Discussiones Mathematicae
Probability and Statistics 25 (2005) 207–215

ESTIMATION OF THE KRONECKER AND INNER PRODUCTS OF TWO MEAN VECTORS IN MULTIVARIATE ANALYSIS

Heinz Neudecker

Department of Econometrics, University of Amsterdam
Roetersstr. 11, NL–1018 WB Amsterdam, The Netherlands

and

Götz Trenkler

Department of Statistics, University of Dortmund
Vogelpothsweg 87, D–44221 Dortmund, Germany

Abstract

In this paper the Kronecker and inner products of mean vectors of two different populations are considered. Using the generalized jackknife approach, estimators for these products are constructed which turn out to be unbiased, provided one can assume multinormal distribution.

Keywords: multivariate normal distribution, Kronecker product, inner product, generalized jackknife, best unbiased estimator.

2000 Mathematics Subject Classification: Primary 62F10, Secondary 62F11.

1. INTRODUCTION

Let $\boldsymbol{y}$ and $\boldsymbol{z}$ be two stochastic vectors with values in $\mathbb{R}^k$ and $\mathbb{R}^l$, respectively, such that $E(\boldsymbol{y}) = \boldsymbol{\nu}$, $E(\boldsymbol{z}) = \boldsymbol{\rho}$ and

$$\mathcal{D}\begin{pmatrix} \boldsymbol{y} \\ \boldsymbol{z} \end{pmatrix} = \boldsymbol{\Sigma} \tag{1.1}$$

exist. Here $E(\cdot)$ and $\mathcal{D}(\cdot)$ denote the mean vector and the dispersion matrix, respectively. Assume that we have independent observations $\boldsymbol{x}_g = (\boldsymbol{y}_g', \boldsymbol{z}_g')', g = 1, \dots, n$, on $\boldsymbol{x} = (\boldsymbol{y}', \boldsymbol{z}')'$. The corresponding $n \times (k+l)$ data matrix will be denoted by $\boldsymbol{X} = (x_{gj})$, $j = 1, \dots, k+l$. Thus $\boldsymbol{X}$ can be written as

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{x}_1' \\ \vdots \\ \boldsymbol{x}_n' \end{pmatrix} = \begin{pmatrix} \boldsymbol{y}_1' & \boldsymbol{z}_1' \\ \vdots & \vdots \\ \boldsymbol{y}_n' & \boldsymbol{z}_n' \end{pmatrix}.$$

In the following we are interested in estimating the Kronecker and inner products of the mean vectors $\boldsymbol{\nu}$ and $\boldsymbol{\rho}$ on the basis of the data matrix $\boldsymbol{X}$.

Let us denote the common mean of $\boldsymbol{y}$ and $\boldsymbol{z}$ by $\boldsymbol{\mu}$, i.e., $\boldsymbol{\mu} = (\boldsymbol{\nu}', \boldsymbol{\rho}')'$. Introducing the matrices $\boldsymbol{J}_1 = (\boldsymbol{I}_k, \boldsymbol{0}_{kl})$ and $\boldsymbol{J}_2 = (\boldsymbol{0}_{lk}, \boldsymbol{I}_l)$ we obtain

$$\boldsymbol{\rho}\boldsymbol{\nu}' = \boldsymbol{J}_2 \boldsymbol{\mu}\boldsymbol{\mu}' \boldsymbol{J}_1'. \tag{1.2}$$

Since $\boldsymbol{\nu} \otimes \boldsymbol{\rho} = \operatorname{vec}(\boldsymbol{\rho}\boldsymbol{\nu}')$ and in case $k = l$, $\boldsymbol{\nu}'\boldsymbol{\rho} = \operatorname{tr}(\boldsymbol{\rho}\boldsymbol{\nu}')$, where $\operatorname{vec}(\cdot)$ and $\operatorname{tr}(\cdot)$ denote the vec– and the trace–operator, respectively, it appears reasonable to develop reliable estimators for $\boldsymbol{\mu}\boldsymbol{\mu}'$ first.

