Estimability of Variance Components when all Model Matrices Commute

R. A. Bailey[a,*], Sandra S. Ferreira[b], Dário Ferreira[b], Célia Nunes[b]

[a] *School of Mathematics and Statistics, University of St Andrews, North Haugh, St Andrews, Fife KY16 9SS, UK and School of Mathematical Sciences, Queen Mary, University of London, Mile End Road, London EC1 4NS, UK (emerita)*

[b] *Department of Mathematics and Center of Mathematics and Applications, University of Beira Interior, Avenida Marquês D'Ávila e Bolama, 6200 Covilhã, Portugal*

Abstract

This paper deals with estimability of variance components in mixed models when all model matrices commute. In this situation, it is well known that the best linear unbiased estimators of fixed effects are the ordinary least squares estimators. If, in addition, the family of possible variance-covariance matrices forms an orthogonal block structure, then there are the same number of variance components as strata, and the variance components are all estimable if and only if there are non-zero residual degrees of freedom in each stratum.

We investigate the case where the family of possible variance-covariance matrices, while still commutative, no longer forms an orthogonal block structure. Now the variance components may or may not all be estimable, but there is no clear link with residual degrees of freedom. Whether or not they are all estimable, there may or may not be uniformly best unbiased quadratic estimators of those that are estimable. Examples are given to demonstrate all four possibilities.

Keywords: analysis of variance, commutativity, mixed model, orthogonal block structure, segregation.
2008 MSC: 62J10, 62K99

*Corresponding author
Email address: rab24@st-andrews.ac.uk (R. A. Bailey)

1. Some assumptions about the linear model

Let $\mathbf{Y}$ be a column vector of N random variables $Y_1, \ldots, Y_N$. Write $\mathbb{E}(\mathbf{Y})$ for the expectation of the vector $\mathbf{Y}$, and $\mathbf{V}$ for its variance-covariance matrix. In this section we present some of the assumptions that are commonly made about $\mathbb{E}(\mathbf{Y})$ and $\mathbf{V}$ in order to have a linear model with good properties.

Assumption 1. [Linear expectation] There is a known integer n, a known $N \times n$ real matrix $\mathbf{X}$ and an unknown column vector $\boldsymbol{\tau}$ of length n such that $\mathbb{E}(\mathbf{Y}) = \mathbf{X}\boldsymbol{\tau}$.

Under Assumption 1, let $\mathbf{T}$ be the $N \times N$ matrix of orthogonal projection onto the column-space of $\mathbf{X}$. Then $\mathbf{T} = \mathbf{X}(\mathbf{X}^\top\mathbf{X})^+\mathbf{X}^\top$, where $^+$ denotes the Moore–Penrose generalized inverse: see texts such as [11, 23]. Also, let $\mathbf{G}_N$ be the matrix of orthogonal projection onto the space W spanned by the all-1 vector $\mathbf{1}$, so that $\mathbf{G}_N = N^{-1}\mathbf{J}_N$, where $\mathbf{J}_N$ is the $N \times N$ matrix whose entries are all equal to 1. It is often the case that $\mathbf{1}$ is in the column-space of $\mathbf{X}$. This happens if and only if $\mathbf{T}\mathbf{G}_N = \mathbf{G}_N\mathbf{T} = \mathbf{G}_N$: see [11, 39].

Assumption 2. [Spectral form of variance-covariance matrix] There are known orthogonal symmetric idempotent matrices $\mathbf{Q}_0, \ldots, \mathbf{Q}_m$ summing to the identity matrix $\mathbf{I}_N$ of order N, and non-negative scalars $\gamma_0, \ldots, \gamma_m$ such that

$$\mathbf{V} = \sum_{i=0}^{m} \gamma_i \mathbf{Q}_i. \tag{1}$$

This assumption says that the scalars $\gamma_0, \ldots, \gamma_m$ are the eigenvalues of $\mathbf{V}$; the corresponding eigenspaces are the column spaces of $\mathbf{Q}_0, \ldots, \mathbf{Q}_m$, and these are known. It is often the case that W is one of the eigenspaces: in that case, we label the spaces so that $\mathbf{Q}_m = \mathbf{G}_N$.

Assumption 3. [No relations among the eigenvalues] Assumption 2 is true and there are no further constraints on the values of $\gamma_0, \ldots, \gamma_m$.

A weaker form of this, given in [37], is that there are no linear constraints on $\gamma_0, \ldots, \gamma_m$. Here is an alternative weaker form.

Assumption 4. [Cone of full dimension] Assumption 2 is true and the family of possible matrices $\mathbf{V}$ forms a positive cone of dimension $m+1$ in the space spanned by $\mathbf{Q}_0, \ldots, \mathbf{Q}_m$.

There are two common ways of justifying Assumption 2. The first starts with known factors with random effects. A factor is simply a function assigning various discrete levels to each of the units 1, ..., N. For $j = 0, \ldots, w$, let $\mathbf{M}_j$ be the $N \times N$ relation matrix for the j-th factor: its (α, β)-entry is equal to 1 if this factor has the same level on units α and β; otherwise it is equal to 0: see [4, 5]. We always include the trivial factor with N different levels and label it as the 0-th factor, so that $\mathbf{M}_0 = \mathbf{I}_N$.

Assumption 5. [Mixed model] There are $w+1$ factors with random effects. The relation matrices $\mathbf{M}_0, \ldots, \mathbf{M}_w$ of these factors are known, and $\mathbf{M}_0 = \mathbf{I}_N$. There are also unknown non-negative numbers $\sigma_0^2, \ldots, \sigma_w^2$ such that

$$\mathbf{V} = \sum_{j=0}^{w} \sigma_j^2 \mathbf{M}_j. \tag{2}$$

No relationships are assumed among $\sigma_0^2, \ldots, \sigma_w^2$.

Assumption 6. [Mixed model with linear independence] Assumption 5 is true and $\mathbf{M}_0, \ldots, \mathbf{M}_w$ are linearly independent.

When Assumption 6 is true, there is a unique expression for the right-hand side of Equation (2). For an example satisfying Assumption 5 but not Assumption 6, use the five factors in a 2×2 Latin square: see Tjur [42, §7.3].

Assumption 7. [Commutativity of relation matrices] Assumption 6 is true and $\mathbf{M}_i\mathbf{M}_j = \mathbf{M}_j\mathbf{M}_i$ for $0 \leqslant i < j \leqslant w$.

Proposition 1. *If Assumption 7 is true, then $\mathbf{M}_0, \ldots, \mathbf{M}_w$ generate a commutative algebra $\mathcal{A}$ of symmetric matrices. If the dimension of $\mathcal{A}$ is $m+1$, then $m \geqslant w$ and Equation (2) can be written as (1), where $\mathbf{Q}_0, \ldots, \mathbf{Q}_m$ are the primitive idempotents of $\mathcal{A}$. Then there are known scalars b_{ij} such that $\gamma_i = \sum_{j=0}^{w} b_{ij}\sigma_j^2$ for $i = 0, \ldots, m$.*

Under Assumptions 1 and 7, Fonseca et al. [17, 18, 19] call the matrices $\mathbf{X}\mathbf{X}^\top$ and $\mathbf{M}_0, \ldots, \mathbf{M}_w$ model matrices. In view of the different roles played by expectation and variance, and to allow for treatments with unequal replication, we prefer to use the term 'model matrices' for $\mathbf{T}$ and $\mathbf{M}_0, \ldots, \mathbf{M}_w$.

