Afrika Statistika

Afrika Statistika
Vol. 10(2), 2015, pages 827–837.
DOI: http://dx.doi.org/10.16929/as/2015.827.77

ISSN 2316-090X

Bilinear regression model with Kronecker and linear structures for the covariance matrix

Joseph Nzabanita[1], Dietrich von Rosen[2,3] and Martin Singull[2]

[1]Department of Mathematics, University of Rwanda, PO.Box 3900 Kigali, Rwanda.
[2]Department of Mathematics, Linköping University, SE–581 83 Linköping, Sweden
[3]Department of Energy and Technology, Swedish University of Agricultural Sciences, SE–750 07 Uppsala, Sweden

Received August 14, 2015; Accepted December 15, 2015

Abstract. In this paper, the bilinear regression model based on normally distributed random matrix is studied. For these models, the dispersion matrix has the so called Kronecker product structure and they can be used for example to model data with spatio-temporal relationships. The aim is to estimate the parameters of the model when, in addition, one covariance matrix is assumed to be linearly structured. On the basis of n independent observations from a matrix normal distribution, estimating equations in a flip-flop relation are established and the consistency of estimators is studied.

Résumé. Nous abordons dans ce papier du model de regession bilinéaire basé sur une matrice aléatoire gaussienne. Dans les modèles que nous étudions, les matrices de dispersion ont la structure du produit de Kronecker si bien qu'elles sont capables de modéliser les données présentant des relations spatio-temporelles. Le but de cette étude est d'estimer les paramètres du modèle, lorsqu'en plus une matrice de covariance est supposée être linéairement structurée. Etant données n observations normalement distribuées, les équations d'estimation sont établies travers une relation flip-flop. La consistence des estimateurs est étudiée.

Key words: Bilinear regression; Estimating equations; Flip-flop algorithm; Kronecker product structure; Linear structured covariance matrix; Maximum likelihood estimation.
AMS 2010 Mathematics Subject Classification : 62G08; 62J05.

*Corresponding author Joseph Nzabanita : j.nzabanita@ur.ac.rw, nzabanita@gmail.com
Dietrich von Rosen : dietrich.von.rosen@liu.se
Martin Singull : martin.singull@liu.se

1. Introduction

Multivariate repeated measures data sets, which correspond to multiple measurements that are taken over time on more than one response variable on each subject or unit, are common in various research fields such as medicine, pharmacy, environment, engineering, business, etc. This kind of data is known to have a covariance matrix with the so called Kronecker product structure $\boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}$, where $\otimes$ stands for the matrix Kronecker product. The positive definite matrices $\boldsymbol{\Psi}$ and $\boldsymbol{\Sigma}$ are often referred to as the temporal and spatial covariance matrix, respectively. The Kronecker product structure may also occur in any spatio-temporal process like multivariate time series or stochastic processes. It comes naturally to base statistical analyses for this kind of data on the matrix normal model. Apart from the Kronecker product covariance structure, one or both of the matrices $\boldsymbol{\Psi}$ and $\boldsymbol{\Sigma}$ may be structured. In this paper we are interested in the case where $\boldsymbol{\Sigma}$ is linearly structured.

In this paper we consider n independent and identically distributed observation matrices $\boldsymbol{X}_i \sim N_{p,q}(\boldsymbol{M}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$, $i = 1, 2, \ldots, n$. It follows that the dispersion matrix of $\boldsymbol{X}_i$ has a Kronecker product structure, i.e, $D[\boldsymbol{X}_i] = \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}$, where $D[\boldsymbol{X}_i] = D[\text{vec}\boldsymbol{X}_i]$ and vec is the usual vec-operator. For the interpretation we note that $\boldsymbol{\Psi}$ describes the covariances between the columns of $\boldsymbol{X}$. These covariances will be the same for each row of $\boldsymbol{X}$. The other covariance matrix $\boldsymbol{\Sigma}$ describes the covariances between the rows of $\boldsymbol{X}$ which will be the same for each column of $\boldsymbol{X}$. The product $\boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}$ takes into account the covariances between columns as well as the covariances between rows. Therefore, $\boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}$ indicates that the overall covariance consists of the products of the covariances in $\boldsymbol{\Psi}$ and in $\boldsymbol{\Sigma}$, respectively, i.e., $\text{Cov}[x_{ij}, x_{kl}] = \sigma_{ik}\psi_{jl}$, where $\boldsymbol{X} = (x_{ij})$, $\boldsymbol{\Sigma} = (\sigma_{ik})$ and $\boldsymbol{\Psi} = (\psi_{jl})$. In Dutilleul (1999) the matrix normal distribution is reviewed and a two-stage algorithm to find maximum likelihood estimators is proposed.

In addition we suppose that the mean $\boldsymbol{M}$ of $\boldsymbol{X}_i$ has a bilinear structure, i.e., $E[\boldsymbol{X}_i] = \boldsymbol{ABC}$, where $\boldsymbol{A} : p \times r$ and $\boldsymbol{C} : s \times q$ are known design matrices of regressors. Throughout this paper, without loss of generality, it is assumed that matrices $\boldsymbol{A}$ and $\boldsymbol{C}$ are of full rank, i.e., $\text{rank}(\boldsymbol{A}) = r$ and $\text{rank}(\boldsymbol{C}) = s$. Also we assume that the covariance matrices $\boldsymbol{\Sigma}$ and $\boldsymbol{\Psi}$ are positive definite.

When $\boldsymbol{\Psi} = \boldsymbol{I}$, the identity matrix, we get the well known growth curve model as introduced in Pothoff and Roy (1964). The growth curve model has been extensively studied and useful references are Khatri (1966); Kollo and von Rosen (2005); Srivastava and Khatri (1979). When $\boldsymbol{\Psi}$ is known, the situation is almost similar to the classical growth curve model. In this case, explicit maximum likelihood estimators (MLEs) were derived by Srivastava *et al.* (2009) and their uniqueness were proved under full rank condition of design matrices.

2. Explicit estimators of linearly structured $\boldsymbol{\Sigma}$ with unknown parameters and known $\boldsymbol{\Psi}$

We observe that for matrices $\boldsymbol{X}_i \sim N_{p,q}(\boldsymbol{ABC}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$, $i = 1, 2, \ldots, n$, we may form a new matrix $\boldsymbol{X} = (\boldsymbol{X}_1 : \boldsymbol{X}_2 : \cdots : \boldsymbol{X}_n)$, such that

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

where $\boldsymbol{1}_n$ is the $n-$dimensional vector of ones, and $\boldsymbol{I}_n$ is the $n \times n$ identity matrix.

