Stockholms universitet

Research Report

Department of Statistics

No. 2013:3

Estimation of Variance Components in Multilevel Models with Block Circular Symmetric Covariance Structures

Yuli Liang

Department of Statistics, Stockholm University, SE-106 91 Stockholm, Sweden

Estimation of Variance Components in Multilevel Models with Block Circular Symmetric Covariance Structures

Yuli Liang

Department of Statistics, Stockholm University, SE-106 91, Stockholm

Abstract

Multilevel model with a block circular symmetric covariance structure is considered. Spectral properties of the corresponding patterned covariance matrix are established. Sufficient conditions for obtaining explicit estimators for the variance-covariance components are presented. Different restricted models are discussed in order to obtain explicit estimators, get interpretable model reparametrizations and keep invariance properties of the block circular symmetric covariance structure.

Keywords: Circular block symmetry, explicit maximum likelihood estimator, restricted model, variance components

1. Introduction

Hierarchically structured data are commonly encountered in many fields of scientific investigations. In educational studies, for example, students' performance in mathematics is measured in various schools. Such data are called hierarchically structured data since students are nested within schools. Another example of hierarchically structured data arise in repeated measurement where outcomes of subjects are repeatedly measured over time. One way to model such data is to use multilevel models (see Raudenbush, 1988; Hox and Kreft, 1994; Goldstein and Woodhouse, 2000; Goldstein, 2003), which can take account the variations from different levels.

The issue of covariance structures from different aspects has been discussed, for example, by Ohlson and von Rosen (2010) in the context of growth curve model, Leiva and Roy (2010) and Louden and Roy (2010) in linear classifications and Pan and Fang (2002) in statistical diagnostic. The presence of symmetry in the data at one or several levels yields a patterned dependence structure within or between the corresponding levels (Dawid, 1988) in the model. Symmetry here means, that dependency between neighboring units remain the same (invariant) even after re-arrangement of units. The properties of the patterned covariance matrices in balanced multilevel models have been studied in Nahtman (2006), Nahtman and von Rosen (2008), von Rosen (2011) and Liang *et al.* (2012).

In this paper, the following multilevel model with a block circular symmetric covariance structure will be considered. Let

$$\mathbf{y} = \mu \mathbf{1}_p + \boldsymbol{Z}_1 \boldsymbol{\gamma}_1 + \boldsymbol{Z}_2 \boldsymbol{\gamma}_2 + \boldsymbol{\epsilon}, \tag{1}$$

where $\mathbf{y}$ is a $p \times 1$ vector of measurements, μ is an unknown constant, $\boldsymbol{\gamma}_1$, $\boldsymbol{\gamma}_2$ and $\boldsymbol{\epsilon}$ are independently normally distributed random variables with zero means and variances-covariance matrices $\boldsymbol{\Sigma}_1$, $\boldsymbol{\Sigma}_2$, and $\sigma^2 \boldsymbol{I}_p$, respectively. Here $p = n_1 n_2$, $\boldsymbol{Z}_1 = \boldsymbol{I}_{n_2} \otimes \mathbf{1}_{n_1}$, $\boldsymbol{Z}_2 = \boldsymbol{I}_{n_2} \otimes \boldsymbol{I}_{n_1}$, $\mathbf{1}_s$ is a column vector of size s with all elements equal to one, $\boldsymbol{I}_s$ is the identity matrix of order s, and $\otimes$ denotes the Kronecker product. Thus,

$$\mathbf{y} \sim N_p(\mu \mathbf{1}_p, \boldsymbol{\Sigma}), \tag{2}$$

$$\boldsymbol{\Sigma} = \boldsymbol{Z}_1 \boldsymbol{\Sigma}_1 \boldsymbol{Z}_1' + \boldsymbol{\Sigma}_2 + \sigma^2 \boldsymbol{I}_p. \tag{3}$$

The covariance matrix $\boldsymbol{\Sigma}$ in (3) may have different structures under different assumptions concerning $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$. With unstructured $\boldsymbol{\Sigma}_2$, the model is obviously overparametrized. In this paper, we will consider $\boldsymbol{\Sigma}$ when the covariance matrix $\boldsymbol{\Sigma}_1 : n_2 \times n_2$ is compound symmetric, i.e.

$$\boldsymbol{\Sigma}_1 = \sigma_1 \boldsymbol{I}_{n_2} + \sigma_2 (\boldsymbol{J}_{n_2} - \boldsymbol{I}_{n_2}), \tag{4}$$

where σ_1 and σ_2 are unknown parameters. Furthermore, the covariance matrix $\boldsymbol{\Sigma}_2 : p \times p$ is assumed to follow a block compound symmetric pattern with a symmetric circular Toeplitz (SC-Toeplitz) matrix in each block, i.e.

$$\boldsymbol{\Sigma}_2 = \boldsymbol{I}_{n_2} \otimes \boldsymbol{\Sigma}^{(1)} + (\boldsymbol{J}_{n_2} - \boldsymbol{I}_{n_2}) \otimes \boldsymbol{\Sigma}^{(2)}, \tag{5}$$

where the SC-Toeplitz matrix $\boldsymbol{\Sigma}^{(h)} = (\sigma_{ij}^{(h)})$ depends on r, $r = [n_1/2]+1$, parameters, the notation $[\bullet]$ stands for the integer part, and for $i, j = 1, \ldots, n_1$, $h = 1, 2$,

$$\sigma_{ij}^{(h)} = \begin{cases} \tau_{|j-i|+1+(h-1)r}, & \text{if } |j-i| \leq r-1, \\ \tau_{n_1-|j-i|+1+(h-1)r}, & \text{otherwise}, \end{cases} \tag{6}$$

and τ_q's are unknown parameters, $q = 1, \ldots, 2r$.

Let $\boldsymbol{\theta} = (\sigma^2, \sigma_1, \sigma_2, \tau_1, \ldots, \tau_{2r})'$ be the vector of unknown (co)variance components. The main problem is to estimate μ and $\boldsymbol{\theta}$ based on independent observed vectors $\mathbf{y}_1, \ldots, \mathbf{y}_n$, each of them having the normal distribution given by (2). Let $\boldsymbol{Y} = vec(\mathbf{y}_1, \ldots, \mathbf{y}_n)$ and it can be found that $\boldsymbol{Y} \sim N_{pn}(\mu \mathbf{1}_{pn}, \boldsymbol{I}_n \otimes \boldsymbol{\Sigma})$, where $vec(\cdot)$ denotes the vectorization operation and $\boldsymbol{\Sigma}$ is given by (3). It has been shown (for example, see Puntanen and Styan, 1989) that the maximum likelihood estimator (MLE) of the mean parameter can be presented in an explicit form if $\boldsymbol{P\Sigma} = \boldsymbol{\Sigma P}$, where $\boldsymbol{P} = \mathbf{1}_{pn}(\mathbf{1}'_{pn}\mathbf{1}_{pn})^{-1}\mathbf{1}'_{pn}$, i.e., $\hat{\mu} = (\mathbf{1}'_{pn}\mathbf{1}_{pn})^{-1}\mathbf{1}'_{pn}\boldsymbol{Y}$.

For any estimation problem, the first question is whether the model parameters can or cannot be uniquely estimated from data, i.e., whether the parameters are estimable or not. Concerning the (co)variance components, in the work of Liang *et al.* (2012), it has been shown that $\boldsymbol{\theta}$ in model (1) are confounded and thus non-estimable. In order to obtain explicit MLEs, the only possibility is to put constraints on the elements of $\boldsymbol{\Sigma}$ and consider a restricted model. The question is what kind of constraints can be imposed on $\boldsymbol{\Sigma}$? Is there any natural way of reparametrizing $\boldsymbol{\Sigma}$? Can these constraints preserve the invariance of $\boldsymbol{\Sigma}$, i.e. keep the structure given in (3)? Do the constraints affect the estimation of the mean parameter? These questions, which has not been dealt with, will be answered in this paper. Firstly, due to invariance, the eigenvalues and eigenvectors of $\boldsymbol{\Sigma}$ become important quantities to study and one of several implications of invariance is that eigenvectors are independent of the parameters $\boldsymbol{\theta}$. Thus, all constraints imposed on $\boldsymbol{\theta}$ are indeed restrictions on the eigenvalues of $\boldsymbol{\Sigma}$. Secondly, altering the dependency structure of $\mathbf{y}$ can result in change of the structure of $\boldsymbol{\Sigma}$ which may violate the assumptions of invariance. Thirdly, it is important to recall that $\boldsymbol{\Sigma}$ is the sum of three matrices characterizing dependency among the three factors. Thus, imposing constraints upon $\boldsymbol{\Sigma}$ can be done via imposing constraints on some or all of these components. i.e., imposing constraints on some or all factors in model (1). This option seems to be meaningful since one often has information about specific factors in the model. The aim of this paper is to discuss the maximum likelihood (ML) estimability of (co)variance components in multilevel models with block circular symmetric covariance structure and propose some possibilities to obtain explicit MLEs of (co)variance components.