The other common way of justifying Assumption 2 comes from considering the pattern of the entries in $\mathbf{V}$, which may well be justified by randomization, as in [1, 2, 27].

Assumption 8. [Patterns of covariance] There are known non-zero symmetric matrices $\mathbf{A}_0, \ldots, \mathbf{A}_w$ summing to $\mathbf{J}_N$, such that $\mathbf{A}_0 = \mathbf{I}_N$ and all entries in $\mathbf{A}_i$ are in $\{0, 1\}$ for $i = 1, \ldots, w$, and also a non-negative number σ^2, and numbers $\rho_1, \ldots, \rho_w$ in $[-1, 1]$ such that

$$\mathbf{V} = \sigma^2 \mathbf{A_0} + \sigma^2 \sum_{i=1}^{w} \rho_i \mathbf{A}_i.$$

The only further conditions assumed on $\rho_1, \ldots, \rho_w$ are that $\mathbf{V}$ is non-negative definite.

Assumption 9. [Commutativity of pattern matrices] Assumption 8 is true and $\mathbf{A}_i\mathbf{A}_j = \mathbf{A}_j\mathbf{A}_i$ for $0 \leqslant i < j \leqslant w$.

Assumption 9 gives a result similar to Proposition 1, with $\mathbf{M}_i$ replaced by $\mathbf{A}_i$. For simplicity, from now on we use the notation $\mathbf{M}_i$ in both cases. Furthermore, let $\mathbf{B}$ be the $(m+1) \times (w+1)$ matrix with entries b_{ij}.

The final assumption in this section relates the expectation part of the model to the variance-covariance part.

Assumption 10. [Commutativity between expectation and variance matrices] Assumptions 1 and 2 are true, and the matrix $\mathbf{T}$ commutes with $\mathbf{Q}_i$ for $i = 0, \ldots, m$.

Under Assumption 10, put $\mathbf{T}_i = \mathbf{T}\mathbf{Q}_i = \mathbf{Q}_i\mathbf{T}$ and $\mathbf{P}_i = \mathbf{Q}_i - \mathbf{T}_i$ for $i = 0, \ldots, m$. If the column-spaces of $\mathbf{X}$ and $\mathbf{Q}_i$ are orthogonal to each other then $\mathbf{T}_i$ is zero; otherwise, $\mathbf{T}_i$ is the matrix of orthogonal projection onto the intersection U_i of these spaces. If U_i is the whole of the column-space W_i of $\mathbf{Q}_i$ then $\mathbf{T}_i = \mathbf{Q}_i$ and $\mathbf{P}_i = \mathbf{0}$; otherwise, $\mathbf{P}_i$ is the matrix of orthogonal projection onto the space $W_i \cap U_i^{\perp}$, which is the orthogonal complement of U_i in W_i. Let d_i be the rank of $\mathbf{P}_i$, so that $d_i = 0$ if $\mathbf{P}_i = \mathbf{0}$ and $d_i = \dim(W_i) - \dim(U_i)$ otherwise.

Under Assumption 1, the ordinary least squares (OLS) estimator $\hat{\boldsymbol{\tau}}_{\mathrm{OLS}}$ of $\boldsymbol{\tau}$ is given by $\hat{\boldsymbol{\tau}}_{\mathrm{OLS}} = (\mathbf{X}^\top\mathbf{X})^+\mathbf{X}^\top\mathbf{Y} = (\mathbf{X}^\top\mathbf{X})^+\mathbf{X}^\top\mathbf{T}\mathbf{Y}$. If $\mathbf{V}$ is known, then the generalized least-squares (GLS) estimator $\hat{\boldsymbol{\tau}}_{\mathrm{GLS}}$ of $\boldsymbol{\tau}$ is given by $\hat{\boldsymbol{\tau}}_{\mathrm{GLS}} = (\mathbf{X}^\top\mathbf{V}^{-1}\mathbf{X})^+\mathbf{X}^\top\mathbf{V}^{-1}\mathbf{Y}$. The importance of Assumption 10 is shown by the following theorem, which can be found in [20, 22, 24, 26, 33, 44], for example.

Theorem 1. *Under Assumptions 1 and 2, Assumption 10 is equivalent to the condition that the OLS estimator of $\boldsymbol{\tau}$ is the same as the GLS estimator of $\boldsymbol{\tau}$ no matter what the values of $\gamma_0, \ldots, \gamma_m$ are.*

More recently, Assumption 10 has been called 'equivalent estimation'; see, for example, [25, 30, 43]. In [8], Brown says that 'ANOVA exists' if and only if Assumption 10 is satisfied.

2. Orthogonal block structure

Many authors have given conditions on $\mathbf{V}$ which ensure that Assumption 2 is true. A factor is said to be balanced if all of its levels occur on the same number of units.

Assumption 11. [Tjur block structure] Assumption 7 is true, all the factors with random effects are balanced, and $\mathbf{M}_0, \ldots, \mathbf{M}_w$ span $\mathcal{A}$.

In [42], Tjur discussed variance-covariance models satisfying Assumption 11. He showed that this implies that $m = w$. Thus the matrix $\mathbf{B}$ is square and invertible, and there is no linear relationship specified among $\gamma_0, \ldots, \gamma_m$, but the non-negativity of $\sigma_0^2, \ldots, \sigma_w^2$ implies that Assumption 3 is not satisfied if $m \geq 1$. He further showed that it is possible to label the relation matrices and the primitive idempotents in such a way that the matrix $\mathbf{B}$ is lower triangular: we adopt this convention in Examples 1–3, 5 and 6.

Nelder [27] and Bailey [4, 6] defined an orthogonal block structure (OBS) to be a collection of factors satisfying some extra conditions in addition to Tjur's, but with Assumption 7 replaced by Assumption 9.

Houtman and Speed [21] generalized the definition of OBS to mean any family of variance-covariance matrices satisfying Assumption 3. They explicitly specified that the family must consist of all positive semi-definite matrices of the form (1), and gave the following two examples, which consider only the structure of $\mathbf{V}$, to clarify this point.

Example 1. Here $N = bk$, and the units are grouped into b blocks of size k. The usual mixed model gives $\mathbf{V} = \sigma_0^2\mathbf{M}_0 + \sigma_1^2\mathbf{M}_1$, where $\mathbf{M}_0 = \mathbf{I}_N$ and $\mathbf{M}_1$ is the relation matrix for blocks. Then

$$\mathbf{V} = \gamma_0\mathbf{Q}_0 + \gamma_1\mathbf{Q}_1, \tag{3}$$

where $\mathbf{Q}_1 = k^{-1}\mathbf{M}_1$, $\mathbf{Q}_0 = \mathbf{I}_N - \mathbf{Q}_1$, $\gamma_0 = \sigma_0^2$ and $\gamma_1 = \sigma_0^2 + k\sigma_1^2$. This is an OBS, apart from the positivity constraint $\gamma_1 \geqslant \gamma_0$.

