Estimation in General Multivariate Linear Models with Kronecker Product Covariance Structure

M. S. Srivastava, T. Nahtman, and D. von Rosen

Research Report
Centre of Biostochastics

Swedish University of Agricultural Sciences

Report 2008:1
ISSN 1651-8543

Estimation in General Multivariate Linear Models with Kronecker Product Covariance Structure

Muni S. Srivastava
Department of Statistics, University of Toronto, Canada

Tatjana Nahtman
Institute of Mathematical Statistics, University of Tartu, Estonia; Department of Statistics, Stockholm University, Sweden

Dietrich von Rosen [1]
Centre of Biostochastics, Swedish University of Agricultural Sciences, Sweden

Abstract

In this article models based on pq-dimensional normally distributed random vectors $\boldsymbol{x}$ are studied with a mean $vec(\boldsymbol{ABC})$, where $\boldsymbol{A}$ and $\boldsymbol{C}$ are known matrices, and a separable covariance matrix $\boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}$, where both $\boldsymbol{\Psi}$ and $\boldsymbol{\Sigma}$ are positive definite and except the estimability condition $\psi_{qq} = 1$, unknown The model may among others be applied when spatial-temporal relationships exist. On the basis of n independent observations on the random vector $\boldsymbol{x}$, we wish to estimate the parameters of the model. In the paper estimation equations for obtaining maximum likelihood estimators are presented. It is shown that there exist only one solution to these equations. Likelihood equations are also considered when $\boldsymbol{FBG} = \boldsymbol{0}$, with $\boldsymbol{F}$ and $\boldsymbol{G}$ known. Moreover, the likelihood ratio test for testing $\boldsymbol{FBG} = \boldsymbol{0}$ against $\boldsymbol{FBG} \neq \boldsymbol{0}$ is considered.

Keywords: Growth Curve model, estimation equations, Kronecker product structure, maximum likelihood estimators, separable covariance.

AMS classification: 62F30, 62J10, 62F99.

1 Introduction

Nowadays the complexity in data is increasing. Classical multivariate statistical analysis is usually based on a vector with correlated components, for

[1]Correspondence author: Dietrich.von.Rosen@bt.slu.se

example, $\boldsymbol{x} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, which means that the p-dimensional vector $\boldsymbol{x}$ is multivariate normally distributed with mean vector equal to $\boldsymbol{\mu}$ and dispersion matrix (variance-covariance matrix) equal to $\boldsymbol{\Sigma}$.

The correlation may be due to time dependence, spatial dependence or some underlying latent process which is not observable. However, in many data sets we may have two processes which generate dependency. For example, in environmental sciences when studying catchment areas we have both spatial and temporal correlations, in neurosciences when evaluating fMRI-voxels, where repeated measurements on each voxels are both temporally and spatially correlated, in array technology, where many genes (antigens) are represented on chips (slides) with repeated observations over time, i.e. we have correlations between genes (antigens) and correlation over time.

The main goal of this paper is to extend the classical model, $\boldsymbol{x} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, to $\boldsymbol{x} \sim N_{pq}(\boldsymbol{\mu}, \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma})$, where $\boldsymbol{x} : pq \times 1$, $\boldsymbol{\mu} : pq \times 1$, $\boldsymbol{\Psi} : n \times n$, $\boldsymbol{\Sigma} : p \times p$, and $\otimes$ denotes the Kronecker product.

Both $\boldsymbol{\Psi}$ and $\boldsymbol{\Sigma}$ are unknown but it will be supposed that they are positive definite. Due to the Kronecker product structure, we may convert $\boldsymbol{x} : pq \times 1$ into a matrix $\boldsymbol{X} : p \times q$ which is matrix normally distributed, i.e. $\boldsymbol{X} \sim N_{p,q}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$, where now $\boldsymbol{\mu}$ is a $p \times q$ matrix. Throughout this paper we will have n independent observations, $\boldsymbol{X}_i \sim N_{p,q}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$, whereas in the classical case one has usually only one observation matrix $\boldsymbol{X} \sim N_{p,q}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{I})$. The paper exploits how the independent "matrix-observations" can be used for estimating $\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Psi}$ with certain bilinear structures on $\boldsymbol{\mu}$.

For some related works we refer to Galecki, A.T. (1994), Shitan & Brockwell (1995), Naik & Rao (2001), Roy & Khattree (2005), Lu & Zimmerman (2005) and Srivastava et al. (2007) who all consider the model $\boldsymbol{X} \sim N_{p,q}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$. In comparison to the mentioned works this paper treats more general mean structures, i.e. the mean structure given in Potthoff & Roy (1964) and some of its extensions. Moreover, in this paper we also pay attention to the problem of showing that the MLEs are unique. More general mean structures like those considered in von Rosen (1989), Srivastava (2002) could also have been treated but it is fairly easy to implement these models in the present framework and therefore they will not be considered.

It is interesting to observe that in the classical case of multivariate analysis, i.e. $\boldsymbol{X} \sim N_{p,q}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{I})$, explicit estimators and tests based on MLEs are available for a large class of structures on $\boldsymbol{\mu}$ (e.g. see Andersson & Perlman, 1993, Kollo & von Rosen, 2005, Chapter 4). When generalizing we have iid. observations $\boldsymbol{X}_i \sim N_{p,q}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$ where $\boldsymbol{\Psi}$ is unknown. In this case no explicit MLEs are available. However, it will be shown in this paper that the likelihood equations provide us with only one solution.

The dispersion matrix of a matrix $\boldsymbol{X}_i$ is defined by a vectorized form, i.e. $D[\boldsymbol{X}_i] = D[vec(\boldsymbol{X}_i)]$, where vec is the usual vec-operator. In our models

$$D[\boldsymbol{X}_i] = \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}.$$

For the interpretation it is noted that $\boldsymbol{\Psi} : q \times q$ describes the covariance structure between the columns of $\boldsymbol{X}_i$. The covariance between the columns will up to proportionality be the same for each row of $\boldsymbol{X}_i$. The other covariance matrix $\boldsymbol{\Sigma} : p \times p$ describes the covariance between the rows in $\boldsymbol{X}_i$ which up to proportionality will be the same for each column. The product $\boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}$ takes into consideration both $\boldsymbol{\Psi}$ and $\boldsymbol{\Sigma}$. Indeed, $\boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}$ tells us that the overall covariance consists of the products of the covariances in $\boldsymbol{\Psi}$ and $\boldsymbol{\Sigma}$, respectively, and we have

$$Cov[x_{kl}, x_{rs}] = \sigma_{kr}\psi_{ls}, \tag{1.1}$$

where $\boldsymbol{X}_i = (x_{kl}), \boldsymbol{\Sigma} = (\sigma_{kr})$ and $\boldsymbol{\Psi} = (\psi_{ls})$. Moreover, if we return to our examples in the beginning $\boldsymbol{\Sigma}$ may consist of the time-dependent covariances and $\boldsymbol{\Psi}$ takes care of the spatial correlation, or for the array data $\boldsymbol{\Psi}$ models the dependency between genes (antigens) and $\boldsymbol{\Sigma}$ represents the correlation over time. Note that (1.1) implies that the correlation of x_{kl} and x_{rs} equals the product

$$corr[x_{kl}, x_{rs}] = \frac{\sigma_{kr}}{\sqrt{\sigma_{kk}\sigma_{rr}}} \frac{\psi_{ls}}{\sqrt{\psi_{ll}\psi_{ss}}},$$

As noted above this paper considers more general mean structures than others. It will be assumed that the mean $\boldsymbol{\mu}$ of $\boldsymbol{X}_i$ follows a bilinear model, i.e.

$$E[\boldsymbol{X}_i] = \boldsymbol{ABC}, \tag{1.2}$$

where $\boldsymbol{A} : p \times r$ and $\boldsymbol{C} : s \times q$ are known design matrices. This type of mean structure was introduced by Potthoff & Roy (1964). Under the assumption that $\boldsymbol{\Psi} = \boldsymbol{I}$ (or $\boldsymbol{\Psi}$ known), i.e. we have independent columns in $\boldsymbol{X}_i$ this will give us the well known Growth Curve model. For details and references connected to the model it is referred to Srivastava & Khatri (1979), Srivastava & von Rosen (1999) or Kollo & von Rosen (2005).

Observe that if the matrix $\boldsymbol{X}_i : p \times q$ is $N_{p,q}(\boldsymbol{ABC}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$ distributed we may form a new matrix

$$\boldsymbol{X} = (\boldsymbol{X}_1 : \boldsymbol{X}_2 : \ldots : \boldsymbol{X}_n),$$

which is

$$N_{p,qn}(\boldsymbol{AB}(\mathbf{1}_n' \otimes \boldsymbol{C}), \boldsymbol{\Sigma}, \boldsymbol{I}_n \otimes \boldsymbol{\Psi}), \tag{1.3}$$

where $\mathbf{1}_n$ is a vector of 1s and of size n, and $\boldsymbol{I}_n$ is the identity matrix of size $n \times n$.

The aim of the paper is to present estimating equations for estimating $\boldsymbol{B}, \boldsymbol{\Sigma}$ and $\boldsymbol{\Psi}$ as well as to show how to estimate the parameters when $\boldsymbol{FBG} = \mathbf{0}$ holds for known matrices $\boldsymbol{F}$ and $\boldsymbol{G}$. Moreover, concerning the MLEs we will show that the proposed equations have a unique solution. This is a property these estimators share with estimators of canonical parameters in the exponential family, although the present model belongs to the curved exponential family. Moreover, since parameters can be estimated when $\boldsymbol{FBG} = \mathbf{0}$ some results for testing the hypothesis $H_0 : \ \boldsymbol{FBG} = \mathbf{0}$ against $H_1 : \ \boldsymbol{FBG} \neq \mathbf{0}$ will be presented.

2 MLEs of $\boldsymbol{\Sigma}$ and $\boldsymbol{B}$ when $\boldsymbol{\Psi}$ is known

In this section we briefly consider n iid observation matrices $\boldsymbol{X}_i \sim N_{p,q}(\boldsymbol{ABC}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$, $i = 1, \ldots, n$, $\boldsymbol{A}$: $p \times r$ and $\boldsymbol{C}$: $s \times q$ are known matrices, $\boldsymbol{B}$: $r \times s$ and $\boldsymbol{\Sigma}$: $p \times p$ are unknown parameters and $\boldsymbol{\Psi}$: $q \times q$ is supposed to be known. In many practical problems we may wish to test that $\boldsymbol{\Psi} = \boldsymbol{\Psi}_0$, where $\boldsymbol{\Psi}_0$ is a specified matrix, so that the usual results available for the growth curve model can easily be modified and used, since now we have n iid observation matrices instead of one. For one observation, the MLEs were given by Khatri (1966) and their uniqueness was proved in Srivastava & Khatri (1979, p. 24). Here, the uniqueness of the MLEs of $\boldsymbol{B}$ and $\boldsymbol{\Sigma}$ is considered in the sense that $\boldsymbol{A}\widehat{\boldsymbol{B}}C$ and $\widehat{\boldsymbol{\Sigma}}$ maximizes the likelihood function, where $\widehat{\boldsymbol{B}}$ and $\widehat{\boldsymbol{\Sigma}}$ are the MLEs of $\boldsymbol{B}$ and $\boldsymbol{\Sigma}$, respectively. However, when $\boldsymbol{A}$ and $\boldsymbol{C}$ are not of full rank, one may find several $\widehat{\boldsymbol{B}}$ giving the same value of $\boldsymbol{A}\widehat{\boldsymbol{B}}C$, see Kollo & von Rosen (2005). From practical view point we need only to calculate $\boldsymbol{A}\widehat{\boldsymbol{B}}C$ which is unique for any version of $\widehat{\boldsymbol{B}}$.

