Jordan Algebras, Generating Pivot Variables and Orthogonal Normal Models

Miguel Fonseca, João Tiago Mexia
Department of Mathematics, Faculty of Science and Technology
New University of Lisbon
Monte da Caparica 2829-516 Caparica Portugal
Email: fonsecamig@yahoo.com

Roman Zmyślony
Faculty of Mathematics, Computer Sciences and Econometrics
University of Zielona Góra - Podgórna 50
65-246 Zielona Góra Poland
Email: r.zmyslony@im.uz.zgora.pl

Abstract

Jordan algebras are used to present normal orthogonal models in a canonical form. It is shown that the usual factor based formulation of such models may, many times, be obtained imposing restrictions on the parameters of the canonical formulation, and examples are presented.

The canonical model formulation is interesting since it leads to complete sufficient statistics. These statistics may be used to obtain pivot variables that induce probability measures in the parameter space. Monte Carlo generated samples, of arbitrary size, may be obtained having the induced probability measures. These samples may be screened so that the restrictions corresponding to the direct model formulations hold. Inference is presented using such samples.

1 Introduction

Jordan algebras were introduced (see [6]) to provide an algebraic foundation for Quantum Mechanics. Later these structures were applied (see [21], [**?**] and [22]) to study estimation problems, namely to obtain minimum variance unbiased estimators. They were now called quadratic vector spaces since they are vector spaces, constituted by symmetric matrices that contain the squares of every matrix in the space. For priority's sake we will use the first name. We are interested in commutative Jordan algebras, where matrices commute. These algebras have (see [22]) unique principal basis constituted by orthogonal projection matrices, all of them mutually orthogonal.

These structures enable an unified presentation of wide classes of normal models. A normal orthogonal model belongs to the class associated to an algebra if:

1. the mean vector $\boldsymbol{\mu}^n$ of the observation vector $\mathbf{y}^n$ belongs to the range space $\mathrm{R}(\mathbf{M})$ of a matrix $\mathbf{M}$ belonging to the algebra;

2. the covariance matrix $\mathbf{V}$ of $\mathbf{y}^n$ also belongs to that commutative Jordan algebra.

We will give a canonical formulation for such classes of models. The parameters in such canonical formulations will be the canonical parameters. Besides the canonical, a direct formulation is given, This formulation takes in account relevant factors either with fixed or with random effects. Usually the direct formulation requires that certain restrictions are satisfied by the canonical parameters.

The canonical formulation of the models will be used to obtain complete sufficient statistics. These statistics may be used to derive pivot variables involving the canonical parameters. The use of pivot variables in statistical inference was greatly enhanced by the introduction of generating pivot variables (see [27] and [28]). This and related concepts have had recently many interesting applications, for instance see [31] and [9][31] and [9]. Moreover, in [1] the Carathéodory theorem is used to show how pivot variables induce probability measures in parameter spaces, allowing the use of Monte Carlo methods to generate distributions. In these distributions, the observations vector acts as a parameter. Thus, these distributions may be considered as *a-posteriori* despite no *a-priori* distribution having been assumed. The assumption of *a-priori* distributions was rendered discardeable by the existence of generating pivot variables involving the canonical parameters. These variables may be considered as structural invariants.

In [25], ancillary statistics are used in inference about parameters, namely in the *plug-in* of estimates of nuisance parameters. In this case, ancillary information could also be used, since the estimators are functions of minimal sufficient statistics.

Monte Carlo methods may be used to obtain arbitrarily large samples with the *a-posteriori* distributions. Inference using such samples can be based on the Glivenko-Cantelli theorem (see [11], page 20). A special version of this theorem may be used when we filter the samples.

In the next section Jordan algebras and pivot variables are presented, and the Glivenko-Cantelli theorem is considered as well as its special version and related results. After this we consider canonical forms for normal orthogonal models deriving the relevant pivot variables thus obtaining a framework that we apply, following [3], to cross-nested balanced models and to the model considered in [31]. The fourth and last sections show how to apply the *posteriori* distributions to carry out inference, first on variance components, and then on the fixed-effects part of mixed models.

2 Preliminary Results

Superscripts will be used to indicate the number of components of vectors. $\mathbf{I}_s$ will be the $s \times s$ identity matrix, $\mathbf{J}_s = \mathbf{1}^s \mathbf{1}^{s\prime}$ and $\bar{\mathbf{J}}_s = \mathbf{I}_s - \frac{1}{s}\mathbf{J}_s$, while $\mathbf{K}_s$ will be obtained deleting the first line equal to $\frac{1}{\sqrt{s}}\mathbf{1}^s$ of an orthogonal $s \times s$ matrix.

Let $g_1, ..., g_w$ be the ranks of the $n \times n$ matrices $\mathbf{Q}_1, ..., \mathbf{Q}_w$ in the principal basis of a commutative Jordan algebra. If $\sum_{j=1}^{w} g_j < n$, we can take

$\mathbf{Q}_{w+1} = \mathbf{I}_n - \sum_{j=1}^{w} \mathbf{Q}_w$ and $g_{w+1} = n - \sum_{j=1}^{w} g_w$ to obtain the principal basis $\{\mathbf{Q}_1, ..., \mathbf{Q}_{w+1}\}$ of a Jordan algebra that contains $\mathbf{I}_n$. In what follows we consider only commutative Jordan algebras that contain $\mathbf{I}_n$.

To see how pivot variables may be used we start with a simple example. The sum of squares S of residuals from the mean of a size n normal sample with variance σ^2, is the product by σ^2 of a central chi-square with $n-1$ degrees of freedom, we put $S \sim \sigma^2 \chi^2_{n-1}$. Then, $\frac{S}{\sigma^2} \sim \chi^2_{n-1}$ and, with $x_{n-1,p}$ the quantile of χ^2_{n-1} for probability p, when $S = s$ we have, with $p_1 < p_2$,

$$\mathbb{P}\left[\frac{s}{x_{n-1,p_1}} \leq \sigma^2 \leq \frac{s}{x_{n-1,p_2}}\right] = p_2 - p_1. \tag{1}$$

We thus obtained probabilities for every closed interval containing σ^2. Now, closed intervals generate the σ-algebra of borelians so that (see [29], pgs. 17 to 20) Carathéodory theorem may be used to show that there is a single probability measure that for closed intervals gives the same results as (1).

When vector parameters $\boldsymbol{\theta}^k$ are considered, some care has to be taken in the restrictions that enter into the definition of the parameter space Ω. In the cases we will consider, the restrictions will be of the type $\mathbf{B}\boldsymbol{\theta}^k \geq \mathbf{r}^l$ with $\mathbf{B}$ a $l \times k$ matrix. Following [1], given the pivot variable $\mathbf{Z}^k = \mathbf{g}^k(\mathbf{t}^d, \boldsymbol{\theta}^h)$, where $\mathbf{t}^d$ is a statistic, we assume that the function $\mathbf{g}^k(\boldsymbol{\theta}^h|\mathbf{t}^d) = \mathbf{g}^k(\mathbf{t}^d, \boldsymbol{\theta}^h)$ has a measurable inverse $\mathbf{m}^h(\mathbf{g}^k|\mathbf{t}^d)$. This inverse function can be used to induce probability measures $\mathbb{P}\big[\cdot \bigm| \mathbf{t}^d\big]$ in the σ-algebra of the k dimensional borelians. If $\Omega \subset \mathbb{R}^k$, we then consider the restrictions $\mathbb{P}^*\big[\cdot|\mathbf{t}^d\big]$ of the $\mathbb{P}\big[\cdot \bigm| \mathbf{t}^d\big]$ to Ω. Since we may assume that Ω is a borelian and that, whatever $\mathbf{t}^d$, $\mathbb{P}\big[\Omega \bigm| \mathbf{t}^d\big] > 0$ the definition of the restrictions is trivial. Besides this, if $\Omega = \mathbb{R}^k$, $\mathbb{P}\big[\cdot \bigm| \mathbf{t}^d\big] = \mathbb{P}^*\big[\cdot|\mathbf{t}^d\big]$ whatever $\mathbf{t}^d$. Let β_Ω be the set of k-dimensional real borelians contained in Ω, and put $G(C, \mathbf{t}^d) = \mathbb{P}^*\big[C|\mathbf{t}^d\big]$ for every $C \in \beta_\Omega$. Since, whatever $\mathbf{t}^d$, $\mathbb{P}^*\big[C|\mathbf{t}^d\big] = G(C, \mathbf{t}^d)$ is a probability measure defined in β_Ω and, for every $C \in \beta_\Omega$, $h(\mathbf{t}^d|C) = G(C, \mathbf{t}^d)$ is measurable (see [7], pg. 19), we will have a probability kernel. If the probability measures $\mathbb{P}^*\big[C|\mathbf{t}^d\big]$ take all values from 0 to 1 we say that we have a regular probability kernel. For the kernel to be regular it is enough that the distributions associated to the $\mathbb{P}\big[C \bigm| \mathbf{t}^d\big]$ are continuous. Let us assume that we have a regular kernel, then, for every $q \in]0,1[$, we can obtain $C_q(\mathbf{t}^d) \in \beta_\Omega$ such that, whatever $\mathbf{t}^d$,

$$\mathbb{P}\big[C_q(\mathbf{t}^d) \bigm| \mathbf{t}^d\big] = q, \tag{2}$$

i.e. $C_q(\mathbf{t}^d)$ is a set whose probability of containing the true parameter value is, whatever the value of the statistic $\mathbf{t}^d$, q. This probability may be considered as a conditional probability since it depends on the value of $\mathbf{t}^d$ and thus on the observations vector. Nevertheless, since the mean value of a constant is the constant itself, when we decondition either in order to the statistic or to the observations vector we obtain, through the procedure used in deriving the sets $C_q(\mathbf{t}^d)$, confidence sets of load q.

