*Econometric Theory*, **19**, 2003, 691–705. Printed in the United States of America.
DOI: 10.1017/S0266466603004122

# PROBLEMS AND SOLUTIONS

## PROBLEMS

### 03.4.1. Normal's Deconvolution and the Independence of Sample Mean and Variance

Karim M. Abadir
*University of York, UK*

Jan R. Magnus
*Tilburg University, The Netherlands*

(a) Let $x_1$ and $x_2$ be independent variates having m.g.f.s $m_1(t_1)$ and $m_2(t_2)$, respectively, and define $y := x_1 + x_2$. Prove that $y$ is normal if and only if $x_1$ and $x_2$ are both normal. Is the existence of m.g.f.s necessary for this result?

(b) Let $\boldsymbol{x} := (x_1, \ldots, x_n)'$ be a vector of independent (but not necessarily identically distributed) components, where $2 \leq n < \infty$. Define $\bar{x} := (1/n)\sum_{i=1}^n x_i$ and $z := \sum_{i=1}^n (x_i - \bar{x})^2$. It is well known that if $\boldsymbol{x} \sim \mathrm{N}(\mu\boldsymbol{\iota}, \sigma^2 \boldsymbol{I}_n)$, then $\bar{x} \sim \mathrm{N}(\mu, \sigma^2/n)$ independently from $z/\sigma^2 \sim \chi^2(n-1)$. For $n \geq 3$, prove that if $\bar{x} \sim \mathrm{N}(\mu, \sigma^2/n)$ and $z/\sigma^2 \sim \chi^2(n-1)$, then $\boldsymbol{x} \sim \mathrm{N}(\mu\boldsymbol{\iota}, \sigma^2 \boldsymbol{I}_n)$.

(c) Why is the last statement in (b) not necessarily true for $n = 2$? What additional conditions are needed to make it hold for $n = 2$?

*ACKNOWLEDGMENT*

We thank Stéphane Gregoir, Paolo Paruolo, and two referees for their help and acknowledge support from the British ESRC grant R000239538.

### 03.4.2. The Asymptotic Distribution of the Dickey–Fuller Statistic under Nonnegativity Constraint

Giuseppe Cavaliere
*University of Bologna, Italy*

Consider the Dickey–Fuller statistic $Z_T := T\hat{\phi}$, where $\hat{\phi} := (\sum_{t=1}^T X_{t-1}^2)^{-1} \times \sum_{t=1}^T X_{t-1} \Delta X_t$. It is well known (see, e.g., Phillips, 1987) that if the data-generating process (d.g.p.) is the random walk $X_t = X_{t-1} + \varepsilon_t$, with initial condition $X_0 = 0$ and with $\{\varepsilon_t\}$ i.i.d. $N(0, \sigma^2)$, then

$$Z_T \xrightarrow{w} \left( 2 \int_0^1 B(s)^2 \, ds \right)^{-1} (B(1)^2 - 1), \tag{1}$$

where $B$ is a standard Brownian motion and $\xrightarrow{w}$ denotes weak convergence with respect to the uniform metric.

© 2003 Cambridge University Press 0266-4666/03 $12.00 

https://doi.org/10.1017/S0266466603004122 Published online by Cambridge University Press

Show that the result (1) holds even if the random walk $\{X_t\}$ is constrained to be nonnegative according to the following d.g.p. (the so-called random walk with one reflecting barrier; see Cox and Miller, 1965)

$$X_t = \begin{cases} X_{t-1} + \varepsilon_t & \text{if } X_{t-1} + \varepsilon_t \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

with initial condition $X_0 = 0$.

*REFERENCES*

Cox, D.R. & H.D. Miller (1965) *The Theory of Stochastic Processes*. London: Chapman and Hall.
Phillips, P.C.B. (1987) Time series regression with a unit root. *Econometrica* 55, 277–301.

## SOLUTIONS

### 02.3.1. Regression with an Evaporating Logarithmic Trend—Solution

Peter C.B. Phillips
*Cowles Foundation, Yale University,*
*University of Auckland, and University of York*

Yixiao Sun
*Yale University*

*Part A.* Let $z_t = 1/\ln t - (1/(n-1))\sum_{t=2}^{n} 1/\ln t$. Then

$$\hat{\beta} - \beta = \frac{\sum_{t=2}^{n} z_t u_t}{\sum_{t=2}^{n} z_t^2},$$

$$\hat{\alpha} - \alpha = \frac{\sum_{t=2}^{n} u_t}{n-1} - (\hat{\beta} - \beta)\frac{\sum_{t=2}^{n} \ln^{-1} t}{n-1}.$$

We start with $\hat{\beta}$ and find an asymptotic representation of the components $\sum_{t=2}^{n} 1/\ln t$ and $\sum_{t=2}^{n} 1/\ln^2 t$ that appear in $\sum_{t=2}^{n} z_t^2$. Using Euler summation (which justifies (1), which follows) and partial integration (which justifies (2)) we obtain an asymptotic series representation as follows:

https://doi.org/10.1017/S0266466603004122 Published online by Cambridge University Press

$$\sum_{t=2}^{n} \frac{1}{\ln t} = \int_2^n \frac{1}{\ln x}\, dx + O(1) \tag{1}$$

$$= n \sum_{k=1}^{m} \frac{(k-1)!}{\ln^k n} - 2 \sum_{k=1}^{m} \frac{(k-1)!}{\ln^k 2} + m! \int_2^n \frac{1}{\ln^{m+1} x}\, dx + O(1) \tag{2}$$

$$= n \sum_{k=1}^{m} \frac{(k-1)!}{\ln^k n} + m! \int_2^n \frac{1}{\ln^{m+1} x}\, dx + O(1). \tag{3}$$

