean-squared error of the variance estimators. Experiment with various error distributions. Note that whenever you pin down distributions for U and X, you know what the true Ω is.

3) (a) For the MD estimator, discussed on problem set 1, establish asymptotic normality of $\widehat{\theta}_n$. [Assume here that $\sqrt{n}(\widehat{\pi}_n - \pi_0) \to_d N(0,V_0)$][1].

(b) Another example of an extremum estimator is the Two-step (TS) Estimator: Suppose the data $\{W_i : i \leq n\}$ are iid, $\widehat{\tau}_n$ is a preliminary consistent estimator of a parameter τ_0, $G_n(\theta,\tau)$ is a random k-vector that should be close to 0 when $\theta = \theta_0$, $\tau = \tau_0$, and n is large (e.g., in the GMM case, $G_n(\theta,\tau) = \frac{1}{n}\sum_{i=1}^n g(W_i,\theta,\tau)$ and in the MD case $G_n(\theta,\tau) = \widehat{\pi}(\tau) - g(\theta,\tau)$), and A_n is a $k \times k$ random weight matrix. Then, the TS estimator $\widehat{\theta}_n$ minimizes

$$Q_n(\theta) = ||A_nG_n(\theta,\widehat{\tau}_n)||^2/2 \tag{5}$$

[1] **Additional hints:** When you differentiate $Q_n(\theta)$ with respect to θ, then $\widehat{\pi}_n$ is to be treated as a vector of constants. The function $g(\theta)$ does not depend on the data. When working out what B_0 and Ω_0 are, replace estimators by their probability limits, e.g. $\widehat{\pi}_n$ by π_0. To show asymptotic normality, let $B_0 \equiv \frac{\partial^2}{\partial\theta\partial\theta'}Q(\theta_0) > 0$. Provide primitive assumptions for the assumption $\sqrt{n}\frac{\partial}{\partial\theta}Q_n(\theta_0) \to_d N(0,\Omega_0)$. What is Ω_0? Then use a Taylor expansion

$$0 = \frac{\partial}{\partial\theta}Q_n(\widehat{\theta}_n) = \frac{\partial}{\partial\theta}Q_n(\theta_0) + \frac{\partial^2}{\partial\theta\partial\theta'}Q_n(\theta_n^*)(\widehat{\theta}_n - \theta_0), \tag{4}$$

where θ_n^* lies between $\widehat{\theta}_n$ and θ_0. Solve for $(\widehat{\theta}_n - \theta_0)$...

over $\theta \in \Theta$. Under appropriate assumptions (including $G_n(\theta,\tau) \xrightarrow{p} G(\theta,\tau)$) what is $Q(\theta)$? Give primitive conditions that imply consistency (in particular, when does θ_0 uniquely minimize $Q(\theta)$ over Θ?) Asymptotic normality will be discussed in the TA section.

4) Investigate how well the asymptotic normal distribution of the two-stage least squares (TSLS) estimator approximates its finite sample distribution. Consider the linear model

$$y_i = x_i\beta + \varepsilon_i \in \mathbb{R} \tag{6}$$

for $i = 1, 2, ..., n$, where the scalar β equals 0. Suppose $z_i \in \mathbb{R}^K$ are *iid* $N(0, I_K)$. Assume that

$$x_i = z_i'\pi + u_i, \tag{7}$$

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 $\pi = (\eta, \eta, \ldots, \eta)'$.

(i) Derive the asymptotic distribution of the TSLS estimator in this model. Does it depend on K, η or ρ?

Let $n = 100$ and simulate $R = 5000$ data samples for all 8 parameter combinations of $K = (1; 10)$, $\eta = (.05; 1)$ and $\rho = (0; .95)$.

(ii) Plot the finite sample distribution of the TSLS estimator. Use separate graphs for each of the eight parameter combinations. Does the distribution change with K, η or ρ?

(iii) Simulate the asymptotic normal distribution you derived in (i) for the TSLS estimator and plot it in the corresponding graphs of (ii). Comment on the goodness of the approximation.