The organization of this article is as follows. In the next section, we introduce some results concerning spectral properties of block circular symmetric covariance matrices and the estimation of the eigenvalues. In Section 3, strategies to find identifiable parameters and obtain explicit MLEs by considering restricted models are presented. Section 4 illustrates the main results as well as an example. Section 5 comprises discussion concerning the obtained results.

2. Preliminaries

In Liang *et al.* (2012), spectral properties of $\mathbf{\Sigma}_1$, $\boldsymbol{Z}_1\mathbf{\Sigma}_1\boldsymbol{Z}_1'$ and $\mathbf{\Sigma}_2$ have been derived. In this section, we will provide some useful results concerning spectral properties and estimation of eigenvalues of $\mathbf{\Sigma}$, given in (3). When estimating $\boldsymbol{\theta}$ in model (1), it is important to derive expressions for the spectral decomposition of the covariance matrix $\mathbf{\Sigma}$. Moreover, the number of distinct eigenvalues of $\mathbf{\Sigma}$ and the linear relation of these different eigenvalues are useful in studying the estimations of (co)variance components.

Theorem 2.1. *Liang* et al. *(2012). Let $\boldsymbol{\theta}$ be the vector of the unknown (co)variance parameters in (1), i.e. $\boldsymbol{\theta} = (\sigma^2, \sigma_1, \sigma_2, \tau_1, \dots, \tau_{2r})'$ and $\boldsymbol{\eta}$ be the vector representing the distinct eigenvalues of $\mathbf{\Sigma}$ given in (3). Then $\boldsymbol{\eta}$ can be expressed as follows:*

$$\boldsymbol{\eta} = \boldsymbol{L}\boldsymbol{\theta}, \tag{7}$$

where

$$\boldsymbol{L} = \left(\boldsymbol{B}_1 \vdots \boldsymbol{B}_2\right), \tag{8}$$

and

$$\boldsymbol{B}_1 = \begin{pmatrix} 1 & n_1 & n_1(n_2-1) \\ \mathbf{1}_{r-1} & \mathbf{0}_{r-1} & \mathbf{0}_{r-1} \\ 1 & n_1 & -n_1 \\ \mathbf{1}_{r-1} & \mathbf{0}_{r-1} & \mathbf{0}_{r-1} \end{pmatrix}, \quad \boldsymbol{B}_2 = \begin{pmatrix} \boldsymbol{A} & (n_2-1)\boldsymbol{A} \\ \boldsymbol{A} & -\boldsymbol{A} \end{pmatrix},$$

and $\mathbf{0}_{[n_1/2]}$ are vectors of length $r-1$, $\boldsymbol{A} = (a_{ij})$ is a $r \times r$ matrix and

$$a_{ij} = 2^{\mathbf{1}_{\{1<j<r\}}} cos(2\pi(i-1)(n_1-j+1)/n_1), \; i, j = 1, \dots, n_1, \tag{9}$$

where $\mathbf{1}_{\{.\}}$ is the indicator function.

As mention previously, under invariance the eigenvectors are independent of the parameters. Therefore, all unknown parameters are connected to the mean μ and the eigenvalues of $\mathbf{\Sigma}$. The maximum likelihood estimators of eigenvalues, as well as their distributions, are now derived.
Let $\boldsymbol{Q}$ be as in the proof of Theorem 2.1. Premultiplying $\boldsymbol{Y}$ by $\boldsymbol{I}_n \otimes \boldsymbol{Q}'$, we obtain,

$$(\boldsymbol{I}_n \otimes \boldsymbol{Q}')\boldsymbol{Y} \sim N_{pn}(\mu \left[\mathbf{1}_n \otimes (\sqrt{p}, 0, \dots, 0)'\right], \boldsymbol{I}_n \otimes \boldsymbol{D}(\boldsymbol{\eta})),$$

where $\boldsymbol{D}(\boldsymbol{\eta})$ is a $p \times p$ diagonal matrix with the eigenvalues of $\boldsymbol{\Sigma}$, given in Theorem 2.1, on the main diagonal, i.e. the model can be splitted into $2r$ independent models. Define $\boldsymbol{\omega}_i = (\boldsymbol{I}_n \otimes \boldsymbol{v}_i')\boldsymbol{Y}$, and let

$$
\begin{aligned}
\tilde{\mathbf{y}}_1 &= \boldsymbol{\omega}_1, \\
\tilde{\mathbf{y}}_i &= vec(\boldsymbol{\omega}_i, \boldsymbol{\omega}_{n_1-i+2}), \\
\tilde{\mathbf{y}}_{r+1} &= vec(\boldsymbol{\omega}_{n_1+1}, \boldsymbol{\omega}_{2n_1+1}, \ldots, \boldsymbol{\omega}_{(n_2-1)n_1+1}), \\
\tilde{\mathbf{y}}_{r+i} &= vec(\boldsymbol{\omega}_{n_1+i}, \boldsymbol{\omega}_{2n_1-i+2}, \boldsymbol{\omega}_{2n_1+i}, \boldsymbol{\omega}_{3n_1-i+2}, \ldots, \boldsymbol{\omega}_{(n_2-1)n_1+i}, \boldsymbol{\omega}_{(n_2-1)n_1-i+2}),
\end{aligned}
$$

where $i = 2, \ldots, r$. Then we have

$$
\begin{aligned}
\tilde{\mathbf{y}}_1 &\sim N_n(\sqrt{p}\mu \mathbf{1}_n, \eta_1 \boldsymbol{I}_n), \\
\tilde{\mathbf{y}}_i &\sim N_{nm_i}(\mathbf{0}, \eta_i \boldsymbol{I}_{nm_i}), \ i = 2, \ldots, 2r,
\end{aligned}
$$

where m_i is the multiplicity of η_i (see Liang *et al.*, 2012). It turns out that μ and η_1 are estimated through $\tilde{\mathbf{y}}_1$. The MLE of η_1 is

$$
\hat{\eta}_1 = \frac{1}{n}\tilde{\mathbf{y}}_1'(\boldsymbol{I}_n - \mathbf{1}_n(\mathbf{1}_n'\mathbf{1}_n)^{-1}\mathbf{1}_n')\tilde{\mathbf{y}}_1, \tag{10}
$$

and the MLE of η_i is

$$
\hat{\eta}_i = (1/nm_i)\tilde{\mathbf{y}}_i'\tilde{\mathbf{y}}_i, \ i = 2, \ldots, 2r. \tag{11}
$$

Now, it is straightforward that $\hat{\eta}_1 \sim \frac{\eta_1}{n}\chi^2_{(n-1)}$, and $\hat{\eta}_i \sim \frac{\eta_i}{nm_i}\chi^2_{(nm_i)}$, $i = 2, \ldots, 2r$.

3. Identifiability of parameters

Szatrowski (1980, Theorem 5) presented a necessary and sufficient condition for checking whether or not explicit MLEs of the variance components in a balanced mixed model exist, which is of the utmost importance in our work. If the sufficient condition in Szatrowski (1980) holds, i.e. the number of elements in $\boldsymbol{\eta}$ equals the number of elements in $\boldsymbol{\theta}$, the MLE of $\boldsymbol{\theta}$ has an explicit expression, which is obtained by solving the linear system in (7) and $\boldsymbol{\eta}$ is replaced by its MLE. If the number of elements in $\boldsymbol{\eta}$ is less than the number elements in $\boldsymbol{\theta}$, $\boldsymbol{\theta}$ is estimable only under some constraints on $\boldsymbol{\theta}$. The next proposition establishes a relationship between the number of elements in $\boldsymbol{\theta}$ and $\boldsymbol{\eta}$.