To show that (3) is a valid asymptotic series, we can ignore the $O(1)$ term, and by a further application of partial integration to the second term we see that the remainder is proportional to

$$R_{m+1} = \int_2^n \frac{1}{\ln^{m+1} x}\, dx = \frac{n}{\ln^{m+1} n} - \frac{2}{\ln^{m+1} 2} + (m+1) \int_2^n \frac{1}{\ln^{m+2} x}\, dx. \tag{4}$$

Let $L = n^\alpha$ for some $\alpha \in (0,1)$ and write

$$\int_2^n \frac{1}{\ln^{m+2} x}\, dx = \int_2^L \frac{1}{\ln^{m+2} x}\, dx + \int_L^n \frac{1}{\ln^{m+2} x}\, dx = O(n^\alpha) + \int_L^n \frac{1}{\ln^{m+2} x}\, dx. \tag{5}$$

But

$$\int_L^n \frac{1}{\ln^{m+2} x}\, dx \le \frac{1}{\ln L} \int_L^n \frac{1}{\ln^{m+1} x}\, dx \le \frac{1}{\ln L} \int_2^n \frac{1}{\ln^{m+1} x}\, dx = \frac{1}{\ln L} R_{m+1}. \tag{6}$$

It follows from (4)–(6) that

$$R_{m+1} \le \frac{n}{\ln^{m+1} n} + (m+1) \int_2^n \frac{1}{\ln^{m+2} x}\, dx$$

$$\le \frac{n}{\ln^{m+1} n} + \int_2^L \frac{m+1}{\ln^{m+2} x}\, dx + \frac{m+1}{\ln L} R_{m+1}$$

$$= \frac{n}{\ln^{m+1} n} + \frac{m+1}{\ln L} R_{m+1} + O(n^\alpha)$$

so that

$$R_{m+1} = O\left( \frac{n}{\ln^{m+1} n} \right), \tag{7}$$

https://doi.org/10.1017/S0266466603004122 Published online by Cambridge University Press

and thus

$$\sum_{t=2}^{n} \frac{1}{\ln t} = n \sum_{k=1}^{m} \frac{(k-1)!}{\ln^k n} + O\left(\frac{n}{\ln^{m+1} n}\right), \tag{8}$$

showing that (3) is a valid asymptotic series.

In a similar way we may establish that

$$\sum_{t=2}^{n} \frac{1}{\ln^2 t} = \int_2^n \frac{1}{\ln^2 x}\, dx + O(1) \tag{9}$$

$$= n \sum_{k=2}^{m} \frac{(k-1)!}{\ln^k n} + O\left(\frac{n}{\ln^{m+1} n}\right) \tag{10}$$

is a valid asymptotic series. Combining (8) and (10), we get

$$\sum_{t=2}^{n} z_t^2 = \sum_{t=2}^{n} \frac{1}{\ln^2 t} - \frac{1}{n-1}\left(\sum_{t=2}^{n} \frac{1}{\ln t}\right)^2$$

$$= \int_2^n \frac{1}{\ln^2 x}\, dx - \frac{1}{n-1}\left(\int_2^n \frac{1}{\ln x}\, dx\right)^2 + O(1) \tag{11}$$

$$= \frac{n}{\ln^2 n} + \frac{2n}{\ln^3 n} + \frac{6n}{\ln^4 n} + O\left(\frac{n}{\ln^5 n}\right) \tag{12}$$

$$- \frac{1}{n}\left\{\frac{n}{\ln n} + \frac{n}{\ln^2 n} + \frac{2n}{\ln^3 n} + O\left(\frac{n}{\ln^4 n}\right)\right\}^2$$

$$= \frac{n}{\ln^4 n} + O\left(\frac{n}{\ln^5 n}\right) \to_{a.s.} \infty, \tag{13}$$

observing that we have to go to the third-order terms in expansion (12) to avoid degeneracy. We can obtain higher order asymptotic expansions by taking the preceding process to further terms, leading to the following explicit expression to order $n/\ln^{10} n$:

$$\sum_{t=2}^{n} z_t^2 = \frac{n}{\ln^4 n}\left(1 + \frac{8}{\ln n} + \frac{56}{\ln^2 n} + \frac{408}{\ln^3 n} + \frac{3{,}228}{\ln^4 n} + \frac{28{,}032}{\ln^5 n} + \frac{267{,}264}{\ln^6 n}\right)$$

$$+ O\left(\frac{n}{\ln^{11} n}\right). \tag{14}$$

It is immediately apparent from the size of the coefficients in (14) that very large values of $n$ are required before these approximations can be expected to work well. Of course, such approximations are hardly necessary because $\sum_{t=2}^{n} z_t^2$ is amenable to direct calculation or to direct approximation using the logarithmic integrals given in (1) and (9). The latter can be evaluated from the

https://doi.org/10.1017/S0266466603004122 Published online by Cambridge University Press

following well known series representation of the logarithmic integral (e.g., Gradshteyn and Ryzhik, 1967, 8.213.2, p. 926):

$$\mathrm{li}(y) = \int_0^y \frac{1}{\ln x}\,dx = \gamma + \ln\ln y + \sum_{k=1}^{\infty} \frac{(\ln y)^k}{kk!}, \quad \text{for } y > 1, \tag{15}$$

where $\gamma$ is Euler's constant. In the present case we have

$$\int_2^n \frac{1}{\ln x}\,dx = \mathrm{li}(n) - \mathrm{li}(2)$$

and

$$\int_2^n \frac{1}{\ln^2 x}\,dx = \frac{2}{\ln 2} - \frac{n}{\ln n} + \int_2^n \frac{1}{\ln x}\,dx = \frac{2}{\ln 2} - \frac{n}{\ln n} + [\mathrm{li}(n) - \mathrm{li}(2)].$$

Observe that $\hat{\beta} - \beta$ is the same as the error in the ordinary least squares (OLS) estimator of $\beta$ in the regression

$$X_t = z_t\beta + u_t,$$

where $\{u_t\}$ is a martingale difference sequence with respect to the natural filtration $\mathcal{F}_t$. The persistent excitation condition (13) holds in this regression, and so $\hat{\beta} \to_{a.s.} \beta$ and $\hat{\beta}$ is strongly consistent.

