Econometric Theory, **18**, 2002, 1007–1017. Printed in the United States of America.
DOI: 10.1017.S026646660218412X

PROBLEMS AND SOLUTIONS

PROBLEMS

02.4.1 *On the Hadamard Product of Square Roots of Correlation Matrices*, proposed by Shuangzhe Liu. Show that for $n \times n$ positive definite correlation matrices A and B (with diagonal entries one)

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

where $\circ$ stands for the Hadamard product, $\geq$ stands for the Loewner ordering, and c and d are positive numbers such that the eigenvalues of $A^{1/2}$ and $B^{1/2}$ are contained in the interval $[d, c]$.

It is a counterpart of the following inequality established by Zhang (2000):

$$A^{1/2} \circ B^{1/2} \leq I.$$

Such observations on Hadamard products involving correlation matrices and applications in multivariate statistical analysis can be traced back to Styan (1973).

REFERENCES

G.P.H. Styan (1973) Hadamard products and multivariate statistical analysis. *Linear Algebra and Its Applications* 6, 217–240.

Zhang, F.-Z. (2000) Hadamard product of square roots of correlation matrices. *Image*, 25, 17.

02.4.2 *On the Rank of a Matrix Useful in Goodness-of-Fit Testing of Structural Equation Models*, proposed by Albert Satorra and Heinz Neudecker. Let A be an $n \times q$ matrix, $M := I_n - AA^+$ and $V(n \times n) \geq 0$. Show that

$$r(MVM) + r(A) = r(V) \quad \textbf{(1)}$$

iff

$$col(A) \subset col(V), \quad \textbf{(2)}$$

where $r(\cdot)$ denotes rank of a matrix, $col(\cdot)$ denotes column space, and A^+ is the Moore–Penrose inverse of A.

Show that an equivalent formulation of the problem is obtained when $M := I_n - A(A'WA)^+A'W$, with W positive definite and MVM changed to MVM'. This formulation of the problem is used in Satorra and Neudecker (2001) to obtain the rank of a goodness-of-fit test for structural equation models when the variance matrix V of sample moments is singular.

© 2002 Cambridge University Press 0266-4666/02 $9.50

https://doi.org/10.1017/S0266466602004127 Published online by Cambridge University Press

REFERENCES

Satorra, A. & H. Neudecker (2001) A Matrix Equality Useful in Goodness-of-Fit Testing of Structural Equation Models. Mimeo, Universitat Pompeu Fabra, Barcellona.

SOLUTIONS

01.4.1. *An Alternative GLS-Like Transformation in Regression Models with AR(1)-Errors—Solution*, proposed by O. Linton and D. Kristensen, the posers of the problem. First we note that the ordinary least squares (OLS) estimators, $\hat{\beta}$ and $\hat{\rho}$, are consistent:

$$\hat{\beta} - \beta = \frac{T^{-1}\sum_t X_t u_t}{T^{-1}\sum_t X_t^2} \xrightarrow{P} \frac{E(X_t u_t)}{E(X_t^2)} = 0,$$

$$\hat{\rho} - \rho = \frac{T^{-1}\sum_t \hat{u}_{t-1}[(\beta - \hat{\beta})X_t + \rho(u_{t-1} - \hat{u}_{t-1}) + \varepsilon_t]}{T^{-1}\sum_{t=2}^{T} \hat{u}_{t-1}^2} \xrightarrow{P} \frac{E(u_{t-1}\varepsilon_t)}{E(u_t^2)} = 0.$$

Thus, $\tilde{\beta}$ is consistent also, because

$$\tilde{\beta} - \beta = \frac{\sum_t X_t \varepsilon_t}{\sum_t X_t^2} + \rho(\hat{\beta} - \beta)\frac{\sum_t X_t X_{t-1}}{\sum_t X_t^2} - (\hat{\rho} - \rho)\frac{\sum_t X_t u_{t-1}}{\sum_t X_t^2}$$

$$+ (\hat{\rho} - \rho)(\hat{\beta} - \beta)\frac{\sum_t X_t X_{t-1}}{\sum_t X_t^2} \xrightarrow{P} 0.$$

