Econometric Theory, **16**, 2000, 143–150. Printed in the United States of America.

PROBLEMS AND SOLUTIONS

PROBLEMS

00.1.1. *Determinant of a Skew–Symmetric Matrix*, proposed by Steve Lawford. Let $S = \{s_{ij}\}$ be an $n \times n$ skew–symmetric matrix ($n \geq 2$). Show by induction that S is singular for n odd and nonsingular for n even.

00.1.2. *A Necessary and Sufficient Condition of a Sequence of Random Variables Converging to a Normal Distribution*, proposed by Hua Liang. Prove the following theorem.

THEOREM. *Let $X_1, \ldots, X_n$ be i.i.d. random variables with non-degenerate distributions. There exist $B_n > 0$ and A_n such that $B_n^{-1}(\sum_{i=1}^n X_i - A_n)$ converges to the standard normal distribution $N(0,1)$ if and only if*

$$\lim_{x\to\infty} \frac{P(|X_1| > \lambda x)}{P(|X_1| > x)} = \lambda^{-2} \textit{ for } \lambda > 0.$$

00.1.3. *A Relationship Satisfied by Two Representations of a Positive Semi-Definite Matrix*, proposed by Heinz Neudecker and Michel van de Velden. Consider (real) matrices F and G of order $m \times n$ and rank r. Let

$$FF' = GG'.$$

Prove then that $F = GQ$, where QQ' (and hence $Q'Q$) is symmetric idempotent of rank r and k can be chosen to be any integer in the closed interval $[r, n]$.

SOLUTIONS

99.2.1. *An Empirical Likelihood Ratio Test for a Unit Root*—Solution,[1] proposed by Nikolay Gospodinov and Victoria Zinde-Walsh. Begin by noting a few results about the sums $\sum \psi_t^l$ for $l = 1,2,3$. First recall

$$\frac{1}{T}\sum \psi_t = \frac{1}{T}\sum y_{t-1}u_t \Rightarrow \frac{\sigma^2}{2}(W(1)^2 - 1); \tag{1}$$

also

$$\frac{1}{T^2}\sum y_{t-1}^2 \Rightarrow \sigma^2 \int W(r)^2\, dr \tag{2}$$

(see, e.g., Phillips, 1987, pp. 296–297).

https://doi.org/10.1017/S0266466600161109 Published online by Cambridge University Press

To obtain limits for $l = 2$ and 3, moments have to be evaluated. Consider a sum

$$\sum y_0^{v_0} u_{i_1}^{v_1} \cdots u_{i_k}^{v_k}$$

such that for each term $i_1 \neq i_2 \neq \ldots \neq i_k$, $1 \leq i_j \leq T$; $v_i \leq 6$ for $i > 0$; and $v_0 + v_1 + \ldots + v_k = m$. Then expectation of a term is zero if any of the $v_1, \ldots, v_k$ equals 1; $E(y_0^{v_0} u_{i_1}^{v_1} \cdots u_{i_k}^{v_k})$ is bounded by

$$B = \max_{v_0 + v_1 + \ldots + v_k = m} y_0^{v_0} E(u^{v_1}) \cdots E(u^{v_k}).$$

Thus, the expectation of the sum is bounded by B times the number of contributing terms (where none of the $v_1, \ldots, v_k$ equals one).

It is easy to check that in the expression for $(\psi_t^2 - y_{t-1}^2 \sigma^2)^2$ there is at most $O(T^2)$ such contributing terms; thus, $E(\frac{1}{T}(\psi_t^2 - y_{t-1}^2 \sigma^2))^2$ is uniformly bounded; by martingale Law of Large Numbers (e.g., White, 1984, p. 58), it follows that

$$\frac{1}{T^2} \sum (\psi_t^2 - y_{t-1}^2 \sigma^2) \xrightarrow{a.s.,p} 0,$$

thus recalling (2)

$$\frac{1}{T^2} \sum \psi_t^2 \Rightarrow \sigma^4 \int W(r)^2 \, dr. \tag{3}$$

In ψ_t^3 the number of contributing terms and, thus, $E(\psi_t^3)$ is at most $O(T)$; similarly, by noting the number of contributing terms

$$E\left[\frac{1}{T^{3/2}} (\psi_t^3 - E(\psi_t^3))\right]^2 = O(1).$$

By martingale Law of Large Numbers,

$$\frac{1}{T^{5/2}} \sum (\psi_t^3 - E(\psi_t^3)) \xrightarrow{a.s.,p} 0,$$

and thus noting the lower order of $E(\psi_t^3)$,

$$\frac{1}{T^{5/2}} \sum \psi_t^3 \xrightarrow{a.s.,p} 0. \tag{4}$$

Now construct the Lagrangian for the maximization problem:

$$L(p_1, \ldots, p_T, \lambda, \mu) = \sum_{t=1}^{T} \log p_t - \lambda \sum_{t=1}^{T} p_t \psi_t - \mu \left(\sum_{t=1}^{T} p_t - 1 \right)$$

https://doi.org/10.1017/S0266466600161109 Published online by Cambridge University Press

over the domain $p_1 > 0, \ldots, p_T > 0$. The first-order conditions are

$$\frac{1}{p_t} - \lambda\psi_t - \mu = 0; \qquad t = 1, \ldots T \tag{5}$$

$$\sum_{t=1}^{T} p_t \psi_t = 0; \tag{6}$$

$$\sum_{t=1}^{T} p_t - 1 = 0. \tag{7}$$

By multiplying the tth equation in (5) by p_t and summing over all t using (6) and (7), we get $\mu = T$ and from (5)

$$p_t = \frac{1}{T}\left(1 + \lambda \frac{\psi_t}{T}\right)^{-1}, \qquad t = 1, \ldots T \tag{8}$$

Substitute (8) into (6) to get

$$\sum \frac{\psi_t}{T}\left(1 + \lambda \frac{\psi_t}{T}\right)^{-1} = 0. \tag{9}$$

By an infinite expansion, we can verify that the left-hand side of (9) can be written as

$$\frac{1}{T}\sum \psi_t - \lambda \frac{1}{T^2}\sum \psi_t^2 + \lambda^2 \frac{1}{T^3}\sum \psi_t^3\left(1 + \lambda \frac{\psi_t}{T}\right)^{-1}. \tag{10}$$

This suggests defining

$$\lambda(\zeta) = \frac{\dfrac{1}{T}\sum \psi_t}{\dfrac{1}{T^2}\sum \psi_t^2} + \zeta; \tag{11}$$

for some ζ with $|\zeta| < K < \infty$. In view of (1) and (3), $\lambda(\zeta)$ in (11) is bounded in probability; and for any ε by Markov's inequality, we have $\left|\lambda(\zeta)\,\dfrac{\psi_t}{T}\right| < \varepsilon$, with probability arbitrarily close to 1 for a large enough T and all $t \le T$; then the values of $\left(1 + \lambda(\zeta)\,\dfrac{\psi_t}{T}\right)^{-1}$ are uniformly bounded in probability. Then the last term in (10) is $o_p(1)$ from (4), and (10) gives

$$\zeta \frac{1}{T^2}\sum \psi_t^2 + o_p(1).$$

This expression has a positive plim for $\zeta > 0$ and a negative one for $\zeta < 0$; thus, for large enough T, there exists a ζ, such that $\tilde{\lambda} = \lambda(\zeta)$ defined by (11) is a solution of (9) and $\zeta = o_p(1)$.

Next,

$$\mathrm{ELR} = -2\sum_{t=1}^{T} \ln Tp_t = 2\sum_{t=1}^{T} \ln\left(1 + \lambda\,\frac{\psi_t}{T}\right)$$

and, expanding each term, obtain

$$\mathrm{ELR} = 2\left[\frac{1}{T}\,\lambda\sum\psi_t - \frac{1}{2}\,\lambda^2\,\frac{1}{T^2}\sum\psi_t^2 + \frac{1}{3}\,\lambda^3\,\frac{1}{T^3}\sum\psi_t^3\left(1+\omega_t\,\lambda\,\frac{\psi_t}{T}\right)^{-3}\right], \tag{12}$$

with each ω_t between 0 and 1. For $\tilde{\lambda}$, the last sum in (12) similarly to that in (10) converges in probability to 0, and (12) provides

$$\mathrm{ELR} = \frac{\left(\dfrac{1}{T}\sum\psi_t\right)^2}{\dfrac{1}{T^2}\sum\psi_t^2} + o_p(T) \Rightarrow \frac{(W(1)^2-1)^2}{4\displaystyle\int W(r)^2\,dr},$$

where the limit is obtained by using (1) and (3).

Remark. There is a typographical error in the statement of the problem: the limit distribution has a 4, not a 2, in the denominator; and then it represents the square of the Dickey–Fuller distribution.

NOTE

1. An excellent solution has been proposed independently by J. Wright, the poser of the problem.

REFERENCES

Phillips, P.C.B. (1987) Time series regression with a unit root. *Econometrica* 55, 277–301.
White, H. (1984) *Asymptotic Theory for Econometricians.* New York: Academic Press.