Now observe that

$$\hat{\alpha} - \alpha = \frac{\sum_{t=2}^{n} u_t}{n-1} - (\hat{\beta} - \beta)\frac{\sum_{t=2}^{n} \ln^{-1} t}{n-1} = o_{a.s.}(1) + o_{a.s.}(1)O_{a.s.}(\ln^{-1} n)$$

$$= o_{a.s.}(1)$$

and so $\hat{\alpha} \to_{a.s.} \alpha$.

Next, turn to the asymptotic distribution of $\hat{\alpha}$ and $\hat{\beta}$. We have

$$\hat{\beta} - \beta = \frac{\sum_{t=2}^{n} z_t u_t}{\sum_{t=2}^{n} z_t^2} = \frac{\sum_{t=2}^{n} z_t u_t}{s_n}\left(\frac{\sqrt{n}}{\ln^2 n}\right)^{-1}(1 + o(1)),$$

$$s_n^2 = \sum_{t=2}^{n} z_t^2.$$

https://doi.org/10.1017/S0266466603004122 Published online by Cambridge University Press

Let $y_{nt} = z_t u_t / s_n$; then $y_{n2}, y_{n3}, \ldots, y_{nn}$ are independent random variables with zero means and variances that sum to $\sigma^2$. To apply the Liapounoff central limit theorem for $\sum_{t=2}^{n} y_{nt}$, we need to show that

$$\sum_{t=2}^{n} E|y_{nt}|^3 \to 0, \quad \text{as } n \to \infty.$$

But

$$\sum_{t=2}^{n} E|y_{nt}|^3 = \frac{\sum_{t=2}^{n} |z_t|^3 E|u_t|^3}{|s_n|^3} \leq \text{Const.} n \left( \frac{\ln^4 n}{n} \right)^{3/2} = o(1),$$

and so

$$\sum_{t=2}^{n} y_{nt} = \frac{\sum_{t=2}^{n} z_t u_t}{s_n} \Rightarrow N(0, \sigma^2).$$

Hence

$$\frac{\sqrt{n}}{\ln^2 n} (\hat{\beta} - \beta) = \frac{\sum_{t=2}^{n} z_t u_t}{s_n} (1 + o_p(1)) \Rightarrow N(0, \sigma^2). \tag{16}$$

Note that

$$\begin{aligned} \hat{\alpha} - \alpha &= \frac{\sum_{t=2}^{n} u_t}{n-1} - (\hat{\beta} - \beta) \frac{\sum_{t=2}^{n} \ln^{-1} t}{n-1} \\ &= O_p(n^{-1/2}) - O_p(n^{-1/2} \ln^2 n \ln^{-1} n) \\ &= O_p(n^{-1/2} \ln n). \end{aligned}$$

Thus, the term $(\hat{\beta} - \beta)(n-1)^{-1} \sum_{t=2}^{n} \ln^{-1} t$ dominates the asymptotics of $\hat{\alpha}$, and we deduce that

$$\sqrt{n} \ln^{-1} n (\hat{\alpha} - \alpha) = -\frac{\sqrt{n}}{\ln^2 n} (\hat{\beta} - \beta) + o_p(1) =_d N(0, \sigma^2). \tag{17}$$

In view of (16) and (17), we have the following joint asymptotics:

$$\begin{pmatrix} n^{1/2} \ln^{-1} n (\hat{\alpha} - \alpha) \\ n^{1/2} \ln^{-2} n (\hat{\beta} - \beta) \end{pmatrix} \Rightarrow N(0, \Sigma) \quad \text{with } \Sigma = \sigma^2 \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}. \tag{18}$$

https://doi.org/10.1017/S0266466603004122 Published online by Cambridge University Press

Remarks

1. The limit distribution (18) is singular, and the components $n^{1/2}\ln^{-1} n(\hat{\alpha}-\alpha)$ and $n^{1/2}\ln^{-2} n(\hat{\beta}-\beta)$ are perfectly negatively correlated as $n\to\infty$. Observe that the rate of convergence of $\hat{\alpha}$ exceeds that of $\hat{\beta}$, by virtue of the fact that the signal from the intercept is stronger than the signal from the evaporating logarithmic regressor $\ln^{-1} t$.
2. Result (18) gives first-order asymptotics. As is apparent from (14), asymptotic series expansions in the present model involve factors of $1/\ln n$ in contrast to the more usual $1/\sqrt{n}$, and, correspondingly, they deliver only very slow improvements in the first-order asymptotics. The asymptotic variance $(\sigma^2\ln^4 n)/n$ and higher order approximations based on (14) are therefore poor approximations to the variance of $\hat{\beta}$ even for quite large $n$. To illustrate, for values of the sample size $n\in[10^2,10^4]$ and for $\sigma^2=1$, Figure 1 provides graphs of the exact variance of $\hat{\beta}$, the asymptotic variance, the two-term and three-term series approximations to the variance based on the first few terms of (14), and direct calculation of the logarithmic integral representations of the variance obtained from (11) and (15). Apparently, only the latter are adequate for sample sizes $n$ in this range.
3. The theory developed here is part of a general theory of regression on slowly varying regressors, a subject that has recently been studied in Phillips (2000).