Using the assumption that $E(u_{t-1}X_t) = 0$ we have that

$$\frac{T^{-1}\sum_t X_t u_{t-1}}{T^{-1}\sum_t X_t^2} \xrightarrow{P} \frac{E(X_t u_{t-1})}{E(X_t^2)} = 0$$

by a law of large numbers for the stationary mixing process $(X_t u_{t-1}, X_t^2)$. Similarly,

$$\frac{T^{-1}\sum_t X_t X_{t-1}}{T^{-1}\sum_t X_t^2} \xrightarrow{P} \frac{E(X_t X_{t-1})}{E(X_t^2)} = \rho_X.$$

https://doi.org/10.1017/S0266466602004127 Published online by Cambridge University Press

Provided also that $\hat{\rho} - \rho = O_p(T^{-1/2})$ and $\hat{\beta} - \beta = O_p(T^{-1/2})$, we have

$$\sqrt{T}(\tilde{\beta} - \beta) = \frac{\frac{1}{\sqrt{T}}\sum_t X_t \varepsilon_t}{\frac{1}{T}\sum_t X_t^2} + \rho\sqrt{T}(\hat{\beta} - \beta)\frac{\frac{1}{T}\sum_t X_t X_{t-1}}{\frac{1}{T}\sum_t X_t^2} + o_p(1)$$

$$= \frac{\frac{1}{\sqrt{T}}\sum_t X_t \varepsilon_t}{\frac{1}{T}\sum_t X_t^2} + \rho\frac{\frac{1}{\sqrt{T}}\sum_t X_t u_t}{T^{-1}\sum_t X_t^2}\rho_X + o_p(1)$$

$$= \frac{1}{\sqrt{T}}\sum_{t=1}^{T} v_t + o_p(1),$$

where $v_t = (X_t/\sigma_X^2)(\varepsilon_t + \rho u_t \rho_X)$ using the fact that $T^{-1}\sum_t X_t^2 \xrightarrow{P} \sigma_X^2$. Hence,

$$\text{asvar}[\sqrt{T}(\tilde{\beta} - \beta)] = \gamma_v(0) + 2\sum_{j=1}^{\infty}\gamma_v(j),$$

where $\gamma_v(\cdot)$ is the autocovariance function of the process v_t.

For comparison, the asymptotic variance of $\hat{\beta}_{FGLS}$ is

$$\text{asvar}[\sqrt{T}(\hat{\beta}_{FGLS} - \beta)] = \frac{\sigma_\varepsilon^2}{E[(X_t - \rho X_{t-1})^2]} = \frac{\sigma_\varepsilon^2}{\sigma_X^2[1 + \rho^2 - 2\rho\rho_X]},$$

and the asymptotic variance of OLS is

$$\text{asvar}[\sqrt{T}(\hat{\beta} - \beta)] = \gamma_{v_*}(0) + 2\sum_{j=1}^{\infty}\gamma_{v_*}(j),$$

where the process $v_{*t} = X_t u_t/\sigma_X^2$.

When X_t is serially uncorrelated, $\tilde{\beta}$ has asymptotic variance $\sigma_\varepsilon^2/\sigma_X^2$, $\hat{\beta}_{FGLS}$ has variance $\sigma_\varepsilon^2/\sigma_X^2(1 + \rho^2)$, and $\hat{\beta}$ has asymptotic variance $\sigma_u^2/\sigma_X^2 = \sigma_\varepsilon^2/\sigma_X^2(1 - \rho^2)$. In this case,

$$\text{asvar}(\hat{\beta}_{FGLS}) < \text{asvar}(\tilde{\beta}) < \text{asvar}(\hat{\beta})$$

whenever $\rho \neq 0$.

https://doi.org/10.1017/S0266466602004127 Published online by Cambridge University Press

It is easy however to find cases when $\text{asvar}(\tilde{\beta}) > \text{asvar}(\hat{\beta})$. Assuming that

$$E[\varepsilon_s \varepsilon_{s'} X_t X_{t'}] = E[\varepsilon_s \varepsilon_{s'}] E[X_t X_{t'}]$$

for all s, s', t, t', we have

$$\gamma_{v_*}(j) = \frac{\gamma_X(j)}{\sigma_X^4} \frac{\rho^j \sigma_\varepsilon^2}{1-\rho^2}$$

$$\gamma_v(0) = \frac{\gamma_X(0)}{\sigma_X^4} \sigma_\varepsilon^2 \left(1 + \frac{\rho^2 \rho_X^2}{1-\rho^2} + 2\rho\rho_X\right)$$

$$\gamma_v(j) = \frac{\gamma_X(j)}{\sigma_X^4} \rho^j \sigma_\varepsilon^2 \left(\rho\rho_X + \frac{\rho^2 \rho_X^2}{1-\rho^2}\right).$$

