Econometric Theory, **17**, 2001, 1025–1031. Printed in the United States of America.

PROBLEMS AND SOLUTIONS

PROBLEMS

01.5.1. *Some Properties of a Generalized Two-Error Components Matrix*, proposed by Franc J.G.M. Klaassen and Jan R. Magnus. The standard two-error components model has a variance matrix where all variances (diagonal elements) are equal to each other and all covariances (off-diagonal elements) are equal. The most convenient way to write the matrix is as $\alpha J + \beta(I_T - J)$, where $J = (1/T)\iota\iota'$, ι denotes a $T \times 1$ vector of ones, and I_T denotes the identity matrix of order T. Consider a generalization of this matrix, namely the $T \times T$ matrix

$$\Omega = \begin{pmatrix} \alpha_1 J_1 + \beta_1(I_{T_1} - J_1) & \gamma\iota_1\iota_2' \\ \gamma\iota_2\iota_1' & \alpha_2 J_2 + \beta_2(I_{T_2} - J_2) \end{pmatrix},$$

where ι_1 and ι_2 denote vectors of ones of dimension T_1 and T_2, respectively, $J_1 = (1/T_1)\iota_1\iota_1'$, $J_2 = (1/T_2)\iota_2\iota_2'$, and $T = T_1 + T_2$. Let $\Delta = \alpha_1\alpha_2 - \gamma^2 T_1 T_2$. Show that:

(a) The determinant of Ω is given by $|\Omega| = \beta_1^{T_1-1}\beta_2^{T_2-1}\Delta$.
(b) Ω is positive definite if and only if α_1, α_2, β_1, β_2, and Δ are all positive.
(c) Assume that Ω is positive definite. Then Ω^p has the same structure as Ω for any real p. In particular,

$$\Omega^p = \begin{pmatrix} \phi_1 J_1 + \beta_1^p(I_{T_1} - J_1) & \delta\iota_1\iota_2' \\ \delta\iota_2\iota_1' & \phi_2 J_2 + \beta_2^p(I_{T_2} - J_2) \end{pmatrix}.$$

(d) Obtain ϕ_1, ϕ_2, and δ for the special cases $p = -1$ and $p = -\frac{1}{2}$.
(e) Consider the transformation $y = \Omega^p x$. Let $x' = (x_1', x_2')$ and $y' = (y_1', y_2')$, where x_1 and y_1 are of dimension T_1 and x_2 and y_2 are of dimension T_2. Then,

$$y_{1t} = y_1^* + \beta_1^p x_{1t} (t = 1, \ldots, T_1), \qquad y_1^* = (\phi_1 - \beta_1^p)\bar{x}_1 + \delta T_2 \bar{x}_2,$$

and

$$y_{2t} = y_2^* + \beta_2^p x_{2t} (t = 1, \ldots, T_2), \qquad y_2^* = (\phi_2 - \beta_2^p)\bar{x}_2 + \delta T_1 \bar{x}_1,$$

where $\bar{x}_1 = (1/T_1)\iota_1' x_1$ and $\bar{x}_2 = (1/T_2)\iota_2' x_2$. This property (for $p = -\frac{1}{2}$) is especially useful in obtaining (feasible) GLS estimates, as in Klaassen and Magnus (2001).

REFERENCES

Klaassen, F.J.G.M. & J.R. Magnus (2001) Are points in tennis independent and identically distributed? Evidence from a dynamic binary panel data model. *Journal of the American Statistical Association* 96, forthcoming.

© 2001 Cambridge University Press 0266-4666/01 $9.50

https://doi.org/10.1017/S0266466601005072 Published online by Cambridge University Press

01.5.2. *The Characteristic Function from a Family of Truncated Normal Distributions*, proposed by William C. Horrace and Clemente Hernandez. Truncated distributions (and in particular truncated normal distributions) are used extensively in the econometrics literature: truncated regression, the Geweke–Hajivassiliou–Keane simulator, and cross-sectional stochastic frontier models, to name a few. Derivation of the characteristic function may prove useful for density estimators in the style of Horowitz and Markatou (1996), where one variate of an error component model is known to be from the family of truncated normal distributions.