99.2.2. *Asymptotic Normality of the Nonparametric Part in Partially Linear Heteroskedastic Models*—Solution, proposed by Hua Liang, Wolfgang Härdle, and Axel Werwatz. We first prove the following two lemmas. The first is a slight version of Theorem 9.1.1 by Chow and Teicher (1988); therefore, we omit its proof.

LEMMA 1. *Let* $\xi_{nk}, k = 1,\ldots,k_n$, *be independent random variables with* $E\xi_{nk} = 0$, *and* $E\xi_{nk}^2 = \sigma_{nk}^2 < \infty$. *Assume that* $\lim_{n\to\infty}\sum_{k=1}^{k_n}\sigma_{nk}^2 = 1$ *and* $\max_{1\le k\le k_n}\sigma_{nk}^2 \to 0$. *Then* $\sum_{k=1}^{k_n}\xi_{nk} \to^{\mathcal{L}} N(0,1)$ *in distribution if and only if*

$$\sum_{k=1}^{k_n} E\xi_{nk}^2 I(|\xi_{nk}| > \delta) \to 0 \quad \textit{for any } \delta > 0 \textit{ as } n \to \infty. \tag{1}$$

LEMMA 2. *Let $V_1, \ldots, V_n$ be independent random variables with $EV_i = 0$ and $\inf_i EV_i^2 > C > 0$ for some constant number C. The function $H(v)$ satisfying $\int_0^\infty vH(v)\,dv < \infty$ such that*

$$P\{|V_k| > v\} \leq H(v) \quad \textit{for large enough } v > 0 \quad \textit{and } k = 1, \ldots, n. \tag{2}$$

Also assume that $\{a_{ni}, i = 1, \ldots, n, n \geq 1\}$ is a sequence of real numbers satisfying $\sum_{i=1}^n a_{ni}^2 = 1$. If $\max_{1\leq i\leq n}|a_{ni}| \to 0$, then for $a'_{ni} = a_{ni}/\sigma_i(V)$,

$$\sum_{i=1}^n a'_{ni} V_i \to^{\mathcal{L}} N(0,1) \quad \textit{as } n \to \infty.$$

Proof. Denote $\xi_{nk} = a'_{nk} V_k$, $k = 1, \ldots, n$. We have $\sum_{k=1}^n E\xi_{nk}^2 = 1$. Moreover, it follows that

$$\begin{aligned}
\sum_{k=1}^n E\xi_{nk}^2 I(|\xi_{nk}| > \delta) &= \sum_{k=1}^n a_{nk}'^2 EV_k^2 I(|a_{nk}V_k| > \delta) \\
&\leq \sum_{k=1}^n \frac{a_{nk}^2}{\sigma_k^2} EV_k^2 I(|a_{nk}V_k| > \delta) \\
&\leq \left(\inf_k \sigma_k^2\right)^{-1} \sup_k E\{V_k^2 I(|a_{nk}V_k| > \delta)\}.
\end{aligned}$$

It follows from the condition (2) that

$$\sup_k E\{V_k^2 I(|a_{nk}V_k| > \delta)\} \to 0 \quad \text{for any } \delta > 0 \text{ as } n \to \infty.$$

Lemma 2 is therefore derived from Lemma 1.

Proof of Theorem. Under assumptions 1–3, a direct calculation derives that

$$\operatorname{Var}\{g_n(t)\} = \sum_{i=1}^n \omega_{ni}^2(t)\sigma_i^2 + o\left\{\sum_{i=1}^n \omega_{ni}^2(t)\sigma_i^2\right\}.$$

Furthermore,

$$g_n(t) - Eg_n(t) - \sum_{i=1}^n \omega_{ni}(t)\varepsilon_i = \sum_{i=1}^n \omega_{ni}(t) X_i^T (\widetilde{\mathbf{X}}^T\widetilde{\mathbf{X}})^{-1}\widetilde{\mathbf{X}}^T\tilde{\varepsilon} = O_P(n^{-1/2}),$$

which yields

$$\frac{\sum_{i=1}^n \omega_{ni}(t) X_i^T (\widetilde{\mathbf{X}}^T\widetilde{\mathbf{X}})^{-1}\widetilde{\mathbf{X}}^T\tilde{\varepsilon}}{\sqrt{\operatorname{Var}\{g_n(t)\}}} = O_P(n^{-1/2}n^{1/3}) = o_P(1).$$

It follows that

$$\frac{g_n(t) - Eg_n(t)}{\sqrt{\mathrm{Var}\{g_n(t)\}}} = \frac{\sum_{i=1}^{n} \omega_{ni}(t)\varepsilon_i}{\sqrt{\sum_{i=1}^{n} \omega_{ni}^2(t)\sigma_i^2}} + o_P(1) \overset{\text{def}}{=} \sum_{i=1}^{n} a_{ni}^* \varepsilon_i + o_P(1),$$

where $a_{ni}^* = (\omega_{ni}(t)) \Big/ \sqrt{\sum_{i=1}^{n} \omega_{ni}^2(t)\sigma_i^*}$. The proof of the theorem immediately follows from the conditions of the theorem and Lemma 2.

