More on the Kronecker
Structured Covariance Matrix

Martin Ohlson[*]
M. Rauf Ahmad[*]
Dietrich von Rosen[†*]

[*] Department of Mathematics, Linköping University, SE–581 83 Linköping, Sweden.
Corresponding author: Martin Ohlson. Tel.: +46 13 281447. E-mail address: martin.ohlson@liu.se
[†] Energy and Technology, Swedish University of Agricultural Sciences, SE–750 07 Uppsala, Sweden.

More on the Kronecker Structured Covariance Matrix

Martin Ohlson*, M. Rauf Ahmad* and Dietrich von Rosen† *

*Department of Mathematics,
Linköping University,
SE–581 83 Linköping, Sweden.
E-mail: martin.ohlson@liu.se, muahm@mai.liu.se.

†Energy and Technology,
Swedish University of Agricultural Sciences,
SE–750 07 Uppsala, Sweden.
E-mail: dietrich.von.rosen@slu.se.

Sammanfattning

In this paper the multivariate normal distribution with a Kronecker product structured covariance matrix is studied. Particularly, estimation of a Kronecker structured covariance matrix of order three, the so called double separable covariance matrix. The estimation procedure, suggested in this paper, is a generalization of the procedure derived by Srivastava et al. (2008), for a separable covariance matrix.

Furthermore, the restrictions imposed by separability and double separability are discussed.

Keywords: Kronecker product structure, Separable covariance, Double separable covariance, Maximum likelihood estimators.

1 Introduction

In this paper we consider estimation of a Kronecker structured covariance matrix of order three. The main goal is to extend the estimation procedure, suggested by Srivastava et al. (2008), for the matrix normal distribution $\text{vec}\boldsymbol{X} \sim N_{pq}(\text{vec}\boldsymbol{M}, \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma})$ to the case where

$$\text{vec}\mathcal{X} \sim N_{pqr}(\text{vec}\mathcal{M}, \boldsymbol{\Theta} \otimes \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}),$$

with some vectorization $\text{vec}\mathcal{X}$ of the third order tensor $\mathcal{X} = (x_{ijk}) : p \times q \times r$, which will be defined in Section 2 and where $\otimes$ denotes the Kronecker product. We will say that the covariance matrix $D(\text{vec}\mathcal{X}) = \boldsymbol{\Theta} \otimes \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}$ is *double separable* (or three-factor separable) compared to the separable covariance matrix $D(\text{vec}\boldsymbol{X}) = \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}$. The Kronecker product restrictions make the family of densities to be curved, i.e., it belongs to the curved exponential family

Recently Roy and Leiva (2011) have studied *d*oubly exchangeable linear models, which are suitable for three-level multivariate data, and closely related to double separability. Doubly exchangeable covariance structure assumes a block circulant covariance structure consisting of three unstructured covariance matrices for three multivariate levels.

Several authors, see for example Naik and Rao (2001); Roy and Khattree (2005a); Lu and Zimmerman (2005); Mitchell et al. (2005, 2006); Srivastava et al. (2008), considered estimation and testing under the separability assumption. Srivastava et al. (2008) discussed estimability of the paramters under the separability assumption. From the likelihood function, constructed of independent observation matrices, Srivastava et al. (2008) proved that the maximum likelihood estimates under the restriction $\psi_{qq} = 1$, where $\boldsymbol{\Psi} = (\psi_{ij}) : q \times q$ are found by an iterative flip-flop algorithm. Srivastava et al. (2008) also showed that the likelihood equations provide unique estimators. A similar algorithm has been suggested by Mardia and Goodall (1993); Dutilleul (1999); Brown et al. (2001) but without the restriction $\psi_{qq} = 1$.

In many applications, different structures of the covariance matrices have been discussed. In Roy and Khattree (2005a,b); Srivastava et al. (2008) the intraclass covariance structure was considered and in Roy and Khattree (2005b) an autoregressive structure hold.

Also a structure on the mean has been considered. In Srivastava et al. (2009) the growth curve model for the mean $\boldsymbol{M} = \boldsymbol{ABC}$, where $\boldsymbol{A} : p \times s$ and $\boldsymbol{C} : t \times q$ are known design matrices and $\boldsymbol{B} : s \times t$ is the parameter matrix, was assumed. Under the restriction $\psi_{qq} = 1$ and some full rank assaumption unique estimatiors for $\boldsymbol{B}, \boldsymbol{\Sigma}, \boldsymbol{\Psi}$ were derived.

This paper is organized as follows. In Section 2 the normal distribution for the third order tensor $\mathcal{X} = (x_{ijk}) : p \times q \times r$ are presented. One of the main points in Section 2 is how to vectorize the third order tensor and how

to permute this vectorization to present the data in a proper way. In Section 3 the estimation procedure are presented and motivated. Section 4 discusses the restrictions imposed on the matrices $\boldsymbol{\Theta}, \boldsymbol{\Psi}$ and $\boldsymbol{\Sigma}$, similar as $\psi_{qq} = 1$, by the Kronecker product structure.

2 Model

Let $\mathcal{X}$ be a tensor of order three, with the dimension p, q and r in the x, y and z direction, respectively, see Figure 1. If $r = 1$ we have a special case with the tensor equal to a $p \times q$ matrix. For such a matrix $\boldsymbol{X} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_q) : p \times q$ the standard way to vectorize is as

$$\mathrm{vec}\boldsymbol{X} = (\boldsymbol{x}_1', \ldots, \boldsymbol{x}_q')'.$$

$$\mathcal{X} = \begin{pmatrix} x_{11k} & \cdots & x_{1qk} \\ \vdots & \ddots & \vdots \\ x_{p1k} & \cdots & x_{pqk} \end{pmatrix} \qquad k = 1, \ldots, r$$
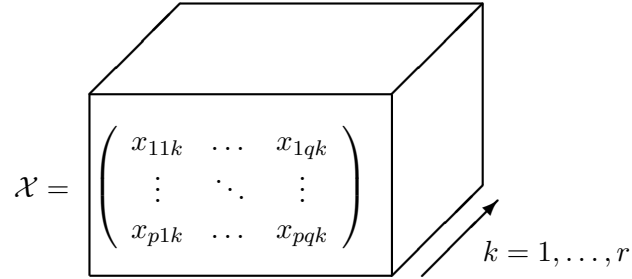
Figur 1: The box visualizes a three dimensional data set as a third order tensor.

