Econometric Theory, **15**, 1999, 629–637. Printed in the United States of America.

PROBLEMS AND SOLUTIONS

PROBLEMS

99.4.1. *The KPSS Test for Cointegration in Case of Bivariate Regressions with Linear Trends*, proposed by Uwe Hassler. Consider the cointegration model

$$y_t = a + b'x_t + e_t, \qquad b \neq 0, \qquad e_t \sim I(0), \qquad t = 1,\ldots,T, \tag{1}$$

where x_t is integrated of order one, $x_t - x_{t-1} = u_t$. Let ϵ_t denote the ordinary least squares (OLS) residuals from the regression of (1). Shin (1994) investigated the limiting distribution of the so-called KPSS statistic based on regression residuals to test for the null of cointegration. It turns out that it depends on

$$\Delta = \sum_{k=0}^{\infty} E(u_t e_{t+k}), \qquad \Omega = \sum_{k=-\infty}^{\infty} E(u_t e_{t+k}).$$

To get rid of these nuisance parameters efficient modifications of OLS are required.

We now assume that x_t is (a) a scalar $I(1)$ series (b) with drift

$$x_t - x_{t-1} = \mu + u_t, \qquad \mu \epsilon \mathcal{R}/\{0\}, \qquad u_t \sim I(0). \tag{2}$$

Prove with assumption (2) that the limiting distribution of the KPSS statistic based on OLS residuals from (1)

(i) does not depend on Δ and Ω (and hence no efficient modifications are necessary);
(ii) is given by Kwiatkowski, Phillips, Schmidt, and Shin (1992) in case of the detrended univariate KPSS test (and hence percentiles from the lower panel in Kwiatkowski et al., 1992, Table 1, are correct instead of those from the middle panel in Shin, 1994, Table 1).

REFERENCES

Kwiatkowski, D., P.C.B. Phillips, P. Schmidt, & Y. Shin (1992) Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics* 54, 159–178.

Shin, Y. (1994) A residual-based test of the null of cointegration against the alternative of no cointegration. *Econometric Theory* 10, 91–115.

99.4.2. *A Modified Estimator of the Error Variance in Linear Errors-in-Variables Models*, proposed by Hua Liang. Let $\{X_i, Y_i, W_i, i = 1,\ldots,n\}$ be independent and identically distributed (i.i.d.) observations from $Y_i = X_i^T\beta + \varepsilon_i$ and $W_i = X_i + U_i$, where X_i are i.i.d. p-dimensional variables with mean zero, ε_i are i.i.d. random errors with mean zero and finite variance σ^2, W_i are the surrogates of X_i, that is, we can not observe X_i but instead W_i. We also assume that the measurement errors

© 1999 Cambridge University Press 0266-4666/99 $9.50

https://doi.org/10.1017/S0266466699154070 Published online by Cambridge University Press

U_i are i.i.d., independent of (Y_i, X_i), with mean zero and covariance matrix Σ_{uu}, and that Σ_{uu} is known.

The moment estimates of β and σ^2 at $\sqrt{n}$-rates of convergence can be explicitly expressed as $\beta_n = (\mathbf{W}^T\mathbf{W} - n\Sigma_{uu})^{-1}(\mathbf{W}^T\mathbf{Y})$ and

$$\sigma_n^2 = n^{-1}\sum_{i=1}^{n}(Y_i - W_i^T\beta_n)^2 - \beta_n^T\Sigma_{uu}\beta_n \tag{1}$$

where $\mathbf{W} = (W_1,\ldots,W_n)^T$ and $\mathbf{Y} = (Y_1,\ldots,Y_n)^T$. Then we can prove that, if $E(\varepsilon^4 + \|U\|^4) < \infty$, then $n^{1/2}(\sigma_n^2 - \sigma^2)$ converges to $N(0,\sigma_*^2)$ in distribution, where $\sigma_*^2 = E\{(\varepsilon - U^T\beta)^2 - (\beta^T\Sigma_{uu}\beta + \sigma^2)\}^2$.

We are interested in whether we can construct estimators for the variance σ^2, whose asymptotic variance is smaller than σ_*^2. The following result gives a positive answer for this issue. The reader is asked to prove it.