*Part B.* Using the Phillips–Solo (1992) device, we have

$$u_t = C(L)\varepsilon_t = C(1)\varepsilon_t + \tilde{\varepsilon}_{t-1} - \tilde{\varepsilon}_t,$$

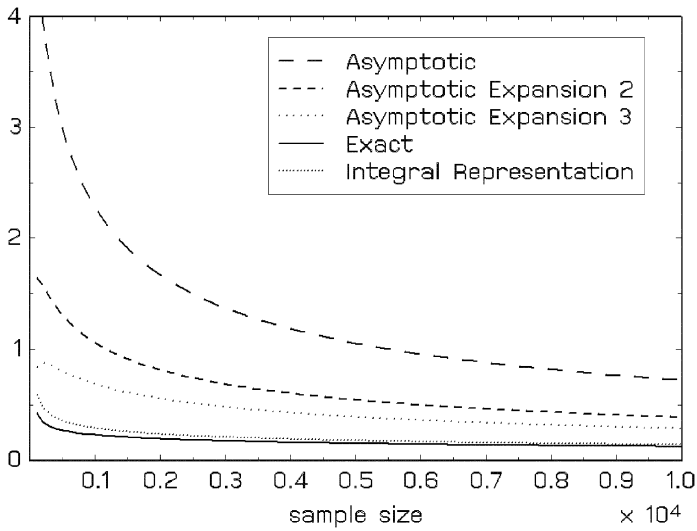



**FIGURE 1.** $\sigma^2/\sum_{t=1}^n z_t^2$ and asymptotic and integral approximations.

https://doi.org/10.1017/S0266466603004122 Published online by Cambridge University Press

where $\tilde{\varepsilon}_t = \tilde{C}(L)\varepsilon_t$, $\tilde{C}(L) = \sum_{u=0}^{n} \tilde{C}_u L^u$, and $\tilde{C}_u = \sum_{s=u+1}^{n} c_s$. Then, we can write

$$\sum_{t=2}^{n} z_t u_t = \sum_{t=2}^{n} z_t C(1)\varepsilon_t + \sum_{t=2}^{n} z_t \tilde{\varepsilon}_{t-1} - \sum_{t=2}^{n} z_t \tilde{\varepsilon}_t$$

$$= C(1)\sum_{t=2}^{n} z_t \varepsilon_t + \sum_{t=2}^{n-1} (z_{t+1} - z_t)\tilde{\varepsilon}_t + z_2\tilde{\varepsilon}_1 - z_n\tilde{\varepsilon}_n.$$

It follows that

$$\hat{\beta} - \beta = \frac{C(1)\sum_{t=2}^{n} z_t \varepsilon_t}{\sum_{t=2}^{n} z_t^2} + \frac{\sum_{t=2}^{n-1} (z_{t+1} - z_t)\tilde{\varepsilon}_t}{\sum_{t=2}^{n} z_t^2} + \frac{z_2\tilde{\varepsilon}_1}{\sum_{t=2}^{n} z_t^2} - \frac{z_n\tilde{\varepsilon}_n}{\sum_{t=2}^{n} z_t^2}.$$

Note that the first term satisfies the limit theory given in the earlier part. If we prove that the last three terms converge strongly to zero, then we are done with the strong consistency of $\hat{\beta}$. To obtain the asymptotic distribution, we need to consider $(\sum_{t=2}^{n} z_t^2)^{1/2}(\hat{\beta} - \beta)$. We prove that the last three terms, so normalized, are $o_p(1)$, and then the asymptotic distribution is determined by the first term. The following three results are given first.

(a) $$\frac{z_2\tilde{\varepsilon}_1}{\sum_{t=2}^{n} z_t^2} = o_{a.s}(1) \quad \text{and} \quad \frac{z_2\tilde{\varepsilon}_1}{\left(\sum_{t=2}^{n} z_t^2\right)^{1/2}} = o_p(1).$$

**Proof.** These are immediate because $\sum_{t=2}^{n} z_t^2 \to \infty$ and $(\sum_{t=2}^{n} z_t^2)^{1/2} \to \infty$. ■

(b) $$\frac{z_n\tilde{\varepsilon}_n}{\sum_{t=2}^{n} z_t^2} = o_{a.s}(1) \quad \text{and} \quad \frac{z_n\tilde{\varepsilon}_n}{\left(\sum_{t=2}^{n} z_t^2\right)^{1/2}} = o_p(1).$$

**Proof.** Note that $z_n$ is bounded and for any $\delta > 0$

$$P\left(\left|\frac{z_n\tilde{\varepsilon}_n}{\sum_{t=2}^{n} z_t^2}\right| > \delta\right) \le \frac{z_n^2 E(\tilde{\varepsilon}_n^2)}{\left(\sum_{t=2}^{n} z_t^2\right)^2 \delta^2} = O\left(\left(\sum_{t=2}^{n} z_t^2\right)^{-2}\right),$$

https://doi.org/10.1017/S0266466603004122 Published online by Cambridge University Press

so $\sum_n P(|z_n\tilde{\varepsilon}_n/\sum_{t=2}^n z_t^2| > \delta) < \infty$ and the first result follows. The second follows because $P(|z_n\tilde{\varepsilon}_n/(\sum_{t=2}^n z_t^2)^{1/2}| > \delta) = O((\sum_{t=2}^n z_t^2)^{-1}) \to 0$. ■

(c) $$\frac{\sum_{t=2}^{n-1}(z_{t+1}-z_t)\tilde{\varepsilon}_t}{\sum_{t=2}^{n} z_t^2} = o_{a.s.}(1) \quad \text{and} \quad \frac{\sum_{t=2}^{n-1}(z_{t+1}-z_t)\tilde{\varepsilon}_t}{\left(\sum_{t=2}^{n} z_t^2\right)^{1/2}} = o_p(1).$$