Vectorization of the three dimensional tensor $\mathcal{X}$ can be done in several ways. Let us use the following definition.

Definition 1 *Let* $\mathcal{X} = (x_{ijk}) : p \times q \times r$ *be a three dimensional tensor. Define the vectorization of* $\mathcal{X}$ *as*

$$\mathrm{vec}\mathcal{X} = \sum_{i=1}^{p} \sum_{j=1}^{q} \sum_{k=1}^{r} x_{ijk} \boldsymbol{e}_k^3 \otimes \boldsymbol{e}_j^2 \otimes \boldsymbol{e}_i^1,$$

where $\boldsymbol{e}_k^3$, $\boldsymbol{e}_j^2$ *and* $\boldsymbol{e}_i^1$ *are the unit basis vectors of size* r, q *and* p, *respectively.*

We will assume that the vectorization of $\mathcal{X}$ follows a multivariate normal distribution with a double separable covariance matrix

$$D(\mathrm{vec}\mathcal{X}) = \boldsymbol{\Theta} \otimes \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma},$$

with $\boldsymbol{\Sigma} : p \times p$, $\boldsymbol{\Psi} : q \times q$ and $\boldsymbol{\Theta} : r \times r$, assumed to be positive definite. This structure is a generalization of the separable covariance matrix discussed by, e.g., Dutilleul (1999); Lu and Zimmerman (2005); Srivastava et al. (2008).

If we use Definition 2.2.3 in Kollo and von Rosen (2005) we can write the double separable model for $\mathcal{X}$ (or $\text{vec}\mathcal{X}$) as

$$\text{vec}\mathcal{X} = \sum_{ijk} \mu_{ijk} \boldsymbol{e}_k^3 \otimes \boldsymbol{e}_j^2 \otimes \boldsymbol{e}_i^1 + \sum_{ijk} \sum_{i'j'k'} \varsigma_{ii'} \tau_{jj'} \vartheta_{kk'} u_{i'j'k'} \boldsymbol{e}_k^3 \otimes \boldsymbol{e}_j^2 \otimes \boldsymbol{e}_i^1,$$

where $\mathcal{M} = (\mu_{ijk}) : p \times q \times r$, $\boldsymbol{\Sigma} = \boldsymbol{\varsigma}\boldsymbol{\varsigma}'$, $\boldsymbol{\Psi} = \boldsymbol{\tau}\boldsymbol{\tau}'$ and $\boldsymbol{\Theta} = \boldsymbol{\vartheta}\boldsymbol{\vartheta}'$ and $u_{i'j'k'} \sim N(0,1)$, iid (independent and identically distributed). The density of $\mathcal{X}$ can now be written

$$(2\pi)^{-pqr/2} |\boldsymbol{\Theta}|^{-pq/2} |\boldsymbol{\Psi}|^{-pr/2} |\boldsymbol{\Sigma}|^{-qr/2}$$

$$\exp\left\{ -\frac{1}{2} \text{vec}'(\mathcal{X} - \mathcal{M})(\boldsymbol{\Theta} \otimes \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma})^{-1} \text{vec}(\mathcal{X} - \mathcal{M}) \right\}$$

and is denoted

$$\mathcal{X} \sim N_{p,q,r}\left(\mathcal{M}, \boldsymbol{\Sigma}, \boldsymbol{\Psi}, \boldsymbol{\Theta}\right). \tag{1}$$

For more details about the multilinear normal distribution (1) see (Kollo and von Rosen, 2005, p. 215). Furthermore, the tensor in Figure 1 can be looked upon from different directions. To understand this we will use the following matrices

$$\boldsymbol{X} = \sum_{ijk} x_{ijk} (\boldsymbol{e}_k^3 \otimes \boldsymbol{e}_j^2)(\boldsymbol{e}_i^1)' : (qr) \times p, \tag{2}$$

$$\boldsymbol{Y} = \sum_{ijk} x_{ijk} (\boldsymbol{e}_i^1 \otimes \boldsymbol{e}_k^3)(\boldsymbol{e}_j^2)' : (pr) \times q, \tag{3}$$

$$\boldsymbol{Z} = \sum_{ijk} x_{ijk} (\boldsymbol{e}_j^2 \otimes \boldsymbol{e}_i^1)(\boldsymbol{e}_k^3)' : (pq) \times r. \tag{4}$$

Using these matrices and the fact that $\text{vec}(\boldsymbol{ab}') = \boldsymbol{b} \otimes \boldsymbol{a}$, for all $\boldsymbol{a} \in \mathbb{R}^p, \boldsymbol{b} \in \mathbb{R}^q$, we will have the following relations

$$\text{vec}\mathcal{X} = \text{vec}\boldsymbol{Z} = \boldsymbol{K}_{qr,p}\text{vec}\boldsymbol{X} = \boldsymbol{K}_{qr,p}\boldsymbol{K}_{pr,q}\text{vec}\boldsymbol{Y},$$

where $\boldsymbol{K}_{p,q} : pq \times pq$ is the commutation matrix. Using properties of the commutation matrix, yields

$$\text{vec}'\mathcal{X}(\boldsymbol{\Theta} \otimes \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma})^{-1}\text{vec}\mathcal{X}$$

$$= \text{vec}'\boldsymbol{Z}(\boldsymbol{\Theta} \otimes \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma})^{-1}\text{vec}\boldsymbol{Z} = \text{tr}\left\{ \boldsymbol{\Theta}^{-1}\boldsymbol{Z}'(\boldsymbol{\Psi} \otimes \boldsymbol{\Sigma})^{-1}\boldsymbol{Z} \right\}$$

$$= \text{vec}'\boldsymbol{X}(\boldsymbol{\Sigma} \otimes \boldsymbol{\Theta} \otimes \boldsymbol{\Psi})^{-1}\text{vec}\boldsymbol{X} = \text{tr}\left\{ \boldsymbol{\Sigma}^{-1}\boldsymbol{X}'(\boldsymbol{\Theta} \otimes \boldsymbol{\Psi})^{-1}\boldsymbol{X} \right\} \tag{5}$$

$$= \text{vec}'\boldsymbol{Y}(\boldsymbol{\Psi} \otimes \boldsymbol{\Sigma} \otimes \boldsymbol{\Theta})^{-1}\text{vec}\boldsymbol{Y} = \text{tr}\left\{ \boldsymbol{\Psi}^{-1}\boldsymbol{Y}'(\boldsymbol{\Sigma} \otimes \boldsymbol{\Theta})^{-1}\boldsymbol{Y} \right\}. \tag{6}$$