Define $\hat{\sigma}_n^2 = \sigma_n^2 - \nu_n/n\xi_n^2\sum_{i=1}^n\hat{\varepsilon}_i$ as the estimator of σ^2, where $\xi_n = 1/n\sum_{i=1}^n\hat{\varepsilon}_i^2$ and $\nu_n = 1/n\sum_{i=1}^n\hat{\varepsilon}_i^3$ with $\hat{\varepsilon}_i = Y_i - W_i^T\beta_n$.

THEOREM. *Suppose $E(WW^T)$ is a positive definite matrix, and $E(\varepsilon^4 + \|U\|^4) < \infty$. Then $\sqrt{n}(\hat{\sigma}_n^2 - \sigma^2)$ converges to $N(0,\sigma_m^2)$ in distribution, where $\sigma_m^2 = \sigma_*^2 - \nu^2a^{-2}$ with $\nu = E(\varepsilon - U^T\beta)^3$ and $a^2 = E(\varepsilon - U^T\beta)^2$.*

99.4.3. *The Relative Efficiency of the Between Estimator with Respect to the Within Estimator*, proposed by Badi H. Baltagi. Consider the simple panel data regression model

$$y_{it} = \alpha + \beta x_{it} + u_{it} \qquad i = 1,2,\ldots,N; \qquad t = 1,2,..,T, \tag{1}$$

where α and β are scalars. Subtract the mean equation to get rid of the constant

$$y_{it} - \bar{y}_{..} = \beta(x_{it} - \bar{x}_{..}) + u_{it} - \bar{u}_{..}, \tag{2}$$

where $\bar{x}_{..} = \sum_{i=1}^N\sum_{t=1}^T x_{it}/NT$ and $\bar{y}_{..}$ and $\bar{u}_{..}$ are similarly defined. Add and subtract $\bar{x}_{i.}$ from the regressor in parentheses and rearrange

$$y_{it} - \bar{y}_{..} = \beta(x_{it} - \bar{x}_{i.}) + \beta(\bar{x}_{1.} - \bar{x}_{..}) + u_{it} - \bar{u}_{..}, \tag{3}$$

where $\bar{x}_{i.} = \sum_{t=1}^T x_{it}/T$. Now run the unrestricted least squares regression

$$y_{it} - \bar{y}_{..} = \beta_w(x_{it} - \bar{x}_{i.}) + \beta_b(\bar{x}_{i.} - \bar{x}_{..}) + u_{it} - \bar{u}_{..}, \tag{4}$$

where β_w is not necessarily equal to β_b.

(a) Show that the least squares estimator of β_w from (4) is the within estimator and that of β_b is the between estimator.

(b) Show that if $u_{it} = \mu_i + \nu_{it}$ where $\mu_i \sim \text{IID}(0,\sigma_\mu^2)$ and $\nu_{it} \sim IID(0,\sigma_\nu^2)$ independent of each other and among themselves, then ordinary least squares (OLS) is equivalent to generalized least squares (GLS) on (4).

https://doi.org/10.1017/S0266466699154070 Published online by Cambridge University Press

(c) Show that for model (1), the relative efficiency of the between estimator with respect to the within estimator is equal to $B_{xx}/W_{xx})[(1 - \rho)/(T\rho + (1 - \rho))]$, where $W_{xx} = \sum_{i=1}^{N}\sum_{t=1}^{T}(x_{it} - \bar{x}_{i.})^2$ denotes the within variation and $B_{xx} = T\sum_{i=1}^{N}(\bar{x}_{i.} - \bar{x}_{..})^2$ denotes the between variation. Also, $\rho = \sigma_\mu^2/(\sigma_\mu^2 + \sigma_\nu^2)$ denotes the equicorrelation coefficient.

(d) Show that the square of the *t*-statistic used to test $H.; \beta_w = \beta_b$ in (4) yields exactly Hausman's (1978) specification test.

REFERENCE

Hausman, J.A. (1978) Specification tests in econometrics. *Econometrica* 46, 1251–1271.