Assume also that $\gamma_X(j) = 0, j > 1$ (e.g., X is an MA(1)). We then get

$$\text{asvar}[\sqrt{T}(\hat{\beta} - \beta)] = \frac{\sigma_\varepsilon^2}{\sigma_X^2} \frac{1 + 2\rho\rho_X}{1-\rho^2},$$

whereas

$$\text{asvar}[\sqrt{T}(\tilde{\beta} - \beta)] = \frac{\sigma_\varepsilon^2}{\sigma_X^2} \left(1 + 2\rho\rho_X + 2\rho^2\rho_X^2 + \frac{1 + 2\rho\rho_X}{1-\rho^2}\right).$$

Therefore, $\text{asvar}(\tilde{\beta}) > \text{asvar}(\hat{\beta})$ if and only if

$$2\rho\rho_X^3 + (3 - 2\rho^2)\rho_X^2 - 2\rho\rho_X - 1 > 0.$$

There are many combinations of ρ, ρ_X that satisfy this inequality. For example, if $\rho = 0$ then any ρ_X with $|\rho_X| > 1/\sqrt{3}$ will do.

Note that assumption (3) may be relaxed. This has to be done carefully however. For instance, if $X_t = Y_{t-1}$ we get that

$$\sqrt{T}(\hat{\beta} - \beta) = \frac{T^{-1/2} \sum_t Y_{t-1} u_t}{T^{-1} \sum_t Y_{t-1}^2}.$$

The denominator is well behaved, $T^{-1} \sum_t Y_{t-1}^2 \xrightarrow{P} E(Y_t^2)$, but the numerator,

$$T^{-1/2} \sum_t Y_{t-1} u_t = T^{-1/2} \sum_t \sum_{i=0}^{\infty} \beta^i u_t u_{t-1-i},$$

is not because u_t and u_{t-1-i} are correlated, $E(u_t u_{t-1-i}) = \sigma_\varepsilon^2[\rho^{i+1}/(1-\rho^2)]$. Thus, we cannot readily obtain asymptotic normality of the OLS estimator; the estimator $\tilde{\beta}$ will inherit this property.

https://doi.org/10.1017/S0266466602004127 Published online by Cambridge University Press

01.4.2. *Nonorthogonal Hilbert Projections in Trend Regression—Solution*, proposed by P.C.B. Phillips and Y. Sun.

Part A: Brownian Motion Case

Define the operator P as

$$PX(t) = \hat{\beta}' Z(t) = \left(\int_0^1 X(s)Z(s)'\, ds \right) \left(\int_0^1 Z(s)Z(s)'\, ds \right)^{-1} Z(t);$$

then

$$\begin{aligned} P^2X &= \left(\int_0^1 PX(s)Z(s)'\, ds \right) \left(\int_0^1 Z(s)Z(s)'\, ds \right)^{-1} Z(t) \\ &= \left[\left(\int_0^1 XZ'\, ds \right) \left(\int_0^1 ZZ'\, ds \right)^{-1} \int_0^1 ZZ'\, ds \right] \left(\int_0^1 ZZ'\, ds \right)^{-1} Z(t) \\ &= \left(\int_0^1 X(s)Z(s)'\, ds \right) \left(\int_0^1 Z(s)Z(s)'\, ds \right)^{-1} Z(t) = PX. \end{aligned}$$

In addition, it is easy to show that $(PX, Y) = (PY, X)$. Thus P is an orthogonal Hilbert projector onto the manifold spanned by $Z(t)$.

Define the operator Q as

$$QX = \tilde{\beta}' Z = \left(\int_0^1 dX(s)Z^{(1)}(s)' \right) \left(\int_0^1 Z^{(1)}(s)Z^{(1)}(s)'\, ds \right)^{-1} Z(t).$$

