Ann Inst Stat Math (2015) 67:773–791
DOI 10.1007/s10463-014-0475-8

On estimation in hierarchical models with block circular covariance structures

Yuli Liang · Dietrich von Rosen · Tatjana von Rosen

Received: 16 February 2013 / Revised: 19 February 2014 / Published online: 31 July 2014

Abstract Hierarchical linear models with a block circular covariance structure are considered. Sufficient conditions for obtaining explicit and unique estimators for the variance–covariance components are derived. Different restricted models are discussed and maximum likelihood estimators are presented. The theory is illustrated through covariance matrices of small sizes and a real-life example.

Keywords Circular block symmetry · Estimation · Identifiability · Maximum likelihood estimator · Restricted model · Variance components

1 Introduction

Mixed linear models offer large flexibility in modeling data structures consisting of multiple levels of nested groups. These models are also known as multilevel models or hierarchical linear models (see e.g. Searle et al. 1992; Demidenko 2004). Hierarchically structured data naturally arise in various applications including sociology, education, biology and life sciences.

Y. Liang (✉) · T. von Rosen
Department of Statistics, Stockholm University, Universitetsvägen 10B,
10691 Stockholm, Sweden
e-mail: yuli.liang@stat.su.se

T. von Rosen
e-mail: Tatjana.vonRosen@stat.su.se

D. von Rosen
Department of Energy and Technology, Swedish University of Agricultural Sciences,
Gerda Nilssons väg 5, 75651 Uppsala, Sweden

D. von Rosen
Department of Mathematics, Linköping University, Campus Valla, 58183 Linköping, Sweden
e-mail: Dietrich.von.Rosen@slu.se

When modeling such data it is important to take into account the sources of variation present at different levels of the hierarchy (see Hox and Kreft 1994; Goldstein 2010). The presence of symmetry in the data at one or several levels yields a patterned dependence structure within or between the corresponding levels. Symmetry here means that dependency between neighboring units remains the same (invariant) after a re-arrangement of units. Properties of patterned covariance matrices in mixed linear models which are obtained under symmetry assumptions have, for example, been studied in Nahtman (2006), Nahtman and Rosen (2008), Rosen (2011) and Liang et al. (2012).

In this article the following hierarchical mixed linear model is considered:

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

where $\mathbf{y}$ is a $p \times 1$ vector of observations, $p = n_1 n_2$, μ is an unknown constant, $\boldsymbol{\gamma}_1 : n_2 \times 1$, $\boldsymbol{\gamma}_2 : p \times 1$ and $\boldsymbol{\epsilon}$ are independently normally distributed random vectors with zero means and variances–covariance matrices $\boldsymbol{\Sigma}_1 \geq 0$, $\boldsymbol{\Sigma}_2 \geq 0$, and $\sigma^2 \boldsymbol{I}_p$, respectively. Here $\boldsymbol{Z}_1 = \boldsymbol{I}_{n_2} \otimes \mathbf{1}_{n_1}$, $\boldsymbol{Z}_2 = \boldsymbol{I}_{n_2} \otimes \boldsymbol{I}_{n_1}$, $\mathbf{1}_s = (1, 1, \ldots, 1)'$, i.e. 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,

$$\begin{aligned} \mathbf{y} &\sim N_p(\mu \mathbf{1}_p, \boldsymbol{\Sigma}), \\ \boldsymbol{\Sigma} &= \boldsymbol{Z}_1 \boldsymbol{\Sigma}_1 \boldsymbol{Z}_1' + \boldsymbol{\Sigma}_2 + \sigma^2 \boldsymbol{I}_p. \end{aligned} \tag{2}$$

The covariance matrix $\boldsymbol{\Sigma}$ in (2) may have different structures depending on $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$. In this article, the covariance matrix $\boldsymbol{\Sigma}_1 : n_2 \times n_2$ is assumed to be compound symmetric, i.e.

$$\boldsymbol{\Sigma}_1 = \sigma_1 \boldsymbol{I}_{n_2} + \sigma_2 (\boldsymbol{J}_{n_2} - \boldsymbol{I}_{n_2}), \tag{3}$$

where σ_1 and σ_2 are unknown parameters, $\boldsymbol{J}_s = \mathbf{1}_s \mathbf{1}_s'$, and the covariance matrix $\boldsymbol{\Sigma}_2 : p \times p$, in (2), is assumed to have a block circular pattern:

$$\boldsymbol{\Sigma}_2 = \sum_{k=0}^{[n_1/2]} SC(k) \otimes \boldsymbol{\Sigma}^{k+1}, \tag{4}$$

where the notation $[\cdot]$ stands for the integer part, $\boldsymbol{\Sigma}^{k+1}$ is unstructured, and $SC(k) : n_1 \times n_1$ is defined in the following way:

$$(SC(k))_{ij} = \begin{cases} 1, & \text{if } |i - j| = k \text{ or } |i - j| = n_1 - k, \\ 0, & \text{otherwise}, \end{cases}$$

where $k \in \{1, \ldots, [n_1/2]\}$. For notational convenience denote $SC(0) = \boldsymbol{I}_{n_1}$.

The covariance structure defined in (4) has been studied by Olkin (1973). In this article, we additionally impose the compound symmetry structure on $\boldsymbol{\Sigma}^{k+1}$ in (4). To facilitate inference about parameters of interest, we will use the following expression of $\boldsymbol{\Sigma}_2$:

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

where $\boldsymbol{\Sigma}^{(h)}$ is a symmetric circular Toeplitz matrix, $h = 1, 2$. The equivalence of the expressions (4) and (5), when $\boldsymbol{\Sigma}^{k+1}$ is compound symmetric, is shown in Liang et al. (2011).

Note that the matrix $\boldsymbol{\Sigma}^{(h)} = (\sigma_{ij}^{(h)})$ in (5) depends on r, $r = [n_1/2]+1$, parameters, and for $i, j = 1, \ldots, n_1, \;\; h = 1, 2$,

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

where the τ_q's are unknown parameters, and taking into account that $h = 1, 2$, the index $q = 1, \ldots, 2r$.

Estimation of parameters in (1) faces several challenges, among others identifiability issues. For example, one cannot distinguish between τ_1 and σ^2. We will show that to obtain explicit maximum likelihood estimators (MLEs) with interpretable parameters, the only possibility is to consider restricted models, i.e. constraints on the elements of the factor(s) included in the model or, equivalently, on the corresponding covariance matrix, must be imposed.

Several different kinds of constraints exist that are sufficient to guarantee the identifiability of the covariance parameters which, however, yield different types of covariance matrices. In particular, certain natural ways of constraining (reparameterizing) a random factor that preserves the structure of the original covariance matrix will be outlined, among others invariance assumptions. Reparameterization issues have been considered, for example in VanLeeuwen (1997) and Nahtman (2006).

The organization of this article is as follows. In the next section, we provide some results concerning spectral properties of block circular symmetric covariance matrices and the estimation of eigenvalues. In Sect. 3, a new approach to finding identifiable parameters and obtaining explicit MLEs for (co)variance parameters by considering restricted models is presented. The model and the proposed approach are illustrated with examples in Sect. 4 and a real data application in Sect. 5.

2 Maximum likelihood estimation

The main goal of the article is to find maximum likelihood estimators for the unknown parameters μ and $\boldsymbol{\theta}$, $\boldsymbol{\theta} = (\sigma^2, \sigma_1, \sigma_2, \tau_1, \ldots, \tau_{2r})'$ in model (1). Let $\mathbf{y}_1, \ldots, \mathbf{y}_n$ be a random sample from $N_p(\mu \mathbf{1}_p, \boldsymbol{\Sigma})$, and define $\boldsymbol{Y} = (\mathbf{y}_1, \ldots, \mathbf{y}_n)$. Then, vec$\boldsymbol{Y} \sim N_{np}(\mu \mathbf{1}_{np}, \boldsymbol{I}_n \otimes \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma}$ is specified in (2), and vec$(\cdot)$ denotes the vectorization operation putting columns underneath starting from the first. Using the loglikelihood function

$$\ln L(\mu, \boldsymbol{\Sigma}) = c - \frac{1}{2}\,|\boldsymbol{I}_n \otimes \boldsymbol{\Sigma}| - \frac{1}{2}(\text{vec}\mathbf{Y} - \mu \mathbf{1}_{np})'(\boldsymbol{I}_n \otimes \boldsymbol{\Sigma})^{-1}(\text{vec}\mathbf{Y} - \mu \mathbf{1}_{np}),$$

where $c = -\frac{1}{2}np\ln(2\pi)$, the MLE for μ is easily derived, i.e.

$$\hat{\mu} = [\mathbf{1}'_{np}(\boldsymbol{I}_n \otimes \boldsymbol{\Sigma}^{-1})\mathbf{1}_{np}]^{-1}\mathbf{1}'_{np}(\boldsymbol{I}_n \otimes \boldsymbol{\Sigma}^{-1})\mathrm{vec}\mathbf{Y}.$$

Since $\mathbf{1}_p$ is an eigenvector of $\boldsymbol{\Sigma}$ and $\mathbf{1}_{np} = \mathbf{1}_n \otimes \mathbf{1}_p$, we get

$$\hat{\mu} = (\mathbf{1}'_{np}\mathbf{1}_{np})^{-1}\mathbf{1}'_{np}\mathrm{vec}\mathbf{Y} = \frac{1}{np}\mathbf{1}'_{np}\mathrm{vec}\mathbf{Y},$$

i.e. $\hat{\mu}$ is an OLS estimator.

