PROBLEMA N. 82

Let $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$ independently distributed with $\boldsymbol{x}_i \sim N_p(\mu_i, \Omega)$, where Ω is nonsingular. It is assumed that $n > p$ and that the matrix $M' = (\mu_1, \dots, \mu_n)$ is of rank 1. Define

$$S = \sum_{i=1}^{n} \boldsymbol{x}_i \boldsymbol{x}_i'$$

Find $E(S^{-1})$ as a second order approximation to the exact solution of Steerneman (1997, 1999), given by

$$E(S^{-1}) = \frac{1}{n-(p+1)}\Omega^{-1} - \frac{1}{n-(p+1)}\left\{\sum_{j=0}^{\infty} \frac{e^{-\frac{1}{2}\lambda}\left(\frac{1}{2}\lambda\right)^j}{j!} \cdot \frac{1}{n+2j}\right\}\Omega^{-1}M'M\Omega^{-1},$$

where $\lambda := \operatorname{tr}\Omega^{-1}M'M$.

(The author invites readers to propose an alternative solution).

References

Steerneman, A.G.M. (1997). «Problem 331», *Statistica. Neerlandica*, 51, 381.

Steerneman, A.G.M. (1999). «Solution 331», *Statistica. Neerlandica*, 53, 252-254.

Heinz Neudecker
Cesaro, Schagen
The Netherlands
heinz@fee.uva.nl

SOLUCIONS ALS PROBLEMES PROPOSATS AL VOLUM 24 N. 1

PROBLEMA N. 82

It is well known that $S \sim W_p(n, \Omega, \Omega^{-1}M'M)$, i.e. S follows a non-central Wishart distribution with scale matrix Ω and non-centrality matrix $\Omega^{-1}M'M$. As M' is of rank 1 we can write $M' = \mu\ell'$, hence $M'M = \lambda\mu\mu'$ where $\lambda := \ell'\ell$.

Define $\widetilde{\Omega} := n\Omega + \lambda\mu\mu'$.

It follows that

$$E(S) = \widetilde{\Omega} \tag{1}$$

$$D(\operatorname{vec} S) = \frac{1}{n}\left(I_{p^2} + K_{pp}\right)\left(\widetilde{\Omega}\otimes\widetilde{\Omega} - \lambda^2\mu\mu'\otimes\mu\mu'\right). \tag{2}$$

See, e.g. Magnus and Neudecker (1979) for these two results and related concepts like the commutation matrix K_{pp}, the vec-operator and the Kronecker product.

We shall now prove the following

Theorem: $E(S^{-1}) \approx \dfrac{n+(p+1)}{n}\widetilde{\Omega}^{-1} - \dfrac{2\lambda^2}{n}\left(\mu'\widetilde{\Omega}^{-1}\mu\right)\widetilde{\Omega}^{-1}\mu\mu'\widetilde{\Omega}^{-1}.$

Proof: Consider the $p \times p$ continuously differentiable symmetric matrix function Z^{-1} and expand this in the point $Z = \widetilde{\Omega}$, with $dZ = S - \widetilde{\Omega}$ and $\Delta Z^{-1} = S^{-1} - \widetilde{\Omega}^{-1}$, viz

$$S^{-1} \approx \widetilde{\Omega}^{-1} - \widetilde{\Omega}^{-1}\left(S - \widetilde{\Omega}\right)\widetilde{\Omega}^{-1} + \widetilde{\Omega}^{-1}\left(S - \widetilde{\Omega}\right)\widetilde{\Omega}^{-1}\left(S - \widetilde{\Omega}\right)\widetilde{\Omega}^{-1}.$$

We used $\Delta Z^{-1} \approx dZ^{-1} + \frac{1}{2}d^2Z^{-1} = -Z^{-1}(dZ)Z^{-1} + Z^{-1}(dZ)Z^{-1}(dZ)Z^{-1}$.

For matrix differentials see Magnus and Neudecker (1999).