When $\boldsymbol{\Psi}$ is known and positive definite, we transform the data by letting $\boldsymbol{Y}_i = \boldsymbol{X}_i\boldsymbol{\Psi}^{-1/2}$, where $\boldsymbol{\Psi}^{-1/2}$ is a symmetric positive definite square root of $\boldsymbol{\Psi}$. This yields $\boldsymbol{Y} = (\boldsymbol{Y}_1 : \boldsymbol{Y}_2 : \cdots : \boldsymbol{Y}_n) : p \times qn$, and the model becomes

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

which is similar to the classical growth curve model.

Now we assume that the matrix $\boldsymbol{\Sigma}$ is linearly structured (see Kollo and von Rosen (2005), Definition 1.3.7). The commonly encountered linear structures for the covariance matrix are the uniform structure, the compound symmetry structure, the banded structure, the Toeplitz structure, etc. The linearly structured covariance matrix will be denoted $\boldsymbol{\Sigma}^{(s)}$. For convenience we define and denote by $\text{vec}\boldsymbol{\Sigma}(K)$ the column-wise vectorization of $\boldsymbol{\Sigma}^{(s)}$ where all 0's and repeated elements (by modulus) have been disregarded. Then there exists, see Kollo and von Rosen (2005), a transformation matrix $\boldsymbol{T}$ such that

$$\text{vec}\boldsymbol{\Sigma}(K) = \boldsymbol{T}\text{vec}\boldsymbol{\Sigma}^{(s)} \quad \text{or} \quad \text{vec}\boldsymbol{\Sigma}^{(s)} = \boldsymbol{T}^{+}\text{vec}\boldsymbol{\Sigma}(K), \tag{1}$$

where $\boldsymbol{T}^{+}$ denotes the Moore-Penrose generalized inverse of $\boldsymbol{T}$.

Put $\boldsymbol{G} = \boldsymbol{1}_n' \otimes \boldsymbol{C}\boldsymbol{\Psi}^{-1/2} : s \times qn$. Then the problem boils down to find estimators in the model

$$\boldsymbol{Y} \sim N_{p,qn}(\boldsymbol{ABG}, \boldsymbol{\Sigma}^{(s)}, \boldsymbol{I}). \tag{2}$$

This problem has been studied by Ohlson and von Rosen (2010). From results in Ohlson and von Rosen (2010), simple manipulations give explicit estimators of parameters in the model (2). The explicit estimator of the mean structure is

$$\boldsymbol{A}\widehat{\boldsymbol{B}}\boldsymbol{G} = \boldsymbol{A}\left(\boldsymbol{A}'(\widehat{\boldsymbol{\Sigma}}_1^{(s)})^{-1}\boldsymbol{A}\right)^{-}\boldsymbol{A}'(\widehat{\boldsymbol{\Sigma}}_1^{(s)})^{-1}\boldsymbol{Y}\boldsymbol{G}'(\boldsymbol{G}\boldsymbol{G}')^{-}\boldsymbol{G}, \tag{3}$$

where $\widehat{\boldsymbol{\Sigma}}_1^{(s)}$ is a consistent estimator of $\boldsymbol{\Sigma}^{(s)}$ obtained from

$$\text{vec}\widehat{\boldsymbol{\Sigma}}_1^{(s)} = \frac{1}{qn-s}\boldsymbol{T}^{+}\left((\boldsymbol{T}^{+})'\boldsymbol{T}^{+}\right)^{-}(\boldsymbol{T}^{+})'\text{vec}\boldsymbol{S},$$

in which $s = \text{rank}(\boldsymbol{G}) = \text{rank}(\boldsymbol{C})$ and $\boldsymbol{S} = \boldsymbol{Y}(\boldsymbol{I} - \boldsymbol{G}'(\boldsymbol{G}\boldsymbol{G}')^{-}\boldsymbol{G})\boldsymbol{Y}'$.

Again, following Ohlson and von Rosen (2010), another consistent estimator of $\boldsymbol{\Sigma}^{(s)}$ is obtained from

$$\text{vec}\widehat{\boldsymbol{\Sigma}}^{(s)} = \boldsymbol{T}^{+}\left((\boldsymbol{T}^{+})'\widehat{\boldsymbol{\Upsilon}}'\widehat{\boldsymbol{\Upsilon}}\boldsymbol{T}^{+}\right)^{-}(\boldsymbol{T}^{+})'\widehat{\boldsymbol{\Upsilon}}'\text{vec}(\boldsymbol{S} + \widehat{\boldsymbol{H}}_1\widehat{\boldsymbol{H}}_1'),$$

where

$$\widehat{\boldsymbol{H}}_1 = \left(\boldsymbol{I} - \boldsymbol{A}\left(\boldsymbol{A}'(\widehat{\boldsymbol{\Sigma}}_1^{(s)})^{-1}\boldsymbol{A}\right)^{-}\boldsymbol{A}'(\widehat{\boldsymbol{\Sigma}}_1^{(s)})^{-1}\right)\boldsymbol{Y}\boldsymbol{G}'(\boldsymbol{G}\boldsymbol{G}')^{-}\boldsymbol{G},$$

$$\widehat{\boldsymbol{\Upsilon}} = (qn-s)\boldsymbol{I} + s\left(\boldsymbol{I} - \boldsymbol{A}\left(\boldsymbol{A}'(\widehat{\boldsymbol{\Sigma}}_1^{(s)})^{-1}\boldsymbol{A}\right)^{-}\boldsymbol{A}'(\widehat{bold\widehat{symbol}\Sigma}_1^{(s)})^{-1}\right) \otimes$$

$$\left(\boldsymbol{I} - \boldsymbol{A}\left(\boldsymbol{A}'(\widehat{\boldsymbol{\Sigma}}_1^{(s)})^{-1}\boldsymbol{A}\right)^{-}\boldsymbol{A}'(\widehat{\boldsymbol{\Sigma}}_1^{(s)})^{-1}\right).$$

From (3) we obtain an unbiased estimator of the parameter matrix $\boldsymbol{B}$,

$$\widehat{\boldsymbol{B}} = \left(\boldsymbol{A}'(\widehat{\boldsymbol{\Sigma}}_1^{(s)})^{-1}\boldsymbol{A}\right)^{-}\boldsymbol{A}'(\widehat{\boldsymbol{\Sigma}}_1^{(s)})^{-1}\boldsymbol{Y}\boldsymbol{G}'(\boldsymbol{G}\boldsymbol{G}')^{-}.$$

3. Estimators of linearly structured $\boldsymbol{\Sigma}$ with unknown parameters and unknown $\boldsymbol{\Psi}$

In this section we consider the model

$$\boldsymbol{X} \sim N_{p,qn}(\boldsymbol{A}\boldsymbol{B}(\boldsymbol{1}_n' \otimes \boldsymbol{C}), \boldsymbol{\Sigma}, \boldsymbol{I}_n \otimes \boldsymbol{\Psi}), \tag{4}$$

in which we assume that $\boldsymbol{\Sigma}$ has a linear structure and is unknown, and $\boldsymbol{\Psi}$ is unknown.