On the other hand, the randomization model of [1, 2] gives $\mathbf{V} = \sigma^2\mathbf{A}_0 + \sigma^2(\rho_1\mathbf{A}_1 + \rho_2\mathbf{A}_2)$, where $\mathbf{A}_0 = \mathbf{I}_N$, $\mathbf{A}_1 = \mathbf{M}_1 - \mathbf{I}_N$ and $\mathbf{A}_2 = \mathbf{J}_N - \mathbf{M}_1$. Then

$$\mathbf{V} = \gamma_2\mathbf{Q}_2 + \gamma_3\mathbf{Q}_3 + \gamma_4\mathbf{Q}_4, \tag{4}$$

where we start the numbering at 2 to avoid confusion with (3). Here $\mathbf{Q}_4 = \mathbf{G}_N$, $\mathbf{Q}_3 = \mathbf{Q}_1 - \mathbf{Q}_4$, $\mathbf{Q}_2 = \mathbf{Q}_0$, $\gamma_2 = \sigma^2(1-\rho_1)$, $\gamma_3 = \sigma^2[1 + (k-1)\rho_1 - k\rho_2]$ and $\gamma_4 = \sigma^2[1 + (k-1)\rho_1 + (N-k)\rho_2]$. This is a different OBS. The usual mixed model gives the constraint that $\gamma_3 = \gamma_4$, and so expression (4) should not be used to show that its variance-covariance structure is an OBS.

Example 2. Here $N = rc$, and the units are arranged in a rectangle with r rows and c columns. The approach of Nelder [27] gives a mixed model with four factors whose effects are random: $\mathbf{M}_0 = \mathbf{I}_N$, $\mathbf{M}_1$ and $\mathbf{M}_2$ are the relation matrices for rows and columns respectively, and $\mathbf{M}_3 = \mathbf{J}_N$. This defines an OBS with $\mathbf{Q}_3 = \mathbf{G}_N$, $\mathbf{Q}_2 = r^{-1}\mathbf{M}_2 - \mathbf{G}_N$, $\mathbf{Q}_1 = c^{-1}\mathbf{M}_1 - \mathbf{G}_N$ and $\mathbf{Q}_0 = \mathbf{I}_N - \mathbf{Q}_1 - \mathbf{Q}_2 - \mathbf{Q}_3$. Moreover, $\gamma_0 = \sigma_0^2$, $\gamma_1 = \sigma_0^2 + c\sigma_1^2$, $\gamma_2 = \sigma_0^2 + r\sigma_2^2$ and $\gamma_3 = \sigma_0^2 + c\sigma_1^2 + r\sigma_2^2 + N\sigma_3^2$. Nelder explicitly allows σ_i^2 to be negative so long as γ_0, …, γ_3 are all non-negative.

An alternative mixed model omits $\mathbf{M}_3$, so that $\sigma_3^2 = 0$. However, $\mathbf{M}_1\mathbf{M}_2 = \mathbf{M}_2\mathbf{M}_1 = \mathbf{M}_3$, and so the spectral decomposition of $\mathbf{V}$ is still $\mathbf{V} = \gamma_0\mathbf{Q}_0 + \gamma_1\mathbf{Q}_1 + \gamma_2\mathbf{Q}_2 + \gamma_3\mathbf{Q}_3$. Now $\gamma_3 = \gamma_2 + \gamma_1 - \gamma_0$, and this linear constraint implies that the family of possible matrices $\mathbf{V}$ does not form an OBS.

Unfortunately, some later authors, such as Bailey [3] and Caliński and Kageyama [9], defined OBS without including the condition of no linear relationship among the γ parameters. There is now some confusion about what an OBS is. Ferreira et al. [16] try to clarify the difference by calling Assumption 3 OBS and Assumption 2 generalized OBS, while Bailey and Brien [7] call them orthogonal variance structure and commutative variance structure respectively. For the remainder of this paper we use Assumption 4 as our definition of OBS, so that it includes classes like (3) with the positivity constraint $\gamma_1 \geqslant \gamma_0$.

The combination of the Nelder–Bailey type of OBS with Assumption 10 is called simply 'orthogonality' in [6]. The combination of Assumption 7 (sometimes without linear independence of $\mathbf{M}_0$, …, $\mathbf{M}_w$) with Assumption 10 is called 'commutative orthogonal block structure' (COBS) in [10, 15, 19, 29].

Estimation under Assumption 10 is straightforward and well-known. By Theorem 1, $\hat{\boldsymbol{\tau}}_{\mathrm{OLS}} = \hat{\boldsymbol{\tau}}_{\mathrm{GLS}}$. The column-space of $\mathbf{Q}_i$ is called the i-th stratum in [2, 6, 27, 28], and in the statistical software GenStat [31, 32], and d_i is called the number of residual degrees of freedom in the i-th stratum.

Consider a value of i such that $\mathbf{T}_i \neq \mathbf{0}$. The standard error of any scalar linear function of $\mathbf{T}_i\mathbf{X}\boldsymbol{\tau}$ is proportional to $\sqrt{\gamma_i}$. Thus we usually want to estimate $\gamma_0, \ldots, \gamma_m$ as well as $\boldsymbol{\tau}$.

Proposition 2. *Suppose that Assumptions 3 and 10 hold.*

(a) If $d_i \neq 0$ then $\|\mathbf{P}_i\mathbf{Y}\|^2/d_i$ is an unbiased estimator for γ_i.

(b) If $d_i \neq 0$, the distribution of $\mathbf{Y}$ is multivariate normal, $t_i = \mathrm{rank}(\mathbf{T}_i)$ and $\mathbf{T}_i\mathbf{X}\boldsymbol{\tau} = \mathbf{0}$, then $d_i\|\mathbf{T}_i\mathbf{Y}\|^2/t_i\|\mathbf{P}_i\mathbf{Y}\|^2$ has an F-distribution on t_i and d_i degrees of freedom.

(c) If $d_i = 0$ then there is no unbiased estimator for γ_i.

Proof . (a) and (b) See [36].

(c) See [37]. ■

From (b), when $d_i \neq 0$ then the quantity $d_i\|\mathbf{T}_i\mathbf{Y}\|^2/t_i\|\mathbf{P}_i\mathbf{Y}\|^2$ is used as the test statistic for testing the null hypothesis that $\mathbf{T}_i\mathbf{X}\boldsymbol{\tau} = \mathbf{0}$. If $d_i = 0$ then this hypothesis cannot be tested.

The following two very simple examples demonstrate two features of this process that are often overlooked. The first feature is discussed by Tjur in [42, §7.3].

Example 3. Consider an $n \times n$ Latin square, with expectation corresponding to the n letters. Then $N = n^2$. The approach of Nelder [27, 28] gives a mixed model with random factors like those in Example 2: $\mathbf{M}_0 = \mathbf{I}_N$, $\mathbf{M}_1$ and $\mathbf{M}_2$ correspond to rows and columns respectively, and $\mathbf{M}_3 = \mathbf{J}_N$. Then $\mathbf{Q}_3 = \mathbf{G}_N$, $\mathbf{Q}_2 = n^{-1}\mathbf{M}_2 - \mathbf{G}_N$, $\mathbf{Q}_1 = n^{-1}\mathbf{M}_1 - \mathbf{G}_N$ and $\mathbf{Q}_0 = \mathbf{I}_N - \mathbf{Q}_1 - \mathbf{Q}_2 - \mathbf{Q}_3$, while $\gamma_0 = \sigma_0^2$, $\gamma_1 = \sigma_0^2 + n\sigma_1^2$, $\gamma_2 = \sigma_0^2 + n\sigma_2^2$ and $\gamma_3 = \sigma_0^2 + n\sigma_1^2 + n\sigma_2^2 + n^2\sigma_3^2$.