**Proof.** Note that $\{\tilde{\varepsilon}_t\}$ has finite fourth moment because

$$\|\tilde{\varepsilon}_t\|_4 := (E(\tilde{\varepsilon}_t^4))^{1/4} = \left\|\sum_{j=0}^{t}\tilde{C}_{t-j}\varepsilon_j\right\|_4 \le \sum_{j=0}^{t}|\tilde{C}_{t-j}|\cdot\|\varepsilon_j\|_4 < \infty.$$

Here we have employed the fact that $\sum_{j=0}^t|\tilde{C}_{t-j}| < \infty$, which follows from the assumption that $\sum_{j=0}^\infty j|c_j| < \infty$.

Next,

$$E\left(\sum_{t=2}^{n}\tilde{\varepsilon}_t^2\right)^2 = E\left(\sum_{t=2}^{n}\tilde{\varepsilon}_t^4 + 2\sum_{\substack{s,\,t=2\\ s<t}}^{n}\tilde{\varepsilon}_s^2\tilde{\varepsilon}_t^2\right)$$

$$\le O(n) + 2\sum_{\substack{s,\,t=2\\ s<t}}^{n}[E(\tilde{\varepsilon}_s^4)]^{1/2}[E(\tilde{\varepsilon}_t^4)]^{1/2} = O(n^2).$$

Therefore, for any $\delta > 0$,

$$P\left(\left|\frac{\sum_{t=2}^{n-1}(z_{t+1}-z_t)\tilde{\varepsilon}_t}{\sum_{t=2}^{n} z_t^2}\right| > \delta\right)$$

$$\le \frac{E\left(\sum_{t=2}^{n-1}(z_{t+1}-z_t)\tilde{\varepsilon}_t\right)^4}{\left(\sum_{t=2}^{n} z_t^2\right)^4\delta^4} \le \frac{E\left[\sum_{t=2}^{n-1}(z_{t+1}-z_t)^2\left(\sum_{t=2}^{n-1}\tilde{\varepsilon}_t^2\right)\right]^2}{\left(\sum_{t=2}^{n} z_t^2\right)^4\delta^4}$$

$$= \frac{E\left(\sum_{t=2}^{n-1}(z_{t+1}-z_t)^2\right)^2\left(\sum_{t=2}^{n-1}\tilde{\varepsilon}_t^2\right)^2}{\left(\sum_{t=2}^{n} z_t^2\right)^4\delta^4} = \frac{\left(\sum_{t=2}^{n-1}(z_{t+1}-z_t)^2\right)^2 O(n^2)}{\left(\sum_{t=2}^{n} z_t^2\right)^4\delta^4}.$$

https://doi.org/10.1017/S0266466603004122 Published online by Cambridge University Press

Because

$$\sum_{t=2}^{n-1}(z_{t+1}-z_t)^2=\sum_{t=2}^{n-1}(\ln^{-1}(t+1)-\ln^{-1}t)^2=\sum_{t=2}^{n-1}\ln^2\left(1+\frac{1}{t}\right)[\ln^{-2}t][\ln^{-2}(t+1)]$$

$$<\text{Const.}\sum_{t=2}^{n}t^{-2}<\infty,$$

we have

$$P\left(\left|\frac{\sum_{t=2}^{n-1}(z_{t+1}-z_t)\tilde{\varepsilon}_t}{\sum_{t=2}^{n}z_t^2}\right|>\delta\right)=O(n^{-2}\ln^8 n),$$

and so $\sum_{t=2}^{n-1}(z_{t+1}-z_t)\tilde{\varepsilon}_t/\sum_{t=2}^{n}z_t^2=o_{a.s.}(1)$. Similarly,

$$P\left(\left|\frac{\sum_{t=2}^{n-1}(z_{t+1}-z_t)\tilde{\varepsilon}_t}{\left(\sum_{t=2}^{n}z_t^2\right)^{1/2}}\right|>\delta\right)$$

$$\leq\frac{E\left[\sum_{t=2}^{n-1}(z_{t+1}-z_t)\tilde{\varepsilon}_t\right]^2}{\sum_{t=2}^{n}z_t^2\delta^2}$$

$$\leq\frac{E\sum_{t=2}^{n-1}(z_{t+1}-z_t)^2\tilde{\varepsilon}_t^2+2\sum_{s>t}(|z_{t+1}-z_t||z_{s+1}-z_s|E|\tilde{\varepsilon}_t\tilde{\varepsilon}_s|)}{\sum_{t=2}^{n}z_t^2\delta^2}$$

$$\leq\frac{E\sum_{t=2}^{n-1}(z_{t+1}-z_t)^2\tilde{\varepsilon}_t^2+2\sum_{s>t}(|z_{t+1}-z_t||z_{s+1}-z_s|(E|\tilde{\varepsilon}_t|^2)^{1/2}(E|\tilde{\varepsilon}_s|^2)^{1/2})}{\sum_{t=2}^{n}z_t^2\delta^2}$$

$$=\text{Const.}\frac{\left(\sum_{t=2}^{n-1}|z_{t+1}-z_t|\right)^2}{\sum_{t=2}^{n}z_t^2\delta^2}=\text{Const.}\frac{\ln^2 n}{\sum_{t=2}^{n}z_t^2\delta^2}\to 0.$$