3 Estimation

All the parameters of $\boldsymbol{\Sigma}$, $\boldsymbol{\Psi}$ and $\boldsymbol{\Theta}$ in the covariance matrix $D(\text{vec}\mathcal{X}) = \boldsymbol{\Theta} \otimes \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}$ are not uniquely defined. Several authors have discussed this for a separable covariance matrix $D(\text{vec}\boldsymbol{X}) = \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}$, e.g., Galecki (1994); Naik and Rao (2001). The parametrization problem is related to the fact that

$$\boldsymbol{\Psi} \otimes \boldsymbol{\Sigma} = (c\boldsymbol{\Psi}) \otimes \left(\frac{1}{c}\boldsymbol{\Sigma}\right)$$

and this leads to estimability problems. Recently, Srivastava et al. (2008) also considered the problem and suggested, without any loss of generality, to set $\psi_{qq} = 1$. For a double separable covariance matrix we have a similar problem since

$$\boldsymbol{\Theta} \otimes \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma} = \left(\frac{1}{ab}\boldsymbol{\Theta}\right) \otimes (a\boldsymbol{\Psi}) \otimes (b\boldsymbol{\Sigma}).$$

In this case to get an unique parametrization, without any loss of generality and similar to Srivastava et al. (2008), we suppose $\boldsymbol{\Sigma}$: $p \times p$ to be unstructured, $\boldsymbol{\Psi} = (\psi_{ij})$: $q \times q$ with $\psi_{qq} = 1$ and $\boldsymbol{\Theta} = (\theta_{ij})$: $r \times r$ with $\theta_{rr} = 1$.

Now, assume that we have n independent observations $\mathcal{X}_j : p \times q \times r$, $j = 1, \ldots, n$, from (1). One can easily see that $\mathcal{M} : p \times q \times r$ will be estimated by averaging. Hence, in the subsequent without any loss of generality, we may put $\mathcal{M} = 0$.

Furthermore, with $\mathcal{M} = 0$ the likelihood function for $\boldsymbol{\Sigma}, \boldsymbol{\Psi}$ and $\boldsymbol{\Theta}$ is proportional to

$$|\boldsymbol{\Theta}|^{-pqn/2}|\boldsymbol{\Psi}|^{-prn/2}|\boldsymbol{\Sigma}|^{-qrn/2}\exp\left\{-\frac{1}{2}\sum_{j=1}^{n}\text{vec}'\mathcal{X}_j\left(\boldsymbol{\Theta} \otimes \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}\right)^{-1}\text{vec}\mathcal{X}_j\right\},$$

which can be written as

$$|\boldsymbol{\Theta}|^{-pqn/2}|\boldsymbol{\Psi}|^{-prn/2}|\boldsymbol{\Sigma}|^{-qrn/2}\exp\left\{-\frac{1}{2}\sum_{j=1}^{n}\text{vec}'\boldsymbol{X}_j\left(\boldsymbol{\Sigma} \otimes \boldsymbol{\Theta} \otimes \boldsymbol{\Psi}\right)^{-1}\text{vec}\boldsymbol{X}_j\right\}$$

$$= |\boldsymbol{\Theta}|^{-pqn/2}|\boldsymbol{\Psi}|^{-prn/2}|\boldsymbol{\Sigma}|^{-qrn/2}\exp\left\{-\frac{1}{2}\sum_{j=1}^{n}\text{tr}\left\{\boldsymbol{\Sigma}^{-1}\boldsymbol{X}_j'\left(\boldsymbol{\Theta} \otimes \boldsymbol{\Psi}\right)^{-1}\boldsymbol{X}_j\right\}\right\},$$

$$\tag{7}$$

where we have used relation (5). Now, the trace in (7) can be rewritten as

$$\text{tr}\left\{\boldsymbol{\Sigma}^{-1}\boldsymbol{X}_j'\left(\boldsymbol{\Theta}\otimes\boldsymbol{\Psi}\right)^{-1}\boldsymbol{X}_j\right\}$$

$$=\text{tr}\left\{\boldsymbol{\Sigma}^{-1}\boldsymbol{X}_j'\left(\boldsymbol{I}_r\otimes\boldsymbol{\Psi}^{-1/2}\right)\left(\boldsymbol{\Theta}^{-1}\otimes\boldsymbol{I}_q\right)\left(\boldsymbol{I}_r\otimes\boldsymbol{\Psi}^{-1/2}\right)\boldsymbol{X}_j\right\}$$

$$=\sum_{l=1}^{q}\text{tr}\left\{\boldsymbol{\Sigma}^{-1}\boldsymbol{X}_j'\left(\boldsymbol{I}_r\otimes\left(\boldsymbol{\Psi}^{-1/2}\boldsymbol{e}_l^2\right)\right)\boldsymbol{\Theta}^{-1}\left(\boldsymbol{I}_r\otimes\left(\left(\boldsymbol{e}_l^2\right)'\boldsymbol{\Psi}^{-1/2}\right)\right)\boldsymbol{X}_j\right\}$$

$$=\sum_{l=1}^{q}\text{tr}\{\boldsymbol{\Sigma}^{-1}\boldsymbol{X}_{jl}'\boldsymbol{\Theta}^{-1}\boldsymbol{X}_{jl}\},$$

where $\boldsymbol{X}_{jl}=\left(\boldsymbol{I}_r\otimes\left((\boldsymbol{e}_l^2))'\boldsymbol{\Psi}^{-1/2}\right)\right)\boldsymbol{X}_j$ which implies that the likelihood function is proportional to

$$|\boldsymbol{\Theta}|^{-pqn/2}|\boldsymbol{\Psi}|^{-prn/2}|\boldsymbol{\Sigma}|^{-qrn/2}\exp\left\{-\frac{1}{2}\sum_{j=1}^{n}\sum_{l=1}^{q}\text{tr}\left\{\boldsymbol{\Sigma}^{-1}\boldsymbol{X}_{jl}'\boldsymbol{\Theta}^{-1}\boldsymbol{X}_{jl}\right\}\right\}. \quad (8)$$