The next step is to estimate $\boldsymbol{\theta}$. This will be achieved using the MLEs for the eigenvalues of $\boldsymbol{\Sigma}$ and the following theorem which concerns the linear representation of distinct eigenvalues of $\boldsymbol{\Sigma}$ in terms of (co)variance components $\boldsymbol{\theta}$ (see Liang et al. 2012).

Theorem 1 *Let $\boldsymbol{\eta}$ be a vector of the $2r$ distinct eigenvalues of $\boldsymbol{\Sigma}$ defined in (2). Then $\boldsymbol{\eta}$ can be expressed as:*

$$\boldsymbol{\eta} = \boldsymbol{L}\boldsymbol{\theta}, \tag{6}$$

where

$$\boldsymbol{L} = (\boldsymbol{B}_1 \vdots \boldsymbol{B}_2), \tag{7}$$

and

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

$\mathbf{0}_{r-1}$ is a column vector of size $r-1$ with all elements equal to zero, and $\boldsymbol{A} = (a_{ij})$ is a square matrix of size r with

$$a_{ij} = \begin{cases} 2^{I(1<j<r)}\cos(2\pi(i-1)(n_1-j+1)/n_1), & \text{if } n_1 \text{ is even}, \\ 2^{I(1<j\leq r)}\cos(2\pi(i-1)(n_1-j+1)/n_1), & \text{if } n_1 \text{ is odd} \end{cases} \tag{8}$$

where $I(\cdot)$ is the indicator function and $i, j = 1, \ldots, r$.

The maximum likelihood estimators of the eigenvalues in $\boldsymbol{\eta}$, given in (6), as well as their distributions will now be derived. Let $\boldsymbol{D}(\boldsymbol{\eta})$ be a diagonal matrix with the distinct eigenvalues $\eta_1, \ldots, \eta_{2r}$ of multiplicities $m_1, \ldots, m_{2r}$, $\sum_{i=1}^{2r} m_i = p$, on the main diagonal. Let $\boldsymbol{Q}$ be the orthogonal matrix where columns $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p$ are the known orthonormal eigenvectors generating the corresponding eigenspace. Thus, $\boldsymbol{\Sigma} = \boldsymbol{Q}\boldsymbol{D}(\boldsymbol{\eta})\boldsymbol{Q}'$. Due to the symmetry assumptions (compound symmetry and symmetric Toeplitz), the eigenvectors $\boldsymbol{v}_i$, $i = 1, \ldots, p$, are completely independent of the $\boldsymbol{\Sigma}$ elements, see Liang et al. (2012).

Replacing μ by its MLE, $\hat{\mu}$, we get for the likelihood function

$$
\begin{aligned}
L(\mu, \boldsymbol{\eta}) &\leq L(\hat{\mu}, \boldsymbol{\eta}) \\
&= (2\pi)^{-\frac{1}{2}pn} |\boldsymbol{D}(\boldsymbol{\eta})|^{-\frac{n}{2}} \mathrm{e}^{-\frac{1}{2}\mathrm{tr}\{[\boldsymbol{D}(\boldsymbol{\eta})]^{-1}\boldsymbol{Q}'(\boldsymbol{Y}-\hat{\mu}\mathbf{1}_p\mathbf{1}_n')(\boldsymbol{Y}-\hat{\mu}\mathbf{1}_p\mathbf{1}_n')'\boldsymbol{Q}\}},
\end{aligned}
$$

where $\mathrm{tr}(\cdot)$ denotes the trace of a matrix.

Let $\boldsymbol{H} = \boldsymbol{Q}'(\boldsymbol{Y}-\hat{\mu}\mathbf{1}_p\mathbf{1}_n')(\boldsymbol{Y}-\hat{\mu}\mathbf{1}_p\mathbf{1}_n')'\boldsymbol{Q}$ and $\boldsymbol{H}_d = \mathrm{diag}(\boldsymbol{H}) = (h_j)$. Then $L(\hat{\mu}, \boldsymbol{\eta})$ can be expressed in terms of η_i, $i = 1, \ldots, 2r$, i.e.,

$$
\begin{aligned}
L(\hat{\mu}, \boldsymbol{\eta}) &= (2\pi)^{-\frac{1}{2}pn} |\boldsymbol{D}(\boldsymbol{\eta})|^{-\frac{n}{2}} \mathrm{e}^{-\frac{1}{2}\mathrm{tr}\{[\boldsymbol{D}(\boldsymbol{\eta})]^{-1}\boldsymbol{H}_d\}} \\
&= (2\pi)^{-\frac{1}{2}pn} \prod_{i=1}^{2r} \eta_i^{-(nm_i/2)} \exp\left\{-\frac{1}{2}\sum_{i=1}^{2r}\eta_i^{-1}\sum_{j=k+1}^{k+m_i} h_j\right\},
\end{aligned}
\tag{9}
$$

where $k = \sum_{l=1}^{i} m_{l-1}$ and $m_0 = 0$. Taking the logarithm in (9) and differentiating with respect to η_i, we obtain the following normal equations

$$
\frac{\partial \ln L(\hat{\mu}, \boldsymbol{\eta})}{\partial \eta_i} = -\frac{nm_i}{2}\frac{1}{\eta_i} + \frac{1}{2\eta_i^2}\sum_{j=k+1}^{k+m_i} h_j = 0. \tag{10}
$$

Solving the equations in (10) yields the MLEs for η_i:

$$
\hat{\eta}_i = \frac{1}{nm_i}\sum_{j=k+1}^{k+m_i} h_j, \ i = 1, \ldots, 2r, \tag{11}
$$

since $L(\boldsymbol{\mu}, \boldsymbol{\eta}) \leq L(\hat{\boldsymbol{\mu}}, \boldsymbol{\eta}) \leq L(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\eta}})$.

In the next proposition, the distribution of MLE for η_i will be derived.

Proposition 1 *The MLEs of* η_i, $\hat{\eta}_i$, *are independent* χ^2 *distributed random variables,* $i = 1, \ldots, 2r$, *with* $\hat{\eta}_1 \sim \frac{\eta_1}{n}\chi^2_{(n-1)}$, *and* $\hat{\eta}_i \sim \frac{\eta_i}{nm_i}\chi^2_{(nm_i)}$, $i = 2, \ldots, 2r$.

Proof Let $\boldsymbol{Q}$ be as in the proof of Theorem 1. Since $\mathbf{1}_p$ is an eigenvector of $\boldsymbol{\Sigma}$, premultiplying $\mathrm{vec}\boldsymbol{Y}$ by $\boldsymbol{I}_n \otimes \boldsymbol{Q}'$ we obtain

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

where $\boldsymbol{D}(\boldsymbol{\eta})$ is a $p \times p$ diagonal matrix with the eigenvalues of $\boldsymbol{\Sigma}$, given in Theorem 1, on the main diagonal, i.e. the model can be split into $2r$ independent models. Define $\boldsymbol{w}_i = (\boldsymbol{I}_n \otimes \boldsymbol{v}_i')\mathrm{vec}\boldsymbol{Y}$, where $\boldsymbol{v}_i$ are the known eigenvectors of $\boldsymbol{\Sigma}$, and let

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

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

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

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 $\hat{\eta}_1 = \frac{1}{n}\tilde{\mathbf{y}}_1'(\boldsymbol{I}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n')\tilde{\mathbf{y}}_1$, and the MLE of η_i is $\hat{\eta}_i = \frac{1}{nm_i}\tilde{\mathbf{y}}_i'\tilde{\mathbf{y}}_i$, $i = 2, \ldots, 2r$. Hence, $\hat{\eta}_1 \sim \frac{\eta_1}{n}\chi^2_{(n-1)}$, and $\hat{\eta}_i \sim \frac{\eta_i}{nm_i}\chi^2_{(nm_i)}$, $i = 2, \ldots, 2r$, which completes the proof. □