As stated in the introduction we will use an extension of the

Theorem 1 (Glivenko-Cantelli Theorem)**.** *With $\tilde{F}_n(\cdot)$ the empirical distribution of a size n sample of scalar observations with distribution $F(\cdot)$ the supreme S_n of $|\tilde{F}_n - F|$ converges almost surely to zero. We denote this by writing $S_n \xrightarrow[n\to\infty]{a.s.} 0$.*

When a filter screen is applied to an initial sample $\mathbf{V}_1^s, ..., \mathbf{V}_n^s$ we only retain the elements that fulfill a condition. If the probability of the condition being fulfilled is q we get a screened sample $\dot{\mathbf{V}}_1^s, ..., \dot{\mathbf{V}}_N^s$ where N has the binomial distribution with parameters n and q from which we may obtain a random screened sample $\dot{X}_1, ..., \dot{X}_N$ with $\dot{X}_i = u(\dot{\mathbf{V}}_i^s)$, $j = 1, ..., N$. As before, let us represent by $\tilde{F}_m(\cdot)$ the empirical distribution of a sample of size m. The empirical sample distribution of the final screened sample will be

$$\tilde{F}_n^*(x) = \sum_{m=0}^{n} \binom{n}{m} q^m (1-q)^{n-m} \tilde{F}_m(x), \tag{3}$$

so that, since $0 \leq S_n \leq 1$, $n \geq 1$, we get

$$\begin{aligned} S_n^* = \sup\{|\tilde{F}_n^*(x) - F(x)|\} &\leq \sum_{m=0}^{n} \binom{n}{m} q^m (1-q)^{n-m} S_m \\ &\leq \sum_{m=0}^{\bar{m}-1} \binom{n}{m} q^m (1-q)^{n-m} + \sum_{m=\bar{m}}^{n} \binom{n}{m} q^m (1-q)^{n-m} S_m. \end{aligned} \tag{4}$$

Now, whatever $\bar{m}$,

$$\sum_{m=0}^{\bar{m}-1} \binom{n}{m} q^m (1-q)^{n-m} \xrightarrow[n\to\infty]{} 0 \tag{5}$$

thus it is now easy to use the Glivenko-Cantelli theorem to establish the

Theorem 2 (Glivenko-Cantelli Theorem for Screened Samples)**.** *S_n^* will almost surely converge to zero when $n \to \infty$.*

It would be interesting to consider multi-dimensional versions of the Glivenko-Cantelli theorem but, in what follows, the one-dimensional versions suffice. About these we point out that, when $q = 1$, the screened version of the Glivenko-Cantelli theorem reduces to the classical one, so that the former is a generalization of the latter.

The reverse of the Glivenko-Cantelli theorem will be useful in deriving sets $C_{n,p}(\mathbf{t}^d)$ with limit confidence levels. We then assume that $F(\cdot)$ has density $f(\cdot)$ and represent by x_p and $\tilde{x}_{n,p}^*$ the quantiles, for probability p, of $F(\cdot)$ and $\tilde{F}_n^*(\cdot)$, in order to establish

Theorem 3 (Reverse Glivenko-Cantelli Theorem)**.** *Let $\tilde{F}_n^*$ and F be distributions with quantiles $\tilde{x}_{n,p}^*$ and x_p. If $\sup\{\tilde{F}_n^*(x) - F(x)\} \xrightarrow[n\to\infty]{a.s.} 0$ and F has continuous density $f > 0$ whenever $0 < F(x) < 1$, whatever $0 < \alpha < 1$*

$$D_n(\alpha) = \sup\left\{|\tilde{x}_{n,p} - x_p|, \frac{\alpha}{2} < p < 1 - \frac{\alpha}{2}\right\} \xrightarrow[n\to\infty]{a.s.} 0$$

Proof. According to the Weierstrass theorem, f will have a minimum $a > 0$ in $\left[x_{\frac{\alpha}{2}}, x_{1-\frac{\alpha}{2}}\right]$. If $\frac{\alpha}{2} < p - \frac{\varepsilon}{a} < p + \frac{\varepsilon}{a} < 1 - \frac{\alpha}{2}$ we will have $F(x_p - \frac{\varepsilon}{a}) < p - \varepsilon < p + \varepsilon < F(x_p + \frac{\varepsilon}{a})$ and, due to $S_n < a$, $\tilde{F}_n^*(x_p - \frac{\varepsilon}{a}) < p < \tilde{F}_n^*(x_p + \frac{\varepsilon}{a})$ so that $x_p - \frac{\varepsilon}{a} < \tilde{x}_{n,p} < x_p + \frac{\varepsilon}{a}$. Since ε is an arbitrary positive number the thesis is established. $\square$

3 Models

3.1 Canonical Forms

We now present orthogonal normal models in a canonical form convenient for applying commutative Jordan Algebras in deriving sufficient complete statistics. From these statistics generating pivot variables will be obtained. In the second section we consider direct formulations for the model. In the third section we study two special cases for which we will derive the canonical form and obtain the relevant statistics.

Model orthogonality is based on orthogonal partitions of the sample space such that the orthogonal projections of the observation vector on the spaces in those partitions are uncorrelated. Let us write $\boxplus$ for orthogonal direct sum of subspaces. Then, we have the orthogonal partitions

$$\mathbb{R}^n = \boxplus_{i=1}^{w+1} \nabla_i. \tag{6}$$

Let the orthogonal projection matrices on $\nabla_i, ..., \nabla_{w+1}$ be $\mathbf{Q}_1, ..., \mathbf{Q}_{w+1}$. Then $g_i = \text{rank}(\mathbf{Q}_i) = \dim(\nabla_i)$, $i = 1, ..., w+1$, as well as $\sum_{i=1}^{w+1} g_i = n$ and as $\sum_{i=1}^{w+1} \mathbf{Q}_i = \mathbf{I}_n$. If the column vectors of $\mathbf{A}_i$ constitute an orthonormal basis for ∇_i we will have $\mathbf{Q}_i = \mathbf{A}_i\mathbf{A}'_i$ as well as $\mathbf{A}'_i\mathbf{A} = \mathbf{I}_{g_i}$, $i = 1, ..., w+1$. Moreover, with $\mathbf{0}_{r,s}$ the null $r \times s$ matrix we will have $\mathbf{A}'_i\mathbf{A}_j = \mathbf{0}_{g_i,g_j}$ and $\mathbf{A}'_i\mathbf{Q}_j = \mathbf{0}_{g_i,n}$, whenever $i \neq j$.

Let us establish

Proposition 1. *A normal orthogonal model associated with the orthogonal partition in* (6) *(as with the corresponding Jordan commutative algebra) has the canonical form*

$$\mathbf{Y}^n = \sum_{i=1}^{w+1} \mathbf{A}_i \boldsymbol{\eta}_i^{g_i},$$

where vectors $\boldsymbol{\eta}_i^{g_i}$, $i = 1, ..., w+1$, are normal, independent, with mean vectors $\boldsymbol{\lambda}_i^{g_i}$, $i = 1, ..., w+1$, and covariance matrices $\gamma_i\mathbf{I}_{g_i}$, $i = 1, ..., w+1$.

We write $\boldsymbol{\eta}_i^{g_i} \sim \mathcal{N}(\boldsymbol{\lambda}_i^{g_i}, \gamma_i\mathbf{I}_{g_i})$, $i = 1, ..., w+1$.

Proof. Let $\mathbf{Q}^*$ be the orthogonal projection matrix on the sub-space that contains the observations mean vector $\boldsymbol{\mu}^n$. Since $\mathbf{Q}^*$ belongs to the algebra we will have $\mathbf{Q}^* = \sum_{i=1}^{w+1} c_i\mathbf{Q}_i$, with $c_i = 0$ or $c_i = 1$, $i = 1, ..., w+1$. We can assume without loss of generality that $\mathbf{Q}^* = \sum_{i=1}^{m} \mathbf{Q}_i$, thus

$$\boldsymbol{\mu}^n = \mathbf{Q}^*\boldsymbol{\mu}^n = \sum_{i=1}^{m} \mathbf{Q}_i\boldsymbol{\mu}^n = \sum_{i=1}^{m} \mathbf{A}_i\mathbf{A}'_i\boldsymbol{\mu}^n = \sum_{i=1}^{m} \mathbf{A}_i\boldsymbol{\lambda}_i^{g_i},$$

where $\boldsymbol{\lambda}_i^{g_i}$ is the mean vector of $\boldsymbol{\eta}_i^{g_i} = \mathbf{A}'_i\mathbf{Y}^n$, $i = 1, ..., m$. Moreover, $\boldsymbol{\lambda}_i^{g_i} = \mathbf{0}^{g_i}$, $i = m+1, ..., w+1$, will be the mean vector of $\boldsymbol{\eta}_i^{g_i} = \mathbf{A}'_i\mathbf{Y}^n$, $i = 1, ..., m$. Then

$$\mathbf{Y}^n = \mathbf{I}_n\mathbf{Y}^n = \sum_{i=1}^{w+1} \mathbf{Q}_i\mathbf{Y}^n = \sum_{i=1}^{w+1} \mathbf{A}_i\mathbf{A}'_i\mathbf{Y}^n = \sum_{i=1}^{w+1} \mathbf{A}_i\boldsymbol{\eta}_i^{g_i}. \tag{7}$$