Hence, it means that we have nq independent observations, $\boldsymbol{X}_{jl}$ $j=1,\ldots,n$ and $l=1,\ldots,q$. From Srivastava et al. (2008) under the restriction $\theta_{rr}=1$ we obtain the likelihood equations

$$\widehat{\boldsymbol{\Sigma}}=\frac{1}{qrn}\sum_{j=1}^{n}\sum_{l=1}^{q}\boldsymbol{X}_{jl}'\widehat{\boldsymbol{\Theta}}^{-1}\boldsymbol{X}_{jl},$$

$$\widehat{\boldsymbol{\Theta}}=\frac{1}{pqn}\sum_{j=1}^{n}\sum_{l=1}^{q}\boldsymbol{X}_{jl}\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{X}_{jl}',$$

which equal

$$\widehat{\boldsymbol{\Sigma}}=\frac{1}{qrn}\sum_{j=1}^{n}\sum_{l=1}^{q}\boldsymbol{X}_{jl}'\widehat{\boldsymbol{\Theta}}^{-1}\boldsymbol{X}_{jl}$$

$$=\frac{1}{qrn}\sum_{j=1}^{n}\sum_{l=1}^{q}\boldsymbol{X}_j'\left(\boldsymbol{I}_r\otimes\widehat{\boldsymbol{\Psi}}^{-1/2}\boldsymbol{e}_l^2\right)\widehat{\boldsymbol{\Theta}}^{-1}\left(\boldsymbol{I}_r\otimes(\boldsymbol{e}_l^2)'\widehat{\boldsymbol{\Psi}}^{-1/2}\right)\boldsymbol{X}_j$$

$$=\frac{1}{qrn}\sum_{j=1}^{n}\boldsymbol{X}_j'\left(\widehat{\boldsymbol{\Theta}}\otimes\widehat{\boldsymbol{\Psi}}\right)^{-1}\boldsymbol{X}_j \quad (9)$$

and

$$\widehat{\boldsymbol{\Theta}}=\frac{1}{pqn}\sum_{j=1}^{n}\sum_{l=1}^{q}\boldsymbol{X}_{jl}\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{X}_{jl}'$$

$$=\frac{1}{pqn}\sum_{j=1}^{n}\sum_{l=1}^{q}\left(\boldsymbol{I}_r\otimes\left((\boldsymbol{e}_l^2)'\widehat{\boldsymbol{\Psi}}^{-1/2}\right)\right)\boldsymbol{X}_j\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{X}_j'\left(\boldsymbol{I}_r\otimes\left(\widehat{\boldsymbol{\Psi}}^{-1/2}\boldsymbol{e}_l^2\right)\right).$$

$$(10)$$

Using (6), the likelihood function (7) can also be expressed as

$$|\boldsymbol{\Theta}|^{-pqn/2}|\boldsymbol{\Psi}|^{-prn/2}|\boldsymbol{\Sigma}|^{-qrn/2}\exp\left\{-\frac{1}{2}\sum_{j=1}^{n}\mathrm{tr}\left\{\boldsymbol{\Psi}^{-1}\boldsymbol{Y}_j'\left(\boldsymbol{\Sigma}\otimes\boldsymbol{\Theta}\right)^{-1}\boldsymbol{Y}_j\right\}\right\}$$

$$=|\boldsymbol{\Theta}|^{-pqn/2}|\boldsymbol{\Psi}|^{-prn/2}|\boldsymbol{\Sigma}|^{-qrn/2}\exp\left\{-\frac{1}{2}\sum_{j=1}^{n}\sum_{l=1}^{r}\mathrm{tr}\left\{\boldsymbol{\Psi}^{-1}\boldsymbol{Y}_{jl}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Y}_{jl}\right\}\right\}$$

$$=|\boldsymbol{\Theta}|^{-pqn/2}|\boldsymbol{\Psi}|^{-prn/2}|\boldsymbol{\Sigma}|^{-qrn/2}\exp\left\{-\frac{1}{2}\sum_{j=1}^{n}\sum_{l=1}^{r}\mathrm{tr}\left\{\boldsymbol{\Sigma}^{-1}\boldsymbol{Y}_{jl}\boldsymbol{\Psi}^{-1}\boldsymbol{Y}_{jl}'\right\}\right\},$$

$$(11)$$

where $\boldsymbol{Y}_{jl}=\left(\boldsymbol{I}_p\otimes\left(\left(\boldsymbol{e}_l^3\right)'\boldsymbol{\Theta}^{-1/2}\right)\right)\boldsymbol{Y}_j$. Hence, since we have $\psi_{qq}=1$ the likelihood equations follows again from Srivastava et al. (2008) as

$$\widehat{\boldsymbol{\Psi}}=\frac{1}{prn}\sum_{j=1}^{n}\sum_{l=1}^{r}\boldsymbol{Y}_{jl}'\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{Y}_{jl}$$

$$=\frac{1}{prn}\sum_{j=1}^{n}\sum_{l=1}^{r}\boldsymbol{Y}_j'\left(\boldsymbol{I}_p\otimes\left(\widehat{\boldsymbol{\Theta}}^{-1/2}\boldsymbol{e}_l^3\right)\right)\widehat{\boldsymbol{\Sigma}}^{-1}\left(\boldsymbol{I}_p\otimes\left(\left(\boldsymbol{e}_l^3\right)'\widehat{\boldsymbol{\Theta}}^{-1/2}\right)\right)\boldsymbol{Y}_j$$

$$=\frac{1}{prn}\sum_{j=1}^{n}\boldsymbol{Y}_j'\left(\widehat{\boldsymbol{\Sigma}}\otimes\widehat{\boldsymbol{\Theta}}\right)^{-1}\boldsymbol{Y}_j \qquad (12)$$

and

$$\widehat{\boldsymbol{\Sigma}}=\frac{1}{qrn}\sum_{j=1}^{n}\sum_{l=1}^{r}\boldsymbol{Y}_{jl}\widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{Y}_{jl}'$$

$$=\frac{1}{qrn}\sum_{j=1}^{n}\sum_{l=1}^{r}\left(\boldsymbol{I}_p\otimes\left(\left(\boldsymbol{e}_l^3\right)'\widehat{\boldsymbol{\Theta}}^{-1/2}\right)\right)\boldsymbol{Y}_j\widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{Y}_j'\left(\boldsymbol{I}_p\otimes\left(\widehat{\boldsymbol{\Theta}}^{-1/2}\boldsymbol{e}_l^3\right)\right).$$

$$(13)$$

The following theorem can now be stated.