99.4.4. *An Upper Bound for the Eigenvalues of the Product of a Symmetric Idempotent and a Non-negative Definite Matrix*, proposed by Jan Graffelman and Michel van de Velden. Let **A** be an $n \times n$ non-negative definite matrix and let **B** be a symmetric idempotent matrix. Show that

$$\lambda_i(\mathbf{AB}) \leq \mu_i(\mathbf{A}) \qquad (i = 1, \ldots, n),$$

where $\lambda_i(\cdot)$ and $\mu_i(\cdot)$ represent the eigenvalues of the respective matrices in decreasing order of magnitude.[1]

NOTE

1. The spectral decomposition of **AB** is at the heart of canonical (restricted) correspondence analysis (see Jongman, Ter Braak, and van Tongeren, 1987, Sect. 5.9), and the eigenvalues are the principal inertias in the restricted analysis. In this context, **A** is the matrix subject to a spectral decomposition in ordinary correspondence analysis (see Greenacre, 1984, p. 92). For such a matrix **A** the eigenvalues are in the interval [0,1]. (See also a recent problem by Neudecker, Satorra, and van de Velden, 1997). The result $\lambda_i \leq \mu_i$ shows that the *i*th principal inertia in canonical correspondence analysis is always smaller than the *i*th principal inertia in ordinary correspondence analysis and consequently the inertias of the restricted analysis are also in the interval [0,1].

REFERENCES

Greenacre, M.J. (1984). *Theory and Applications of Correspondence Analysis*. New York: Academic Press.

Jongman, R.H.G., C.J.F. Ter Braak, & O.F.R. van Tongeren (1987) *Data Analysis in Community and Landscape Ecology*. Pudoc Wagenigen.

Neudecker, H., A. Satorra, & M. van de Velden (1997) A fundamental matrix result on scaling in multivariate analysis. *Econometric Theory* 13 (6), 890.

99.4.5. *Asymptotic Bias of the OLS Estimator for a Censored Pareto Regression Model*, proposed by S.K. Sapra. Greene (1981) derived the asymptotic bias in ordinary least squares (OLS) estimator for the censored normal regression model under the assumption of multivariate normality of the right-hand side vari-

https://doi.org/10.1017/S0266466699154070 Published online by Cambridge University Press

ables. Chung and Goldberger (1984) derived general results for a wide class of limited dependent variable models under weaker distributional assumptions.

Consider the censored regression model

$$y_i^* = \alpha + \beta' x_i + u_i, \qquad i = 1,2,\ldots,n, \tag{1}$$

where u_i are independent and identically distributed (i.i.d.) with mean 0 and variance σ^2 and are independent of x_i for $i = 1,2,\ldots,n$, α is an unknown scalar, β is an $(m \times 1)$ vector of unknown constants and $x_i = (x_{i1}, x_{i2},\ldots,x_{im})'$. The variable y^* is unobserved, and the observed response is $y = \min(k, y^*)$, where k is a known constant. Assume that $(y_i^*, x_i')', i = 1,2,\ldots,n$ are independently distributed with a joint multivariate $(m+1)$ Pareto density (suppressing the subscript i for convenience)

$$f(y^*,x) = a(a+1)\ldots(a+m)\left(y^* + \sum_{j=1}^{m} x_j - m\right)^{-(a+m+1)},$$
$$a > 0, \qquad y^* > 1, \qquad x_j > 1, \qquad j = 1,2,\ldots,m. \tag{2}$$

(a) Show that each element of $\hat{\beta}$, the OLS estimate of β, has the same proportional asymptotic bias as $n \to \infty$. Derive the common proportionality constant.

(b) Show that $\hat{\beta}$ and $\hat{\alpha}$, the OLS estimates of β and α, respectively, are asymptotically unbiased as $k \to \infty$ (as the degree of censoring approaches 0) and $n \to \infty$.

REFERENCES

Chung, C. & A.S. Goldberger (1984) Proportional projections in limited dependent variable models. *Econometrica* 52, 531–534.

Greene, W.H. (1981) On the asymptotic bias of the ordinary least squares estimator of the Tobit model. *Econometrica* 49, 505–514.

SOLUTIONS

98.4.1. *A Determinantal Identity*—Solution,[1] proposed by Simo Puntanen, George P.H. Styan, and Hans J. Werner. We will present two solutions.