To complete the proof we have only to point out that $\boldsymbol{\eta}_i^{g_i}$, $i = 1, ..., w+1$, will be normal and independent since, as it is easily seen, their cross-covariance matrices are null and their covariance matrices are $\gamma_i \mathbf{I}_{g_i}$, $i = 1, ..., w+1$, while their joint distribution is normal. $\square$

In the previous proof we showed that

$$\boldsymbol{\mu}^n = \sum_{i=1}^{m} \mathbf{A}_i \boldsymbol{\lambda}_i^{g_i}. \tag{8}$$

Moreover, since the $\boldsymbol{\eta}_i^{g_i}$ are independent, the covariance matrix of $\mathbf{Y}^n$ will be

$$\mathbf{V} = \sum_{i=1}^{w+1} \gamma_i \mathbf{Q}_i \tag{9}$$

thus (see [1])

$$\begin{cases} \det(\mathbf{V}) = \prod\limits_{i=1}^{w+1} \gamma_i^{g_i} \\ \mathbf{V}^{-1} = \sum\limits_{i=1}^{w+1} \gamma_i^{-1} \mathbf{Q}_i \end{cases}. \tag{10}$$

The following result will be useful since it shows that, for every $\boldsymbol{\mu}^n$ and $\mathbf{V}$, the $\boldsymbol{\lambda}_i^{g_i}$, $i = 1, ..., m$, and the γ_i, $i = 1, ..., w+1$, are unique.

Proposition 2. *We have* $\sum_{i=1}^m \mathbf{A}_i \mathbf{a}_i^{g_i} = \sum_{i=1}^m \mathbf{A}_i \mathbf{b}_i^{g_i}$ *if and only if* $\mathbf{a}_i^{g_i} = \mathbf{b}_i^{g_i}$, $i = 1, ..., m$, *and* $\sum_{i=m+1}^{w+1} u_i \mathbf{Q}_i = \sum_{i=m+1}^{w+1} v_i \mathbf{Q}_i$ *when and only when* $u_i = v_i$, $i = m+1, ..., w+1$.

Proof. For either part of the thesis it suffices to establish the necessary conditions since the corresponding sufficient one is self-evident. Starting with the first part, since $\nabla_i \bigcap \boxplus_{j \neq i} \nabla_j = \{\mathbf{0}^n\}$, if $\sum_{j=1}^{w+1} \mathbf{A}_i \mathbf{a}_i^{g_i} = \sum_{j=1}^{w+1} \mathbf{A}_i \mathbf{b}_i^{g_i}$, *i.e.*, if $\mathbf{A}_i(\mathbf{b}_i^{g_i} - \mathbf{a}_i^{g_i}) = -\sum_{j \neq i} \mathbf{A}_j(\mathbf{b}_j^{g_j} - \mathbf{a}_j^{g_j}) \in \nabla_i \bigcap \boxplus_{j \neq i} \nabla_j$, we have $\mathbf{A}_i(\mathbf{b}_i^{g_i} - \mathbf{a}_i^{g_i}) = \mathbf{0}^n$ as well as $\mathbf{b}_i^{g_i} - \mathbf{a}_i^{g_i} = \mathbf{A}_i' \mathbf{A}_i (\mathbf{b}_i^{g_i} - \mathbf{a}_i^{g_i}) = \mathbf{0}^n$, $i = 1, ..., m$, so the first part is established. Moreover, if $\sum_{i=m+1}^{w+1} u_i \mathbf{Q}_i = \sum_{i=m+1}^{w+1} v_i \mathbf{Q}_i$ we have $u_i \mathbf{Q}_i = \mathbf{Q}_i \left(\sum_{j=m+1}^{w+1} u_j \mathbf{Q}_j \right) = \mathbf{Q}_i \left(\sum_{j=m+1}^{w+1} v_j \mathbf{Q}_j \right) = v_i \mathbf{Q}_i$, thus $u_i = v_i$, $i = m+1, ..., w+1$, and the proof is complete. $\square$

As we saw,$\mathbf{A}_i' \mathbf{Y}^n = \boldsymbol{\eta}_i^{g_i}$, $i = 1, ..., w+1$, $\mathbf{A}_i' \boldsymbol{\mu}^n = \boldsymbol{\lambda}_i^{g_i}$, $i = 1, ..., m$, and $\mathbf{A}_i' \boldsymbol{\mu}^n = \mathbf{0}^{g_i}$, $i = m+1, ..., w+1$. We also get

$$\mathbf{A}_i' \mathbf{V}^{-1} \mathbf{A}_i = \mathbf{A}_i' \left(\sum_{j=1}^{w+1} \gamma_j^{-1} \mathbf{Q}_j \right) \mathbf{A}_i = \gamma_i^{-1} \mathbf{A}_i' \mathbf{A}_i' = \gamma_i^{-1} \mathbf{I}_{g_i}, i = 1, ..., w+1 \tag{11}$$

so that

$$\begin{aligned} (\mathbf{y}^n - \boldsymbol{\mu}^n)' \mathbf{V}^{-1} (\mathbf{y}^n - \boldsymbol{\mu}^n) &= \sum_{i=1}^{w+1} \frac{1}{\gamma_i} (\mathbf{y}^n - \boldsymbol{\mu}^n)' \mathbf{A}_i \mathbf{A}_i' (\mathbf{y}^n - \boldsymbol{\mu}^n) \\ &= \sum_{i=1}^{m} \frac{\|\tilde{\boldsymbol{\eta}}_i^{g_i} - \boldsymbol{\lambda}_i^{g_i}\|^2}{\gamma_i} + \sum_{i=m+1}^{w+1} \frac{s_i}{\gamma_i}, \end{aligned} \tag{12}$$

where $\tilde{\boldsymbol{\eta}}_i^{g_i} = \mathbf{A}_i'\mathbf{y}^n$ and $s_i = \|\mathbf{A}_i'\mathbf{y}^n\|^2$. Thus, the model's density will be

$$n(\mathbf{y}^n|\boldsymbol{\mu}^n, \mathbf{V}) = \frac{\exp\left(-\frac{1}{2}\left(\sum_{i=1}^{m} \frac{\|\tilde{\boldsymbol{\eta}}_i^{g_i} - \boldsymbol{\lambda}_i^{g_i}\|^2}{\gamma_i} + \sum_{i=m+1}^{w+1} \frac{s_i}{\gamma_i}\right)\right)}{(2\pi)^n \prod_{i=1}^{w+1} \gamma_i^{\frac{g_i}{2}}}, \tag{13}$$

and (see [26], pg. 31 and 32) we have the set of complete sufficient statistics $\tilde{\boldsymbol{\eta}}_i^{g_i}$, $i = 1, ..., m$, and s_i, $i = m+1, ..., w+1$. Moreover, according to the Blackwell-Lehmann-Scheffé theorem, the $\tilde{\boldsymbol{\eta}}_i^{g_i}$, $i = 1, ..., m$, and the $\tilde{\gamma}_i = \frac{s_i}{g_i}$, $i = m+1, ..., w+1$ are UMVUE for the $\boldsymbol{\lambda}_i^{g_i}$, $i = 1, ..., m$, and the γ_i, $i = m+1, ..., w+1$. To avoid over-parametrization, we now assume that

$$\gamma_i = \sum_{j=m+1}^{w+1} b_{i,j}\gamma_j, i = 1, ..., m, \tag{14}$$

so that we will also have the UMVUE

$$\tilde{\gamma}_i = \sum_{j=m+1}^{w+1} b_{i,j}\tilde{\gamma}_j, i = 1, ..., m. \tag{15}$$

The estimable vectors will be the $\boldsymbol{\psi}_i^{r_i} = \mathbf{B}_i\boldsymbol{\lambda}_i^{g_i}$, $i = 1, ...m$, for which we have the UMVUE $\tilde{\boldsymbol{\psi}}_i^{r_i} = \mathbf{B}_i\boldsymbol{\eta}_i^{g_i}$, $i = 1, ...m$.

The joint distribution of the $\mathbf{A}_i'\mathbf{y}^n$, $i = 1, ..., w+1$, is normal and, since their cross covariance matrices are null, they will be (see [14], pg. 42) independent. Thus the $\boldsymbol{\eta}_i^{g_i} = \mathbf{A}_i'\mathbf{y}^n$, $i = 1, ..., m$, and the $\tilde{\gamma}_i = \frac{1}{g_i}\|\mathbf{A}_i'\mathbf{y}^n\|^2$, $i = m+1, .., w+1$, will be independent. Moreover, the $\boldsymbol{\eta}_i^{g_i} = \mathbf{A}_i'\mathbf{y}^n$ and the $\tilde{\gamma}_i = \frac{1}{g_i}\|\mathbf{A}_i'\mathbf{y}^n\|^2$, $i = 1, ..., m$, will also be independent, as well as the $\boldsymbol{\psi}_i^{r_i}$ and the $\tilde{\gamma}_i$, $i = 1, ..., m$, with $\boldsymbol{\psi}_i^{r_i} = \mathbf{B}_i\boldsymbol{\lambda}_i^{g_i}$ an estimable vector, $i = 1, ..., m$. It may be interesting to point out that we may take $\mathbf{B}_i = \mathbf{I}_{g_i}$, so that $\boldsymbol{\lambda}_i^{g_i}$ is itself an estimable vector. If $\text{rank}(\mathbf{B}_i) = r_i$, $\mathbf{B}_i\mathbf{B}_i'$ will be positive definite and $\boldsymbol{\psi}_i^{r_i} = \mathbf{B}_i\boldsymbol{\lambda}_i^{g_i}$ will be a regular estimable vector. In what follows we restrict ourselves to such estimable vectors.