In the unstructured case with the only additional estimability condition $\psi_{qq} = 1$, the maximum likelihood estimation equations were derived in Srivastava *et al.* (2009). Those are

$$n\widehat{\boldsymbol{ABC}} = \boldsymbol{A}(\boldsymbol{A}'\widehat{\boldsymbol{S}}^{-1}\boldsymbol{A})^{-}\boldsymbol{A}'\widehat{\boldsymbol{S}}^{-1}\boldsymbol{X}(\boldsymbol{1}_n \otimes \widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}'(\boldsymbol{C}\widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}')^{-}\boldsymbol{C}), \tag{5}$$

$$nq\widehat{\boldsymbol{\Sigma}} = (\boldsymbol{X} - \boldsymbol{A}\widehat{\boldsymbol{B}}(\boldsymbol{1}_n' \otimes \boldsymbol{C}))(\boldsymbol{I} \otimes \widehat{\boldsymbol{\Psi}}^{-1})(\boldsymbol{X} - \boldsymbol{A}\widehat{\boldsymbol{B}}(\boldsymbol{1}_n' \otimes \boldsymbol{C}))', \tag{6}$$

and

$$\widehat{\boldsymbol{\Psi}} = \frac{1}{np}\sum_{i=1}^{n}(\boldsymbol{X}_i - \widehat{\boldsymbol{ABC}})'\widehat{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{X}_i - \widehat{\boldsymbol{ABC}}), \tag{7}$$

where $\widehat{\boldsymbol{S}}$ and $\widehat{\boldsymbol{B}}$ are given by

$$n\widehat{\boldsymbol{B}} = (\boldsymbol{A}'\widehat{\boldsymbol{S}}^{-1}\boldsymbol{A})^{-}\boldsymbol{A}'\widehat{\boldsymbol{S}}^{-1}\boldsymbol{X}(\boldsymbol{1}_n \otimes \widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}'(\boldsymbol{C}\widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}')^{-}), \tag{8}$$

$$\widehat{\boldsymbol{S}} = \boldsymbol{X}(\boldsymbol{I} \otimes \widehat{\boldsymbol{\Psi}}^{-1} - n^{-1}\boldsymbol{1}_n\boldsymbol{1}_n' \otimes \widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}'(\boldsymbol{C}\widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}')^{-}\boldsymbol{C}\widehat{\boldsymbol{\Psi}}^{-1})\boldsymbol{X}', \tag{9}$$

and $\widehat{\boldsymbol{S}}$ is assumed to be positive definite.

Moreover, it has been shown in Srivastava *et al.* (2009) that solving these equations, using the flip-flop algorithm, the estimates in the algorithm converge to the unique maximum likelihood estimators of the parameters. However, when $\boldsymbol{\Sigma}$ is linearly structured this approach will hardly produce an estimator with the desired structure. In fact, we would achieve the original structure if n is large enough which is not the case for real data sets. The aim of this paper is to build up a flip-flop relation that will handle the linear structure of $\boldsymbol{\Sigma}$.

To find the estimating equation for the linearly structured $\boldsymbol{\Sigma}$, hereafter denoted $\boldsymbol{\Sigma} := \boldsymbol{\Sigma}^{(s)}$, let $\boldsymbol{\Psi}$ in (4) be fixed. Then, we know that

$$\boldsymbol{S} = \boldsymbol{X}(\boldsymbol{I} \otimes \boldsymbol{\Psi}^{-1} - n^{-1}\boldsymbol{1}_n\boldsymbol{1}_n' \otimes \boldsymbol{\Psi}^{-1}\boldsymbol{C}'(\boldsymbol{C}\boldsymbol{\Psi}^{-1}\boldsymbol{C}')^{-}\boldsymbol{C}\boldsymbol{\Psi}^{-1})\boldsymbol{X}' \sim W_p(\boldsymbol{\Sigma}^{(s)}, nq - s),$$

and hence $E[\boldsymbol{S}] = (nq-s)\boldsymbol{\Sigma}^{(s)}$. Here $W_p(\cdot,\cdot)$ denotes the Wishart distribution. So it is natural to use $\boldsymbol{S}$ when finding an estimator of $\boldsymbol{\Sigma}^{(s)}$. We apply a least squares approach, i.e., we minimize

$$\mathrm{tr}\left\{\left(\boldsymbol{S}-(nq-s)\boldsymbol{\Sigma}^{(s)}\right)'\left(\boldsymbol{S}-(nq-s)\,\boldsymbol{\Sigma}^{(s)}\right)\right\} \tag{10}$$

with respect to $\boldsymbol{\Sigma}^{(s)}$. Here, tr stands for the trace of a matrix.

To find the minimizer of (10), we use techniques based on differentiations. For more details on matrix differentiation one can consult Kollo and von Rosen (2005). The expression (10) can be rewritten as

$$\mathrm{tr}\left\{(\boldsymbol{S}-(nq-s)\boldsymbol{\Sigma}^{(s)})'(\,)\right\} = (\mathrm{vec}\boldsymbol{S}-(nq-s)\mathrm{vec}\boldsymbol{\Sigma}^{(s)})'(\,), \tag{11}$$

where the notation $(\boldsymbol{Q})'()$ stands for $(\boldsymbol{Q})'(\boldsymbol{Q})$. We differentiate (11) with respect to $\mathrm{vec}\boldsymbol{\Sigma}(K)$ and equalize to $\mathbf{0}$ to get

$$-2(nq-s)\frac{\mathrm{d}\boldsymbol{\Sigma}^{(s)}}{\mathrm{d}\boldsymbol{\Sigma}(K)}\mathrm{vec}(\boldsymbol{S}-(nq-s)\,\boldsymbol{\Sigma}^{(s)}) = \mathbf{0}, \tag{12}$$

in which $\frac{\mathrm{d}\boldsymbol{\Sigma}^{(s)}}{\mathrm{d}\boldsymbol{\Sigma}(K)}$ is given (see Kollo and von Rosen, 2005) by

$$\frac{\mathrm{d}\boldsymbol{\Sigma}^{(s)}}{\mathrm{d}\boldsymbol{\Sigma}(K)} = (\boldsymbol{T}^{+})'. \tag{13}$$