Proposition 3.1. *Let s_1 be the number of distinct eigenvalues of $\boldsymbol{\Sigma}$ defined in (3), and s_2 be the number of unknown parameters in $\boldsymbol{\Sigma}$, then $\Delta \equiv s_2 - s_1 = 3$.*

Proof. The result is obtained via (8). Recall that $\boldsymbol{\Sigma}$ given in (3) is the sum of three matrices:

$$\boldsymbol{\Sigma} = \underbrace{\boldsymbol{Z}_1\boldsymbol{\Sigma}_1\boldsymbol{Z}_1'}_{\text{2 parameters}} + \underbrace{\boldsymbol{\Sigma}_2}_{2r \text{ parameters}} + \underbrace{\sigma^2\boldsymbol{I}}_{\text{1 parameter}}. \tag{12}$$

According to Liang *et al.* (2012), it follows that there are $2r$ distinct eigenvalues. So, $\Delta = 3$. ■

Thus, according to Theorem 3.1, model (1) is overparametrized, and hence explicit MLEs do not exist for $\boldsymbol{\theta}$. Hence, we have to put at least three restrictions on the parameter space in order to use the MLEs of the eigenvalues to estimate $\boldsymbol{\theta}$ explicitly. Since σ^2 in (12) is connected to the random error, a restriction on σ^2, i.e. σ^2 is a known constant, will not be considered. Thus, there are two scenarios. Either one can put one constraint on the spectrum of $\boldsymbol{\Sigma}_1$ and two constraints on the spectrum of $\boldsymbol{\Sigma}_2$ or alternatively one can put three constraints on the spectrum of $\boldsymbol{\Sigma}_2$.

Observe that $\boldsymbol{\eta}$ specified in (7) is not only a function of unknown covariance parameters in $\boldsymbol{\theta}$, $\boldsymbol{\eta} = \boldsymbol{\eta}(\boldsymbol{\theta})$, but also a function of the distinct eigenvalues

$$\boldsymbol{\lambda}^{\boldsymbol{\Sigma}_1} = (\lambda_1^{\boldsymbol{\Sigma}_1}, \lambda_2^{\boldsymbol{\Sigma}_1}),\ \boldsymbol{\lambda}^{\boldsymbol{\Sigma}_2} = (\lambda_{11}^{\boldsymbol{\Sigma}_2}, \ldots, \lambda_{1r}^{\boldsymbol{\Sigma}_2}, \lambda_{21}^{\boldsymbol{\Sigma}_2}, \ldots, \lambda_{2r}^{\boldsymbol{\Sigma}_2}) \text{ and } \lambda^{\boldsymbol{I}}$$

of $\boldsymbol{\Sigma}_1$, $\boldsymbol{\Sigma}_2$ and $\sigma^2\boldsymbol{I}$, respectively, i.e.

$$\eta_i = \lambda^{\boldsymbol{I}} + n_1\lambda_h^{\boldsymbol{\Sigma}_1} I(i \in \{1, r+1\}) + \lambda_{hi}^{\boldsymbol{\Sigma}_2}, \tag{13}$$

where $h = 1 + I(i = r+1)$, $i = 1, \ldots, 2r$, and $I(\bullet)$ is the indicator function.

It turns out that instead of putting constraints on $\boldsymbol{\theta}$, it is more reasonable to impose constraints on the eigenvalues of the covariance matrices $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$. The advantage of this approach is that the corresponding eigenvectors will specify the constraints to be imposed on the corresponding factors which can be easily interpretable and the original symmetry assumptions will be preserved.

Scenario 1: One constraint on the spectrum of $\boldsymbol{\Sigma}_1$ and two constraints on the spectrum of $\boldsymbol{\Sigma}_2$.
Two possibilities for imposing restrictions on the spectra are given by
(i) $\lambda_i^{\boldsymbol{\Sigma}_1} = 0$, $\lambda_{i1}^{\boldsymbol{\Sigma}_2} = 0$, $\lambda_{i'1}^{\boldsymbol{\Sigma}_2} = 0$, $i, i' \in \{1, 2\}$, $i \neq i'$;
(ii) $\lambda_i^{\boldsymbol{\Sigma}_1} = 0$, $\lambda_{i'1}^{\boldsymbol{\Sigma}_2} = 0$, $\lambda_{kj}^{\boldsymbol{\Sigma}_2} = 0$, $i, i', k \in \{1, 2\}$, $i \neq i'$, $j \in \{2, \ldots, r\}$.

Remark 3.1. In fact, the condition $\lambda_2^{\boldsymbol{\Sigma}_1} = 0$ in (i) is very restrictive, since in this case $\sigma_1 = \sigma_2 = cov(\gamma_{1i}, \gamma_{1l})$, $i, l = 1, \ldots, n_2$, i.e. the covariance matrix $\boldsymbol{\Sigma}_1$ in

(4) becomes equal to $\sigma_1 \boldsymbol{J}_{n_2}$. In the subsequent, we will only consider the case when $\lambda_1^{\boldsymbol{\Sigma}_1} = 0$ in (i).

Using the relationship between the eigenvalues $\boldsymbol{\lambda}^{\boldsymbol{\Sigma}_2}$ of $\boldsymbol{\Sigma}_2$ and the elements of $\boldsymbol{\Sigma}_2$ (see Liang *et al.*, 2012, Corollary 2.6), conditions (i) and (ii) can be expressed in terms of constraints on $\boldsymbol{\theta}$ as $\boldsymbol{K}_i\boldsymbol{\theta} = 0$, $i = 1, 2$, where

$$\boldsymbol{K}_1 = \left(\begin{array}{ccc|cc} 0 & 1 & (n_2-1) & \mathbf{0}_r & \mathbf{0}_r \\ \hline 0 & 0 & 0 & \boldsymbol{a}_1 & (n_2-1)\boldsymbol{a}_1 \\ 0 & 0 & 0 & \boldsymbol{a}_1 & -\boldsymbol{a}_1 \end{array}\right), \quad (14)$$

$$\boldsymbol{K}_2 = \left(\begin{array}{ccc|cc} 0 & 1 & (n_2-1) & \mathbf{0}_r & \mathbf{0}_r \\ \hline 0 & 0 & 0 & \boldsymbol{a}_1 & -\boldsymbol{a}_1 \\ 0 & 0 & 0 & \boldsymbol{a}_j & -(1-n_2)^{2-k}\boldsymbol{a}_j \end{array}\right), \quad (15)$$

and the vectors $\boldsymbol{a}_1$ and $\boldsymbol{a}_j$ are the corresponding rows of the matrix $\boldsymbol{A}$ defined in (9).

Scenario 2: Three constraints on the spectrum of $\boldsymbol{\Sigma}_2$.
Two possibilities of putting restrictions on the spectrum are given by
(iii) $\lambda_{i1}^{\boldsymbol{\Sigma}_2} = 0$ and $\lambda_{i'j}^{\boldsymbol{\Sigma}_2} = 0$, $i = 1, 2$, $i' \in \{1, 2\}$, $j \in \{2, \ldots, r\}$.
Conditions in (iii) can also be expressed as $\boldsymbol{K}_3\boldsymbol{\theta} = 0$, where

$$\boldsymbol{K}_3 = \begin{pmatrix} 0 & 0 & 0 & \boldsymbol{a}_1 & (n_2-1)\boldsymbol{a}_1 \\ 0 & 0 & 0 & \boldsymbol{a}_1 & -\boldsymbol{a}_1 \\ 0 & 0 & 0 & \boldsymbol{a}_j & -(1-n_2)^{2-i'}\boldsymbol{a}_j \end{pmatrix}, \quad (16)$$

and the vectors $\boldsymbol{a}_1$ and $\boldsymbol{a}_j$ are the corresponding rows of matrix $\boldsymbol{A}$ defined in (9). In order to have better understanding the meaning of restrictions (i)-(iii), their implications for the factors $\boldsymbol{\gamma}_1$ and $\boldsymbol{\gamma}_2$ in model (1) will be studied.
The most commonly used constraints when performing data analysis are "sum-to-zero" and "set-to-zero" constraints: if $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_l)'$ is a vector of parameters, then the "sum-to-zero" constraint for $\boldsymbol{\gamma}$ equals $\sum_{i=1}^{l} \gamma_i = 0$ and the "set-to-zero" constraint for $\boldsymbol{\gamma}$, for example, equals $\gamma_l = 0$. Nahtman (2006) showed that the restriction $n_1\sigma_1 + n_1(n_2-1)\sigma_2 = 0$ in Scenario 1 is equivalent to the "sum-to-zero" restriction of the random factor $\boldsymbol{\gamma}_1$ in model (1) and described the structure of $\boldsymbol{\Sigma}_1$ given in (4) under this restriction.