Writing $Z \sim \chi_r^2$ $[\sim \chi_{r,\delta}^2]$ to indicate that Z is distributed as a central chi-square with r degrees of freedom [non-central chi-square with r degrees of freedom and non-centrality parameter δ], we have

$$\begin{cases} Z_i = \dfrac{S_i}{\gamma_i} \sim \chi_{g_i}^2, i = m+1, ..., w+1 \\ U_i(\boldsymbol{\psi}_i^{r_i}) = \dfrac{(\boldsymbol{\psi}_i^{r_i} - \tilde{\boldsymbol{\psi}}_i^{r_i})'(\mathbf{B}_i\mathbf{B}_i')^{-1}(\boldsymbol{\psi}_i^{r_i} - \tilde{\boldsymbol{\psi}}_i^{r_i})}{\gamma_i} \sim \chi_{r_i}^2, i = 1, ..., m \\ U_i(\boldsymbol{\psi}_{i,0}^{r_i}) = \dfrac{(\boldsymbol{\psi}_{i,0}^{r_i} - \tilde{\boldsymbol{\psi}}_i^{r_i})'(\mathbf{B}_i\mathbf{B}_i')^{-1}(\boldsymbol{\psi}_{i,0}^{r_i} - \tilde{\boldsymbol{\psi}}_i^{r_i})}{\gamma_i} \sim \chi_{r_i,\delta_{i,0}}^2, i = 1, ..., m \end{cases} \tag{16}$$

with

$$\delta_{i,0} = \frac{(\boldsymbol{\psi}_{i,0}^{r_i} - \boldsymbol{\psi}_i^{r_i})'(\mathbf{B}_i\mathbf{B}_i')^{-1}(\boldsymbol{\psi}_{i,0}^{r_i} - \boldsymbol{\psi}_i^{r_i})}{\gamma_i}, i = 1, ..., m, \tag{17}$$

this non-centrality parameter being null when and only when $\boldsymbol{\psi}_{i,0}^{r_i} = \boldsymbol{\psi}_i^{r_i}$, $i = 1, ..., m$. While Z_i, $i = m+1, ...w+1$, and $U_i(\boldsymbol{\psi}_i^{r_i})$ are pivot variables, $U_i(\boldsymbol{\psi}_{i,0}^{r_i})$, $i = 1, ..., m$, are statistics. Since $\boldsymbol{\lambda}_i^{g_i}$ are regular estimable vectors, $i = 1, ..., m$, the pivot variables now defined relate to all canonical parameters.

3.2 Direct Forms

We now consider the action of one or more factors. This approach is followed usually, for instance see [8]. Then the model may be written as

$$\mathbf{Y}^n = \sum_{i=1}^{t} \mathbf{X}_i \boldsymbol{\beta}_i^{c_i} + \mathbf{e}^n, \tag{18}$$

with $\boldsymbol{\beta}_i^{c_i} = \boldsymbol{\delta}_i^{c_i}$ fixed, $i = 1, ..., d$, and $\boldsymbol{\beta}_i^{c_i} \sim \mathcal{N}(\mathbf{0}^{c_i}, \sigma_i^2 \mathbf{I}_{c_i})$, $i = d+1, ..., t$, and $\mathbf{e}^n \sim \mathcal{N}(\mathbf{0}^n, \sigma^2 \mathbf{I}_n)$, these vectors being independent. The mean vector and covariance matrix of $\mathbf{Y}^n$ will be

$$\begin{cases} \boldsymbol{\mu}^n = \sum\limits_{i=1}^{d} \mathbf{X}_i \boldsymbol{\beta}_i^{c_i} \\ \mathbf{V} = \sum\limits_{i=d+1}^{t} \sigma_i^2 \mathbf{M}_i + \sigma^2 \mathbf{I}_n \end{cases}, \tag{19}$$

where $\mathbf{M}_i = \mathbf{X}_i \mathbf{X}_i'$, $i = 1, ..., t$. The σ_i^2, $i = 1, ..., t$, and σ^2 will be the variance components.

The necessary and sufficient condition for the matrices $\mathbf{M}_i$, $i = 1, ..., t$, and $\mathbf{I}_n$ to generate to a commutative Jordan algebra is (see [1]) that they commute. Then they will generate a commutative Jordan algebra $\mathcal{A}$ with principal basis $\{\mathbf{Q}_1, ..., \mathbf{Q}_{w+1}\}$.

We now point out that, with $\mathbf{X} = [\mathbf{X}_1', ..., \mathbf{X}_d']'$, we have

$$\boldsymbol{\mu}^n \in \mathrm{R}(\mathbf{X}) = \mathrm{R}(\mathbf{X}\mathbf{X}') = \sum_{i=1}^{d} \mathrm{R}(\mathbf{M}_i)\,. \tag{20}$$

Since $\mathbf{M}_1, ..., \mathbf{M}_t \in \mathcal{A}$, we will have

$$\mathbf{M}_i = \sum_{j=1}^{w+1} a_{i,j} \mathbf{Q}_j, i = 1, ..., t, \tag{21}$$

and so,

$$\mathbf{V} = \sum_{i=d+1}^{t} \sigma_i^2 \sum_{j=1}^{w+1} \mathbf{Q}_j + \sigma^2 \sum_{j=1}^{w+1} \mathbf{Q}_j = \sum_{j+1}^{w+1} \gamma_j \mathbf{Q}_j, \tag{22}$$

where

$$\gamma_j = \sigma^2 + \sum_{i=d+1}^{t} a_{i,j} \sigma_i^2, j = 1, ..., w+1. \tag{23}$$

We now establish

Proposition 3. *If* $\mathrm{R}(\mathbf{X}) \subset \mathbb{R}^n$ *we may order the matrices in the principal basis so that the orthogonal projection matrix* $\mathbf{Q}^+$ *on* $\mathrm{R}(\mathbf{X})$ *will be* $\sum_{j=1}^{m} \mathbf{Q}_j$, *with* $m < w+1$.

Proof. Since $\mathbf{X}\mathbf{X}' = \sum_{i=1}^{d} \mathbf{M}_i \in \mathcal{A}$, it will be a linear combination of the $\mathbf{Q}_1, ..., \mathbf{Q}_{w+1}$. Now $\mathrm{R}(\mathbf{X}\mathbf{X}') = \mathrm{R}(\mathbf{X}) \subset \mathbb{R}^n$ so that not all the coefficients in that linear combination may be non null allowing us to reorder the matrices in the principal basis to get $\mathbf{X}\mathbf{X}' = \sum_{j=1}^{m} u_j \mathbf{Q}_j$ with $u_j \neq 0$, $j = 1, ..., m$. Since the matrices in the principal basis are mutually orthogonal so that $\mathbf{Q}_j\mathbf{Q}_l$ is null whenever $j \neq l$ it is easy to show that $(\mathbf{XX}')^+ = \sum_{j=1}^{m} u_j^{-1}\mathbf{Q}_j$, and so (see [1], page 16), $\mathbf{Q}^+ = (\mathbf{XX}')(\mathbf{XX}')^+ = \sum_{j=1}^{m}\mathbf{Q}_j$. $\square$

Moreover we can, when $\mathrm{R}(\mathbf{X}) \subset \mathbb{R}^n$, and the matrices $\mathbf{Q}_1, ..., \mathbf{Q}_{w+1}$ have been reordered, consider the expressions of the γ_j, $j = 1, ..., w+1$, as a system of equations on the σ_i^2, $i = 1, ..., t$, and on σ^2. When this system has an unique solution we can use the s_j, $j = m+1, ..., w+1$, to estimate the γ_j, $j = 1, ..., w+1$, and nextly obtain UMVUE for the σ_i^2, $i = 1, ..., t$, and for σ^2. We will then have

$$\begin{cases} \sigma_i^2 = \sum_{j=m+1}^{w+1} b_{i,j}\gamma_j, i = 1, ..., t \\ \tilde{\sigma}_i^2 = \sum_{j=m+1}^{w+1} b_{i,j}\tilde{\gamma}_j, i = 1, ..., t \end{cases}. \tag{24}$$

As it is well known we may obtain negative estimators for all variance components but σ^2. Now the first expression in (24) shows that the variance components are linear combinations of the canonical parameters $\gamma_{m+1}, ..., \gamma_{w+1}$. Thus requiring $\sigma_i^2 \geq 0$, $i = 1, ..., t$, may be seen as imposing restrictions on the canonical formulation. Inference based on the less restrictive formulation will be more robust. This will also lead to a better understanding of negative estimators. They can be seen as a strong indication of nullity for the corresponding variance component or, in more extreme cases, suggesting a rethinking of the model formulation.