2. ESTIMATION OF $\boldsymbol{\mu}\boldsymbol{\mu}'$

Define the sample mean of the j–th column of the data matrix $\boldsymbol{X}$ as

$$\bar{x}_j = \frac{1}{n} \sum_{g=1}^{n} x_{gj},$$

$j = 1, 2, \ldots, k+l$. The corresponding vector of sample means is

$$\bar{\boldsymbol{x}} = (\bar{x}_1, \ldots, \bar{x}_{k+l})'$$

which alternatively can be expressed as

$$\bar{\boldsymbol{x}} = \frac{1}{n} \sum_{g=1}^{n} \boldsymbol{x}_g = \frac{1}{n} \boldsymbol{X}' \mathbf{1}_n \,, \tag{2.1}$$

where $\mathbf{1}_n$ is the $n \times 1$ vector of ones.

As $E(\bar{\boldsymbol{x}}) = \boldsymbol{\mu}$ and also $E(\boldsymbol{x}_g) = \boldsymbol{\mu}, g = 1, \ldots, n$, candidate estimators for $\boldsymbol{\mu}\boldsymbol{\mu}'$ could be

$$\boldsymbol{R} = \bar{\boldsymbol{x}}\bar{\boldsymbol{x}}' = \frac{1}{n^2} \boldsymbol{X}' \mathbf{1}_n \mathbf{1}_n' \boldsymbol{X}$$

and

$$\boldsymbol{S} = \frac{1}{n} \sum_{g=1}^{n} \boldsymbol{x}_g \boldsymbol{x}_g' = \frac{1}{n} \boldsymbol{X}' \boldsymbol{X} \,.$$

Now we have

$$E(\boldsymbol{R}) = \mathcal{D}(\bar{\boldsymbol{x}}) + E(\bar{\boldsymbol{x}}) E(\bar{\boldsymbol{x}})' = \frac{1}{n} \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}' \tag{2.2}$$

and

$$E(\boldsymbol{S}) = \frac{1}{n} \sum_{g=1}^{n} [\mathcal{D}(\boldsymbol{x}_g) + E(\boldsymbol{x}_g) E(\boldsymbol{x}_g)'] = \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}' \,. \tag{2.3}$$

Hence both estimators of $\boldsymbol{\mu}\boldsymbol{\mu}'$ are biased. The first estimator, $\boldsymbol{R} = \bar{\boldsymbol{x}}\bar{\boldsymbol{x}}'$, however, is asymptotically unbiased.

Using the generalized jackknife procedure (cf. Gray and Schucany, 1972), we obtain as an unbiased estimator for $\boldsymbol{\mu}\boldsymbol{\mu}'$:

$$\begin{aligned}\boldsymbol{T} &= \frac{1}{n-1}\,(n\boldsymbol{R}-\boldsymbol{S})\\ &= \frac{1}{n-1}\left(\frac{1}{n}\,\boldsymbol{X}'\mathbf{1}_n\mathbf{1}_n'\boldsymbol{X}-\frac{1}{n}\,\boldsymbol{X}'\boldsymbol{X}\right)\\ &= \frac{1}{n(n-1)}\,\boldsymbol{X}'(\mathbf{1}_n\mathbf{1}_n'-\boldsymbol{I}_n)\boldsymbol{X}\,.\end{aligned}$$

By using (2.2) and (2.3), we can easily verify that $\boldsymbol{T}$ is unbiased for $\boldsymbol{\mu\mu}'$, indeed.

Let us now assume that $\mathrm{vec}\boldsymbol{X}'$ is multinormally distributed, i.e.,

$$\mathrm{vec}\boldsymbol{X}' \sim N(\mathbf{1}_n\otimes\boldsymbol{\mu},\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}) \tag{2.4}$$

(see Section 1). This can be reexpressed as

$$\mathrm{vec}\boldsymbol{X}' \sim N(\mathrm{vec}\boldsymbol{M}',\boldsymbol{U}\otimes\boldsymbol{V})\,,$$

where $\boldsymbol{M}'=\boldsymbol{\mu}\mathbf{1}_n'$, $\boldsymbol{U}=\boldsymbol{I}_n$ and $\boldsymbol{V}=\boldsymbol{\Sigma}$.