Theorem 3.2. *Nahtman (2006). Let $\boldsymbol{\gamma}_1 = (\gamma_{1,1}, \ldots, \gamma_{1,n_2})'$, where $\gamma_{1,i} \neq \gamma_{1,j}$ a.s., $i \neq j$. Let $E(\boldsymbol{\gamma}_1) = 0$ and assume that $D(\boldsymbol{\gamma}_1)$ has the structure specified in (4). The following conditions are equivalent:*
(i) $\mathbf{1}'_{n_2}\gamma_1 = 0$ *a.s. ("sum-to-zero" restriction),*
(ii) $\boldsymbol{\Sigma}_1 = \frac{\sigma_1 n_2}{n_2-1}(\boldsymbol{I}_{n_2} - \frac{1}{n_2}\boldsymbol{J}_{n_2})$, *where* $\sigma_1 = D(\boldsymbol{\gamma}_{1,i}), i = 1, \ldots, n_2,$
(iii) $n_1\sigma_1 + n_1(n_2-1)\sigma_2 = 0.$

In the following theorem, equivalent conditions obtained from imposing two restrictions on the spectrum of $\boldsymbol{\Sigma}_2$ are presented. By putting different restrictions on the spectrum of $\boldsymbol{\Sigma}_2$, the corresponding eigenvectors provide us the reparametrization of $\boldsymbol{\gamma}_2$.

Theorem 3.3. *Let $\boldsymbol{\gamma}_2 = (\gamma_{2,11}, \dots, \gamma_{2,n_2n_1})'$ be the nested factor in model (1) with $E(\boldsymbol{\gamma}_2) = 0$ and $D(\boldsymbol{\gamma}_2) = \boldsymbol{\Sigma}_2$ as defined in (3). Let $\boldsymbol{v}$: $n_2 \times 1$ be any non-zero vector satisfying $\boldsymbol{v}'\mathbf{1}_{n_2} = 0$, and let $\{\boldsymbol{v}_j\}$ be the eigenvectors corresponding to r distinct eigenvalues of $\boldsymbol{\Sigma}^{(h)}$ specified in (5), $h = 1, 2$, and $\boldsymbol{v}_j \neq \mathbf{1}_{n_1}$. Assume that $\gamma_{2,ij} \neq \gamma_{2,kj}$ for all j and $\gamma_{2,ij} \neq \gamma_{2,is}$ for all i a.s. Then the following conditions hold:*
(i) $\mathbf{1}_p'\gamma_2 = 0$ *and* $(\boldsymbol{v} \otimes \mathbf{1}_{n_1})'\gamma_2 = 0$ *a.s., iff* $\lambda_{11}^{\boldsymbol{\Sigma}_2} = 0$ *and* $\lambda_{21}^{\boldsymbol{\Sigma}_2} = 0$,
(ii) $(\boldsymbol{v} \otimes \mathbf{1}_{n_1})'\boldsymbol{\gamma}_2 = 0$ *and* $(\boldsymbol{v}^{k-1} \otimes \boldsymbol{v}_j)'\boldsymbol{\gamma}_2 = 0$ *a.s., iff* $\lambda_{21}^{\boldsymbol{\Sigma}_2} = 0$ *and* $\lambda_{kj}^{\boldsymbol{\Sigma}_2} = 0$, $k \in \{1, 2\}$, $j \in \{2, \dots, r\}$,
(iii) $(\boldsymbol{v}^{i-1} \otimes \mathbf{1}_{n_1})'\boldsymbol{\gamma}_2 = 0$ *and* $(\boldsymbol{v}^{i'-1} \otimes \boldsymbol{v}_j)'\boldsymbol{\gamma}_2 = 0$ *a.s., iff* $\lambda_{i1}^{\boldsymbol{\Sigma}_2} = 0$ *and* $\lambda_{i'j}^{\boldsymbol{\Sigma}_2} = 0$, $i = 1, 2$, $i' \in \{1, 2\}$, $j \in \{2, \dots, r\}$.

Proof. Let us show that (i) holds. Using (i) we get

$$\mathbf{0} = D\left(\mathbf{1}_p'\gamma_2\right) = D\left[(\boldsymbol{v}' \otimes \mathbf{1}_{n_1}')\boldsymbol{\gamma}_2\right],$$

and the structure of $\boldsymbol{\Sigma}$ given in (3) implies

$$\mathbf{1}_p'\left[\boldsymbol{I}_{n_2} \otimes \boldsymbol{\Sigma}^{(1)} + (\boldsymbol{J}_{n_2} - \boldsymbol{I}_{n_2}) \otimes \boldsymbol{\Sigma}^{(2)}\right]\mathbf{1}_p = \mathbf{0}, \tag{17}$$

$$(\boldsymbol{v}' \otimes \mathbf{1}_{n_1}')\left[\boldsymbol{I}_{n_2} \otimes \boldsymbol{\Sigma}^{(1)} + (\boldsymbol{J}_{n_2} - \boldsymbol{I}_{n_2}) \otimes \boldsymbol{\Sigma}^{(2)}\right](\boldsymbol{v} \otimes \mathbf{1}_{n_1}) = \mathbf{0}. \tag{18}$$

From (17) and (18), we have

$$\begin{aligned}
&n_2\mathbf{1}_{n_1}'\boldsymbol{\Sigma}^{(1)}\mathbf{1}_{n_1} + (n_2^2 - n_2)\mathbf{1}_{n_1}'\boldsymbol{\Sigma}^{(2)}\mathbf{1}_{n_1} = 0,\\
&\boldsymbol{v}'\boldsymbol{v}(\mathbf{1}_{n_1}'\boldsymbol{\Sigma}^{(1)}\mathbf{1}_{n_1} - \mathbf{1}_{n_1}'\boldsymbol{\Sigma}^{(2)}\mathbf{1}_{n_1}) = 0,
\end{aligned}$$

which are equivalent to

$$\begin{aligned}
\mathbf{1}_{n_1}'\boldsymbol{\Sigma}^{(1)}\mathbf{1}_{n_1} &= -(n_2 - 1)\mathbf{1}_{n_1}'\boldsymbol{\Sigma}^{(2)}\mathbf{1}_{n_1},\\
\mathbf{1}_{n_1}'\boldsymbol{\Sigma}^{(1)}\mathbf{1}_{n_1} &= \mathbf{1}_{n_1}'\boldsymbol{\Sigma}^{(2)}\mathbf{1}_{n_1}.
\end{aligned}$$

From the spectral decomposition, we have

$$\begin{aligned}
\lambda_{11}^{\boldsymbol{\Sigma}_2} &= \frac{1}{n_1}\mathbf{1}_{n_1}'\boldsymbol{\Sigma}^{(1)}\mathbf{1}_{n_1} + \frac{1}{n_1}(n_2 - 1)\mathbf{1}_{n_1}'\boldsymbol{\Sigma}^{(2)}\mathbf{1}_{n_1} = 0,\\
\lambda_{21}^{\boldsymbol{\Sigma}_2} &= \frac{1}{n_1}\mathbf{1}_{n_1}'\boldsymbol{\Sigma}^{(1)}\mathbf{1}_{n_1} - \frac{1}{n_1}\mathbf{1}_{n_1}'\boldsymbol{\Sigma}^{(2)}\mathbf{1}_{n_1} = 0,
\end{aligned}$$

which completes the proof of the first implication in (i).
Now, the condition $\lambda_{11}^{\boldsymbol{\Sigma}_2} = 0$ and $\lambda_{21}^{\boldsymbol{\Sigma}_2} = 0$ in (i) implies that, the eigenvectors corresponding to the zero eigenvalues are, $\boldsymbol{v}_1 = \mathbf{1}_p$ and $\boldsymbol{v}_2 = \boldsymbol{v} \otimes \mathbf{1}_{n_1}$. Since $E(\boldsymbol{\gamma}_2) = 0$, we have $E(\boldsymbol{v}_1'\boldsymbol{\gamma}_2) = 0$ and $E(\boldsymbol{v}_2'\boldsymbol{\gamma}_2) = 0$. Moreover, $D(\boldsymbol{v}_1'\boldsymbol{\gamma}_2) = \boldsymbol{v}_1' D(\boldsymbol{\gamma}_2)\boldsymbol{v}_1 = 0$ and $D(\boldsymbol{v}_2'\boldsymbol{\gamma}_2) = \boldsymbol{v}_2' D(\boldsymbol{\gamma}_2)\boldsymbol{v}_2 = 0$. Thus, $\boldsymbol{v}_1'\boldsymbol{\gamma}_2 = \boldsymbol{v}_2'\boldsymbol{\gamma}_2 = 0$.