The main reason for presenting results for known $\boldsymbol{\Psi}$ is that they will be used later. Since $\boldsymbol{\Psi}$ is positive definite, the data may be transformed, i.e. $\boldsymbol{Y}_i = \boldsymbol{X}_i\boldsymbol{\Psi}^{-1/2}$, where $\boldsymbol{\Psi}^{1/2}$ is a symmetric positive definite square root of $\boldsymbol{\Psi}$. Let $\boldsymbol{Y} = (\boldsymbol{Y}_1 : \boldsymbol{Y}_2 : \ldots : \boldsymbol{Y}_n) : p \times qn$, and then

$$\boldsymbol{Y} \sim N_{p,qn}(\boldsymbol{AB}(\mathbf{1}_n' \otimes \boldsymbol{C\Psi}^{-1/2}), \boldsymbol{\Sigma}, \boldsymbol{I}),$$

where $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ are as given above. From results in Srivastava & Khatri

(1979) or Kollo & von Rosen (2005) it follows directly that

$$\begin{aligned} n\widehat{\boldsymbol{B}} &= (\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{A})^{-}\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{Y}(\boldsymbol{1}_n\otimes\boldsymbol{\Psi}^{-1/2}\boldsymbol{C}'(\boldsymbol{C}\boldsymbol{\Psi}^{-1}\boldsymbol{C}')^{-}) \\ &\quad +(\boldsymbol{A}')^{\circ}\boldsymbol{Z}_1+\boldsymbol{A}'\boldsymbol{Z}_2\boldsymbol{C}^{\circ'}, \qquad (2.1) \\ \boldsymbol{S} &= \boldsymbol{Y}(\boldsymbol{I}-n^{-1}\boldsymbol{1}_n\boldsymbol{1}_n'\otimes\boldsymbol{\Psi}^{-1/2}\boldsymbol{C}'(\boldsymbol{C}\boldsymbol{\Psi}^{-1}\boldsymbol{C}')^{-}\boldsymbol{C}\boldsymbol{\Psi}^{-1/2})\boldsymbol{Y}', \qquad (2.2) \end{aligned}$$

where $\boldsymbol{A}'^{\circ}$ and $\boldsymbol{C}^{\circ}$ are any matrices which generate $\mathcal{C}(\boldsymbol{A}')^{\perp}$ and $\mathcal{C}(\boldsymbol{C})^{\perp}$, i.e. the orthogonal complements of $\mathcal{C}(\boldsymbol{A}')$ and $\mathcal{C}(\boldsymbol{C})$, respectively, and $\mathcal{C}(\cdot)$ denotes the column vector space. In (2.1) and (2.2) $^{-}$ denotes an arbitrary g-inverse, and $\boldsymbol{Z}_1$ and $\boldsymbol{Z}_2$ are arbitrary matrices of proper sizes. Furthermore,

$$\begin{aligned} nq\widehat{\boldsymbol{\Sigma}} &= (\boldsymbol{Y}-\boldsymbol{A}\widehat{\boldsymbol{B}}(\boldsymbol{1}_n'\otimes\boldsymbol{C}\boldsymbol{\Psi}^{-1/2}))(\boldsymbol{Y}-\boldsymbol{A}\widehat{\boldsymbol{B}}(\boldsymbol{1}_n'\otimes\boldsymbol{C}\boldsymbol{\Psi}^{-1/2}))' \\ &= \boldsymbol{S}+n^{-1}\boldsymbol{S}\boldsymbol{A}^{\circ}(\boldsymbol{A}^{\circ'}\boldsymbol{S}\boldsymbol{A}^{\circ})^{-}\boldsymbol{A}^{\circ'} \\ &\quad\times\boldsymbol{Y}(\boldsymbol{1}_n\boldsymbol{1}_n'\otimes\boldsymbol{\Psi}^{-1/2}\boldsymbol{C}(\boldsymbol{C}'\boldsymbol{\Psi}^{-1}\boldsymbol{C})^{-}\boldsymbol{C}'\boldsymbol{\Psi}^{-1/2})\boldsymbol{Y}' \\ &\quad\times\boldsymbol{A}^{\circ}(\boldsymbol{A}^{\circ'}\boldsymbol{S}\boldsymbol{A}^{\circ})^{-}\boldsymbol{A}^{\circ'}\boldsymbol{S} \\ &= \boldsymbol{S}+n^{-1}(\boldsymbol{I}-\boldsymbol{A}(\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{A})^{-}\boldsymbol{A}'\boldsymbol{S}^{-1}) \\ &\quad\times\boldsymbol{Y}(\boldsymbol{1}_n\boldsymbol{1}_n'\otimes\boldsymbol{\Psi}^{-1/2}\boldsymbol{C}(\boldsymbol{C}'\boldsymbol{\Psi}^{-1}\boldsymbol{C})^{-}\boldsymbol{C}'\boldsymbol{\Psi}^{-1/2})\boldsymbol{Y}' \\ &\quad\times(\boldsymbol{I}-\boldsymbol{S}^{-1}\boldsymbol{A}(\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{A})^{-}\boldsymbol{A}'), \end{aligned} \qquad (2.3)$$

where $\boldsymbol{S}$ is given in (2.2). If in (1.2) $\text{rank}(\boldsymbol{A})=r$ and $\text{rank}(\boldsymbol{C})=s$ then $\widehat{\boldsymbol{B}}$ is uniquely estimated, i.e.

$$n\widehat{\boldsymbol{B}}=(\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{Y}(\boldsymbol{1}_n\otimes\boldsymbol{\Psi}^{-1/2}\boldsymbol{C}'(\boldsymbol{C}\boldsymbol{\Psi}^{-1}\boldsymbol{C}')^{-1}).$$

Note that $\widehat{\boldsymbol{\Sigma}}$ is always uniquely estimated.

Turning to the restrictions $\boldsymbol{F}\boldsymbol{B}\boldsymbol{G}=\boldsymbol{0}$ it is observed that these restrictions are equivalent to the relation

$$\boldsymbol{B}=(\boldsymbol{F}')^{\circ}\boldsymbol{\theta}_1+\boldsymbol{F}'\boldsymbol{\theta}_2\boldsymbol{G}^{\circ'},$$

where $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ may be regarded as new unknown parameters. From Theorem 4.1.15 in Kollo & von Rosen (2005) it follows that

$$\widehat{\boldsymbol{B}}=(\boldsymbol{F}')^{\circ}\widehat{\boldsymbol{\theta}}_1+\boldsymbol{F}'\widehat{\boldsymbol{\theta}}_2\boldsymbol{G}^{\circ'},$$

where

$$\begin{aligned} &\widehat{\boldsymbol{\theta}}_2=(\boldsymbol{F}\boldsymbol{A}'\boldsymbol{T}_1'\boldsymbol{S}_2^{-1}\boldsymbol{T}_1\boldsymbol{A}\boldsymbol{F}')^{-}\boldsymbol{F}\boldsymbol{A}'\boldsymbol{T}_1'\boldsymbol{S}_2^{-1}\boldsymbol{T}_1\boldsymbol{Y}(\boldsymbol{1}_n \\ &\otimes\boldsymbol{\Psi}^{-1/2}\boldsymbol{C}'\boldsymbol{G}^{\circ}(\boldsymbol{G}^{\circ'}\boldsymbol{C}\boldsymbol{\Psi}^{-1}\boldsymbol{C}'\boldsymbol{G}^{\circ})^{-})+(\boldsymbol{F}\boldsymbol{A}'\boldsymbol{T}_1')^{\circ}\boldsymbol{Z}_{11}+\boldsymbol{F}\boldsymbol{A}'\boldsymbol{T}_1'\boldsymbol{Z}_{12}(\boldsymbol{G}^{\circ'}\boldsymbol{C})^{\circ'} \end{aligned} \qquad (2.4)$$

with

$$
\begin{aligned}
\boldsymbol{T}_1 &= \boldsymbol{I} - \boldsymbol{A}(\boldsymbol{F'})^{\circ}((\boldsymbol{F'})^{\circ'}\boldsymbol{A}'\boldsymbol{S}_1^{-1}\boldsymbol{A}(\boldsymbol{F'})^{\circ})^{-}(\boldsymbol{F'})^{\circ'}\boldsymbol{A}'\boldsymbol{S}_1^{-1},\\
\boldsymbol{S}_1 &= \boldsymbol{Y}(\boldsymbol{I} - n^{-1}\mathbf{1}_n\mathbf{1}_n' \otimes \boldsymbol{\Psi}^{-1/2}\boldsymbol{C}'(\boldsymbol{C}\boldsymbol{\Psi}^{-1}\boldsymbol{C}')^{-}\boldsymbol{C}\boldsymbol{\Psi}^{-1/2})\boldsymbol{Y}',\\
\boldsymbol{S}_2 &= \boldsymbol{S}_1 + \boldsymbol{T}_1\boldsymbol{Y}(n^{-1}\mathbf{1}_n\mathbf{1}_n' \otimes \boldsymbol{\Psi}^{-1/2}\boldsymbol{C}'(\boldsymbol{C}\boldsymbol{\Psi}^{-1}\boldsymbol{C}')^{-}\boldsymbol{C}\boldsymbol{\Psi}^{-1/2})\\
&\quad \times (\boldsymbol{I} - n^{-1}\mathbf{1}_n\mathbf{1}_n' \otimes \boldsymbol{\Psi}^{-1/2}\boldsymbol{C}'\boldsymbol{G}^{\circ}(\boldsymbol{G}^{\circ'}\boldsymbol{C}\boldsymbol{\Psi}^{-1}\boldsymbol{C}'\boldsymbol{G}^{\circ})^{-}\boldsymbol{G}^{\circ'}\boldsymbol{C}\boldsymbol{\Psi}^{-1/2})\\
&\quad \times (n^{-1}\mathbf{1}_n\mathbf{1}_n' \otimes \boldsymbol{\Psi}^{-1/2}\boldsymbol{C}'(\boldsymbol{C}\boldsymbol{\Psi}^{-1}\boldsymbol{C}')^{-}\boldsymbol{C}\boldsymbol{\Psi}^{-1/2})\boldsymbol{Y}'\boldsymbol{T}_1',\\
\widehat{\boldsymbol{\theta}}_1 &= ((\boldsymbol{F}')^{\circ'}\boldsymbol{A}'\boldsymbol{S}_1^{-1}\boldsymbol{A}(\boldsymbol{F'})^{\circ})^{-}(\boldsymbol{F}')^{\circ'}\boldsymbol{A}'\boldsymbol{S}_1^{-1}\\
&\quad \times (\boldsymbol{Y} - \boldsymbol{A}\boldsymbol{F}'\widehat{\boldsymbol{\theta}}_2\boldsymbol{G}^{\circ'}(\mathbf{1}_n' \otimes \boldsymbol{C}\boldsymbol{\Psi}^{-1/2}))(\boldsymbol{I} \otimes \boldsymbol{\Psi}^{-1/2}\boldsymbol{C}'(\boldsymbol{C}\boldsymbol{\Psi}^{-1}\boldsymbol{C}')^{-})\\
&\quad + ((\boldsymbol{F})'^{\circ'}\boldsymbol{A})^{\circ'}\boldsymbol{Z}_{21} + (\boldsymbol{F}')^{\circ'}\boldsymbol{A}'\boldsymbol{Z}_{22}(\mathbf{1}_n' \otimes \boldsymbol{C})^{\circ'},
\end{aligned}
\tag{2.5}
$$

where $\boldsymbol{S}_1$ is assumed to be positive definite and $\boldsymbol{Z}_{ij}$ are arbitrary matrices. Furthermore,

$$
\begin{aligned}
nq\widehat{\boldsymbol{\Sigma}} &= (\boldsymbol{Y} - \boldsymbol{A}\widehat{\boldsymbol{B}}(\mathbf{1}_n' \otimes \boldsymbol{C}\boldsymbol{\Psi}^{-1/2}))(\boldsymbol{Y} - \boldsymbol{A}\widehat{\boldsymbol{B}}(\mathbf{1}_n' \otimes \boldsymbol{C}\boldsymbol{\Psi}^{-1/2}))'\\
&= \boldsymbol{S}_1 + \boldsymbol{S}_2 + n^{-1}\boldsymbol{T}_2\boldsymbol{T}_1\boldsymbol{Y}\big(\mathbf{1}_n\mathbf{1}_n'\\
&\qquad \otimes \boldsymbol{\Psi}^{-1/2}\boldsymbol{C}'\boldsymbol{G}^{o}(\boldsymbol{G}^{o'}\boldsymbol{C}\boldsymbol{\Psi}^{-1}\boldsymbol{C}'\boldsymbol{G}^{o})^{-}G^{o'}\boldsymbol{C}\boldsymbol{\Psi}^{-1/2}\big)\boldsymbol{Y}'\boldsymbol{T}_1'\boldsymbol{T}_2',
\end{aligned}
\tag{2.6}
$$

where

$$
\boldsymbol{T}_2 = \boldsymbol{I} - \boldsymbol{T}_1\boldsymbol{A}\boldsymbol{D}'(\boldsymbol{D}\boldsymbol{A}'\boldsymbol{T}_1'(\boldsymbol{S}_1 + \boldsymbol{S}_2)^{-1}\boldsymbol{T}_1\boldsymbol{A}\boldsymbol{D}')^{-}\boldsymbol{D}\boldsymbol{A}'\boldsymbol{T}_1'(\boldsymbol{S}_1 + \boldsymbol{S}_2)^{-1}.
$$

3 Explicit estimators when diag($\boldsymbol{\Psi}$) is known

In this section we present some easily obtained explicit estimators under the condition that diag($\boldsymbol{\Psi}$) is known. Among others they can be used as starting values when solving the likelihood equations presented in the next section.