From Ghazal and Neudecker (2000), Corollary 1, we know that the dispersion matrix of $\mathrm{vec}\boldsymbol{S_A}$, where $\boldsymbol{S_A}=\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$ and $\boldsymbol{A}$ is an $n\times n$–matrix, is given by

$$\begin{aligned}\mathcal{D}(\mathrm{vec}\boldsymbol{S}_A) &= (\mathrm{tr}\,\boldsymbol{A}'\boldsymbol{U}\boldsymbol{A}\boldsymbol{U})\boldsymbol{V}\otimes\boldsymbol{V}+\boldsymbol{M}'\boldsymbol{A}'\boldsymbol{U}\boldsymbol{A}\boldsymbol{M}\otimes\boldsymbol{V}\\ &\quad+\boldsymbol{V}\otimes\boldsymbol{M}'\boldsymbol{A}\boldsymbol{U}\boldsymbol{A}'\boldsymbol{M}+\boldsymbol{K}_{pp}\{(\mathrm{tr}\,\boldsymbol{A}\boldsymbol{U}\boldsymbol{A}\boldsymbol{U})\boldsymbol{V}\otimes\boldsymbol{V}\\ &\quad+\boldsymbol{V}\otimes\boldsymbol{M}'\boldsymbol{A}'\boldsymbol{U}\boldsymbol{A}'\boldsymbol{M}+\boldsymbol{M}'\boldsymbol{A}\boldsymbol{U}\boldsymbol{A}\boldsymbol{M}\otimes\boldsymbol{V}\}\,,\end{aligned} \tag{2.5}$$

where $p=k+l$.

Here $\boldsymbol{K}_{pp}$ denotes the commutation matrix of the type $p^2\times p^2$ (cf. Magnus and Neudecker, 1988, Ch. 3).

By letting $\boldsymbol{A}=\frac{1}{n^2}\,\mathbf{1}_n\mathbf{1}_n'$, $\boldsymbol{A}=\frac{1}{n}\,\boldsymbol{I}_n$ and $\boldsymbol{A}=\frac{1}{n(n-1)}\,(\mathbf{1}_n\mathbf{1}_n'-\boldsymbol{I}_n)$ we get the following result.

Theorem 1. *Let* $\text{vec}\boldsymbol{X}'$ *be multinormally distributed as in* (2.4). *Then we have*

(i) $\mathcal{D}(\text{vec}\boldsymbol{R}) = \frac{1}{n}(\boldsymbol{I}_{p^2} + \boldsymbol{K}_{pp})\ (\frac{1}{n}\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma} + \boldsymbol{\Sigma} \otimes \boldsymbol{\mu}\boldsymbol{\mu}' + \boldsymbol{\mu}\boldsymbol{\mu}' \otimes \boldsymbol{\Sigma})$,

(ii) $\mathcal{D}(\text{vec}\boldsymbol{S}) = \frac{1}{n}\ (\boldsymbol{I}_{p^2} + \boldsymbol{K}_{pp})(\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma} + \boldsymbol{\Sigma} \otimes \boldsymbol{\mu}\boldsymbol{\mu}' + \boldsymbol{\mu}\boldsymbol{\mu}' \otimes \boldsymbol{\Sigma})$,

(iii) $\mathcal{D}(\text{vec}\boldsymbol{T}) = \frac{1}{n}(\boldsymbol{I}_{p^2} + \boldsymbol{K}_{pp})\ (\frac{1}{n-1}\ \boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma} + \boldsymbol{\Sigma} \otimes \boldsymbol{\mu}\boldsymbol{\mu}' + \boldsymbol{\mu}\boldsymbol{\mu}' \otimes \boldsymbol{\Sigma})$.

Proof.

(i) Letting $\boldsymbol{A} = \frac{1}{n^2}\ \mathbf{1}_n\mathbf{1}_n'$, from (2.5) it follows that
$\mathcal{D}(\text{vec}\boldsymbol{R}) = \frac{1}{n^2}\ \boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma} + \frac{1}{n}\ \boldsymbol{\mu}\boldsymbol{\mu}' \otimes \boldsymbol{\Sigma} + \frac{1}{n}\ \boldsymbol{\Sigma} \otimes \boldsymbol{\mu}\boldsymbol{\mu}' + \boldsymbol{K}_{pp}\ (\frac{1}{n^2}\ \boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma}$
$+\ \frac{1}{n}\ \boldsymbol{\Sigma} \otimes \boldsymbol{\mu}\boldsymbol{\mu}' + \frac{1}{n}\ \boldsymbol{\mu}\boldsymbol{\mu}' \otimes \boldsymbol{\Sigma})$
which yields (i).