Next, we prove (ii). It follows from the condition (ii) that

$$\mathbf{0} = D\left[(\boldsymbol{v}' \otimes \mathbf{1}_{n_1}')\boldsymbol{\gamma}_2\right] = D\left[(\boldsymbol{v}^{k-1\prime} \otimes \mathbf{1}_{n_1}')\boldsymbol{\gamma}_2\right],$$

and for $j \in \{2, \dots, r\}$, we have

$$\begin{aligned} &\mathbf{1}_{n_1}'\boldsymbol{\Sigma}^{(1)}\mathbf{1}_{n_1} - \mathbf{1}_{n_1}'\boldsymbol{\Sigma}^{(2)}\mathbf{1}_{n_1} = 0, \\ &\boldsymbol{v}_j'\boldsymbol{\Sigma}^{(1)}\boldsymbol{v}_j + (n_2 - 1)\boldsymbol{v}_j'\boldsymbol{\Sigma}^{(2)}\boldsymbol{v}_j = 0, \quad \text{if } k = 1, \\ &\boldsymbol{v}_j'\boldsymbol{\Sigma}^{(1)}\boldsymbol{v}_j - \boldsymbol{v}_j'\boldsymbol{\Sigma}^{(2)}\boldsymbol{v}_j = 0, \quad \text{if } k = 2. \end{aligned}$$

Thus, $\lambda_{21}^{\boldsymbol{\Sigma}_2} = 0$ and $\lambda_{kj}^{\boldsymbol{\Sigma}_2} = 0$, $k \in \{1, 2\}$, $j \in \{2, \dots, r\}$, hold.
To complete the proof of (ii), suppose that $\lambda_{21}^{\boldsymbol{\Sigma}_2} = 0$ and $\lambda_{kj}^{\boldsymbol{\Sigma}_2} = 0$, $k \in \{1, 2\}$, $j \in \{2, \dots, r\}$. The eigenvectors corresponding to the zero eigenvalues are $\boldsymbol{v}_1 = \boldsymbol{v} \otimes \mathbf{1}_{n_1}$ and $\boldsymbol{v}_2 = \boldsymbol{v}^{k-1} \otimes \boldsymbol{v}_j$, $k \in \{1, 2\}$, $j \in \{2, \dots, r\}$. Since $E(\boldsymbol{v}_i'\boldsymbol{\gamma}_2) = 0$ and $D(\boldsymbol{v}_i'\boldsymbol{\gamma}_2) = 0, i = 1, 2$, we have $\boldsymbol{v}_i'\boldsymbol{\gamma}_2 = 0, i = 1, 2$. This completes the proof of (ii).

Finally, we prove the statement in (iii). The condition (iii) implies that

$$\begin{aligned} &\mathbf{1}_{n_1}'\boldsymbol{\Sigma}^{(1)}\mathbf{1}_{n_1} + (n_2 - 1)\mathbf{1}_{n_1}'\boldsymbol{\Sigma}^{(2)}\mathbf{1}_{n_1} = 0, \\ &\mathbf{1}_{n_1}'\boldsymbol{\Sigma}^{(1)}\mathbf{1}_{n_1} - \mathbf{1}_{n_1}'\boldsymbol{\Sigma}^{(2)}\mathbf{1}_{n_1} = 0, \\ &\boldsymbol{v}_j'\boldsymbol{\Sigma}^{(1)}\boldsymbol{v}_j + (n_2 - 1)\boldsymbol{v}_j'\boldsymbol{\Sigma}^{(2)}\boldsymbol{v}_j = 0, \quad \text{if } i' = 1, \\ &\boldsymbol{v}_j'\boldsymbol{\Sigma}^{(1)}\boldsymbol{v}_j - \boldsymbol{v}_j'\boldsymbol{\Sigma}^{(2)}\boldsymbol{v}_j = 0, \quad \text{if } i' = 2, \end{aligned}$$

which is followed by $\lambda_{i1}^{\boldsymbol{\Sigma}_2} = 0$ and $\lambda_{i'j}^{\boldsymbol{\Sigma}_2} = 0$, $i = 1, 2$, $i' \in \{1, 2\}$, $j \in \{2, \dots, r\}$.
On the other hand, for the eigenvalues $\lambda_{i1}^{\boldsymbol{\Sigma}_2} = 0$ and $\lambda_{i'j}^{\boldsymbol{\Sigma}_2} = 0$, $i = 1, 2$, $i' \in \{1, 2\}$, $j \in \{2, \dots, r\}$, the corresponding eigenvectors are $\boldsymbol{v}_1 = \boldsymbol{v}^{i-1} \otimes \mathbf{1}_{n_1}$ and $\boldsymbol{v}_2 = \boldsymbol{v}^{i'-1} \otimes \boldsymbol{v}_j$. It is straightforward to verify that $\boldsymbol{v}_i = 0$, $i = 1, 2$. ■

Remark 3.2. Consider the condition (i). If $\boldsymbol{v}$ is any non-zero orthonormal vector such that $\boldsymbol{v}'\mathbf{1}_{n_2} = 0$, it implies that $(\frac{1}{\sqrt{n_2}}\mathbf{1}_{n_2} \otimes \frac{1}{\sqrt{n_1}}\mathbf{1}_{n_1})'\boldsymbol{\gamma}_2 = 0$ and $(\boldsymbol{v} \otimes \frac{1}{\sqrt{n_1}}\boldsymbol{I}_{n_1})'\boldsymbol{\gamma}_2 = 0$. Premultiplying theses expressions with $(\frac{1}{\sqrt{n_2}}\mathbf{1}_{n_2} \otimes \sqrt{n_1}\boldsymbol{I}_{n_1})$ and $(\boldsymbol{v} \otimes \sqrt{n_1}\boldsymbol{I}_{n_1})$, respectively, we obtain $(\frac{1}{n_2}\boldsymbol{J}_{n_2} \otimes \mathbf{1}_{n_1}')\boldsymbol{\gamma}_2 = 0$ and $(\boldsymbol{v}\boldsymbol{v}' \otimes \mathbf{1}_{n_1}')\boldsymbol{\gamma}_2 = 0$, which yields $(\boldsymbol{I}_{n_2} \otimes \mathbf{1}_{n_1}')\boldsymbol{\gamma}_2 = 0$. This result provides the condition which leads to the classical "sum-to-zero" reparametrization for $\boldsymbol{\gamma}_2$.

4. Explicit MLEs of (co)variance parameters

In the previous section, different restrictions were described in order to derive estimable (co)variance components. The results of the previous section will now be applied to model (1). We first state the estimability conditions in the following theorem. In the rest of this section, different constraints provided in Theorem 4.1 will be illustrated.

Theorem 4.1. *Let* $\boldsymbol{v}$: $n_2 \times 1$ *be any non-zero vector satisfying* $\boldsymbol{v}'\mathbf{1}_{n_2}= 0$*, and let* $\{\boldsymbol{v}_j\}$ *be the eigenvectors corresponding to* r *distinct eigenvalues of* $\boldsymbol{\Sigma}^{(h)}$ *specified in (5),* $h = 1, 2$ *and* $\boldsymbol{v}_j \neq \mathbf{1}_{n_1}$*. Model (1) has explicit MLEs for* $\boldsymbol{\theta}$ *if one of the following conditions hold:*
In Scenario 1,
(i) $\mathbf{1}'_{n_2}\boldsymbol{\gamma}_1 = 0$, $\mathbf{1}'_p\boldsymbol{\gamma}_2 = 0$ *and* $(\boldsymbol{v} \otimes \mathbf{1}_{n_1})'\boldsymbol{\gamma}_2 = 0$,
(ii) $\mathbf{1}'_{n_2}\boldsymbol{\gamma}_1 = 0$, $(\boldsymbol{v} \otimes \mathbf{1}_{n_1})'\boldsymbol{\gamma}_2 = 0$ *and* $(\boldsymbol{v}^{k-1} \otimes \boldsymbol{v}_j)'\boldsymbol{\gamma}_2 = 0$, $k \in \{1, 2\}$, $j \in \{2, \ldots, r\}$,
In Scenario 2,
(iii) $(\boldsymbol{v}^{i-1} \otimes \mathbf{1}_{n_1})'\boldsymbol{\gamma}_2 = 0$ *and* $(\boldsymbol{v}^{i'-1} \otimes \boldsymbol{v}_j)'\boldsymbol{\gamma}_2 = 0$, $i = 1, 2$, $i' \in \{1, 2\}$, $j \in \{2, \ldots, r\}$.