Combining equations in (1), (12) and (13) we get the linear equation

$$(\boldsymbol{T}^{+})'\mathrm{vec}\boldsymbol{S} = (nq-s)\,(\boldsymbol{T}^{+})'\boldsymbol{T}^{+}\mathrm{vec}\boldsymbol{\Sigma}(K),$$

which is consistent. Its general solution is given by

$$\mathrm{vec}\boldsymbol{\Sigma}(K) = \frac{1}{nq-s}\left((\boldsymbol{T}^{+})'\boldsymbol{T}^{+}\right)^{-}(\boldsymbol{T}^{+})'\mathrm{vec}\boldsymbol{S} + \left((\boldsymbol{T}^{+})'\boldsymbol{T}^{+}\right)^{o}\boldsymbol{z},$$

where $\boldsymbol{z}$ is an arbitrary vector and the notation $\boldsymbol{Q}^{o}$ stands for any matrix of full rank spanning the orthogonal complement of the column space of $\boldsymbol{Q}$. Hence, using (1) we obtain the unique minimizer of (10) given by

$$\mathrm{vec}\boldsymbol{\Sigma}^{(s)} = \boldsymbol{T}^{+}\mathrm{vec}\boldsymbol{\Sigma}(K) = \frac{1}{nq-s}\boldsymbol{T}^{+}\left((\boldsymbol{T}^{+})'\boldsymbol{T}^{+}\right)^{-}(\boldsymbol{T}^{+})'\mathrm{vec}\boldsymbol{S}.$$

Thus, a first estimator for $\boldsymbol{\Sigma}^{(s)}$ is given by

$$\mathrm{vec}\widehat{\boldsymbol{\Sigma}}_1^{(s)} = \frac{1}{nq-s}\boldsymbol{T}^{+}\left((\boldsymbol{T}^{+})'\boldsymbol{T}^{+}\right)^{-}(\boldsymbol{T}^{+})'\mathrm{vec}\widehat{\boldsymbol{S}}, \tag{14}$$

where $\boldsymbol{\Psi}$ in $\boldsymbol{S}$ has been replaced with its estimator $\widehat{\boldsymbol{\Psi}}$ to get $\widehat{\boldsymbol{S}}$ as in (9).

Now we suppose that $\widehat{\boldsymbol{\Sigma}}_1^{(s)}$ is positive definite (which always holds for large n) and use it in (8) instead of $\widehat{\boldsymbol{S}}$ to find an estimator of $\boldsymbol{B}$, given by

$$n\widehat{\boldsymbol{B}} = (\boldsymbol{A}'(\widehat{\boldsymbol{\Sigma}}_1^{(s)})^{-1}\boldsymbol{A})^{-}\boldsymbol{A}'(\widehat{\boldsymbol{\Sigma}}_1^{(s)})^{-1}\boldsymbol{X}(\mathbf{1}_n\otimes\widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}'(\boldsymbol{C}\widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}')^{-}). \tag{15}$$

To derive the final estimator of $\boldsymbol{\Sigma}^{(s)}$, we follow similar ideas as in Nzabanita *et al.* (2012) or Ohlson and von Rosen (2010). Let

$$\boldsymbol{Q}_1 = \boldsymbol{I} - \boldsymbol{A}\left(\boldsymbol{A}'\boldsymbol{S}^{-1}\boldsymbol{A}\right)^{-}\boldsymbol{A}'\boldsymbol{S}^{-1}$$

and

$$\boldsymbol{S}_1 = n^{-1}\boldsymbol{Q}_1\boldsymbol{X}(\mathbf{1}_n\mathbf{1}_n' \otimes \boldsymbol{\Psi}^{-1}\boldsymbol{C}'(\boldsymbol{C}\boldsymbol{\Psi}^{-1}\boldsymbol{C}')^{-}\boldsymbol{C}\boldsymbol{\Psi}^{-1})\boldsymbol{X}'\boldsymbol{Q}_1'.$$

We want to use the sum of $\boldsymbol{S}$ and $\boldsymbol{S}_1$ to find the final estimator of $\boldsymbol{\Sigma}^{(s)}$. We first observe that $\boldsymbol{X}(\mathbf{1}_n\mathbf{1}_n' \otimes \boldsymbol{\Psi}^{-1}\boldsymbol{C}'(\boldsymbol{C}\boldsymbol{\Psi}^{-1}\boldsymbol{C}')^{-}\boldsymbol{C}\boldsymbol{\Psi}^{-1})\boldsymbol{X}'$ is independent of $\boldsymbol{S}$ and thus

$$\boldsymbol{S}_1|\boldsymbol{S} \sim W_p(\boldsymbol{Q}_1\boldsymbol{\Sigma}^{(s)}\boldsymbol{Q}_1', s).$$

Again we apply a least squares approach and minimize

$$\operatorname{tr}\left\{(\boldsymbol{S} + \boldsymbol{S}_1 - [(nq-s)\boldsymbol{\Sigma}^{(s)} + s\boldsymbol{Q}_1\boldsymbol{\Sigma}^{(s)}\boldsymbol{Q}_1'])'(\,)\right\}$$

with respect to $\boldsymbol{\Sigma}^{(s)}$. This is equivalent to the minimization with respect to $\boldsymbol{\Sigma}^{(s)}$ of

$$\left(\operatorname{vec}(\boldsymbol{S} + \boldsymbol{S}_1) - \boldsymbol{\Upsilon}\operatorname{vec}\boldsymbol{\Sigma}^{(s)}\right)'\left(\operatorname{vec}(\boldsymbol{S} + \boldsymbol{S}_1) - \boldsymbol{\Upsilon}\operatorname{vec}\boldsymbol{\Sigma}^{(s)}\right),$$

where

$$\boldsymbol{\Upsilon} = (nq-s)\boldsymbol{I} + s\boldsymbol{Q}_1 \otimes \boldsymbol{Q}_1.$$

Using differentiation techniques, as above, a unique minimizer is obtained

$$\operatorname{vec}\boldsymbol{\Sigma}^{(s)} = \boldsymbol{T}^{+}\left((\boldsymbol{T}^{+})'\boldsymbol{\Upsilon}'\boldsymbol{\Upsilon}\boldsymbol{T}^{+}\right)^{-}(\boldsymbol{T}^{+})'\boldsymbol{\Upsilon}'\operatorname{vec}(\boldsymbol{S} + \boldsymbol{S}_1).$$