The last equality follows as $\sum_{t=2}^{n-1}|z_{t+1}-z_t|=O(\sum_{t=2}^{n}1/t)=O(\ln n)$. Hence $(\sum_{t=2}^{n}z_t^2)^{-1/2}(\sum_{t=2}^{n-1}(z_{t+1}-z_t)\tilde{\varepsilon}_t)=o_p(1)$. ■

https://doi.org/10.1017/S0266466603004122 Published online by Cambridge University Press

Combining results (a)–(c), we have

$$\hat{\beta} - \beta = o_{a.s.}(1), \tag{19}$$

and then

$$\left(\sum_{t=2}^{n} z_t^2\right)^{1/2} (\hat{\beta} - \beta) = \left(\sum_{t=2}^{n} z_t^2\right)^{1/2} \frac{C(1) \sum_{t=2}^{n} z_t \varepsilon_t}{\sum_{t=2}^{n} z_t^2} + o_p(1) =_d N(0, C(1)^2 \sigma^2) + o_p(1). \tag{20}$$

With (19), (20), and $\sum_{t=2}^{n} u_t/(n-1) \to_{a.s.} 0$ (e.g., Phillips and Solo, 1992), the previous arguments can now be repeated to obtain the strong consistency and the asymptotic distribution of $\hat{\alpha}$. Specifically

$$\begin{pmatrix} n^{1/2} \ln^{-1} n(\hat{\alpha} - \alpha) \\ n^{1/2} \ln^{-2} n(\hat{\beta} - \beta) \end{pmatrix} \Rightarrow N(0, \Sigma) \quad \text{with } \Sigma = (C(1))^2 \sigma^2 \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

*REFERENCES*

Gradshteyn I.S. & I.M. Ryzhik (1965) *Tables of Integrals, Series, and Products*. New York: Academic Press.

Phillips, P.C.B. (2000) Regression with Slowly Varying Regressors. Cowles Foundation Discussion paper 1310, Yale University.

Phillips, P.C.B. & V. Solo (1992) Asymptotics for linear processes. *Annals of Statistics* 20, 971–1001.

## 02.3.2. Badly Weighted Least Squares—Solution[1]

Douglas P. Wiens
*University of Alberta, Edmonton, Alberta*

Let $\mathbf{X}_{n\times p}$ have full column rank $p < n$ and let $\mathbf{\Omega}$ be a diagonal matrix whose diagonal elements $\{\omega_i\}_{i=1}^{n}$ are positive and not all equal. The discussion by Professors Koenker and Portnoy suggested that, for any (nonzero) $\boldsymbol{\delta} \in \mathbb{R}^p$, the function

$$e_{a,\boldsymbol{\delta}} = \boldsymbol{\delta}'(\mathbf{X}'\mathbf{\Omega}^{-2a}\mathbf{X})^{-1}\mathbf{X}'\mathbf{\Omega}^{2-4a}\mathbf{X}(\mathbf{X}'\mathbf{\Omega}^{-2a}\mathbf{X})^{-1}\boldsymbol{\delta}$$

satisfies $e_{0,\boldsymbol{\delta}} \geq e_{0.5,\boldsymbol{\delta}}$ and more generally that $e_{a,\boldsymbol{\delta}}$ is a decreasing function of $a \in [0,1]$.

We shall show that the second suggestion, and hence the first also, is true for $p = 1$. Both can fail if $p > 1$; we illustrate this by an example in which $e_{a,\boldsymbol{\delta}}$ is not only nonmonotone but is *maximized* at $a = 0.5$.

https://doi.org/10.1017/S0266466603004122 Published online by Cambridge University Press

First let $p = 1$, so that $\mathbf{X} = \mathbf{x}$, a column vector. We calculate that

$$\frac{d}{da} e_{a,\delta} = \frac{-2\delta^2 M}{\mathbf{x}'\boldsymbol{\Omega}^{-2a}\mathbf{x}}, \tag{1}$$

where

$$M = 2\mathbf{v}'\boldsymbol{\Omega}^{2-2a}\mathbf{L}_{\boldsymbol{\Omega}}\mathbf{v} - 2(\mathbf{v}'\boldsymbol{\Omega}^{2-2a}\mathbf{v})(\mathbf{v}'\mathbf{L}_{\boldsymbol{\Omega}}\mathbf{v}),$$

$\mathbf{L}_{\boldsymbol{\Omega}}$ is the diagonal matrix with diagonal elements $\{\log \omega_i\}_{i=1}^n$, and $\mathbf{v} = \boldsymbol{\Omega}^{-a}\mathbf{x}/\|\boldsymbol{\Omega}^{-a}\mathbf{x}\|$ has unit norm.

Because $M$, and hence the derivative (1), vanishes when $a = 1$, our assertion is equivalent to the statement that $M > 0$ for $a \in [0,1)$. Define $f(\omega) = \omega^{2-2a}$ and set $f_i = f(\omega_i)$. Then

$$\begin{aligned} 2(1-a)M &= 2\sum_{i=1}^n (f_i \log f_i)v_i^2 - 2\sum_{i,j=1}^n (f_i \log f_j)v_i^2 v_j^2 \\ &= \sum_{i,j=1}^n (f_i - f_j)(\log f_i - \log f_j)v_i^2 v_j^2. \end{aligned} \tag{2}$$

Note that

$$(f_i - f_j)(\log f_i - \log f_j) > 0 \quad \text{whenever } \omega_i \neq \omega_j. \tag{3}$$

Furthermore, if any element $x_i$ of $\mathbf{x}$ is 0, it can be removed (and the value of $n$ reduced) without affecting the value of $\mathbf{x}'\boldsymbol{D}\boldsymbol{x} = \sum_{i=1}^n d_i x_i^2$ for a diagonal $\mathbf{D}$ and hence without affecting the value of $e_{a,\delta}$. Thus we can assume that all $x_i \neq 0$ and hence that all $v_i \neq 0$. This together with (3) ensures that (2) is positive (because $\omega_i \neq \omega_j$ at least once), thus establishing the monotonicity claim when $p = 1$.