(ii) If $\boldsymbol{A} = \frac{1}{n}\ \boldsymbol{I}_n$, from (2.5) we get
$\mathcal{D}(\text{vec}\boldsymbol{S}) = \frac{1}{n}\ \boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma} + \frac{1}{n}\ \boldsymbol{\mu}\boldsymbol{\mu}' \otimes \boldsymbol{\Sigma} + \frac{1}{n}\ \boldsymbol{\Sigma} \otimes \boldsymbol{\mu}\boldsymbol{\mu}' + \boldsymbol{K}_{pp}\ (\frac{1}{n}\ \boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma}$
$+\ \frac{1}{n}\ \boldsymbol{\mu}\boldsymbol{\mu}' \otimes \boldsymbol{\Sigma} + \boldsymbol{\Sigma} \otimes\ \frac{1}{n}\ \boldsymbol{\mu}\boldsymbol{\mu}')$
which implies (ii).

(iii) Let now $\boldsymbol{A} = \frac{1}{n(n-1)}\ (\mathbf{1}_n\mathbf{1}_n' - \boldsymbol{I}_n)$, in (2.5).
Then
$\mathcal{D}(\text{vec}\boldsymbol{T}) = \frac{1}{n(n-1)}\ \boldsymbol{\Sigma}\otimes\boldsymbol{\Sigma} + \frac{1}{n}\ \boldsymbol{\Sigma}\otimes\boldsymbol{\mu}\boldsymbol{\mu}' + \frac{1}{n}\ \boldsymbol{\mu}\boldsymbol{\mu}'\otimes\boldsymbol{\Sigma} + \boldsymbol{K}_{pp}\ (\frac{1}{n(n-1)}\ \boldsymbol{\Sigma}\otimes\boldsymbol{\Sigma}$
$+\ \frac{1}{n}\boldsymbol{\Sigma} \otimes \boldsymbol{\mu}\boldsymbol{\mu}' + \frac{1}{n}\ \boldsymbol{\mu}\boldsymbol{\mu}' \otimes \boldsymbol{\Sigma})$
so that (iii) follows. ■

3. Estimation of the Kronecker product

Since

$$\boldsymbol{\nu} \otimes \boldsymbol{\rho} = \text{vec}(\boldsymbol{\rho}\boldsymbol{\nu}') = \text{vec}(\boldsymbol{J}_2\boldsymbol{\mu}\boldsymbol{\mu}'\boldsymbol{J}_1') = (\boldsymbol{J}_1 \otimes \boldsymbol{J}_2)\text{vec}\boldsymbol{\mu}\boldsymbol{\mu}'$$

(cf. Magnus and Neudecker, 1988, p. 30), using the results of the preceding section, we have available three estimators of $\boldsymbol{\nu} \otimes \boldsymbol{\rho}$, namely
$\boldsymbol{R}_{\boldsymbol{\nu}\otimes\boldsymbol{\rho}} = (\boldsymbol{J}_1 \otimes \boldsymbol{J}_2)\text{vec}\boldsymbol{R}$, $\boldsymbol{S}_{\boldsymbol{\nu}\otimes\boldsymbol{\rho}} = (\boldsymbol{J}_1 \otimes \boldsymbol{J}_2)\text{vec}\boldsymbol{S}$ and $\boldsymbol{T}_{\boldsymbol{\nu}\otimes\boldsymbol{\rho}} = (\boldsymbol{J}_1 \otimes \boldsymbol{J}_2)\text{vec}\boldsymbol{T}$.