An alternative method to estimate η_i, $i = 1, \ldots, 2r$, is to use restricted maximum likelihood estimation (REML), in which the basic idea is to eliminate the bias due to mean parameters when estimating variance components (see Patterson and Thompson 1971; Searle et al. 1992; LaMotte 2007). In model (1) with only one mean parameter μ, the REMLs for η_i's are the following

$$\hat{\eta}_{1,\mathrm{REML}} = \frac{1}{n-1}\tilde{\mathbf{y}}_1'\left(\boldsymbol{I}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n'\right)\tilde{\mathbf{y}}_1,$$

and, for $i = 2, \ldots, 2r$,

$$\hat{\eta}_{i,\mathrm{REML}} = \frac{1}{nm_i}\tilde{\mathbf{y}}_i'\tilde{\mathbf{y}}_i.$$

It is noteworthy that for $i = 2, \ldots, 2r$, $\hat{\eta}_{i,\mathrm{REML}} = \hat{\eta}_{i,\mathrm{ML}}$ and $\hat{\eta}_{1,\mathrm{REML}} = \frac{n}{n-1}\hat{\eta}_{1,\mathrm{ML}}$. Hence, the difference between MLEs and REMLs of variance components in model (1) is very small. Moreover, the problem of over parametrization always exists, no matter which method, MLEs or REMLs, is applied. Therefore, in the subsequent, only the MLEs of variance components will be considered.

Now using (11) we get the following estimator of $\boldsymbol{\Sigma}$:

$$\hat{\boldsymbol{\Sigma}} = \boldsymbol{Q}\boldsymbol{D}(\hat{\boldsymbol{\eta}})\boldsymbol{Q}'.$$

Since the covariance matrix $\boldsymbol{\Sigma}$ is a linear combination of the three covariance matrices $\boldsymbol{\Sigma}_1$, $\boldsymbol{\Sigma}_2$ and $\sigma^2\boldsymbol{I}$, the elements of $\boldsymbol{\Sigma}$ are functions of the unknown parameters in $\boldsymbol{\theta}$, i.e. $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\boldsymbol{\theta})$. If the number of unknown parameters in $\boldsymbol{\Sigma}$ (size of $\boldsymbol{\theta}$) equals the number distinct eigenvalues of $\boldsymbol{\Sigma}$ (size of $\boldsymbol{\eta}$), the MLE for $\boldsymbol{\theta}$ has an explicit expression, e.g. see Szatrowski (1980), which can be obtained by solving the system of linear equations (6) when $\boldsymbol{\eta}$ has been replaced by its MLE, and $\hat{\boldsymbol{\Sigma}}$ is a MLE.

The next proposition establishes a relationship between the number of elements in $\boldsymbol{\theta}$ and $\boldsymbol{\eta}$.

Proposition 2 *The difference between the number of unknown parameters in $\boldsymbol{\Sigma}$ and the number of distinct eigenvalues of the covariance matrix $\boldsymbol{\Sigma}$ defined in* (2) *equals 3.*

Proof Recall that $\boldsymbol{\Sigma}$ given in (2) is a linear combination of the three covariance matrices, i.e.

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

Thus there are $2r + 3$ unknown parameters in $\boldsymbol{\Sigma}$, whereas there are only $2r$ distinct eigenvalues of $\boldsymbol{\Sigma}$. □

According to Proposition 2 we have to put at least three restrictions on the parameter space 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 $\boldsymbol{\Sigma}_1$ and two constraints on $\boldsymbol{\Sigma}_2$ or alternatively one can put three constraints on $\boldsymbol{\Sigma}_2$.

Observe that $\boldsymbol{\eta}$ in (6) 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. $\boldsymbol{\eta} = \boldsymbol{\eta}(\boldsymbol{\lambda}^{\boldsymbol{\Sigma}_1}, \boldsymbol{\lambda}^{\boldsymbol{\Sigma}_2}, \lambda^{\boldsymbol{I}})$:

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

where $h = 1 + I(i \geq r+1)$, $j = i - r(h-1)$ and $i = 1, \ldots, 2r$.

It turns out that instead of putting constraints on $\boldsymbol{\theta}$, it is 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 factor, which usually are interpretable, and at the same time keep the original symmetry assumptions. However, in practise the constraints are part of the model assumptions which have to be verified against data. Below, we present model restrictions which are called Scenario 1 and Scenario 2. Observe, that the maximum of the likelihood will not be affected by the choice of the different scenarios. Therefore, the likelihood cannot guide us which model should be used, only the observed data.

Scenario 1 One constraint is imposed on the spectrum of $\boldsymbol{\Sigma}_1$ and two constraints on the spectrum of $\boldsymbol{\Sigma}_2$. Two possibilities for imposing constraints are given by

(i) $\lambda_g^{\boldsymbol{\Sigma}_1} = 0$, $\lambda_{g1}^{\boldsymbol{\Sigma}_2} = 0$ and $\lambda_{h1}^{\boldsymbol{\Sigma}_2} = 0$, $g, h \in \{1, 2\}$, $g \neq h$;
(ii) $\lambda_g^{\boldsymbol{\Sigma}_1} = 0$, $\lambda_{h1}^{\boldsymbol{\Sigma}_2} = 0$ and $\lambda_{ij}^{\boldsymbol{\Sigma}_2} = 0$, $g, h, i \in \{1, 2\}$, $g \neq h$, $j \in \{2, \ldots, r\}$.

In fact, the condition $\lambda_2^{\boldsymbol{\Sigma}_1} = 0$ in (i) is very restrictive, since in this case $\sigma_1 = \sigma_2 = \text{cov}(\gamma_{1k}, \gamma_{1l})$, $k, l = 1, \ldots, n_2$, i.e. the covariance matrix $\boldsymbol{\Sigma}_1$ in (3) 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 $\boldsymbol{K}_i\boldsymbol{\theta} = \boldsymbol{0}$, $i = 1, 2$, where

$$\boldsymbol{K}_1 = \begin{pmatrix} 0 & 1 & (n_2-1) & \mathbf{0}_r & \mathbf{0}_r \\ 0 & 0 & 0 & \boldsymbol{a}_1 & (n_2-1)\boldsymbol{a}_1 \\ 0 & 0 & 0 & \boldsymbol{a}_1 & -\boldsymbol{a}_1 \end{pmatrix}, \tag{13}$$

$$\boldsymbol{K}_2 = \begin{pmatrix} 0 & 1 & (n_2-1) & \mathbf{0}_r & \mathbf{0}_r \\ 0 & 0 & 0 & \boldsymbol{a}_1 & -\boldsymbol{a}_1 \\ 0 & 0 & 0 & \boldsymbol{a}_j & -(1-n_2)^{2-h}\boldsymbol{a}_j \end{pmatrix}, \tag{14}$$

and $\boldsymbol{a}_1 : r \times 1$ and $\boldsymbol{a}_j : r \times 1$ are the corresponding rows of the matrix $\boldsymbol{A}$ defined in (8), $h \in \{1, 2\}$ and $j \in \{2, \ldots, r\}$.

Scenario 2 Three constraints are imposed on the spectrum of $\boldsymbol{\Sigma}_2$:

(iii) $\lambda_{g1}^{\boldsymbol{\Sigma}_2} = 0$ and $\lambda_{hj}^{\boldsymbol{\Sigma}_2} = 0$, $g = 1, 2$, $h \in \{1, 2\}$, $j \in \{2, \ldots, r\}$.

Condition (iii) can also be expressed as $\boldsymbol{K}_3\boldsymbol{\theta} = \mathbf{0}$, where

$$\boldsymbol{K}_3 = \begin{pmatrix} 0 & 0 & 0 & \boldsymbol{a}_1 & (n_2-1)\boldsymbol{a}_1 \\ 0 & 0 & 0 & \boldsymbol{a}_1 & -\boldsymbol{a}_1 \\ 0 & 0 & 0 & \boldsymbol{a}_j & -(1-n_2)^{2-h}\boldsymbol{a}_j \end{pmatrix}, \tag{15}$$

and $\boldsymbol{a}_1 : r \times 1$ and $\boldsymbol{a}_j : r \times 1$ are the corresponding rows of matrix $\boldsymbol{A}$ defined in (8).

