Discussiones Mathematicae
Probability and Statistics 35 (2015) 29–44
doi:10.7151/dmps.1176

ORTHOGONAL MODELS: ALGEBRAIC STRUCTURE AND EXPLICIT ESTIMATORS FOR ESTIMABLE VECTORS

Artur Pereira, Miguel Fonseca

and

João Tiago Mexia

Center of Mathematics and Applications
Faculty of Sciences and Technology, NOVA University of Lisbon
Portugal

e-mail: amc.pereira@campus.fct.unl.pt
fmig@fct.unl.pt
jtm@fct.unl.pt

Abstract

We study the algebraic structure of orthogonal models thus of mixed models whose variance covariance matrices are all positive semi definite, linear combinations of known pairwise orthogonal projection matrices, POOPM, and whose least square estimators, LSE, of estimable vectors are best linear unbiased estimator, BLUE, whatever the variance components, so they are uniformly BLUE, UBLUE. From the results of the algebraic structure we will get explicit expression for the LSE of these models.

Keywords: linear models, mixed models, inference, orthogonal models, UBLUE.

2010 Mathematics Subject Classification: 62J10.

1. Introduction

Nelder [12] introduced the important class of models with orthogonal block structure, OBS, which continues to play an important role in the analysis of randomized block designer, see Calinski and Kageyama [2]. There are many interesting papers on OBS such as Houtman and Speed [8] and Mejza [10]. A mixed model has OBS if its variance covariance matrices are all positive semi definite linear combinations, so, it is written of following form:

$$\sum_{j=1}^{m} \boldsymbol{\gamma}_j \mathbf{K}_j$$

where $\mathbf{K}_1, \ldots, \mathbf{K}_m$ are known pairwise orthogonal orthogonal projection matrices, POOPM, that add up to $\mathbf{I}_n$, thus,

$$\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_m) \in \mathbb{R}_+^m,$$

with ∇_+ the family of vectors of subspace ∇ with non negative components. We are interested in mixed models

$$\mathbf{Y} = \sum_{i=0}^{w} \mathbf{X}_i \boldsymbol{\beta}_i$$

where $\boldsymbol{\beta}_\circ$ is fixed and the $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_w$ have null mean vectors as well as null cross covariance matrices and variance covariance matrices $\boldsymbol{\theta}_1 \mathbf{I}_{c_1}, \ldots, \boldsymbol{\theta}_w \mathbf{I}_{c_w}$, that have:

1. OBS;
2. LSE, for estimable vectors that are UBLUE, this is, are BLUE whatever the variance components.

Following, Vanleeuwen *et al.* [16], we say that these models are orthogonal, ORT.

In the next section we study the algebraic structure of these models.

Moreover, when normality holds, models with Normal OBS, NOBS, have, as we shall see, complete and sufficient statistics from which UMVUE are obtained not only for estimable vectors but also for variance components.

2. Algebraic structure

2.1. Commutative Jordan Algebras, CJA

In the study of the mixed models, we will use Commutative Jordan Algebras of symmetric matrices, CJA, these are linear subspaces constituted by symmetric matrices that commute and containing the squares of its matrices. Each CJA $\mathcal{A}$ has an unique basis, the principal basis $pb(\mathcal{A})$, constituted by pairwise orthogonal orthogonal projection matrices, POOPM. It is easy to show that orthogonal projection matrices of $\mathcal{A}$ are sums of matrices of $pb(\mathcal{A})$ and when they have rank 1, they belong to $pb(\mathcal{A})$.

If $\frac{1}{n}\mathbf{J}_n$ belong to $pb(\mathcal{A})$, $\mathcal{A}$ will be regular and if the matrices of $pb(\mathcal{A})$ add to $\mathbf{I}_n$, $\mathcal{A}$ will be complete.

The symmetric matrices $\mathbf{M}_1, \ldots, \mathbf{M}_w$ commutes, see Schott [14], if and only if they are diagonalized by the same orthogonal matrix $\mathbf{P}$, then they belong to $\mathcal{A}(\mathbf{P})$ the CJA of matrices diagonalized by $\mathbf{P}$, thus $\mathbf{M} = \{\mathbf{M}_1, \ldots, \mathbf{M}_w\}$ is contained in a CJA if and only if its matrices commutes.

Moreover the intersection of CJA gives CJA, so if the matrices of $\mathbf{M}$ commute, intersecting all CJA that contains $\mathbf{M}$, we obtain the smallest CJA, $\mathcal{A}(\mathbf{M})$, that contains $\mathbf{M}$. This CJA will be generated by $\mathbf{M}$.

2.2. Orthogonal Block Structure, OBS

The mixed models we presented in the introduction will have following mean vector

$$\boldsymbol{\mu} = \mathbf{X}_\circ \boldsymbol{\beta}_\circ$$

and variance covariance matrix

$$V(\boldsymbol{\theta}) = \sum_{i=1}^{w} \boldsymbol{\theta}_i \mathbf{M}_i,$$

where the

$$\mathbf{M}_i = \mathbf{X}_i \mathbf{X}_i^T.$$

commute. They generate a CJA with principal basis $\mathbf{K} = \{\mathbf{K}_1, \ldots, \mathbf{K}_m\}$.