Hence, replacing $\boldsymbol{\Psi}$ with its estimator we get an estimating equation for $\boldsymbol{\Sigma}^{(s)}$ given by

$$\operatorname{vec}\widehat{\boldsymbol{\Sigma}}^{(s)} = \boldsymbol{T}^{+}\left((\boldsymbol{T}^{+})'\widehat{\boldsymbol{\Upsilon}}'\widehat{\boldsymbol{\Upsilon}}\boldsymbol{T}^{+}\right)^{-}(\boldsymbol{T}^{+})'\widehat{\boldsymbol{\Upsilon}}'\operatorname{vec}(\widehat{\boldsymbol{S}} + \widehat{\boldsymbol{S}}_1), \tag{16}$$

where

$$\widehat{\boldsymbol{S}}_1 = n^{-1}\widehat{\boldsymbol{Q}}_1\boldsymbol{X}(\mathbf{1}_n\mathbf{1}_n' \otimes \widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}'(\boldsymbol{C}\widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}')^{-}\boldsymbol{C}\widehat{\boldsymbol{\Psi}}^{-1})\boldsymbol{X}'\widehat{\boldsymbol{Q}}_1', \tag{17}$$

$$\widehat{\boldsymbol{Q}}_1 = \boldsymbol{I} - \boldsymbol{A}\left(\boldsymbol{A}'(\widehat{\boldsymbol{\Sigma}}_1^{(s)})^{-1}\boldsymbol{A}\right)^{-}\boldsymbol{A}'(\widehat{\boldsymbol{\Sigma}}_1^{(s)})^{-1}, \tag{18}$$

and

$$\widehat{\boldsymbol{\Upsilon}} = (nq-s)\boldsymbol{I} + s\widehat{\boldsymbol{Q}}_1 \otimes \widehat{\boldsymbol{Q}}_1. \tag{19}$$

The relation (16) gives us an estimating equation for $\boldsymbol{\Sigma}^{(s)}$. Thus, modifying (5) using (15), replacing (6) with (16) and modifying (7) using (15) we obtain the following theorem which is the main result of this paper. Note that the estimators to be uniquely determined, the estimability condition is added.

Theorem 1. *Let* $\boldsymbol{X} \sim N_{p,qn}(\boldsymbol{AB}(\boldsymbol{1}_n' \otimes \boldsymbol{C}), \boldsymbol{\Sigma}^{(s)}, \boldsymbol{I}_n \otimes \boldsymbol{\Psi})$. *Assume that* $\boldsymbol{\sigma}_{pp}^{(s)} = 1$. *The estimation equations for the parameters* $\boldsymbol{\Sigma}^{(s)}$ *and* $\boldsymbol{\Psi}$ *are given by*

$$n\widehat{\boldsymbol{ABC}} = \boldsymbol{A}\left(\boldsymbol{A}'(\widehat{\boldsymbol{\Sigma}}_1^{(s)})^{-1}\boldsymbol{A}\right)^{-}\boldsymbol{A}'(\widehat{\boldsymbol{\Sigma}}_1^{(s)})^{-1}\boldsymbol{X}(\boldsymbol{1}_n \otimes \widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}'(\boldsymbol{C}\widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}')^{-}\boldsymbol{C}), \tag{20}$$

$$\mathrm{vec}\widehat{\boldsymbol{\Sigma}}^{(s)} = \boldsymbol{T}^{+}\left((\boldsymbol{T}^{+})'\widehat{\boldsymbol{\Upsilon}}'\widehat{\boldsymbol{\Upsilon}}\boldsymbol{T}^{+}\right)^{-}(\boldsymbol{T}^{+})'\widehat{\boldsymbol{\Upsilon}}'\mathrm{vec}(\widehat{\boldsymbol{S}} + \widehat{\boldsymbol{S}}_1), \tag{21}$$

and

$$\widehat{\boldsymbol{\Psi}} = \frac{1}{np}\sum_{i=1}^{n}(\boldsymbol{X}_i - \widehat{\boldsymbol{ABC}})'(\widehat{\boldsymbol{\Sigma}}^{(s)})^{-1}(\boldsymbol{X}_i - \widehat{\boldsymbol{ABC}}), \tag{22}$$

where $\widehat{\boldsymbol{\Sigma}}_1^{(s)}$, $\widehat{\boldsymbol{S}}$, $\widehat{\boldsymbol{S}}_1$ *and* $\widehat{\boldsymbol{\Upsilon}}$ *are given by (3), (9), (17) and (19) respectively.*

These equations are nested and cannot be solved explicitly. Therefore an iterative algorithm, like the so called flip-flop algorithm, is required to get estimates of parameters. In addition, these equations are modification of equations presented in Srivastava *et al.* (2009) for an unstructured dispersion matrix where it has been shown that only one solution exists. By construction this still holds true. Few simulations (not included in this paper) have supported that and the solution does not depend on the starting values. Next, we show the consistency of the proposed estimators.

Theorem 2. *The estimators* $\widehat{\boldsymbol{ABC}}$, $\widehat{\boldsymbol{\Psi}}$ *and* $\widehat{\boldsymbol{\Sigma}}^{(s)}$, *given in Theorem 1, are consistent estimators of* $\boldsymbol{ABC}$, $\boldsymbol{\Psi}$ *and* $\boldsymbol{\Sigma}^{(s)}$, *respectively.*

In the sequel, for two random sequences X_n and Y_n, the notation $X_n \overset{a}{=} Y_n$ means that $X_n - Y_n \xrightarrow{p} 0$, $n \to \infty$ and the notation "$\xrightarrow{p}$" means "converges in probablity". The following lemma will be utilized in the proof of Theorem 2.

Lemma 1. *Let* $\widehat{\boldsymbol{\Sigma}}_1^{(s)}$, $\widehat{\boldsymbol{\Sigma}}^{(s)}$ *and* $\widehat{\boldsymbol{\Psi}}$ *be given in (14), (21) and (22), respectively. Then, the following hold*

$$\widehat{\boldsymbol{\Psi}} \overset{a}{=} \boldsymbol{\Psi}\frac{1}{p}\mathrm{tr}(\boldsymbol{\Sigma}^{(s)}(\widehat{\boldsymbol{\Sigma}}^{(s)})^{-1}), \tag{23}$$

$$\widehat{\boldsymbol{\Sigma}}_1^{(s)} \overset{a}{=} \boldsymbol{\Sigma}^{(s)}\left(\frac{1}{p}\mathrm{tr}(\boldsymbol{\Sigma}^{(s)}(\widehat{\boldsymbol{\Sigma}}^{(s)})^{-1})\right)^{-1}, \tag{24}$$

$$\widehat{\boldsymbol{\Sigma}}^{(s)} \overset{a}{=} \boldsymbol{\Sigma}^{(s)}\left(\frac{1}{p}\mathrm{tr}(\boldsymbol{\Sigma}^{(s)}(\widehat{\boldsymbol{\Sigma}}^{(s)})^{-1})\right)^{-1}. \tag{25}$$