Because the treatments are applied in a Latin square, $\mathbf{T}_1 = \mathbf{T}_2 = \mathbf{0}$ and $\mathbf{T}_3 = \mathbf{G}_N = \mathbf{Q}_3$. Hence $d_3 = 0$, and so it is impossible to estimate γ_3 or σ_3^2.

Put $\bar{\tau} = n^{-1}(\tau_1 + \cdots + \tau_n)$, so that $\mathbf{T}_3\mathbf{X}\boldsymbol{\tau} = \bar{\tau}\mathbf{1}$. It is impossible to give a standard error for the estimate of $\bar{\tau}$, and there is no test of the hypothesis that $\bar{\tau} = 0$. This may be why the line for the overall mean is often omitted from the analysis-of-variance table.

Example 4. Consider the two-sample t-test. Then $n = 2$. Suppose that treatment 1 is applied to the first r units and that treatment 2 is applied to the remaining s units, where $r + s = N$, $r \geqslant 2$ and $s \geqslant 2$. Let $\mathbf{Q}_0$ be the diagonal matrix whose first r diagonal entries are equal to 1, the remainder being 0, and put $\mathbf{Q}_1 = \mathbf{I}_N - \mathbf{Q}_0$.

Assume that

$$\mathbf{V} = \gamma_0\mathbf{Q}_0 + \gamma_1\mathbf{Q}_1. \tag{5}$$

(Usually we would write γ_i as σ_i^2, but we want to be consistent with (1).) If we believe that γ_0 and γ_1 are different then we have OBS with Assumption 10, and we estimate the standard error of the difference $\hat{\tau}_1 - \hat{\tau}_2$ as

$$\sqrt{\frac{\hat{\gamma}_0}{r} + \frac{\hat{\gamma}_1}{s}},$$

where $\hat{\gamma}_0 = \|\mathbf{P}_0\mathbf{Y}\|^2/(r-1)$ and $\hat{\gamma}_1 = \|\mathbf{P}_1\mathbf{Y}\|^2/(s-1)$.

On the other hand, if we believe that $\gamma_0 = \gamma_1 = \gamma$ then expression (5) is not OBS. In this case, $\|\mathbf{P}_0\mathbf{Y}\|^2/(r-1)$ and $\|\mathbf{P}_1\mathbf{Y}\|^2/(s-1)$ are both unbiased estimators of γ. They are usually combined to give the estimator

$$\frac{\|\mathbf{P}_0\mathbf{Y}\|^2 + \|\mathbf{P}_1\mathbf{Y}\|^2}{N-2} = \frac{\|(\mathbf{I}_N - \mathbf{T})\mathbf{Y}\|^2}{N-2},$$

which comes directly from writing $\mathbf{V} = \gamma\mathbf{I}_N$, which is indeed an OBS.

3. Linear relationships among the eigenvalues

From now on, we assume Assumption 10 and either Assumption 7 or Assumption 9, so that the matrix $\mathbf{B}$ is defined and its columns are linearly independent. For simplicity, we discuss only Assumption 7, but the results hold whenever the matrices $\mathbf{M}_i$ are symmetric and commute with each other. In particular, writing $\mathbf{M}_i$ as $\mathbf{A}_i$ gives Assumption 9. Finally, we assume that $m > w$, so that there is at least one linear relationship among $\gamma_0, \ldots, \gamma_m$.

Theorem 2. *Suppose that there are exactly $t+1$ values of i for which $d_i \neq 0$. Label the primitive idempotents in such a way that $d_i = 0$ if and only if $t < i \leqslant m$. For $i = 0, \ldots, t$, put $\hat{\gamma}_i = \|\mathbf{P}_i\mathbf{Y}\||^2/d_i$, so that $\hat{\gamma}_i$ is an unbiased estimator for γ_i. Let $\tilde{\mathbf{B}}$ be the $(t+1)\times(w+1)$ matrix obtained from the first $(t+1)$ rows of $\mathbf{B}$. Let $\theta = \sum_{i=0}^{t}\lambda_i\gamma_i$, where $\lambda_0, \ldots, \lambda_t$ are known scalars.*

(i) If the rows of $\tilde{\mathbf{B}}$ are linearly independent and the columns of $\tilde{\mathbf{B}}$ are linearly independent, then $t = w$ and $\tilde{\mathbf{B}}$ is invertible. Thus $\tilde{\mathbf{B}}^{-1}(\hat{\gamma}_0, \ldots, \hat{\gamma}_w)^\top$ gives an unbiased estimate of $(\sigma_0^2, \ldots, \sigma_w^2)^\top$ and so $\mathbf{B}\tilde{\mathbf{B}}^{-1}(\hat{\gamma}_0, \ldots, \hat{\gamma}_w)^\top$ gives an unbiased estimate of $(\gamma_0, \ldots, \gamma_m)^\top$.

(ii) The quantity $\sum_{i=0}^{t} \lambda_i \hat{\gamma}_i$ is an unbiased estimator for θ. If the rows of $\tilde{\mathbf{B}}$ are linearly independent, then no other linear combination of $\hat{\gamma}_0$, …, $\hat{\gamma}_t$ is an unbiased estimator for θ.

(iii) If the rows of $\tilde{\mathbf{B}}$ are not linearly independent, then there is at least one i with $0 \leqslant i \leqslant t$ such that there is a linear combination of $\hat{\gamma}_0$, …, $\hat{\gamma}_t$ other than $\hat{\gamma}_i$ which is an unbiased estimator for γ_i. This non-uniqueness of estimation then propogates to some of σ_0^2, …, σ_w^2 and hence may propogate to some of γ_{t+1}, …, γ_m.

(iv) If the columns of $\tilde{\mathbf{B}}$ are linearly independent, then each of σ_0^2, …, σ_w^2, and hence each of γ_0, …, γ_m, can be estimated unbiasedly by at least one linear combination of $\hat{\gamma}_0$, …, $\hat{\gamma}_t$.

(v) If the columns of $\tilde{\mathbf{B}}$ are not linearly independent, then there is at least one value of i such that no linear combination of $\hat{\gamma}_0$, …, $\hat{\gamma}_t$ gives an unbiased estimate of γ_i.

Proof . (i), (ii) and (iii) are immediate.

(iv) See Ferreira [12].

(v) If the columns of $\tilde{\mathbf{B}}$ are not linearly independent, then there is at least one value of j such that σ_j^2 is not a linear combination of γ_0, …, γ_t. Hence no linear combination of $\hat{\gamma}_0$, …, $\hat{\gamma}_t$ gives an unbiased estimate of σ_j^2. There is at least one value of i for which $b_{ij} \neq 0$: for any such i, there is no linear combination of $\hat{\gamma}_0$, …, $\hat{\gamma}_t$ which gives an unbiased estimate of γ_i. ■

When $\tilde{\mathbf{B}}$ is invertible, the only difference from Section 2 is that, when $w < i \leqslant m$, the estimate of γ_i may be a linear combination of two or more mean squares, and so Satterthwaite's approximation [35] must be used for performing an F-test. Since $m > w$, there may be at least one such value of i.

When the columns of $\tilde{\mathbf{B}}$ are linearly independent, then each parameter σ_i^2 has at least one unbiased estimator which is a linear combination of the residual mean squares. In [8], Brown says that $(\sigma_0^2, \ldots, \sigma_w^2)^\top$ is 'estimable' if and only if this happens. This property is called 'segregation' in [13, 14, 15, 16].