Proof. In Section 3, it has been shown that conditions (i)-(iii) are equivalent to the imposing restrictions on the parameter $\boldsymbol{\theta}$, i.e. $\boldsymbol{K}_i\boldsymbol{\theta} = \mathbf{0}$, $i = 1, \ldots, 3$. According to (7), the equality $\boldsymbol{\eta} = \boldsymbol{L}\boldsymbol{\theta}$ with the restriction $\boldsymbol{K}_i\boldsymbol{\theta} = \mathbf{0}$, is equivalent to $\boldsymbol{\theta} = (\boldsymbol{K}'_i)^o\boldsymbol{\theta}^*$, where $(\boldsymbol{K}'_i)^o$ is a matrix which columns generate the orthogonal complement to the column vector space of $\boldsymbol{K}'_i$ and $\boldsymbol{\theta}^* : 2r \times 1$ is the vector of unknown parameters in model (1) which is determined by restrictions (i)-(iii). If $\boldsymbol{L}(\boldsymbol{K}'_i)^o$ is invertible, then $\boldsymbol{\theta}^*$ can be estimated. Since $r(\boldsymbol{L}(\boldsymbol{K}'_i)^o) = r(\boldsymbol{L} \vdots \boldsymbol{K}'_i) - r(\boldsymbol{K}'_i)$, and $r(\boldsymbol{K}'_i) = 3$, where $r(\bullet)$ denotes matrix rank. Due to the structure of $\boldsymbol{L}$ and $\boldsymbol{K}_i$, given in (7) and (14)-(16), respectively, we find that $r(\boldsymbol{L} \vdots \boldsymbol{K}'_i) = 2r + 3$, i.e. $\boldsymbol{L}(\boldsymbol{K}'_i)^o$ is of full rank and therefore invertible. Thus, $\boldsymbol{\theta}^* = (\boldsymbol{L}(\boldsymbol{K}'_i)^o)^{-1}\boldsymbol{\eta}$. ■

Remark 4.2. Consider the reparametrization $(\mathbf{1}'_{n_2} \otimes \boldsymbol{I}_{n_1})\boldsymbol{\gamma}_2 = 0$. In this case,

$$\begin{aligned}\mathbf{0} &= D\left[(\mathbf{1}'_{n_2} \otimes \boldsymbol{I}_{n_1})\boldsymbol{\gamma}_2\right] \\ &= (\mathbf{1}'_{n_2} \otimes \boldsymbol{I}_{n_1})\left[\boldsymbol{I}_{n_2} \otimes \boldsymbol{\Sigma}^{(1)} + (\boldsymbol{J}_{n_2} - \boldsymbol{I}_{n_2}) \otimes \boldsymbol{\Sigma}^{(2)}\right](\mathbf{1}_{n_2} \otimes \boldsymbol{I}_{n_1}) \\ &= \boldsymbol{\Sigma}^{(1)} + (n_2 - 1)\boldsymbol{\Sigma}^{(2)}. \end{aligned} \quad (19)$$

Hence, $\boldsymbol{\Sigma}^{(1)} = -(n_2 - 1)\boldsymbol{\Sigma}^{(2)}$. Let $\boldsymbol{\tau} = (\boldsymbol{\tau}'_1, \boldsymbol{\tau}'_2)'$ be the vector of the unknown parameters in $\boldsymbol{\Sigma}_2$, where $\boldsymbol{\tau}_1 = (\tau_1, \ldots, \tau_r)'$ and $\boldsymbol{\tau}_2 = (\tau_{r+1}, \ldots, \tau_{2r})'$ are the unknown parameters in $\boldsymbol{\Sigma}^{(1)}$ and $\boldsymbol{\Sigma}^{(2)}$ given in (5), respectively. The equality $\boldsymbol{\Sigma}^{(1)} = -(n_2 - 1)\boldsymbol{\Sigma}^{(2)}$ implies that $\boldsymbol{\tau}_2 = -\frac{1}{n_2-1}\boldsymbol{\tau}_1$. Thus, $\boldsymbol{B}_2\boldsymbol{\tau}$ with $\boldsymbol{B}_2$ defined in Theorem 2.1, becomes

$$\boldsymbol{B}_2\boldsymbol{\tau} = \begin{pmatrix} \boldsymbol{A} & (n_2-1)\boldsymbol{A} \\ \boldsymbol{A} & -\boldsymbol{A} \end{pmatrix} \begin{pmatrix} \boldsymbol{\tau}_1 \\ -\frac{1}{n_2-1}\boldsymbol{\tau}_1 \end{pmatrix} = \begin{pmatrix} \mathbf{0}_r \\ \frac{n_2}{n_2-1}\boldsymbol{A}\boldsymbol{\tau}_1 \end{pmatrix},$$

and $\lambda_{1j}^{\boldsymbol{\Sigma}_2} = 0$, $j = 1, \dots, r$. In this case, s_1 and s_2 defined in Proposition 3.1 are the following:

$$s_1 = \begin{cases} r+1, & \text{with the condition } \mathbf{1}_p'\boldsymbol{\gamma}_2 = 0, \\ r+2, & \text{otherwise }, \end{cases}$$
$$s_2 = \begin{cases} r+2, & \text{with the condition } \mathbf{1}_p'\boldsymbol{\gamma}_2 = 0, \\ r+3, & \text{otherwise }. \end{cases}$$

Therefore, the sufficient condition for obtaining the explicit MLEs of the variance components in Szatrowski (1980) does not hold.
In the next example, we illustrate the obtained results concerning the explicit estimators of the parameters in model (1).

Example

Let $n_1 = 4$. In this case for the model (1) we have

$$\begin{aligned} \mathbf{y} &\sim N_{4n_2}(\mu \mathbf{1}_{4n_2}, \boldsymbol{\Sigma}), \\ \boldsymbol{\Sigma} &= \boldsymbol{Z}_1\boldsymbol{\Sigma}_1\boldsymbol{Z}_1' + \boldsymbol{\Sigma}_2 + \sigma^2\boldsymbol{I}_{4n_2}, \end{aligned}$$

where $\boldsymbol{\Sigma}_1 = \sigma_1\boldsymbol{I}_{n_2} + \sigma_2(\boldsymbol{J}_{n_2} - \boldsymbol{I}_{n_2})$, $\boldsymbol{\Sigma}_2 = \boldsymbol{I}_{n_2} \otimes \boldsymbol{\Sigma}^{(1)} + (\boldsymbol{J}_{n_2} - \boldsymbol{I}_{n_2}) \otimes \boldsymbol{\Sigma}^{(2)}$, and

$$\boldsymbol{\Sigma}^{(1)} = \left(\begin{array}{cc|cc} \tau_1 & \tau_2 & \tau_3 & \tau_2 \\ \tau_2 & \tau_1 & \tau_2 & \tau_3 \\ \hline \tau_3 & \tau_2 & \tau_1 & \tau_2 \\ \tau_2 & \tau_3 & \tau_2 & \tau_1 \end{array}\right), \quad \boldsymbol{\Sigma}^{(2)} = \left(\begin{array}{cc|cc} \tau_4 & \tau_5 & \tau_6 & \tau_5 \\ \tau_5 & \tau_4 & \tau_5 & \tau_6 \\ \hline \tau_6 & \tau_5 & \tau_4 & \tau_5 \\ \tau_5 & \tau_6 & \tau_5 & \tau_4 \end{array}\right).$$