Then Q is idempotent because

$$\begin{aligned} Q^2X &= \left(\int_0^1 d[QX]Z^{(1)}(s)' \right) \left(\int_0^1 Z^{(1)}(s)Z^{(1)}(s)'\, ds \right)^{-1} Z(t) \\ &= \left(\int_0^1 dX(s)Z^{(1)}(s)' \right) \left(\int_0^1 Z^{(1)}(s)Z^{(1)}(s)'\, ds \right)^{-1} \\ &\quad \times \left(\int_0^1 dZ(s)Z^{(1)}(s)' \right) \left(\int_0^1 Z^{(1)}(s)Z^{(1)}(s)'\, ds \right)^{-1} Z(t) \\ &= \left(\int_0^1 dX(s)Z^{(1)}(s)' \right) \left(\int_0^1 Z^{(1)}(s)Z^{(1)}(s)'\, ds \right)^{-1} Z(t) \\ &= QX. \end{aligned}$$

But Q is not orthogonal because

$$(QX, Y) = \left(\int_0^1 dX(s)Z^{(1)}(s)'\, ds \right) \left(\int_0^1 Z^{(1)}(s)Z^{(1)}(s)'\, ds \right)^{-1} \left(\int_0^1 Z(s)Y(s)\, ds \right),$$

https://doi.org/10.1017/S0266466602004127 Published online by Cambridge University Press

whereas

$$(X, QY) = \left(\int_0^1 dY(s) Z^{(1)}(s)' \, ds \right) \left(\int_0^1 Z^{(1)}(s) Z^{(1)}(s)' \, ds \right)^{-1} \left(\int_0^1 Z(s) X(s) \, ds \right),$$

and so $(QX, Y) \neq (X, QY)$ in general.

Therefore both PX and QX are Hilbert projections of $X(t)$ onto the space spanned by $Z(t)$. The difference lies in that PX is an orthogonal projection whereas QX is a nonorthogonal projection. We now investigate the statistical properties of these two projections.

First consider the distribution of $\hat{\beta}$. Because

$$\hat{\beta} - \beta = \left(\int_0^1 Z(s) Z(s)' \, ds \right)^{-1} \int_0^1 Z(s) W(s) \, ds$$

in a linear functional of a Brownian motion, we know that $\hat{\beta} - \beta \sim N(0, V_1)$, where

$$V_1 = \left(\int_0^1 Z(s) Z(s)' \, ds \right)^{-1} \left(\int_0^1 \int_0^1 Z(t)(s \wedge t) Z(s)' \, dt \, ds \right) \left(\int_0^1 Z(s) Z(s)' \, ds \right)^{-1}.$$

Next consider the distribution of $\tilde{\beta}$. Because

$$\tilde{\beta} - \beta = \left(\int_0^1 Z^{(1)}(s) Z^{(1)}(s)' \, ds \right)^{-1} \int_0^1 Z^{(1)}(s) \, dW(s)$$

is also a linear functional of a Brownian motion, we know immediately that $\tilde{\beta} - \beta \sim N(0, V_2)$, where

$$V_2 = \left(\int_0^1 Z^{(1)}(s) Z^{(1)}(s)' \, ds \right)^{-1}.$$

We now compare the distributions of $\hat{\beta}$ and $\tilde{\beta}$ when $p = 1$. In this case $V_1 = \frac{6}{5}$ and $V_2 = \int_0^1 dr = 1$. Therefore $\tilde{\beta}$ is a more efficient estimator than $\hat{\beta}$ and the efficiency gain is 20%.

Remark. Note that $\hat{\beta}' Z(t)$ is the best $L_2[0,1]$ functional approximation to $\beta' Z(t)$ and it is delivered by an orthogonal projection of $X(t)$ onto the manifold spanned by $Z(t)$. On the other hand, $\tilde{\beta}' Z(t)$ is the better (more efficient) statistical estimator of $\beta' Z(t)$, and it is delivered by a nonorthogonal projection onto the same space.

Part B: The Diffusion Process Case

Both $\hat{\beta}' Z(t)$ and $\tilde{\beta}' Z(t)$ are still Hilbert projections. The proofs are essentially the same as before and are omitted.

https://doi.org/10.1017/S0266466602004127 Published online by Cambridge University Press

We proceed to compute the distributions of $\hat{\beta}$ and $\tilde{\beta}$. First

$$\hat{\beta} - \beta = \left(\int_0^1 Z(s)Z(s)'\, ds \right)^{-1} \int_0^1 Z(t)J_c(t)\, dt \sim N(0, V_3),$$

where

$$\begin{aligned} V_3 &= \left(\int_0^1 Z(s)Z(s)'\, ds \right)^{-1} \mathrm{Var}\left[\int_0^1 Z(t)J_c(t)\, dt \right] \left(\int_0^1 Z(s)Z(s)'\, ds \right)^{-1} \\ &= \left(\int_0^1 ZZ'\, ds \right)^{-1} \left\{ \int_0^1 \left(\int_0^1 Z(r)\Lambda_{r,s}Z(s)'\, ds \right) dr \right\} \left(\int_0^1 ZZ'\, ds \right)^{-1} \end{aligned}$$

and $\Lambda_{r,s} = EJ_c(r)J_c(s) = (e^{(r+s)c} - e^{|r-s|c})/(2c)$ is the covariance kernel of the O-U process.