If $\boldsymbol{X}_i \sim N_{p,q}(\boldsymbol{ABC}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$ it follows that

$$
D[\boldsymbol{X}_i] = \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}.
$$

Since $c\boldsymbol{\Psi} \otimes c^{-1}\boldsymbol{\Sigma} = \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}$, we are in a situation when the parameterization can not be uniquely interpreted, i.e. the model is overparameterized. Therefore, if $\boldsymbol{\Psi}$ is unknown we will always suppose that $\psi_{qq} = 1$. However, if we assume all diagonal elements in $\boldsymbol{\Psi}$ to equal 1, it is easy to produce some heuristic estimators. The main idea is to produce estimators of $\boldsymbol{B}$ and $\boldsymbol{\Sigma}$ by neglecting the dependency among columns. Thereafter the off-diagonal elements in $\boldsymbol{\Psi}$ are estimated.

Theorem 3.1. *Let* $\boldsymbol{X} \sim N_{p,qn}(\boldsymbol{AB}(\boldsymbol{1}_n' \otimes \boldsymbol{C}), \boldsymbol{\Sigma}, \boldsymbol{I}_n \otimes \boldsymbol{\Psi})$*, where diag(*$\boldsymbol{\Psi}$*)=*$\boldsymbol{I}$*. Unbiased estimators of* $\boldsymbol{ABC}$ *and* $\boldsymbol{\Sigma}$ *are given by*

$$\begin{aligned} \boldsymbol{A}\widehat{\boldsymbol{B}}\boldsymbol{C} &= n^{-1}\boldsymbol{A}(\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{A})^{-}\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{X}(\boldsymbol{1}_n \otimes \boldsymbol{C}'(\boldsymbol{CC}')^{-}\boldsymbol{C}), \\ \widehat{\boldsymbol{\Sigma}} &= (q(n-1))^{-1}\boldsymbol{S} = (q(n-1))^{-1}\boldsymbol{X}(\boldsymbol{1}_n^{\circ}(\boldsymbol{1}_n^{\circ\prime}\boldsymbol{1}_n^{\circ})^{-}\boldsymbol{1}_n^{\circ\prime} \otimes \boldsymbol{I})\boldsymbol{X}'. \end{aligned}$$

Proof. Since $\boldsymbol{1}_n'\boldsymbol{1}_n^{\circ} = \boldsymbol{0}$, $\boldsymbol{X}(\boldsymbol{1}_n \otimes \boldsymbol{C}')$ and $\boldsymbol{X}(\boldsymbol{1}_n^{\circ} \otimes \boldsymbol{I})$ are independent. Thus,

$$\begin{aligned} E[\boldsymbol{A}\widehat{\boldsymbol{B}}\boldsymbol{C}] &= n^{-1}E[\boldsymbol{A}(\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{A})^{-}\boldsymbol{A}'\boldsymbol{S}^{-1}]E[\boldsymbol{X}(\boldsymbol{1}_n \otimes \boldsymbol{C}'(\boldsymbol{CC}')^{-}\boldsymbol{C})] \\ &= n^{-1}E[\boldsymbol{A}(\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{A})^{-}\boldsymbol{A}'\boldsymbol{S}^{-1}]\boldsymbol{AB}(\boldsymbol{1}_n' \otimes \boldsymbol{C})(\boldsymbol{1}_n \otimes \boldsymbol{C}'(\boldsymbol{CC}')^{-}\boldsymbol{C}) \\ &= \boldsymbol{ABC}. \end{aligned}$$

Moreover, by using that if $\boldsymbol{Y} \sim N_{p,q}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$, then for any $\boldsymbol{A}$ of proper size,

$$E[\boldsymbol{YAY'}] = tr(\boldsymbol{A\Psi})\boldsymbol{\Sigma} + \boldsymbol{\mu A \mu'},$$

and since $\boldsymbol{AB}(\boldsymbol{1}_n' \otimes \boldsymbol{C})(\boldsymbol{1}_n^{\circ} \otimes \boldsymbol{I}) = \boldsymbol{0}$,

$$\begin{aligned} E[\boldsymbol{S}] &= tr((\boldsymbol{1}_n^{\circ}(\boldsymbol{1}_n^{\circ\prime}\boldsymbol{1}_n^{\circ})^{-}\boldsymbol{1}_n^{\circ\prime} \otimes \boldsymbol{I})(\boldsymbol{I} \otimes \boldsymbol{\Psi}))\boldsymbol{\Sigma}, \\ &= tr(\boldsymbol{1}_n^{\circ}(\boldsymbol{1}_n^{\circ\prime}\boldsymbol{1}_n^{\circ})^{-}\boldsymbol{1}_n^{\circ\prime})tr(\boldsymbol{\Psi})\boldsymbol{\Sigma} = (n-1)q\boldsymbol{\Sigma}, \end{aligned} \tag{3.1}$$

where we have used that for any square matrices $\boldsymbol{M}$ and $\boldsymbol{N}$, $\mathrm{tr}(\boldsymbol{M} \otimes \boldsymbol{N}) = \mathrm{tr}(\boldsymbol{M})\mathrm{tr}(\boldsymbol{N})$, $\boldsymbol{1}_n^{\circ}$ is of size $n \times n-1$ and by assumption $\mathrm{tr}(\boldsymbol{\Psi}) = q$. Thus the theorem is established. □

In order to find estimators of the unknown parameters in $\boldsymbol{\Psi}$ we observe that from the Law of Large Numbers it follows that if $n \to \infty$ a consistent estimator of $\boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}$ is given by

$$\widehat{\boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}} = \frac{1}{n}\sum_{i=1}^{n} vec(\boldsymbol{X}_i - \widehat{\boldsymbol{ABC}})vec(\boldsymbol{X}_i - \widehat{\boldsymbol{ABC}})',$$

where as $\widehat{\boldsymbol{ABC}}$ we apply Theorem 3.1 and use

$$\widehat{\boldsymbol{ABC}} = n^{-1}\boldsymbol{A}(\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{A})^{-}\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{X}(\boldsymbol{1}_n \otimes \boldsymbol{C}'(\boldsymbol{CC}')^{-}\boldsymbol{C}).$$

We may observe that $\widehat{\boldsymbol{ABC}} \to \boldsymbol{ABC}$ in probability, if $n \to \infty$, since $\boldsymbol{X}(\boldsymbol{1}_n \otimes \boldsymbol{C}'(\boldsymbol{CC}')^{-}\boldsymbol{C}) - \boldsymbol{ABC} \to \boldsymbol{0}$ in probability and $(qn)^{-1}\boldsymbol{S} \to \boldsymbol{\Sigma}$ in probability. The last relation holds because of (3.1) and since $D[\boldsymbol{S}] \to \boldsymbol{0}$ as $n \to \infty$. Let $\boldsymbol{e}_k$

denote the unit base vector of size q with 1 in the kth position and 0 elsewhere. Then, since $(\boldsymbol{e}'_k \otimes \boldsymbol{I})(\boldsymbol{\Psi} \otimes \boldsymbol{\Sigma})(\boldsymbol{e}_l \otimes \boldsymbol{I}) = \boldsymbol{e}'_k \boldsymbol{\Psi} \boldsymbol{e}_l \otimes \boldsymbol{\Sigma} = \psi_{kl}\boldsymbol{\Sigma}$,

$$
\begin{aligned}
\widehat{\psi_{kl}\boldsymbol{\Sigma}} &= \frac{1}{n}\sum_{i=1}^{n}(\boldsymbol{e}'_k \otimes \boldsymbol{I})vec(\boldsymbol{X}_i - \widehat{\boldsymbol{ABC}})vec(\boldsymbol{X}_i - \widehat{\boldsymbol{ABC}})'(\boldsymbol{e}_l \otimes \boldsymbol{I}), \\
&= \frac{1}{n}\sum_{i=1}^{n}(\boldsymbol{x}_{ik} - \widehat{\boldsymbol{ABc}_k})(\boldsymbol{x}_{il} - \widehat{\boldsymbol{ABc}_l})',
\end{aligned}
$$

where $\boldsymbol{x}_{ik}$ and $\boldsymbol{c}_k$ denote the kth columns in $\boldsymbol{X}_i$ and $\boldsymbol{C}$, respectively.

Theorem 3.2. *Let $\boldsymbol{X} \sim N_{p,qn}(\boldsymbol{AB}(\boldsymbol{1}'_n \otimes \boldsymbol{C}), \boldsymbol{\Sigma}, \boldsymbol{I}_n \otimes \boldsymbol{\Psi})$, where diag($\boldsymbol{\Psi}$)=$\boldsymbol{I}$. A consistent estimator of the unknown elements in $\boldsymbol{\Psi} = (\psi_{kl})$ is given by*

$$
\begin{aligned}
\widehat{\psi}_{kl} &= (pn)^{-1}tr(\widehat{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{X}(\boldsymbol{I} \otimes \boldsymbol{e}_k) - \boldsymbol{A}\widehat{\boldsymbol{B}}(\boldsymbol{1}'_n \otimes \boldsymbol{c}_k)) \\
&\qquad \times (\boldsymbol{X}(\boldsymbol{I} \otimes \boldsymbol{e}_l) - \boldsymbol{A}\widehat{\boldsymbol{B}}(\boldsymbol{1}'_n \otimes \boldsymbol{c}_l))', \ k \neq l,
\end{aligned}
$$

where

$$
\begin{aligned}
\boldsymbol{A}\widehat{\boldsymbol{B}}\boldsymbol{c}_k &= n^{-1}\boldsymbol{A}(\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{A})^{-}\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{X}(\boldsymbol{1}_n \otimes \boldsymbol{C}'(\boldsymbol{CC}')^{-}\boldsymbol{c}_k), \\
\widehat{\boldsymbol{\Sigma}} &= (qn)^{-1}\boldsymbol{S}
\end{aligned}
$$

and $\boldsymbol{S}$ is given in Theorem 3.1.

Proof. The proof follows from the fact that $\widehat{\boldsymbol{\Sigma}}$ is a consistent estimator of $\boldsymbol{\Sigma}$ and

$$
\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n} tr(\widehat{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{x}_{ik} - \widehat{\boldsymbol{ABc}_k})(\boldsymbol{x}_{il} - \widehat{\boldsymbol{ABc}_l}) \\
&= n^{-1}tr(\widehat{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{X}(\boldsymbol{I} \otimes \boldsymbol{e}_k) - \boldsymbol{AB}(\boldsymbol{1}'_n \otimes \boldsymbol{c}_k))(\boldsymbol{X}(\boldsymbol{I} \otimes \boldsymbol{e}_l) - \boldsymbol{AB}(\boldsymbol{1}'_n \otimes \boldsymbol{c}_l))').
\end{aligned}
$$

□

Now we turn to the case where restrictions $\boldsymbol{FBG} = \boldsymbol{0}$ hold on the mean parameters.