For a 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. 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 $\boldsymbol{v}_j \neq \mathbf{1}_{n_1}$. Then Scenario 1 and Scenario 2 can be formulated as follows:

Scenario 1: (alternative formulation)

(iv) $\mathbf{1}_{n_2}'\boldsymbol{\gamma}_1 = 0$, $\mathbf{1}_p'\boldsymbol{\gamma}_2 = 0$ and $(\boldsymbol{v} \otimes \mathbf{1}_{n_1})'\boldsymbol{\gamma}_2 = 0$;

(v) $\mathbf{1}_{n_2}'\boldsymbol{\gamma}_1 = 0$, $(\boldsymbol{v} \otimes \mathbf{1}_{n_1})'\boldsymbol{\gamma}_2 = 0$ and $(\boldsymbol{v}^{h-1} \otimes \boldsymbol{v}_j)'\boldsymbol{\gamma}_2 = 0$, $h \in \{1, 2\}$, $j \in \{2, \ldots, r\}$.

Scenario 2: (alternative formulation)

(vi) $(\boldsymbol{v}^{g-1} \otimes \mathbf{1}_{n_1})'\boldsymbol{\gamma}_2 = 0$ and $(\boldsymbol{v}^{h-1} \otimes \boldsymbol{v}_j)'\boldsymbol{\gamma}_2 = 0$, $g = 1, 2$, $h \in \{1, 2\}$, $j \in \{2, \ldots, r\}$.

It is noteworthy that the restrictions in Scenario 1 and Scenario 2 only yield different reparameterizations of $\boldsymbol{\theta}$ due to $\boldsymbol{\theta} = (\boldsymbol{K}_i')^o\boldsymbol{\theta}_i^*$, but they result in the same estimate of $\boldsymbol{\eta}$ (invariance to reparameterization of $\boldsymbol{\theta}$), since $\boldsymbol{\eta} = \boldsymbol{L}(\boldsymbol{K}_i')^o\boldsymbol{\theta}_i^*$ and $\boldsymbol{L}(\boldsymbol{K}_i')^o$ is of full rank, $i = 1, 2, 3$.

3 Explicit MLEs of variance parameters

In the previous section, different types of restrictions were described to derive estimable (co)variance components. These restrictions will be shown to yield explicit MLEs of the (co)variance parameters. Let $\boldsymbol{\theta}^*$ be the vector of the unknown parameters in model (1) under any restriction given by Scenario 1 or Scenario 2, i.e., (i)–(iii).

Theorem 2 *Model* (1) *has explicit MLEs for* $\boldsymbol{\theta}$ *if one of the conditions (i)–(iii) holds.*

Proof The restriction $\boldsymbol{K}_i\boldsymbol{\theta} = \mathbf{0}$, $i = 1, 2, 3$, on $\boldsymbol{\theta}$ in $\boldsymbol{\eta} = \boldsymbol{L}\boldsymbol{\theta}$ is equivalent to $\boldsymbol{\theta} = (\boldsymbol{K}_i')^o\boldsymbol{\theta}_i^*$, where $(\boldsymbol{K}_i')^o$: $(2r + 3) \times 2r$ is a matrix in which columns generate the orthogonal complement to the column vector space of $\boldsymbol{K}_i'$ and $\boldsymbol{\theta}_i^*$: $2r \times 1$ is the vector of unknown parameters in model (1) which is determined by restrictions (i)–(iii) given in Scenario 1 and Scenario 2. If $\boldsymbol{L}(\boldsymbol{K}_i')^o$ is invertible then $\boldsymbol{\theta}_i^*$ can be estimated. Now observe that $r(\boldsymbol{L}(\boldsymbol{K}_i')^o) = r(\boldsymbol{L}' : \boldsymbol{K}_i') - r(\boldsymbol{K}_i')$, and $r(\boldsymbol{K}_i') = 3$. Due to the structure of $\boldsymbol{L}$ and $\boldsymbol{K}_i$, given in (7) and (13)–(15), respectively, e.g.

$$(\boldsymbol{L}' \vdots \boldsymbol{K}_1') = \left(\begin{array}{cccc:ccc} 1 & \mathbf{1}_{r-1}' & 1 & \mathbf{1}_{r-1}' & 0 & 0 & 0 \\ n_1 & \mathbf{0}_{r-1}' & n_1 & \mathbf{0}_{r-1}' & 1 & 0 & 0 \\ n_1(n_2-1) & \mathbf{0}_{r-1}' & -n_1 & \mathbf{0}_{r-1}' & (n_2-1) & 0 & 0 \\ \hdashline \multicolumn{2}{c}{\boldsymbol{A}'} & \multicolumn{2}{c:}{\boldsymbol{A}'} & \mathbf{0}_r' & \boldsymbol{a}_1' & \boldsymbol{a}_1' \\ \multicolumn{2}{c}{(n_2-1)\boldsymbol{A}'} & \multicolumn{2}{c:}{-\boldsymbol{A}'} & \mathbf{0}_r' & (n_2-1)\boldsymbol{a}_1' & -\boldsymbol{a}_1' \end{array} \right)$$

we find that $r(\boldsymbol{L}' \vdots \boldsymbol{K}_i') = 2r + 3$. Thus, $r(\boldsymbol{L}(\boldsymbol{K}_i')^o) = (2r+3) - 3 = 2r$, i.e. the matrix $\boldsymbol{L}(\boldsymbol{K}_i')^o$ is of full rank, and therefore is invertible. Hence, $\boldsymbol{\eta} = \boldsymbol{L}\boldsymbol{\theta} = \boldsymbol{L}(\boldsymbol{K}_i')^o \boldsymbol{\theta}_i^*$ and $\boldsymbol{\theta}_i^* = (\boldsymbol{L}(\boldsymbol{K}_i')^o)^{-1}\boldsymbol{\eta}$. □

Corollary 1 *The MLEs for the vector of the unknown parameters $\boldsymbol{\theta}_i^*$, $i = 1, 2, 3$, in model (1) under any restriction given by Scenario 1 or Scenario 2 are the following*

$$\hat{\boldsymbol{\theta}}_i^* = (\boldsymbol{L}(\boldsymbol{K}_i')^o)^{-1}\hat{\boldsymbol{\eta}}, \tag{16}$$

where $\boldsymbol{K}_i$, $i \in \{1, 2, 3\}$ is given in (13)–(15).

Corollary 2 *The estimator $\hat{\boldsymbol{\theta}}_i^*$, $i = 1, 2, 3$, is a linear combination of independent χ^2-distributed random variables.*

Proof Follows from (16) and Proposition 1. □

Theorem 3 *The MLEs obtained for $\boldsymbol{\theta}$ under conditions* (i)–(iii) *given in Scenario* 1 *and Scenario* 2 *are unique.*

Proof In $\boldsymbol{\eta} = \boldsymbol{L}(\boldsymbol{K}_i')^o \boldsymbol{\theta}_i^*$ the vector of unknown parameters $\boldsymbol{\theta}_i^*$ is determined by restrictions (i)–(iii) given in Scenario 1 and Scenario 2, and $\boldsymbol{\eta}$ and $\boldsymbol{L}$ are fixed by the structure of the covariance matrix $\boldsymbol{\Sigma}$ specified in (2). Uniqueness of $(\boldsymbol{K}_i')^o$ yields the uniqueness of MLEs for $\boldsymbol{\theta}$. Assume, that there exist two matrices $(\boldsymbol{K}_i')_1^o$ and $(\boldsymbol{K}_i')_2^o$ such that $(\boldsymbol{K}_i')_j^o \boldsymbol{\theta}_i^* = \boldsymbol{\theta}$, $j = 1, 2$. Then, $\boldsymbol{L}[(\boldsymbol{K}_i')_1^o - (\boldsymbol{K}_i')_2^o]\boldsymbol{\theta}_i^* = \mathbf{0}$, and since $\boldsymbol{L} : (2r+3) \times 2r$ with $r(\boldsymbol{L}) = 2r$, $[(\boldsymbol{K}_i')_1^o - (\boldsymbol{K}_i')_2^o]\boldsymbol{\theta}_i^* = \mathbf{0}$ which means that $(\boldsymbol{K}_i')_1^o = (\boldsymbol{K}_i')_2^o$. Therefore, the matrix $(\boldsymbol{K}_i')^o$ is unique, and the theorem is established. □

Up to now it has been shown that there exist one unique $(\boldsymbol{K}_i')^o$, $i = 1, 2, 3$, but no explicit expression for it has been given.