Let W be a random variate whose density function can be parameterized in terms of location μ and scale σ^2. Let X be a doubly truncated random variate constructed from the arbitrary truncation of W. Formulate an expression for the characteristic function of X and specialize that expression to derive the characteristic function of a doubly truncated normal distribution. Use the truncated normal characteristic function to calculate the mean of the distribution.

REFERENCE

Horowitz, J.L. & M. Markatou (1996) Semiparametric estimation of regression models for panel data. *Review of Economic Studies* 63, 145–168.

01.5.3. *Serial Correlation and Asymptotic Variance*, proposed by Stanislav Anatolyev. Consider the stationary time series regression

$$y_t = \alpha + \beta x_t + e_t,$$

where the zero mean error e_t is uncorrelated with the regressor x_t. The object of interest is the asymptotic variance V_β of the OLS estimator of β.

(a) Construct an example where e_t is serially correlated, but nevertheless $V_\beta = (Var(e_t)/Var(x_t))$. That is, one does *not* have to correct the standard error for serial correlation, even though serial correlation is present.
(b) Construct an example where e_t is serially uncorrelated but nevertheless $V_\beta \neq (Var(x_t e_t)/[Var(x_t)]^2)$. That is, one *has* to correct the standard error for serial correlation, even though there is no serial correlation in the error term.
(c) What implication do the results in (a) and (b) have regarding pretesting for serial correlation?

SOLUTIONS

00.5.1. *An Exact Fitting Estimator for Linear Regression*—Solution[1] proposed by Oliver Linton. If the ranking was based on the true errors rather than the residuals, we would have

https://doi.org/10.1017/S0266466601005072 Published online by Cambridge University Press

$$\tilde{\beta} - \beta = \frac{u_{(1)}}{x_{(1)}}.$$

In this case, consistency immediately follows from the fact that $\min_{1\le i\le n}|u_i| \to_p 0$. In fact, we have $\tilde{\beta} - \beta = O_p(n^{-1})$. Unfortunately, this superconsistency vanishes when we use only the data, as we know from standard regression theory. When we use the OLS estimator $\hat{\beta}$ to construct the residuals, we have instead

$$\tilde{\beta} - \beta = \frac{y_{(1)}}{x_{(1)}} - \frac{\beta x_{(1)}}{x_{(1)}} = \frac{u_{(1)}}{x_{(1)}},$$

where the ordering is on the value of the least squares residuals. We next substitute $u_{(1)} = \hat{u}_{(1)} + (\hat{\beta} - \beta)x_{(1)}$ to obtain

$$\tilde{\beta} - \beta = \frac{\hat{u}_{(1)}}{x_{(1)}} + \hat{\beta} - \beta,$$

where the ordering of the numerator is now aligned, i.e., $\hat{u}_{(1)} = \min_{1\le i\le n}|\hat{u}_i|$. By the triangle inequality

$$\min_{1\le i\le n}|\hat{u}_i| \le \min_{1\le i\le n}|u_i| + |\hat{\beta} - \beta| \min_{1\le i\le n}|x_i| \to_p 0,$$

which implies that $\tilde{\beta}$ is consistent. We now find the rate of convergence and limiting distribution. The triangle inequality argument will not suffice because $|\hat{\beta} - \beta| \min_{1\le i\le n}|x_i| = O_p(n^{-1/2})$. Write

$$\sqrt{n}(\tilde{\beta} - \beta) = \frac{\sqrt{n}\min_{1\le i\le n}\left|u_i - \frac{1}{\sqrt{n}}x_i\sqrt{n}(\tilde{\beta} - \beta)\right|}{x_{(1)}} + \sqrt{n}(\hat{\beta} - \beta).$$

We show that

$$\sqrt{n}\min_{1\le i\le n}\left|u_i - \frac{1}{\sqrt{n}}x_i\sqrt{n}(\hat{\beta} - \beta)\right| = o_p(1), \tag{1}$$

in which case

$$\sqrt{n}(\tilde{\beta} - \beta) = \sqrt{n}(\hat{\beta} - \beta) + o_p(1)$$

because $x_{(1)}$ is bounded away from zero.

