Variance Components Estimation in Mixed Linear Model—The Sub-diagonalization Method

A. Silva, M. Fonseca and J. Mexia

Abstract This work aims to introduce a new method of estimating the variance components in mixed linear models. The approach will be done firstly for models with 3 variances components and secondly attention will be devoted to general case of models with an arbitrary number of variance components. In our approach, we construct and apply a finite sequence of orthogonal matrices to the mixed linear model variance-covariance structure in order to produce a set of Gauss–Markov sub-models which will be used to create pooled estimators for the variance components. Numerical results will be given, comparing the performance of our proposed estimator to the one based on likelihood procedure.

Keywords Mixed linear model · Variance components · Orthogonal matrices · Simultaneous diagonalization

1 Introduction

Mixed linear models (*MLM*) arise due to the necessity of assessing the amount of variation caused by certain sources in a statistical designs with fixed effects (see Khuri [7]), for example, the amount of variations that are not controlled by the experimenters and those whose levels are selected at random. The variances of such sources of variation, currently refereed to as variance components, has been widely investigated in the last fifty years of the last century (see Khuri and Sahai [8], Searle [13, 14], among others) and during the period ranging somewhat from early 1960

A. Silva (✉)
UniCV, Praia, Cabo Verde
e-mail: adilson.dasilva@docente.unicv.edu.cv; ad.silva@campus.fct.unl.pt

A. Silva · M. Fonseca · J. Mexia
UNL, Lisbon, Portugal
e-mail: fmig@fct.unl.pt

J. Mexia
e-mail: jtm@fct.unl.pt

N. Bebiano (ed.), *Applied and Computational Matrix Analysis*,
Springer Proceedings in Mathematics & Statistics 192,
DOI 10.1007/978-3-319-49984-0_21

to 1990, due to the proliferation of investigation on genetic and animal breeding as well as industrial quality control and improvement (for more details, see Anderson [1–3], Anderson and Crump [4], Searle [13], among others), several techniques of estimation have been proposed. Among those techniques we highlight the ANOVA and the maximum likelihood - based methods (see, for example, Searle et al. [15] and Casella and Berger [5]). Nevertheless, notwithstanding the ANOVA method adapt readily to mixed models with balanced data and save the unbiasedness, it does not adapt in situation with unbalanced data (mostly because it use computations derived from fixed effect models rather than mixed models). On its turn, the maximum likelihood - based methods, highlighting the ML and the restricted ML (REML) methods, provide estimators with several statistical optimal properties such as consistency and asymptotic normality either for models with balanced data, or for those with unbalanced data. For these optimal properties we recommend Miller [9], and for some details on applications of such methods we recommend, for example, Anderson [2] and Hartley and Rao [6].

This paper is organized as follows. In Sect. 2 (notation and basic concepts on matrix theory) we review some needed notions and results on matrix theory, mainly on matrix diagonalization. A new method to estimate the variance components in the *MLM* is summarized in Sect. 3, and numerical results ensuring their optimality will be available in Sect. 4.

2 Notation and Basic Concepts on Matrix Theory

In this section we summarize a few needed notions and results on matrix diagonalization. The proofs for the results can be found in Schott [12].

Let $\mathscr{M}^{n\times m}$ and $\mathscr{S}^n = \{A : A \in \mathscr{M}^{n\times n}, A = A^\top\}$ stands for the set of the matrices with n rows and m columns and the set of the $n \times n$ symmetric matrices, respectively. The *range* and the *rank* of a matrix A will be respectively denoted by $R(A)$ and $r(A)$, and the *projection matrix* onto the range space of A denoted by $P_{R(A)}$ (see Schott [12, Chap. 2, Sect. 7] for *projection matrix* notion). We will denote by $tr(A)$ the *trace* of A.

If the eigenvalues $\lambda_1, \ldots, \lambda_r$ of the matrix $M \in \mathscr{M}^{r\times r}$ are all distinct, it follows from the Theorem 3.6 of Schott [12] that the matrix X, whose columns are the eigenvectors associated to those eigenvalues, is non-singular. Thus, by the eigenvalue - eigenvector equation $MX = XD$ or, equivalently, $X^{-1}MX = D$, with $D = diag(\lambda_1 \ldots \lambda_r)$, and the Theorem 3.2.$(d)$ of Schott [12], the eigenvalues of D are the same as those of M. Meanwhile, since M can be transformed into a diagonal matrix by postmultiplication by the non-singular matrix X and premultiplication by its inverse X^{-1} it is said to be diagonalizable.

If the matrix M is symmetric we will have that the eigenvectors associated to its different eigenvalues will be orthogonal (see Schott [12]). Indeed, if we consider two different eigenvalues λ_i and λ_j whose associated eigenvectors are $\mathbf{x}_i$ and $\mathbf{x}_j$, respectively, we see that, since M is symmetric,

$$\lambda_i \mathbf{x}_i^\top \mathbf{x}_j = (M\mathbf{x}_i)^\top \mathbf{x}_j = \mathbf{x}_i^\top (M\mathbf{x}_j) = \lambda_j \mathbf{x}_i^\top \mathbf{x}_j.$$

So, since $\lambda_i \neq \lambda_j$, we must have $\mathbf{x}_i^\top \mathbf{x}_j = 0$.

According with Theorem 3.10 of Schott [12], without lost in generality, the columns of the matrix X can be taken to be orthonormal so that X is an orthogonal matrix. Thus, the eigenvalue - eigenvector equation can now be written as

$$X^\top MX = D \text{ or, equivalently, } M = XDX^\top,$$

which is known as spectral decomposition of M.

Definition 1 Let

$$A = \begin{bmatrix} A_{11} & \dots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \dots & A_{nn} \end{bmatrix}$$

be a diagonal blockwise matrix. We say that a matrix T sub-diagonalizes A if the TA produces a blockwise matrix whose matrices in the diagonal are all diagonal matrices, that is T diagonalizes the matrices $A_{11}, \dots, A_{nn}$ in the diagonal of A.

3 Inference

Variance components estimation in linear models (with mixed and/or fixed effects) have been widely investigated and consequently several methods for estimation with important properties have been derived. Some of this methods are summarized in Searle et al. [15].

In this section we will sub-diagonalize the variance-covariance matrix

$$V = \sum_{d=1}^{r+1} \gamma_d N_d$$

in the Normal *MLM*

$$z \sim \mathcal{N}_m (X\beta,\ V), \tag{1}$$

with $\gamma_d > 0, d = 1, \dots, r$, unknown parameters, $N_d = X_d X_d^\top \in \mathcal{S}^m$, $X_d \in \mathcal{M}^{m\times s}$ known matrices, and $N_{r+1} = I_m$, and develop optimal estimators for the variance components $\gamma_1, \dots, \gamma_{r+1}$.

Since the components we want to estimate depends only on the random effect part, it is of our interest to remove the dependence of the distribution of z on the fixed effect part. With $P_o = P_{R(X)}$ denoting the projection matrix onto the column space of the matrix X, so that $I_m - P_o$ will be the projection matrix onto its orthogonal complement, there is a matrix B_o whose columns are the eigenvectors associated to the null eigenvalues of P_o such that

$$B_o^\top B_o = I_{m-r(P_o)} \text{ and } B_o B_o^\top = I_m - P_o.$$

Thus, instead of the model (1) we will approach the restricted model:

$$y = B_o^\top z \sim \mathcal{N}_n\left(\mathbf{0}_n, \sum_{d=1}^{r+1} \gamma_d M_d\right), \tag{2}$$

with $M_d = B_o^\top N_d B_o$, $n = m - r(P_o)$, and $\mathbf{0}_n$ denotes an $n \times 1$ vector of zeros; that is, we will diagonalize the variance-covariance matrix

$$V^* = \sum_{d=1}^{r+1} \gamma_d M_d$$

instead of V.

3.1 The Case $r = 2$

In this subsection we will sub-diagonalize the variance-covariance matrix in the *MLM* for $r = 2$ (recall the general model in (2)), that is

$$y \sim \mathcal{N}_n\left(\mathbf{0}_n, \ \gamma_1 M_1 + \gamma_2 M_2 + \gamma_3 I_n\right). \tag{3}$$

There exists (see Schott [12, Chap. 4, Sects. 3 and 4]) an orthogonal matrix $P_1 = \begin{bmatrix} A_{11} \\ \vdots \\ A_{1h_1} \end{bmatrix} \in \mathcal{M}^{\left(\sum_{i=1}^{h_1} g_i\right) \times n}$, with $A_{1i} \in \mathcal{M}^{g_i \times n}$ ($\sum_{i=1}^{h_1} g_i = n$), such that $M_1 = P_1^\top D_1 P_1$, or equivalently $P_1 M_1 P_1^\top = D_1$, where

$$D_1 = \begin{bmatrix} \theta_{11} I_{g_1} & 0 & \dots & 0 \\ 0 & \theta_{12} I_{g_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \theta_{1h_1} I_{g_{h_1}} \end{bmatrix} \tag{4}$$

is a diagonal matrix whose diagonal entries θ_{1i}, $i = 1, \dots, h_1$, are the eigenvalues of the matrix M_1 with corresponding roots $g_i = r(A_{1i}^\top)$, $i = 1, \dots, h_1$. It must be noted that the set of columns of each matrix $A_{1i}^\top$ forms a set of g_i orthonormal vectors associated to the eigenvalue θ_{1i} of the matrix M_1 (Theorem 3.10. of Schott [12] guarantees the existence of such matrix $A_{1i}^\top$), so that $A_{1i} A_{1i}^\top = I_{g_i}$ and $A_{1i}^\top A_{1i} =$

$P_{R(A_{1i}^\top)}$. Hence $P_1 P_1^\top = I_n$, and

$$\begin{aligned} P_1^\top P_1 &= A_{11}^\top A_{11} + \cdots + A_{1h_1}^\top A_{1h_1} \\ &= P_{R(A_{11}^\top)} + \cdots + P_{R(A_{1h_1}^\top)} \\ &= I_n. \end{aligned} \tag{5}$$