Proposition 3 *The structure of the matrix $(\boldsymbol{K}_1')^o$ is uniquely determined by the choice of restrictions given in (i) in Scenario 1 and equals*

$$(\boldsymbol{K}_1')^o = \begin{pmatrix} \boxed{1} & & & \mathbf{0} \\ & \boxed{\begin{matrix} -(n_2-1) \\ 1 \end{matrix}} & & \\ & & \boxed{\begin{matrix} -2 & -2 & \dots & -1 \\ & \boldsymbol{I}_{r-1} & & \end{matrix}} & \\ \mathbf{0} & & & \boxed{\begin{matrix} -2 & -2 & \dots & -1 \\ & \boldsymbol{I}_{r-1} & & \end{matrix}} \end{pmatrix}. \tag{17}$$

Proof Due to the relationship $\boldsymbol{\lambda}^{\boldsymbol{\Sigma}_2} = \boldsymbol{B}_2\boldsymbol{\tau}$ (Liang et al. 2012), we have

$$\lambda_{ij}^{\boldsymbol{\Sigma}_2} = \boldsymbol{a}_j(\boldsymbol{\tau}_1 - (1-n_2)^{2-i}\boldsymbol{\tau}_2),$$

where $\boldsymbol{a}_j$ is the jth row of the matrix $\boldsymbol{A}$ defined in (8), $i = 1, 2$, $j = 1, \dots, 2r$, and $\boldsymbol{\tau} = (\boldsymbol{\tau}_1', \boldsymbol{\tau}_2')'$ with $\boldsymbol{\tau}_1 = (\tau_1, \dots, \tau_r)'$, $\boldsymbol{\tau}_2 = (\tau_{r+1}, \dots, \tau_{2r})'$. According to (i) in Scenario 1, $\lambda_{11}^{\boldsymbol{\Sigma}_2} = 0$ and $\lambda_{21}^{\boldsymbol{\Sigma}_2} = 0$, which yields the following system of linear equations

$$\begin{aligned} \lambda_{11}^{\boldsymbol{\Sigma}_2} &= \boldsymbol{a}_1(\boldsymbol{\tau}_1 - (1-n_2)\boldsymbol{\tau}_2) = 0, \\ \lambda_{21}^{\boldsymbol{\Sigma}_2} &= \boldsymbol{a}_1(\boldsymbol{\tau}_1 - \boldsymbol{\tau}_2) = 0, \end{aligned} \tag{18}$$

where $\boldsymbol{a}_1 = (1, 2, \dots, 2, 1)'$.

Solving the system of equations (18) yields

$$\boldsymbol{a}_1\boldsymbol{\tau}_1 = 0, \tag{19}$$

$$\boldsymbol{a}_1\boldsymbol{\tau}_2 = 0, \tag{20}$$

which is the equivalent representation of restrictions $\lambda_{11}^{\boldsymbol{\Sigma}_2} = 0$ and $\lambda_{21}^{\boldsymbol{\Sigma}_2} = 0$ yet in terms of $\boldsymbol{\tau} = (\boldsymbol{\tau}_1, \boldsymbol{\tau}_2)'$.

The restrictions on the elements of $\boldsymbol{\Sigma}_2$ in (19)–(20) can specifically be expressed as

$$\tau_1 = -2\sum_{i=2}^{r-1}\tau_i - \tau_r, \tag{21}$$

$$\tau_{r+1} = -2\sum_{i=2}^{r-1}\tau_{r+i} - \tau_{2r}, \tag{22}$$

which together with the constraint $\sigma_1 = -(n_2-1)\sigma_2$, as a consequence of $\lambda_1^{\boldsymbol{\Sigma}_1} = 0$, yields a restricted vector of parameters

$$\boldsymbol{\theta}_1^* = (\sigma^2, \sigma_2, \tau_2, \dots, \tau_r, \tau_{r+2}, \dots, \tau_{2r})'$$

of length $2r$.

The matrix $(\boldsymbol{K}_1')^o$ satisfying $\boldsymbol{\theta} = (\boldsymbol{K}_1')^o\boldsymbol{\theta}_1^*$ is then constructed as a block-diagonal matrix comprising four blocks which correspond to the unknown parameters σ^2, (σ_1, σ_2), $\boldsymbol{\tau}_1$ and $\boldsymbol{\tau}_2$ in $\boldsymbol{\theta}$. The structure of these blocks is determined by the restrictions in (i), i.e. $\sigma_1 = -(n_2 - 1)\sigma_2$ and those obtained in (21)–(22).

It is easy to check that in model (1) the vector of unknown (co)variance parameters

$$\boldsymbol{\theta} = (\sigma^2, \sigma_1, \sigma_2, \tau_1, \ldots, \tau_r, \tau_{r+1}, \ldots, \tau_{2r})',$$

under restrictions (i) given in Scenario 1, i.e. $\boldsymbol{K}_1\boldsymbol{\theta} = \mathbf{0}$, where

$$\boldsymbol{K}_1 = \left(\begin{array}{c|cc|cc} 0 & 1 & (n_2-1) & \mathbf{0}_r & \mathbf{0}_r \\ 0 & 0 & 0 & \boldsymbol{a}_1 & (n_2-1)\boldsymbol{a}_1 \\ 0 & 0 & 0 & \boldsymbol{a}_1 & -\boldsymbol{a}_1 \end{array}\right)$$
$$= \left(\begin{array}{c|cc|ccccccccc} 0 & 1 & (n_2-1) & 0 & 0 \ldots 0 & 0 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 \ldots 2 & 1 & (n_2-1) & 2(n_2-1) & \ldots & 2(n_2-1) & (n_2-1) \\ 0 & 0 & 0 & 1 & 2 \ldots 2 & 1 & -1 & -2 & \ldots & -2 & -1 \end{array}\right),$$

becomes

$$\boldsymbol{\theta}_1^* = (\sigma^2, \sigma_2, \tau_2, \ldots, \tau_r, \tau_{r+2}, \ldots, \tau_{2r})',$$

and $(\boldsymbol{K}_1')^o\boldsymbol{\theta}_1^* = \boldsymbol{\theta}$, where the matrix $(\boldsymbol{K}_1')^o$ is defined in (17). □

As an illustration consider

$$(\boldsymbol{K}_1')^o\boldsymbol{\theta}_1^* = \left(\begin{array}{cc|cc|cc} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -(n_2-1) & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & -2 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & -2 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right)\begin{pmatrix} \sigma^2 \\ \sigma_2 \\ \tau_2 \\ \tau_3 \\ \tau_5 \\ \tau_6 \end{pmatrix} = \begin{pmatrix} \sigma^2 \\ -(n_2-1)\sigma_2 \\ \sigma_2 \\ -\tau_2 - \tau_3 \\ \tau_2 \\ \tau_3 \\ -\tau_5 - \tau_6 \\ \tau_5 \\ \tau_6 \end{pmatrix}.$$

Let us now consider (ii) in Scenario 1, i.e. $\lambda_{21}^{\boldsymbol{\Sigma}_2} = 0$ and $\lambda_{kj}^{\boldsymbol{\Sigma}_2} = 0$, $k \in \{1, 2\}$, $j \in \{2, \ldots, r\}$ (when $\lambda_1^{\boldsymbol{\Sigma}_1} = 0$) which due to $\boldsymbol{\lambda}^{\boldsymbol{\Sigma}_2} = \boldsymbol{B}_2\boldsymbol{\tau}$ equal

$$\begin{aligned} \lambda_{21}^{\boldsymbol{\Sigma}_2} &= \boldsymbol{a}_1(\boldsymbol{\tau}_1 - \boldsymbol{\tau}_2) = 0, \\ \lambda_{kj}^{\boldsymbol{\Sigma}_2} &= \boldsymbol{a}_j(\boldsymbol{\tau}_1 - (1-n_2)^{2-k}\boldsymbol{\tau}_j) = 0. \end{aligned} \tag{23}$$

From (23) the following two restrictions on the elements of $\boldsymbol{\Sigma}_2$ can be obtained

$$\tau_{r+1} = \frac{2-\tilde{n}a_{j2}}{\tilde{n}(2-a_{j2})}\tau_1 + 2\sum_{i=3}^{r-1}\frac{(a_{ji}-\tilde{n}a_{j2})}{\tilde{n}(2-a_{j2})}\tau_i + \frac{(2a_{jr}-\tilde{n}a_{j2})}{\tilde{n}(2-a_{j2})}\tau_r$$
$$-2\sum_{i=3}^{r-1}\left(\frac{a_{ji}-a_{j2}}{2-a_{j2}}\right)\tau_{r+i} - \left(\frac{2a_{jr}-a_{j2}}{2-a_{j2}}\right)\tau_{2r}, \tag{24}$$
$$\tau_{r+2} = \frac{\tilde{n}-1}{\tilde{n}(2-a_{j2})}\tau_1 + \sum_{i=2}^{r-1}\frac{2\tilde{n}-a_{ji}}{\tilde{n}(2-a_{j2})}\tau_i + \frac{\tilde{n}-a_{jr}}{\tilde{n}(2-a_{j2})}\tau_r$$
$$-\sum_{i=3}^{r-1}\frac{2-a_{ji}}{2-a_{j2}}\tau_{r+i} - \frac{1-a_{jr}}{2-a_{j2}}\tau_{2r}, \tag{25}$$

where $\tilde{n} = (1-n_2)^{2-k}$, $k \in \{1, 2\}$, and a_{ji} is defined in (8), $i, j = 1, \dots, r$.