In accordance with (1.1) let us write

$$\mathcal{D}\begin{pmatrix} \boldsymbol{y} \\ \boldsymbol{z} \end{pmatrix} = \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}'_{12} & \boldsymbol{\Sigma}_{22} \end{pmatrix}$$

such that $\boldsymbol{\Sigma}_{11} = \mathcal{D}(\boldsymbol{y}), \boldsymbol{\Sigma}_{22} = \mathcal{D}(\boldsymbol{z})$ and $\boldsymbol{\Sigma}_{12} = \mathrm{Cov}(\boldsymbol{y}, \boldsymbol{z})$. Let us now calculate the expected values of the preceding three estimators of $\boldsymbol{\nu} \otimes \boldsymbol{\rho}$. By (2.2) we get

$$E(\boldsymbol{R}_{\boldsymbol{\nu}\otimes\boldsymbol{\rho}}) = (\boldsymbol{J}_1 \otimes \boldsymbol{J}_2)\mathrm{vec}E(\boldsymbol{R}) = \frac{1}{n}\mathrm{vec}\boldsymbol{\Sigma}'_{12} + \boldsymbol{\nu} \otimes \boldsymbol{\rho}\,.$$

Similarly we get

$$E(\boldsymbol{S}_{\boldsymbol{\nu}\otimes\boldsymbol{\rho}}) = \mathrm{vec}\boldsymbol{\Sigma}'_{12} + \boldsymbol{\nu} \otimes \boldsymbol{\rho}$$

and

$$E(\boldsymbol{T}_{\boldsymbol{\nu}\otimes\boldsymbol{\rho}}) = \boldsymbol{\nu} \otimes \boldsymbol{\rho}\,.$$

Hence $\boldsymbol{T}_{\boldsymbol{\nu}\otimes\boldsymbol{\rho}}$ is an unbiased estimator for $\boldsymbol{\nu} \otimes \boldsymbol{\rho}$. Both $\boldsymbol{R}_{\boldsymbol{\nu}\otimes\boldsymbol{\rho}}$ and $\boldsymbol{S}_{\boldsymbol{\nu}\otimes\boldsymbol{\rho}}$ are biased estimators for $\boldsymbol{\nu} \otimes \boldsymbol{\rho}$, the former being asymptotically unbiased.

Since by Theorem 1 the dispersion matrices $\mathcal{D}(\boldsymbol{R}), \mathcal{D}(\boldsymbol{S})$ and $\mathcal{D}(\boldsymbol{T})$ are known, the dispersion matrices of $\boldsymbol{R}_{\boldsymbol{\nu}\otimes\boldsymbol{\rho}}, \boldsymbol{S}_{\boldsymbol{\nu}\otimes\boldsymbol{\rho}}$ and $\boldsymbol{T}_{\boldsymbol{\nu}\otimes\boldsymbol{\rho}}$ are obtainable from the identities

$$\mathcal{D}(\boldsymbol{R}_{\boldsymbol{\nu}\otimes\boldsymbol{\rho}}) = (\boldsymbol{J}_1 \otimes \boldsymbol{J}_2)\mathcal{D}(\mathrm{vec}\boldsymbol{R})(\boldsymbol{J}'_1 \otimes \boldsymbol{J}'_2),$$

$$\mathcal{D}(\boldsymbol{S}_{\boldsymbol{\nu}\otimes\boldsymbol{\rho}}) = (\boldsymbol{J}_1 \otimes \boldsymbol{J}_2)\mathcal{D}(\mathrm{vec}\boldsymbol{S})(\boldsymbol{J}'_1 \otimes \boldsymbol{J}'_2)$$

and

$$\mathcal{D}(\boldsymbol{T}_{\boldsymbol{\nu}\otimes\boldsymbol{\rho}}) = (\boldsymbol{J}_1 \otimes \boldsymbol{J}_2)\mathcal{D}(\mathrm{vec}\boldsymbol{T})(\boldsymbol{J}'_1 \otimes \boldsymbol{J}'_2).$$