REFERENCE

Chow, Y.S. & H. Teicher. (1988) *Probability Theory*, 2nd ed. New York: Springer-Verlag.

99.2.3. *Asymptotic Efficiency of OLS Estimator in Models with Deterministic Regressors*—Solution, proposed by Jinyoung Hahn. The joint likelihood equals

$$\prod_{i=1}^{n} \frac{\sqrt{\omega}}{\sqrt{2\pi}} \exp\left[-\frac{\omega}{2}(y_i - x_i'\beta)^2\right],$$

where $\omega \equiv \sigma^{-2}$. Consider parameter values around (β_0, ω_0):

$$\beta = \beta(n) = \beta_0 + \frac{1}{\sqrt{n}}\delta, \quad \omega = \omega(n) = \omega_0 + \frac{1}{\sqrt{n}}\xi.$$

Writing the joint likelihood at $h = (\delta, \xi)$ as $P_{n,h}$, we obtain

$$\log \frac{dP_{n,h}}{dP_{n,h_0}} = \frac{n}{2}\log\frac{\omega}{\omega_0} - \frac{\omega}{2}\sum_{i=1}^{n}(y_i - x_i'\beta)^2 + \frac{\omega_0}{2}\sum_{i=1}^{n}(y_i - x_i'\beta_0)^2,$$

where $h_0 = (0,0)$. After some algebra, it can be shown that

$$\log \frac{dP_{n,h}}{dP_{n,h_0}} = h'V_n - \frac{1}{2}h'\Omega h + o_p(1),$$

where

$$\Omega = \begin{bmatrix} \omega_0 \cdot \Upsilon & 0 \\ 0 & \dfrac{1}{2\omega_0^2} \end{bmatrix},$$

and

$$V_n' \equiv \left(\omega_0 \cdot \frac{1}{\sqrt{n}}\sum_{i=1}^{n} x_i' \cdot \varepsilon_i, \ -\frac{1}{2\sqrt{n}}\sum_{i=1}^{n}(\varepsilon_i^2 - \sigma_0^2)\right)$$

https://doi.org/10.1017/S0266466600161109 Published online by Cambridge University Press

converges in distribution to $N(0,\Omega)$. This shows that the model is locally asymptotically normal. Now, notice that the sequence of parameters $\kappa_n(\delta,\xi) = \beta$ is regular by observing that

$$\sqrt{n}(\kappa_n(h) - \kappa_n(h_0)) = \sqrt{n}\cdot\left(\frac{\delta}{\sqrt{n}}\right) = \delta = \dot{\kappa}(h).$$

The convolution theorem can then be obtained by the application, for example, of van der Vaart and Wellner's (1996) Theorem 3.11.2.

REFERENCE

van der Vaart, A.W. & J.A. Wellner (1996) *Weak Convergence and Empirical Processes with Applications to Statistics*. New York: Springer-Verlag.

99.2.4. *Prediction in the Spatially Autocorrelated Error Component Model—*Solution,[1] proposed by Seuck Heun Song and Byoung Cheol Jung. We order the observations such that all the time periods observed in the first regional unit are stacked on top of those observed in the second regional unit, and so on. In vector form, equation (3) can be rewritten as

$$\phi = [(I_N - \lambda W)^{-1} \otimes I_T]\nu = (B^{-1} \otimes I_T)\nu, \tag{4}$$

where $B = I_N - \lambda W$, $\nu' = (\nu_{11},\dots,\nu_{1T},\dots,\nu_{NT})'$ and $\phi' = (\phi_{11},\dots,\phi_{1T},\dots,\phi_{NT})'$, I_N and I_T are identity matrix of dimension N and T, respectively, and $\otimes$ denotes the Kronecker product. Therefore, the overall disturbance vector of (2) becomes

$$\varepsilon = (I_N \otimes i_T)\mu + (B^{-1} \otimes I_T)\nu, \tag{5}$$

where $\mu = (\mu_1,\dots,\mu_N)'$ and i_T is a vector of ones of dimension T. In matrix form, the variance covariance matrix of ε is given by

$$\Omega = \sigma_\mu^2(I_N \otimes J_T) + \sigma_\nu^2[(B'B)^{-1} \otimes I_T], \tag{6}$$

where J_T is a matrix of ones of dimension T.