Theorem 1 *The likelihood equations that are maximizing the likelihood function (7) under the conditions $\psi_{qq}=1$ and $\theta_{rr}=1$ are given by*

$$\widehat{\boldsymbol{\Sigma}}=\frac{1}{qrn}\sum_{j=1}^{n}\boldsymbol{X}_j'\left(\widehat{\boldsymbol{\Theta}}\otimes\widehat{\boldsymbol{\Psi}}\right)^{-1}\boldsymbol{X}_j, \qquad (14)$$

$$\widehat{\boldsymbol{\Psi}}=\frac{1}{prn}\sum_{j=1}^{n}\boldsymbol{Y}_j'\left(\widehat{\boldsymbol{\Sigma}}\otimes\widehat{\boldsymbol{\Theta}}\right)^{-1}\boldsymbol{Y}_j, \qquad (15)$$

$$\widehat{\boldsymbol{\Theta}}=\frac{1}{pqn}\sum_{j=1}^{n}\boldsymbol{Z}_j'\left(\widehat{\boldsymbol{\Psi}}\otimes\widehat{\boldsymbol{\Sigma}}\right)^{-1}\boldsymbol{Z}_j. \qquad (16)$$

9

Furthermore, equation (9) *equals equation* (13).

Proof Since we maximize the same likelihood function, (8) and (11), twice with respect to $\boldsymbol{\Sigma}$, equation (9) and (13) must be the same. Let the observations be $\mathcal{X}_d = \left(x_{ijk}^d\right)$, $d = 1, \ldots, n$. Using (3) the expression for $\boldsymbol{Y}_d$ and (2) the expression for $\boldsymbol{X}_d$ in (13) one can show that this is the case, i.e., we have

$$
\begin{aligned}
\widehat{\boldsymbol{\Sigma}} &= \frac{1}{qrn} \sum_{d=1}^{n} \sum_{l=1}^{r} (\boldsymbol{I}_p \otimes ((\boldsymbol{e}_l^3)'\widehat{\boldsymbol{\Theta}}^{-1/2})) \boldsymbol{Y}_d \widehat{\boldsymbol{\Psi}}^{-1} \boldsymbol{Y}_d' (\boldsymbol{I}_p \otimes (\widehat{\boldsymbol{\Theta}}^{-1/2} \boldsymbol{e}_l^3)) \\
&= \frac{1}{qrn} \sum_{d=1}^{n} \sum_{l=1}^{r} \sum_{ijk} \sum_{i'j'k'} x_{ijk}^d x_{i'j'k'}^d (\boldsymbol{I}_p \otimes ((\boldsymbol{e}_l^3)'\widehat{\boldsymbol{\Theta}}^{-1/2})) \\
&\qquad (\boldsymbol{e}_i^1 \otimes \boldsymbol{e}_k^3) \left[(\boldsymbol{e}_j^2)'\widehat{\boldsymbol{\Psi}}^{-1} \boldsymbol{e}_{j'}^2 \right] ((\boldsymbol{e}_{i'}^1)' \otimes (\boldsymbol{e}_{k'}^3)')(\boldsymbol{I}_p \otimes (\widehat{\boldsymbol{\Theta}}^{-1/2} \boldsymbol{e}_l^3)) \\
&= \frac{1}{qrn} \sum_{d=1}^{n} \sum_{ijk} \sum_{i'j'k'} \left\{ (x_{ijk}^d x_{i'j'k'}^d \left[(\boldsymbol{e}_j^2)'\widehat{\boldsymbol{\Psi}}^{-1} \boldsymbol{e}_{j'}^2 \right] \right. \\
&\qquad \left. \sum_{l=1}^{r} \left(\boldsymbol{I}_p \otimes ((\boldsymbol{e}_l^3)'\widehat{\boldsymbol{\Theta}}^{-1/2}) \right) \left(\boldsymbol{e}_i^1(\boldsymbol{e}_{i'}^1)' \otimes \boldsymbol{e}_k^3(\boldsymbol{e}_{k'}^3)' \right) \left(\boldsymbol{I}_p \otimes (\widehat{\boldsymbol{\Theta}}^{-1/2} \boldsymbol{e}_l^3) \right) \right\} \\
&= \frac{1}{qrn} \sum_{d=1}^{n} \sum_{ijk} \sum_{i'j'k'} \left\{ x_{ijk}^d x_{i'j'k'}^d \left[(\boldsymbol{e}_j^2)'\widehat{\boldsymbol{\Psi}}^{-1} \boldsymbol{e}_{j'}^2 \right] \right. \\
&\qquad \left. \sum_{l=1}^{r} \left(\boldsymbol{e}_i^1(\boldsymbol{e}_{i'}^1)' \otimes \left(\left[(\boldsymbol{e}_l^3)'\widehat{\boldsymbol{\Theta}}^{-1/2} \boldsymbol{e}_k^3 \right] \left[(\boldsymbol{e}_{k'}^3)'\widehat{\boldsymbol{\Theta}}^{-1/2} \boldsymbol{e}_l^3 \right] \right) \right) \right\} \\
&= \frac{1}{qrn} \sum_{d=1}^{n} \sum_{ijk} \sum_{i'j'k'} \left\{ x_{ijk}^d x_{i'j'k'}^d \left[(\boldsymbol{e}_j^2)'\widehat{\boldsymbol{\Psi}}^{-1} \boldsymbol{e}_{j'}^2 \right] \right. \\
&\qquad \left. (\boldsymbol{e}_{k'}^3)'\widehat{\boldsymbol{\Theta}}^{-1/2} \left(\sum_{l=1}^{r} \boldsymbol{e}_l^3(\boldsymbol{e}_l^3)' \right) \widehat{\boldsymbol{\Theta}}^{-1/2} \boldsymbol{e}_k^3 \boldsymbol{e}_i^1(\boldsymbol{e}_{i'}^1)' \right\}.
\end{aligned}
$$