4. ESTIMATION OF THE INNER PRODUCT

Let us assume that the mean vectors $\boldsymbol{\mu}$ and $\boldsymbol{\rho}$ have the same dimension, i.e., $k = l$. Since the inner product $\boldsymbol{\nu}'\boldsymbol{\rho}$ of $\boldsymbol{\nu}$ and $\boldsymbol{\rho}$ equals $\mathrm{tr}(\boldsymbol{\rho}\boldsymbol{\nu}')$, by (1.2) it is reasonable to estimate $\boldsymbol{\nu}'\boldsymbol{\rho}$ by $\mathrm{tr}(\boldsymbol{J}_2\widehat{\boldsymbol{\mu}\boldsymbol{\mu}'}\boldsymbol{J}'_1)$, where $\widehat{\boldsymbol{\mu}\boldsymbol{\mu}'}$ is one of the estimators $\boldsymbol{R}, \boldsymbol{S}$ or $\boldsymbol{T}$. Hence we obtain three estimators

(4.1) $$\boldsymbol{R}_{\boldsymbol{\nu}'\boldsymbol{\rho}} = \mathrm{tr}(\boldsymbol{J}_2\boldsymbol{R}\boldsymbol{J}_1'),$$

(4.2) $$\boldsymbol{S}_{\boldsymbol{\nu}'\boldsymbol{\rho}} = \mathrm{tr}(\boldsymbol{J}_2\boldsymbol{S}\boldsymbol{J}_1')$$

and

(4.3) $$\boldsymbol{T}_{\boldsymbol{\nu}'\boldsymbol{\rho}} = \mathrm{tr}(\boldsymbol{J}_2\boldsymbol{T}\boldsymbol{J}_1')\,.$$

From (2.2), (2.3) and the fact that $\boldsymbol{T}$ is unbiased for $\boldsymbol{\mu}\boldsymbol{\mu}'$ it follows that

$$E(\boldsymbol{R}_{\boldsymbol{\nu}'\boldsymbol{\rho}}) = \frac{1}{n}\,\mathrm{tr}\,\boldsymbol{\Sigma}_{12} + \boldsymbol{\nu}'\boldsymbol{\rho},$$

$$E(\boldsymbol{S}_{\boldsymbol{\nu}'\boldsymbol{\rho}}) = \mathrm{tr}\,\boldsymbol{\Sigma}_{12} + \boldsymbol{\nu}'\boldsymbol{\rho}$$

and

$$E(\boldsymbol{T}_{\boldsymbol{\nu}'\boldsymbol{\rho}}) = \boldsymbol{\nu}'\boldsymbol{\rho}\,.$$

Using the identity $\mathrm{tr}(\boldsymbol{ABC}) = (\mathrm{vec}\boldsymbol{I})'(\boldsymbol{C}' \otimes \boldsymbol{A})\mathrm{vec}\boldsymbol{B}$ (cf. Magnus and Neudecker, 1988, p. 31), the dispersion matrices of the three estimators of $\boldsymbol{\nu}'\boldsymbol{\rho}$ are given by

$$\mathcal{D}(\boldsymbol{R}_{\boldsymbol{\nu}'\boldsymbol{\rho}}) = (\mathrm{vec}\boldsymbol{I})'(\boldsymbol{J}_1 \otimes \boldsymbol{J}_2)\mathcal{D}(\mathrm{vec}\boldsymbol{R})(\boldsymbol{J}_1' \otimes \boldsymbol{J}_2')\mathrm{vec}\boldsymbol{I},$$

$$\mathcal{D}(\boldsymbol{S}_{\boldsymbol{\nu}'\boldsymbol{\rho}}) = (\mathrm{vec}\boldsymbol{I})'(\boldsymbol{J}_1 \otimes \boldsymbol{J}_2)\mathcal{D}(\mathrm{vec}\boldsymbol{S})(\boldsymbol{J}_1' \otimes \boldsymbol{J}_2')\mathrm{vec}\boldsymbol{I},$$

$$\mathcal{D}(\boldsymbol{T}_{\boldsymbol{\nu}'\boldsymbol{\rho}}) = (\mathrm{vec}\boldsymbol{I})'(\boldsymbol{J}_1 \otimes \boldsymbol{J}_2)\mathcal{D}(\mathrm{vec}\boldsymbol{T})(\boldsymbol{J}_1' \otimes \boldsymbol{J}_2')\mathrm{vec}\boldsymbol{I}$$