Assuming that $\boldsymbol{\theta}_i \geq 0$, $i = \{1, \ldots, w\}$, and that the model has OBS, we get

$$\mathbf{M}_i \in \mathcal{A} = \left\{ \sum_{j=1}^{m} \gamma_j \mathbf{K}_j;\ \boldsymbol{\gamma} \in \mathbb{R}_+^m \right\}.$$

Now $\mathcal{A}$ is a linear space constituted by symmetric matrices that, given the $\mathbf{K}_1, \ldots, \mathbf{K}_m$ are POOPM, commute. The squares of the matrices of $\mathcal{A}$ will, since the $\mathbf{K}_1, \ldots, \mathbf{K}_m$ are POOPM, belong to $\mathcal{A}$. So $\mathcal{A}$ will be a commutative Jordan algebra, CJA. Now, see Seely [15], every CJA $\mathcal{A}^\circ$ has an unique basis, this is, the principal basis $pb(\mathcal{A}^\circ)$, constituted by POOPM. It is easy to see that

$$pb(\mathcal{A}) = \mathbf{K} = \{\mathbf{K}_1, \ldots, \mathbf{K}_m\}.$$

We say that the mixed model that we are considering previously, with OBS and

$$\mathbf{K} = pb(\mathcal{A}), \text{ is } OBS(\mathcal{A}).$$

Since,

$$\mathbf{M}_i \in \mathcal{A},\ i = \{1, \ldots, w\},$$

we also have

$$\mathbf{M}_i = \sum_{j=1}^{m} b_{ij}\mathbf{K}_j,\ i = \{1,\dots,w\}$$

and so

$$V(\boldsymbol{\theta}) = \sum_{i=1}^{w} \boldsymbol{\theta}_i \sum_{j=1}^{m} b_{ij}\mathbf{K}_j = \sum_{j=1}^{m} \boldsymbol{\gamma}_j \mathbf{K}_j = V(\gamma)$$

with

$$\boldsymbol{\gamma}_j = \sum_{i=1}^{w} b_{ij}\boldsymbol{\theta},\ j = \{1,\dots,m\}.$$

The b_{ij}, $i = \{1,\dots,w\}$, $j = \{1,\dots,m\}$, are non negative, and with $\boldsymbol{\delta}_i$ the vector

$$\boldsymbol{\delta}_i = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix},$$

the $\mathbf{M}_i = V(\boldsymbol{\delta})$, $i = \{1,\dots,w\}$ are positive semi definite and the $w \times m$ matrices $\mathbf{B} = [b_{ij}]$ will be the transition matrix. For $V(\boldsymbol{\theta}_1) = V(\boldsymbol{\theta}_2)$, implying $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_2$, the matrices $\mathbf{M}_1,\dots,\mathbf{M}_w$ must be linearly independent, as shown in the following proposition

Proposition 1. *We have*

$$[(V(\theta_1) = V(\theta_2) \Longrightarrow (\theta_1 = \theta_2)] \Longleftrightarrow \boldsymbol{M}_1,\dots,\boldsymbol{M}_w \text{ are linearly independent.}$$

Proof.

1. We will prove that

$$[(V(\theta_1) = V(\theta_2) \Longrightarrow (\theta_1 = \theta_2)] \Longrightarrow \mathbf{M}_1,\dots,\mathbf{M}_w \text{ are linearly independent.}$$

   Let $V(\theta(l)) = \sum\limits_{i=1}^{w} \theta_i(l)\mathbf{M}_i$, $l = 1, 2$ be, then

$$\begin{aligned} V(\theta_1) = V(\theta_2) \Longleftrightarrow 0 &= V(\theta_2) - V(\theta_1) \\ &= \sum_{i=1}^{w} \theta_i(2)\mathbf{M}_i - \sum_{i=1}^{w} \theta_i(1)\mathbf{M}_i \\ &= \sum_{i=1}^{w} (\theta_i(2) - \theta_i(1))\mathbf{M}_i. \end{aligned} \tag{1}$$

If this equality implies $\theta_i(2) = \theta_i(1)$; $i = 1, \ldots, w$ then $\mathbf{M}_1, \ldots, \mathbf{M}_w$ will are linearly independent.

2. We will establish the reverse.

   If the $\mathbf{M}_1, \ldots, \mathbf{M}_w$ are linearly independent, then

$$V(\theta_2) = V(\theta_1),$$

   implies that

$$\theta_1 = \theta_2.$$

■

Which is equivalent to the linear independence of the row vectors of $\mathbf{B}$ and implies $w \leq m$. Now we see that for $\boldsymbol{\gamma} \in \mathbb{R}_+^m$, matrix $\mathbf{B}$ must be invertible, thus we will have $w = m$ and $\mathbf{M} = \{\mathbf{M}_1, \ldots, \mathbf{M}_m\}$ will be a basis for $\mathcal{A}$, $\mathbf{M} = b(\mathcal{A})$. We now have the following proposition:

Proposition 2. *A mixed model*

$$\boldsymbol{Y} = \sum_{i=0}^{w} \boldsymbol{X}_i \boldsymbol{\beta}_i$$

is OBS $(\mathcal{A})$ *if and only if* $\boldsymbol{M} = b(\mathcal{A})$ (*which implies* $w = m$).