When the columns of $\tilde{\mathbf{B}}$ are not linearly independent then there is one or more values of i for which there is no unbiased estimator of γ_i. In such cases, clearly $i > t$, and so $\mathbf{T}_i = \mathbf{Q}_i \neq \mathbf{0}$. Thus there is no estimator of a standard error for any scalar linear function of $\mathbf{T}_i\mathbf{X}\boldsymbol{\tau}$, and there is no test of the hypothesis that $\mathbf{T}_i\mathbf{X}\boldsymbol{\tau} = \mathbf{0}$. If this occurs for only one value of i, and the corresponding primitive idempotent is $\mathbf{G}_N$, as in Example 3, then this may not be regarded as problematic.

When the rows of $\tilde{\mathbf{B}}$ are not linearly independent, which estimator should we use? Should we combine the different estimators in some way? In Example 4, there was a simple answer, given by replacing model (5) by an OBS, but this solution is not available in general.

The set of residual mean squares forms a minimal sufficient set of statistics for $\gamma_0, \ldots, \gamma_m$ when the columns of $\tilde{\mathbf{B}}$ are linearly independent: see [8]. However, Seely showed in [38] that it is not complete if the rows of $\tilde{\mathbf{B}}$ are linearly dependent, because no unbiased linear combination is uniformly better than any other.

One possibility is to maximize the likelihood of $(\mathbf{I}_N - \mathbf{T})\mathbf{Y}$ under the assumption of multivariate normality. However, Szatrowski [40] and Szatrowski and Miller [41] showed that there is no closed-form solution when the rows of $\tilde{\mathbf{B}}$ are linearly dependent.

When neither the rows nor the columns of $\tilde{\mathbf{B}}$ are linearly independent, we have both problems at the same time: inestimability of some variance components but multiple estimates of others.

There are four possible combinations of linear independence/dependence of the rows of $\tilde{\mathbf{B}}$ with linear independence/dependence of the columns of $\tilde{\mathbf{B}}$. We show in Section 4 that all four possiblities can occur.

4. Examples

This section contains two examples to demonstrate all of the behaviour described in Section 3. Each considers a variance model satisfying Assumption 7 with $m > w$, and compares it with the model satisfying Assumption 3 with the same idempotents. Several different models for expectation are used in each case, all satisfying Assumption 10 and all defined by the combinations of levels of one or more factors. In order to maintain the same notation for the primitive idempotents while using different models for expectation, it is not always possible for the strata for which $d_i \neq 0$ to be labelled $0, \ldots, t$.

Example 5. Suppose that $N = 96$, and that the units are partitioned into six blocks, each of which is a 4×4 array, so that the 24 rows and 24 columns are nested within blocks: see Fig. 1.

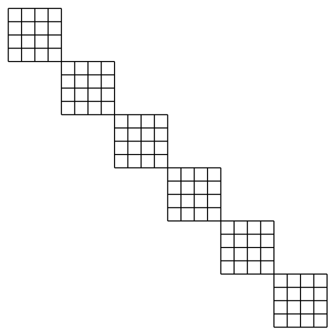

Figure 1: Structure of the 96 units in Example 5

One practical instance of this occurs in consumer testing. An experiment uses 24 volunteer consumers during 24 weeks. The weeks are divided into six groups of four weeks each, so that each group is approximately one month. The consumers are also partitioned into six groups of four. For $i = 1, \ldots, 6$, the consumers in group i participate during all the weeks in month i, and in no others. Each volunteer is given a packet of a type of coffee during each week in which he or she participates. He or she uses that coffee all week, and gives it a score at the end of the week. Here the months are the blocks, the weeks are the rows and the consumers are the columns.

A second instance of this structure is an experiment on feeds for lactating cows. Six different pens are used, one in each four-week 'month'. Twenty-four cows are used, four per pen. Each cow is given one type of feed throughout each week, and her milk production for that week is recorded. Of course, if cows in the same pen get different feeds in the same week then they must be fed individually. Now the pens, weeks and cows are the blocks, rows and columns respectively.

In most statistical software, this structure is created by first declaring factors blocks, row and columns with six, four and four levels respectively. Then one unit is created for each combination of levels. Finally, the experimental structure is declared by a formula such as

$$\text{blocks}/(\text{rows} * \text{columns}),$$

which is used in GenStat [31, 32] and R [34].

Put $\mathbf{M}_0 = \mathbf{I}_{96}$. Let $\mathbf{M}_1$, $\mathbf{M}_2$ and $\mathbf{M}_3$ be the relation matrices corresponding to rows, columns and blocks respectively. Then $\mathbf{M}_1\mathbf{M}_2 = \mathbf{M}_2\mathbf{M}_1 = \mathbf{M}_3$. The primitive idempotents of the algebra generated by $\mathbf{M}_0$, $\mathbf{M}_1$, $\mathbf{M}_2$ and $\mathbf{M}_3$ are $\mathbf{Q}_0$, $\mathbf{Q}_1$, $\mathbf{Q}_2$ and $\mathbf{Q}_3$, where $\mathbf{Q}_3 = 16^{-1}\mathbf{M}_3$, $\mathbf{Q}_2 = 4^{-1}\mathbf{M}_2 - \mathbf{Q}_3$, $\mathbf{Q}_1 = 4^{-1}\mathbf{M}_1 - \mathbf{Q}_3$ and $\mathbf{Q}_0 = \mathbf{M}_0 - \mathbf{Q}_1 - \mathbf{Q}_2 - \mathbf{Q}_3$. The Appendix shows the R code which generates these matrices, for one systematic ordering of the units.

First we consider the orthogonal block structure defined by

$$\begin{aligned}\mathbf{V} &= \sigma_0^2\mathbf{M}_0 + \sigma_1^2\mathbf{M}_1 + \sigma_2^2\mathbf{M}_2 + \sigma_3^2\mathbf{M}_3 \\ &= \gamma_0\mathbf{Q}_0 + \gamma_1\mathbf{Q}_1 + \gamma_2\mathbf{Q}_2 + \gamma_3\mathbf{Q}_3,\end{aligned}$$

where $\gamma_0 = \sigma_0^2$, $\gamma_1 = \sigma_0^2 + 4\sigma_1^2$, $\gamma_2 = \sigma_0^2 + 4\sigma_2^2$ and $\gamma_3 = \sigma_0^2 + 4\sigma_1^2 + 4\sigma_2^2 + 16\sigma_3^2$. There is no assumed linear relationship among $\gamma_0, \ldots, \gamma_3$, nor among $\sigma_0^2, \ldots, \sigma_3^2$.

We consider four different models for expectation.

(a) Treatment factors F and G each have four levels. Within each block, levels of F are randomly applied to the four rows and levels of G are randomly applied to the four columns. Each combination of levels of F and G gives an entry in $\boldsymbol{\tau}$, so that $n = 16$. Then $\mathbf{T}_0$ corresponds to the interaction $F.G$, $\mathbf{T}_1$ to the main effect of F, $\mathbf{T}_2$ to the main effect of G, and $\mathbf{T}_3$ to the overall mean.

Part (a) of Table 1 gives the skeleton analysis of variance, including the overall mean. It shows that $d_0 = 45$, $d_1 = d_2 = 15$ and $d_3 = 5$. There is one residual mean square for each γ-parameter.

(b) Treatment factor A has three levels, randomly applied to two whole blocks each. Each level of A gives an entry in $\boldsymbol{\tau}$, so that $n = 3$. Now $\mathbf{T} = \mathbf{T}_3$, which includes both the main effect of A and the overall mean. As part (b) of Table 1 shows, $d_0 = 54$, $d_1 = d_2 = 18$, $d_3 = 3$, and there is still one residual mean square for each γ-parameter.