Nevertheless, if we are confident about the direct formulation we may accept the restriction imposed on the $\gamma_{m+1}, ..., \gamma_{w+1}$. This will lead, as we shall see, to the use of screened samples.

It may be interesting to point out that the characteristic function of linear combinations of independent central chi-squares has been recently studied in [30]. Moreover numerical methods obtained in [5] may be useful in inverting those characteristic functions.

3.3 Special Models

We start with balanced cross-nested models. We assume that there are L groups of $u_1, .., u_L$ factors. Let $a_l(1)$ be the number of levels of the first factor in the l-th group, $l = 1, ..., L$. If $u_l > 1$, there is nesting in the l-th group, and for each level of the first factor there will be $a_l(2)$ levels of the second and so on. There will be $c_l(h) = \prod_{t=1}^{h} a_l(t)$ of the first h factor levels, $h = 1, ..., u_l$, each nesting $b_l(h) = c_l(u_l)/c_l(h)$ level combinations of the remaining factors in the l-th group. We will also put $c_l(-1) = 0$ and $c_l(0) = 1$ as well as $g_l(h) = c_l(h) - c_l(h-1)$, $h = 1, ..., h_l$, $l = 1, ..., L$. For all factors we have $\prod_{l=1}^{L} = c_l(u_l)$ level combinations, which define the treatments in the design. For each treatment we will have r replicates so the number of observations will be $n = r\prod_{l=1}^{L} c_l(u_l)$.

Factors in the same group do not interact so that we may indicate the general mean, the factor effects and factor interactions by the vectors $\mathbf{h}^L$ with components $h_l = 0, ..., u_l$, $l = 1, ..., L$. The general mean will correspond to $\mathbf{0}^L$, and if $h_{l'}$ is the only non-null component of the vector $\mathbf{h}^L$ this vector will correspond to the factor of group l' with index $h_{l'}$. When $\mathbf{h}^L$ has several non-null components it will associated to the interaction between the factors labelled by these components. We will represent by Γ the set of vectors $\mathbf{h}^L$. Then (see [3]) the balanced cross-nested model can be written as

$$\mathbf{Y}^n = \sum_{\mathbf{h}^L \in \Gamma} \mathbf{X}(\mathbf{h}^L)\boldsymbol{\beta}^{c(\mathbf{h}^L)}(\mathbf{h}^L) + \mathbf{e}^n, \tag{25}$$

where $c(\mathbf{h}^L) = \prod_{l=1}^{L} = c_l(h_l)$ and, representing by $\otimes$ the Kronecker matrix product, we have

$$\mathbf{X}(\mathbf{h}^L) = \bigotimes_{l=1}^{L} \left(\mathbf{I}_{c_l(h_l)} \otimes \mathbf{1}^{b_l(h_l)}\right) \otimes \mathbf{1}^r, \mathbf{h}^L \in \Gamma. \tag{26}$$

This formulation is more detailed than the one given in [19] , pages 144–185. To the $\mathbf{h}^L \in \Gamma$ we can associate (see [3]) the, mutually orthogonal, orthogonal projection matrices

$$\mathbf{Q}(\mathbf{h}^L) = \bigotimes_{l=1}^{L} \left(\mathbf{I}_{c_(h_l-1)} \otimes \bar{\mathbf{J}}_{a_l(h_l)} \otimes \frac{1}{b_l(h_l)}\mathbf{J}_{b_l(h_l)}\right) \otimes \frac{1}{r}\mathbf{J}_r, \mathbf{h}^L \in \Gamma, \tag{27}$$

such that, with $b(\mathbf{h}^L) = \prod_{l=1}^{L} b_l(h_l)$,

$$\mathbf{X}(\mathbf{h}^L)\mathbf{X}'(\mathbf{h}^L) = b(\mathbf{h}^L) \sum_{\mathbf{k}^L:\mathbf{k}^L \leq \mathbf{h}^L} \mathbf{Q}(\mathbf{k}^L), \mathbf{h}^L \in \Gamma. \tag{28}$$

In order to obtain the principal basis of a convenient commutative Jordan algebra, besides the $\mathbf{Q}(\mathbf{h}^L)$, with $\mathbf{h}^L \in \Gamma$, we consider

$$\mathbf{Q}^{\perp} = \mathbf{I}_n - \sum_{\mathbf{h}^L \in \Gamma} \mathbf{Q}(\mathbf{h}^L). \tag{29}$$

If besides $\boldsymbol{\mu}^n$ we have only variance components as parameters, *i.e.*, $\boldsymbol{\beta}^1(\mathbf{0}^L) = \mu$ and $\boldsymbol{\beta}^{c(\mathbf{h}^L)} \sim \mathcal{N}(\mathbf{0}^{c(\mathbf{h}^L)}, \sigma^2(\mathbf{h}^L)\mathbf{I}_{c(\mathbf{h}^L)})$ for $\mathbf{h}^L \neq \mathbf{0}^L$ we get $\mathbf{Y}^n \sim \mathcal{N}(\mathbf{1}^n\mu, \mathbf{V})$ with (see [3])

$$\mathbf{V} = \sum_{\mathbf{h}^L \in \Gamma} \gamma(\mathbf{h}^L)\mathbf{Q}(\mathbf{h}^L) + \sigma^2\mathbf{Q}^{\perp}, \tag{30}$$

where

$$\gamma(\mathbf{h}^L) = \sigma^2 + \sum_{\mathbf{k}^L:\mathbf{h}^L \leq \mathbf{k}^L} b(\mathbf{h}^L)\sigma^2(\mathbf{k}^L). \tag{31}$$

With $m(\mathbf{k}^L, \mathbf{h}^L)$ the number of components of $\mathbf{k}^L$ exceeded by the corresponding components of $\mathbf{h}^L$ we have (see [3])

$$\sigma^2(\mathbf{h}^L) = b(\mathbf{h}^L)^{-1} \sum_{\mathbf{k}^L \in \circledast(\mathbf{h}^L)} (-1)^{m(\mathbf{k}^L,\mathbf{h}^L)}\gamma(\mathbf{k}^L), \mathbf{h}^L \in \Gamma \backslash \{\mathbf{0}^L\}, \tag{32}$$

where

$$\circledast(\mathbf{h}^L) = \left\{\mathbf{h}^L : h_l \leq k_l \leq \min\{h_l + 1, u_l\}, l = 1, ..., L\right\}. \tag{33}$$

The sufficient statistics will be, besides the general mean, the

$$S(\mathbf{h}^L) = \|\mathbf{Q}(\mathbf{h}^L)\|^2 \sim \chi^2_{g(\mathbf{h}^L)}, \mathbf{h}^L \in \Gamma \backslash \{\mathbf{0}^L\}, \tag{34}$$

with $g(\mathbf{h}^L) = \prod_{i=1}^{L} g(h_l)$, $\mathbf{h}^L \in \Gamma \backslash \{\mathbf{0}^L\}$. From these we get the UMVUE $\tilde{\gamma}(\mathbf{h}^L) = \frac{S(\mathbf{h}^L)}{g(\mathbf{h}^L)}$, $\mathbf{h}^L \in \Gamma \backslash \{\mathbf{0}^L\}$ and the

$$\tilde{\sigma}^2(\mathbf{h}^L) = \sum_{\mathbf{k}^L \in \circledast(\mathbf{h}^L)} (-1)^{m(\mathbf{k}^L, \mathbf{h}^L)} \tilde{\gamma}(\mathbf{k}^L), \mathbf{h}^L \in \Gamma \backslash \{\mathbf{0}^L\}. \tag{35}$$

It may be interesting to point out that, if $\circledast(\mathbf{h}^L)^+$ $[\circledast(\mathbf{h}^L)^-]$ is the set of the $\mathbf{k}^L \in \circledast(\mathbf{h}^L)$ for which $m(\mathbf{k}^L, \mathbf{h}^L)$ is even [odd], we have

$$\tilde{\sigma}^2(\mathbf{h}^L) = \tilde{\sigma}^2(\mathbf{h}^L)^+ - \tilde{\sigma}^2(\mathbf{h}^L)^-, \mathbf{h}^L \in \Gamma \backslash \{\mathbf{0}^L\}, \tag{36}$$

with

$$\begin{cases} \tilde{\sigma}^2(\mathbf{h}^L)^+ = \sum_{\mathbf{k}^L \in \circledast(\mathbf{h}^L)^+} \tilde{\gamma}(\mathbf{k}^L), \mathbf{h}^L \in \Gamma \backslash \{\mathbf{0}^L\} \\ \tilde{\sigma}^2(\mathbf{h}^L)^- = \sum_{\mathbf{k}^L \in \circledast(\mathbf{h}^L)^-} \tilde{\gamma}(\mathbf{k}^L), \mathbf{h}^L \in \Gamma \backslash \{\mathbf{0}^L\} \end{cases}. \tag{37}$$

Following [15] and [16] we could apply the generalized F tests with statistics

$$\mathcal{F}(\mathbf{h}^L) = \frac{\tilde{\sigma}^2(\mathbf{h}^L)^+}{\tilde{\sigma}^2(\mathbf{h}^L)^-}, \mathbf{h}^L \in \Gamma \backslash \{\mathbf{0}^L\}, \tag{38}$$

to test

$$H_0(\mathbf{h}^L) : \sigma^2(\mathbf{h}^L) = 0, \mathbf{h}^L \in \Gamma \backslash \{\mathbf{0}^L\}, \tag{39}$$

but we will use other tests. Actually, if the $g(\mathbf{k}^L)$, with $\mathbf{k}^L \in \circledast(\mathbf{h}^L)^+$ $[\mathbf{k}^L \in \circledast(\mathbf{h}^L)^+]$ are even, the exact distribution of $\mathcal{F}(\mathbf{h}^L)$ (see [2]) is known. It may be interesting to point out (see [2]) that at least one of these evenness conditions holds in many cases. Moreover, generalized F tests were applied successfully in the study of grapevine genetic homogeneity breakdown (see [2]).