Proof. The Cramér-Slutsky's theorem Cramér (1946) will be utilized several times. From equation (22),

$$\mathrm{vec}\widehat{\boldsymbol{\Psi}} = \frac{1}{np}\sum_{i=1}^{n}[(\boldsymbol{X}_i - \widehat{\boldsymbol{ABC}})' \otimes (\boldsymbol{X}_i - \widehat{\boldsymbol{ABC}})']\mathrm{vec}(\widehat{\boldsymbol{\Sigma}}^{(s)})^{-1}.$$

Since, $\widehat{\boldsymbol{ABC}} \xrightarrow{p} \boldsymbol{ABC}$ (see the proof later), this equation gives

$$\mathrm{vec}\widehat{\boldsymbol{\Psi}} \overset{a}{=} \frac{1}{np}\sum_{i=1}^{n}[(\boldsymbol{X}_i - \boldsymbol{ABC})' \otimes (\boldsymbol{X}_i - \boldsymbol{ABC})']\mathrm{vec}(\widehat{\boldsymbol{\Sigma}}^{(s)})^{-1}. \tag{26}$$

As $E[(\boldsymbol{X}_i - \boldsymbol{ABC}) \otimes (\boldsymbol{X}_i - \boldsymbol{ABC})] = \text{vec}\boldsymbol{\Sigma}^{(s)}\text{vec}'\boldsymbol{\Psi}$, the law of large numbers yields

$$\frac{1}{n}\sum_{i=1}^{n}(\boldsymbol{X}_i - \boldsymbol{ABC})' \otimes (\boldsymbol{X}_i - \boldsymbol{ABC})' \xrightarrow{p} \text{vec}\boldsymbol{\Psi}\text{vec}'\boldsymbol{\Sigma}^{(s)}.$$

Hence, the relation (26) becomes

$$\text{vec}\widehat{\boldsymbol{\Psi}} \overset{a}{=} \text{vec}\boldsymbol{\Psi}\frac{1}{p}\text{tr}(\boldsymbol{\Sigma}^{(s)}(\widehat{\boldsymbol{\Sigma}}^{(s)})^{-1}),$$

or equivalently

$$\widehat{\boldsymbol{\Psi}} \overset{a}{=} \boldsymbol{\Psi}\frac{1}{p}\text{tr}(\boldsymbol{\Sigma}^{(s)}(\widehat{\boldsymbol{\Sigma}}^{(s)})^{-1}),$$

which establishes (23).

Since

$$\boldsymbol{S} = \boldsymbol{X}(\boldsymbol{I} \otimes \boldsymbol{\Psi}^{-1} - n^{-1}\boldsymbol{1}_n\boldsymbol{1}_n' \otimes \boldsymbol{\Psi}^{-1}\boldsymbol{C}'(\boldsymbol{C}\boldsymbol{\Psi}^{-1}\boldsymbol{C}')^{-}\boldsymbol{C}\boldsymbol{\Psi}^{-1})\boldsymbol{X}' \tag{27}$$

and

$$\frac{1}{nq-s}\boldsymbol{S} \xrightarrow{p} \boldsymbol{\Sigma}^{(s)}, \tag{28}$$

using (23) in (9), it follows that

$$\widehat{\boldsymbol{S}} \overset{a}{=} (nq-s)\boldsymbol{\Sigma}^{(s)}\left(\frac{1}{p}\text{tr}(\boldsymbol{\Sigma}^{(s)}(\widehat{\boldsymbol{\Sigma}}^{(s)})^{-1})\right)^{-1}. \tag{29}$$

From (1) and use of (29) in (14) yield

$$\begin{aligned}
\text{vec}\widehat{\boldsymbol{\Sigma}}_1^{(s)} &\overset{a}{=} \boldsymbol{T}^{+}\left((\boldsymbol{T}^{+})'\boldsymbol{T}^{+}\right)^{-}(\boldsymbol{T}^{+})'\text{vec}\boldsymbol{\Sigma}^{(s)}\left(\frac{1}{p}\text{tr}(\boldsymbol{\Sigma}^{(s)}(\widehat{\boldsymbol{\Sigma}}^{(s)})^{-1})\right)^{-1} \\
&= \boldsymbol{T}^{+}\left((\boldsymbol{T}^{+})'\boldsymbol{T}^{+}\right)^{-}(\boldsymbol{T}^{+})'\boldsymbol{T}^{+}\text{vec}\boldsymbol{\Sigma}(K)\left(\frac{1}{p}\text{tr}(\boldsymbol{\Sigma}^{(s)}(\widehat{\boldsymbol{\Sigma}}^{(s)})^{-1})\right)^{-1} \\
&= \boldsymbol{T}^{+}\text{vec}\boldsymbol{\Sigma}(K)\left(\frac{1}{p}\text{tr}(\boldsymbol{\Sigma}^{(s)}(\widehat{\boldsymbol{\Sigma}}^{(s)})^{-1})\right)^{-1} \\
&= \text{vec}\boldsymbol{\Sigma}^{(s)}\left(\frac{1}{p}\text{tr}(\boldsymbol{\Sigma}^{(s)}(\widehat{\boldsymbol{\Sigma}}^{(s)})^{-1})\right)^{-1}.
\end{aligned}$$

Hence

$$\widehat{\boldsymbol{\Sigma}}_1^{(s)} \overset{a}{=} \boldsymbol{\Sigma}^{(s)}\left(\frac{1}{p}\text{tr}(\boldsymbol{\Sigma}^{(s)}(\widehat{\boldsymbol{\Sigma}}^{(s)})^{-1})\right)^{-1}.$$

Using (24) in (18), it follows that

$$\widehat{\boldsymbol{Q}}_1 \overset{a}{=} \boldsymbol{I} - \boldsymbol{A}\left(\boldsymbol{A}'(\boldsymbol{\Sigma}^{(s)})^{-1}\boldsymbol{A}\right)^{-}\boldsymbol{A}'(\boldsymbol{\Sigma}^{(s)})^{-1} =: \widetilde{\boldsymbol{Q}}_1. \tag{30}$$

