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

\[
D \left( \begin{array}{c} y \\ z \end{array} \right) = \Sigma
\]

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

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

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

Let us denote the common mean of \( y \) and \( z \) by \( \mu \), i.e., \( \mu = (\nu', \rho')' \). Introducing the matrices \( J_1 = (I_k, 0_{kl}) \) and \( J_2 = (0_{lk}, I_l) \) we obtain

\[
\rho \nu' = J_2 \mu \mu' J_1'.
\]

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

2. Estimation of \( \mu \mu' \)

Define the sample mean of the \( j \)-th column of the data matrix \( X \) as

\[
\bar{x}_j = \frac{1}{n} \sum_{g=1}^{n} x_{gj},
\]
Estimation of the Kronecker and inner products of two ... 209

\[ j = 1, 2, \ldots, k + l. \]  

The corresponding vector of sample means is

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

which alternatively can be expressed as

\[
(2.1) \quad \bar{x} = \frac{1}{n} \sum_{g=1}^{n} x_g = \frac{1}{n} X' 1_n,
\]

where \(1_n\) is the \(n \times 1\) vector of ones.

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

\[
R = \bar{x}\bar{x}' = \frac{1}{n^2} X' 1_n 1_n' X
\]

and

\[
S = \frac{1}{n} \sum_{g=1}^{n} x_g'x_g = \frac{1}{n} X' X.
\]

Now we have

\[
(2.2) \quad E(R) = D(\bar{x}) + E(\bar{x})E(\bar{x})' = \frac{1}{n} \Sigma + \mu\mu'
\]

and

\[
(2.3) \quad E(S) = \frac{1}{n} \sum_{g=1}^{n} [D(x_g) + E(x_g)E(x_g)'] = \Sigma + \mu\mu'.
\]

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

Using the generalized jackknife procedure (cf. Gray and Schucany, 1972), we obtain as an unbiased estimator for \(\mu\mu'\):
\[ T = \frac{1}{n-1} (nR - S) \]
\[ = \frac{1}{n-1} \left( \frac{1}{n} X'1_n1'_nX - \frac{1}{n} X'X \right) \]
\[ = \frac{1}{n(n-1)} X'(1_n1'_n - I_n)X . \]

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

Let us now assume that \( \text{vec}X' \) is multinormally distributed, i.e.,

\[ (2.4) \quad \text{vec}X' \sim N(1_n \otimes \mu, I_n \otimes \Sigma) \]

(see Section 1). This can be reexpressed as

\[ \text{vec}X' \sim N(\text{vec}M', U \otimes V) , \]

where \( M' = \mu 1'_n \), \( U = I_n \) and \( V = \Sigma \).

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

\[ D(\text{vec}S_A) = (\text{tr} A'AUAU)V \otimes V + M'A'UAM \otimes V + M'AUA'M + K_{pp} \{ (\text{tr} AUAU) V \otimes V \} \]

(2.5)

where \( p = k + l \).

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

By letting \( A = \frac{1}{n^2} 1_n1'_n \), \( A = \frac{1}{n} I_n \) and \( A = \frac{1}{n(n-1)} (1_n1'_n - I_n) \)

we get the following result.
Theorem 1. Let \( \text{vec}X' \) be multinormally distributed as in (2.4). Then we have

(i) \( D(\text{vec}R) = \frac{1}{n}(I_{p^2} + K_{pp}) (\frac{1}{n} \Sigma \otimes \Sigma + \Sigma \otimes \mu\mu' + \mu\mu' \otimes \Sigma), \)

(ii) \( D(\text{vec}S) = \frac{1}{n} (I_{p^2} + K_{pp})(\Sigma \otimes \Sigma + \Sigma \otimes \mu\mu' + \mu\mu' \otimes \Sigma), \)

(iii) \( D(\text{vec}T) = \frac{1}{n}(I_{p^2} + K_{pp}) (\frac{1}{n-1} \Sigma \otimes \Sigma + \Sigma \otimes \mu\mu' + \mu\mu' \otimes \Sigma). \)

Proof.

(i) Letting \( A = \frac{1}{n} 1_n 1_n' \), from (2.5) it follows that
\[
D(\text{vec}R) = \frac{1}{n^2} \Sigma \otimes \Sigma + \frac{1}{n} \mu\mu' \otimes \Sigma + \frac{1}{n} \Sigma \otimes 1_n \mu\mu' + K_{pp} (\frac{1}{n^2} \Sigma \otimes \Sigma + \frac{1}{n} \Sigma \otimes \mu\mu' + \frac{1}{n} \mu\mu' \otimes \Sigma)
\]
which yields (i).

(ii) If \( A = \frac{1}{n} I_n \), from (2.5) we get
\[
D(\text{vec}S) = \frac{1}{n} \Sigma \otimes \Sigma + \frac{1}{n} \mu\mu' \otimes \Sigma + \frac{1}{n} \Sigma \otimes \mu\mu' + K_{pp} (\frac{1}{n} \Sigma \otimes \Sigma + \frac{1}{n} \mu\mu' \otimes \Sigma + \frac{1}{n} \mu\mu' \otimes \Sigma)
\]
which implies (ii).

(iii) Let now \( A = \frac{1}{n(n-1)} (1_n 1_n' - I_n) \), in (2.5).

Then
\[
D(\text{vec}T) = \frac{1}{n(n-1)} \Sigma \otimes \Sigma + \frac{1}{n} \Sigma \otimes \mu\mu' + \frac{1}{n} \mu\mu' \otimes \Sigma + K_{pp} (\frac{1}{n(n-1)} \Sigma \otimes \Sigma + \frac{1}{n} \Sigma \otimes \mu\mu' + \frac{1}{n} \mu\mu' \otimes \Sigma)
\]
so that (iii) follows.