The extension of this model to the case when the first $L' < L$ groups are of fixed effects factors was carried out by [17] using the same Jordan algebra. We then get $\mathbf{Q}(\mathbf{h}^L) = \mathbf{A}'(\mathbf{h}^L)\mathbf{A}(\mathbf{h}^L)$, $\mathbf{h}^L \in \Gamma$, as well as the UMVUE given by

$$\tilde{\gamma}(\mathbf{h}^L) = \sum_{\mathbf{k}^L \in \Gamma'^C} b(\mathbf{h}^L, \mathbf{k}^L) \frac{S(\mathbf{k}^L)}{g(\mathbf{k}^L)}, \mathbf{h}^L \in \Gamma'. \tag{40}$$

We omit the expression of the $b(\mathbf{h}^L, \mathbf{k}^L)$ since it is quite cumbersome. For more details see [17].

With Γ' the sub-set of Γ of vectors with components $h_{L'+1}, ..., h_L = 0$ and $\Gamma'^C = \Gamma \backslash \Gamma'$, we can use expression (35) to estimate the $\sigma^2(\mathbf{h}^L)$, $\mathbf{h}^L \in \Gamma'^C$ and expression (38) to obtain the generalized F test for $H_0(\mathbf{h}^L)$, with $\mathbf{h}^L \in \Gamma'^C$. As for the fixed factors, using the

$$\mathbf{A}(\mathbf{h}^L) = \frac{1}{\sqrt{rb(\mathbf{h}^L)}} \bigotimes_{l=1}^{L} \left(\mathbf{I}_{c_l(h_l-1)} \otimes \bar{\mathbf{J}}_{a_l(h_l)} \otimes \mathbf{1}^{b_l(h_l)'} \right) \otimes \mathbf{1}^{r'}, \mathbf{h}^L \in \Gamma, \tag{41}$$

it was possible to derive generalized F tests for hypothesis in the fixed effects part of the model.

Nextly we consider a variant of the model, presented by [31]. Now, a first factor nests a second and a third that cross. Each of the a levels of the first factor will nest different bc level combinations of the second and third factors. There will be a, b and c levels for the first, second and third factors and r replicates for all level combinations so that the model is balanced. We now have $n = abcr$, and

$$\mathbf{Y}^n = \mathbf{1}^n\mu+\mathbf{X}(1)\boldsymbol{\beta}^a(1)+\mathbf{X}(2)\boldsymbol{\beta}^{ab}(2)+\mathbf{X}(3)\boldsymbol{\beta}^{ac}(3)+\mathbf{X}(2,3)\boldsymbol{\beta}^{abc}(2,3)+\mathbf{e}^n, \quad (42)$$

were single indexes refer to factors while $(2,3)$ refers to the interaction between the second and third factors, and

$$\begin{cases} \mathbf{X}(1) = \mathbf{I}_a \otimes \mathbf{1}^b \otimes \mathbf{1}^c \otimes \mathbf{1}^r \\ \mathbf{X}(2) = \mathbf{I}_a \otimes \mathbf{I}_b \otimes \mathbf{1}^c \otimes \mathbf{1}^r \\ \mathbf{X}(3) = \mathbf{I}_a \otimes \mathbf{1}^b \otimes \mathbf{I}_c \otimes \mathbf{1}^r \\ \mathbf{X}(2,3) = \mathbf{I}_a \otimes \mathbf{I}_b \otimes \mathbf{I}_c \otimes \mathbf{1}^r \end{cases}. \quad (43)$$

In this model it is assumed that μ is fixed and unknown, while $\boldsymbol{\beta}^a(1) \sim \mathcal{N}(\mathbf{0}^a, \sigma^2(1)\mathbf{I}_a)$, $\boldsymbol{\beta}^b(2) \sim \mathcal{N}(\mathbf{0}^b, \sigma^2(2)\mathbf{I}_b)$, $\boldsymbol{\beta}^c(3) \sim \mathcal{N}(\mathbf{0}^c, \sigma^2(3)\mathbf{I}_c)$, $\boldsymbol{\beta}^{bc}(2,3) \sim \mathcal{N}(\mathbf{0}^{bc}, \sigma^2(2,3)\mathbf{I}_{bc})$ and $\mathbf{e}^n \sim \mathcal{N}(\mathbf{0}^n, \sigma^2\mathbf{I}_n)$, all these variables being independent. Moreover, with $\bar{\mathbf{J}}_v = \mathbf{I}_v - \frac{1}{v}\mathbf{J}_v$, it is easy to see that the orthogonal projection matrices

$$\begin{cases} \mathbf{Q}(0) = \dfrac{1}{n}\mathbf{J}_a \otimes \mathbf{J}_b \otimes \mathbf{J}_c \otimes \mathbf{J}_r \\ \mathbf{Q}(1) = \dfrac{1}{bcr}\bar{\mathbf{J}}_a \otimes \mathbf{J}_b \otimes \mathbf{J}_c \otimes \mathbf{J}_r \\ \mathbf{Q}(2) = \dfrac{1}{cr}\mathbf{I}_a \otimes \bar{\mathbf{J}}_b \otimes \mathbf{J}_c \otimes \mathbf{J}_r \\ \mathbf{Q}(3) = \dfrac{1}{br}\mathbf{I}_a \otimes \mathbf{J}_b \otimes \bar{\mathbf{J}}_c \otimes \mathbf{J}_r \\ \mathbf{Q}(2,3) = \dfrac{1}{r}\mathbf{I}_a \otimes \bar{\mathbf{J}}_b \otimes \bar{\mathbf{J}}_c \otimes \mathbf{J}_r \end{cases}, \quad (44)$$

jointly with the

$$\mathbf{Q}^\perp = \mathbf{I}_n - \mathbf{Q}(0) - \mathbf{Q}(1) - \mathbf{Q}(2) - \mathbf{Q}(3) - \mathbf{Q}(2,3), \quad (45)$$

constitute the principal basis of the Jordan algebra associated with the model, and that

$$\mathbf{V} = \boldsymbol{\Sigma}(\mathbf{Y}^n) = \gamma(1)\mathbf{Q}(1) + \gamma(2)\mathbf{Q}(2) + \gamma(3)\mathbf{Q}(3) + \gamma(2,3)\mathbf{Q}(2,3) + \sigma^2\mathbf{Q}^\perp, \quad (46)$$

with

$$\begin{cases} \gamma(1) = \sigma^2 + r(\sigma^2(2,3) + b\sigma^2(3) + c\sigma^2(2) + bc\sigma^2(1)) \\ \gamma(2) = \sigma^2 + r(\sigma^2(2,3) + c\sigma^2(2)) \\ \gamma(3) = \sigma^2 + r(\sigma^2(2,3) + b\sigma^2(3)) \\ \gamma(2,3) = \sigma^2 + r\sigma^2(2,3) \end{cases}, \quad (47)$$

so that

$$\begin{cases} \sigma^2(1) = \dfrac{1}{rbc}(\gamma(1) - \gamma(2) - \gamma(3) + \gamma(2,3)) \\ \sigma^2(2) = \dfrac{1}{rc}(\gamma(2) - \gamma(2,3)) \\ \sigma^2(3) = \dfrac{1}{rb}(\gamma(3) - \gamma(2,3)) \\ \sigma^2(2,3) = \dfrac{1}{r}(\gamma(2,3) - \sigma^2) \end{cases} . \tag{48}$$

Since the ranks of the matrices in the principal basis are $g(1) = a - 1$, $g(2) = a(b-1)$, $g(3) = a(c-1)$, $g(2,3) = a(b-1)(c-1)$ and $g = abc(r-1)$, we get

$$\begin{cases} S(1) = \|\mathbf{Q}(1)\mathbf{Y}^n\|^2 \sim \gamma(1)\chi^2_{g(1)} \\ S(2) = \|\mathbf{Q}(2)\mathbf{Y}^n\|^2 \sim \gamma(2)\chi^2_{g(2)} \\ S(3) = \|\mathbf{Q}(3)\mathbf{Y}^n\|^2 \sim \gamma(3)\chi^2_{g(3)} \\ S(2,3) = \|\mathbf{Q}(2,3)\mathbf{Y}^n\|^2 \sim \gamma(2,3)\chi^2_{g(2,3)} \\ S = \|\mathbf{Q}^{\perp}\mathbf{Y}^n\|^2 \sim \sigma^2\chi^2_g \end{cases} , \tag{49}$$

thus it is straightforward to get UMVUE for the variance components. The models in this section may be trimmed when the nullity of some of the variance components is strongly suggested.