Using Theorem 2.1, the distinct eigenvalues of $\boldsymbol{\Sigma}_2$ are given by $\boldsymbol{\eta} = \boldsymbol{L}\boldsymbol{\theta}$, where $\boldsymbol{\theta} = (\sigma^2, \sigma_1, \sigma_2, \tau_1, \tau_2, \tau_3, \tau_4, \tau_5, \tau_6)'$ and

$$\boldsymbol{L} = \left(\begin{array}{ccc|cccccc} 1 & 4 & 4\,(n_2-1) & 1 & 2 & 1 & n_2-1 & 2\,(n_2-1) & n_2-1 \\ 1 & 0 & 0 & 1 & 0 & -1 & n_2-1 & 0 & -\,(n_2-1) \\ 1 & 0 & 0 & 1 & -2 & 1 & n_2-1 & -2\,(n_2-1) & n_2-1 \\ \hline 1 & 4 & -4 & 1 & 2 & 1 & -1 & -2 & -1 \\ 1 & 0 & 0 & 1 & 0 & -1 & -1 & 0 & 1 \\ 1 & 0 & 0 & 1 & -2 & 1 & -1 & 2 & -1 \end{array}\right).$$

According to Proposition 3.1, the number of distinct eigenvalues of $\boldsymbol{\Sigma}$ equals to 6, and the number of unknown parameters in $\boldsymbol{\Sigma}$ is 9. Thus, as noted previously, the model is overparametrized. Let now apply the restriction (i) from Scenario 1, i.e. $\lambda_1^{\boldsymbol{\Sigma}_1}=0$, $\lambda_{11}^{\boldsymbol{\Sigma}_2}=0$ and $\lambda_{21}^{\boldsymbol{\Sigma}_2}=0$. In this case, $\sigma_1=-(n_2-1)\sigma_2$, $\tau_1=-2\tau_2-\tau_3$ and $\tau_4=-2\tau_5-\tau_6$. For this restricted model, the vector of unknown parameters is the following. $\boldsymbol{\theta}^*=(\sigma^2,\sigma_2,\tau_2,\tau_3,\tau_5,\tau_6)'$. The vector of the distinct eigenvalues of $\boldsymbol{\Sigma}$, $\boldsymbol{\eta}$, can be written in the form $\boldsymbol{\eta}=\boldsymbol{L}(\boldsymbol{K}_1')^o\boldsymbol{\theta}^*$, where

$$
(\boldsymbol{K}_1')^o=\begin{pmatrix}
\boxed{1} & & & \\
 & \boxed{\begin{array}{c}-(n_2-1)\\ 1\end{array}} & & 0\\
 & & \boxed{\begin{array}{cc}-2 & -1\\ \multicolumn{2}{c}{\boldsymbol{I}_2}\end{array}} & \\
0 & & & \boxed{\begin{array}{cc}-2 & -1\\ \multicolumn{2}{c}{\boldsymbol{I}_2}\end{array}}
\end{pmatrix},
$$

which yields

$$
\begin{pmatrix}\eta_1\\ \eta_2\\ \eta_3\\ \eta_4\\ \eta_5\\ \eta_6\end{pmatrix}=
\begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0\\
1 & 0 & -2 & -2 & 2(1-n_2) & 2(1-n_2)\\
1 & 0 & -4 & 0 & 4(1-n_2) & 0\\
1 & -4n_2 & 0 & 0 & 0 & 0\\
1 & 0 & -2 & -2 & 2 & 2\\
1 & 0 & -4 & 0 & 4 & 0
\end{pmatrix}
\begin{pmatrix}\sigma^2\\ \sigma_2\\ \tau_2\\ \tau_3\\ \tau_5\\ \tau_6\end{pmatrix}.
$$

The explicit MLE of $\boldsymbol{\theta}^*$ equals

$$
\hat{\boldsymbol{\theta}}^*=\begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0\\
\frac{1}{4n_2} & 0 & 0 & -\frac{1}{4n_2} & 0 & 0\\
\frac{1}{4} & 0 & -\frac{1}{4n_2} & 0 & 0 & \frac{1}{4}\left(\frac{1}{n_2}-1\right)\\
\frac{1}{4} & -\frac{1}{2n_2} & \frac{1}{4n_2} & 0 & \frac{1}{2}\left(\frac{1}{n_2}-1\right) & \frac{1}{4}(1-\frac{1}{4n_2})\\
0 & 0 & -\frac{1}{4n_2} & 0 & 0 & \frac{1}{4n_2}\\
0 & -\frac{1}{2n_2} & \frac{1}{4n_2} & 0 & \frac{1}{2n_2} & -\frac{1}{4n_2}
\end{pmatrix}\hat{\boldsymbol{\eta}},
$$

and $\hat{\boldsymbol{\eta}}$ is given in (10)-(11). Namely, $\hat{\theta}_1^*=\hat{\sigma}^2=\hat{\eta}_1$.
Let us now apply the restriction (ii) in Scenario 1, for example, $\lambda_1^{\boldsymbol{\Sigma}_1}=0$, $\lambda_{21}^{\boldsymbol{\Sigma}_2}=0$ and $\lambda_{22}^{\boldsymbol{\Sigma}_2}=0$. Now, $\sigma_1=-(n_2-1)\sigma_2$, $\tau_4=\tau_1-\tau_3+\tau_6$ and $\tau_5=\tau_2+\tau_3-\tau_6$. The vector of unknown parameters is now $\boldsymbol{\theta}^*=(\sigma^2,\sigma_2,\tau_1,\tau_2,\tau_3,\tau_6)'$. The vector of the distinct eigenvalues of $\boldsymbol{\Sigma}$, $\boldsymbol{\eta}$, can be expressed with the help of

$$
(\boldsymbol{K}_2')^o=\begin{pmatrix}
\boxed{1} & & & \\
 & \boxed{\begin{array}{c}-(n_2-1)\\ 1\end{array}} & & 0\\
 & & \boxed{\boldsymbol{I}_3} & \\
 & & \boxed{\begin{array}{cccc}1 & 0 & 1 & 1\\ 0 & 1 & 1 & -1\end{array}} & \\
0 & & & \boxed{1}
\end{pmatrix},
$$

as $\boldsymbol{\eta} = \boldsymbol{L}(\boldsymbol{K}_2')^o\boldsymbol{\theta}^*$, or

$$\begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \\ \eta_5 \\ \eta_6 \end{pmatrix} = \begin{pmatrix} 1 & 0 & n_2 & 2n_2 & n_2 & 0 \\ 1 & 0 & n_2 & 0 & -n_2 & 0 \\ 1 & 0 & n_2 & -2n_2 & 4-3n_2 & 4n_2-4 \\ 1 & -4n_2 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 4 & -4 \end{pmatrix} \begin{pmatrix} \sigma^2 \\ \sigma_2 \\ \tau_1 \\ \tau_2 \\ \tau_3 \\ \tau_6 \end{pmatrix}.$$

The explicit MLE of $\boldsymbol{\theta}^*$ is obtained as follows

$$\hat{\boldsymbol{\theta}}^* = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -\frac{1}{4n_2} & \frac{1}{4n_2} & 0 \\ \frac{1}{4n_2} & \frac{1}{2n_2} & \frac{1}{4n_2} & 0 & -\frac{3+n_2}{4n_2} & \frac{1}{4}(1-\frac{1}{n_2}) \\ \frac{1}{4n_2} & 0 & -\frac{1}{4n_2} & 0 & \frac{1}{4}(1-\frac{1}{n_2}) & \frac{1}{4}\left(\frac{1}{n_2}-1\right) \\ \frac{1}{4n_2} & -\frac{1}{2n_2} & \frac{1}{4n_2} & 0 & \frac{1}{4}\left(\frac{1}{n_2}-1\right) & \frac{1}{4}(1-\frac{1}{n_2}) \\ \frac{1}{4n_2} & -\frac{1}{2n_2} & \frac{1}{4n_2} & 0 & \frac{1}{4n_2} & -\frac{1}{4n_2} \end{pmatrix} \hat{\boldsymbol{\eta}},$$