Next, the error process $J_c(t) = \int_0^t e^{(t-s)c}\, dW(s)$ in (2) satisfies the linear differential equation

$$dJ_c(t) = cJ_c(t)\, dt + dW(t),$$

so we can write $\tilde{\beta} - \beta$ as

$$\begin{aligned} \tilde{\beta} - \beta &= \left(\int_0^1 Z^{(1)}(s)Z^{(1)}(s)'\, ds \right)^{-1} \left(Z^{(1)}(1)J_c(1) - \int_0^1 Z^{(2)}(s)J_c(s)\, ds \right) \\ &\sim N(0, V_4), \end{aligned}$$

where the equality follows from integration by parts and

$$\begin{aligned} V_4 = \left(\int_0^1 Z^{(1)}Z^{(1)\prime} \right)^{-1} \Bigg(& \int_0^1 \int_0^1 Z^{(2)}(r)\Lambda_{r,s}Z^{(2)}(s)'\, dsdr \\ & + Z^{(1)}(1)\Lambda_{1,1}Z^{(1)}(1)' - 2\int_0^1 Z^{(1)}(1)\Lambda_{1,s}Z^{(2)}(s)'\, ds \Bigg) \\ \times \left(\int_0^1 Z^{(1)}Z^{(1)\prime} \right)^{-1}. \end{aligned}$$

When $p = 1$, some algebra gives $V_3 = (3/2c^5)[2c^3 + 3c^2 - 3 + 3(c - 1)^2e^{2c}]$ and $V_4 = \Lambda_{1,1} = (e^{2c} - 1)/2c$. So $V_3/V_4 = 3[2c^3 + 3c^2 - 3 + 3(c - 1)^2e^{2c}]/[c^4(e^{2c} - 1)]$. Figure 1 graphs the relative efficiency of $\tilde{\beta}$ and $\hat{\beta}$, i.e., V_3/V_4 against c. It reveals that $\hat{\beta}$ is more efficient than $\tilde{\beta}$ for small values of c. As c approaches zero, $\tilde{\beta}$ becomes more efficient, which is consistent with our results in Part A. As c approaches zero from below, the linear diffusion J_c behaves like the underlying Brownian motion and the estimator delivered by the nonorthogonal projection is more efficient. The cutoff value of c is approximately -3.80. When $c < -3.80$, $V_3 < V_4$; when $-3.80 \leq c < 0$, $V_3 \geq V_4$.

https://doi.org/10.1017/S0266466602004127 Published online by Cambridge University Press

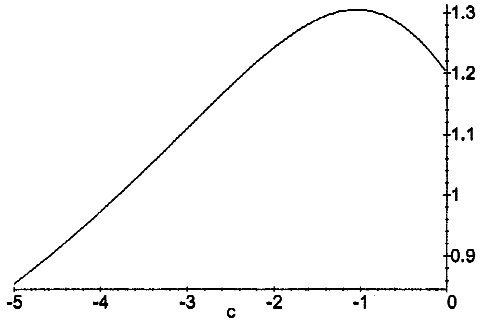


FIGURE 1. V_3/V_4 as a function of c.

A. More Efficient Estimator

We now propose another unbiased estimator $\bar{\beta}$, which is more efficient than both $\hat{\beta}$ and $\tilde{\beta}$. Define $Z_c(t) = Z^{(1)}(t) - cZ(t)$ and

$$\bar{\beta} = \left(\int_0^1 Z_c(t) Z_c(t)' \, dt \right)^{-1} \left(\int_0^1 Z_c(t) \, dX(t) - c \int_0^1 Z_c(t) X(t) \, dt \right). \tag{3}$$