4 Pivot Variables

4.1 Variance Components

Taking $k = w + 1 - m$, let $\boldsymbol{\gamma}^k$ have components $\gamma_{m+1}, ..., \gamma_{w+1}$. In many cases the direct formulation $\boldsymbol{\gamma}^k$ will have to belong to a parameter space

$$\Omega = \{\boldsymbol{\gamma}^k : \mathbf{M}\boldsymbol{\gamma}^k \geq \mathbf{0}^d\}, \tag{50}$$

with $\mathbf{M} = [\mathbf{A}' \quad \mathbf{I}_k]'$. Then, besides $\boldsymbol{\gamma}^k \geq \mathbf{0}^k$, the direct formulation requires $\mathbf{A}\boldsymbol{\gamma}^k \geq \mathbf{0}^k$.

Let the s_i be the values taken by the $\|\mathbf{A}_i'\mathbf{Y}^n\|^2$, $i = 1, ..., w+1$, and consider

$$\theta = \sum_{i=m+1}^{w+1} b_i\gamma_i \tag{51}$$

be the parameter of interest. We point out that θ may be a variance component or its positive or negative part. To such parameters we have the UMVUE

$$\tilde{\theta} = \sum_{i=m+1}^{w+1} b_i\tilde{\gamma}_i. \tag{52}$$

These estimators may take negative values even when we require, in the direct formulation, that $\theta = 0$. This is a well known problem, there being two main positions (for instance, see [10], pg. 151):

1. consider that getting a negative estimator is a clean indicator of having θ, at least approximately, null;

2. use alternative estimation that ensures obtaining non-negative estimators. These techniques nonetheless have no small sample theoretical basis.

A third approach may rest in basing inference on the canonical formulation. An advantage we thus get in carrying out more robust inference since it is based in a less restrictive formulation of the models used. Now parameters θ will not necessarily be non-negative.

Since $S_i \sim \gamma_i \chi^2_{g_i}$, $i = m+1, ..., w+1$, we have the pivot variables $\frac{S_i}{\gamma_i} \sim \chi^2_{g_i}$, $i = m+1, ..., w+1$, and may induce probability measures for the γ_i, $i = m+1, ..., w+1$, using the

$$W_i = \frac{s_i}{W_i}, i = m+1, ..., w+1, \tag{53}$$

where $W_i \sim \chi^2_{g_i}$, $i = m+1, ..., w+1$. The probability generation would be carried out using

$$U = \sum_{i=1}^{w+1} b_i \frac{s_i}{W_i}. \tag{54}$$

The chi squares W_i are assumed to be independent. It is possible to generate sets $\{W_{m+1,l}, ..., W_{w+1,l}\}$ of independent chi-squares with $g_{m+1}, ..., g_{w+1}$ degrees of freedom, $l = 1, ..., \bar{n}$, and obtain the

$$U_l = \sum_{i=1}^{w+1} b_i \frac{s_i}{W_{i,l}}, l = 1, ...\bar{n}. \tag{55}$$

thus obtaining a sample of arbitrarily large size $\bar{n}$, where the observations have the distribution generated for θ.

If we want to assume that $\theta \geq 0$ we may screen the sample retaining only the sets $\{W_{m+1,l}, ..., W_{w,l}\}$ that originate non-negative U_l. Without screening the point density of the Z_i, $i = m+1, ..., w$ will be

$$f(\mathbf{z}^k) = \prod_{i=m+1}^{w+1} \frac{\left(\frac{z_i}{2}\right)^{\frac{g_i}{2}-1} e^{-\frac{z_i}{2}}}{2\Gamma\left(\frac{g_i}{2}\right)} \tag{56}$$

and, if we consider the restriction, for the filtered sample we get the density

$$f^*(\mathbf{z}^k) = \frac{f(\mathbf{z}^k)}{\int_\Omega f(\mathbf{u}^k) \prod_{i=m+1}^{w+1}}, \mathbf{z}^k \in \Omega. \tag{57}$$

Apply the transformation $z_{m+1} = u - \sum_{i=m+2}^{w+1} b_i z_i$, $z_i = z_i$, $i = m+2, ..., w+1$, with jacobian $\frac{1}{b_{m+1}}$ and integrate in order to z_i, $i = m+2, ..., w+1$. Thus, the conditions in proposition 3 will hold for the density of U whether we assume or not that $\theta \geq 0$. Then with $u_{\bar{n},1-\alpha}$ the empirical quantile for probability $1-\alpha$, we will have

$$\mathbb{P}[\theta \leq u_{\bar{n},1-\alpha}] \xrightarrow[\bar{n}\to\infty]{a.s.} 1-\alpha. \tag{58}$$

In this way we obtain limit level $1-\alpha$ confidence intervals $]-\infty; u_{\bar{n},1-\alpha}]$, $[u_{\bar{n},\frac{\alpha}{2}}; u_{\bar{n},1-\frac{\alpha}{2}}]$, $[u_{\bar{n},\alpha}; +\infty[$ for θ. If we assume $\theta \geq 0$ the first of these must be rewritten as $[0; u_{\bar{n},1-\alpha}]$, and screened samples must be used. These intervals can be used to obtain, through duality, limit level α right-sided, two-sided and left-sided tests for

$$H_0(\theta_0) : \theta = \theta_0. \tag{59}$$

These tests reject $H_0(\theta_0)$ when θ_0 is not covered by the corresponding confidence intervals.

It may be of interest to connect our previous discussion with the notion of generalized test variable (see [28], pg. 115). To avoid repetitions we represent the parameter of interest by θ and the tested hypothesis by $H_0(\theta_0)$, keeping the notation we have introduced. Now, given the sufficient statistic vector $\mathbf{S}^k$ which takes values $\mathbf{s}^k$ and the nuissance parameter $\boldsymbol{\gamma}^k$, $g(\mathbf{S}^k, \mathbf{s}^k\boldsymbol{\gamma}^k, \theta)$ is a generalized test variable if:

1. $g(\mathbf{s}^k, \mathbf{s}^k\boldsymbol{\gamma}^k, \theta)$ does not depend on unknown parameters;
2. the distribution of $g(\mathbf{S}^k, \mathbf{s}^k\boldsymbol{\gamma}^k, \theta)$, given θ, is known;
3. $g(\mathbf{S}^k, \mathbf{s}^k\boldsymbol{\gamma}^k, \theta)$ is stochastically non-decreasing with θ.

Taking

$$g(\mathbf{S}^k, \mathbf{s}^k\boldsymbol{\gamma}^k, \theta) = \theta - \sum_{i=m+1}^{w+1} b_i\gamma_i\frac{s_i}{S_i} \tag{60}$$

we get

$$\begin{cases} g(\mathbf{s}^k, \mathbf{s}^k\boldsymbol{\gamma}^k, \theta) = 0 \\ g(\mathbf{S}^k, \mathbf{s}^k\boldsymbol{\gamma}^k, \theta) = \theta - \sum_{i=m+1}^{w+1} b_i\frac{s_i}{W_i} \end{cases}, \tag{61}$$

with $W_i \sim \chi^2_{g_i}$, $i = m+1, ...w+1$. It is now easy to see that $g(\mathbf{S}^k, \mathbf{s}^k\boldsymbol{\gamma}^k, \theta)$ is a generalized test variable well connected with our previous discussion since

$$g(\mathbf{S}^k, \mathbf{s}^k\boldsymbol{\gamma}^k, \theta) = \theta - U. \tag{62}$$

Since $g(\mathbf{S}^k, \mathbf{s}^k\boldsymbol{\gamma}^k, \theta)$ increases stochastically with θ it will be used to test $H_0(\theta_0)$ against

$$H_1(\theta_0) : \theta > \theta_0 \tag{63}$$

Let $\mathbb{P}[\theta_0]$ be the generalized P value. Then

$$\mathbb{P}[\theta_0] = \mathbb{P}\left[\sum_{i=m+1}^{d} b_i\frac{s_i}{W_i} \le \theta - \sum_{i=d+1}^{w+1} b_i\frac{s_i}{W_i}\right] \tag{64}$$

if $b_i > 0$, $i = m+1, ...d$, and $b_i < 0$, $i = d+1, ...w+1$.

If we want to test $H_0(\theta_0)$ against

$$H_1(\theta_0) : \theta < \theta_0 \tag{65}$$

we rewrite the generalized test variable as

$$g(\mathbf{S}^k, \mathbf{s}^k\boldsymbol{\gamma}^k, \theta) = \sum_{i=m+1}^{w+1} b_i\gamma_i\frac{s_i}{S_i} - \theta = \sum_{i=m+1}^{w+1} b_i\frac{s_i}{W_i} - \theta. \tag{66}$$

The generalized p-value is now

$$\mathbb{P}[\theta_0] = \mathbb{P}\left[\sum_{i=m+1}^{d} b_i\frac{s_i}{W_i} \ge \theta - \sum_{i=d+1}^{w} b_i\frac{s_i}{W_i}\right]. \tag{67}$$

Thus, generalized test variables and the corresponding generalized p-values enabled us to better understand one sided testing.