Theorem 3.3. *Let $\boldsymbol{X} \sim N_{p,qn}(\boldsymbol{AB}(\boldsymbol{1}'_n \otimes \boldsymbol{C}), \boldsymbol{\Sigma}, \boldsymbol{I}_n\boldsymbol{\Psi})$, where diag($\boldsymbol{\Psi}$)=$\boldsymbol{I}$ and $\boldsymbol{FBG} = \boldsymbol{0}$. Unbiased estimators of $\boldsymbol{ABC}$ and $\boldsymbol{\Sigma}$ are given by*

$$
\begin{aligned}
\widehat{\boldsymbol{ABC}} &= \boldsymbol{A}(\boldsymbol{F}')^{\circ}\widehat{\boldsymbol{\theta}}_1\boldsymbol{C} + \boldsymbol{AF}'\widehat{\boldsymbol{\theta}}_2\boldsymbol{G}^{\circ\prime}\boldsymbol{C}, \\
q(n-1)\widehat{\boldsymbol{\Sigma}} &= \boldsymbol{S} = \boldsymbol{X}(\boldsymbol{1}_n^{\circ}(\boldsymbol{1}_n^{\circ\prime}\boldsymbol{1}_n^{\circ})^{-}\boldsymbol{1}_n^{\circ\prime} \otimes \boldsymbol{I})\boldsymbol{X}',
\end{aligned}
$$

where

$$\begin{aligned}
n\widehat{\boldsymbol{\theta}}_2 &= (\boldsymbol{F}\boldsymbol{A}'\boldsymbol{T}_1'\boldsymbol{S}^{-1}\boldsymbol{T}_1\boldsymbol{A}\boldsymbol{F}')^{-}\boldsymbol{F}\boldsymbol{A}'\boldsymbol{T}_1'\boldsymbol{S}^{-1}\boldsymbol{T}_1\boldsymbol{X}(\mathbf{1}_n\otimes\boldsymbol{C}'\boldsymbol{G}^{\circ}(\boldsymbol{G}^{\circ\prime}\boldsymbol{C}\boldsymbol{C}'\boldsymbol{G}^{\circ})^{-}) \\
&\quad + (\boldsymbol{F}\boldsymbol{A}'\boldsymbol{T}_1')^{\circ}\boldsymbol{Z}_{11} + \boldsymbol{F}\boldsymbol{A}'\boldsymbol{T}_1'\boldsymbol{Z}_{12}(\boldsymbol{G}^{\circ\prime}\boldsymbol{C})^{\circ\prime}, \qquad (3.2)\\
\boldsymbol{T}_1 &= \boldsymbol{I} - \boldsymbol{A}(\boldsymbol{F}')^{\circ}((\boldsymbol{F}')^{\circ\prime}\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{A}(\boldsymbol{F}')^{\circ})^{-}(\boldsymbol{F}')^{\circ\prime}\boldsymbol{A}'\boldsymbol{S}^{-1}, \\
n\widehat{\boldsymbol{\theta}}_1 &= ((\boldsymbol{F}')^{\circ\prime}\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{A}(\boldsymbol{F}')^{\circ})^{-}\boldsymbol{F}'^{\circ\prime}\boldsymbol{A}'\boldsymbol{S}^{-1}(\boldsymbol{X} - \boldsymbol{A}\boldsymbol{F}'\widehat{\boldsymbol{\theta}}_2\boldsymbol{G}^{\circ\prime}\boldsymbol{C}) \\
&\quad \times \boldsymbol{\Psi}^{-1}\boldsymbol{C}'(\boldsymbol{C}\boldsymbol{\Psi}^{-1}\boldsymbol{C}')^{-} + ((\boldsymbol{F}')^{\circ\prime}\boldsymbol{A})^{\circ\prime}\boldsymbol{Z}_{21} + (\boldsymbol{F}')^{\circ\prime}\boldsymbol{A}'\boldsymbol{Z}_{22}\boldsymbol{C}^{\circ\prime},
\end{aligned}$$

where $\boldsymbol{Z}_{ij}$ *are arbitrary matrices.*

Proof. Observe that $\boldsymbol{X}(\mathbf{1}_n\otimes\boldsymbol{C}'\boldsymbol{G}^{\circ})$ as well as $\boldsymbol{X}(\mathbf{1}_n\otimes\boldsymbol{C}')$ are independent of $\boldsymbol{S}$. From the proof of Theorem 3.1. it follows that $(q(n-1))^{-1}\boldsymbol{S}$ is an unbiased estimator of $\boldsymbol{\Sigma}$. Hence, it remains to show that $\widehat{\boldsymbol{ABC}}$ is unbiased.
It follows that

$$\begin{aligned}
\widehat{\boldsymbol{ABC}} &= n^{-1}(\boldsymbol{I}-\boldsymbol{T}_1)\boldsymbol{X}(\mathbf{1}_n\otimes\boldsymbol{C}'(\boldsymbol{C}\boldsymbol{C}')^{-}\boldsymbol{C}) - (\boldsymbol{I}-\boldsymbol{T}_1)\boldsymbol{A}\boldsymbol{F}'\widehat{\boldsymbol{\theta}}_2\boldsymbol{G}^{\circ\prime}\boldsymbol{C} \\
&\qquad + \boldsymbol{A}\boldsymbol{F}'\widehat{\boldsymbol{\theta}}_2\boldsymbol{G}^{\circ\prime}\boldsymbol{C} \\
&= n^{-1}\boldsymbol{X}(\mathbf{1}_n\otimes\boldsymbol{C}'(\boldsymbol{C}\boldsymbol{C}')^{-}\boldsymbol{C}) - n^{-1}\boldsymbol{T}_1\boldsymbol{X}(\mathbf{1}_n\otimes\boldsymbol{C}'(\boldsymbol{C}\boldsymbol{C}')^{-}\boldsymbol{C}) \\
&\qquad + \boldsymbol{T}_1\boldsymbol{A}\boldsymbol{F}'\widehat{\boldsymbol{\theta}}_2\boldsymbol{G}^{\circ\prime}\boldsymbol{C}.
\end{aligned}$$

Now

$$\begin{aligned}
n^{-1}E[\boldsymbol{X}(\mathbf{1}_n\otimes\boldsymbol{C}'(\boldsymbol{C}\boldsymbol{C}')^{-}\boldsymbol{C})] &= \boldsymbol{A}(\boldsymbol{F}')^{\circ}\boldsymbol{\theta}_1\boldsymbol{C} + \boldsymbol{A}\boldsymbol{F}'\boldsymbol{\theta}_2\boldsymbol{G}^{\circ\prime}\boldsymbol{C}, \\
n^{-1}E[\boldsymbol{T}_1\boldsymbol{X}(\mathbf{1}_n\otimes\boldsymbol{C}'(\boldsymbol{C}\boldsymbol{C}')^{-}\boldsymbol{C})] &= E[\boldsymbol{T}_1]\boldsymbol{A}\boldsymbol{F}'\boldsymbol{\theta}_2\boldsymbol{G}^{\circ\prime}\boldsymbol{C}, \\
E[\boldsymbol{T}_1\boldsymbol{A}\boldsymbol{F}'\widehat{\boldsymbol{\theta}}_2\boldsymbol{G}^{\circ\prime}\boldsymbol{C}] &= E[\boldsymbol{T}_1]\boldsymbol{A}\boldsymbol{F}'\boldsymbol{\theta}_2\boldsymbol{G}^{\circ\prime}\boldsymbol{C}.
\end{aligned}$$

Thus

$$E[\widehat{\boldsymbol{ABC}}] = \boldsymbol{A}(\boldsymbol{F}')^{\circ}\boldsymbol{\theta}_1\boldsymbol{C} + \boldsymbol{A}\boldsymbol{F}'\boldsymbol{\theta}_2\boldsymbol{G}^{\circ\prime}\boldsymbol{C}$$

and the theorem is verified. □

4 MLEs of $\boldsymbol{B}$, $\boldsymbol{\Sigma}$ and $\boldsymbol{\Psi}$

In this section the problem of finding maximum likelihood estimators of the parameters in the model

$$N_{p,qn}(\boldsymbol{A}\boldsymbol{B}(\mathbf{1}_n'\otimes\boldsymbol{C}), \boldsymbol{\Sigma}, \boldsymbol{I}_n\otimes\boldsymbol{\Psi})$$

will be studied. The estimability condition $\psi_{qq} = 1$ will be supposed to hold throughout the section. The other diagonal elements of $\boldsymbol{\Psi}$ will be positive but unknown which is a somewhat different assumption than in the previous section where $\mathrm{diag}(\boldsymbol{\Psi}) = \boldsymbol{I}$.

The likelihood equals

$$L = c|\boldsymbol{\Sigma}|^{-1/2qn}|\boldsymbol{\Psi}|^{-1/2np}e^{-\frac{1}{2}tr\{\boldsymbol{\Sigma}^{-1}(\boldsymbol{X}-\boldsymbol{AB}(\boldsymbol{1}_n'\otimes\boldsymbol{C}))(\boldsymbol{I}\otimes\boldsymbol{\Psi})^{-1}(\boldsymbol{X}-\boldsymbol{AB}(\boldsymbol{1}_n'\otimes\boldsymbol{C}))'\}}, \tag{4.1}$$

where c is the proportionality constant which does not depend on the parameters. If differentiating (4.1) with respect to the parameter matrix $\boldsymbol{B}$ and the upper triangle of $\boldsymbol{\Sigma}^{-1}$ it follows from Section 2 that

$$\begin{aligned}
n\widehat{\boldsymbol{B}} &= (\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{A})^{-}\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{X}(\boldsymbol{1}_n\otimes\widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}'(\boldsymbol{C}\widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}')^{-}) \\
&\qquad + (\boldsymbol{A}')^{\circ}\boldsymbol{Z}_1 + \boldsymbol{A}'\boldsymbol{Z}_2\boldsymbol{C}^{\circ'}, && (4.2) \\
\boldsymbol{S} &= \boldsymbol{X}(\boldsymbol{I}\otimes\widehat{\boldsymbol{\Psi}}^{-1} - n^{-1}\boldsymbol{1}_n\boldsymbol{1}_n'\otimes\widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}'(\boldsymbol{C}\widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}')^{-}\boldsymbol{C}\widehat{\boldsymbol{\Psi}}^{-1})\boldsymbol{X}', && (4.3) \\
nq\widehat{\boldsymbol{\Sigma}} &= (\boldsymbol{X}-\boldsymbol{A}\widehat{\boldsymbol{B}}(\boldsymbol{1}_n'\otimes\boldsymbol{C}))(\boldsymbol{I}\otimes\widehat{\boldsymbol{\Psi}}^{-1})(\boldsymbol{X}-\boldsymbol{A}\widehat{\boldsymbol{B}}(\boldsymbol{1}_n'\otimes\boldsymbol{C}))' \\
&= \boldsymbol{S} + n^{-1}\boldsymbol{S}\boldsymbol{A}^{o}(\boldsymbol{A}^{o'}\boldsymbol{S}\boldsymbol{A}^{o})^{-}\boldsymbol{A}^{o'}\boldsymbol{X}(\boldsymbol{1}_n\boldsymbol{1}_n'\otimes\widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}'(\boldsymbol{C}\widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}')^{-}\boldsymbol{C}\widehat{\boldsymbol{\Psi}}^{-1}) \\
&\quad \times \boldsymbol{X}'\boldsymbol{A}^{o}(\boldsymbol{A}^{o'}\boldsymbol{S}\boldsymbol{A}^{o})^{-}\boldsymbol{A}^{o'}\boldsymbol{S}. && (4.4)
\end{aligned}$$

It remains to find estimation equations for $\boldsymbol{\Psi}$. Put

$$\boldsymbol{X}_{ic} = \boldsymbol{X}_i - \boldsymbol{A}\widehat{\boldsymbol{B}}\boldsymbol{C}$$

and we will use the partition

$$\boldsymbol{X}_{ic} = (\boldsymbol{X}_{ic1} : \boldsymbol{X}_{icq}).$$

Due to the constraint $\psi_{qq} = 1$ we can not differentiate the likelihood with respect to $\boldsymbol{\Psi}$. Instead we will first perform a suitable decomposition of $\boldsymbol{\Psi}$ and thereafter maximize the likelihood. Let

$$\boldsymbol{\Psi} = \begin{pmatrix} \boldsymbol{\Psi}_{11} & \underline{\boldsymbol{\psi}}_{1q} \\ \underline{\boldsymbol{\psi}}_{1q}' & \psi_{qq} \end{pmatrix}, \quad \boldsymbol{\Psi}_{11} : (q-1)\times(q-1). \tag{4.5}$$

From Srivastava and Khatri (1979, Corollary 1.4.2 (i), p. 8), it follows since $\psi_{qq} = 1$, that

$$\boldsymbol{\Psi}^{-1} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} \boldsymbol{I}_{q-1} \\ -\underline{\boldsymbol{\psi}}_{1q}' \end{pmatrix} \boldsymbol{\Psi}_{1\bullet q}^{-1} \begin{pmatrix} \boldsymbol{I}_{q-1} : & -\underline{\boldsymbol{\psi}}_{1q} \end{pmatrix},$$

where

$$\boldsymbol{\Psi}_{1\bullet q} = \boldsymbol{\Psi}_{11} - \underline{\boldsymbol{\psi}}_{1q}\underline{\boldsymbol{\psi}}'_{1q} : \ (q-1) \times (q-1). \tag{4.6}$$

Moreover,

$$|\boldsymbol{\Psi}| = |\boldsymbol{\Psi}_{1\bullet q}| = |\boldsymbol{\Psi}_{11} - \underline{\boldsymbol{\psi}}_{1q}\underline{\boldsymbol{\psi}}'_{1q}|.$$

Thus, in order to find the MLE for $\boldsymbol{\Psi}$ the likelihood (4.1) may be rewritten as

$$\begin{aligned} L &= c|\boldsymbol{\Sigma}|^{-\frac{1}{2}qn}|\boldsymbol{\Psi}_{1\bullet q}|^{-\frac{1}{2}pn} etr(-\tfrac{1}{2}\sum_{i=1}^{n}\boldsymbol{\Sigma}^{-1}(\boldsymbol{X}_{icq}\boldsymbol{X}'_{icq} \\ &\quad + \boldsymbol{X}_{ic}\begin{pmatrix} \boldsymbol{I}_{q-1} \\ -\underline{\boldsymbol{\psi}}'_{1q} \end{pmatrix}\boldsymbol{\Psi}^{-1}_{1\bullet q}\begin{pmatrix} \boldsymbol{I}_{q-1} : & -\underline{\boldsymbol{\psi}}_{1q} \end{pmatrix}\boldsymbol{X}'_{ic})). \end{aligned}$$