Recall that $dJ_c(t) = cJ_c(t)\,dt + dW(t)$, so we may linearly transform the model (2) to

$$dX(t) - cX(t)\,dt = \beta' [Z^{(1)}(t) - cZ(t)]\,dt + dW(t), \qquad t \in [0,1], \tag{4}$$

in which the innovations $dW(t)$ are independent. The estimator $\bar{\beta}$ in (3) is the "least squares" estimator of β in the system (4). It follows from (3) and (4) that

$$\bar{\beta} - \beta = \left(\int_0^1 Z_c(t) Z_c(t)' \, dt \right)^{-1} \left(\int_0^1 Z_c(t) \, dW(t) \right) \sim N(0, V_5).$$

where

$$V_5 = \left(\int_0^1 Z_c(t) Z_c(t)' \, dt \right)^{-1}.$$

When $p = 1$, $V_5 = (\int_0^1 (ct - 1)^2 \, dt)^{-1} = 3/(c^2 - 3c + 3)$. Figure 2 graphs V_3/V_5 and V_4/V_5 against c. It shows that $\bar{\beta}$ is indeed more efficient than both $\hat{\beta}$ and $\tilde{\beta}$ for all values of $c < 0$.

Define the operator M as

$$MX = \left(\int_0^1 dX(t) Z_c(t)' - c \int_0^1 X(t) Z_c(t)' \, dt \right) \left(\int_0^1 Z_c(t) Z_c(t)' \, dt \right)^{-1} Z(t). \tag{5}$$

https://doi.org/10.1017/S0266466602004127 Published online by Cambridge University Press

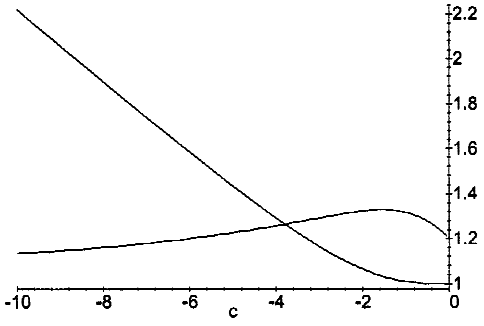


FIGURE 2. V_3/V_5 and V_4/V_5 as a function of c.

Then, because $Z(t)$ is continuously differentiable,

$$\int_0^1 dZ(t)Z_c(t)' - c\int_0^1 Z(t)Z_c(t)'\,dt = \int_0^1 Z_c(t)Z_c(t)'\,dt,$$

we have

$$\begin{aligned} MZ &= \left(\int_0^1 dZ(t)Z_c(t)' - c\int_0^1 Z(t)Z_c(t)'\,dt\right)\left(\int_0^1 Z_c(t)Z_c(t)'\,dt\right)^{-1} Z(t) \\ &= Z(t) \end{aligned}$$

and by virtue of the definition (5)

$$\begin{aligned} M^2X &= MMX \\ &= \left(\int_0^1 dX(t)Z_c(t)' - c\int_0^1 X(t)Z_c(t)'\,dt\right)\left(\int_0^1 Z_c(t)Z_c(t)'\,dt\right)^{-1} MZ(t) \\ &= \left(\int_0^1 dX(t)Z_c(t)' - c\int_0^1 X(t)Z_c(t)'\,dt\right)\left(\int_0^1 Z_c(t)Z_c(t)'\,dt\right)^{-1} Z(t) \\ &= MX. \end{aligned}$$

So M is a Hilbert projection. In addition, it is easy to see that $(MX, Y) \neq (X, MY)$ in general. Therefore, $\bar{\beta}'Z(t) = MX$ is a nonorthogonal Hilbert projection. However it delivers the minimum variance unbiased trend estimator in this case.

Remarks.

(a) The transformed system (4) is the continuous time version of the traditional Cochrane–Orcutt transformation of a discrete time regression model to remove autoregressive errors.

(b) The results obtained here clearly apply for any differentiable deterministic sequence $Z(t)$ for which $\int_0^1 ZZ'\,dt$, $\int_0^1 Z^{(1)}Z^{(1)}$, and $\int_0^1 Z_c Z_c'\,dt$ are positive definite.

https://doi.org/10.1017/S0266466602004127 Published online by Cambridge University Press

01.4.3 *The Limit Distribution of Cointegration Rank Tests of "Wald" Type—Solution*,[1] proposed by Anders Rahbek. Notation and results in Johansen (1996, Chs. 10 and 11) are used in the solution.