Note that

$$\widetilde{\boldsymbol{S}}_1 = n^{-1}\widetilde{\boldsymbol{Q}}_1\boldsymbol{X}(\boldsymbol{1}_n\boldsymbol{1}_n' \otimes \boldsymbol{\Psi}^{-1}\boldsymbol{C}'(\boldsymbol{C}\boldsymbol{\Psi}^{-1}\boldsymbol{C}')^{-}\boldsymbol{C}\boldsymbol{\Psi}^{-1})\boldsymbol{X}'\widetilde{\boldsymbol{Q}}_1'$$

$$= \widetilde{\boldsymbol{Q}}_1(\boldsymbol{X}(\boldsymbol{I}_n \otimes \boldsymbol{\Psi}^{-1})\boldsymbol{X}' - \boldsymbol{S})\widetilde{\boldsymbol{Q}}_1'$$

$$= \widetilde{\boldsymbol{Q}}_1 \left(\sum_{i=1}^{n} \boldsymbol{X}_i \boldsymbol{\Psi}^{-1} \boldsymbol{X}_i' - \boldsymbol{S} \right) \widetilde{\boldsymbol{Q}}_1',$$

where $\boldsymbol{S}$ is given in (27). But, by the law of large numbers

$$\frac{1}{n} \sum_{i=1}^{n} \boldsymbol{X}_i \boldsymbol{\Psi}^{-1} \boldsymbol{X}_i' \xrightarrow{p} q\boldsymbol{\Sigma}^{(s)},$$

which together with (28) imply that

$$\widetilde{\boldsymbol{S}}_1 \overset{a}{=} \widetilde{\boldsymbol{Q}}_1(nq\boldsymbol{\Sigma}^{(s)} - (nq-s)\boldsymbol{\Sigma}^{(s)})\widetilde{\boldsymbol{Q}}_1'$$

$$= s\widetilde{\boldsymbol{Q}}_1 \boldsymbol{\Sigma}^{(s)} \widetilde{\boldsymbol{Q}}_1'.$$

Thus, use (23) and (30) in (17) to get

$$\widehat{\boldsymbol{S}}_1 \overset{a}{=} s\widetilde{\boldsymbol{Q}}_1 \boldsymbol{\Sigma}^{(s)} \widetilde{\boldsymbol{Q}}_1' \left(\frac{1}{p} \mathrm{tr}(\boldsymbol{\Sigma}^{(s)} (\widehat{\boldsymbol{\Sigma}}^{(s)})^{-1}) \right)^{-1}. \tag{31}$$

From (29) and (31),

$$\mathrm{vec}(\widehat{\boldsymbol{S}} + \widehat{\boldsymbol{S}}_1) \overset{a}{=} \left((nq-s)\boldsymbol{I} + s\widetilde{\boldsymbol{Q}}_1 \otimes \widetilde{\boldsymbol{Q}}_1 \right) \mathrm{vec}\boldsymbol{\Sigma}^{(s)} \left(\frac{1}{p} \mathrm{tr}(\boldsymbol{\Sigma}^{(s)} (\widehat{\boldsymbol{\Sigma}}^{(s)})^{-1}) \right)^{-1}$$

$$= \widetilde{\boldsymbol{\Upsilon}} \mathrm{vec}\boldsymbol{\Sigma}^{(s)} \left(\frac{1}{p} \mathrm{tr}(\boldsymbol{\Sigma}^{(s)} (\widehat{\boldsymbol{\Sigma}}^{(s)})^{-1}) \right)^{-1},$$

where $\widetilde{\boldsymbol{\Upsilon}} = (nq-s)\boldsymbol{I} + s\widetilde{\boldsymbol{Q}}_1 \otimes \widetilde{\boldsymbol{Q}}_1$. Hence, from (21) and (1),

$$\mathrm{vec}\widehat{\boldsymbol{\Sigma}}^{(s)} \overset{a}{=} \boldsymbol{T}^{+} \left((\boldsymbol{T}^{+})' \widetilde{\boldsymbol{\Upsilon}}' \widetilde{\boldsymbol{\Upsilon}} \boldsymbol{T}^{+} \right)^{-} (\boldsymbol{T}^{+})' \widetilde{\boldsymbol{\Upsilon}}' \widetilde{\boldsymbol{\Upsilon}} \mathrm{vec}\boldsymbol{\Sigma}^{(s)} \left(\frac{1}{p} \mathrm{tr}(\boldsymbol{\Sigma}^{(s)} (\widehat{\boldsymbol{\Sigma}}^{(s)})^{-1}) \right)^{-1}$$

$$= \boldsymbol{T}^{+} \left((\boldsymbol{T}^{+})' \widetilde{\boldsymbol{\Upsilon}}' \widetilde{\boldsymbol{\Upsilon}} \boldsymbol{T}^{+} \right)^{-} (\boldsymbol{T}^{+})' \widetilde{\boldsymbol{\Upsilon}}' \widetilde{\boldsymbol{\Upsilon}} \boldsymbol{T}^{+} \mathrm{vec}\boldsymbol{\Sigma}(K) \left(\frac{1}{p} \mathrm{tr}(\boldsymbol{\Sigma}^{(s)} (\widehat{\boldsymbol{\Sigma}}^{(s)})^{-1}) \right)^{-1}$$

$$= \boldsymbol{T}^{+} \mathrm{vec}\boldsymbol{\Sigma}(K) \left(\frac{1}{p} \mathrm{tr}(\boldsymbol{\Sigma}^{(s)} (\widehat{\boldsymbol{\Sigma}}^{(s)})^{-1}) \right)^{-1} = \mathrm{vec}\boldsymbol{\Sigma}^{(s)} \left(\frac{1}{p} \mathrm{tr}(\boldsymbol{\Sigma}^{(s)} (\widehat{\boldsymbol{\Sigma}}^{(s)})^{-1}) \right)^{-1},$$

since $\widetilde{\boldsymbol{\Upsilon}}$ has full rank and hence

$$\widehat{\boldsymbol{\Sigma}}^{(s)} \overset{a}{=} \boldsymbol{\Sigma}^{(s)} \left(\frac{1}{p} \mathrm{tr}(\boldsymbol{\Sigma}^{(s)} (\widehat{\boldsymbol{\Sigma}}^{(s)})^{-1}) \right)^{-1} \overset{a}{=} \widehat{\boldsymbol{\Sigma}}_1^{(s)},$$

which proves (25).