Solution 1—Factoring out $|\mathbf{A}|$, we may write the right-hand side of (1) as

$$|\mathbf{A}+\mathbf{B}|/|\mathbf{A}| = |\mathbf{I} + \mathbf{A}^{-1}\mathbf{B}| = |\mathbf{I} + \mathbf{A}^{-1}\mathbf{X}\mathbf{X}'\mathbf{A} - \mathbf{X}\mathbf{X}'| = |\mathbf{I} + \mathbf{F}\mathbf{G}'|, \tag{2}$$

say, where $\mathbf{F} = (\mathbf{A}^{-1}\mathbf{X} : -\mathbf{X})$ and $\mathbf{G} = (\mathbf{A}\mathbf{X} : \mathbf{X})$. It is well known that

$$|\mathbf{I} + \mathbf{K}\mathbf{L}| = |\mathbf{I} + \mathbf{L}\mathbf{K}| \tag{3}$$

for any conformable matrices $\mathbf{K}$ and $\mathbf{L}$ (cf., e.g., Harville, 1997, p. 416). Putting $\mathbf{K} = \mathbf{F}$ and $\mathbf{L} + \mathbf{G}'$ in (3) yields

$$|\mathbf{I} + \mathbf{F}\mathbf{G}'| = |\mathbf{I} + \mathbf{G}'\mathbf{F}| = \begin{vmatrix} 2\mathbf{I}_k & -\mathbf{X}'\mathbf{A}\mathbf{X} \\ \mathbf{X}'\mathbf{A}^{-1}\mathbf{X} & \mathbf{0} \end{vmatrix}$$
$$= |2\mathbf{I}_k| \cdot |\mathbf{X}'\mathbf{A}\mathbf{X}(2\mathbf{I}_k)^{-1}\mathbf{X}'\mathbf{A}^{-1}\mathbf{X}|, \tag{4}$$

https://doi.org/10.1017/S0266466699154070 Published online by Cambridge University Press

as determinant is multiplicative on the Schur complement (cf., e.g., Harville, 1997, p. 189). Because the right-hand side of (4) simplifies at once to the left-hand side of (1), our first solution is complete.

Solution 2—There exists an $n \times (n-k)$ matrix $\mathbf{Y}$ so that the matrix $\mathbf{P} = (\mathbf{X} : \mathbf{Y})$ is orthogonal, and so $\mathbf{X}'\mathbf{Y} = \mathbf{0}, \mathbf{Y}'\mathbf{X} = \mathbf{0}$, and $\mathbf{Y}'\mathbf{Y} = \mathbf{I}_{n-k}$. Let $\mathbf{M} = \mathbf{I}_n + \mathbf{A}^{-1}\mathbf{X}\mathbf{X}'\mathbf{A} - \mathbf{X}\mathbf{X}'$ and let $\mathbf{W} = \mathbf{P}'\mathbf{M}\mathbf{P}$. Clearly

$$|\mathbf{W}| = |\mathbf{M}| = |\mathbf{A} + \mathbf{B}|/|\mathbf{A}|,$$

the right-hand side of (1). Moreover, because

$$\mathbf{W} = \begin{pmatrix} \mathbf{X}'\mathbf{A}^{-1}\mathbf{X}\mathbf{X}'\mathbf{A}\mathbf{X} & \mathbf{0} \\ \mathbf{Y}'\mathbf{A}^{-1}\mathbf{X}\mathbf{X}'\mathbf{A}\mathbf{X} & \mathbf{I}_{n-k} \end{pmatrix} \begin{pmatrix} \mathbf{I}_k & (\mathbf{X}'\mathbf{A}\mathbf{X})^{-1}\mathbf{X}'\mathbf{A}\mathbf{Y} \\ \mathbf{0} & \mathbf{I}_{n-k} \end{pmatrix},$$

also

$$|\mathbf{W}| = |\mathbf{X}'\mathbf{A}^{-1}\mathbf{X}| \cdot |\mathbf{X}'\mathbf{A}\mathbf{X}|,$$

the left-hand side of (1), and our second solution is complete.

NOTE

1. Excellent solutions have been proposed independently by R.W. Farebrother, by B.D. McCullough, and by H. Neudecker and M. van de Velden, the posers of the problem.