(a) For $p = 1$ $S_{ij} = M_{ij}$, and thus (2) becomes $|\rho S_{00} - (\rho + 1)S_{01}S_{11}^{-1}S_{10}|$, showing that $\lambda = \rho/(\rho + 1)$ and hence $\rho = \lambda/(1 - \lambda)$; this relation is monotonically increasing on $[0,1)$.

(b) In Johansen (1996) it is shown that the r biggest eigenvalues $\hat{\lambda}_1, \ldots, \hat{\lambda}_r$ converge in probability to the eigenvalues of the matrix $\Sigma_{\beta\beta}^{-1/2}\Sigma_{\beta 0}\Sigma_{00}^{-1}\Sigma_{0\beta}\Sigma_{\beta\beta}^{-1/2}$ and that the remaining $(n - r)$ eigenvalues λ_j, $j = r + 1, \ldots, n$ converge in probability to zero. Furthermore, $T\hat{\lambda}_{r+1}, \ldots, T\hat{\lambda}_n$ converge weakly to the eigenvalues of $\int_0^1 dBB'(\int_0^1 BB'\,du)^{-1}\int_0^1 BdB'$, where B is a standard Brownian motion of dimension $n - r$. Using the relation $\hat{\rho} = \hat{\lambda}/(1 - \hat{\lambda})$ from (a), therefore $\hat{\rho}_i$ converges in probability for $i = 1, \ldots, r$, whereas a first-order expansion (see, e.g., Brockwell and Davies 1991, p. 201) of $\lambda/(1 - \lambda)$ gives $T\hat{\rho}_j = T\hat{\lambda}_j + o_p(1)$ for $j = r + 1, \ldots, n$.

(c) Note initially that the eigenvalue problem in (2) has the same eigenvalues as the dual one given by

$$|S(\rho)| := |\rho S_{11}M_{11}^{-1}S_{11} - S_{10}S_{00.1}^{-1}S_{01}| = 0 \qquad \textbf{(2′)}$$

and then follow the outline of the proof of Theorem 11.1 in Johansen (1996). Consider first the case of rank $r = 0$ and use the following simple results:

$$M_{11}^{-1}S_{11} = I_p + O_P(T^{-1}), \qquad S_{10} = S_{1\varepsilon} \xrightarrow{w} \int_0^1 GdW', \quad \text{and}$$

$$S_{00.1} = S_{00} + o_P(T^{-1/2}) \xrightarrow{p} \Sigma_{00} = \Omega,$$

where, in general, $G = \beta_\perp' CW$ and W is an n-dimensional Brownian motion with covariance Ω. When $r = 0$, $G = \Gamma^{-1}W$ (see Johansen, for definition of Γ, 1996, p. 45). In the limit ρ_i solve $|\rho\int_0^1 GG'\,du| = 0$ and are hence zero. With $\psi = T\rho$, likewise as $T \to \infty$, the eigenvalue problem tends to

$$\left|\psi\int_0^1 GG'\,du - \int_0^1 GdW'\Omega^{-1}\int_0^1 dWG'\right| = 0,$$

and the result follows by the definition $B = \Omega^{-1/2}W$.

For the general case of rank r, note initially that the $n - r$ smallest eigenvalues tend to zero and that the r largest are positive in the limit. To see this, define $A_T = (\beta, \beta_\perp T^{-1/2})$ and $\mu_{\beta\beta} = \mathrm{Var}(\beta' X_t)$ and mimic Johansen (1996, p. 158), using the following results:

$$\begin{aligned}\beta' S_{11}M_{11}^{-1}S_{11}\beta &= \beta' S_{11}A_T(A_T'M_{11}A_T)^{-1}A_T'S_{11}\beta \\ &= \Sigma_{\beta\beta}\mu_{\beta\beta}^{-1}\Sigma_{\beta\beta} + o_P(1),\end{aligned}$$

$$\begin{aligned}S_{00.1} &= S_{00} - S_{01}A_T(A_T'S_{11}A_T)^{-1}A_T'S_{10} \xrightarrow{p} \Sigma_{00.\beta} \\ &= \Sigma_{00} - \Sigma_{0\beta}\Sigma_{\beta\beta}^{-1}\Sigma_{\beta 0},\end{aligned}$$

$$T^{-1}\beta_\perp' S_{11}M_{11}^{-1}S_{11}\beta_\perp \xrightarrow{w} \int_0^1 GG'\,du,$$

$$T^{-1/2}\beta_\perp' S_{11}M_{11}^{-1}S_{11}\beta = o_P(1).$$

https://doi.org/10.1017/S0266466602004127 Published online by Cambridge University Press