With

$$A_{1i} M_2 A_{1s}^\top = \begin{cases} M_{ii}^2 & i = s \\ W_{is}^2 & i \neq s \end{cases} \tag{6}$$

and $cov(v)$ denoting the variance-covariance matrix of a random vector v, we will have that

$$\begin{aligned} cov(P_1 y) &= \gamma_1 P_1 M_1 P_1^\top + \gamma_2 P_1 M_2 P_1^\top + \gamma_3 P_1 P_1^\top \\ &= \gamma_1 \begin{bmatrix} \theta_{11} I_{g_1} & 0 & \dots & 0 \\ 0 & \theta_{12} I_{g_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \theta_{1h_1} I_{g_{h_1}} \end{bmatrix} + \gamma_2 \begin{bmatrix} M_{11}^2 & W_{12}^2 & \dots & W_{1h_1}^2 \\ W_{21}^2 & M_{22}^2 & \dots & W_{2h_1}^2 \\ \vdots & \vdots & \ddots & \vdots \\ W_{h_1 1}^2 & W_{h_1 2}^2 & \dots & M_{h_1 h_1}^2 \end{bmatrix} \\ &\quad + \gamma_3 \begin{bmatrix} I_{g_1} & 0 & \dots & 0 \\ 0 & I_{g_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & I_{g_{h_1}} \end{bmatrix} \\ &= \gamma_1 D(\theta_1 I_{g_1} \dots \theta_{h1} I_{g_{h_1}}) + \gamma_2 \Gamma + \gamma_3 D(I_{g_1} \dots I_{g_{h_1}}), \end{aligned} \tag{7}$$

where

$$\Gamma = \begin{bmatrix} M_{11}^2 & W_{12}^2 & \dots & W_{1h_1}^2 \\ W_{21}^2 & M_{22}^2 & \dots & W_{2h_1}^2 \\ \vdots & \vdots & \ddots & \vdots \\ W_{h_1 1}^2 & W_{h_1 2}^2 & \dots & M_{h_1 h_1}^2 \end{bmatrix}.$$

It is clear that for the three matrices $D(\theta_1 I_{g_1} \dots \theta_{h1} I_{g_{h_1}})$, $D(I_{g_1} \dots I_{g_{h_1}})$ and Γ appearing in (7), the blockwise matrix Γ is the only one which is not a diagonal matrix.

Next we diagonalize the symmetric matrices M_{ii}^2, $i = 1, \dots, h_1$, that appear in the diagonal of the matrix Γ, i.e, we sub-diagonalize the matrix Γ.

Since M_{ii}^2 is symmetric there exists (see Schott [12, Chap. 4, Sects. 3 and 4]) an orthogonal matrix $P_{2i} = \begin{bmatrix} A_{2i1} \\ \vdots \\ A_{2ih_{2i}} \end{bmatrix} \in \mathscr{M}^{\left(\sum_{j=1}^{h_{2i}} g_{ij}\right) \times g_i}$, where $A_{2ij} \in \mathscr{M}^{g_{ij} \times g_i}$ $(\sum_{j=1}^{h_{2i}} g_{ij} = g_i)$, such that

$$D_{ii}^2 = P_{2i} M_{ii}^2 P_{2i}^\top = \begin{bmatrix} \theta_{2i1} I_{g_{i1}} & 0 & \dots & 0 \\ 0 & \theta_{2i2} I_{g_{i2}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \theta_{2ih_{2i}} I_{g_{ih_{2i}}} \end{bmatrix}, \; i = 1, \dots, h_1. \tag{8}$$

It must be noted that the matrix $A_{2ij}^\top, i = 1, \dots, h_1, j = 1, \dots, h_{2i}$, is an orthogonal matrix whose columns form a set of $g_{ij} = r(A_{2ij}^\top)$ orthonormal eigenvectors associated to the eigenvalue θ_{2ij} of the matrix M_{ii}^2; that is, g_{ij} is the multiplicity of the eigenvalues θ_{2ij}, and $A_{2ij}^T A_{2ij} = P_{R\left(A_{2ij}^T\right)}$ and $A_{2ij} A_{2ij}^T = I_{g_{ij}}$.

Thus, with

$$P_2 = \begin{bmatrix} P_{21} & 0 & \dots & 0 \\ 0 & P_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & P_{2h_1} \end{bmatrix} \in \mathscr{M}^{\left(\sum_{i=1}^{h_1} \sum_{j=1}^{h_{2i}} g_{ij}\right) \times \left(\sum_{i=1}^{h_1} g_i\right)},$$

the new model $w_2 = P_2 P_1 y$ will have variance-covariance matrix

$$\begin{aligned} cov(w_2) = \Sigma(P_2 P_1 y) &= \gamma_1 P_2 D(\theta_{11} I_{g_1} \dots \theta_{1h1} I_{g_{h_1}}) P_2^\top + \gamma_2 P_2 \Gamma P_2^\top + \gamma_3 P_2 D(I_{g_1} \dots I_{g_{h_1}}) P_2^\top \\ &= \gamma_1 \begin{bmatrix} \theta_{11} P_{21} P_{21}^\top & 0 & \dots & 0 \\ 0 & \theta_{12} P_{22} P_{22}^\top & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \theta_{1h_1} P_{2h_1} P_{2h_1}^\top \end{bmatrix} \\ &+ \gamma_2 \begin{bmatrix} D_{11}^2 & P_{21} W_{12}^2 P_{22}^\top & \dots & P_{21} W_{1h_1}^2 P_{2h_1}^\top \\ P_{22} W_{21}^2 P_{21}^\top & D_{22}^2 & \dots & P_{22} W_{2h_1}^2 P_{2h_1}^\top \\ \vdots & \vdots & \ddots & \vdots \\ P_{2h_1} W_{h_1 1}^2 P_{21}^\top & P_{2h_1} W_{h_1 2}^2 P_{22}^\top & \dots & D_{h_1 h_1}^2 \end{bmatrix} \\ &+ \gamma_3 \begin{bmatrix} P_{21} P_{21}^\top & 0 & \dots & 0 \\ 0 & P_{22} P_{22}^\top & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & P_{2h_1} P_{2h_1}^\top \end{bmatrix}, \end{aligned} \tag{9}$$

where

$$P_{2i} P_{2i}^\top = \begin{bmatrix} A_{2i1} A_{2i1}^\top & 0 & \dots & 0 \\ 0 & A_{2i2} A_{2i2}^\top & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A_{2ih_{2i}} A_{2ih_{2i}}^\top \end{bmatrix} = \begin{bmatrix} I_{g_{i1}} & 0 & \dots & 0 \\ 0 & I_{g_{i2}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & I_{g_{ih_{2i}}} \end{bmatrix},$$

and, with $i \neq s$,

$$P_{2i} W_{is}^2 P_{2s}^\top = \begin{bmatrix} A_{2i1} W_{is}^2 A_{2s1}^\top & A_{2i1} W_{is}^2 A_{2s2}^\top & \dots & A_{2i1} W_{is}^2 A_{2sh_{2s}}^\top \\ A_{2i2} W_{is}^2 A_{2s1}^\top & A_{2i2} W_{is}^2 A_{2s2}^\top & \dots & A_{2i2} W_{is}^2 A_{2sh_{2s}}^\top \\ \vdots & \vdots & \ddots & \vdots \\ A_{2ih_{2i}} W_{is}^2 A_{2s1}^\top & A_{2ih_{2i}} W_{is}^2 A_{2s2}^\top & \dots & A_{2ih_{2i}} W_{is}^2 A_{2sh_{2s}}^\top \end{bmatrix}.$$

The matrix $D_{ii}^2 = P_{2i} M_{ii}^2 P_{2i}^\top$, $i = 1, \dots, h_1$, appearing in the diagonal at the right side of (9) is defined in (8).

Note that

$$w_2 = P_2 P_1 y = \begin{bmatrix} A_{211} A_{11} y \\ \vdots \\ A_{21h_{21}} A_{11} y \\ A_{221} A_{12} y \\ \vdots \\ A_{22h_{22}} A_{12} y \\ \vdots \\ \vdots \\ A_{2h_1 1} A_{1h_1} y \\ \vdots \\ A_{2h_1 h_{2h_1}} A_{1h_1} y \end{bmatrix}.$$

The distribution of the sub-models

$$y_{ij} = A_{2ij} A_{1i} y, \ i = 1, \dots, h_1, \ j = 1, \dots, h_{2i}$$

is summarized in the following result.

Proposition 1

$$y_{ij} \sim \mathcal{N}_{g_{ij}} \left(\boldsymbol{0}_{g_{ij}}, \ \lambda_{ij} I_{g_{ij}} \right), \ i = 1, \dots, h_1; \ j = 1, \dots, h_{2i},$$

where $\lambda_{ij} = \gamma_1 \theta_{1i} + \gamma_2 \theta_{2ij} + \gamma_3$.