where $\hat{\boldsymbol{\eta}}$ is given in (10)-(11).
Suppose condition (iii) in Scenario 2 is used. For example, $\lambda_{11}^{\boldsymbol{\Sigma}_2} = 0$, $\lambda_{12}^{\boldsymbol{\Sigma}_2} = 0$ and $\lambda_{21}^{\boldsymbol{\Sigma}_2} = 0$. This implies that, $\tau_1 = -2\tau_2 - \tau_3$, $\tau_4 = -\frac{2}{1-n_2}\tau_2 - \frac{2}{1-n_2}\tau_3 + \tau_6$ and $\tau_5 = \frac{1}{1-n_2}\tau_2 + \frac{1}{1-n_2}\tau_3 - \tau_6$, and $\boldsymbol{\theta}^* = (\sigma^2, \sigma_1, \sigma_2, \tau_2, \tau_3, \tau_6)'$. Now, the equality of $\boldsymbol{\eta} = \boldsymbol{L}\boldsymbol{\theta}$, can be written as $\boldsymbol{\eta} = \boldsymbol{L}(\boldsymbol{K}_3')^o\boldsymbol{\theta}^*$, where

$$(\boldsymbol{K}_3')^o = \begin{pmatrix} \boxed{1} & & & \mathbf{0} \\ & \boxed{\boldsymbol{I}_2} & & \\ & & \begin{array}{|cc|c} \hline -2 & -1 & \\ \multicolumn{2}{|c|}{\boldsymbol{I}_2} & \\ \hline -\frac{2}{1-n_2} & -\frac{2}{1-n_2} & 1 \\ \frac{1}{1-n_2} & \frac{1}{1-n_2} & -1 \\ \hline \end{array} & \\ \mathbf{0} & & & \boxed{1} \end{pmatrix},$$

and we obtain

$$\begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \\ \eta_5 \\ \eta_6 \end{pmatrix} = \begin{pmatrix} 1 & 4 & 5n_2-5 & 0 & 0 & 0 \\ 1 & 0 & n_2-1 & 0 & 0 & 0 \\ 1 & 0 & n_2-1 & 0 & 4 & 4n_2-4 \\ 1 & 4 & -5 & 0 & 0 & 0 \\ 1 & 0 & -1 & \frac{2n_2}{1-n_2} & \frac{2n_2}{1-n_2} & 0 \\ 1 & 0 & -1 & \frac{4n_2}{1-n_2} & \frac{4}{1-n_2} & -4 \end{pmatrix} \begin{pmatrix} \sigma^2 \\ \sigma_1 \\ \sigma_2 \\ \tau_2 \\ \tau_3 \\ \tau_6 \end{pmatrix}.$$

The explicit MLE of $\boldsymbol{\theta}^*$ is obtained as

$$\hat{\boldsymbol{\theta}}^* = \begin{pmatrix} \frac{1}{5}\left(\frac{1}{n_2}-1\right) & 1 & 0 & \frac{1}{5}(1-\frac{1}{n_2}) & 0 & 0 \\ \frac{1}{20}+\frac{1}{5n_2} & -\frac{1}{4} & 0 & \frac{1}{5}(1-\frac{1}{n_2}) & 0 & 0 \\ \frac{1}{5n_2} & 0 & 0 & -\frac{1}{5n_2} & 0 & 0 \\ \frac{1}{20}\left(\frac{1}{n_2}-1\right) & \frac{1}{4} & -\frac{1}{4n_2} & \frac{n_2-1}{20n_2} & 0 & \frac{1}{4}\left(\frac{1}{n_2}-1\right) \\ \frac{1}{20}\left(\frac{1}{n_2}-1\right) & \frac{1}{4}-\frac{1}{2n_2} & \frac{1}{4n_2} & \frac{n_2-1}{20n_2} & \frac{1}{2}\left(\frac{1}{n_2}-1\right) & \frac{1}{4}(1-\frac{1}{n_2}) \\ \frac{1}{20n_2} & -\frac{1}{2n_2} & \frac{1}{4n_2} & -\frac{1}{20n_2} & \frac{1}{2n_2} & -\frac{1}{4n_2} \end{pmatrix} \hat{\boldsymbol{\eta}},$$

where $\hat{\boldsymbol{\eta}}$ is given by (10)-(11).

5. Discussion

In order to get explicit MLE for all (co)variance parameters in model (1), certain restrictions should be imposed on the covariance structure $\boldsymbol{\Sigma}$, or equivalent on the unknown covariance parameters $\boldsymbol{\theta}$. Choosing appropriate restrictions on the parameters $\boldsymbol{\theta}$ for model (1) may sometimes be a difficult task. One reason is that the existence of an explicit estimator for the mean parameter (μ) should not be affected in the restricted models. Imposing restrictions on the spectra of the covariance matrices is beneficial from a $\boldsymbol{\Sigma}$-invariant point of view, since in this case we can find the restrictions which can preserve the block structure of $\boldsymbol{\Sigma}$, and then the explicit estimator of μ exists. Even for more complicated mean structures (for example, bilinear structures), it is still possible to find out what kind of design matrices can lead the existence of explicit estimates of mean parameters.

Seely (1971) proved that Jordan algebra implies that there exists an optimal unbiased estimator (uniform minimum variance quadratic unbiased estimator) for any linear combination of variance components. We may show that the set of block circular symmetric matrices is a Jordan algebra. It can also be shown that not all restrictions on $\boldsymbol{\theta}$ can preserve the block circular structure. For example, a "set-to-zero" reparameterization, i.e., $\gamma_{2,p} = 0$, will not preserve the structure of $\boldsymbol{\Sigma}$.

An interesting finding is (see Remark 3.2.) that the "sum-to-zero" reparametrization $(\boldsymbol{I}_{n_2} \otimes \mathbf{1}'_{n_1})\boldsymbol{\gamma}_2$=0 can give the explicit MLEs of $\boldsymbol{\theta}$. However, not all "sum-to-zero" conditions lead explicit MLEs of $\boldsymbol{\theta}$ in model (1). For example, there is another "sum-to-zero" condition, i.e. $(\mathbf{1}'_{n_2} \otimes \boldsymbol{I}_{n_1})\boldsymbol{\gamma}_2 = 0$ or equivalently $\sum_{i=1}^{n_2} \gamma_{2,ij} = 0$ a.s. for all $j = 1, \ldots, n_1$, without explicit MLEs. Moreover, it has been shown (see Remark 4.2.) that under this reparameterized model (1), the sufficient condition for the existence of the explicit MLEs of the variance components in Szatrowski (1980) does not hold.

References

Dawid, A. P. (1988). Symmetry models and hypotheses for structured data layouts. *Journal of the Royal Statistical Society. Series B*, **50**, 1–34.

Goldstein, H. (2003). *Multilevel statistical models*. Wiley, New York.

Goldstein, H. and Woodhouse, G. (2000). School effectiveness research and educational policy. *Oxford Review of Education*, **26**, 353–363.

Hox, J. J. and Kreft, I. G. (1994). Multilevel analysis methods. *Sociological Methods and Research*, **22**, 283–299.

Leiva, R. and Roy, A. (2010). Linear discrimination for multi-level multivariate data with separable means and jointly equicorrelated covariance structure. *Journal of Statistical Planning and Inference*, **141**, 1910–1924.

Liang, Y., von Rosen, D. and von Rosen, T. (2012). On estimation in multilevel models with block circular symmetric covariance structures. *Acta et Commentationes Universitatis Tartuensis de Mathematica*, **16**, 1–14.

Louden, C. and Roy, A. (2010). Classification rules under autoregressive and general circulant covariance. *Communications in Statistics-Theory and Methods*, **39**, 3294–3315.

Nahtman, T. (2006). Marginal permutation invariant covariance matrices with applications to linear models. *Linear Algebra and its Applications*, **417**, 183–210.

Nahtman, T. and von Rosen, D. (2008). Shift permutation invariance in linear random factor models. *Mathematical Methods of Statistics*, **17**, 173–185.

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

Pan, J. X. and Fang, K. T. (2002). *Growth curve models and statistical diagnostics*. Springer Verlag, New York.

Puntanen, S. and Styan, G. P. H. (1989). The equality of the ordinary least squares estimator and the best linear unbiased estimator. *American Statistician*, **43**, 153–161.

Raudenbush, S. W. (1988). Educational applications of hierarchical linear models: A review. *Journal of Educational and Behavioral Statistics*, **13**, 85–116.

von Rosen, T. (2011). On the inverse of certain block structured matrices generated by linear combinations of kronecter products. *Linear and Multilinear Algebra*, **59**, 595–606.

Seely, J. (1971). Quadratic subspaces and completeness. *The Annals of Mathematical Statistics*, **42**, 710–721.

Szatrowski, T. H. (1980). Necessary and sufficient conditions for explicit solutions in the multivariate normal estimation problem for patterned means and covariances. *The Annals of Statistics*, **8**, 802–810.