REFERENCE

Harville, David A. (1977) *Matrix Algebra from a Statistician's Perspective*. New York: Springer-Verlag.

98.4.2. *An Alternative Way to Calculate the SUR Estimator*—Solution, proposed by Paolo Paruolo. I give two solutions, the first one based on the associated likelihood function, the second one on explicit matrix calculations.

Solution 1—The SUR estimator is the ML estimator for fixed $\Omega = \hat{\Omega}$, and it is unique. We want to show that steps 1–3 define an alternative way to compute the ML estimator. The likelihood function associated with (1) under (2) can be concentrated with respect to η by performing step 1 (see, e.g., Johansen, 1996, Lemma A.17). The concentrated likelihood as a function of ξ_1 and Ω corresponds to the equations $y_{t.v} = \gamma z_{t.v} + \text{error}$, under the given restrictions $\text{vec}(\gamma) = H_{11}\xi_1$; this likelihood can be maximized for fixed $\Omega = \hat{\Omega}$ by calculating the SUR estimator (4) for ξ_1 as in step 2. The corresponding estimator of η is found substituting $\gamma = \hat{\gamma}$ in equation (1), leading to the regression estimator of η in step 3.

Solution 2—In this solution we exploit standard properties of the vec operator (see, e.g., Magnus and Neudecker, 1988). Rewrite (3) as $\hat{\xi} = A^{-1}B$ and partition $A = (H'(M_{xx} \otimes \hat{\Omega}^{-1})H)$ and $B = (B_1', B_2')' = H'\text{vec}(\hat{\Omega}^{-1}M_{yx})$ conformably with

https://doi.org/10.1017/S0266466699154070 Published online by Cambridge University Press

$H = (H_1, H_2)$; this defines $A_{ij} = H_i'(M_{xx} \otimes \hat{\Omega}^{-1})H_j, i, j = 1,2$. Denote also by A^{ij} the blocks of A^{-1}. Observe that

$$H_1' = H_{11}'((I_m, 0_{m\times u}) \otimes I_p) \qquad H_2' = ((0_{u\times m}, I_u) \otimes I_p); \tag{6}$$

hence $A_{11} = H_{11}'(M_{zz} \otimes \hat{\Omega}^{-1})H_{11}$, $A_{12} = A_{21}' = H_{11}'(M_{zv} \otimes \hat{\Omega}^{-1})$, $A_{22} = (M_{vv} \otimes \hat{\Omega}^{-1})$. By the inversion formula for partitioned matrices one obtains $A^{11} = (H_{11}'(M_{zz.v} \otimes \hat{\Omega}^{-1})H_{11})^{-1}$, $A^{12} = (A^{21})' = -A^{11}C$, $A^{22} = (M_{vv}^{-1} \otimes \hat{\Omega}) + C'A^{11}C$, where $C = H_{11}'(M_{zv}M_{vv}^{-1} \otimes I_p)$. From (6) one also obtains $B_1 = H_{11}'\text{vec}(\hat{\Omega}^{-1}M_{yz})$ and $B_2 = \text{vec}(\hat{\Omega}^{-1}M_{yv})$; note that $B_1 - CB_2 = H_{11}'\text{vec}(\hat{\Omega}^{-1}M_{yz.v})$. Multiplying $A^{-1}B$ in (3) one obtains for the first block

$$\begin{aligned}\hat{\xi}_1 &= A^{11}B_1 + A^{12}B_2 = A^{11}(B_1 - CB_2)\\ &= (H_{11}'(M_{zz.v} \otimes \hat{\Omega}^{-1})H_{11})^{-1}H_{11}'\text{vec}(\hat{\Omega}^{-1}M_{yz.v}).\end{aligned} \tag{7}$$

This expression equals the one in (4). For the second block of rows in (3) one has, recalling that $H_{11}\xi_1 = \text{vec}(\gamma)$ and using (7),

$$\begin{aligned}\hat{\xi}_2 &= A^{21}B_1 + A^{22}B_2 = -C'A^{11}B_1 + ((M_{vv}^{-1} \otimes \hat{\Omega}) + C'A^{11}C)B_2\\ &= (M_{vv}^{-1} \otimes \hat{\Omega})B_2 - C'\hat{\xi}_1 = \text{vec}(M_{yv}M_{vv}^{-1}) - (M_{vv}^{-1}M_{vz} \otimes I_p)H_{11}\hat{\xi}_1\\ &= \text{vec}(M_{yv}M_{vv}^{-1}) - (M_{vv}^{-1}M_{vz} \otimes I_p)\text{vec}(\hat{\gamma}) = \text{vec}((M_{yv} - \hat{\gamma}M_{zv})M_{vv}^{-1}),\end{aligned}$$

which is identical to (5).