Proof Recalling that $A_{2ij} A_{1i} \in \mathcal{M}^{g_{ij} \times n}$ and $g_{ij} \leq n$, according with Moser [10, Theorem 2.1.2] we will have that

$$y_{ij} \sim \mathcal{N}_{g_{ij}} \left(\mathbf{0}_{g_{ij}}, \ \sum_{d=1}^{2} \gamma_d A_{2ij} A_{1i} M_d A_{1i}^\top A_{2ij}^\top + \gamma_3 A_{2ij} A_{1i} A_{1i}^\top A_{2ij}^\top \right).$$

The portions $\sum_{d=1}^{2} \gamma_d A_{2ij} A_{1i} M_d A_{1i}^\top A_{2ij}^\top$ and $\gamma_3 A_{2ij} A_{1i} A_{1i}^\top A_{2ij}^\top$ in the variance-covariance matrix yield:

$$\sum_{d=1}^{2} \gamma_d A_{2ij} A_{1i} M_d A_{1i}^\top A_{2ij}^\top = \gamma_1 A_{2ij} \left(\theta_{1i} I_{g_i}\right) A_{2ij}^\top + \gamma_2 A_{2ij} M_{ii}^2 A_{2ij}^\top$$
$$= \gamma_1 \theta_{1i} I_{g_{ij}} + \gamma_2 \theta_{2ij} I_{g_{ij}};$$

and

$$\gamma_3 A_{2ij} A_{1i} A_{1i}^\top A_{2ij}^\top = \gamma_3 A_{2ij} I_{g_i} A_{2ij}^\top = \gamma_3 I_{g_{ij}}$$

which, clearly, completes the proof. □

With $\mathbf{0}$ denoting an adequate null matrix and $cov(\nu, \upsilon)$ denoting the cross-covariance between the random vectors ν and υ, from (9) one might note that the cross-covariance matrix between the sub-models $y_{ij} = A_{2ij} A_i y$ and $y_{sk} = A_{2sk} A_s y$, $i, s = 1, \ldots, h_1$, $j, k = 1, \ldots, h_{2i}$ is given by

$$cov(y_{ij},\ y_{sk}) = \gamma_2 A_{2ij} A_{1i} M_2 A_{1s}^\top A_{2sk}^\top = \begin{cases} \mathbf{0} & i = s;\ j \neq k \\ \lambda_{ij} & i = s;\ j = k \\ \gamma_2 A_{2ij} W_{is}^2 A_{2sk}^\top & i \neq s \end{cases} \quad (10)$$

with $i \leq s$, $j \leq k$ (symmetry applies), so that, for $i \neq s$, the sub-models y_{ij} and y_{sk} are correlated and for $i = s$ they are not.

3.2 Estimation for r = 2

From the Sect. 3.1 we see that (with i and j respectively replaced by i_1 and i_2, for convenience) $w_2 = P_2 P_1 y$ produces the following sub-models

$$y_{i_1 i_2} \sim \mathcal{N}_{g_{i_1 i_2}}(\mathbf{0}_{g_{i_1 i_2}},\ \lambda_{i_1 i_2} I_{g_{i_1 i_2}}),\ i_1 = 1, \ldots, h_1,\ i_2 = 1, \ldots, h_{2i_1}, \quad (11)$$

of the model $y \sim \mathcal{N}_n(\mathbf{0}_n,\ \gamma_1 M_1 + \gamma_2 M_2 + \gamma_3 I_n)$, where

$$\lambda_{i_1 i_2} = \gamma_1 \theta_{1 i_1} + \gamma_2 \theta_{2 i_1 i_2} + \gamma_3.$$

An unbiased estimator of $\lambda_{i_1 i_2}$ for model (11) is (one based on its maximum likelihood estimator $\hat{\lambda}_{i_1 i_2}$)

$$S_{i_1 i_2}^2 = \frac{y_{i_1 i_2}^\top y_{i_1 i_2}}{g_{i_1 i_2}},$$
$$i_1 = 1, \ldots, h_1,\ i_2 = 1, \ldots, h_{2i_1}.$$

Indeed (see Rencher and Schaalje [11, Theorem 5.2a]),

$$
\begin{aligned}
E(S_{i_1 i_2}^2) &= \frac{1}{g_{i_1 i_2}} tr\left\{\lambda_{i_1 i_2} I_{g_{i_1 i_2}}\right\} \\
&= \lambda_{i_1 i_2}.
\end{aligned} \tag{12}
$$

Thus

$$
E(S_{i_1 i_2}^2) = \lambda_{i_1 i_2} = \gamma_1 \theta_{1 i_1} + \gamma_2 \theta_{2 i_1 i_2} + \gamma_3, \quad i_1 = 1, \ldots, h_1, \; i_2 = 1, \ldots, h_{2 i_1}
$$

so that, with $S = \begin{bmatrix} S_{11}^2 \\ \cdots \\ S_{1h_{21}}^2 \\ S_{21}^2 \\ \cdots \\ S_{2h_{22}}^2 \\ \cdots \\ \cdots \\ S_{h_1 1}^2 \\ \cdots \\ S_{h_1 h_{2h_1}}^2 \end{bmatrix}$, $\Theta = \begin{bmatrix} \theta_{11} & \theta_{211} & 1 \\ \cdots & \cdots & \cdots \\ \theta_{11} & \theta_{21h_{21}} & 1 \\ \theta_{12} & \theta_{221} & 1 \\ \cdots & \cdots & \cdots \\ \theta_{12} & \theta_{22h_{22}} & 1 \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ \theta_{1h_1} & \theta_{2h_1 1} & 1 \\ \cdots & \cdots & \cdots \\ \theta_{1h_1} & \theta_{2h_1 h_{2h_1}} & 1 \end{bmatrix}$, and $\gamma = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{bmatrix}$, we will have

$$
E(S) = \Theta \gamma. \tag{13}
$$

Thus, for $i_1 = 1, \ldots, h_1, \; i_2 = 1, \ldots, h_{2i_1}$, equalizing the variances $\lambda_{i_1 i_2}$ to the correspondent estimators $S_{i_1 i_2}^2$ it yields the following system of equations:

$$
\begin{aligned}
S_{11}^2 &= \gamma_1 \theta_{11} + \gamma_2 \theta_{211} + \gamma_3; \\
\cdots \quad & \quad \cdots\cdots\cdots\cdots\cdots; \\
S_{1h_{21}}^2 &= \gamma_1 \theta_{11} + \gamma_2 \theta_{21h_{21}} + \gamma_3; \\
S_{21}^2 &= \gamma_1 \theta_{12} + \gamma_2 \theta_{221} + \gamma_3; \\
\cdots \quad & \quad \cdots\cdots\cdots\cdots\cdots \\
S_{2h_{22}}^2 &= \gamma_1 \theta_{12} + \gamma_2 \theta_{22h_{22}} + \gamma_3; \\
\cdots \quad & \quad \cdots\cdots\cdots\cdots\cdots; \\
\cdots \quad & \quad \cdots\cdots\cdots\cdots\cdots; \\
S_{h_1 1}^2 &= \gamma_1 \theta_{1h_1} + \gamma_2 \theta_{2h_1 1} + \gamma_3; \\
\cdots \quad & \quad \cdots\cdots\cdots\cdots\cdots; \\
S_{h_1 h_{2h_1}}^2 &= \gamma_1 \theta_{1h_1} + \gamma_2 \theta_{2h_1 h_{2h_1}} + \gamma_3;
\end{aligned}
$$

which in matrix notation becomes

$$
S = \Theta \gamma. \tag{14}
$$

Since by construction $\theta_{1i_1} \neq \theta_{1i_1'}$, $i_1 \neq i_1' = 1, \ldots, h_1$ (they are the different eigenvalues of M_1) and $\theta_{2i_1i_2} \neq \theta_{2i_1i_2'}$, $i_2 \neq i_2' = 1, \ldots, h_{2i_1}$ (they are the distinct eigenvalues of $M_{ii}^2 = A_{1i_1} M_2 A_{1i_1}^\top$), it is easily seen that the matrix Θ is a full rank one; that is $r(\Theta) = 3$.

By Rencher and Schaalje [11, Theorem 2.6*d*] the matrix

$$\Theta^\top \Theta = \begin{bmatrix} \sum_{i_1}^{h_1} \sum_{i_2}^{h_{2i_1}} \theta_{1i_1}^2 & \sum_{i_1}^{h_1} \sum_{i_2}^{h_{2i_1}} \theta_{1i_1}\theta_{2i_1i_2} & \sum_{i_1}^{h_1} \sum_{i_2}^{h_{2i_1}} \theta_{1i_1} \\ \sum_{i_1}^{h_1} \sum_{i_2}^{h_{2i_1}} \theta_{1i_1}\theta_{2i_1i_2} & \sum_{i_1}^{h_1} \sum_{i_2}^{h_{2i_1}} \theta_{2i_1i_2}^2 & \sum_{i_1}^{h_1} \sum_{i_2}^{h_{2i_1}} \theta_{2i_1i_2} \\ \sum_{i_1}^{h_1} \sum_{i_2}^{h_{2i_1}} \theta_{1i_1} & \sum_{i_1}^{h_1} \sum_{i_2}^{h_{2i_1}} \theta_{2i_1i_2} & \sum_{i_1}^{h_1} \sum_{i_2}^{h_{2i_1}} \end{bmatrix}$$

is positive-definite, and by Rencher and Schaalje [11, Corollary 1], $\Theta^\top \Theta$ is non-singular; we, thus, take its inverse to be $(\Theta^\top \Theta)^{-1}$.

Now, premultiplying the system (14) in both side by $\Theta^\top$ the resulting system of equations will be

$$\Theta^\top S = \Theta^\top \Theta \gamma, \tag{15}$$

whose unique solution (and therefore an estimator of γ) is

$$\hat{\gamma} = (\Theta^\top \Theta)^{-1} \Theta^\top S. \tag{16}$$

$\hat{\gamma} = \begin{bmatrix} \hat{\gamma}_1 \\ \hat{\gamma}_2 \\ \hat{\gamma}_3 \end{bmatrix}$ will be referred to as *Sub-D estimator* and the underlying method referred to as *Sub-D method*.

Proposition 2 *$\hat{\gamma}$ is an unbiased estimator of γ, with $\gamma = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{bmatrix}$.*

Proof Indeed, $E(\hat{\gamma}) = E\left((\Theta^\top \Theta)^{-1} \Theta^\top S\right) = (\Theta^\top \Theta)^{-1} \Theta^\top E(S) = (\Theta^\top \Theta)^{-1}$ $\Theta^\top \Theta \gamma = \gamma$. □

Proposition 3 *With $i \leq i^*$, $j \leq j^*$ (symmetry applies),*

$$cov\left(S_{ij}^2,\ S_{i^*j^*}^2\right) = \begin{cases} (\boldsymbol{a})\ i = i^*;\ j \neq j^* : & 0, \\ (\boldsymbol{b})\ i = i^*;\ j = j^* : & 2\frac{\lambda_{ij}^2}{g_{ij}}, \\ (\boldsymbol{c})\ i \neq i^* : & 2\gamma_2^2 tr(\Omega M_2), \end{cases}$$

where $\Omega = \nabla_{ij} M_2 \nabla_{i^*j^*}$, with $\nabla_{ij} = \frac{A_{1i}^\top A_{2ij}^\top A_{2ij} A_{1i}}{g_{ij}}$.