By differentiation with respect to the upper triangle of $\boldsymbol{\Psi}^{-1}_{1\bullet q}$ as well as differentiation with respect to $\underline{\boldsymbol{\psi}}_{1q}$ we obtain after some manipulations

$$np\widehat{\boldsymbol{\Psi}}_{1\bullet q} = \sum_{i=1}^{n}\begin{pmatrix} \boldsymbol{I}_{q-1} : & -\widehat{\underline{\boldsymbol{\psi}}}_{1q} \end{pmatrix}\boldsymbol{X}'_{ic}\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{X}_{ic}\begin{pmatrix} \boldsymbol{I}_{q-1} \\ -\widehat{\underline{\boldsymbol{\psi}}}'_{1q} \end{pmatrix} \tag{4.7}$$

and

$$\widehat{\underline{\boldsymbol{\psi}}}_{1q} = \sum_{i=1}^{n}\boldsymbol{X}'_{ic1}\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{X}_{icq}(\sum_{i=1}^{n}\boldsymbol{X}'_{icq}\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{X}_{icq})^{-1}. \tag{4.8}$$

It is interesting to observe that (4.8) can be simplified. Post-multiplying (4.4) by $\widehat{\boldsymbol{\Sigma}}^{-1}$ and then taking the trace implies

$$\begin{aligned} nqp &= \sum_{i=1}^{n}\boldsymbol{X}'_{icq}\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{X}_{icq} \\ &\quad + tr(\sum_{i=1}^{n}\widehat{\boldsymbol{\Psi}}^{-1}_{1\bullet q}\begin{pmatrix} \boldsymbol{I}_{q-1} : & -\widehat{\underline{\boldsymbol{\psi}}}_{1q} \end{pmatrix}\boldsymbol{X}'_{ic}\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{X}_{ic}\begin{pmatrix} \boldsymbol{I}_{q-1} \\ -\widehat{\underline{\boldsymbol{\psi}}}'_{1q} \end{pmatrix}) \\ &= \sum_{i=1}^{n}\boldsymbol{X}'_{icq}\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{X}_{icq} + nptr(\boldsymbol{I}_{q-1}), \end{aligned} \tag{4.9}$$

where (4.7) has been used. Thus,

$$\sum_{i=1}^{n}\boldsymbol{X}'_{icq}\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{X}_{icq} = np. \tag{4.10}$$

and (4.8) equals

$$\widehat{\underline{\boldsymbol{\psi}}}_{1q} = \frac{1}{np}\sum_{i=1}^{n}\boldsymbol{X}'_{ic1}\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{X}_{icq}. \tag{4.11}$$

Next, we simplify (4.7). Using (4.11) we get after some calculations

$$np\widehat{\boldsymbol{\Psi}}_{1\bullet q} = \sum_{i=1}^{n}\boldsymbol{X}'_{ic1}\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{X}_{ic1} - np\widehat{\underline{\boldsymbol{\psi}}}_{1q}\widehat{\underline{\boldsymbol{\psi}}}'_{1q}. \tag{4.12}$$

Thus,

$$np\widehat{\boldsymbol{\Psi}}_{11} = \sum_{i=1}^{n}\boldsymbol{X}'_{ic1}\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{X}_{ic1}. \tag{4.13}$$

Hence, using (4.8) and (4.13) as well as (4.10), we get

$$\begin{aligned}
\widehat{\boldsymbol{\Psi}} &= \begin{pmatrix} \widehat{\boldsymbol{\Psi}}_{11} & \widehat{\underline{\boldsymbol{\psi}}}_{1q} \\ \widehat{\underline{\boldsymbol{\psi}}}'_{1q} & 1 \end{pmatrix} & (4.14)\\
&= \frac{1}{np}\begin{pmatrix} \sum_{i=1}^{n}\boldsymbol{X}'_{ic1}\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{X}_{ic1} & \sum_{i=1}^{n}\boldsymbol{X}'_{ic1}\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{X}_{icq} \\ \sum_{i=1}^{n}\boldsymbol{X}'_{icq}\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{X}_{ic1} & \sum_{i=1}^{n}\boldsymbol{X}'_{icq}\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{X}_{icq} \end{pmatrix} \\
&= \frac{1}{np}\sum_{i=1}^{n}\boldsymbol{X}'_{ic}\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{X}_{ic}. & (4.15)
\end{aligned}$$

It is also noted that (4.4) can be written as

$$\widehat{\boldsymbol{\Sigma}} = \frac{1}{nq}\sum_{i=1}^{n}\boldsymbol{X}_{ic}\widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{X}'_{ic}. \tag{4.16}$$

In the next theorem the above given calculations are summarized and it presents a "flip-flop" relation which will be utilized in Theorem 4.2.

Theorem 4.1. *Let $\boldsymbol{X} \sim N_{p,qn}(\boldsymbol{A}\boldsymbol{B}(\mathbf{1}'_n \otimes \boldsymbol{C}), \boldsymbol{\Sigma}, \boldsymbol{I}_n \otimes \boldsymbol{\Psi})$, where $\psi_{qq} = 1$. Maximum likelihood estimation equations for the parameters $\boldsymbol{B}$, $\boldsymbol{\Sigma}$ and $\boldsymbol{\Psi}$ are given by*

$$\begin{aligned}
n\widehat{\boldsymbol{ABC}} &= \boldsymbol{A}(\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{A})^{-}\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{X}(\mathbf{1}_n \otimes \widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}'(\boldsymbol{C}\widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}')^{-}\boldsymbol{C}),\\
nq\widehat{\boldsymbol{\Sigma}} &= (\boldsymbol{X} - \boldsymbol{A}\widehat{\boldsymbol{B}}(\mathbf{1}'_n \otimes \boldsymbol{C}))(\boldsymbol{I} \otimes \boldsymbol{\Psi}^{-1})(\boldsymbol{X} - \boldsymbol{A}\widehat{\boldsymbol{B}}(\mathbf{1}'_n \otimes \boldsymbol{C}))',\\
\widehat{\boldsymbol{\Psi}} &= \frac{1}{np}\sum_{i=1}^{n}(\boldsymbol{X}_i - \widehat{\boldsymbol{ABC}})'\widehat{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{X}_i - \widehat{\boldsymbol{ABC}}),
\end{aligned}$$

where $\boldsymbol{S}$ is given by (4.3).

Although the estimators in Theorem 4.1 are rather involved there is still so much structure that they can be considered theoretically. The next theorem comprises the main theoretical result of the paper.

Theorem 4.2. *Let $\boldsymbol{X} \sim N_{p,qn}(\boldsymbol{A}\boldsymbol{B}(\mathbf{1}'_n \otimes \boldsymbol{C}), \boldsymbol{\Sigma}, \boldsymbol{I}_n \otimes \boldsymbol{\Psi})$, where $\psi_{qq} = 1$. If $n > \max(p, q)$ the maximum likelihood estimation equations given in Theorem 4.1 will always converge to the unique maximum provided the starting value $\tilde{\psi}_{qq}$ equals $\tilde{\psi}_{qq} = 1$.*

Proof. Suppose that there exist two different solutions to the equations given in Theorem 4.1, $(\widehat{\boldsymbol{B}}_1, \widehat{\boldsymbol{\Psi}}_1, \widehat{\boldsymbol{\Sigma}}_1)$ and $(\widehat{\boldsymbol{B}}_2, \widehat{\boldsymbol{\Psi}}_2, \widehat{\boldsymbol{\Sigma}}_2)$. Put

$$\begin{aligned} \boldsymbol{X}_{c1} &= \boldsymbol{X} - \boldsymbol{A}\widehat{\boldsymbol{B}_1}(\mathbf{1}'_n \otimes \boldsymbol{C}), \\ \boldsymbol{X}_{c2} &= \boldsymbol{X} - \boldsymbol{A}\widehat{\boldsymbol{B}_2}(\mathbf{1}'_n \otimes \boldsymbol{C}). \end{aligned}$$

Thus, from Theorem 4.1 we obtain

$$nq\widehat{\boldsymbol{\Sigma}}_1 = \boldsymbol{X}_{c1}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Psi}}_1^{-1})\boldsymbol{X}'_{c1}, \tag{4.17}$$

$$\widehat{\boldsymbol{\Psi}}_1 = \frac{1}{np}(vec'(\boldsymbol{I}_n) \otimes \boldsymbol{I}_q)(\boldsymbol{I}_n \otimes \boldsymbol{X}'_{c1}\widehat{\boldsymbol{\Sigma}}_1^{-1}\boldsymbol{X}_{c1})(vec(\boldsymbol{I}_n) \otimes \boldsymbol{I}_q). \tag{4.18}$$

From (4.17) and (4.18) the next crucial relation is obtained:

$$\frac{p}{q}\boldsymbol{I}_q = (vec'(\boldsymbol{I}_n) \otimes \boldsymbol{I}_q)(\boldsymbol{I}_n \otimes \boldsymbol{P}_1)(vec(\boldsymbol{I}_n) \otimes \boldsymbol{I}_q), \tag{4.19}$$

where the projection $\boldsymbol{P}_1$ is given by

$$\boldsymbol{P}_1 = \boldsymbol{X}'_{c1}(\boldsymbol{X}_{c1}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Psi}}_1^{-1})\boldsymbol{X}'_{c1})^{-1}\boldsymbol{X}_{c1}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Psi}}_1^{-1}). \tag{4.20}$$

Similarly for the second set of solutions we have

$$\frac{p}{q}\boldsymbol{I}_q = (vec'(\boldsymbol{I}_n) \otimes \boldsymbol{I}_q)(\boldsymbol{I}_n \otimes \boldsymbol{P}_2)(vec(\boldsymbol{I}_n) \otimes \boldsymbol{I}_q), \tag{4.21}$$

where

$$\boldsymbol{P}_2 = \boldsymbol{X}'_{c2}(\boldsymbol{X}_{c2}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Psi}}_2^{-1})\boldsymbol{X}'_{c2})^{-1}\boldsymbol{X}_{c2}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Psi}}_2^{-1}). \tag{4.22}$$

By subtracting (4.21) from (4.19) the relation

$$\mathbf{0} = (vec'(\boldsymbol{I}_n) \otimes \boldsymbol{I}_q)(\boldsymbol{I}_n \otimes (\boldsymbol{P}_1 - \boldsymbol{P}_2))(vec(\boldsymbol{I}_n) \otimes \boldsymbol{I}_q) \tag{4.23}$$

is obtained, which is equivalent to

$$\mathbf{0} = (\boldsymbol{P} \otimes \boldsymbol{I}_q)(\boldsymbol{I}_n \otimes (\boldsymbol{P}_1 - \boldsymbol{P}_2))(\boldsymbol{P} \otimes \boldsymbol{I}_q), \tag{4.24}$$

where

$$\boldsymbol{P} = n^{-1} vec(\boldsymbol{I}_n) vec'(\boldsymbol{I}_n) \tag{4.25}$$

which is a symmetric projection operator on $C(vec(\boldsymbol{I}_n))$. The next lemma is fundamental.