Then

$$E(S^{-1}) \approx \widetilde{\Omega}^{-1} + E\left\{\widetilde{\Omega}^{-1}\left(S - \widetilde{\Omega}\right)\widetilde{\Omega}^{-1}\left(S - \widetilde{\Omega}\right)\widetilde{\Omega}^{-1}\right\}, \quad \text{as}$$

$$E(S - \widetilde{\Omega}) = 0.$$

Vectorization of $E\left\{\left(S-\widetilde{\Omega}\right)\widetilde{\Omega}^{-1}\left(S-\widetilde{\Omega}\right)\right\}$ yields

$$\begin{aligned}
&\operatorname{vec}E\left\{\left(S-\widetilde{\Omega}\right)\widetilde{\Omega}^{-1}\left(S-\widetilde{\Omega}\right)\right\}=E\left\{\left(S-\widetilde{\Omega}\right)\otimes\left(S-\widetilde{\Omega}\right)\right\}\operatorname{vec}\widetilde{\Omega}^{-1}\\
&=\operatorname{vec}\left[E\left\{\left(S-\widetilde{\Omega}\right)\otimes\left(S-\widetilde{\Omega}\right)\right\}\operatorname{vec}\widetilde{\Omega}^{-1}\right]\\
&=\left(\operatorname{vec}\widetilde{\Omega}^{-1}\otimes I_{p^2}\right)'E\operatorname{vec}\left\{\left(S-\widetilde{\Omega}\right)\otimes\left(S-\widetilde{\Omega}\right)\right\}\\
&=\left(\operatorname{vec}\widetilde{\Omega}^{-1}\otimes I_{p^2}\right)'C_2^p\operatorname{vec}E\left[\left\{\operatorname{vec}\left(S-\widetilde{\Omega}\right)\right\}\left\{\operatorname{vec}\left(S-\widetilde{\Omega}\right)\right\}'\right]\\
&=\left(\operatorname{vec}\widetilde{\Omega}^{-1}\otimes I_{p^2}\right)'C_2^p\operatorname{vec}D\left\{\operatorname{vec}\left(S-\widetilde{\Omega}\right)\right\}\\
&=\left(\operatorname{vec}\widetilde{\Omega}^{-1}\otimes I_{p^2}\right)'C_2^p\operatorname{vec}D\left(\operatorname{vec}S\right)\\
&=\frac{1}{n}\left(\operatorname{vec}\widetilde{\Omega}^{-1}\otimes I_{p^2}\right)'C_2^p\operatorname{vec}\left(\widetilde{\Omega}\otimes\widetilde{\Omega}-\lambda^2\mu\mu'\otimes\mu\mu'\right)+\\
&+\frac{1}{n}\left(\operatorname{vec}\widetilde{\Omega}^{-1}\otimes I_{p^2}\right)'C_2^pC_3^p\operatorname{vec}\left(\widetilde{\Omega}\otimes\widetilde{\Omega}-\lambda^2\mu\mu'\otimes\mu\mu'\right),
\end{aligned}$$

where $C_2^p := I_p\otimes K_{pp}\otimes I_p$ and $C_3^p := I_{p^2}\otimes K_{pp}$.

See Neudecker and Wansbeek (1987) for these definitions and some applications.

Further

$$\begin{aligned}
&\frac{1}{n}\left(\operatorname{vec}\widetilde{\Omega}^{-1}\otimes I_{p^2}\right)'C_2^p\operatorname{vec}\left(\widetilde{\Omega}\otimes\widetilde{\Omega}-\lambda^2\mu\mu'\otimes\mu\mu'\right)\\
&\quad=\frac{1}{n}\left(\operatorname{vec}\widetilde{\Omega}^{-1}\otimes I_{p^2}\right)'\left(\operatorname{vec}\widetilde{\Omega}\otimes\operatorname{vec}\widetilde{\Omega}-\lambda^2\operatorname{vec}\mu\mu'\otimes\operatorname{vec}\mu\mu'\right)\\
&\quad=\frac{1}{n}\left(\operatorname{tr}\widetilde{\Omega}^{-1}\widetilde{\Omega}\right)\operatorname{vec}\widetilde{\Omega}-\frac{\lambda^2}{n}\left(\mu'\widetilde{\Omega}^{-1}\mu\right)\operatorname{vec}\mu\mu'\\
&\quad=\frac{p}{n}\operatorname{vec}\widetilde{\Omega}-\frac{\lambda^2}{n}\left(\mu'\widetilde{\Omega}^{-1}\mu\right)\operatorname{vec}\mu\mu'
\end{aligned}$$