Proof We have that

$$
\begin{aligned}
cov\left(S_{ij}^2, S_{i^*j^*}^2\right) &= cov\left(\frac{y_{ij}^\top y_{ij}}{g_{ij}}, \ \frac{y_{i^*j^*}^\top y_{i^*j^*}}{g_{i^*j^*}}\right) \\
&= cov\left(y^\top\left(\frac{A_{1i}^\top A_{2ij}^\top A_{2ij}A_{1i}}{g_{ij}}\right)y, \ y^\top\left(\frac{A_{1i^*}^\top A_{2i^*j^*}^\top A_{2i^*j^*}A_{1i^*}}{g_{i^*j^*}}\right)y\right) \\
&= cov\left(y^\top \nabla_{ij} y, \ y^\top \nabla_{i^*j^*} y\right) \\
&= 2tr\left(\nabla_{ij} V \nabla_{i^*j^*} V\right) \\
&= 2\gamma_1^2 tr(\nabla_{ij} M_1 \nabla_{i^*j^*} M_1) + 2\gamma_1\gamma_2 tr(\nabla_{ij} M_1 \nabla_{i^*j^*} M_2) + 2\gamma_1\gamma_3 tr(\nabla_{ij} M_1 \nabla_{i^*j^*}) \\
&+ 2\gamma_2\gamma_1 tr(\nabla_{ij} M_2 \nabla_{i^*j^*} M_1) + 2\gamma_2^2 tr(\nabla_{ij} M_2 \nabla_{i^*j^*} M_2) + 2\gamma_2\gamma_3 tr(\nabla_{ij} M_2 \nabla_{i^*j^*}) \\
&+ 2\gamma_3\gamma_1 tr(\nabla_{ij} \nabla_{i^*j^*} M_1) + 2\gamma_3\gamma_2 tr(\nabla_{ij} \nabla_{i^*j^*} M_2) + 2\gamma_3^2 tr(\nabla_{ij} \nabla_{i^*j^*}) \\
&= \begin{cases} i = i^*; \, j \neq j^* : & 0, \\ i = i^*; \, j = j^* : & 2\frac{\lambda_{ij}^2}{g_{ij}}, \\ i \neq i^* : & 2\gamma_2^2 tr(\nabla_{ij} M_2 \nabla_{i^*j^*} M_2). \end{cases}
\end{aligned}
$$

For the case (**a**), that is $i = i^*; \, j \neq j^*$, we have that

$$
\begin{aligned}
\nabla_{ij} M_1 \nabla_{ij^*} &= \frac{1}{g_{ij} g_{ij^*}} A_{1i}^\top A_{2ij}^\top A_{2ij} A_{1i} M_1 A_{1i}^\top A_{2ij^*}^\top A_{2ij^*} A_{1i} \\
&= \frac{1}{g_{ij} g_{ij^*}} A_{1i}^\top A_{2ij}^\top A_{2ij} \left(\theta_{1i} I_{g_i}\right) A_{2ij^*}^\top A_{2ij^*} A_{1i} \\
&= \mathbf{0}_{g_i \times g_i} \ \text{(see (4) for the explanation);} \qquad (17)
\end{aligned}
$$

$$
\begin{aligned}
\nabla_{ij} M_2 \nabla_{ij^*} &= \frac{1}{g_{ij} g_{ij^*}} A_{1i}^\top A_{2ij}^\top A_{2ij} A_{1i} M_2 A_{1i}^\top A_{2ij^*}^\top A_{2ij^*} A_{1i} \\
&= \frac{1}{g_{ij} g_{ij^*}} A_{1i}^\top A_{2ij}^\top A_{2ij} \left(M_{ii}^2\right) A_{2ij^*}^\top A_{2ij^*} A_{1i} \\
&= \mathbf{0}_{g_i \times g_i} \ \text{(see (8) for the explanation);} \qquad (18)
\end{aligned}
$$

$$
\begin{aligned}
\nabla_{ij} \nabla_{ij^*} &= \frac{1}{g_{ij} g_{ij^*}} A_{1i}^\top A_{2ij}^\top \left(\mathbf{0}_{g_{ii} \times g_{ij^*}}\right) A_{2ij^*} A_{1i} \\
&= \mathbf{0}_{g_i \times g_i}. \qquad (19)
\end{aligned}
$$

Therefore, (17)–(19) together with Schott [12, Theorem 1.3.(d)] proves the case (**a**).

For the case (**c**), that is $i \neq i^*$, the desired result becomes clear if use the Theorem 1.3.(d) of Schott [12] and note that

$$
A_{1i} M_1 A_{1i^*} = A_{1i} A_{1i^*} = \mathbf{0}_{g_i \times g_{i^*}}.
$$

Finally, for the case (**b**), that is $i = i^*$; $j = j^*$, recalling $y_{ij} \sim \mathcal{N}_n\left(\mathbf{0}_{g_{ij}},\ \lambda_{ij} I_{g_{ij}}\right)$, it holds

$$
\begin{aligned}
cov\left(S_{ij}^2\right) &= \Sigma\left(\frac{y_{ij}^\top y_{ij}}{g_{ij}}, \frac{y_{ij}^\top y_{ij}}{g_{ij}}\right) = 2tr\left\{\frac{\lambda_{ij}}{g_{ij}} I_{g_{ij}} \frac{\lambda_{ij}}{g_{ij}} I_{g_{ij}}\right\} = 2\frac{\lambda_{ij}^2}{g_{ij}^2} tr\left\{I_{g_{ij}}\right\} \\
&= 2\frac{\lambda_{ij}^2}{g_{ij}},
\end{aligned}
\tag{20}
$$

and therefore the proof is complete. □

The next result introduce the variance-covariance matrix of the sub-diagonalization estimator:

$$\hat{\gamma} = (\Theta^\top \Theta)^{-1} \Theta^\top S.$$

Proposition 4 *In order to simplify the notation, let* $\Sigma_{S_{ij} S_{kl}}$ *denote* $cov(S_{ij}^2,\ S_{kl}^2)$. *Then,*

$$cov(\hat{\gamma}) = (\Theta^\top \Theta)^{-1} \Theta^\top cov(S) \Theta (\Theta^\top \Theta)^{-1}, \tag{21}$$

where $cov(S) = \begin{bmatrix} D_1 & \Lambda_{12} & \Lambda_{13} & \dots & \Lambda_{1h_1} \\ \Lambda_{21} & D_2 & \Lambda_{23} & \dots & \Lambda_{2h_1} \\ \Lambda_{31} & \Lambda_{32} & D_3 & \dots & \Lambda_{3h_1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Lambda_{h_1 1} & \Lambda_{h_1 2} & \Lambda_{h_1 3} & \dots & D_{h_1} \end{bmatrix}$, *with* $D_i = 2 \begin{bmatrix} \frac{\lambda_{i1}^2}{g_{i1}} & 0 & \dots & 0 \\ 0 & \frac{\lambda_{i2}^2}{g_{i2}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{\lambda_{ih_{2i}}^2}{g_{ih_{2i}}} \end{bmatrix}$ *and*

$$\Lambda_{ks} = \begin{bmatrix} \Sigma_{S_{k1} S_{s1}} & \Sigma_{S_{k1} S_{s2}} & \dots & \Sigma_{S_{k1} S_{sh_{2s}}} \\ \Sigma_{S_{k2} S_{s1}} & \Sigma_{S_{k2} S_{s2}} & \dots & \Sigma_{S_{k2} S_{sh_{2s}}} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{S_{kh_{2k}} S_{s1}} & \Sigma_{S_{kh_{2k}} S_{s2}} & \dots & \Sigma_{S_{kh_{2k}} S_{sh_{2s}}} \end{bmatrix}.$$

Proof The proof is a consequence of the Proposition 3. □

3.3 The General Case: $r \geq 1$

Now, without lost in generality, lets consider the general *MLM* in (2):

$$y \sim \mathcal{N}_n\left(\mathbf{0}_n, \sum_{d=1}^{r+1} \gamma_d M_d\right), \text{ with } M_d = X_d X_d^\top \in \mathcal{S}^n \text{ and } M_{r+1} = I_n.$$

One may note that $y = \sum_{d=1}^{r+1} B_o^\top X_d \beta_d$, where $\beta_d \sim \mathcal{N}(0,\ \gamma_d I)$, $d = 1, \ldots, r$, $\beta_{r+1} \sim \mathcal{N}(0,\ \gamma_d I_n)$, and $\beta_1, \ldots, \beta_{r+1}$ are not correlated.