*Counterexample when $p > 1$.* For general $p$ the preceding method results in

$$\frac{d}{da} e_{a,\boldsymbol{\delta}} = -2\boldsymbol{\delta}'(\mathbf{X}'\boldsymbol{\Omega}^{-2a}\mathbf{X})^{-1/2}\mathbf{M}(\mathbf{X}'\boldsymbol{\Omega}^{-2a}\mathbf{X})^{-1/2}\boldsymbol{\delta},$$

where

$$\mathbf{M} = 2\mathbf{V}'\boldsymbol{\Omega}^{2-2a}\mathbf{L}_{\boldsymbol{\Omega}}\mathbf{V} - (\mathbf{V}'\boldsymbol{\Omega}^{2-2a}\mathbf{V})(\mathbf{V}'\mathbf{L}_{\boldsymbol{\Omega}}\mathbf{V}) - (\mathbf{V}'\mathbf{L}_{\boldsymbol{\Omega}}\mathbf{V})(\mathbf{V}'\boldsymbol{\Omega}^{2-2a}\mathbf{V})$$

and $\mathbf{V} = \boldsymbol{\Omega}^{-a}\mathbf{X}(\mathbf{X}'\boldsymbol{\Omega}^{-2a}\mathbf{X})^{-1/2}$ satisfies $\mathbf{V}'\mathbf{V} = \sum_{i=1}^n \mathbf{v}_i\mathbf{v}_i' = \mathbf{I}_p$. The analogue of (2) is that

$$2(1-a)\mathbf{M} = \sum_{i,j=1}^n (f_i - f_j)(\log f_i - \log f_j)\mathbf{v}_i\mathbf{v}_i'\mathbf{v}_j\mathbf{v}_j',$$

and the monotonicity claim is equivalent to the positive definiteness of $\mathbf{M}$. However, for $p > 1$, $\mathbf{M}$ often has at least one negative eigenvalue. In such cases the monotonicity properties will depend on the choice of $\boldsymbol{\delta}$. A particular counter-

https://doi.org/10.1017/S0266466603004122 Published online by Cambridge University Press

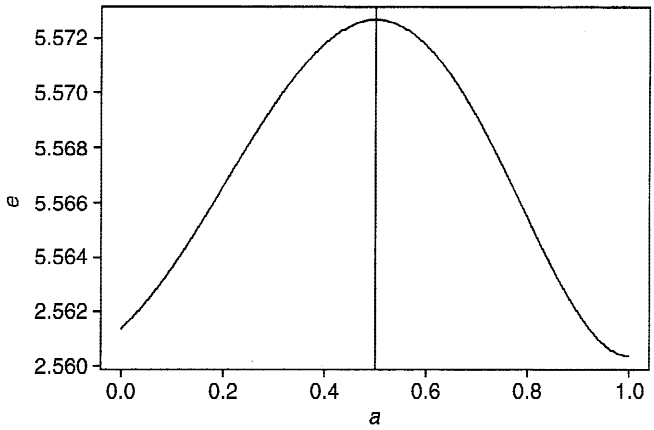



**FIGURE 1.** Plot of $e_{a,\boldsymbol{\delta}}$ against $a$ with $\boldsymbol{\delta}$ chosen so that the maximum is attained at $a = 0.5$.

example to both suggestions of Professors Koenker and Portnoy has $n = 3$, $p = 2$, and

$$\boldsymbol{\Omega} = \begin{pmatrix} 5.4 & 0 & 0 \\ 0 & 2.9 & 0 \\ 0 & 0 & 1.8 \end{pmatrix}, \qquad \mathbf{X} = \begin{pmatrix} 0.2 & -1 \\ 1.2 & 0.4 \\ 0.3 & 0.5 \end{pmatrix}.$$

For these choices the eigenvalues of $\mathbf{M}$ are 1.49 and $-0.003$ when $a = 0.5$. We choose $\boldsymbol{\delta} = (1,0.26291)'$; then $e_{a,\boldsymbol{\delta}}$ is nonmonotone and is *maximized* at $a = 0.5$. See Figure 1.

This counterexample can be extended to all $n > 3$ by adding $n - 3$ rows of zeros to the bottom of $\mathbf{X}$ and continuing the diagonal of $\boldsymbol{\Omega}$ arbitrarily.

*NOTE*

1. An excellent solution was independently proposed by Geert Dhaene. A proof for the case $p = 1$ based on the convexity of Laplace transforms was also provided by Koenker and Portnoy, the posers of the problem.

## 02.4.1. On Hadamard Product of Square Roots of Correlation Matrices—Solution

Shuangzhe Liu (poser of the problem)
*Australian National University, Canberra, Australia*

We have the following matrix Kantorovich inequality (see, e.g., Liu and Neudecker, 1997)

https://doi.org/10.1017/S0266466603004122 Published online by Cambridge University Press

$$Y'NY(Y'M^{-1}Y)^{-1}Y'NY \geq \frac{4pq}{(p+q)^2}\, Y'NMNY, \tag{1}$$

where $N > 0$ and $M > 0$ are of the same order and $p > 0$ and $q > 0$ are the largest and smallest eigenvalues of $NM$, respectively.