REFERENCES

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

Magnus, J.R. & H. Neudecker (1988) *Matrix Differential Calculus with Applications in Statistics and Econometrics*. Chichester: Wiley.

98.4.3. *Regression Specification Error Test as a Gauss–Newton Regression*—Solution,[1] proposed by Giovanni Baiocchi and Walter Distaso. This problem can be addressed in the more general problem of variable addition tests (Pagan and Hall 1983). Whenever we test the general linear model

$$y_t = X_t'\beta + u_t \tag{1}$$

against

$$y_t = X_t'\beta + Z_t'\tau + e_t, \tag{2}$$

with Z' a $T \times l$ matrix and τ an $l \times 1$ vector, we can compute an ordinary F-test for $\tau = \mathbf{0}$ as

$$\frac{(RSSR - USSR)/l}{USSR/(T - k - l)}, \tag{3}$$

https://doi.org/10.1017/S0266466699154070 Published online by Cambridge University Press

where *RSSR* and *USSR* are the sums of squared residuals from (1) and (2), respectively.

If we compute a GNR regression on (2), evaluated at $\hat{\epsilon} \equiv \begin{pmatrix} \hat{\beta} \\ \hat{\tau} \end{pmatrix} = \begin{pmatrix} \hat{\beta} \\ \mathbf{0} \end{pmatrix}$, the estimates under the null hypothesis, we get

$$y_t - X_t' \hat{\beta} = X_t' \mathbf{b} + Z_t' \mathbf{c} + \text{residuals}. \tag{4}$$

Applying the Frisch–Waugh–Lovell (FWL) theorem (Frisch and Waugh, 1933; Lovell, 1963) to both (2) and (4) we can see that the two equations will yield the same SSR as

$$M_{X'} \mathbf{y} \equiv \hat{\mathbf{u}} = M_{X'} Z' \mathbf{c} + \text{residuals}, \tag{5}$$

where $M_{X'}$ is the matrix that projects off the subspace spanned by the columns of X'. Hence, the F-statistics based on (4) will be identical to (3), the F-statistics based on the ordinary regression (2). In the case of interest

$$Z_t' = [(X_t' \hat{\beta})^2, (X_t' \hat{\beta})^3, (X_t' \hat{\beta})^4] \quad \text{and} \quad \tau = \begin{pmatrix} \theta \\ \gamma \\ \lambda \end{pmatrix}.$$

So equation (4) becomes

$$y_t - X_t' \hat{\beta} = X_t' \mathbf{b} + (X_t' \hat{\beta})^2 c_1 + (X_t' \hat{\beta})^3 c_2 + (X_t' \hat{\beta})^4 c_3 + \text{residuals}, \tag{6}$$

and the F-statistics on

$$\mathbf{c} \equiv \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \mathbf{0}$$

are equivalent to the F on $\tau = \mathbf{0}$ from regression

$$y_t = X_t' \beta + (X_t' \hat{\beta})^2 \theta + (X_t' \hat{\beta})^3 \gamma + (X_t' \hat{\beta})^4 \lambda + \text{residuals}. \tag{7}$$

NOTE

1. Excellent solutions have been proposed independently by K.G. Stewart, and by Badi H. Baltagi, the poser of the problem.

REFERENCES

Frisch, R. & F.V. Waugh (1933) Partial time regression as compared with individual trends. *Econometrica* 2, 387–401.

Lovell, M.C. (1963) Seasonal adjustments of economic time series. *Journal of the American Statistical Association* 58, 993–1010.