Proof of (1). We can restrict our attention to the set $\mathcal{A} = \{|\sqrt{n}(\hat{\beta} - \beta)| \le c \log n\}$ for some $c > 0$ because for any event $\mathcal{B}$ we have

$$\Pr[\mathcal{B}] \le \Pr[\mathcal{B} \cap \mathcal{A}] + \Pr[\mathcal{A}^c] = \Pr[\mathcal{B} \cap \mathcal{A}] + o(1).$$

https://doi.org/10.1017/S0266466601005072 Published online by Cambridge University Press

Define the stochastic process

$$G_n(\gamma) = \sqrt{n} \min_{1\le i\le n} \left| u_i - \frac{1}{\sqrt{n}} x_i \gamma \right|$$

for any γ. Then, on the set $\mathcal{A}$, we have

$$\sqrt{n} \min_{1\le i\le n} \left| u_i - \frac{1}{\sqrt{n}} x_i \sqrt{n}(\hat{\beta} - \beta) \right| \le \sup_{|\gamma|\le c \log n} G_n(\gamma).$$

Now let $\gamma_1, \ldots, \gamma_J$ be the centers of balls of radius $c \log n/J(n)$ that cover the set $\{|\gamma| \le c \log n\}$. Then

$$\sup_{|\gamma|\le c \log n} G_n(\gamma) \le \max_{1\le j\le J} G_n(\gamma_j) + \max_{1\le j\le J} \sup_{|\gamma-\gamma_j|<c \log n/J(n)} |G_n(\gamma_j) - G_n(\gamma)|.$$

We now use the inequalities

$$\left| \min_{1\le i\le n} |a_i| - \min_{1\le i\le n} |b_i| \right| \le \max_{1\le i\le n} ||a_i| - |b_i|| \tag{2}$$

$$||c - d| - |c|| \le |d|, \tag{3}$$

which are true for any real numbers a_i, b_i, c, and d. Taking $a_i = \sqrt{n}(u_i - (1/\sqrt{n})x_i\gamma)$, $b_i = \sqrt{n}(u_i - n^{-1/2}x_i\gamma_j)$, $c = \sqrt{n}u_i - x_i\gamma$, and $d = x_i(\gamma_j - \gamma)$, we obtain for any γ, γ_j

$$|G_n(\gamma_j) - G_n(\gamma)| \le \max_{1\le i\le n} |x_i| \cdot |\gamma - \gamma_j|$$

so that

$$\max_{1\le j\le J} \sup_{|\gamma-\gamma_j|<c \log n/J(n)} |G_n(\gamma_j) - G_n(\gamma)| \le \frac{c \log n}{J(n)}. \tag{4}$$

Provided $J(n)/\log n \to \infty$, the random variable on the left-hand side of (4) is $o_p(1)$.

Now, by the Bonferroni inequality

$$\Pr\left[\max_{1\le j\le J} G_n(\gamma_j) > \eta \right] \le \sum_{j=1}^{J} \Pr[G_n(\gamma_j) > \eta].$$

We prove that

$$\Pr[G_n(\gamma) > \eta] = o(n^{-\kappa}) \quad \text{for some } \kappa > 0$$

for any fixed number γ. We have

$$\Pr[G_n(\gamma) \ge \eta] = \Pr\left[\sqrt{n} \left| u_i - \frac{1}{\sqrt{n}} x_i \cdot \gamma \right| \ge \eta \right]^n,$$

https://doi.org/10.1017/S0266466601005072 Published online by Cambridge University Press

where

$$\Pr\left[\sqrt{n}\left|u_i - \frac{1}{\sqrt{n}}\, x_i\cdot\gamma\right| \geq \eta\right]$$