We use $N = A^{1/2} \otimes B^{1/2}$, $M = I \otimes B^{-1}$, and $Y = J$, where $\otimes$ is the Kronecker product and $J$ is the $n^2 \times n$ selection matrix (see, e.g., Rao and Rao, 1998, Sect. 6.3; Zhang, 1999, Sect. 6.5). Clearly, $Y'NY = J'(A^{1/2} \otimes B^{1/2})J = A^{1/2} \circ B^{1/2}$, $Y'M^{-1}Y = J'(I \otimes B)J = I \circ B = I$, and $Y'NMNY = J'(A \otimes I)J = A \circ I = I$ for $n \times n$ matrices $A$, $B$, and $I$. Inserting these terms into (1), we get

$$(A^{1/2} \circ B^{1/2})^2 \geq \frac{4pq}{(p+q)^2}\, I. \tag{2}$$

Note that $NM = A^{1/2} \otimes B^{-1/2}$ and the eigenvalues of $A^{1/2}$ and $B^{1/2}$ are contained in $[d, c]$. Therefore $p \leq c/d$ and $q \geq d/c$. We can then replace $p$ and $q$ in (2) by $c/d$ and $d/c$, respectively, and establish

$$A^{1/2} \circ B^{1/2} \geq \frac{2cd}{c^2 + d^2}\, I.$$

*REFERENCES*

Liu, S. & H. Neudecker (1997) Kantorovich inequalities and efficiency comparisons for several classes of estimators in linear models. *Statistica Neerlandica* 51, 345–355.

Rao, C.R. & M.B. Rao (1998) *Matrix Algebra and Its Applications to Statistics and Econometrics*. Singapore: World Scientific.

Zhang, F.-Z. (1999) *Matrix Theory: Basic Results and Techniques*. New York: Springer-Verlag.

## 02.4.2. On the Rank of a Matrix Useful in Goodness-of-Fit Testing of Structural Equation Models — Solution[1]

Simo Puntanen
*University of Tampere, Tampere, Finland*

George P.H. Styan
*McGill University, Montréal, Québec, Canada*

Hans Joachim Werner
*University of Bonn, Bonn, Germany*

We present two proofs. From Marsaglia and Styan (1974, Theorem 5, p. 274) we have

$$\operatorname{rank}(A : V) = \operatorname{rank}(A) + \operatorname{rank}\big((I - AA^-)V\big) \tag{1}$$

for any generalized inverse $A^-$. Because $M$ is symmetric idempotent and $V$ is nonnegative definite we have $\operatorname{rank}(MVM) = \operatorname{rank}(MV)$. We are, therefore, asked to show that $\operatorname{rank}(MV) + \operatorname{rank}(A) = \operatorname{rank}(V)$ iff $\operatorname{rank}(A : V) = \operatorname{rank}(V)$. This follows at once from (1) replacing $A^-$ with $A^+$. When $W$ is nonnegative

https://doi.org/10.1017/S0266466603004122 Published online by Cambridge University Press

definite $(A'WA)^{+}A'W$ is a generalized inverse of $A$, and so $\text{rank}(MVM) + \text{rank}(A) = \text{rank}(V) \Leftrightarrow \text{col}(A) \subseteq \text{col}(V)$ with $M := I_n - A(A'WA)^{+}A'W$. This completes our first proof. We see that $W$ need only be nonnegative definite, not necessarily positive definite.

For our second proof we use (2.6) in Werner and Yapar (1996) (see also Werner, 1985, Lemma 3.2):

$$\text{col}(A,V) = \text{col}(A) \oplus \text{col}(VQ), \tag{2}$$

where $\oplus$ indicates direct sum and where $Q$ is any matrix such that $\text{col}(Q) = \text{null}(A')$, with $\text{null}(\cdot)$ denoting the null space of $(\cdot)$. In view of (2), clearly

$$\text{col}(A) \oplus \text{col}(VQ) = \text{col}(V) \tag{3}$$

iff

$$\text{col}(A) \subseteq \text{col}(V). \tag{4}$$

Because $\text{col}(M) = \text{null}(A')$, $M = M'$, $V \geq 0$, $MVM \geq 0$, we have $\text{rank}(MVM) = \text{rank}(VM) = \text{rank}(VQ)$. As (3) is equivalent to $\text{rank}(A) + \text{rank}(VQ) = \text{rank}(V)$, the first assertion is established. To see that an equivalent formulation of the problem is obtained when $M := I_n - A(A'WA)^{+}A'W$, with $W$ nonnegative definite and $MVM$ changed to $MVM'$, we note that $MVM' \geq 0$, $\text{rank}(MVM') = \text{rank}(VM')$, and $\text{col}(M') = \text{null}(A')$ (see, e.g., Werner and Yapar, 1996, Theorem 2.5 and (2.2) and (2.3)). Therefore, $\text{col}(VM') = \text{col}(VQ)$, and the assertion again is an immediate consequence of (3) $\Leftrightarrow$ (4).

It is easy to see that the preceding equivalent formulation can be generalized further to any square (not necessarily nonnegative definite) matrix $W$ such that $\text{rank}(A'WA) = \text{rank}(A)$ for then $(A'WA)^{-}A'W$ is a generalized inverse of $A$ and $\text{col}(M') = \text{null}(A')$.

## NOTE

1. An excellent solution was independently proposed by A. Satorra and H. Neudecker, the posers of the problem.

## REFERENCES

Marsaglia, G. & G.P.H. Styan (1974) Equalities and inequalities for rank of matrices. *Linear and Multilinear Algebra* 2, 269–292.

Werner, H.J. (1985) More on the BLUE estimation in regression models with possibly singular covariances. *Linear Algebra and Its Applications* 67, 207–214.

Werner, H.J. & C. Yapar (1996) On inequality constrained generalized least squares selections in the general possibly singular Gauss-Markov model: A projector theoretical approach. *Linear Algebra and Its Applications* 237/238, 359–393.

https://doi.org/10.1017/S0266466603004122 Published online by Cambridge University Press