With $i_1 = 1, \ldots, h_1,\ i_j = 1, \ldots, h_{j,i_1,\ldots,i_{j-1}}$, consider the finite sequence of r matrices P_1, P_2, …, P_r defined as follow:

$$P_1 = \begin{bmatrix} A_{11} \\ A_{12} \\ \vdots \\ A_{1h_1} \end{bmatrix} \in \mathscr{M}^{\left(\sum_{i_1}^{h_1} g_{i_1}\right)\times n},\ \text{with } A_{1i_1} \in \mathscr{M}^{(g_{i_1})\times n} \left(\text{note:} \sum_{i_1}^{h_1} g_{i_1} = n\right); \quad (22)$$

$$P_2 = \begin{bmatrix} P_{21} & 0 & \ldots & 0 \\ 0 & P_{22} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & P_{2h_1} \end{bmatrix} \in \mathscr{M}^{\left(\sum_{i_1}^{h_1} \sum_{i_2}^{h_{2,i_1}} g_{i_1 i_2}\right)\times\left(\sum_{i_1}^{h_1} g_{i_1}\right)},\ \text{where}$$

$$P_{2i_1} = \begin{bmatrix} A_{2i_1 1} \\ A_{2i_1 2} \\ \vdots \\ A_{2i_1 h_{2i_1}} \end{bmatrix} \in \mathscr{M}^{\left(\sum_{i_2}^{h_{2,i_1}} g_{i_1 i_2}\right)\times g_{i_1}},\ \text{with} \sum_{i_2}^{h_{2,i_1}} g_{i_1 i_2} = g_{i_1} \text{ and } A_{2i_1 i_2} \in \mathscr{M}^{g_{i_1 i_2}\times g_{i_1}};$$

$$P_3 = \begin{bmatrix} P_{31} & 0 & \ldots & 0 \\ 0 & P_{32} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & P_{3h_1} \end{bmatrix} \in \mathscr{M}^{\left(\sum_{i_1}^{h_1} \sum_{i_2}^{h_{2,i_1}} \sum_{i_3}^{h_{3,i_1,i_2}} g_{i_1 i_2 i_3}\right)\times\left(\sum_{i_1}^{h_1} \sum_{i_2}^{h_{2,i_1}} g_{i_1 i_2}\right)},$$

$$\text{where } P_{3i_1} = \begin{bmatrix} P_{3i_1 1} & 0 & \ldots & 0 \\ 0 & P_{3i_1 2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & P_{3i_1 h_{2,i_1}} \end{bmatrix} \in \mathscr{M}^{\left(\sum_{i_2}^{h_{2,i_1}} \sum_{i_3}^{h_{3,i_1,i_2}} g_{i_1 i_2 i_3}\right)\times\left(\sum_{i_2}^{h_{2,i_1}} g_{i_1 i_2}\right)} \text{ and}$$

$$P_{3i_1 i_2} = \begin{bmatrix} A_{3i_1 i_2 1} \\ A_{3i_1 i_2 2} \\ \vdots \\ A_{3i_1 i_2 h_{3,i_1,i_2}} \end{bmatrix} \in \mathscr{M}^{\left(\sum_{i_3}^{h_{3,i_1,i_2}} g_{i_1 i_2 i_3}\right)\times g_{i_1 i_2}},\ \text{with} \sum_{i_3}^{h_{3,i_1,i_2}} g_{i_1 i_2 i_3} = g_{i_1 i_2} \text{ and}$$

$$A_{3i_1 i_2 i_3} \in \mathscr{M}^{g_{i_1 i_2 i_3}\times g_{i_1 i_2}};$$

Thus, for $r \geq 2$, each matrix P_r will be given by (P_1 is given in (22)):

$$P_r = \begin{bmatrix} P_{r1} & 0 & \dots & 0 \\ 0 & P_{r2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & P_{rh_1} \end{bmatrix} \in \mathscr{M}^{\left(\sum_{i_1}^{h_1} \cdots \sum_{i_r}^{h_{r,i_1,\dots,i_{r-1}}} g_{i_1 \dots i_r}\right) \times \left(\sum_{i_1}^{h_1} \cdots \sum_{i_{(r-1)}}^{h_{(r-1),i_1,\dots,i_{r-2}}} g_{i_1 \dots i_{(r-1)}}\right)}, \tag{23}$$

where

$$P_{ri_1} = \begin{bmatrix} P_{ri_1 1} & 0 & \dots & 0 \\ 0 & P_{ri_1 2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & P_{ri_1 h_{2,i_1}} \end{bmatrix} \in \mathscr{M}^{\left(\sum_{i_2}^{h_{2,i_1}} \cdots \sum_{i_r}^{h_{r,i_1,\dots,i_{r-1}}} g_{i_1 \dots i_r}\right) \times \left(\sum_{i_2}^{h_{2,i_1}} \cdots \sum_{i_{(r-1)}}^{h_{(r-1),i_1,\dots,i_{r-2}}} g_{i_1 \dots i_{(r-1)}}\right)},$$

$$\dots \dots \dots \dots \dots$$

$$P_{ri_1 \dots i_{(r-2)}} = \begin{bmatrix} P_{ri_1 \dots i_{(r-2)} 1} & 0 & \dots & 0 \\ 0 & P_{ri_1 \dots i_{(r-2)} 2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & P_{ri_1 \dots i_{(r-2)} h_{r-1,i_1,\dots,i_{r-2}}} \end{bmatrix} \in \mathscr{M}^{\left(\sum_{i_{(r-1)}}^{h_{(r-1),i_1,\dots,i_{r-2}}} \sum_{i_r}^{h_{r,i_1,\dots,i_{r-1}}} g_{i_1 \dots i_r}\right) \times \left(\sum_{i_{(r-1)}}^{h_{(r-1),i_1,\dots,i_{r-2}}} g_{i_1 \dots i_{(r-1)}}\right)},$$

$$\text{and} \quad P_{ri_1 \dots i_{(r-1)}} = \begin{bmatrix} A_{ri_1 \dots i_{(r-1)} 1} \\ A_{ri_1 \dots i_{(r-1)} 2} \\ \vdots \\ A_{ri_1 \dots i_{(r-1)} h_{r,i_1,\dots,i_{r-1}}} \end{bmatrix} \in \mathscr{M}^{\left(\sum_{i_r}^{h_{r,i_1,\dots,i_{r-1}}} g_{i_1 \dots i_r}\right) \times g_{i_1 \dots i_{(r-1)}}},$$

$$\text{with} \quad \sum_{i_r}^{h_{r,i_1,\dots,i_{r-1}}} g_{i_1 \dots i_r} = g_{i_1 \dots i_{(r-1)}}, \sum_{i_1}^{h_1} g_{i_1} = n, \; A_{ri_1 \dots i_r} \in \mathscr{M}^{g_{i_1 \dots i_r} \times g_{i_1 \dots i_{(r-1)}}};$$

Theorem 1 *Let the matrices $P_1, P_2, \dots, P_r$ defined above be such that:*

(c_1) The columns of $A_{1i_1}^{\top}$, $i_1 = 1, \dots, h_1$, form a set of $g_{i_1} = r(A_{1i_1}^{\top})$ orthonormal eigenvectors associated to the eigenvalues θ_{1i_1} of the matrix M_1 (θ_{1i_1} has multiplicity g_{i_1});

(c_2) *The columns of* $A^{\top}_{2i_1i_2}$, $i_2 = 1, \ldots, h_{2,i_1}$, *form a set of* $g_{i_1i_2} = r(A^{\top}_{2i_1i_2})$ *orthonormal eigenvectors associated to the eigenvalues* $\theta_{2i_1i_2}$ *of the matrix* $M^2_{i_1i_1} = A_{1i_1}M_2A^{\top}_{1i_1}$ ($\theta_{2i_1i_2}$ *has multiplicity* $g_{i_1i_2}$);

(c_3) *The columns of* $A^{\top}_{3i_1i_2i_3}$, $i_3 = 1, \ldots, h_{3,i_1,i_2}$, *form a set of* $g_{i_1i_2i_3} = r(A^{\top}_{3i_1i_2i_3})$ *orthonormal eigenvectors associated to the eigenvalues* $\theta_{3i_1i_2i_3}$ *of the matrix*

$$A_{2i_1i_2}M^3_{i_1i_1}A^{\top}_{2i_1i_2} = A_{2i_1i_2}A_{1i_1}M_3A^{\top}_{1i_1}A_{2i_1i_2}$$

($\theta_{3i_1i_2i_3}$ *has multiplicity* $g_{i_1i_2i_3}$);

…………

(c_r) *The columns of* $A^{\top}_{ri_1\ldots i_r}$, $i_r = 1, \ldots, h_{r,i_1,\ldots,i_{r-1}}$, *form a set of* $g_{i_1\ldots i_r} = r(A^{\top}_{ri_1\ldots i_r})$ *orthonormal eigenvectors associated to the eigenvalues* $\theta_{ri_1\ldots i_r}$ *of the matrix*

$$A_{(r-1)i_1\ldots i_{(r-1)}} \ldots A_{1i_1}M_rA^{\top}_{1i_1} \ldots A^{\top}_{(r-1)i_1\ldots i_{(r-1)}}$$

($\theta_{ri_1\ldots i_r}$ *has multiplicity* $g_{i_1\ldots i_r}$).

Then each matrix P_d, $d = 1, \ldots, r$, in the finite sequence of matrices $P_1, P_2, \ldots, Pr$ will be an orthogonal matrix.