The specified restrictions in (24)–(25) together with the constraint $\sigma_1 = -(n_2 - 1)\sigma_2$, as a consequence of $\lambda_1^{\boldsymbol{\Sigma}_1} = 0$, yield a restricted vector of parameters

$$\boldsymbol{\theta}_2^* = (\sigma^2, \sigma_2, \tau_1, \dots, \tau_r, \tau_{r+3}, \dots, \tau_{2r})'$$

of length $2r$.

The matrix $(\boldsymbol{K}_2')^o$ is then constructed as a block-matrix comprising four blocks which correspond to unknown parameters σ^2, (σ_1, σ_2), $\boldsymbol{\tau}_1$ and $\boldsymbol{\tau}_2$. The structure of these blocks is determined by the restrictions in (ii), i.e. $\sigma_1 = -(n_2 - 1)\sigma_2$ and those relations presented in (24)–(25). The matrix $(\boldsymbol{K}_2')^o$ is presented in the next proposition.

Proposition 4 *The structure of the matrix $(\boldsymbol{K}_2')^o$ is uniquely determined by the choice of restrictions given in* (ii) *in Scenario* 1 *and equals*

$$(\boldsymbol{K}_2')^o = \begin{pmatrix} \boxed{1} & & & \\ & \begin{matrix}-(n_2-1)\\ 1\end{matrix} & & \mathbf{0} \\ & & \boldsymbol{I}_r & \\ & & \begin{matrix} b_1 & b_2 & \dots & b_{r-1} & b_r \\ c_1 & c_2 & \dots & c_{r-1} & c_r \end{matrix} & \begin{matrix} b_{r+3} & \dots & b_{2r-1} & b_{2r} \\ c_{r+3} & \dots & c_{2r-1} & c_{2r} \end{matrix} \\ \mathbf{0} & & & \boldsymbol{I}_{r-2} \end{pmatrix},$$

where b_i and c_i are the coefficients defined in (24)–(25) *for τ_{r+1} and τ_{r+2}, respectively.*

Finally let us consider (iii) in Scenario 2, i.e. $\lambda_{11}^{\boldsymbol{\Sigma}_2} = 0$, $\lambda_{21}^{\boldsymbol{\Sigma}_2} = 0$, and $\lambda_{kj}^{\boldsymbol{\Sigma}_2} = 0$, $k \in \{1, 2\}$, $j \in \{2, \dots, r\}$ which due to $\boldsymbol{\lambda}^{\boldsymbol{\Sigma}_2} = \boldsymbol{B}_2\boldsymbol{\tau}$ can be expressed as follows

$$\begin{aligned} \lambda_{11}^{\boldsymbol{\Sigma}_2} &= \boldsymbol{a}_1(\boldsymbol{\tau}_1 + (n_2-1)\boldsymbol{\tau}_2) = 0, \\ \lambda_{21}^{\boldsymbol{\Sigma}_2} &= \boldsymbol{a}_1(\boldsymbol{\tau}_1 - \boldsymbol{\tau}_2) = 0, \\ \lambda_{kj}^{\boldsymbol{\Sigma}_2} &= \boldsymbol{a}_j(\boldsymbol{\tau}_1 - (1-n_2)^{2-k}\boldsymbol{\tau}_j) = 0. \end{aligned} \tag{26}$$

The restrictions on the eigenvalues in (26) can be expressed in terms of the elements of $\boldsymbol{\Sigma}_2$ in the following way

$$\tau_1 = -2\sum_{i=2}^{r-1}\tau_i - \tau_r, \tag{27}$$

$$\begin{aligned}\tau_{r+1} = &-\frac{2}{\tilde{n}}\tau_2 - \frac{2}{\tilde{n}}\sum_{i=3}^{r-1}\frac{a_{ji}-2}{(a_{j2}-2)}\tau_i - \frac{2(a_{jr}-1)}{\tilde{n}(a_{j2}-2)}\tau_r \\ &+2\sum_{i=3}^{r-1}\frac{a_{ji}-a_{j2}}{a_{j2}-2}\tau_{r+i} + \frac{2a_{jr}-a_{j2}}{a_{j2}-2}\tau_{2r},\end{aligned} \tag{28}$$

$$\tau_{r+2} = \frac{1}{\tilde{n}}\tau_2 + \frac{1}{\tilde{n}}\sum_{i=3}^{r-1}\frac{a_{ji}-2}{(a_{j2}-2)}\tau_i + \frac{a_{jr}-1}{\tilde{n}(a_{j2}-2)}\tau_r \tag{29}$$

$$-\sum_{i=3}^{r-1}\frac{a_{ji}-2}{a_{j2}-2}\tau_{r+i} - \frac{a_{jr}-1}{a_{j2}-2}\tau_{2r}, \tag{30}$$

where $\tilde{n} = (1-n_2)^{2-k}$, $k \in \{1, 2\}$, and a_{ji} is defined in (8), $i, j = 1, \ldots, r$.

Specified restrictions in (27)–(30) yield a restricted vector of parameters

$$\boldsymbol{\theta}_3^* = (\sigma^2, \sigma_1, \sigma_2, \tau_2, \ldots, \tau_r, \tau_{r+3}, \ldots, \tau_{2r})'$$

of length $2r$.

The matrix $(\boldsymbol{K}_3')^o$ is then constructed as a block-matrix comprising four blocks which correspond to the unknown parameters σ^2, (σ_1, σ_2), $\boldsymbol{\tau}_1$ and $\boldsymbol{\tau}_2$. The structure of these blocks is determined by the restrictions in (iii), i.e. those presented in (27)–(30). Hence, the following proposition can be stated.

Proposition 5 *The structure of the matrix $(\boldsymbol{K}_3')^o$ is uniquely determined by the choice of restrictions given in* (iii) *in Scenario* 2 *and equals*

$$(\boldsymbol{K}_3')^o = \begin{pmatrix} \boxed{1} & & & \\ & \boxed{\boldsymbol{I}_2} & & \mathbf{0} \\ & & \boldsymbol{I}_r & \\ & & b_1\ b_2\ \ldots\ b_{r-1}\ b_r & b_{r+3}\ \ldots\ b_{2r-1}\ b_{2r} \\ & & c_1\ c_2\ \ldots\ c_{r-1}\ c_r & c_{r+3}\ \ldots\ c_{2r-1}\ c_{2r} \\ & & d_1\ d_2\ \ldots\ d_{r-1}\ d_r & d_{r+3}\ \ldots\ d_{2r-1}\ d_{2r} \\ \mathbf{0} & & & \boldsymbol{I}_{r-3} \end{pmatrix},$$

where b_i, c_i and d_i are the coefficients defined in (27)–(30) *for τ_1, τ_{r+1} and τ_{r+2}, respectively.*

4 Example

In the next example, Propositions 3, 4 and 5 are illustrated, i.e., based on the MLEs for the eigenvalues $\boldsymbol{\eta}$, via $\boldsymbol{L}$ and $(\boldsymbol{K}_i')^o$, $i = 1, 2, 3$, the explicit estimators of the parameters in model (1) are obtained. When $n_1 = 4$, model (1) becomes

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

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

$$\boldsymbol{\Sigma}^{(1)} = \begin{pmatrix} \tau_1 & \tau_2 & \tau_3 & \tau_2 \\ \tau_2 & \tau_1 & \tau_2 & \tau_3 \\ \tau_3 & \tau_2 & \tau_1 & \tau_2 \\ \tau_2 & \tau_3 & \tau_2 & \tau_1 \end{pmatrix}, \quad \boldsymbol{\Sigma}^{(2)} = \begin{pmatrix} \tau_4 & \tau_5 & \tau_6 & \tau_5 \\ \tau_5 & \tau_4 & \tau_5 & \tau_6 \\ \tau_6 & \tau_5 & \tau_4 & \tau_5 \\ \tau_5 & \tau_6 & \tau_5 & \tau_4 \end{pmatrix}.$$

Using Theorem 1, the distinct eigenvalues of $\boldsymbol{\Sigma}$ 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

$$\boldsymbol{L} = \begin{pmatrix} 1 & 4 & 4\,(n_2-1) & 1 & 2 & 1 & n_2-1 & 2\,(n_2-1) & n_2-1 \\ 1 & 0 & 0 & 1 & 0 & -1 & n_2-1 & 0 & -\,(n_2-1) \\ 1 & 0 & 0 & 1 & -2 & 1 & n_2-1 & -2\,(n_2-1) & n_2-1 \\ 1 & 4 & -4 & 1 & 2 & 1 & -1 & -2 & -1 \\ 1 & 0 & 0 & 1 & 0 & -1 & -1 & 0 & 1 \\ 1 & 0 & 0 & 1 & -2 & 1 & -1 & 2 & -1 \end{pmatrix}.$$