(c) Treatment factor H is like treatment factor A, except that it has six levels. As part (c) of Table 1 shows, $t = 2$ and there is no estimator for γ_3.

stratum	df	source	df	source	df	source	df	source	df
$\mathbf{Q}_3$	6	mean	1	mean	1	mean	1	mean	1
				A	2	H	5	H	5
		residual	5	residual	3				
$\mathbf{Q}_2$	18	G	3					G	3
								$G.H$	15
		residual	15	residual	18	residual	18		
$\mathbf{Q}_1$	18	F	3						
		residual	15	residual	18	residual	18	residual	18
$\mathbf{Q}_0$	54	$F.G$	9						
		residual	45	residual	54	residual	54	residual	54
Expectation		model (a)		model (b)		model (c)		model (d)	

Table 1: Skeleton analysis of variance tables in Example 5

(d) The treatment factors are G, as in (a), and H, as in (c). Expectation effects correspond to the overall mean, the main effects of G and H, and their interaction $G.H$. Part (d) of Table 1 shows that there is no estimator for γ_3 or γ_2.

Now we suppose that $\sigma_3^2 = 0$. The primitive idempotents of the algebra generated by $\mathbf{M}_0$, $\mathbf{M}_1$ and $\mathbf{M}_2$ are still $\mathbf{Q}_0$, …, $\mathbf{Q}_3$. Now

$$\begin{aligned} \mathbf{V} &= \sigma_0^2\mathbf{M}_0 + \sigma_1^2\mathbf{M}_1 + \sigma_2^2\mathbf{M}_2 \\ &= \gamma_0\mathbf{Q}_0 + \gamma_1\mathbf{Q}_1 + \gamma_2\mathbf{Q}_2 + \gamma_3\mathbf{Q}_3, \end{aligned}$$

where $\gamma_0 = \sigma_0^2$, $\gamma_1 = \sigma_0^2 + 4\sigma_1^2$, $\gamma_2 = \sigma_0^2 + 4\sigma_2^2$ and $\gamma_3 = \sigma_0^2 + 4\sigma_1^2 + 4\sigma_2^2 = \gamma_1 + \gamma_2 - \gamma_0$.

If the expectation follows model (a), then part (a) of Table 1 shows that there are four independent residual mean squares whose expectations are γ_0, …, γ_3. There is therefore the problem of deciding, for example, whether to estimate γ_3 by $\hat{\gamma}_3$ or $\hat{\gamma}_1 + \hat{\gamma}_2 - \hat{\gamma}_0$ or some weighted average of these. Model (b) has the same problem.

If the expectation follows model (c) then part (c) of Table 1 shows that $t = w = 2$. Moreover,

$$\tilde{\mathbf{B}} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 4 & 0 \\ 1 & 0 & 4 \end{bmatrix},$$

which is invertible. The unique estimator of γ_3 is $\hat{\gamma}_1 + \hat{\gamma}_2 - \hat{\gamma}_0$. This can be used to estimate standard errors of differences between levels of H, and for an approximate F-test for H, even though $d_3 = 0$.

If the expectation follows model (d) then part (d) of Table 1 shows that $t = 1$. The parameters γ_0 and γ_1 can be estimated. Hence σ_0^2 and σ_1^2 can be estimated, but σ_2^2, γ_2 and γ_3 cannot. So there are no F-tests for any treatment effects, and no estimators of standard errors of differences.

Example 6. In a modified version of the cow-feeding experiment, all six pens are used at the same time for a single 'month' of four weeks. Thus rows are now crossed with blocks, and the formula for the experimental structure is

$$\text{rows} * (\text{blocks}/\text{columns}) :$$

see Figure 2.

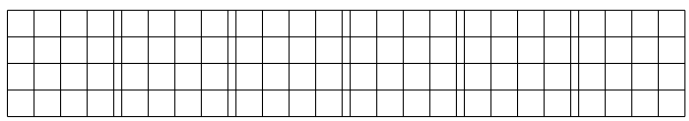

Figure 2: Structure of the 96 units in Example 6

Put $\mathbf{M}_0 = \mathbf{I}_{96}$. Let $\mathbf{M}_1$ be the relation matrix for row-block intersections. Let $\mathbf{M}_2$, $\mathbf{M}_3$ and $\mathbf{M}_4$ be the relation matrices for columns, blocks and whole rows, respectively, and let $\mathbf{M}_5 = \mathbf{J}_{96}$. The primitive idempotents of the algebra generated by $\mathbf{M}_0, \ldots, \mathbf{M}_5$ are $\mathbf{Q}_0, \ldots, \mathbf{Q}_5$, where $\mathbf{Q}_5 = 96^{-1}\mathbf{M}_5 = \mathbf{G}_{96}$, $\mathbf{Q}_4 = 24^{-1}\mathbf{M}_4 - \mathbf{Q}_5$, $\mathbf{Q}_3 = 16^{-1}\mathbf{M}_3 - \mathbf{Q}_5$, $\mathbf{Q}_2 = 4^{-1}\mathbf{M}_2 - 16^{-1}\mathbf{M}_3$, $\mathbf{Q}_1 = 4^{-1}\mathbf{M}_1 - \mathbf{Q}_3 - \mathbf{Q}_4 - \mathbf{Q}_5$, and $\mathbf{Q}_0 = \mathbf{M}_0 - \mathbf{Q}_1 - \mathbf{Q}_2 - \mathbf{Q}_3 - \mathbf{Q}_4 - \mathbf{Q}_5$.

The corresponding orthogonal block structure has

$$\begin{aligned}
\mathbf{V} &= \sigma_0^2\mathbf{M}_0 + \sigma_1^2\mathbf{M}_1 + \sigma_2^2\mathbf{M}_2 + \sigma_3^2\mathbf{M}_3 + \sigma_4^2\mathbf{M}_4 + \sigma_5^2\mathbf{M}_5 \\
&= \gamma_0\mathbf{Q}_0 + \gamma_1\mathbf{Q}_1 + \gamma_2\mathbf{Q}_2 + \gamma_3\mathbf{Q}_3 + \gamma_4\mathbf{Q}_4 + \gamma_5\mathbf{Q}_5,
\end{aligned}$$

where $\gamma_0 = \sigma_0^2$, $\gamma_1 = \sigma_0^2 + 4\sigma_1^2$, $\gamma_2 = \sigma_0^2 + 4\sigma_2^2$, $\gamma_3 = \sigma_0^2 + 4\sigma_1^2 + 4\sigma_2^2 + 16\sigma_3^2$, $\gamma_4 = \sigma_0^2 + 4\sigma_1^2 + 24\sigma_4^2$ and $\gamma_5 = \sigma_0^2 + 4\sigma_1^2 + 4\sigma_2^2 + 16\sigma_3^2 + 24\sigma_4^2 + 96\sigma_5^2$.

This time, we consider three different models for expectation.