Proof By the way P_d is defined (see (23)), since

$$P_{di_1\ldots i_{(d-1)}} = \begin{bmatrix} A_{di_1\ldots i_{(d-1)}1} \\ A_{di_1\ldots i_{(d-1)}2} \\ \vdots \\ A_{di_1\ldots i_{(d-1)}h_{d,i_1,\ldots,i_{d-1}}} \end{bmatrix}, \ i_{(d-1)} = 1, \ldots, h_{(d-1),i_1,\ldots,i_{d-2}},$$

and according with condition c_d we see that the matrices $P_{di_1\ldots i_{(d-1)}}$ are orthogonal. Thus, the desired result comes if we see that $P_d^{\top}P_d$ will be a diagonal blockwise matrix whose diagonal entries are $P^{\top}_{di_1}P_{di_1}$, $i_1 = 1, \ldots, h_1$. The diagonal entries $P^{\top}_{di_1}P_{di_1}$ will be diagonal blockwise matrices whose diagonal entries will be $P^{\top}_{di_1i_2}P_{di_1i_2}$, $i_2 = 1, \ldots, h_{2,i_1}$. Proceeding this way $d-2$ times, we will find that the diagonal entries of the blockwise matrices $P^{\top}_{di_1\ldots i_{(d-2)}}P_{di_1\ldots i_{(d-2)}}$, $i_{(d-2)} = 1, \ldots, h_{(d-2),i_1,\ldots,i_{d-3}}$, will be

$$\begin{aligned} P^{\top}_{di_1\ldots i_{(d-1)}}P_{di_1\ldots i_{(d-1)}} &= A^{\top}_{di_1\ldots i_{(d-1)}1}A_{di_1\ldots i_{(d-1)}1} \\ &\quad + \cdots + A^{\top}_{di_1\ldots i_{(d-1)}h_{d,i_1,\ldots,i_{d-1}}}A_{di_1\ldots i_{(d-1)}h_{d,i_1,\ldots,i_{d-1}}} \\ &= I_{g_{i_1\ldots i_{(d-1)}}}, \end{aligned}$$

reaching, therefore, the desired result. Proceeding in same way we would also see that $P_{di_1\ldots i_{(d-1)}}P^{\top}_{di_1\ldots i_{(d-1)}}$ is a Blockwise diagonal matrix whose diagonal entries are $A_{di_1\ldots i_{(d-1)}1}A^{\top}_{di_1\ldots i_{(d-1)}j}$, $j = 1, \ldots, h_{d,i_1,\ldots,i_{d-1}}$, so that $P_dP_d^{\top}$ is an identity matrix. □

The model $w_r = P_r \dots P_2 P_1 y$ will produces the following sub - models:

$$y_{i_1 \dots i_r} = A_{r i_1 \dots i_r} A_{(r-1) i_1 \dots i_{(r-1)}} \dots A_{2 i_1 i_2} A_{1 i_1} y,$$
$$i_1 = 1, \dots, h_1, i_j = 1, \dots, h_{j, i_1, \dots, i_{j-1}}.$$

We summarize the distribution of each of the sub-model $y_{i_1 \dots i_r}$ in the following result.

Proposition 5

$$y_{i_1 \dots i_r} \sim \mathcal{N}_{g_{i_1 \dots i_r}} \left(0_{g_{i_1 \dots i_r}},\ \lambda_{i_1 \dots i_r} I_{g_{i_1 \dots i_r}}\right),$$

where $\lambda_{i_1 \dots i_r} = \sum_{d=1}^{r} \gamma_d \theta_{d i_1 \dots i_d} + \gamma_{r+1}$.

Proof The proof becomes obvious after looking to the proofs of the Proposition 1. □

From the results about cross-covariance on the preceding sections we easily conclude that the cross-covariance matrix between the sub-models $y_{i_1 \dots i_r}$ and $y_{i_1^* \dots i_r^*}$, with $i_1, i_1^* = 1, \dots, h_1; i_j, i_j^* = 1, \dots, h_{j, i_1, \dots, i_{j-1}}$, is given by

$$cov(y_{i_1 \dots i_r},\ y_{i_1^* \dots i_r^*}) = \begin{cases} 0 & i_1 = i_1^*, \\ \lambda_{i_1 \dots i_r} & i_j = i^* j \\ & j = 1, \dots, h_{j, i_1, \dots, i_{j-1}} \\ \sum_{d=2}^{r} \gamma_d A_{r i_1 \dots i_r} \dots A_{1 i} M_d A_{1 i_1^*}^{\top} \dots A_{r i_1^* \dots i_r^*} & i_1 \neq i_1^* \end{cases}$$

so that, for $i_1 \neq i_1^*$, the sub-models $y_{i_1 \dots i_r}$ and $y_{i_1^* \dots i_r^*}$ are correlated and for $i_1 = i_1^*$ they are not.

3.4 *Estimation for the General Case: r ≥ 1*

Recalling that for the *MLM* in (1), $P_r \dots P_2 P_1 y$ produces the following sub-models

$$y_{i_1 i_2 \dots i_r} \sim \mathcal{N}_{g_{i_1 i_2 \dots i_r}} (0_{g_{i_1 \dots i_r}},\ \lambda_{i_1 i_2 \dots i_r} I_{g_{i_1 i_2 \dots i_r}}),$$
$$i_1 = 1, \dots, h_1, i_j = 1, \dots, h_{j, i_1, \dots, i_{j-1}} \quad (24)$$

where

$$\lambda_{i_1 i_2 \dots i_r} = \sum_{d=1}^{r} \gamma_d \theta_{d i_1 \dots i_d} + \gamma_{r+1}.$$

The matrices P_d, $d = 1, \dots, r$, are defined in the Sect. 3.3.

An unbiased estimator of $\lambda_{i_1 i_2 \ldots i_r}$ in the sub-model (24) is (the one based on its maximum likelihood estimator $\hat{\lambda}_{i_1 i_2 \ldots i_r}$)

$$S^2_{i_1 i_2 \ldots i_r} = \frac{1}{g_{i_1 i_2 \ldots i_r}} y^{\top}_{i_1 i_2 \ldots i_r} y_{i_1 i_2 \ldots i_r}$$

Indeed (see Rencher and Schaalje [11], Theorem 5.2(a), and the explanation for (12)),

$$\begin{aligned} E\left(S^2_{i_1 i_2 \ldots i_r}\right) &= \frac{\lambda_{i_1 i_2 \ldots i_r}}{g_{i_1 i_2 \ldots i_r}} tr\left[I_{g_{i_1 i_2 \ldots i_r}}\right] \\ &= \lambda_{i_1 i_2 \ldots i_r}. \end{aligned} \tag{25}$$

For convenience, in what follows, instead of $S^2_{i_1 i_2 \ldots i_r}$, we may sometimes use the notation $S^2_{i_1 i_2 \ldots i_{(r-1)} i_r}$.

Thus

$$\begin{aligned} E(S^2_{i_1 i_2 \ldots i_{(r-1)} i_r}) &= \sum_{d=1}^{r} \gamma_d \theta_{d i_1 \ldots i_d} + \gamma_{r+1} \\ &= \gamma_1 \theta_{1 i_1} + \gamma_2 \theta_{2 i_1 i_2} + \cdots + \gamma_r \theta_{r i_1 i_2 \ldots i_{(r-1)} i_r} + \gamma_{r+1}, \end{aligned}$$

$$i_1 = 1, \ldots, h_1; i_j = 1, \ldots, h_{j, i_1, \ldots, i_{j-1}}$$

so that, with $S = \begin{bmatrix} S^2_{11\ldots11} \\ S^2_{11\ldots12} \\ \ldots \\ S^2_{11\ldots1h_{r,1,\ldots,1}} \\ S^2_{11\ldots21} \\ \ldots \\ S^2_{11\ldots2h_{r,1,\ldots,2}} \\ \ldots \\ \ldots \\ \ldots \\ S^2_{h_1 1\ldots11} \\ \ldots \\ \ldots \\ \ldots \\ S^2_{h_1 h_{2,h_1} \ldots h_{r,h_1,\ldots,h_{r-1}}} \end{bmatrix}$,

$$\Theta = \begin{bmatrix} \theta_{11} & \theta_{211} & \theta_{3111} & \dots & \theta_{r11\dots11} & 1 \\ \theta_{11} & \theta_{211} & \theta_{3111} & \dots & \theta_{r11\dots12} & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \theta_{11} & \theta_{211} & \theta_{3111} & \dots & \theta_{r11\dots1h_{r,1,\dots,1,h_{r-1}}} & 1 \\ \theta_{11} & \theta_{211} & \theta_{3111} & \dots & \theta_{r11\dots21} & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \theta_{11} & \theta_{211} & \theta_{3111} & \dots & \theta_{r11\dots2h_{r,1,\dots,2,h_{r-1}}} & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \theta_{1h_1} & \theta_{2h_1 1} & \theta_{3h_1 11} & \dots & \theta_{rh_1 1\dots11} & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \theta_{1h_1} & \theta_{2h_1h_{2,h_1}} & \theta_{3h_1h_{2,h_1}h_{3,h_1,h_2}} & \dots & \theta_{rh_1h_{2,h_1}\dots h_{(r-1),h_1,\dots,h_{r-2}}h_{r,h_1,\dots,h_{r-1}}} & 1 \end{bmatrix},$$

and $\gamma = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \dots \\ \dots \\ \gamma_r \\ \gamma_{(r+1)} \end{bmatrix}$, we will have

$$E(S) = \Theta\gamma. \tag{26}$$

Thus, for $i_1 = 1, \dots, h_1$, $i_j = 1, \dots, h_{j,i_1,\dots,i_{j-1}}$, $j > 1$, equalizing the variances $\lambda_{i_1 i_2 \dots i_r}$ to the correspondent estimators $S^2_{i_1 i_2 \dots i_r}$ it yields the following system of equations (in matrix notation)

$$S = \Theta\gamma. \tag{27}$$

Since by construction $\theta_{1i_1} \neq \theta_{1i_1'}$ (they are the different eigenvalues of M_1), $\theta_{2i_1i_2} \neq \theta_{2i_1i_2'}$ (they are the distinct eigenvalues of $M^2_{ii} = A_{1i_1}M_2A^\top_{1i_1}$), $\theta_{3i_1i_2i_3} \neq \theta_{3i_1i_2i_3'}$ (they are the distinct eigenvalues of $A_{2i_1i_2}A_{1i_1}M_2A^\top_{1i_1}A^\top_{2i_1i_2}$), ..., $\theta_{ri_1i_2\dots i_{(r-1)}i_r} \neq \theta_{ri_1i_2\dots i_{(r-1)}i_r'}$ (they are the distinct eigenvalues of $A_{(r-1)i_1i_2\dots i_{(r-1)}} \dots A_{1i_1}M_rA^\top_{1i_1} \dots A^\top_{(r-1)i_1i_2\dots i_{(r-1)}}$) where $i_j \neq i_j^{'}$, $j = 1, \dots, r$, it is easily seen that the matrix Θ is of full rank; that is $r(\Theta) = r + 1$.