According to Proposition 2, the number of distinct eigenvalues of $\boldsymbol{\Sigma}$ equals 6, and the number of unknown parameters in $\boldsymbol{\Sigma}$ is 9. Thus, as noted previously, the model is overparametrized. Let now impose the restriction (i) in 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 $\tau_4 = -2\tau_5 - \tau_6$. The condition $\lambda_1^{\boldsymbol{\Sigma}_1} = 0$ implies that we have the "smallest possible" covariance between the elements in $\boldsymbol{\Sigma}$. Moreover, the eigenvalue restriction on $\boldsymbol{\Sigma}_2$ implies that both $\boldsymbol{\Sigma}^{(1)}$ and $\boldsymbol{\Sigma}^{(2)}$ are weakly diagonally dominant. Diagonal dominance is connected to stability of a system and has many applications. In this restricted model, let $\boldsymbol{\theta}_1^* = (\sigma^2, \sigma_2, \tau_2, \tau_3, \tau_5, \tau_6)'$. The distinct eigenvalues of $\boldsymbol{\Sigma}$, $\boldsymbol{\eta}$, can be written in the form $\boldsymbol{\eta} = \boldsymbol{L}(\boldsymbol{K}_1')^o \boldsymbol{\theta}_1^*$, where, using the expression in (17),

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

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 $\boldsymbol{\theta}_1^*$ equals

$$\hat{\boldsymbol{\theta}}_1^* = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0\\ \frac{1}{4n_2} & 0 & 0 & -\frac{1}{4n_2} & 0 & 0\\ \frac{1}{4} & 0 & -\frac{1}{4n_2} & 0 & 0 & \frac{1}{4}\left(\frac{1}{n_2}-1\right)\\ \frac{1}{4} & -\frac{1}{2n_2} & \frac{1}{4n_2} & 0 & \frac{1}{2}\left(\frac{1}{n_2}-1\right) & \frac{1}{4}(1-\frac{1}{n_2})\\ 0 & 0 & -\frac{1}{4n_2} & 0 & 0 & \frac{1}{4n_2}\\ 0 & -\frac{1}{2n_2} & \frac{1}{4n_2} & 0 & \frac{1}{2n_2} & -\frac{1}{4n_2} \end{pmatrix} \hat{\boldsymbol{\eta}}, \tag{31}$$

and $\hat{\boldsymbol{\eta}}$ was presented in (11).

Let us now impose 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 $\tau_5 = \tau_2 + \tau_3 - \tau_6$. Then, $\tau_6 - \tau_3 = \tau_4 - \tau_1 = \tau_2 - \tau_5$. Thus, the absolute difference between the elements in $\boldsymbol{\Sigma}^{(1)}$ and $\boldsymbol{\Sigma}^{(2)}$ is the same, meaning that random mechanisms generating $\boldsymbol{\Sigma}^{(1)}$ and $\boldsymbol{\Sigma}^{(2)}$ are similar and proportional to each other. In this restricted model, let $\boldsymbol{\theta}_2^* = (\sigma^2, \sigma_2, \tau_1, \tau_2, \tau_3, \tau_6)'$. The distinct eigenvalues of $\boldsymbol{\Sigma}$, i.e. $\boldsymbol{\eta}$, can be written with the help of

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

which is followed by $\boldsymbol{\eta} = \boldsymbol{L}(\boldsymbol{K}_2')^o\boldsymbol{\theta}_2^*$, or

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

The explicit MLE of $\boldsymbol{\theta}_2^*$ is

$$\hat{\boldsymbol{\theta}}_2^* = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -\frac{1}{4n_2} & \frac{1}{4n_2} & 0 \\ \frac{1}{4n_2} & \frac{1}{2n_2} & \frac{1}{4n_2} & 0 & -\frac{3+n_2}{4n_2} & \frac{1}{4}\left(1-\frac{1}{n_2}\right) \\ \frac{1}{4n_2} & 0 & -\frac{1}{4n_2} & 0 & \frac{1}{4}\left(1-\frac{1}{n_2}\right) & \frac{1}{4}\left(\frac{1}{n_2}-1\right) \\ \frac{1}{4n_2} & -\frac{1}{2n_2} & \frac{1}{4n_2} & 0 & \frac{1}{4}\left(\frac{1}{n_2}-1\right) & \frac{1}{4}\left(1-\frac{1}{n_2}\right) \\ \frac{1}{4n_2} & -\frac{1}{2n_2} & \frac{1}{4n_2} & 0 & \frac{1}{4n_2} & -\frac{1}{4n_2} \end{pmatrix} \hat{\boldsymbol{\eta}}, \tag{32}$$

where $\hat{\boldsymbol{\eta}}$ is given in (11).

Suppose condition (iii) in Scenario 2 holds, for example, $\lambda_{11}^{\boldsymbol{\Sigma}_2} = 0$, $\lambda_{12}^{\boldsymbol{\Sigma}_2} = 0$ and $\lambda_{21}^{\boldsymbol{\Sigma}_2} = 0$. Hence, $\tau_1 = -2\tau_2 - \tau_3$, $\tau_4 = -2\tau_5 - \tau_6$, $\tau_4 = -\frac{1}{n_2-1}\tau_1 + \frac{1}{n_2-1}\tau_3 + \tau_6$ and $\tau_5 = -\frac{1}{n_2-1}\tau_2 - \frac{1}{n_2-1}\tau_3 - \tau_6$. The two last conditions are equivalent to $\tau_4 + \frac{1}{n_2-1}\tau_1 = \frac{1}{n_2-1}\tau_3 + \tau_6$ and $\tau_5 + \frac{1}{n_2-1}\tau_2 = -(\frac{1}{n_2-1}\tau_3 + \tau_6)$. Thus, $\boldsymbol{\Sigma}^{(1)}$ and $\boldsymbol{\Sigma}^{(2)}$ are weakly diagonal dominant, and the absolute value of the difference between the covariances in $\boldsymbol{\Sigma}^{(1)}$ and $\boldsymbol{\Sigma}^{(2)}$ is the same. For this model, let $\boldsymbol{\theta}_3^* = (\sigma^2, \sigma_1, \sigma_2, \tau_2, \tau_3, \tau_6)'$. The distinct eigenvalues of $\boldsymbol{\Sigma}$, $\boldsymbol{\eta}$, can be written in the form of $\boldsymbol{\eta} = \boldsymbol{L}(\boldsymbol{K}_3')^o\boldsymbol{\theta}_3^*$, where

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

and we have

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

The explicit MLE of $\boldsymbol{\theta}_3^*$ is

$$\hat{\boldsymbol{\theta}}_3^* = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ \frac{1}{4n_2} & -\frac{1}{4} & 0 & \frac{1}{4}\left(1-\frac{1}{n_2}\right) & 0 & 0 \\ \frac{1}{4n_2} & 0 & 0 & -\frac{1}{4n_2} & 0 & 0 \\ 0 & \frac{1}{4} & -\frac{1}{4n_2} & 0 & 0 & \frac{1}{4}\left(\frac{1}{n_2}-1\right) \\ 0 & \frac{1}{4}\left(1-\frac{2}{n_2}\right) & \frac{1}{4n_2} & 0 & \frac{1}{2}\left(\frac{1}{n_2}-1\right) & \frac{1}{4}\left(1-\frac{1}{n_2}\right) \\ 0 & -\frac{1}{2n_2} & \frac{1}{4n_2} & 0 & \frac{1}{2n_2} & -\frac{1}{4n_2} \end{pmatrix} \hat{\boldsymbol{\eta}}, \tag{33}$$

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

This section shows that analysis of hierarchical models is a complicated issue which has to be further exploited. Usually model parameters are the objects which we would like to receive knowledge about. Moreover, it has been shown that we need to put restrictions on the parameter space. However, if we do not have a specific data set with specific hypotheses it is not clear what restrictions should be imposed. In this article, a few different restrictions have been exploited. It appeared that there exist non-trivial differences among them. For example, consider the estimate of $\hat{\sigma}^2$ in (31), (32) and (33). In (31) $\hat{\sigma}^2 = \hat{\eta}_1$, in (32) $\hat{\sigma}^2 = \hat{\eta}_5$ and in (33) $\hat{\sigma}^2 = \hat{\eta}_2$. Thus, depending on the model assumption we obtain different estimators, which is in accordance with the different interpretations of the parameters of the models.