Then the r largest eigenvalues tend to the positive eigenvalues of

$$(\Sigma_{\beta\beta}\mu_{\beta\beta}^{-1}\Sigma_{\beta\beta})^{-1/2}\Sigma_{\beta 0}\Sigma_{00.\beta}^{-1}\Sigma_{0\beta}(\Sigma_{\beta\beta}\mu_{\beta\beta}^{-1}\Sigma_{\beta\beta})^{-1/2}.$$

Note that these are different from the ones in Johansen (1996). For the study of the limit of the scaled smallest eigenvalues, $\psi = T\rho$ use the results

$$\beta_\perp' S(\psi T^{-1})\beta = -\beta_\perp' S_{10}\Sigma_{00.\beta}^{-1}\Sigma_{0\beta} + o_P(1),$$
$$\beta_\perp' S(\psi T^{-1})\beta_\perp = \psi T^{-1}\beta_\perp' S_{11}M_{11}^{-1}S_{11}\beta_\perp - \beta_\perp' S_{10}\Sigma_{00.\beta}^{-1}S_{01}\beta_\perp + o_P(1),$$
$$\beta' S(\psi T^{-1})\beta = -\Sigma_{\beta 0}\Sigma_{00.\beta}^{-1}\Sigma_{0\beta} + o_P(1).$$

Asymptotically as in Johansen (1996, p. 159), ψ_i solve the limit of

$$|\psi T^{-1}\beta_\perp' S_{11}M_{11}^{-1}S_{11}\beta_\perp - \beta_\perp' S_{10}\Sigma S_{01}\beta_\perp| = 0,$$

where

$$\begin{aligned}\Sigma &= \Sigma_{00.\beta}^{-1} - \Sigma_{00.\beta}^{-1}\Sigma_{0\beta}(\Sigma_{\beta 0}\Sigma_{00.\beta}^{-1}\Sigma_{0\beta})^{-1}\Sigma_{0\beta}\Sigma_{00.\beta}^{-1}\\ &= \Sigma_{00.\beta}^{-1}[I - \alpha(\alpha'\Sigma_{00.\beta}^{-1}\alpha)^{-1}\alpha'\Sigma_{00.\beta}^{-1}] \quad \text{(Johansen, 1996, Lemma 10.1)}\\ &= \Sigma_{00.\beta}^{-1}[\Sigma_{00.\beta}\alpha_\perp(\alpha_\perp'\Sigma_{00.\beta}\alpha_\perp)^{-1}\alpha_\perp'] \quad \text{(skew projection)}\\ &= \alpha_\perp(\alpha_\perp'\Omega\alpha_\perp)^{-1}\alpha_\perp'.\end{aligned}$$

The last line follows by the identity $\alpha_\perp'\Sigma_{00.\beta} = \alpha_\perp'\Sigma_{00} - \alpha_\perp'\Sigma_{0\beta}\Sigma_{\beta\beta}^{-1}\Sigma_{\beta 0} = \alpha_\perp'\Omega - \alpha_\perp'\alpha\Sigma_{\beta 0} = \alpha_\perp'\Omega$. Hence by finally using

$$\beta_\perp' S_{10}\alpha_\perp = \beta_\perp' S_{1\varepsilon}\alpha_\perp \xrightarrow{w} \int_0^1 GdW'\alpha_\perp$$

(cf. Johansen, 1996, Lemma 10.3) one finds as desired that ψ_i in the limit solves

$$\left|\psi\int_0^1 GG'\,du - \int_0^1 GdW'\alpha_\perp(\alpha_\perp'\Omega\alpha_\perp)^{-1}\alpha_\perp'\int_0^1 dWG'\right| = 0.$$

NOTE

1. An independent solution has been proposed by Paolo Paruolo, the poser of the problem.

REFERENCES

Brockwell, P. & R. Davis (1991) *Time Series: Theory and Methods*, 2nd edition. New York: Springer Verlag.

Johansen, S. (1996) *Likelihood-Based Inference in Cointegrated Vector Auto-regressive Models*. Oxford: Oxford University Press.

https://doi.org/10.1017/S0266466602004127 Published online by Cambridge University Press