Lemma 4.3. *Let $\boldsymbol{P}_1$, $\boldsymbol{P}_2$ and $\boldsymbol{P}$ given by (4.20), (4.22) and (4.25), respectively. Then, the only solution to (4.24) is when $\boldsymbol{P}_1 = \boldsymbol{P}_2$.*

Proof. It follows that (4.24) is equivalent to

$$\mathcal{C}((\boldsymbol{P} \otimes \boldsymbol{I}_q)(\boldsymbol{I}_n \otimes (\boldsymbol{P}_1 - \boldsymbol{P}_2))(\boldsymbol{P} \otimes \boldsymbol{I}_q)) = \{\mathbf{0}\} \tag{4.26}$$

and from a relation concerning projections (see Kollo and von Rosen, 2005; Theorem 1.2.16) (4.26) is identical to

$$\mathcal{C}(\boldsymbol{P} \otimes \boldsymbol{I}_q) \cap \{\mathcal{C}((\boldsymbol{P} \otimes \boldsymbol{I}_q))^{\perp} + \mathcal{C}((\boldsymbol{I}_n \otimes (\boldsymbol{P}_1 - \boldsymbol{P}_2))(\boldsymbol{P} \otimes \boldsymbol{I}_q))\} = \{\mathbf{0}\}.$$

Taking the orthogonal complement to this expression gives

$$\mathcal{C}(\boldsymbol{P} \otimes \boldsymbol{I}_q)^{\perp} + \mathcal{C}(\boldsymbol{P} \otimes \boldsymbol{I}_q) \cap \mathcal{C}((\boldsymbol{I}_n \otimes (\boldsymbol{P}_1 - \boldsymbol{P}_2))(\boldsymbol{P} \otimes \boldsymbol{I}_q))^{\perp} = \mathcal{V},$$

where $\mathcal{V}$ denotes the whole space. Thus, in order for this relation to hold, we have to show that

$$\mathcal{C}((\boldsymbol{P} \otimes \boldsymbol{I}_q)) \subseteq \mathcal{C}((\boldsymbol{I}_n \otimes (\boldsymbol{P}_1 - \boldsymbol{P}_2))(\boldsymbol{P} \otimes \boldsymbol{I}_q))^{\perp} \tag{4.27}$$

which implies that it is reasonable to study

$$\mathcal{C}(\boldsymbol{P} \otimes \boldsymbol{I}_q) \subseteq \mathcal{C}(\boldsymbol{I}_n \otimes (\boldsymbol{P}_1 - \boldsymbol{P}_2))^{\perp} \subseteq \mathcal{C}((\boldsymbol{I}_n \otimes (\boldsymbol{P}_1 - \boldsymbol{P}_2))(\boldsymbol{P} \otimes \boldsymbol{I}_q))^{\perp}.$$

However, we show now that $\mathcal{C}((\boldsymbol{P} \otimes \boldsymbol{I}_q)) \subseteq \mathcal{C}(\boldsymbol{I}_n \otimes (\boldsymbol{P}_{\boldsymbol{X}} - \boldsymbol{P}_{\boldsymbol{X}}))^{\perp}$ if and only if

$$\boldsymbol{P}_1 - \boldsymbol{P}_2 = \mathbf{0},$$

and then (4.27) is also true if and only if $\boldsymbol{P}_1 = \boldsymbol{P}_2$. Observe that $\mathcal{C}(\boldsymbol{P} \otimes \boldsymbol{I}_q) = \mathcal{C}(\sum_{i=1}^{n} (\boldsymbol{e}_i \otimes \boldsymbol{e}_i) \otimes \boldsymbol{I}_q)$ and $\mathcal{C}((\boldsymbol{I}_n \otimes (\boldsymbol{P}_1 - \boldsymbol{P}_2))^{\perp} = \mathcal{C}(\boldsymbol{I}_n \otimes \boldsymbol{H})$ for some $\boldsymbol{H}$ which generates $\mathcal{C}(\boldsymbol{P}_1 - \boldsymbol{P}_2)^{\perp}$. If $\mathcal{C}(\boldsymbol{P} \otimes \boldsymbol{I}_q) \subseteq \mathcal{C}(\boldsymbol{I}_n \otimes \boldsymbol{H})$ then $\mathcal{C}(\boldsymbol{e}_i \otimes \boldsymbol{I}_q) \subseteq \mathcal{C}(\boldsymbol{H})$ which implies that

$$\mathcal{C}((\boldsymbol{e}_1 : \boldsymbol{e}_2 : \cdots : \boldsymbol{e}_n) \otimes \boldsymbol{I}_q) \subseteq \mathcal{C}(\boldsymbol{H})$$

and this in turn gives that $\mathcal{C}(\boldsymbol{I}_n \otimes \boldsymbol{I}_q) \subseteq \mathcal{C}(\boldsymbol{H})$. Thus, $\mathcal{C}(\boldsymbol{H})$ generates the whole space and therefore $\mathcal{C}(\boldsymbol{P}_1 - \boldsymbol{P}_2) = \{\boldsymbol{0}\}$ which is equivalent to the matrix relation $(\boldsymbol{P}_1 - \boldsymbol{P}_2) = \boldsymbol{0}$. □

Hence, we have two projectors $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ which are equal and in the next we shortly show the implication of this fact which also proves the theorem. First it is noted that the projectors have to project on the same spaces and thus (let $\boldsymbol{Q}_2 = \boldsymbol{I} - \boldsymbol{P}_2$)

$$\boldsymbol{0} = (\boldsymbol{P}_1 - \boldsymbol{P}_2) = \boldsymbol{P}_1(\boldsymbol{I} - \boldsymbol{P}_2) = \boldsymbol{P}_1\boldsymbol{Q}_2.$$

This is equivalent to

$$\boldsymbol{X}'_{c1}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Psi}}_1^{-1}\widehat{\boldsymbol{\Psi}}_2)\boldsymbol{X}^o_{c1} = \boldsymbol{0}. \tag{4.28}$$

There are two possibilities for (4.28) to hold. Either $\widehat{\boldsymbol{\Psi}}_1 = \widehat{\boldsymbol{\Psi}}_2$ or the column space generated by $\boldsymbol{X}^o_{c1}$ is invariant with respect to $\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Psi}}_1^{-1}\widehat{\boldsymbol{\Psi}}_2$, i.e. the space is generated by the eigenvectors of $\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Psi}}_1^{-1}\widehat{\boldsymbol{\Psi}}_2$. However, since the matrix of eigenvectors is of the form $\boldsymbol{I}_n \otimes \boldsymbol{\Gamma}$ for some $\boldsymbol{\Gamma}$ it shows, since the column space of $\boldsymbol{X}^o_{c1}$ is a function of the observations, that $\boldsymbol{I}_n \otimes \boldsymbol{\Gamma}$ does not generate the space, unless $n = 1$. Thus, in order for (4.28) to hold $\widehat{\boldsymbol{\Psi}}_1 = \widehat{\boldsymbol{\Psi}}_2$. This in turn implies that $\widehat{\boldsymbol{\Sigma}}_1 = \widehat{\boldsymbol{\Sigma}}_2$ and then also $\widehat{\boldsymbol{B}}_1 = \widehat{\boldsymbol{B}}_2$ and hereby the theorem is proved. □

In comparison to Lu & Zimmerman (2005) we have mathematically shown, under a much more general mean structure, that the MLEs satisfy the "flip-flop" algorithm and that they are unique.

Finally we consider the MLEs when there exist restrictions $\boldsymbol{FBG} = \boldsymbol{0}$ on the parameter $\boldsymbol{B}$. For example when testing hypothesis. It follows from (4.1) and Section 2 that the MLEs for $\boldsymbol{B}$ and $\boldsymbol{\Sigma}$ under the restrictions $\boldsymbol{FBG} = \boldsymbol{0}$ satisfy

$$\boldsymbol{A}\widehat{\boldsymbol{B}}\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{F}')^{\circ}\widehat{\boldsymbol{\theta}}_1\boldsymbol{C} + \boldsymbol{A}\boldsymbol{F}'\widehat{\boldsymbol{\theta}}_2\boldsymbol{G}^{\circ\prime}\boldsymbol{C}, \tag{4.29}$$

where

$$\widehat{\boldsymbol{\theta}}_2 = (\boldsymbol{F}\boldsymbol{A}'\boldsymbol{T}'_1\boldsymbol{S}_2^{-1}\boldsymbol{T}_1\boldsymbol{A}\boldsymbol{F}')^{-}\boldsymbol{F}\boldsymbol{A}'\boldsymbol{T}'_1\boldsymbol{S}_2^{-1}\boldsymbol{T}_1\boldsymbol{X}(\boldsymbol{1}_n \\ \otimes \widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}'\boldsymbol{G}^{\circ}(\boldsymbol{G}^{\circ\prime}\boldsymbol{C}\widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}'\boldsymbol{G}^{\circ})^{-}) + (\boldsymbol{F}\boldsymbol{A}'\boldsymbol{T}'_1)^{\circ}\boldsymbol{Z}_{11} + \boldsymbol{F}\boldsymbol{A}'\boldsymbol{T}'_1\boldsymbol{Z}_{12}(\boldsymbol{G}^{\circ\prime}\boldsymbol{C})^{\circ\prime}$$

with

$$\begin{aligned}
\boldsymbol{T}_1 &= \boldsymbol{I} - \boldsymbol{A}(\boldsymbol{F}')^{\circ}((\boldsymbol{F}')^{\circ'}\boldsymbol{A}'\boldsymbol{S}_1^{-1}\boldsymbol{A}(\boldsymbol{F}')^{\circ})^{-}(\boldsymbol{F}')^{\circ'}\boldsymbol{A}'\boldsymbol{S}_1^{-1},\\
\boldsymbol{S}_1 &= \boldsymbol{X}(\boldsymbol{I}\otimes\widehat{\boldsymbol{\Psi}}^{-1} - n^{-1}\mathbf{1}_n\mathbf{1}_n'\otimes\widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}'(\boldsymbol{C}\widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}')^{-}\boldsymbol{C}\widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{X}',\\
\boldsymbol{S}_2 &= \boldsymbol{S}_1 + \boldsymbol{T}_1\boldsymbol{X}(n^{-1}\mathbf{1}_n\mathbf{1}_n'\otimes\widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}'(\boldsymbol{C}\widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}')^{-}\boldsymbol{C}\widehat{\boldsymbol{\Psi}}^{-1})\\
&\quad\times(\boldsymbol{I}\otimes\widehat{\boldsymbol{\Psi}}^{-1} - n^{-1}\mathbf{1}_n\mathbf{1}_n'\otimes\widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}'\boldsymbol{G}^{\circ}(\boldsymbol{G}^{\circ'}\boldsymbol{C}\widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}'\boldsymbol{G}^{\circ})^{-}\boldsymbol{G}^{\circ'}\boldsymbol{C}\widehat{\boldsymbol{\Psi}}^{-1})\\
&\quad\times(n^{-1}\mathbf{1}_n\mathbf{1}_n'\otimes\widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}'(\boldsymbol{C}\widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}')^{-}\boldsymbol{C}\widehat{\boldsymbol{\Psi}}^{-1})\boldsymbol{X}'\boldsymbol{T}_1',\\
\widehat{\boldsymbol{\theta}}_1 &= (\boldsymbol{F}'^{\circ'}\boldsymbol{A}'\boldsymbol{S}_1^{-1}\boldsymbol{A}(\boldsymbol{F}')^{\circ})^{-}\boldsymbol{F}'^{\circ'}\boldsymbol{A}'\boldsymbol{S}_1^{-1}(\boldsymbol{X} - \boldsymbol{A}\boldsymbol{F}'\widehat{\boldsymbol{\theta}}_2\boldsymbol{G}^{\circ'}(\mathbf{1}_n'\otimes\boldsymbol{C}))\\
&\quad\times(\boldsymbol{I}\otimes\widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}'(\boldsymbol{C}\widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}')^{-}) + (\boldsymbol{F}'^{\circ'}\boldsymbol{A})^{\circ'}\boldsymbol{Z}_{21} + \boldsymbol{F}'^{\circ'}\boldsymbol{A}'\boldsymbol{Z}_{22}(\mathbf{1}_n'\otimes\boldsymbol{C})^{\circ'},
\end{aligned}$$

where $\boldsymbol{S}_1$ is assumed to be positive definite and $\boldsymbol{Z}_{ij}$ are arbitrary matrices. Furthermore,

$$nq\widehat{\boldsymbol{\Sigma}} = (\boldsymbol{X} - \boldsymbol{A}\widehat{\boldsymbol{B}}(\mathbf{1}_n'\otimes\boldsymbol{C}))(\boldsymbol{I}\otimes\widehat{\boldsymbol{\Psi}}^{-1})(\boldsymbol{X} - \boldsymbol{A}\widehat{\boldsymbol{B}}(\mathbf{1}_n'\otimes\boldsymbol{C}))'. \quad (4.30)$$

By copying the arguments when there were no restrictions on $\boldsymbol{B}$ we may state that $\widehat{\boldsymbol{\Psi}}$ satisfies

$$\widehat{\boldsymbol{\Psi}} = \begin{pmatrix} \widehat{\boldsymbol{\Psi}}_{11} & \underline{\widehat{\boldsymbol{\psi}}}_{1q} \\ \underline{\widehat{\boldsymbol{\psi}}}_{1q}' & 1 \end{pmatrix} = \frac{1}{np}\sum_{i=1}^{n}\boldsymbol{X}_{ic}'\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{X}_{ic}.$$

Since

$$\widehat{\boldsymbol{\Sigma}} = \frac{1}{nq}\sum_{i=1}^{n}\boldsymbol{X}_{ic}\widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{X}_{ic}', \quad (4.31)$$

where

$$\boldsymbol{X}_{ic} = \boldsymbol{X}_i - \boldsymbol{A}(\boldsymbol{F}')^{\circ}\widehat{\boldsymbol{\theta}}_1\boldsymbol{C} - \boldsymbol{A}\boldsymbol{F}'\widehat{\boldsymbol{\theta}}_2\boldsymbol{G}^{\circ'}\boldsymbol{C}$$

we see that we can apply a "flip-flop" algorithm. Moreover, once again, by discussing projectors it can be shown that there exist only one $\widehat{\boldsymbol{\Psi}}$ and one $\widehat{\boldsymbol{\Sigma}}$. The uniqueness of these estimators implies that also $\widehat{\boldsymbol{ABC}}$ is uniquely estimated, i.e., under some full rank conditions, $\widehat{\boldsymbol{B}}$ is also unique.