But $\sum_{l=1}^{r} \boldsymbol{e}_l^3(\boldsymbol{e}_l^3)' = \boldsymbol{I}_r$, hence

$$
\begin{aligned}
\widehat{\boldsymbol{\Sigma}} &= \frac{1}{qrn} \sum_{d=1}^{n} \sum_{ijk} \sum_{i'j'k'} x_{ijk}^d x_{i'j'k'}^d \left[(\boldsymbol{e}_j^2)'\widehat{\boldsymbol{\Psi}}^{-1} \boldsymbol{e}_{j'}^2 \right] \left[(\boldsymbol{e}_{k'}^3)'\widehat{\boldsymbol{\Theta}}^{-1} \boldsymbol{e}_k^3 \right] \boldsymbol{e}_i^1(\boldsymbol{e}_{i'}^1)' \\
&= \frac{1}{qrn} \sum_{d=1}^{n} \sum_{ijk} \sum_{i'j'k'} x_{ijk}^d x_{i'j'k'}^d ((\boldsymbol{e}_k^3 \otimes \boldsymbol{e}_j^2)(\boldsymbol{e}_i^1)')'(\widehat{\boldsymbol{\Theta}} \otimes \widehat{\boldsymbol{\Psi}})^{-1}(\boldsymbol{e}_{k'}^3 \otimes \boldsymbol{e}_{j'}^2)(\boldsymbol{e}_{i'}^1)' \\
&= \frac{1}{qrn} \sum_{d=1}^{n} \boldsymbol{X}_d'(\widehat{\boldsymbol{\Theta}} \otimes \widehat{\boldsymbol{\Psi}})^{-1} \boldsymbol{X}_d,
\end{aligned}
$$

i.e., equation (9) and (13) are equal. Equation (10) can be rewritten in the same way as above, using (2) and (4) in (10). Hence, we have

$$\widehat{\boldsymbol{\Theta}} = \frac{1}{pqn} \sum_{d=1}^{n} \boldsymbol{Z}_j' \left(\widehat{\boldsymbol{\Psi}} \otimes \widehat{\boldsymbol{\Sigma}} \right)^{-1} \boldsymbol{Z}_j \qquad (17)$$

and the proof is complete. □

Note that the likelihood equations (14)-(16) are nested and there exist no explicit solution. Thus, we can solve (14)-(16) using the so called *flip-flop* algorithm.

4   Restrictions imposed by the Kronecker product

Many statistical hypotheses can be formulated in terms of polynomial equalities and inequalities in the unknown parameters. Hence, under the null hypothesis the parameter space correspond to semi-algebraic subsets of the parameter space. In statistical testing it is important to consider the parameter space under the null hypothesis careful, see for example (Rao, 1973, p. 415-420) for some general classes of large sample tests, or Self and Liang (1987); Drton (2009) for more details when problems can arise.

The double separable covariance matrix, $D(\boldsymbol{X}) = \boldsymbol{\Omega} = \boldsymbol{\Theta} \otimes \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}$, imposes a number of restrictions on the parameter space of the variances and covariances. Hence, the hypothesis

$$H_0 : \boldsymbol{\Omega} = \boldsymbol{\Theta} \otimes \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma} \quad \text{vs.} \quad A : \boldsymbol{\Omega} > \boldsymbol{0},$$

can be written as

$$H_0 : R_i(\boldsymbol{\Omega}) = 0 \text{ for } i = 1, \dots, k \quad \text{vs.} \quad A : \text{not } H_0,$$

where $R_i(\boldsymbol{\Omega})$, $i = 1, \dots, k$ are some functions of the variances and covariances.

The restrictions imposed by separability, are shortly discussed by Lu and Zimmerman (2005) and are given as

$$\frac{\omega_{11}}{\omega_{ip+1,ip+1}} = \frac{\omega_{22}}{\omega_{ip+2,ip+2}} = \cdots = \frac{\omega_{pp}}{\omega_{ip+p,ip+p}}, \quad i = 1, \dots, q-1,$$

$$\rho_{[ii][kl]} = \rho_{[11][kl]}, \quad i = 2, \dots, q; \ k = 1, \dots, p; \ l = k+1, \dots, p,$$

$$\rho_{[ij][kk]} = \rho_{[ij][11]}, \quad i = 1, \dots, q; \ j = i+1, \dots, q; \ k = 2, \dots, p, \qquad (18)$$

$$\rho_{[ij][kl]} = \rho_{[ij][lk]}, \quad i = 1, \dots, q; \ j = i+1, \dots, q; \ l = k+1, \dots, p,$$

$$\rho_{[ij][kl]} = \rho_{[ij][11]}\rho_{[11][kl]}, \quad i = 1, \dots, q; \ j = i+1, \dots, q;$$
$$k = 1, \dots, p; \ l = k+1, \dots, p,$$

where $\rho_{[ij][kl]}$ is the (k, l)th element of (i, j)th $p \times p$ block of the correlation matrix $\boldsymbol{R}$. Since the nature of the functions $R_i(\boldsymbol{\Omega})$, $i = 1, \dots k$ are important we state the following Proposition.