(a) Treatment factor A has three levels, each applied to two whole blocks, so that $n = 3$. Part (a) of Table 2 gives the skeleton analysis of variance. It shows that there is one residual mean square for each of $\gamma_0, \ldots, \gamma_4$. As in

stratum	df	source	df	source	df	source	df
$\mathbf{Q}_5$	1	mean	1	mean	1	mean	1
$\mathbf{Q}_4$	3			F	3		
		residual	3			residual	3
$\mathbf{Q}_3$	5	A	2			H	5
		residual	3	residual	5		
$\mathbf{Q}_2$	18			G	3		
		residual	18	residual	15	residual	18
$\mathbf{Q}_1$	15	residual	15	residual	15	residual	18
$\mathbf{Q}_0$	54			$F.G$	9		
		residual	54	residual	45	residual	54
Expectation		model (a)		model (b)		model (c)	

Table 2: Skeleton analysis of variance tables in Example 6

Example 3, there is no estimator for γ_5 and no test of the hypothesis that $\bar{\tau} = 0$.

(b) Treatment factors F and G each have four levels. Each level of F is applied to one whole row. Levels of G are applied to the four columns within each block. There is one entry in $\boldsymbol{\tau}$ for each combination of levels of F and G, so that $n = 16$. Part (b) of Table 2 shows that $d_4 = d_5 = 0$, so that there is no estimate of γ_4 or γ_5, no test for $\bar{\tau} = 0$ and no test for the main effect of F.

(c) Treatment factor H has six levels, each of which is applied to one whole block. Now part (c) of Table 2 shows that there is no estimator for γ_3 or γ_5.

A mixed model for this structure might have $\sigma_3^2 = \sigma_5^2 = 0$. (The labelling of the matrices $\mathbf{M}_0, \ldots, \mathbf{M}_5$ is chosen so that $\mathbf{M}_0, \ldots, \mathbf{M}_3$ denote the same matrices in Examples 5 and 6.) However, $\mathbf{M}_1\mathbf{M}_2 = \mathbf{M}_2\mathbf{M}_1 = \mathbf{M}_3$ and $\mathbf{M}_2\mathbf{M}_4 = \mathbf{M}_4\mathbf{M}_2 = \mathbf{M}_5$, so the algebra generated by $\mathbf{M}_0$, $\mathbf{M}_1$, $\mathbf{M}_2$ and $\mathbf{M}_4$ still has primitive idempotents $\mathbf{Q}_0, \ldots, \mathbf{Q}_5$. Now

$$\begin{aligned} \mathbf{V} &= \sigma_0^2\mathbf{M}_0 + \sigma_1^2\mathbf{M}_1 + \sigma_2^2\mathbf{M}_2 + \sigma_4^2\mathbf{M}_4 \\ &= \gamma_0\mathbf{Q}_0 + \gamma_1\mathbf{Q}_1 + \gamma_2\mathbf{Q}_2 + \gamma_3\mathbf{Q}_3 + \gamma_4\mathbf{Q}_4 + \gamma_5\mathbf{Q}_5, \end{aligned}$$

where $\gamma_0 = \sigma_0^2$, $\gamma_1 = \sigma_0^2 + 4\sigma_1^2$, $\gamma_2 = \sigma_0^2 + 4\sigma_2^2$, $\gamma_3 = \sigma_0^2 + 4\sigma_1^2 + 4\sigma_2^2$, $\gamma_4 = \sigma_0^2 + 4\sigma_1^2 + 24\sigma_4^2$ and $\gamma_5 = \sigma_0^2 + 4\sigma_1^2 + 4\sigma_2^2 + 24\sigma_4^2$.

If the expectation follows model (a) then part (a) of Table 2 shows that there are five independent residual mean squares whose expectations are $\gamma_0, \ldots, \gamma_4$. However, $\gamma_3 = \gamma_1 + \gamma_2 - \gamma_0$, so each of $\gamma_0, \ldots, \gamma_3$ can be unbiasedly estimated by several linear combinations of these mean squares. Since $\gamma_5 = \gamma_3 + \gamma_4 - \gamma_1$, it is also possible to estimate γ_5, and so a standard error can be given for $\bar{\tau}$.

Under model (b), part (b) of Table 2 shows that $t = w = 3$, with $d_i \neq 0$ when $i = 0$, 1, 2 and 3. Thus

$$\tilde{\mathbf{B}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 4 & 0 & 0 \\ 1 & 0 & 4 & 0 \\ 1 & 4 & 4 & 0 \end{bmatrix}.$$

Neither the rows nor the columns of $\tilde{\mathbf{B}}$ are linearly independent. There is no estimator for σ_4^2, γ_4 or γ_5, and hence no test for the main effect of F. However, there are multiple estimators for γ_0, γ_1, γ_2, γ_3, σ_0^2, σ_1^2 and σ_2^2.

On the other hand, if the expectation follows model (c) then part (c) of Table 2 shows that again $t = w = 3$, with $d_i \neq 0$ when $i = 0$, 1, 2 and 4. Labelling the rows of $\tilde{\mathbf{B}}$ by γ_0, γ_1, γ_2 and γ_4 gives

$$\tilde{\mathbf{B}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 4 & 0 & 0 \\ 1 & 0 & 4 & 0 \\ 1 & 4 & 0 & 24 \end{bmatrix},$$

which is invertible. Hence there are unique linear combinations of the residual mean squares which provide unbiased estimators for each of $\gamma_0, \ldots, \gamma_5$, σ_0^2, σ_1^2, σ_2^2 and σ_4^2. All treatment effects, including the overall mean, can be tested and assigned standard errors.

5. Conclusion

It is widely believed that estimation and inference are straightforward for mixed models in which all possible variance-covariance matrices commute with each other and with the matrix of orthogonal projection onto the space of all possible fitted values of the expectation vector. As summarized in Section 2, this is indeed true when the dimension $m + 1$ of the commutative algebra spanned by all possible variance-covariance matrices is equal to the number $w+1$ of linearly independent unknown variance components. However, when $m > w$ then there are four possibilities for estimability of variance components. We recommend summarizing the data in an analysis-of-variance table in every case. However, the criterion for unique estimability of all variance components is no longer a function of how many residual degrees of freedom are non-zero: what is needed is that the matrix $\tilde{\mathbf{B}}$ introduced in Section 3 be invertible.

The examples discussed in Section 4 show realistic experimental situations where different assumptions about the variance-covariance matrix, combined with different simple methods of assigning treatments to experimental units, can lead to all four types of behaviour: variance components may be all estimable or not, and their estimators may or may not be unique linear combinations of the residual mean squares. This shows that, when an experiment is being planned, it is advisable to construct the skeleton analysis-of-variance table and find the properties of the matrix $\tilde{\mathbf{B}}$, before the experiment is carried out.

Acknowledgements

This work was partially supported by national funds of FCT - Foundation for Science and Technology under UID/MAT/00212/2013.

Appendix: R code

```
a=4
b=4
c=6

Ia= as.matrix(diag(a))
Ib= as.matrix(diag(b))
Ic= as.matrix(diag(c))

Ja= rep(1,a)
Jb= rep(1,b)
Jc= rep(1,c)

X0=Ic%x%Ib%x%Ia
X1=Ic%x%Jb%x%Ia
```

```
X2=Ic%x%Ib%x%Ja
X3=Ic%x%Jb%x%Ja

M0=X0%*%t(X0)
M1=X1%*%t(X1)
M2=X2%*%t(X2)
M3=X3%*%t(X3)

Q3=(1/b)*(1/a)*M3
Q2=(1/a)*M2-Q3
Q1=(1/b)*M1-Q3
Q0=M0-Q1-Q2-Q3
```

References

[1] R. A. Bailey, A unified approach to design of experiments. J. Roy. Statist. Soc. A 144 (1981) 214–223.