Pagan, A.R. & A.D. Hall (1983) Diagnostic tests as residual analysis (with discussion). *Econometric Reviews* 2, 159–254.

https://doi.org/10.1017/S0266466699154070 Published online by Cambridge University Press

98.4.4. *Regression with Missing Observations on an Explanatory Variable*—Solution, proposed by Enrique Sentana. First, let us find the maximum likelihood (ML) estimators based on the combined sample. By assumption

$$\begin{pmatrix} y_i \\ x_i \\ z_i \end{pmatrix} \sim \text{i.i.d. } N\left[\begin{pmatrix} \mu_y \\ \mu_x \\ \mu_z \end{pmatrix}, \begin{pmatrix} \sigma_{yy} & \sigma_{yx} & \sigma_{yz} \\ \sigma_{yx} & \sigma_{xx} & \sigma_{xz} \\ \sigma_{yz} & \sigma_{xz} & \sigma_{zz} \end{pmatrix}\right] \equiv N(\boldsymbol{\mu}, \boldsymbol{\Sigma}),$$

so

$$y_i | x_i, z_i \sim N(\beta_0 + \beta_x x_i + \beta_z z_i; \omega^2),$$

where

$$(\beta_x, \beta_z) = (\sigma_{yx}, \sigma_{yz}) \begin{pmatrix} \sigma_{xx} & \sigma_{xz} \\ \sigma_{xz} & \sigma_{zz} \end{pmatrix}^{-1}, \tag{1}$$

$$\beta_0 = \mu_y - \beta_x \mu_x - \beta_z \mu_z, \tag{2}$$

$$\omega^2 = \sigma_{yy} - (\sigma_{yx}, \sigma_{yz}) \begin{pmatrix} \sigma_{xx} & \sigma_{xz} \\ \sigma_{xz} & \sigma_{zz} \end{pmatrix}^{-1} \begin{pmatrix} \sigma_{yx} \\ \sigma_{yz} \end{pmatrix}. \tag{3}$$

Because the reparameterization of the joint distribution into marginal and conditional is one to one, the invariance property of ML estimation implies that if we have ML estimators of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$, say, ML estimators of $(\beta_0, \beta_x, \beta_z)$ and ω^2 can simply be obtained by inserting them in (1)–(3).

The joint likelihood function for the N_1 observations in the first sample is the joint distribution of y_i and x_i. On the other hand, the joint likelihood function for the N_2 observations in the second sample is the joint distribution of y_i, x_i, and z_i, which can be factorized as the conditional distribution of z_i given y_i and x_i times the joint distribution of y_i and x_i. Both components are normal. Specifically,

$$z_i | x_i, y_i \sim N(\gamma_0 + \gamma_x x_i + \gamma_y y_i; \lambda^2)$$

with

$$(\gamma_x, \gamma_y) = (\sigma_{xz}, \sigma_{yz}) \begin{pmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{xy} & \sigma_{yy} \end{pmatrix}^{-1}$$

$$\gamma_0 = \mu_z - \gamma_x \mu_x - \gamma_y \mu_y$$

and

$$\lambda^2 = \sigma_{zz} - (\sigma_{xz}, \sigma_{yz}) \begin{pmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{xy} & \sigma_{yy} \end{pmatrix}^{-1} \begin{pmatrix} \sigma_{xz} \\ \sigma_{yz} \end{pmatrix}.$$

Hence, ML estimates of the parameters in the joint distribution of y_i and x_i, are simply their combined smaple means and covariance matrix (with denominator $N = N_1 + N_2$), whereas ML estimates of $(\gamma_0, \gamma_x, \gamma_y)$ and λ^2 can be obtained from

https://doi.org/10.1017/S0266466699154070 Published online by Cambridge University Press

the ordinary least squares (OLS) regression of z_i on (a constant) x_i and y_i over the second sample. Let us call $(\hat{\gamma}_0, \hat{\gamma}_x, \hat{\gamma}_y)$ and $\hat{\lambda}^2$ such estimators, and $\bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\Sigma}}$ the standard ML estimators of means and variances over the second sample. Then, the invariance property implies that