and since $\mathrm{vec}\boldsymbol{ABC} = (\boldsymbol{C}' \otimes \boldsymbol{A})\mathrm{vec}\boldsymbol{B}$:

$$\mathcal{D}(\boldsymbol{R}_{\boldsymbol{\nu}'\boldsymbol{\rho}}) = \boldsymbol{a}'\mathcal{D}(\mathrm{vec}\boldsymbol{R})\boldsymbol{a},$$

$$\mathcal{D}(\boldsymbol{S}_{\boldsymbol{\nu}'\boldsymbol{\rho}}) = \boldsymbol{a}'\mathcal{D}(\mathrm{vec}\boldsymbol{S})\boldsymbol{a},$$

$$\mathcal{D}(\boldsymbol{T}_{\boldsymbol{\nu}'\boldsymbol{\rho}}) = \boldsymbol{a}'\mathcal{D}(\mathrm{vec}\boldsymbol{T})\boldsymbol{a}\,,$$

where $\boldsymbol{a} = \mathrm{vec}(\boldsymbol{J}_2'\boldsymbol{J}_1)$.

5. Concluding remarks

In Sections 2, 3 and 4 we assumed that $\widetilde{\boldsymbol{x}} := \mathrm{vec}\boldsymbol{X}'$ is multinormally distributed with $E(\widetilde{\boldsymbol{x}}) = \mathbf{1}_n \otimes \boldsymbol{\mu}$ and $\mathcal{D}(\widetilde{\boldsymbol{x}}) = \boldsymbol{I}_n \otimes \boldsymbol{\Sigma}$. Under this assumption it is well-known that the statistics $\bar{\boldsymbol{x}}$ from (2.1) and

$$\begin{aligned}\widehat{\boldsymbol{\Sigma}} &= \frac{1}{n-1}\left[\Sigma_{g=i}^{n}\boldsymbol{x}_g\boldsymbol{x}_g' - n\bar{\boldsymbol{x}}\bar{\boldsymbol{x}}'\right] \\ &= \frac{1}{n-1}\left[\boldsymbol{X}'\boldsymbol{X} - \frac{1}{n}\boldsymbol{X}'\mathbf{1}_n\mathbf{1}_n'\boldsymbol{X}\right] \\ &= \frac{1}{n-1}\boldsymbol{X}'\left(\boldsymbol{I}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n'\right)\boldsymbol{X}\end{aligned}$$

are complete sufficient statistics for $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, cf. Anderson (1984, p. 78).

It is easy to see that

$$\boldsymbol{T} = \bar{\boldsymbol{x}}\bar{\boldsymbol{x}}' - \tfrac{1}{n}\widehat{\boldsymbol{\Sigma}}\ .$$

Hence, since $\boldsymbol{T}$ is unbiased for $\boldsymbol{\mu}\boldsymbol{\mu}'$, $\boldsymbol{T}$ is the best unbiased estimator for $\boldsymbol{\mu}\boldsymbol{\mu}'$, i.e., any other unbiased estimator of $\boldsymbol{\mu}\boldsymbol{\mu}'$ has a dispersion matrix exceeding that of $\boldsymbol{T}$ by a nonnegative definite matrix (see Giri, 1996, Sec. 5.2). In a similar fashion the estimators $\boldsymbol{T}_{\boldsymbol{\nu}\otimes\boldsymbol{\rho}}$ and $\boldsymbol{T}_{\boldsymbol{\nu}'\boldsymbol{\rho}}$ are best.

References

[1] T.W. Anderson, *An Introduction to Multivariate Statistical Analysis*, John Wiley, New York 1984.

[2] A.G. Ghazal and H. Neudecker, *On second-order and fourth-order moments of jointly distributed random matrices. A survey*, Linear Algebra and Its Applications **321** (2000), 63–93.

[3] N.C. Giri, *Multivariate Statistical Analysis*, Marcel Dekker, New York 1996.

[4] H.L. Gray and W.R. Schucany, *The Generalized Jackknife Statistics*, Marcel Dekker, New York 1972.

[5] J.R. Magnus and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*, Wiley, New York 1988.

Received 14 May 2004