According with Theorem 2.6*d* (Rencher and Schaalje [11]), with $\sum$ denoting $\sum_{i_1}^{h_1}\sum_{i_2}^{h_{2,i_1}} \dots \sum_{i_r}^{h_{r,i_1,\dots,i_{r-1}}}$, the matrix

$$\Theta^\top\Theta = \begin{bmatrix} \sum\theta_{1i_1}^2 & \sum\theta_{1i_1}\theta_{2i_1i_2} & \sum\theta_{1i_1}\theta_{3i_1i_2i_3} & \cdots & \sum\theta_{1i_1}\theta_{ri_1\ldots ir} & \sum\theta_{1i_1} \\ \sum\theta_{1i_1}\theta_{2i_1i_2} & \sum\theta_{2i_1i_2}^2 & \theta_{2i_1i_2}\theta_{3i_1i_2i_3} & \cdots & \sum\theta_{2i_1i_2}\theta_{ri_1\ldots ir} & \sum\theta_{2i_1i_2} \\ \sum\theta_{1i_1}\theta_{3i_1i_2i_3} & \sum\theta_{2i_1i_2}\theta_{3i_1i_2i_3} & \sum\theta_{3i_1i_2i_3}^2 & \cdots & \sum\theta_{3i_1i_2i_3}\theta_{ri_1\ldots ir} & \sum\theta_{3i_1i_2i_3} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \sum\theta_{1i_1}\theta_{ri_1\ldots ir} & \sum\theta_{2i_1i_2}\theta_{ri_1\ldots ir} & \sum\theta_{3i_1i_2i_3}\theta_{ri_1\ldots ir} & \cdots & \sum\theta_{ri_1\ldots ir}^2 & \sum\theta_{ri_1\ldots ir} \\ \sum\theta_{1i_1} & \sum\theta_{2i_1i_2} & \sum\theta_{3i_1i_2i_3} & \cdots & \sum\theta_{ri_1\ldots ir} & \sum \end{bmatrix}$$

is positive-definite, and according with Corollary 1 of (Rencher and Schaalje [11], p. 27) $\Theta^\top\Theta$ is non-singular; that is, it is invertible. We denote its inverse by $(\Theta^\top\Theta)^{-1}$.

Now, premultiplying the system (27) in both side by $\Theta^\top$ the resulting system of equations will be

$$\Theta^\top S = \Theta^\top\Theta\gamma, \tag{28}$$

whose unique solution (and therefore an estimator of γ) will be the *Sub-D* estimator

$$\hat{\gamma} = (\Theta^\top\Theta)^{-1}\Theta^\top S. \tag{29}$$

Proposition 6 $\hat{\gamma} = (\Theta^\top\Theta)^{-1}\Theta^\top S$ *is an unbiased estimator of*

$$\gamma = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \cdots \\ \cdots \\ \gamma_r \\ \gamma_{(r+1)} \end{bmatrix}, \textit{ where } \begin{bmatrix} \hat{\gamma}_1 \\ \hat{\gamma}_2 \\ \hat{\gamma}_3 \\ \cdots \\ \cdots \\ \hat{\gamma}_r \\ \hat{\gamma}_{(r+1)} \end{bmatrix}.$$

Indeed, $E(\hat{\gamma}) = E\left((\Theta^\top\Theta)^{-1}\Theta^\top S\right) = (\Theta^\top\Theta)^{-1}\Theta^\top E(S) = (\Theta^\top\Theta)^{-1}\Theta^\top\Theta\gamma = \gamma$.

4 Numerical Results

In this section we carry numerical tests to the sub-diagonalization method for the case $r = 2$, that is for a model with 3 variances components. For this case we pick the particular model $z \sim \mathcal{N}_{21}\,(X\beta, \gamma_1 N_1 + \gamma_2 N_2 + \gamma_3 I_{21})$, where $N_j = X_j X_j^\top$, $j = 1, 2$, with design matrices

$$X_1 = \begin{bmatrix} 1_5 & 0_5 & 0_5 \\ 0_9 & 1_9 & 0_9 \\ 0_7 & 0_7 & 1_7 \end{bmatrix}, X_2 = \begin{bmatrix} 1_2 & 0_2 & 0_2 \\ 0_4 & 1_4 & 0_4 \\ 0_8 & 0_8 & 1_8 \\ 1_4 & 0_4 & 0_4 \\ 0_3 & 1_3 & 0_3 \end{bmatrix},$$

and $X = 1_{21}$. 1_k and 0_k denote, respectively, $k \times 1$ vectors of 1 and 0.

Let B_o be a matrix whose columns are the eigenvectors associated to the null eigenvalues of $\frac{1}{21}J_{21}$. Then $B_o B_o^\top = I_{21} - \frac{1}{21}J_{21}$ and $B_o^\top B_o = I_{20}$, and so the new model will be

$$y = B_o^\top z \sim \mathcal{N}_{20}\left(\mathbf{0}_{20},\ \gamma_1 M_1 + \gamma_2 M_2 + \gamma_3 I_{20}\right),$$

where $M_d = B_o^\top N_d B_o$.

Since $r(N_1) = 3$ we have that (see Schott [12, Theorem 2.10c]) $r(M_1) = r(B_0^\top N_1 B_0) = 3$. The eigenvalues of M_1 are $\theta_{11} = 7.979829$, $\theta_{12} = 5.639219$, and $\theta_{13} = 0$ (θ_{13} with multiplicity (root) equal to 18). Thus we have that $M_{11}^2 = A_{11} M_2 A_{11}^\top = 5.673759$ and $M_{22}^2 = A_{12} M_2 A_{12}^\top = 0.6246537$ will be 1×1 matrices, and $M_{33}^2 = A_{13} M_2 A_{13}^\top$ an 18×18 matrix.

We have the following: M_{11}^2 has eigenvalue $\theta_{211} = 5.673759$; M_{22}^2 has eigenvalue $\theta_{221} = 0.6246537$; M_{33}^2 has 3 eigenvalues: $\theta_{231} = 6.390202$; $\theta_{232} = 1.216148$; $\theta_{233} = 0$ (θ_{233} with multiplicity equal to 16).

Finally we found that

$$S^\top = [190.779246 \ \ 8.866357 \ \ 5.234293 \ \ 53.654627 \ \ 1.334877]$$

and $$\Theta = \begin{bmatrix} 7.979829 & 5.6737590 & 1 \\ 5.639219 & 0.6246537 & 1 \\ 0 & 6.3902016 & 1 \\ 0 & 1.2161476 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

With $\beta_k \sim \mathcal{N}_{20}(\mathbf{0}_3, \gamma_k I_3)$, $k = 1, 2$, and $e \sim \mathcal{N}_{20}(\mathbf{0}_{20}, \gamma_3 I_{20})$, and taking $\gamma_3 = 1$, the model can be rewritten as $y = B_o^\top X_1 \beta_1 + B_o^\top X_2 \beta_2 + B_o^\top e$.

We consider γ_1 and γ_2 taking values in $\{0.1, 0.25, 0.5, 0.75, 1, 2, 5, 10\}$. Thus, for each possible combination of γ_1 and γ_2, the model y is observed 1000 time, and for each observation the sub-diagonalization method is applied and the variance components estimated for each observed y. The Tables 1 and 3 present the average of the estimated values of γ_1 and γ_2, respectively. In order to compare the sub-diagonalization method performance with the REML, for the same 1000 observations of y, the REML method is applied and the results presented in both Tables 2 and 4.

Taking a look at tables, and comparing the averages estimated values from the sub-diagonalization method to the ones of the REML methods (see Tables 1, 2, 3, and 4), the reader may easily concludes that the results provided by the sub-diagonalization method are in general slightly more realistic. In other hand, the averages variability of the sub-diagonalization methods is relatively higher than those of REML method (see Tables 5, 6, 7, and 8); this is because of the correlation between the sub-models. This gap will be fixed in future works.

5 Concluding Remarks

Besides its simple and fast computational implementation once it depends only on the information retained on the eigenvalues of the design matrices and the quadratic errors of the model, *Sub-D* provides centered estimates whether for balanced or unbalanced designs, which is not the case of estimators based on ANOVA methods. As seen at Sect. 4, *Sub-D* provides a slightly more realistic estimates than the REML estimator, but with more variability (when the model is balanced they have a comparable variability). However, since in any computational program (source code) when we are interested in share the code, create package or use it repeatedly, we might consider its efficiency and, for this matter, the code run-time constitutes a good start point. Doing so, to compute the estimates and the corresponding variance for each pair γ_1 and γ_2 taking values in $\{0.25,\ 0.5,\ 1,\ 2,\ 5,\ 10\}$, for 1000 observations of the model, we found that the *Sub-D* run-time is about 0.25 s while the REML estimator run-time is about 35.53 s, which means that the code for *Sub-D* is more than 70 times faster than the one for REML. The code was run using R software.

It seems that the problem of the little higher variability in *Sub-D* comparing to REML estimator is due to the correlation between the sub-models (for the case of models with three variance components, for example) $y_{ij},\ i = 1, \ldots, h_1,\ j = 1, \ldots, h_{2h_i}$. From (10) we see that the variance components matrix of the model $w_2 = P_2 P_1 y$ is a blockwise matrix whose diagonal matrices are $D_1, \ldots, D_{h_1}$, where $D_i = diag(\lambda_{i1}, \ldots, \lambda_{ih_{2i}})$, corresponding to $cov(y_{ij}, y_{sk})$ for $i = s,\ j = k$, and the off diagonal matrices are the non-null matrices $\gamma_2 A_{2ij} W_{is}^2 A_{2sk}$, corresponding to $cov(y_{ij}, y_{sk})$ for $i \neq s$. This problem will be handled in future work. Confidence region will be obtained and tests of Hypothesis for the variance components will be derived in future works.

Acknowledgements This work was partially supported by the Fundaç ao para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) through PEst-OE/MAT/UI0297/2011 (CMA), and by the Fundaç ao Calouste Gulbenkian Through a PhD Grants. It was also partially supported by Universidade de Cabo Verde.