$$(\hat{\sigma}_{xz}, \hat{\sigma}_{yz}) = (\hat{\gamma}_x, \hat{\gamma}_y)\begin{pmatrix} \hat{\sigma}_{xx} & \hat{\sigma}_{xy} \\ \hat{\sigma}_{xy} & \hat{\sigma}_{yy} \end{pmatrix} = (\bar{\sigma}_{xz}, \bar{\sigma}_{yz})\begin{pmatrix} \bar{\sigma}_{xx} & \bar{\sigma}_{xy} \\ \bar{\sigma}_{xy} & \bar{\sigma}_{yy} \end{pmatrix}^{-1}\begin{pmatrix} \hat{\sigma}_{xx} & \hat{\sigma}_{xy} \\ \hat{\sigma}_{xy} & \hat{\sigma}_{yy} \end{pmatrix},$$

$$\hat{\mu}_z = \hat{\gamma}_0 + \hat{\gamma}_x \hat{\mu}_x + \hat{\gamma}_y \hat{\mu}_y = \bar{\mu}_z + \hat{\gamma}_x(\hat{\mu}_x - \bar{\mu}_x) + \hat{\gamma}_y(\hat{\mu}_y - \bar{\mu}_y)$$

and

$$\begin{aligned} \hat{\sigma}_{zz} &= \hat{\lambda}^2 + (\hat{\sigma}_{xz}, \hat{\sigma}_{yz})\begin{pmatrix} \hat{\sigma}_{xx} & \hat{\sigma}_{xy} \\ \hat{\sigma}_{xy} & \hat{\sigma}_{yy} \end{pmatrix}^{-1}\begin{pmatrix} \hat{\sigma}_{xz} \\ \hat{\sigma}_{yz} \end{pmatrix} \\ &= \bar{\sigma}_{zz} + (\hat{\gamma}_x, \hat{\gamma}_y)\left[\begin{pmatrix} \hat{\sigma}_{xx} & \hat{\sigma}_{xy} \\ \hat{\sigma}_{xy} & \hat{\sigma}_{yy} \end{pmatrix} - \begin{pmatrix} \bar{\sigma}_{xx} & \bar{\sigma}_{xy} \\ \bar{\sigma}_{xy} & \bar{\sigma}_{yy} \end{pmatrix}\right]\begin{pmatrix} \hat{\gamma}_x \\ \hat{\gamma}_y \end{pmatrix}. \end{aligned}$$

Therefore, the ML estimator computed from the pooled sample is using the information in the first sample to "improve" the estimates of μ_z, σ_{yz}, σ_{xz}, and σ_{zz} obtained from the second sample. But if the estimates of μ_x, μ_y, σ_{yy}, σ_{xx}, and σ_{xy} are the same in both samples, no such improvement takes place.

CORRIGENDUM

97.5.3. *A Fundamental Matrix Result on Scaling in Multivariate Analysis*—Solution, proposed by Heinz Neudecker, Albert Satorra, and Michel van de Velden. The following solution was omitted by mistake from the set of published solutions to Problem 97.5.3. The Editor is grateful to the authors for pointing out this mistake.

The singular values of $T = D_r^{-1/2} F D_c^{-1/2}$ are the (positive) square roots of the positive eigenvalues of $T'T$ and hence of TT'.

The eigenvalues are nonnegative as TT' (and $T'T$) are positive semidefinite. Clearly TT' and $T'T$ have the same eigenvalues as $D_c^{-1} F' D_r^{-1} F$. Note that $D_c^{-1} F' D_r^{-1} F 1_p = 1_p$.

From this follows that the matrix has a unit eigenvalue, and hence T has a unit singular value. Further,

$$\| D_c^{-1} F' D_r^{-1} F \| = 1,$$

where $\|Z\|_t = \max_i \sum_{j=1}^{p} |z_{ij}|$ is the row sum norm for any matrix $Z = (z_{ij})$. The nonnegativity of $D_c^{-1} F' D_r^{-1} F$ is of strategic importance.)

As no eigenvalue can exceed (in absolute value) the row sum norm, it follows that all eigenvalues of $D_c^{-1} F' D_r^{-1} F$ are in the (0,1) interval. So are the singular values of $D_r^{-1/2} F D_c^{-1/2}$.

https://doi.org/10.1017/S0266466699154070 Published online by Cambridge University Press