and

$$\begin{aligned}
&\frac{1}{n}\left(\operatorname{vec}\widetilde{\Omega}^{-1}\otimes I_{p^2}\right)'C_2^pC_3^p\operatorname{vec}\left(\widetilde{\Omega}\otimes\widetilde{\Omega}-\lambda^2\mu\mu'\otimes\mu\mu'\right)\\
&\quad=\frac{1}{n}K_{pp}\left(\operatorname{vec}\widetilde{\Omega}^{-1}\otimes I_{p^2}\right)'C_3^pC_2^pC_3^p\operatorname{vec}\left(\widetilde{\Omega}\otimes\widetilde{\Omega}-\lambda^2\mu\mu'\otimes\mu\mu'\right)
\end{aligned}$$

$$= \frac{1}{n} K_{pp} \left(\text{vec}\,\widetilde{\Omega}^{-1} \otimes I_{p^2} \right)' C_2^p C_3^p C_2^p \text{vec} \left(\widetilde{\Omega} \otimes \widetilde{\Omega} - \lambda^2 \mu\mu' \otimes \mu\mu' \right)$$

$$= \frac{1}{n} K_{pp} \left(\text{vec}\,\widetilde{\Omega}^{-1} \otimes I_{p^2} \right)' C_2^p C_3^p \left(\text{vec}\,\widetilde{\Omega} \otimes \text{vec}\,\widetilde{\Omega} - \lambda^2 \text{vec}\,\mu\mu' \otimes \text{vec}\,\mu\mu' \right)$$

$$= \frac{1}{n} K_{pp} \left(\text{vec}\,\widetilde{\Omega}^{-1} \otimes I_{p^2} \right)' C_2^p \left(\text{vec}\widetilde{\Omega} \otimes \text{vec}\,\widetilde{\Omega} - \lambda^2 \text{vec}\,\mu\mu' \otimes \text{vec}\,\mu\mu' \right)$$

$$= \frac{1}{n} K_{pp} \left(\text{vec}\,\widetilde{\Omega}^{-1} \otimes I_{p^2} \right)' \text{vec} \left(\widetilde{\Omega} \otimes \widetilde{\Omega} - \lambda^2 \mu\mu' \otimes \mu\mu' \right)$$

$$= \frac{1}{n} K_{pp} \,\text{vec} \left\{ \left(\widetilde{\Omega} \otimes \widetilde{\Omega} - \lambda^2 \mu\mu' \otimes \mu\mu' \right) \text{vec}\,\widetilde{\Omega}^{-1} \right\}$$

$$= \frac{1}{n} K_{pp} \left(\widetilde{\Omega} \otimes \widetilde{\Omega} - \lambda^2 \mu\mu' \otimes \mu\mu' \right) \text{vec}\,\widetilde{\Omega}^{-1}$$

$$= \frac{1}{n} \left(\widetilde{\Omega} \otimes \widetilde{\Omega} - \lambda^2 \mu\mu' \otimes \mu\mu' \right) \text{vec}\,\widetilde{\Omega}^{-1}$$

$$= \frac{1}{n} \text{vec}\,\widetilde{\Omega} - \frac{\lambda^2}{n} \left(\mu' \widetilde{\Omega}^{-1} \mu \right) \text{vec}\,\mu\mu'.$$

Use was made of the property

$$C_3^p C_2^p C_3^p = C_2^p C_3^p C_2^p.$$

See Ghazal and Neudecker (1998).