$$= 1 - \int \left[F_u\left(\frac{1}{\sqrt{n}}\,(x\cdot\gamma + \eta)\right) - F_u\left(\frac{1}{\sqrt{n}}\,(x\cdot\gamma - \eta)\right)\right] f_x(x)\, dx$$

$$\simeq 1 - \frac{2\eta}{\sqrt{n}}\, f_u(0)$$

by Taylor series expansion about the origin using the cancellation that eliminates $x\gamma/\sqrt{n}$. Therefore,

$$\Pr\left[\sqrt{n}\left|u - \frac{1}{\sqrt{n}}\, x\cdot\gamma\right| \geq \eta\right]^n \simeq \left(1 - \frac{2\eta}{\sqrt{n}}\, f_u(0)\right)^n = o(n^{-\kappa})$$

for any $\kappa > 0$. Therefore, $\Pr[\max_{1\leq j\leq J} G_n(\gamma_j) > \eta] = o(1)$ and we have shown that $\sup_{|\gamma|\leq c\log n} G_n(\gamma) = o_p(1)$.

Simulations reveal that the estimator is indeed very close to the OLS estimator in this case. However, when x has support that includes zero or the error distribution is zero at the origin, the rate of convergence and limiting distribution can be different and can be worse than OLS.

NOTE

1. An excellent solution was proposed independently by Xing Ming.

00.5.2. *When Are Nested Reduced Rank Autoregressive Processes Integrated?*—Solution proposed by Paolo Paruolo. Let $\Phi(L) := I - \sum_{i=1}^{p} \Phi_i L^i$, and $\bar{a} := a(a'a)^{-1}$ for any full column rank matrix a. Let also $\phi := \sum_{i=1}^{p} \xi_i$, $\zeta := \sum_{i=1}^{p} i\xi_i$ and $s := rank(\phi)$, such that $\phi' = \rho\eta'$, where ρ is $q \times s$ and η is $m \times s$, both of full rank s. Note that the decomposition $\phi' = \rho\eta'$ includes the case of full rank $s = q$ and all cases of deficient rank. Define $\tau := \gamma\rho$ such that $\gamma\phi' = \tau\eta'$, where τ is $m \times s$ and of full rank s. The following proposition states necessary and sufficient conditions under which the AR process Y_t in (1) under the restriction (2) is I(d) at frequency zero.

PROPOSITION 1. *The process Y_t is $I(d)$, $d = 1,2,\ldots$, at frequency zero if and only if at least one eigenvalue of the $s \times s$ matrix $\eta'\tau$ is equal to 1. Let u be the number of such eigenvalues equal to 1, $u \leq s$; then $r := rank(\Phi(1)) = m - u$. If $u > 0$, then $I_s - \eta'\tau = ab'$ with a and b $(s \times u)$ matrices of full column rank, and $\Phi(1)$ has representation $\Phi(1) = \alpha\beta'$, with $\alpha = (\bar{\eta}a, \eta_\perp)$, $\beta =$*

https://doi.org/10.1017/S0266466601005072 Published online by Cambridge University Press

$(\bar{\tau}b, \tau_\perp)c$ full rank $m \times r$ matrices, where c is some $r \times r$ nonsingular matrix. In this case, the process Y_t is $I(1)$ at frequency zero if and only if the matrix

$$a_\perp' \eta' \gamma \zeta' \tau b_\perp \tag{4}$$

has full rank $s - u$.