5 Likelihood ratio test

In Theorem 4.1 MLEs of $\widehat{\boldsymbol{B}}$, $\widehat{\boldsymbol{\Sigma}}$ and $\widehat{\boldsymbol{\Psi}}$ were presented and in (4.29) – (4.31) estimators of the same parameters were estimated under the condition $\boldsymbol{FBG} = \mathbf{0}$.

Thus, the likelihood ratio, L, for testing H_0: $\boldsymbol{FBG} = \mathbf{0}$ versus H_1: $\boldsymbol{FBG} \neq \mathbf{0}$ can be obtained. From standard asymptotic theory it follows that $-2\ln L \sim \chi^2(f)$, where $f = \dim(\mathcal{C}(\boldsymbol{A}') \cap \mathcal{C}(\boldsymbol{F}'))\dim(\mathcal{C}(\boldsymbol{C}) \cap \mathcal{C}(\boldsymbol{G}))$. We are going to manipulate the likelihood so that one can see the correspondence with the usual likelihood ratio test when $\boldsymbol{\Psi} = \boldsymbol{I}$. First observe that the likelihood ratio statistic is equivalent to

$$\lambda^{-1} = \frac{|\widetilde{\boldsymbol{\Sigma}}|^q |\widetilde{\boldsymbol{\Psi}}|^p}{|\widehat{\boldsymbol{\Sigma}}|^q |\widehat{\boldsymbol{\Psi}}|^p}, \tag{5.1}$$

where $\widehat{\boldsymbol{\Sigma}}$ and $\widehat{\boldsymbol{\Psi}}$ are MLEs under the alternative, and $\widetilde{\boldsymbol{\Sigma}}$ and $\widetilde{\boldsymbol{\Psi}}$ are the MLEs under H_0.

We begin by noting that

$$|nq\widehat{\boldsymbol{\Sigma}}| = |\widehat{\boldsymbol{S}}_1 + \widehat{\boldsymbol{V}}\widehat{\boldsymbol{V}}'|,$$

where

$$\begin{aligned}
\widehat{\boldsymbol{Y}} &= \boldsymbol{X}(\boldsymbol{I} \otimes \widehat{\boldsymbol{\Psi}}^{-1/2}),\\
\widehat{\boldsymbol{C}} &= \boldsymbol{C}\widehat{\boldsymbol{\Psi}}^{-1/2},\\
\widehat{\boldsymbol{S}}_1 &= \widehat{\boldsymbol{Y}}(\boldsymbol{I} - \widehat{\boldsymbol{C}}'(\widehat{\boldsymbol{C}}\widehat{\boldsymbol{C}}')^{-}\widehat{\boldsymbol{C}})\widehat{\boldsymbol{Y}}'\\
\widehat{\boldsymbol{V}} &= n^{-1/2}(\boldsymbol{I} - \boldsymbol{A}(\boldsymbol{A}'\widehat{\boldsymbol{S}}_1^{-1}\boldsymbol{A})^{-}\boldsymbol{A}'\widehat{\boldsymbol{S}}_1^{-1})\widehat{\boldsymbol{Y}}(\mathbf{1}_n \otimes \widehat{\boldsymbol{C}}'(\widehat{\boldsymbol{C}}\widehat{\boldsymbol{C}}')^{-}\widehat{\boldsymbol{C}})
\end{aligned}$$

and

$$\begin{aligned}
&|nq\widetilde{\boldsymbol{\Sigma}}|\\
&\quad = |\widetilde{\boldsymbol{S}}_1 + \boldsymbol{S}_2 + n^{-1}\boldsymbol{T}_2\boldsymbol{T}_1\widetilde{\boldsymbol{Y}}(\mathbf{1}_n\mathbf{1}_n' \otimes \widetilde{\boldsymbol{C}}'\boldsymbol{G}^o(\boldsymbol{G}^{o\prime}\widetilde{\boldsymbol{C}}\widetilde{\boldsymbol{C}}'\boldsymbol{G}^o)^{-}\boldsymbol{G}^{o\prime}\widetilde{\boldsymbol{C}})\widetilde{\boldsymbol{Y}}'\boldsymbol{T}_1'\boldsymbol{T}_2'|,
\end{aligned}$$

where

$$\begin{aligned}
\widetilde{\boldsymbol{Y}} &= \boldsymbol{X}(\boldsymbol{I} \otimes \widetilde{\boldsymbol{\Psi}}^{-1/2}),\\
\widetilde{\boldsymbol{C}} &= \boldsymbol{C}\widetilde{\boldsymbol{\Psi}}^{-1/2},\\
\widetilde{\boldsymbol{S}}_1 &= \widetilde{\boldsymbol{Y}}(\boldsymbol{I} - \widetilde{\boldsymbol{C}}'(\widetilde{\boldsymbol{C}}\widetilde{\boldsymbol{C}}')^{-}\widetilde{\boldsymbol{C}})\widetilde{\boldsymbol{Y}}',\\
\boldsymbol{T}_1 &= \boldsymbol{I} - \boldsymbol{A}(\boldsymbol{F}')^{\circ}((\boldsymbol{F}')^{\circ\prime}\boldsymbol{A}'\widetilde{\boldsymbol{S}}_1^{-1}\boldsymbol{A}(\boldsymbol{F}')^{\circ})^{-}(\boldsymbol{F}')^{\circ\prime}\boldsymbol{A}'\widetilde{\boldsymbol{S}}_1^{-1},\\
\widetilde{\boldsymbol{S}}_2 &= \widetilde{\boldsymbol{S}}_1 + \boldsymbol{T}_1\widetilde{\boldsymbol{Y}}(n^{-1}\mathbf{1}_n\mathbf{1}_n' \otimes \widetilde{\boldsymbol{C}}'(\widetilde{\boldsymbol{C}}\widetilde{\boldsymbol{C}}')^{-}\widetilde{\boldsymbol{C}}\\
&\quad \times(\boldsymbol{I} - n^{-1}\mathbf{1}_n\mathbf{1}_n' \otimes \widetilde{\boldsymbol{C}}'\boldsymbol{G}^{\circ}(\boldsymbol{G}^{\circ\prime}\widetilde{\boldsymbol{C}}\widetilde{\boldsymbol{C}}'\boldsymbol{G}^{\circ})^{-}\boldsymbol{G}^{\circ\prime}\widetilde{\boldsymbol{C}})\\
&\quad \times(n^{-1}\mathbf{1}_n\mathbf{1}_n' \otimes \widetilde{\boldsymbol{C}}'(\widetilde{\boldsymbol{C}}\widetilde{\boldsymbol{C}}')^{-}\widetilde{\boldsymbol{C}})\widetilde{\boldsymbol{Y}}'\boldsymbol{T}_1'.
\end{aligned}$$

From now on we start to manipulate $|nq\widetilde{\boldsymbol{\Sigma}}|$:

$$
\begin{aligned}
|nq\widetilde{\boldsymbol{\Sigma}}| &= |\widetilde{\boldsymbol{S}}_1+\boldsymbol{S}_2||\boldsymbol{I}+\big((\boldsymbol{T}_1\boldsymbol{A}\boldsymbol{F}')^{o'}(\widetilde{\boldsymbol{S}}_1+\boldsymbol{S}_2)(\boldsymbol{T}_1\boldsymbol{A}\boldsymbol{F}')^{o}\big)^{-}(\boldsymbol{T}_1\boldsymbol{A}\boldsymbol{F}')^{o'}\boldsymbol{T}_1 \\
&\quad \times\widetilde{\boldsymbol{Y}}\widetilde{\boldsymbol{C}}'\boldsymbol{G}^{o}(\boldsymbol{G}^{o'}\widetilde{\boldsymbol{C}}\widetilde{\boldsymbol{C}}'\boldsymbol{G}^{o})^{-}\boldsymbol{G}^{o'}\widetilde{\boldsymbol{C}}\widetilde{\boldsymbol{Y}}'\boldsymbol{T}_1'(\boldsymbol{T}_1\boldsymbol{A}\boldsymbol{F}')^{o}|. \qquad (5.2)
\end{aligned}
$$

This expression does not depend on the choice of $(\boldsymbol{T}_1\boldsymbol{A}\boldsymbol{F}')^{o}$. Let $\boldsymbol{M}$ and $\boldsymbol{N}$ be matrices satisfying

$$
\mathcal{C}(\boldsymbol{M})=\mathcal{C}(\boldsymbol{F}')\cap\mathcal{C}(\boldsymbol{A}'),\qquad \mathcal{C}(\boldsymbol{N})=\mathcal{C}(\boldsymbol{G})\cap\mathcal{C}(\boldsymbol{C}).
$$

Then we may use

$$
\begin{aligned}
&(\boldsymbol{T}_1\boldsymbol{A}\boldsymbol{F}')^{o} \\
&\quad =\boldsymbol{I}-\widetilde{\boldsymbol{S}}_1^{-1}\boldsymbol{A}(\boldsymbol{A}'\widetilde{\boldsymbol{S}}_1^{-1}\boldsymbol{A})^{-}\boldsymbol{M}(\boldsymbol{M}'(\boldsymbol{A}'\widetilde{\boldsymbol{S}}_1^{-1}\boldsymbol{A})^{-}\boldsymbol{M})^{-}\boldsymbol{M}'(\boldsymbol{A}'\widetilde{\boldsymbol{S}}_1^{-1}\boldsymbol{A})^{-}\boldsymbol{A}',
\end{aligned}
$$

which implies

$$
\begin{aligned}
&\boldsymbol{T}_1'(\boldsymbol{T}_1\boldsymbol{A}\boldsymbol{F}')^{o}=\boldsymbol{I}-\widetilde{\boldsymbol{S}}_1^{-1}\boldsymbol{A}(\boldsymbol{A}'\widetilde{\boldsymbol{S}}_1^{-1}\boldsymbol{A})^{-}\boldsymbol{A}, \\
&\boldsymbol{T}_1'(\boldsymbol{I}-\widetilde{\boldsymbol{S}}_1^{-1}\boldsymbol{A}(\boldsymbol{A}'\widetilde{\boldsymbol{S}}_1^{-1}\boldsymbol{A})^{-}\boldsymbol{A}')=\boldsymbol{I}-\widetilde{\boldsymbol{S}}_1^{-1}\boldsymbol{A}(\boldsymbol{A}'\widetilde{\boldsymbol{S}}_1^{-1}\boldsymbol{A})^{-}\boldsymbol{A}', \\
&(\boldsymbol{T}_1\boldsymbol{A}\boldsymbol{F}')^{o'}\widetilde{\boldsymbol{S}}_2(\boldsymbol{T}_1\boldsymbol{A}\boldsymbol{F}')^{o}=(\boldsymbol{I}-\boldsymbol{A}(\boldsymbol{A}'\widetilde{\boldsymbol{S}}_1^{-1}\boldsymbol{A})^{-}\boldsymbol{A}'\widetilde{\boldsymbol{S}}_1^{-1})\widetilde{\boldsymbol{Y}}\widetilde{\boldsymbol{C}}'(\widetilde{\boldsymbol{C}}\widetilde{\boldsymbol{C}}')^{-} \\
&\quad\times\boldsymbol{N}(\boldsymbol{N}'(\widetilde{\boldsymbol{C}}\widetilde{\boldsymbol{C}}')^{-}\boldsymbol{N})^{-}\boldsymbol{N}'(\widetilde{\boldsymbol{C}}\widetilde{\boldsymbol{C}}')^{-}\widetilde{\boldsymbol{C}}\widetilde{\boldsymbol{Y}}'(\boldsymbol{I}-\widetilde{\boldsymbol{S}}_1^{-1}\boldsymbol{A}(\boldsymbol{A}'\widetilde{\boldsymbol{S}}_1^{-1}\boldsymbol{A})^{-}\boldsymbol{A}'), \\
&(\boldsymbol{T}_1\boldsymbol{A}\boldsymbol{F}')^{o'}\widetilde{\boldsymbol{S}}_1(\boldsymbol{T}_1\boldsymbol{A}\boldsymbol{F}')^{o}=\widetilde{\boldsymbol{S}}_1 \\
&\quad-\boldsymbol{A}(\boldsymbol{A}'\widetilde{\boldsymbol{S}}_1^{-1}\boldsymbol{A})^{-}\boldsymbol{M}(\boldsymbol{M}'(\boldsymbol{A}'\widetilde{\boldsymbol{S}}_1^{-1}\boldsymbol{A})^{-}\boldsymbol{M})^{-}\boldsymbol{M}'(\boldsymbol{A}'\widetilde{\boldsymbol{S}}_1^{-1}\boldsymbol{A})^{-}\boldsymbol{A}'.
\end{aligned}
$$