Collecting the two sub-results and de-vectorizing yields

$$E \left(S - \widetilde{\Omega} \right) \widetilde{\Omega}^{-1} \left(S - \widetilde{\Omega} \right) = \frac{p+1}{n} \widetilde{\Omega} - \frac{2\lambda^2}{n} \left(\mu' \widetilde{\Omega}^{-1} \mu \right) \mu\mu'.$$

Hence

$$E \left(S^{-1} \right) \approx \frac{n + (p+1)}{n} \widetilde{\Omega}^{-1} - \frac{2\lambda^2}{n} \left(\mu' \widetilde{\Omega}^{-1} \mu \right) \widetilde{\Omega}^{-1} \mu\mu' \widetilde{\Omega}^{-1}.$$

■

Note

In the **centrality** case ($\mu = 0$, hence $\widetilde{\Omega} = n\Omega$)

we get

$$E \left(S^{-1} \right) \approx \frac{n + (p+1)}{n^2} \Omega^{-1}.$$

The exact result is

$$E\left(S^{-1}\right)=\frac{1}{n-(p+1)}\Omega^{-1}.$$

See, e.g., Giguère and Styan (1978). The exact and approximate expressions are close together when $n \gg p$.

Also note that the exact solution to the problem is

$$E\left(S^{-1}\right)=\frac{1}{n-(p+1)}\Omega^{-1}-\frac{1}{n-(p+1)}\left\{\sum_{j=0}^{\infty} e^{-\frac{1}{2}\lambda}\frac{\left(\frac{1}{2}\lambda\right)^{j}}{j!}\frac{1}{n+2j}\right\}\Omega^{-1}M'M\Omega^{-1}$$

where λ is the non zero eigenvalue of $\Omega^{-\frac{1}{2}}M'M\Omega^{-\frac{1}{2}}$, hence $\lambda=\mathrm{tr}\Omega^{-1}M'M$.

References

Ghazal, A.G. and H. Neudecker (1998). «On second-order and fourth-order moments of jointly distributed random matrices», submitted to *Linear Algebra and Its Applications*.

Giguère, M.A. and G.P.H. Styan (1978). «Multivariate normal estimation with missing data on several variates». *Transactions of the Seventh Prague Conference on Information Theory, Statistical Decision Functions, Random Processes*. Academica, Prague & Reidel, Dordrecht, vol. B, 129-139.

Magnus, J.R. and H. Neudecker (1979). «The commutation matrix: some properties and applications». *Annals of Statistics*, 7, 381-394.

Magnus, J.R. and H. Neudecker (1999). *Matrix Differential Calculus with Applications in Statistics and Econometrics*. Revised Edition, John Wiley & Sons, Chichester.

Neudecker, H. and T.J. Wansbeek (1987). «Fourth-order properties of normally distributed random matrices». *Linear Algebra and Its Applications*, 97, 13-24.

Steerneman, A.G.M. (1997). «Problem 331», *Statistica. Neerlandica*, 51, 381.

Steerneman, A.G.M. (1999). «Solution 331», *Statistica. Neerlandica*, 53, 252-254.

Heinz Neudecker
Cesaro, Schagen
The Netherlands
heinz@fee.uva.nl

PROBLEMA N. 84

1) La función de densidad de X_β, variable X condicionada a $X \leqslant \beta$ es

$$\begin{aligned} f_\beta(x) &= \frac{d}{dx} P(X \leqslant x / X \leqslant \beta) \\ &= \frac{d}{dx} \frac{F(x)}{F(\beta)} \\ &= \frac{f(x)}{F(\beta)} \qquad 0 \leqslant x \leqslant \beta \end{aligned}$$