Appendix

Table 1 Sub-diagonalization method - average estimate for γ_1

γ_1/γ_2	0.1	0.25	0.5	0.75	1	2	5	10
0.1	0.0917	0.0984	0.0828	0.1162	0.0833	0.1052	0.1102	0.1053
0.25	0.2716	0.2954	0.2698	0.2538	0.3041	0.2882	0.1993	0.3322
0.5	0.5010	0.5127	0.4929	0.5088	0.5297	0.4613	0.5314	0.5569
0.75	0.7279	0.7683	0.7685	0.7755	0.7693	0.7504	0.6982	0.8215
1	1.0305	1.0293	1.0143	0.9971	1.0309	1.0013	1.0046	1.0809
2	1.9844	2.0004	2.0032	1.9702	2.0827	2.0893	2.0643	2.2640
5	5.1864	5.0386	4.9128	5.0722	5.2111	5.0170	4.8472	5.1269
10	9.6167	10.1588	10.2468	10.1263	9.6940	9.9046	10.0246	9.8474

Table 2 REML method - average estimate for γ_1

γ_1/γ_2	0.1	0.25	0.5	0.75	1	2	5	10
0.1	0.1431	0.1683	0.1779	0.1884	0.1975	0.2154	0.2189	0.2156
0.25	0.2872	0.3157	0.3379	0.3286	0.3416	0.3316	0.3740	0.3480
0.5	0.5191	0.5546	0.5244	0.5637	0.6110	0.5897	0.6469	0.6281
0.75	0.7271	0.7620	0.7587	0.7908	0.8159	0.8245	0.8373	0.8241
1	1.0300	1.0026	1.0245	1.0172	1.0138	1.0726	1.0352	1.0515
2	1.9343	1.9884	1.9565	2.0178	2.1510	2.1482	2.0774	2.2323
5	5.1267	4.9747	4.7743	5.0955	5.1395	4.9907	4.8066	4.8150
10	9.5043	10.0881	10.1912	10.0269	9.4706	9.7784	9.9445	9.6754

Table 3 Sub-diagonalization method - average estimate for γ_2

γ_1/γ_2	0.1	0.25	0.5	0.75	1	2	5	10
0.1	0.1026	0.2643	0.5147	0.7147	1.0286	1.9595	4.9390	9.9718
0.25	0.1051	0.2589	0.4918	0.7827	1.0172	2.0427	4.8713	9.7690
0.5	0.0903	0.2323	0.5043	0.7865	1.0117	1.9496	4.8136	9.8913
0.75	0.0855	0.3068	0.5144	0.7676	1.1207	2.0762	4.7910	9.7847
1	0.0581	0.2746	0.5052	0.7969	1.0035	2.1009	5.0871	10.2702
2	0.0902	0.2966	0.6198	0.7870	0.9909	1.9605	5.217	9.7318
5	0.1759	0.3403	0.5565	0.7276	1.0007	2.036	4.8617	9.7160
10	0.1614	0.2562	0.5649	0.7481	0.9934	2.1402	5.1631	10.1369

Table 4 REML method - average estimate for γ_2

γ_1/γ_2	0.1	0.25	0.5	0.75	1	2	5	10
0.1	0.1539	0.2701	0.5143	0.7095	0.9992	1.9007	4.9153	9.9579
0.25	0.1630	0.2965	0.5165	0.7840	1.0271	2.0990	4.7929	9.5820
0.5	0.1867	0.3061	0.5490	0.7964	1.0400	1.9358	4.7022	9.6481
0.75	0.1976	0.3501	0.5480	0.8079	1.0678	2.1196	4.6759	9.7793
1	0.2008	0.3289	0.5488	0.8134	1.0282	2.0205	5.0126	10.3663
2	0.2186	0.3379	0.5703	0.8469	1.0249	1.9900	5.4291	9.5900
5	0.2198	0.3799	0.5603	0.7773	1.0027	2.0142	4.7727	9.6886
10	0.2284	0.3551	0.5906	0.7792	1.1087	2.0735	4.9235	10.0843

Table 5 Sub-diagonalization method - variation of the estimated γ_1

γ_1/γ_2	0.1	0.25	0.5	0.75	1	2	5	10
0.1	0.1264	0.2253	0.4626	0.8296	1.2005	4.3832	19.6631	83.6993
0.25	0.2637	0.3814	0.6248	1.0775	1.5931	4.7676	20.1332	72.7948
0.5	0.5737	0.7863	1.1830	1.7217	2.3142	4.7103	22.8545	78.2997
0.75	0.9224	1.2110	1.5779	2.0896	3.3078	7.4140	20.7793	77.7225
1	77.7225	1.8328	2.4022	2.9417	3.8380	7.6562	27.1356	101.9337
2	4.8401	5.6613	6.9492	6.8652	8.4356	13.2666	37.4524	107.8436
5	30.5767	31.3904	34.2362	36.0102	36.5273	43.1085	72.8085	157.0055
10	111.1505	117.9503	114.2234	120.8808	124.3445	138.0213	192.7288	288.9592

Table 6 Sub-diagonalization method - variation of the estimated γ_2

γ_1/γ_2	0.1	0.25	0.5	0.75	1	2	5	10
0.1	0.1532	0.2972	0.6524	1.1154	2.0379	6.4364	33.8728	138.7916
0.25	0.2379	0.4537	0.7838	1.3616	2.0686	7.7435	32.4170	112.701
0.5	0.5232	0.7162	1.1545	1.7515	2.7932	6.1609	31.2810	117.2392
0.75	0.7703	1.0841	1.4314	1.9380	3.3226	7.6266	35.7370	139.0834
1	1.1496	1.4291	1.8988	2.6630	3.6221	8.7960	39.6377	159.5489
2	3.8362	4.5207	4.6976	5.5365	6.9396	11.6933	47.5170	140.7587
5	21.0152	22.2408	24.2194	24.0984	29.4643	34.2175	65.9059	176.7041
10	81.3183	82.3035	89.9235	85.9040	85.1849	93.4313	153.1855	265.6179

Table 7 REML method - variation of the estimated γ_1

γ_1/γ_2	0.1	0.25	0.5	0.75	1	2	5	10
0.1	0.07807	0.0880	0.1324	0.1579	0.1801	0.2524	0.2679	0.2052
0.25	0.20365	0.2229	0.2729	0.2676	0.3350	0.3365	0.4485	0.3235
0.5	0.4747	0.6030	0.5822	0.7576	0.8165	0.7607	0.8321	0.9255
0.75	0.8896	0.9458	1.0035	1.1702	1.2667	1.2627	1.2131	1.4153
1	1.4500	1.4368	1.7622	1.7407	1.8813	1.9144	1.8597	1.9659
2	4.6049	4.9522	4.8249	5.6586	6.0638	6.3735	6.0565	7.8698
5	28.4367	29.6686	29.0413	32.1312	29.1439	28.4656	28.1731	29.3058
10	106.6903	108.3732	106.734	105.7222	106.4887	101.2775	111.1112	104.9005

Table 8 REML method - variation of the estimated γ_2

γ_1/γ_2	0.1	0.25	0.5	0.75	1	2	5	10
0.1	0.0833	0.1798	0.5192	0.7836	1.4306	4.8877	27.2749	100.2321
0.25	0.0914	0.2295	0.5842	0.9688	1.5517	6.1586	25.9314	92.9996
0.5	0.1260	0.2744	0.5607	1.2902	1.8142	4.4948	23.3488	94.9688
0.75	0.1534	0.3081	0.6120	1.2712	1.6747	5.9940	26.5791	110.6777
1	0.1732	0.3270	0.6852	1.2331	1.8197	5.2857	29.3231	126.1761
2	0.2289	0.3608	0.7416	1.5226	1.7834	5.7763	31.7812	101.8187
5	0.2399	0.4452	0.8946	1.2738	1.6384	5.2879	26.9691	97.7408
10	0.2280	0.4149	0.7789	1.2234	2.1941	5.7251	31.2616	98.4346

References

1. Anderson, R.L.: Use of variance component analysis in the interpretation of biological experiments. Bull. Int. Stat. Inst. **37**, 71–90 (1960)
2. Anderson, R.L.: Designs and estimators for variance components. Statistical Design and Linear Model, pp. 1–30. North-Holland, Amsterdam (1975)
3. Anderson, R.L.: Recent developments in designs and estimators for variance components. Statistics and Related Topics, pp. 3–22. North-Holland, Amsterdam (1981)
4. Anderson, R.L., Crump, P.P.: Comparisons of designs and estimation procedures for estimating parameters in a two-stages nested process. Tecnometrics **9**, 499–516 (1967)
5. Casella, G., Berger, R.L.: Statistical Inference. Duxbury Pacific Grove (2002)
6. Hartley, H.O., Rao, J.K.: Maximum likelihood estimation for the mixed analysis of variance model. Biometrika **54**, 93–108 (1967)
7. Khuri, A.I.: Design for variance components estimation: past and present. Int. Stat. Rev. **68**, 311–322 (2000)
8. Khuri, A.I., Sahai, H.: Variance components analysis: a selective literature survey. Int. Stat. Rev. **53**, 279–300 (1985)
9. Miller, J.J.: Asymptotic properties and computation of maximum likelihood estimates in the mixed model of the analysis of variance. Technical report 12, Department of Statistics, Stanford University, Stanford, California (1973)
10. Moser, B.: Linear Models: A Mean Model Approach. Elsevier, New York (1996)
11. Rencher, A.C., Schaalje, G.B.: Linear Models in Statisitcs. Wiley, New York (2008)

12. Schott, J.R.: Matrix Analysis for Statistics. Wiley, New York (1997)
13. Searle, S.: Topics in variance component estimation. Biometrics **27**, 1–76 (1971)
14. Searle, S.: An overview of variance component estimation. Metrika **42**, 215–230 (1995)
15. Searle, S., Casella, G., McCulloch, C.: Variance Components. Wiley, New York (2009)