3. Estimation of the Kronecker product

Since
\[
\nu \otimes \rho = \text{vec}(\rho\nu') = \text{vec}(J_2 \mu\mu' J_1') = (J_1 \otimes J_2) \text{vec}\mu\mu'
\]
(cf. Magnus and Neudecker, 1988, p. 30), using the results of the preceding section, we have available three estimators of \( \nu \otimes \rho \), namely
\[
R_{\nu \otimes \rho} = (J_1 \otimes J_2) \text{vec}R, \ S_{\nu \otimes \rho} = (J_1 \otimes J_2) \text{vec}S \quad \text{and} \quad T_{\nu \otimes \rho} = (J_1 \otimes J_2) \text{vec}T.\]
In accordance with (1.1) let us write

\[ D \begin{pmatrix} y \\ z \end{pmatrix} = \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}^\prime & \Sigma_{22} \end{pmatrix} \]

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

\[ E(R_{\nu \otimes \rho}) = (J_1 \otimes J_2)\text{vec}(R) = \frac{1}{n}\text{vec}\Sigma_{12}^\prime + \nu \otimes \rho. \]

Similarly we get

\[ E(S_{\nu \otimes \rho}) = \text{vec}\Sigma_{12}^\prime + \nu \otimes \rho \]

and

\[ E(T_{\nu \otimes \rho}) = \nu \otimes \rho. \]

Hence \( T_{\nu \otimes \rho} \) is an unbiased estimator for \( \nu \otimes \rho \). Both \( R_{\nu \otimes \rho} \) and \( S_{\nu \otimes \rho} \) are biased estimators for \( \nu \otimes \rho \), the former being asymptotically unbiased.

Since by Theorem 1 the dispersion matrices \( D(R) \), \( D(S) \) and \( D(T) \) are known, the dispersion matrices of \( R_{\nu \otimes \rho} \), \( S_{\nu \otimes \rho} \) and \( T_{\nu \otimes \rho} \) are obtainable from the identities

\[ D(R_{\nu \otimes \rho}) = (J_1 \otimes J_2)D(\text{vec}(R))(J_1^\prime \otimes J_2^\prime), \]

\[ D(S_{\nu \otimes \rho}) = (J_1 \otimes J_2)D(\text{vec}(S))(J_1^\prime \otimes J_2^\prime) \]

and

\[ D(T_{\nu \otimes \rho}) = (J_1 \otimes J_2)D(\text{vec}(T))(J_1^\prime \otimes J_2^\prime). \]

4. Estimation of the inner product

Let us assume that the mean vectors \( \mu \) and \( \rho \) have the same dimension, i.e., \( k = l \). Since the inner product \( \nu' \rho \) of \( \nu \) and \( \rho \) equals \( \text{tr}(\rho'\nu) \), by (1.2) it is reasonable to estimate \( \nu' \rho \) by \( \text{tr}(J_2\mu'\nu'J_1) \), where \( \mu' \) is one of the estimators \( R, S \) or \( T \). Hence we obtain three estimators.
Estimation of the Kronecker and inner products of two \ldots

(4.1) \[ R_{\nu' \rho} = \text{tr}(J_2 R J_1'), \]

(4.2) \[ S_{\nu' \rho} = \text{tr}(J_2 S J_1') \]

and

(4.3) \[ T_{\nu' \rho} = \text{tr}(J_2 T J_1'). \]

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

\[ E(\text{tr}(J_2 R J_1')) = \frac{1}{n} \text{tr} \Sigma_{12} + \nu' \rho, \]

\[ E(\text{tr}(J_2 S J_1')) = \text{tr} \Sigma_{12} + \nu' \rho \]

and

\[ E(\text{tr}(J_2 T J_1')) = \nu' \rho. \]

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

\[ D(\text{vec}R_{\nu' \rho}) = (\text{vec}I)'(J_1 \otimes J_2)D(\text{vec}R)(J_1' \otimes J_2')\text{vec}I, \]

\[ D(\text{vec}S_{\nu' \rho}) = (\text{vec}I)'(J_1 \otimes J_2)D(\text{vec}S)(J_1' \otimes J_2')\text{vec}I, \]

\[ D(\text{vec}T_{\nu' \rho}) = (\text{vec}I)'(J_1 \otimes J_2)D(\text{vec}T)(J_1' \otimes J_2')\text{vec}I \]

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

\[ D(\text{vec}R_{\nu' \rho}) = a' D(\text{vec}R)a, \]

\[ D(\text{vec}S_{\nu' \rho}) = a' D(\text{vec}S)a, \]
\[ D(T_{\nu \rho}) = a' D(\text{vec} T) a, \]

where \( a = \text{vec}(I_2 J_1) \).

5. Concluding remarks

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

\[
\hat{\Sigma} = \frac{1}{n-1} \left[ \sum_{g=1}^{n} x_g x_g' - n \bar{x} \bar{x}' \right]
\]

\[
= \frac{1}{n-1} \left[ X'X - \frac{1}{n} X'1_n 1_n' X \right]
\]

\[
= \frac{1}{n-1} X' \left( I_n - \frac{1}{n} 1_n 1_n' \right) X
\]

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

It is easy to see that

\[
T = \bar{x} \bar{x}' - \frac{1}{n} \hat{\Sigma}.
\]

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

References


Estimation of the Kronecker and inner products of two ...


Received 14 May 2004