Goldberger (1962) showed that, for any Ω, the best linear unbiased predictor (BLUP) for $y_{i,T+S}$ is given by

$$\hat{y}_{i,T+S} = x'_{i,T+S}\hat{\beta}_{\text{GLS}} + w'\Omega^{-1}\hat{\varepsilon}_{\text{GLS}}, \tag{7}$$

where $w' = E(\varepsilon_{i,T+S}\varepsilon)'$, $\varepsilon_{i,T+S} = \mu_i + \phi_{i,T+S} = \mu_i + b_i'\nu_{T+S}$, where ν_{T+S} is the $N \times 1$ vector of $(T+S)$th time period and b_i' is the ith row of B^{-1}, and $\varepsilon = (I_N \otimes i_T)\mu + (B^{-1} \otimes I_T)\nu$ and $\hat{\varepsilon}_{\text{GLS}} = y - X\hat{\beta}_{\text{GLS}}$.

Using the spectral decomposition of Ω derived by Baltagi (1980, p. 1548), it follows that

$$\Omega^{-1} = (Z \otimes \bar{J}_T) + \frac{1}{\sigma_\nu^2}(B'B \otimes E_T), \tag{8}$$

where $\bar{J}_T = J_T/T$, $E_T = I_T - \bar{J}_T$ and $Z = [T\sigma_\mu^2 I_N + \sigma_\nu^2 (B'B)^{-1}]^{-1}$, and the last term of equation (7) is given by

$$\begin{aligned}
w'\Omega^{-1}\hat{\varepsilon}_{\text{GLS}} &= E(\varepsilon_{i,T+S}\,\varepsilon)'\Omega^{-1}\hat{\varepsilon}_{\text{GLS}} \\
&= E\{(\mu_i + b_i'\nu_{T+S})[(I_N \otimes i_T)\mu + (B^{-1} \otimes I_T)\nu]'\}\Omega^{-1}\hat{\varepsilon}_{\text{GLS}} \\
&= \sigma_\mu^2 (l_i \otimes i_T)'\Omega^{-1}\hat{\varepsilon}_{\text{GLS}} \\
&= \sigma_\mu^2 (l_i \otimes i_T)' \left(Z \otimes \bar{J}_T + \frac{1}{\sigma_\nu^2} B'B \otimes E_T \right) \hat{\varepsilon}_{\text{GLS}} \\
&= \sigma_\mu^2 (l_i' Z \otimes i_T')\hat{\varepsilon}_{\text{GLS}} \\
&= \sigma_\mu^2 \sum_{k=1}^{N} z_{ik} \sum_{t=1}^{T} \hat{\varepsilon}_{kt},
\end{aligned} \tag{9}$$

where l_i in the third equation is the ith column of I_N, and the third equality follows from the fact that μ_i is independent of ν and ν_{T+S} is independent of ν. Also, the fifth equality follows from the fact that $i_T'\bar{J}_T = i_T'$ and $i_T'E_T = 0$, and z_{ik} in the last equation is the (i,k)th elements of $Z = [T\sigma_\mu^2 I_N + \sigma_\nu^2 (B'B)^{-1}]^{-1}$. Therefore, the BLUP for $\hat{y}_{i,T+S}$ is given by

$$\hat{y}_{i,T+S} = x_{i,T+S}'\hat{\beta}_{\text{GLS}} + \sigma_\mu^2 \sum_{k=1}^{N} z_{ik} \sum_{t=1}^{T} \hat{\varepsilon}_{kt}. \tag{10}$$

NOTE

1. An excellent solution has been proposed independently by B. Baltagi and D. Li, the posers of the problem.

REFERENCES

Baltagi, B.H. (1980) On seemingly unrelated regression with error components. *Econometrica* 48, 1547–1551.

Goldberger, A.S. (1962) Best linear unbiased prediction in the generalized linear regression model. *Journal of the American Statistical Association* 57, 369–375.

ERRATUM

Neudecker, Heinz, and Michel van de Velden. Relationship between Two Eigenmatrices of a (Real) Symmetric Matrix

Michel van de Velden's name was omitted from the title of Problem 99.5.2.

https://doi.org/10.1017/S0266466600161109 Published online by Cambridge University Press