2) En el problema n. 49 (*Qüestiió*, 17(2), 1993), se prueba que si X es una variable aleatoria con densidad $f(x)$ y

$$\Gamma_\beta = \inf_{0 \leqslant x \leqslant \beta} f(x)$$

entonces se cumple la desigualdad

$$\beta^2 \Gamma_\beta \leqslant \sqrt{12}\sigma$$

siendo σ la desviación típica. Aplicando esta desigualdad a la función $f_\beta(x)$ obtenemos

$$\beta^2 \inf_{0 \leqslant x \leqslant \beta} \frac{f(x)}{F(\beta)} \leqslant \sqrt{12}\sigma(\beta)$$

siendo $\sigma(\beta)$ la desviación típica de X_β. Obsérvese que hay igualdad si y sólo si X es uniforme en $(0, \alpha)$, con $\beta \leqslant \alpha$.

C.M. Cuadras
Universitat de Barcelona

PROBLEMA N. 85

1) Supongamos cierta la hipótesis nula

$$H_0 : (X,Y) \text{ tiene la misma distribución que } (Y,X).$$

Debemos probar que $P(Z \leqslant a) = P(-Z \leqslant a)$, siendo $Z = X - Y$. Tenemos:

$$\begin{aligned} P(X-Y \leqslant a) &= P\big((X,Y) \in \{(x,y)/x-y \leqslant a\}\big) \\ &= P\big((Y,X) \in \{(x,y)/x-y \leqslant a\}\big) \qquad \text{(por } H_0\text{)} \\ &= P\big((X,Y) \in \{(x,y)/y-x \leqslant a\}\big) \qquad \text{(intercambiando } x,y\text{)} \\ &= P(Y-X \leqslant a) \end{aligned}$$

Luego $Z = X - Y$ tiene la misma distribución que $-Z = Y - X$, y la distribución de Z es simétrica respecto del origen.

2) Aceptar la hipótesis

$$H_1 : \text{ la mediana de } Z \text{ es positiva}$$

implica rechazar H_0. En efecto, si H_0 es cierta, entonces

$$P(Z \leqslant 0) = P(Z \geqslant 0) = \frac{1}{2}$$

y la mediana de Z es 0. Luego H_1 es incompatible con H_0.

3) Sea $(X_1,Y_1),\ldots,(X_n,Y_n)$ una muestra aleatoria simple de (X,Y). Si H_0 es cierta

$$P(X-Y < 0) = P(Y-X < 0) = \frac{1}{2},$$

luego el número de veces k tal que

$$X_i - Y_i > 0$$

sigue la distribución binomial $\mathrm{B}\left(n,\frac{1}{2}\right)$. El test de los signos sería un test no paramétrico adecuado para contrastar H_0 frente H_1, pues la distribución de k, bajo H_0, no depende de la distribución de (X,Y). También podríamos utilizar el test del signo-rango de Wilcoxon.

C.M. Cuadras y D. Cuadras
Universitat de Barcelona

PROBLEMA PROPOSAT

PROBLEMA N. 86

Sigui $H(x,y)$ la funció de distribució bivariant del parell de variables aleatòries (X,Y). Siguin $F(x)$, $G(y)$ les funcions de distribució marginals. El coeficient de correlació de Spearman ρ_s és el coeficient de correlació ordinari entre $F(X)$ i $G(Y)$. És ben conegut que

$$\begin{aligned} \rho_s &= \operatorname{corr}\left(F(X),G(Y)\right) \\ &= 12\int_{\mathbb{R}^2} \left(H(x,y)-F(x)\,G(y)\right) dF(x)\,dG(y) \end{aligned}$$

Per tant, si X,Y són estocàsticament independents, aleshores $\rho_s = 0$, és a dir,

$$H(x,y) = F(x)\,G(y), \quad \forall x,y \in \mathbb{R} \quad \Rightarrow \quad \rho_s = 0.$$

Es demana provar que, en general, el recíproc no és cert: X,Y poden ser estocàsticament dependents però $\rho_s = 0$.

C.M. Cuadras
Universitat de Barcelona