[2] R. A. Bailey, Strata for randomized experiments (with discussion). J. Roy. Statist. Soc. B 53 (1991) 27–78.

[3] R. A. Bailey, General balance: artificial theory or practical relevance?, in: T. Caliński, R. Kala (Eds), Proceedings of the International Conference on Linear Statistical Inference LINSTAT '93, Kluwer, Dordrecht, Netherlands, 1994, pp. 171–184.

[4] R. A. Bailey, Orthogonal partitions in designed experiments. Designs, Codes and Cryptography, 8 (1996) 45–77.

[5] R. A. Bailey, Association Schemes. Cambridge Studies in Advanced Mathematics, 84, Cambridge University Press, Cambridge, 2004.

[6] R. A. Bailey, Design of Comparative Experiments. Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, Cambridge, 2008.

[7] R. A. Bailey, C. J. Brien, Randomization-based models for multitiered experiments: I. A chain of randomizations. Ann. Statist. (2016), to appear.

[8] K. G. Brown, On analysis of variance in the mixed model. Ann. Statist. 12 (1984) 1488–1499.

[9] T. Caliński, S. Kageyama, Block Designs: A Randomization Approach. Vol. I: Analysis. Lecture Notes in Statistics, 150, Springer-Verlag, New York, 2000.

[10] F. Carvalho, J. T. Mexia, C. Santos, C. Nunes, Inference for types and structured families of commutative orthogonal block structures. Metrika 78 (2015) 337–372.

[11] R. Christensen, Plane Answers to Complex Questions. Springer Texts in Statistics, Springer-Verlag, New York, 1987.

[12] S. S. Ferreira, Inferência para modelos ortogonais com segregação. Unpublished PhD thesis, Universidade da Beira Interior, 2006.

[13] S. S. Ferreira, D. Ferreira, E. Moreira, J. T. Mexia, Inference for L orthogonal models. J. Interdiscip. Math. 12 (2009) 815–824.

[14] S. S. Ferreira, D. Ferreira, E. Moreira, J. T. Mexia, Exact estimators for normal linear mixed models. Far East J. Theor. Statist. 31 (2010) 33–48.

[15] S. S. Ferreira, D. Ferreira, C. Nunes, J. T. Mexia, Crossing segregated models with commutative orthogonal block structure, in: Proceedings of the 8^{th} International Conference on Numerical Analysis and Applied Mathematics, ICNAAM 2010, AIP Conf. Proc., 1281, 2010, pp. 1229–1232.

[16] S. S. Ferreira, D. Ferreira, C. Nunes, J. T. Mexia, Estimation of variance components in linear mixed models with commutative orthogonal block structure. Revista Colombiana de Estadística 36 (2013) 261–271.

[17] M. Fonseca, J. T. Mexia, R. Zmyślony, Binary operations on Jordan algebras and orthogonal normal models. Linear Algebra Appl. 417 (2006) 75–86.

[18] M. Fonseca, J. T. Mexia, R. Zmyślony, Jordan algebras, generating pivot variables and orthogonal normal models. J. Interdiscip. Math. 10 (2007) 305–326.

[19] M. Fonseca, J. T. Mexia, R. Zmyślony, Inference in normal models with commutative orthogonal block structure. Acta Comment. Univ. Tartu. Math. 12 (2008) 3–16.

[20] M. Fonseca, J. T. Mexia, R. Zmyślony, Least squares and generalized least squares in models with orthogonal block structure. J. Statist. Plan. Infer. 140 (2010) 1346–1352.

[21] A. M. Houtman, T. P. Speed, Balance in designed experiments with orthogonal block structure. Ann. Statist. 11 (1983) 1069–1085.

[22] J. Isotalo, S. Puntanen, G. P. H. Styan, The BLUE's covariance matrix revisited: a review. J. Statist. Plan. Infer. 138 (2008) 2722–2737.

[23] T. Kariya, H. Kurata, Generalized Least Squares. Wiley, New York, 2004.

[24] W. Kruskal, When are Gauss-Markov and least squares estimators identical? A coordinate-free approach. Ann. Math. Statist. 39 (1968) 70–75.

[25] H. Macharia, P. Goos, D-optimal and D-efficient equivalent-estimation second-order split-plot designs. J. Qual. Techno. 42 (2010) 358–372.

[26] A. Markiewicz, S. Puntanen, G. P. H. Styan, A note on the interpretation of the equality of OLSE and BLUE. Pakistan J. Statist. 26 (2010) 127–134.

[27] J. A. Nelder, The analysis of randomized experiments with orthogonal block structure. I. Block structure and the null analysis of variance. Proc. Roy. Soc. Lond. Ser. A 273 (1965) 147–162.

[28] J. A. Nelder, The analysis of randomized experiments with orthogonal block structure. II. Treatment structure and the general analysis of variance. Proc. Roy. Soc. Lond. Ser. A 273 (1965) 163–178.

[29] C. Nunes, C. Santos, J. T. Mexia, Relevant statistics for models with commutative orthogonal block structure and unbiased estimator for variance components. J. Interdiscip. Math. 11 (2008) 553–564.

[30] P. Parker, S. M. Kowalski, G. G. Vining, Construction of balanced equivalent estimation second-order split-plot designs. Technometrics 49 (2007) 56–65.

[31] R. W. Payne, GenStat. Wiley Interdisciplinary Reviews: Computational Statistics 1 (2009) 255–258. DOI:10.1002/wics.32.

[32] R. W. Payne, D. A. Murray, S. A. Harding, D. B. Baird, D. M. Soutar, GenStat for Windows: Introduction (12th Edition). VSN International, Hemel Hempstead, 2009.

[33] S. Puntanen, G. P. Styan, The equality of the ordinary least squares estimator and the best linear unbiased estimator (with discussion). Amer. Statist. 43 (1989) 153–164.

[34] R Core Team, R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, 2013.

[35] F. E. Satterthwaite, An approximate distribution of estimates of variance components. Biometrics Bulletin 2 (1946) 110–114.

[36] H. Scheffé, The Analysis of Variance. John Wiley & Sons Inc., New York, 1959.

[37] J. Seely, Linear spaces and unbiased estimators. Applications to a mixed model. Ann. Math. Statist. 41 (1970) 1735–1748.

[38] J. Seely, Quadratic subspaces and completeness. Ann. Math. Statist. 42 (1971) 710–721.

[39] J. Seely, G. Zyskind, Linear spaces and minimum variance unbiased estimators. Ann. Math. Statist. 42 (1971) 691–703.

[40] T. H. Szatrowski, Necessary and sufficient conditions for explicit solutions in the multivariate normal estimation problem for patterned means and covariances. Ann. Statist. 8 (1980) 802–810.

[41] T. H. Szatrowski, J. J. Miller, Explicit maximum likelihood estimates from balanced data in the mixed model of the analysis of variance. Ann. Statist. 8 (1980) 811–819.

[42] T. Tjur, Analysis of variance models in orthogonal designs (with discussion). International Statistical Review 52 (1984) 33–81.

[43] G. G. Vining, S. M. Kowalski, D. C. Montgomery, Response surface designs within a split-plot structure. J. Qual. Techno. 37 (2005) 115–129.

[44] R. Zmyślony, A characterization of best linear unbiased estimators in the general linear model, in: Mathematical Statistics and Probability Theory, Proceedings of the Sixth International Conference, Wisla 1978, Lecture Notes in Statistics, 2, Springer, New York-Berlin, 1980, pp. 365–373.