4.2 Fixed Effects

With $\mathrm{rank}(\mathbf{B}_i) = a_i$ we take $\boldsymbol{\psi}_i^{a_i} = \mathbf{B}_i \boldsymbol{\eta}_i^{g_i}$ and $\tilde{\boldsymbol{\psi}}_i^{a_i} = \mathbf{B}_i \tilde{\boldsymbol{\eta}}_i^{g_i}$ so that $\tilde{\boldsymbol{\psi}}_i^{a_i} \sim \mathcal{N}(\boldsymbol{\psi}_i^{a_i}, \gamma_i \mathbf{B}_i \mathbf{B}_i')$ with $\mathrm{rank}(\mathbf{B}_i \mathbf{B}_i') = a_i$ and $\gamma_i = \sum_{j=1}^{w} b_{i,j}\gamma_j$, $i = 1, ..., m$. Thus (see [14], pg. 52)

$$(\boldsymbol{\psi}_i^{a_i} - \tilde{\boldsymbol{\psi}}_i^{a_i})'(\mathbf{B}_i \mathbf{B}_i')^{-1}(\boldsymbol{\psi}_i^{a_i} - \tilde{\boldsymbol{\psi}}_i^{a_i}) \sim \gamma_i \chi^2_{a_i}, i = 1, ..., m. \tag{68}$$

On the right side of the last expression we have the product by γ_i of a central chi-square with a_i degrees of freedom. Let U_l, $l = 1, ..., \bar{n}$ be the elements in the sample generated according to the probability measure induced by $\tilde{\gamma}_i$, $D_l = U_l W$ with $W \sim \chi^2_{a_i}$, $i = 1, ..., m$. Let $d_{\bar{n},1-\alpha}$ be the empirical quantile for probability $1 - \alpha$ of this last sample we will have

$$\mathbb{P}\big[(\boldsymbol{\psi}_i^{a_i} - \tilde{\boldsymbol{\psi}}_i^{a_i})'(\mathbf{B}_i \mathbf{B}_i')^{-1}(\boldsymbol{\psi}_i^{a_i} - \tilde{\boldsymbol{\psi}}_i^{a_i}) \leq d_{\bar{n},1-\alpha}\big] \xrightarrow[\bar{n}\to\infty]{} 1 - \alpha, \tag{69}$$

thus getting a limit level confidence $1 - \alpha$ ellipsoid for $\boldsymbol{\psi}_i^{a_i}$. Through duality we will also get a limit level α test for

$$H_{0,i}(\boldsymbol{\psi}_{0,i}^{a_i}) : \boldsymbol{\psi}_i^{a_i} = \boldsymbol{\psi}_{0,i}^{a_i}. \tag{70}$$

This hypothesis is rejected when and only when $\boldsymbol{\psi}_{0,i}^{a_i}$ is not covered by the ellipsoid. Moreover (see [18], pg. 406 to 411) we will also have

$$\mathbb{P}\left[\bigcap_{\mathbf{c}^{a_i} \in \mathbb{R}^{a_i}} \left\{ |\mathbf{c}^{a_i\prime}\boldsymbol{\psi}_i^{a_i} - \mathbf{c}^{a_i\prime}\tilde{\boldsymbol{\psi}}_i^{a_i}| \leq \sqrt{d_{\bar{n},1-\alpha}\mathbf{c}^{a_i\prime}\mathbf{B}_i\mathbf{B}_i'\mathbf{c}^{a_i}} \right\}\right] \xrightarrow[\bar{n}\to\infty]{} 1 - \alpha. \tag{71}$$

In this way we get simultaneous confidence intervals for $\mathbf{c}^{a_i\prime}\boldsymbol{\psi}_i^{a_i}$, $i = 1, ..., m$. This last expression may be considered as a limit version of the well known Scheffé theorem.

The tests considered in this and the previous section were obtained through duality from UMVUE. To see that the optimality of the estimators ensures good behavior of the resulting tests we establish

Proposition 4. *If, for every level, the confidence regions obtained from a set of estimators are contained in those obtained from another set of estimators, the tests derived, through duality, from the first set of estimators have, for all levels, at least the power of those derived from the second set.*

Proof. It suffices to point out that whenever a test in the second set rejects an hypothesis the test, of the same size, in the first set also rejects. □

As for the tests considered in this section in defining $H_{0,i}(\boldsymbol{\psi}_{0,i}^{a_i})$, we must take $\boldsymbol{\psi}_{0,i}^{a_i} \in \mathrm{R}(\mathbf{B}_i)$, otherwise the test hypothesis may never hold, $i = 1, ..., m$. We then have $\boldsymbol{\psi}_{0,i}^{a_i} = \mathbf{B}_i \boldsymbol{\eta}_{0,i}^{a_i}$ and, with Δ_i the kernel of $\mathbf{B}_i$, $H_{0,i}(\boldsymbol{\psi}_{0,i}^{a_i})$ holds if and only if $\boldsymbol{\eta}_i^{a_i} - \boldsymbol{\eta}_{0,i}^{a_i} \in \Delta_i$, *i.e.*, when $\boldsymbol{\eta}_i^{a_i}$ and $\boldsymbol{\eta}_{0,i}^{a_i}$ belong to the same congruence class in the quotient vector space $\mathbb{R}^{a_i}/_{\Delta_i}$, $i = 1, ..., m$. A similar formulation of tested hypothesis for fixed effects models is given in [13].

Lastly we point out that the application of the results in these two sections, when direct model formulation is used, is also quite straightforward, and that right-sided tests for the nullity of variance components can be used as a basis for model trimming in the integrated approach.

References

[1] Covas R. (2003). *Inferência Semi-Bayesiana e Modelos de Componentes de Variância*. Master's Degree Thesis in Statistics and Optimization. FCT/UNL.

[2] Fonseca M., Mexia J.T., Zmyślony R. (2002). Exact Distributions for the Generalized F Statistic. *Discussiones Mathematicae – Probability and Statistics* **22**, 37–51

[3] Fonseca M., Mexia J.T., Zmyślony R. (2003). Estimators and Tests for Variance Components in Cross Nested Orthogonal Models. *Discussiones Mathematicae – Probability and Statistics* **23**, 173–201

[4] Fonseca M., Mexia J.T., Zmyślony R. (2003). Estimating and Testing of Variance Components: an Application to a Grapevine Experiment. *Biometrical Letters* **40**, 1–7

[5] Garcia J. (2002). *Laplace and Fourier Transforms in Risk Theory*. Ph.D. Thesis. Lisbon Technical Univ.

[6] Jordan P., von Neumann J., Wigner E.P. (1934). On an algebraic generalization of the quantum mechanical formalism. *Ann. Math. II. Ser.* **35**, 29–64

[7] Kallenberg O. (1997). *Foundations of Modern Probability*. Springer

[8] Khuri A.I., Mathew T., Sinha B.K. (1997). *Statistical Tests for Mixed Linear Models*. Wiley

[9] Krishnamoorthy K., Mathew T. (2004). One-Sided Tolerance Limits in Balanced and Unbalanced One-Way Random Models based in Generalized Confidence Intervals. To be published in *JASA*

[10] Lehmann E.L., Casella G. (1998). *Theory of point Estimation. 2nd. Ed.*. Springer

[11] Loéve M. (1960). *Probability Theory. 2nd. Ed.*. D. Van Nostrand

[12] Malley J.D. (1994). *Statistical applications of Jordan algebras*. Springer-Verlag

[13] Mexia J.T. (1989). Controlled Heteroscedasticity, Quotient Vector Spaces and F Tests for Hypothesis on Mean Vectors. *Trabalhos de Investigação* **1**. FCT/UNL

[14] Mexia J.T. (1995). *Introdução à Inferência Estatística Linear*. Ed. Lusófonas

[15] Michalski A., Zmyślony R. (1996). Testing Hypothesis for Variance Components in Mixed Linear Models. *Statistics* **27**, 297–310

[16] Michalski A., Zmyślony R. (1999). Testing Hypothesis for Linear Functions of Parameters in Mixed Linear Models. *Tatra Mt. Math. Publ.* **17**, 103–110

[17] Nunes C. (2004). *F Tests and Selective F Tests in Orthogonal Mixed Linear Models*. Ph.D. Thesis. Beira Interior Univ.

[18] Scheffé H. (1959). *The Analysis of Variance*. John Wiley & Sons

[19] Searle R.S., Casella G., McCullock C.E. (1992). *Variance Components*. John Wiley & Sons

[20] Seber G.A.F. (1980). *The Linear Hypothesis: a General Theory. 2nd. Ed.*. Charles Grifith & Co.

[21] Seely J. (1970). Linear Spaces and Unbiased Estimation. *Ann. Math. Stat.* **41**, 1725–1734

[22] Seely J. (1971). Quadratic Subspaces and Completeness. *Ann. Math. Stat.* **42**, 710–721

[23] Seely J., Zyskind G. (1971). Linear Spaces and Minimum Variance Unbiased Estimation. *Ann. Math. Stat.* **42**, 691–703

[24] Seely J., Zyskind G. (1971). Linear Spaces and Minimum Variance Unbiased Estimation. *Ann. Math. Stat.* **42**, 691–703

[25] Severini T.A. (2000). *Likelihood Methods in Statistics*. Oxford University Press

[26] Silvey S.D. (1975). *Statistical Inference. Reprinted*. Chapman & Hall

[27] Weerahandi S. (1993). Generalized Confidence Intervals. *J. Am. Stat. Assoc.* **88**, 899–905

[28] Weerahandi S. (1996). *Exact Statistical Methods for Data Analysis. 2nd. Print*. Springer-Verlag

[29] Williams D. (1991). *Probability with Martingales*. Cambridge Univ. Press

[30] Witkovský V. (2001). Computing the distribution of a linear combination of inverted gamma variables. *Kybernetika* **37**, 79–90

[31] Zhou, L., Mathew T. (1994). Some tests for variance components using generalized p values. *Technometrics* **36**, 394–402