5 A real data example

To illustrate the results presented in this article, data were analyzed which consist of petal length measurements of a specific Kalanchoe plant. In Table 1 the petal length measurements made on 11 plants from the same greenhouse are presented. From each plant, there have been randomly chosen three flowers where each flower has four petals. Table 1 displays the measurements.

As the arrangement of petals is circular within each Kalanchoe flower, it is reasonable to assume that the correlation between the observations on any two petals within a single flower is a function of the number of petals between them. Therefore, a circular covariance structure is applied to describe the intra-flower correlation. The compound-symmetric covariance structure is assumed to describe the inter-flower correlation. It is supposed that all Kalanchoe plants, since they come from the same greenhouse, have the same mean. Hence, the following mixed linear model is fit to data:

$$\mathbf{y}_i = \mu \mathbf{1}_{12} + \boldsymbol{Z}_1 \boldsymbol{\gamma}_1 + \boldsymbol{Z}_2 \boldsymbol{\gamma}_2 + \boldsymbol{\epsilon}_i, \tag{34}$$

Table 1 Petal length measurements (mm) made on flowers from a specific Kalanchoe plant are presented

Plants/petals F	1a	1b	1c	1d	2a	2b	2c	2d	3a	3b	3c	3d	4a	4b	4c	4d	5a	5b	5c	5d	6a	6b	6c	6d
I	8.6	8.6	7.8	8.0	7.8	8.0	8.0	7.1	6.7	7.7	8.1	7.3	7.6	7.0	7.9	6.7	5.8	6.9	6.7	6.4	6.5	7.2	7.0	6.5
II	7.1	6.7	7.5	8.5	6.4	7.8	6.6	6.8	7.7	7.5	7.8	6.4	7.5	7.6	6.9	8.0	6.8	8.0	7.2	6.6	7.9	7.5	6.9	7.6
III	7.2	6.5	6.8	7.3	7.7	7.3	6.5	6.6	7.5	7.9	7.1	7.2	6.2	6.5	6.9	5.9	9.4	8.5	8.1	9.4	7.2	6.0	6.7	7.8

Plants/petals F	7a	7b	7c	7d	8a	8b	8c	8d	9a	9b	9c	9d	10a	10b	10c	10d	11a	11b	11c	11d
I	7.6	7.0	8.0	7.7	6.6	6.2	7.0	6.1	7.3	6.6	6.4	6.0	6.5	7.3	7.9	6.9	7.4	5.9	6.9	6.5
II	8.0	6.7	7.1	8.6	7.2	6.7	7.1	6.7	7.5	8.0	9.1	7.8	7.6	7.8	7.8	7.9	7.2	8.6	7.8	8.0
III	7.7	7.8	7.7	7.2	6.6	6.7	6.2	7.0	7.5	7.3	8.2	8.1	7.5	7.6	7.6	7.7	8.2	8.0	9.3	8.7

Plants, labeled 1, 2, ..., 11, where on each plant three flowers (F), labeled I, II, III, with four petals, labeled a, b, c, d, have been collected where $\mathbf{y}_i$: 12×1 is a vector of observations on plant i, $i = 1, \ldots, 11$, μ is a general mean, $\boldsymbol{Z}_1 = \boldsymbol{I}_3 \otimes \mathbf{1}_4$ and $\boldsymbol{Z}_2 = \boldsymbol{I}_3 \otimes \boldsymbol{I}_4$ are the incidence matrices for the random effects, $\boldsymbol{\gamma}_1$ and $\boldsymbol{\gamma}_2$, respectively, $\boldsymbol{\gamma}_1 = (\gamma_1, \gamma_2, \gamma_3)'$: 3×1 is the random vector representing the effect of the hth flower, $h = 1, 2, 3$, of the ith plant, $\boldsymbol{\gamma}_2 = (\gamma_{11}, \ldots, \gamma_{34})'$: 12×1 is the random vector representing the effect of the jth petal, $j = 1, 2, 3, 4$, on the hth flower from the ith plant and $\boldsymbol{\epsilon}_i$ is the vector of random errors, including model errors. Furthermore, assume that $\boldsymbol{\gamma}_1$, $\boldsymbol{\gamma}_2$ and $\boldsymbol{\epsilon}_i$ are independently distributed as $N_3(\mathbf{0}, \boldsymbol{\Sigma}_1)$, $N_{12}(\mathbf{0}, \boldsymbol{\Sigma}_2)$ and $N_{12}(\mathbf{0}, \sigma^2 \boldsymbol{I}_{12})$, correspondingly. The matrices $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ are defined in (3) and (5), respectively and the covariance matrix of $\mathbf{y}_i$ equals $\boldsymbol{\Sigma} = \boldsymbol{Z}_1 \boldsymbol{\Sigma}_1 \boldsymbol{Z}_1' + \boldsymbol{\Sigma}_2 + \sigma^2 \boldsymbol{I}_{12}$.

To illuminate the proposed technique for estimation of variance components in model (34) we apply the restrictions specified in Scenario 2 (iii). The MLE for the general mean μ_L is $\hat{\mu}_L = 7.341$ and for the distinct eigenvalues $\boldsymbol{\eta}_L$ we obtain

$$\hat{\boldsymbol{\eta}}_L = (1.013, 0.250, 0.119, 1.582, 0.346, 0.238)'.$$

Restrictions (iii) in Scenario 2 imply that the following parameters are estimated: $\boldsymbol{\theta}^* = (\sigma^2, \sigma_1, \sigma_2, \tau_2, \tau_3, \tau_6)'$. The MLE of $\boldsymbol{\theta} = (\sigma^2, \sigma_1, \sigma_2, \tau_1, \tau_2, \tau_3, \tau_4, \tau_5, \tau_6)'$ equals

$$\hat{\boldsymbol{\theta}} = (0.250, 0.286, -0.047, 0.019, 0.013, -0.045, -0.026, 0.010, 0.006)',$$

where $\hat{\tau}_1 = -2\hat{\tau}_2 - \hat{\tau}_3$, $\hat{\tau}_4 = (\hat{\tau}_3 - \hat{\tau}_1)/2 + \hat{\tau}_6$ and $\hat{\tau}_5 = -(\hat{\tau}_2 + \hat{\tau}_3)/2 - \hat{\tau}_6$. All estimates make sense, i.e. none of the variances is negatively estimated. When comparing the estimates with the estimates of Scenario 1 (i) and (ii) with estimates of Scenario 2 (iii) it is seen that $\hat{\sigma}^2$ is smallest when Scenario 2 (iii) holds ($\hat{\sigma}^2 = 0.250$). For the other two cases $\hat{\sigma}^2$ equals either 1.013 or 0.346. Since $\hat{\sigma}^2$ is independent of the other parameter estimates, $\hat{\sigma}^2$ mimics the model error. Therefore, it may be stated that among the alternatives discussed in this article Scenario 2 (iii) fits the Kalanchoe plant data best.

Acknowledgments The authors thank two anonymous referees for their helpful comments and suggestions which have helped to improve the paper. Y. Liang's work was in part supported by Hierta-Retzius Foundation grant FOA12H-026 from The Royal Swedish Academy of Sciences. The research of T. von Rosen was supported by the Swedish Research Council grant 2010-18915-75688-45 and by the Estonian Science Foundation grant ETF8294. Those supports are gratefully acknowledged.

References

Demidenko, E. (2004). *Mixed models: theory and applications*. New York: Wiley.

Goldstein, H. (2010). *Multilevel statistical models*. New York: Wiley.

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

LaMotte, L.R. (2007). A direct derivation of the REML likelihood function. *Statistical Papers*, *48*, 321–327.

Liang, Y., von Rosen, T., von Rosen, D. (2011). Block circular symmetry in multilevel models. In: Research Report 2011, vol. 3, Department of Statistics, Stockholm University.

Liang, Y., von Rosen, D., von Rosen, T. (2012). On estimation in multilevel models with block circular symmetric covariance structures. *Acta et Commentationes Universitatis Tartuensis de Mathematica*, *16*, 83–96.

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

Nahtman, T., von Rosen, D. (2008). Shift permutation invariance in linear random factor models. *Mathematical Methods of Statistics*, *17*, 173–185.
Olkin, I. (1973). Testing and estimation for structures which are circularly symmetric in blocks. In D. G. Kabe, R. P. Gupta (Eds.), *Multivariate statistical inference* (pp. 183–195). Amsterdam: North-Holland.
Patterson, H. D., Thompson, R. (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika*, *58*, 545–554.
Searle, S. R., Casella, G., McCulloch, C.E. (1992). *Variance components*. New York: Wiley.
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.
VanLeeuwen, D. M. (1997). A note on the covariance structure in a linear model. *The American Statistician*, *51*, 140–144.
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.