Proof (Proof of Theorem 2). From (20) we have

$$\widehat{\boldsymbol{ABC}} = \frac{1}{n}\boldsymbol{A}\left(\boldsymbol{A}'(\widehat{\boldsymbol{\Sigma}}_1^{(s)})^{-1}\boldsymbol{A}\right)^{-}\boldsymbol{A}'(\widehat{\boldsymbol{\Sigma}}_1^{(s)})^{-1}\boldsymbol{X}(\mathbf{1}_n \otimes \widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}'(\boldsymbol{C}\widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}')^{-}\boldsymbol{C})$$

$$= \boldsymbol{A}\left(\boldsymbol{A}'(\widehat{\boldsymbol{\Sigma}}_1^{(s)})^{-1}\boldsymbol{A}\right)^{-}\boldsymbol{A}'(\widehat{\boldsymbol{\Sigma}}_1^{(s)})^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{X}_i\right)\widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}'(\boldsymbol{C}\widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}')^{-}\boldsymbol{C}$$

$$\stackrel{a}{=} \boldsymbol{A}\left(\boldsymbol{A}'(\widehat{\boldsymbol{\Sigma}}_1^{(s)})^{-1}\boldsymbol{A}\right)^{-}\boldsymbol{A}'(\widehat{\boldsymbol{\Sigma}}_1^{(s)})^{-1}\boldsymbol{ABC}\widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}'(\boldsymbol{C}\widehat{\boldsymbol{\Psi}}^{-1}\boldsymbol{C}')^{-}\boldsymbol{C}$$

$$= \boldsymbol{ABC},$$

since the law of large numbers gives

$$\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{X}_i \xrightarrow{p} E\left[\boldsymbol{X}_i\right] = \boldsymbol{ABC}$$

and, $\boldsymbol{A}$ and $\boldsymbol{C}$ are of full column rank and full row rank, respectively. This proves that $\widehat{\boldsymbol{ABC}} \xrightarrow{p} \boldsymbol{ABC}$.

A pre-multiplication of (25) with $(\boldsymbol{\Sigma}^{(s)})^{-1}$ gives

$$(\boldsymbol{\Sigma}^{(s)})^{-1}\widehat{\boldsymbol{\Sigma}}^{(s)} - \boldsymbol{I}\left(\frac{1}{p}\mathrm{tr}(\boldsymbol{\Sigma}^{(s)}(\widehat{\boldsymbol{\Sigma}}^{(s)})^{-1})\right)^{-1} \stackrel{a}{=} \mathbf{0},$$

which implies that

$$\mathrm{tr}\left((\boldsymbol{\Sigma}^{(s)})^{-1}\widehat{\boldsymbol{\Sigma}}^{(s)}\right) - p^2\left(\mathrm{tr}(\boldsymbol{\Sigma}^{(s)}(\widehat{\boldsymbol{\Sigma}}^{(s)})^{-1})\right)^{-1} \stackrel{a}{=} 0.$$

This is equivalent to

$$\sum_{i=1}^{p}\hat{\lambda}_i - p^2\left(\sum_{i=1}^{p}\hat{\lambda}_i^{-1}\right)^{-1} \stackrel{a}{=} 0,$$

where $\hat{\lambda}_i$, $i = 1, 2, \ldots, p$ are eigenvalues of $(\boldsymbol{\Sigma}^{(s)})^{-1}\widehat{\boldsymbol{\Sigma}}^{(s)}$ and $\hat{\lambda}_i > 0$ for all i with probability one. So, we must have $\hat{\lambda}_i \xrightarrow{p} 1$ for all i, which implies that

$$\boldsymbol{\Sigma}^{(s)}(\widehat{\boldsymbol{\Sigma}}^{(s)})^{-1} \stackrel{a}{=} \boldsymbol{I}. \tag{32}$$

Using (32) and Lemma 1 we get that $\widehat{\boldsymbol{\Sigma}}^{(s)} \xrightarrow{p} \boldsymbol{\Sigma}^{(s)}$ and $\widehat{\boldsymbol{\Psi}} \xrightarrow{p} \boldsymbol{\Psi}$ and the proof of Theorem 2 is complete.

In principle the reasoning in the proof of Theorem 2 can be used to prove the consistency of estimators in Srivastava *et al.* (2009) for the unstructured $\boldsymbol{\Sigma}$ (what has not been done yet) or any other estimators based on the flip-flop algorithm. For example, equation (6) gives

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

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

$$\stackrel{a}{=} \sum_{i=1}^{n} [(\boldsymbol{X}_i - \boldsymbol{ABC}) \boldsymbol{\Psi}^{-1} (\boldsymbol{X}_i - \boldsymbol{ABC})'] \left(\frac{1}{p} \mathrm{tr} \left(\boldsymbol{\Sigma} \widehat{\boldsymbol{\Sigma}}^{-1} \right) \right)^{-1}.$$

Using the law of large numbers and

$$E \left[(\boldsymbol{X}_i - \boldsymbol{ABC}) \boldsymbol{\Psi}^{-1} (\boldsymbol{X}_i - \boldsymbol{ABC}) \right] = q \boldsymbol{\Sigma}$$

we have

$$\widehat{\boldsymbol{\Sigma}} \stackrel{a}{=} \boldsymbol{\Sigma} \left(\frac{1}{p} \mathrm{tr} \left(\boldsymbol{\Sigma} \widehat{\boldsymbol{\Sigma}}^{-1} \right) \right)^{-1},$$

which can be used in a similar way as in the proof of Theorem 2 to prove the consistency of estimators in Srivastava *et al.* (2009).

Acknowledgements

The research of Joseph Nzabanita has been supported through the UR-Sweden Program for Research, Higher Learning and Institution Advancement and all persons and institutions involved are hereby acknowledged.

References

Dutilleul, P., 1999. The MLE algorithm for the matrix normal distribution. *J. Stat. Comput. Simul.,* **64**(2), 105-123.

Pothoff, R.F. and Roy, S.N., 1964. A generalized multivariate analysis of variance model useful especially for growth curve problems. *Biometrika,* **51**, 313-326.

Khatri, C.G., 1966. A note on a MANOVA model applied to problems in growth curve, *Ann. Inst. Statist. Math.*, **18**(1), 75-86.

Kollo, T. and von Rosen, D. 2005. *Advanced Multivariate Statistics with Matrices.* Mathematics and Its Applications (New York), 579. Springer, Dordrecht.

Srivastava, M.S. and Khatri, C.G., 1979. *An Introduction to Multivariate Statistics.* North-Holland, New York-Oxford.

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

Ohlson, M. and von Rosen, D., 2010. Explicit estimators of parameters in the growth curve model with linearly structured covariance matrices, *J. Multivariate Anal.*, **101**(5), 1284-1295.

Nzabanita, J., Singull, M. and von Rosen, D., 2012. Estimation of parameters in the extended growth curve model with a linearly structured covariance matrix, *Acta Comment. Univ. Tartu. Math.*, **16**(1), 13-32.

Cramér, H., 1946. *Mathematical Methods of Statistics.* Princeton Mathematical Series, vol. 9. Princeton University Press, Princeton, N. J.