COROLLARY 2. *If ϕ is of full rank q (such that $s = q$, $\phi = \eta$, $\gamma = \tau$), then matrix (4) reduces to $a_\perp' \zeta' \tau b_\perp$.*

Proof. A necessary and sufficient condition for an AR process to be $I(d)$, $d > 0$, is that the characteristic equation $|\Phi(z)| = 0$ has solution $z = 1$, i.e., $\Phi(1)$ has reduced rank. By the above definitions, one finds that $\Phi(1) = I_m - \gamma\phi' = I_m - \tau\eta'$. By orthogonal projections on *span*(τ), one finds

$$\Phi(1) = (\tau, \tau_\perp)(\bar{\tau}, \bar{\tau}_\perp)'\Phi(1) = (\tau, \tau_\perp)(\bar{\tau} - \eta, \bar{\tau}_\perp)'(\tau, \tau_\perp)(\bar{\tau}, \bar{\tau}_\perp)'$$

$$= (\tau, \tau_\perp)\begin{pmatrix} I_s - \eta'\tau & \eta'\tau_\perp \\ 0 & I_{m-s} \end{pmatrix}(\bar{\tau}, \bar{\tau}_\perp)', \tag{5}$$

which shows that $\Phi(1)$ has reduced rank iff $I_s - \eta'\tau$ has reduced rank, i.e., iff 1 is an eigenvalue of $\eta'\tau$. If $I_s - \eta'\tau$ has reduced rank, then it admits representation $I_s - \eta'\tau = ab'$ with a and b, $m \times u$ matrices of rank u, and by (5) one obtains the claims of the proposition regarding α. A similar approach based on orthogonal projections on *span*(η) gives the results about β. Observe that $\alpha_\perp = \eta a_\perp$ and $\beta_\perp = \tau b_\perp$.

As proved in Johansen (1996), Theorem 4.2, an integrated process is $I(1)$ iff $\Phi(1) = \alpha\beta'$ has reduced rank r and $\alpha_\perp' \dot{\Phi}(1)\beta_\perp$ has full rank $p - r$, where $\dot{\Phi}(z) = d\Phi(z)/dz$. It is simple to verify that $\dot{\Phi}(1) = -\gamma\zeta'$, such that $-\alpha_\perp' \dot{\Phi}(1)\beta_\perp$ has the form in (4). This completes the proof. ■

Thus, AR processes subject to the constraint (2) may be $I(1)$ under the conditions stated in Proposition 1, or even $I(d)$, $d > 1$, if matrix (4) is not of full rank.

As an illustration of Proposition 1, consider the numerical example of Section 5.4 in Chapter 5 of Reinsel and Velu (1998, pp. 126–129). Reinsel and Velu estimated a 5-dimensional AR(2) with $\hat{F} := (\hat{\Phi}_1, \hat{\Phi}_2) = \hat{\gamma}\hat{\xi}'$, where $\hat{\gamma}$ is of dimension 5×3 and $\hat{\xi}'$ is of dimension 3×10; their estimates appear on page 129. The corresponding $\hat{\phi} = \hat{\xi}_1 + \hat{\xi}_2$ is of rank 3 and thus $\hat{\gamma} = \hat{\tau}$, $\hat{\eta} = \hat{\phi}$; one also finds the following value of $\hat{\eta}'\hat{\tau}$:

$$\hat{\eta}'\hat{\tau} = \begin{pmatrix} 0.9897 & 0.0203 & -0.0050 \\ 0.0376 & 0.7413 & -0.2932 \\ -0.0273 & -0.1638 & 0.4499 \end{pmatrix}.$$

https://doi.org/10.1017/S0266466601005072 Published online by Cambridge University Press

The eigenvalues of $\hat{\eta}'\hat{\tau}$ are real and equal to 0.996743, 0.851581, and 0.332576; the first one and possibly the second one are close to 1. Thus, the estimated reduced rank AR suggests that this dataset may well have been generated by a nonstationary integrated process. This finding is consistent with later calculations by Reinsel and Velu regarding the cointegration rank on the same data, performed in Chapter 5.5, where they selected $r = 3$.

REFERENCES

Johansen, S. (1996) *Likelihood Based Inference on Cointegrated Vector Autoregressive Models*. Oxford: Oxford University Press.

Reinsel, G.C. & P. Velu (1998) *Multivariate Reduced Rank Regression, Theory and Applications*. Lecture Notes in Statistics 136. New York: Springer-Verlag.

https://doi.org/10.1017/S0266466601005072 Published online by Cambridge University Press