11

Proposition 1 *The functions $R_i(\boldsymbol{\Omega})$, $i = 1, \ldots, k$ imposed by the Kronecker product structure $\boldsymbol{\Omega} = \boldsymbol{\Theta} \otimes \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}$, are smooth functions.*

Proof We will consider the simple cases $p = q = 2$ and $p = q = r = 2$ for a separable and double separable covariance matrix and use these examples to understand the restrictions imposed by the double separability. We start with separability and $p = q = 2$. Consider the Kronecker product

$$\boldsymbol{\Omega} = \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma} = \begin{pmatrix} \psi_{11}\sigma_{11} & \psi_{11}\sigma_{12} & \psi_{12}\sigma_{11} & \psi_{12}\sigma_{12} \\ \cdot & \psi_{11}\sigma_{22} & \psi_{12}\sigma_{21} & \psi_{12}\sigma_{22} \\ \cdot & \cdot & \psi_{22}\sigma_{11} & \psi_{22}\sigma_{12} \\ \cdot & \cdot & \cdot & \psi_{22}\sigma_{22} \end{pmatrix} = (\omega_{ij}),$$

where $\boldsymbol{\Sigma} = (\sigma_{ij})$ and $\boldsymbol{\Psi} = (\psi_{ij})$. We can directly identify one restriction since

$$\frac{\omega_{11}}{\omega_{22}} = \frac{\omega_{33}}{\omega_{44}}. \tag{19}$$

More restrictions can be found from the correlation matrix $\boldsymbol{R}$. The correlation matrix $\boldsymbol{R}$ is nothing else than the Kronecker product of the correlation matrices corresponding to $\boldsymbol{\Psi}$ and $\boldsymbol{\Sigma}$, i.e.,

$$\boldsymbol{R} = \boldsymbol{R_\Psi} \otimes \boldsymbol{R_\Sigma} = \begin{pmatrix} 1 & \rho_{\boldsymbol{\Sigma}} & \rho_{\boldsymbol{\Psi}} & \rho_{\boldsymbol{\Psi}}\rho_{\boldsymbol{\Sigma}} \\ \cdot & 1 & \rho_{\boldsymbol{\Psi}}\rho_{\boldsymbol{\Sigma}} & \rho_{\boldsymbol{\Psi}} \\ \cdot & \cdot & 1 & \rho_{\boldsymbol{\Sigma}} \\ \cdot & \cdot & \cdot & 1 \end{pmatrix} = (\rho_{ij}), \tag{20}$$

where

$$\boldsymbol{R_\Sigma} = \begin{pmatrix} 1 & \rho_{\boldsymbol{\Sigma}} \\ \cdot & 1 \end{pmatrix} \quad \text{and} \quad \boldsymbol{R_\Psi} = \begin{pmatrix} 1 & \rho_{\boldsymbol{\Psi}} \\ \cdot & 1 \end{pmatrix},$$

are the correlation matrices corresponding to $\boldsymbol{\Sigma}$ and $\boldsymbol{\Psi}$, respectively. From (20) we see that we have the following restrictions

$$\rho_{12} = \rho_{34}, \quad \rho_{13} = \rho_{24}, \quad \rho_{14} = \rho_{23} \quad \text{and} \quad \rho_{14} = \rho_{12}\rho_{13}. \tag{21}$$

These restrictions (19) and (21) are of course nothing else than the restrictions (18) given by Lu and Zimmerman (2005). Written in the original covariances ω_{ij}, the restrictions are

$$\omega_{11}\omega_{14} = \omega_{12}\omega_{13}, \quad \omega_{23} = \omega_{14}, \quad \omega_{11}\omega_{24} = \omega_{13}\omega_{22},$$
$$\omega_{11}\omega_{34} = \omega_{12}\omega_{33}, \quad \omega_{11}\omega_{44} = \omega_{22}\omega_{33},$$

i.e., the functions $R_i(\boldsymbol{\Omega})$ for $i = 1, \ldots, 5$ can be formulated as

$$R_1(\boldsymbol{\Omega}) = \omega_{11}\omega_{14} - \omega_{12}\omega_{13}, \quad R_2(\boldsymbol{\Omega}) = \omega_{23} - \omega_{14}, \quad R_3(\boldsymbol{\Omega}) = \omega_{11}\omega_{24} - \omega_{13}\omega_{22},$$
$$R_4(\boldsymbol{\Omega}) = \omega_{11}\omega_{34} - \omega_{12}\omega_{33}, \quad R_5(\boldsymbol{\Omega}) = \omega_{11}\omega_{44} - \omega_{22}\omega_{33},$$

These functions, $R_i(\boldsymbol{\Omega})$ for $i = 1, \ldots, 5$, are smooth, i.e., they have derivatives of all orders.

Similar argument as above can be used when considering the double separable covariance matrix. For the case with three Kronecker products, and with $p = q = r = 2$, we have the following covariance matrix

$$\boldsymbol{\Omega} = \boldsymbol{\Theta} \otimes \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma} = (\omega_{ij}). \tag{22}$$

The covariance matrix (22) directly gives the following restrictions

$$\frac{\omega_{11}}{\omega_{33}} = \frac{\omega_{22}}{\omega_{44}} = \frac{\omega_{55}}{\omega_{77}} = \frac{\omega_{66}}{\omega_{88}} \quad \text{and} \quad \frac{\omega_{11}}{\omega_{55}} = \frac{\omega_{22}}{\omega_{66}}. \tag{23}$$

Furthermore, the correlation matrix

$$\boldsymbol{R} = \boldsymbol{R_\Theta} \otimes \boldsymbol{R_\Psi} \otimes \boldsymbol{R_\Sigma} = (\rho_{ij}),$$

where

$$\boldsymbol{R_\Theta} = \begin{pmatrix} 1 & \rho_{\boldsymbol{\Theta}} \\ \cdot & 1 \end{pmatrix},$$

gives all the other restrictions

$$
\begin{aligned}
\rho_{12} = \rho_{34} = \rho_{56} = \rho_{78}, \quad & \rho_{13} = \rho_{24} = \rho_{57} = \rho_{68}, \\
\rho_{14} = \rho_{23} = \rho_{58} = \rho_{67}, \quad & \rho_{15} = \rho_{26} = \rho_{37} = \rho_{48}, \\
\rho_{16} = \rho_{25} = \rho_{38} = \rho_{47}, \quad & \rho_{17} = \rho_{28} = \rho_{35} = \rho_{46}, \\
\rho_{18} = \rho_{27} = \rho_{36} = \rho_{45} &
\end{aligned}
\tag{24}
$$

and

$$\rho_{14} = \rho_{12}\rho_{13}, \quad \rho_{16} = \rho_{12}\rho_{15}, \quad \rho_{17} = \rho_{13}\rho_{15}, \quad \rho_{18} = \rho_{12}\rho_{13}\rho_{15}. \tag{25}$$

These 29 restrictions, (23), (24) and (25), are similar and a direct generalization of the restrictions imposed by separability (19) and (21). Since they have the same form, the functions $R_i(\boldsymbol{\Omega})$, $i = 1, \ldots, 29$ given by the double separability will also be smooth functions. For general dimensions, p, q and r, the smoothness can be shown using induction. $\qquad\square$

From the Proposition above we specifically have that the functions $R_i(\boldsymbol{\Omega})$, $i = 1, \ldots, k$ have continuous partial derivatives of the first order and this will facilitate the asymptotics, see (Rao, 1973, p. 415-420) for more details.