Let

$$
\begin{aligned}
&\boldsymbol{K}=\boldsymbol{I}-\boldsymbol{A}(\boldsymbol{A}'\widetilde{\boldsymbol{S}}_1^{-1}\boldsymbol{A})^{-}\boldsymbol{M}(\boldsymbol{M}'(\boldsymbol{A}'\widetilde{\boldsymbol{S}}_1^{-1}\boldsymbol{A})^{-}\boldsymbol{M})^{-}\boldsymbol{M}'(\boldsymbol{A}'\widetilde{\boldsymbol{S}}_1^{-1}\boldsymbol{A})^{-}\boldsymbol{A}'\widetilde{\boldsymbol{S}}_1^{-1}, \\
&\boldsymbol{U}=\boldsymbol{I}-\boldsymbol{A}(\boldsymbol{A}'\widetilde{\boldsymbol{S}}_1^{-1}\boldsymbol{A})^{-}\boldsymbol{A}'\widetilde{\boldsymbol{S}}_1^{-1} \\
&\widetilde{\boldsymbol{V}}=n^{-1/2}\boldsymbol{U}\widetilde{\boldsymbol{Y}}(\boldsymbol{1}_n\otimes\widetilde{\boldsymbol{C}}'(\widetilde{\boldsymbol{C}}\widetilde{\boldsymbol{C}}')^{-}\widetilde{\boldsymbol{C}}).
\end{aligned}
$$

Some calculations show that (5.2) equals

$$
\begin{aligned}
|nq\widetilde{\boldsymbol{\Sigma}}| &= |\widetilde{\boldsymbol{S}}_1+\boldsymbol{S}_2||\boldsymbol{I}+(\boldsymbol{T}_1\boldsymbol{A}\boldsymbol{F}')^{o'}\widetilde{\boldsymbol{S}}_2(\boldsymbol{T}_1\boldsymbol{A}\boldsymbol{F}')^{o'}|^{-1}|\widetilde{\boldsymbol{S}}_1+\widetilde{\boldsymbol{V}}\widetilde{\boldsymbol{V}}'||\boldsymbol{I} \\
&\quad+(\widetilde{\boldsymbol{S}}_1+\widetilde{\boldsymbol{V}}\widetilde{\boldsymbol{V}}')^{-1}\boldsymbol{U}(\boldsymbol{K}^{-}-\boldsymbol{I})\boldsymbol{U}'\widetilde{\boldsymbol{Y}}\widetilde{\boldsymbol{C}}'\boldsymbol{G}^{o}(\boldsymbol{G}^{o'}\widetilde{\boldsymbol{C}}\widetilde{\boldsymbol{C}}'\boldsymbol{G}^{o'})^{-}\boldsymbol{G}^{o'}\widetilde{\boldsymbol{C}}\widetilde{\boldsymbol{Y}}'|. \qquad (5.3)
\end{aligned}
$$

However, since $\boldsymbol{U}\boldsymbol{K}=\boldsymbol{U}$ and $\boldsymbol{K}\boldsymbol{U}=\boldsymbol{U}$, it is established that

$$
\boldsymbol{U}(\boldsymbol{K}^{-}-\boldsymbol{I})\boldsymbol{U}=\boldsymbol{0}.
$$

Thus, (5.3) equals

$$|nq\widetilde{\mathbf{\Sigma}}| = |\widetilde{\boldsymbol{S}}_1 + \boldsymbol{S}_2||\boldsymbol{I} + (\boldsymbol{T}_1\boldsymbol{A}\boldsymbol{F}')^{o'}\widetilde{\boldsymbol{S}}_2(\boldsymbol{T}_1\boldsymbol{A}\boldsymbol{F}')^{o'}|^{-1}|\widetilde{\boldsymbol{S}}_1 + \widetilde{\boldsymbol{V}}\widetilde{\boldsymbol{V}}'|$$

and the likelihood ratio is of the form

$$\begin{aligned}
&\lambda^{-1}\\
&= \frac{|\widetilde{\boldsymbol{S}}_1 + \boldsymbol{S}_2|^q(|\boldsymbol{I} + (\boldsymbol{T}_1\boldsymbol{A}\boldsymbol{F}')^{o'}\widetilde{\boldsymbol{S}}_2(\boldsymbol{T}_1\boldsymbol{A}\boldsymbol{F}')^{o'}|^{-1})^q|\widetilde{\boldsymbol{S}}_1 + \widetilde{\boldsymbol{V}}\widetilde{\boldsymbol{V}}'|^q|\widetilde{\mathbf{\Psi}}|^p}{|\widehat{\boldsymbol{S}}_1 + \widehat{\boldsymbol{V}}\widehat{\boldsymbol{V}}'|^q|\widehat{\mathbf{\Psi}}|^p}\\
&= \frac{|\widetilde{\boldsymbol{S}}_1 + \widetilde{\boldsymbol{V}}\widetilde{\boldsymbol{V}}'|^q|\widetilde{\mathbf{\Psi}}|^p}{|\widehat{\boldsymbol{S}}_1 + \widehat{\boldsymbol{V}}\widehat{\boldsymbol{V}}'|^q|\widehat{\mathbf{\Psi}}|^p}\\
&\times|\boldsymbol{I} + \widetilde{\boldsymbol{S}}_1^{-1}\boldsymbol{U}\widetilde{\boldsymbol{Y}}\widetilde{\boldsymbol{C}}'(\widetilde{\boldsymbol{C}}\widetilde{\boldsymbol{C}}')^{-}\boldsymbol{N}(\boldsymbol{N}'(\widetilde{\boldsymbol{C}}\widetilde{\boldsymbol{C}}')^{-}\boldsymbol{N})^{-}\boldsymbol{N}'(\widetilde{\boldsymbol{C}}\widetilde{\boldsymbol{C}}')^{-}\widetilde{\boldsymbol{C}}\widetilde{\boldsymbol{Y}}'\boldsymbol{U}'|^q\\
&\times(|\boldsymbol{I} + \widetilde{\boldsymbol{S}}_1^{-1}\boldsymbol{U}\widetilde{\boldsymbol{Y}}\widetilde{\boldsymbol{C}}'(\widetilde{\boldsymbol{C}}\widetilde{\boldsymbol{C}}')^{-}\boldsymbol{N}(\boldsymbol{N}'(\widetilde{\boldsymbol{C}}\widetilde{\boldsymbol{C}}')^{-}\boldsymbol{N})^{-})^{-}\boldsymbol{N}'(\widetilde{\boldsymbol{C}}\widetilde{\boldsymbol{C}}')^{-}\boldsymbol{C}\widetilde{\boldsymbol{Y}}'\boldsymbol{U}`'|^{-1})^q.
\end{aligned} \tag{5.4}$$

Hitherto, no full rank conditions have been assumed but in order to simplify (5.4) and compare the statistic with some well known one it is assumed that $\boldsymbol{A}$, $\boldsymbol{C}$ $\boldsymbol{F}$ and $\boldsymbol{G}$ are of full rank. Therefore, in (5.4) we may replace $\boldsymbol{N}$ by $\boldsymbol{G}'$ and later also $\boldsymbol{M}$ by $\boldsymbol{F}'$. After some more calculations it follows that (5.4) can be written

$$\lambda^{-1} = \frac{|\widetilde{\boldsymbol{S}}_1 + \widetilde{\boldsymbol{V}}\widetilde{\boldsymbol{V}}'|^q|\widetilde{\mathbf{\Psi}}|^p}{|\widehat{\boldsymbol{S}}_1 + \widehat{\boldsymbol{V}}\widehat{\boldsymbol{V}}'|^q|\widehat{\mathbf{\Psi}}|^p}\frac{|\boldsymbol{S}_E + S_H|^q}{|\boldsymbol{S}_E|^q}, \tag{5.5}$$

where

$$\begin{aligned}
\boldsymbol{S}_E &= \boldsymbol{F}(\boldsymbol{A}'\widetilde{\boldsymbol{S}}_1^{-1}\boldsymbol{A})^{-1}\boldsymbol{F}'\\
\boldsymbol{S}_H &= \boldsymbol{F}(\boldsymbol{A}'\widetilde{\boldsymbol{S}}_1^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}'\widetilde{\boldsymbol{S}}_1^{-1}\widetilde{\boldsymbol{Y}}\widetilde{\boldsymbol{C}}'(\widetilde{\boldsymbol{C}}\widetilde{\boldsymbol{C}}')^{-1}\boldsymbol{G}(\boldsymbol{G}'\boldsymbol{R}\boldsymbol{G})^{-1}\\
&\quad\times\boldsymbol{G}'(\widetilde{\boldsymbol{C}}\widetilde{\boldsymbol{C}}')^{-1}\widetilde{\boldsymbol{C}}'\widetilde{\boldsymbol{Y}}'\widetilde{\boldsymbol{S}}_1^{-1}\boldsymbol{A}(\boldsymbol{A}'\widetilde{\boldsymbol{S}}_1^{-1}\boldsymbol{A})^{-1}\boldsymbol{F}',\\
\boldsymbol{R} &= (\widetilde{\boldsymbol{C}}\widetilde{\boldsymbol{C}}')^{-1} + (\widetilde{\boldsymbol{C}}\widetilde{\boldsymbol{C}}')^{-1}\widetilde{\boldsymbol{C}}\widetilde{\boldsymbol{Y}}'\widetilde{\boldsymbol{S}}_1^{-1}\boldsymbol{U}\widetilde{\boldsymbol{Y}}\widetilde{\boldsymbol{C}}'(\widetilde{\boldsymbol{C}}\widetilde{\boldsymbol{C}}')^{-1}.
\end{aligned}$$

Concerning (5.5) we can make several observations. If $\mathbf{\Psi}$ would be known then there would be explicit MLEs and in this case (5.5) equals

$$\lambda^{-1} = \frac{|\boldsymbol{S}_E + S_H|^q}{|\boldsymbol{S}_E|^q}, \tag{5.6}$$

which is equivalent to the likelihood obtained by Khatri (1966). In this case also distribution expansions of the likelihood ratio are available. Moreover, for large n, (5.5) will approach Khatri's. Indeed, even for an unknown $\boldsymbol{\Psi}$, the essential part of the likelihood ratio of testing the hypothesis $\boldsymbol{FBG} = \boldsymbol{0}$ is given by (5.6) since it reflects the differences between H_0 and H_1 concerning the parameter of interest, i.e., $\boldsymbol{B}$.

Acknowledgements

The work of Tatjana Nahtman is supported by the grant GMTMS6702 of the Estonian Research Foundation.

References

[1] Andersson, S.A. & Perlman, M.D. (1993) Lattice models for conditional independence in a multivariate normal distribution. *Annals of Statistics* **21**, 1318–1358.

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

[3] Khatri, C.G. (1966). A note on a manova model applied to problem in growth curve. *Annals of the Institute of Statistical Mathematics*, **18**, 75–86.

[4] Kollo, T. and von Rosen, D. (2005). *Advanced Multivariate Statistics with Matrices*, Dordrecht: Springer.

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

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

[7] von Rosen, D. (1989). Maximum likelihood estimators in multivariate linear normal models. *Journal of Multivariate Analysis*, **31**, 187–200.

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

[9] Shitan, M. and Brockwell, P.J. (1995). An asymptotic test for separability of a spatial autoregressive model. *Comunications in Statistics - Theory and Methods*, **24**, 2027–2040.

[10] Srivastava, M.S. (2002). Nested growth curve models. *Sankyā, Series A*, **64**, 379–408.

[11] Srivastava, M.S, Nahtman, T. and von Rosen, D. (2007). Models with a Kronecker Product Covariance Structure: Estimation and Testing. *Submitted.*

[12] Srivastava, M.S. and Khatri, C.G. (1979). *An Introduction to Multivariate Statistics*, North Holland, New York.

[13] Srivastava, M.S. and von Rosen, D. (1999). Growth curve models. *Multivariate Analysis, Design of Experiments, and Survey Sampling*, 547–578. Textbooks Monogr. **159** Dekker, New York.