Under separability the covariance matrix $\boldsymbol{\Omega} = \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}$ has $\frac{1}{2}(p \times (p+1) + q \times (q+1))$ parameters. Under the hypothesis A, the covariance matrix is the unstructured matrix $\boldsymbol{\Omega} : pq \times pq$ which has $\frac{1}{2}pq \times (pq+1)$ parameters. Hence, under separability and $p = q = 2$, the covariance matrix $\boldsymbol{\Omega} = \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}$ has six parameters, and under the alternative the covariance matrix $\boldsymbol{\Omega} > \boldsymbol{0} : 4 \times 4$

has ten parameters. Express the parameters $\omega_{14}, \omega_{23}, \omega_{24}, \omega_{34}$ and ω_{44} as functions of $\omega_{11}, \omega_{12}, \omega_{13}, \omega_{22}$ and ω_{33}, i.e.,

$$\omega_{14} = \frac{\omega_{12}\omega_{13}}{\omega_{11}}, \quad \omega_{23} = \omega_{14}, \quad \omega_{24} = \frac{\omega_{13}\omega_{22}}{\omega_{11}},$$
$$\omega_{34} = \frac{\omega_{12}\omega_{33}}{\omega_{11}}, \quad \omega_{44} = \frac{\omega_{22}\omega_{33}}{\omega_{11}}$$

and of course if we express the five parameters $\omega_{11}, \omega_{12}, \omega_{13}, \omega_{22}$ and ω_{33} in the six parameters $\sigma_{11}, \sigma_{12}, \sigma_{22}, \psi_{11}, \psi_{12}$ and ψ_{22} we have

$$\omega_{11} = \psi_{11}\sigma_{11}, \quad \omega_{12} = \psi_{11}\sigma_{12}, \quad \omega_{13} = \psi_{12}\sigma_{11},$$
$$\omega_{22} = \psi_{11}\sigma_{22}, \quad \omega_{33} = \psi_{22}\sigma_{11}.$$

Hence, we have six parameters but only five equations. This is of course related to the fact that all the parameters of $\mathbf{\Psi}$ and $\mathbf{\Sigma}$ are not defined uniquely

$$\mathbf{\Psi} \otimes \mathbf{\Sigma} = (c\mathbf{\Psi}) \otimes \left(\frac{1}{c}\mathbf{\Sigma}\right)$$

and can be overcome with the restriction $\psi_{22} = 1$.

Furthermore, for the double separablity case with $p = q = r = 2$, we have 29 equations and 36 parameters in the unstructured covariance matrix $\mathbf{\Omega} > \mathbf{0}$, i.e., we have seven free parameters instead of nine as in the Kronecker product $\mathbf{\Omega} = \mathbf{\Theta} \otimes \mathbf{\Psi} \otimes \mathbf{\Sigma}$. Hence, we can again set the restrictions $\psi_{22} = 1$ and $\theta_{22} = 1$.

Referenser

Brown, P. J., Kenward, M. G., and Basset, E. E. (2001). Bayesian discrimination with longitudinal data. *Biostatistics*, 2(4):417–432.

Drton, M. (2009). Likelihood ratio tests and singularities. *The Annals of Statistics*, 37(2):979–1012.

Dutilleul, P. (1999). The MLE algorithm for the matrix normal distribution. *Journal of Statistical Computation and Simulation*, 64(2):105–123.

Galecki, A. T. (1994). General class of covariance structures for two or more repeated factors in longitudinal data analysis. *Communications in Statistics - Theory and Methods*, 23:3105–3119.

Kollo, T. and von Rosen, D. (2005). *Advanced Multivariate Statistics with Matrices*. Springer, Dordrecht, The Netherlands.

Lu, N. and Zimmerman, D. (2005). The likelihood ratio test for a separable covariance matrix. *Statistics and Probability Letters*, 73:449–457.

Mardia, K. V. and Goodall, C. R. (1993). Spatial-temporal analysis of multivariate environmental monitoring data. *Multivariate Environmental Statistics*, 6:347–385.

Mitchell, M. W., Genton, M. G., and Gumpertz, M. L. (2005). Testing for separability of space-time covariances. *Environmetrics*, 16(8):819–831.

Mitchell, M. W., Genton, M. G., and Gumpertz, M. L. (2006). A likelihood ratio test for separability of covariances. *Journal of Multivariate Analysis*, 97(5):1025–1043.

Naik, D. and Rao, S. (2001). Analysis of multivariate repeated measures data with a kronecker product structured covariance matrix. *Journal of Applied Statistics*, 28:91–105.

Rao, C. (1973). *Linear Statistical Inference and Its Applications*. Wiley, New York, USA.

Roy, A. and Khattree, R. (2005a). On implementation of a test for kronecker product covariance structure for multivariate repeated measures data. *Statistical Methodology*, 2:297–306.

Roy, A. and Khattree, R. (2005b). Testing the hypothesis of a Kroneckar product covariance matrix in multivariate repeated measures data. *In Proceedings of the Thirty Annual SAS® Users Group International (SUGI) Conference, Philadelphia, Pennsylvania*, pages 1–11.

Roy, A. and Leiva, R. (2011). Estimating and testing a structured covariance matrix for three-level multivariate data. *To appear in "Communications in Statistics - Theory and Methods"*.

Self, S. G. and Liang, K.-Y. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association*, 82(398):605–610.

Srivastava, M., von Rosen, T., and von Rosen, D. (2008). Models with a Kronecker product covariance structure: estimation and testing. *Mathematical Methods in Statistics*, 17:357–370.

Srivastava, M., von Rosen, T., and von Rosen, D. (2009). Estimation in general multivariate linear models with kronecker product covariance structure. *Sankhyā*, 71, Ser. A(2):137–163.