Proof. We already showed that, if the model is $OBS(\mathcal{A})$, $\mathbf{M} = b(\mathcal{A})$. Now, when $\mathbf{M} = b(\mathcal{A})$, $w = m$ and the $V(\boldsymbol{\theta}) = \sum_{i=1}^{m} \boldsymbol{\theta}_i \mathbf{M}_i$ will be all the positive semi definite matrices belonging to $\mathcal{A}$, since $V(\boldsymbol{\theta}) = V(\boldsymbol{\gamma})$ with $\boldsymbol{\gamma} = \mathbf{B}^T \boldsymbol{\theta} \in \mathbb{R}(\mathbf{B}^T)$. ■

2.3. Commutative Orthogonal Block Structure, COBS

Now, see Zmyślony [17], a mixed model with:

1. $\mathbf{T}$ the Orthogonal projection matrix, OPM, on the space spanned by the mean vector;

2. The family $\boldsymbol{\nu}$ of the variance covariance matrices has LSE that are UBLUE if and only if

$$\mathbf{T}\boldsymbol{\nu} = \boldsymbol{\nu}\mathbf{T},$$

   this is, if $\mathbf{T}$ commutes with the variance covariance matrices.

Now if

$$\mathbf{T}\mathbf{K}_j = \mathbf{K}_j\mathbf{T}, \; j = \{1, \ldots, m\},$$

following Fonseca *et al.* [7], we say that these models have commutative orthogonal block structure, COBS, associated with $\mathcal{A}$, $COBS(\mathcal{A})$.

Combining proposition (2) with the Zmyślony theorem, we get

Proposition 3. The orthogonal models are the $COBS(\mathcal{A})$.

This characterization of orthogonal models has the advantage of establishing an equivalence relation between these models writing

$$\mathbf{M}_1 = \mathbf{M}_2$$

when both $\mathbf{M}_1$ and $\mathbf{M}_2$ are $COBS(\mathcal{A})$. Inside one such equivalence class, the models are identified by the orthogonal projection matrices on the spaces spanned by their mean vectors or, equivalently, by these spaces. ■

We now point out that if the g_j row vectors of $\mathbf{A}_j$ constitute an orthonormal basis for the range space $R(\mathbf{K}_j)$ of $\mathbf{K}_j$, $j = 1, \ldots, m$, we have

$$\begin{cases} \mathbf{K}_j = \mathbf{A}_j^T \mathbf{A}_j, \; j = \{1, \ldots, m\} \\ \mathbf{I}_{g_j} = \mathbf{A}_j \mathbf{A}_j^T, \; j = \{1, \ldots, m\} \end{cases},$$

and

$$\mathbf{Y}_j = \mathbf{A}_j \mathbf{Y}, \; j = \{1, \ldots, m\}$$

will have mean vector

$$\boldsymbol{\mu}_j = \mathbf{A}_j = \mathbf{X}_{0j} \boldsymbol{\beta}_\circ$$

with

$$\mathbf{X}_{0j} = \mathbf{A}_j \mathbf{X}_\circ$$

and variance covariance matrix

$$\boldsymbol{\gamma}_j \mathbf{I}_{g_j}, \; j = \{1, \ldots, m\}.$$

These sub-vectors have null cross-covariance matrices so, when normality is assumed, they are independent.

Let $\mathbf{P}_j$ be the orthogonal projection matrix, OPM, on

$$\Omega_j = R(\mathbf{X}_{0j}),$$

with $R(\mathbf{X}_{0j})$, the range space of matrix $\mathbf{X}_{0j}$, and put

$$r_j = rank(\mathbf{P}_j), \; j = \{1, \ldots, m\}.$$

Since $\mathbf{T}$ and $\mathbf{K}_j$ commute, $\mathbf{K}_j^\circ = \mathbf{T}\mathbf{K}_j = \mathbf{K}_j\mathbf{T}$ will be symmetric and idempotent, so they are OPM. We point out that when this product matrix is null it may be considered as the orthogonal projection matrix, OPM, on the subspace $\{\mathbf{0}\}$ whose only vector is $\mathbf{0}$. Given that

$$R(\mathbf{K}_j^{\circ}) \subset R(\mathbf{K}_j),\ j = \{1, \dots, m\},$$

the $\mathbf{K}_j^{\circ}$ will be pairwise orthogonal orthogonal projection matrices, POOPM.

Reordering, if necessary, assume that the $\mathbf{K}_1^{\circ}, \dots, \mathbf{K}_z^{\circ}$ are the non null product matrices, $z \leq m$. Now

$$\mathbf{T} = \mathbf{T}\sum_{j=1}^{n}\mathbf{K}_j = \sum_{j=1}^{m}\mathbf{K}_j^{\circ} = \sum_{j=1}^{z}\mathbf{K}_j^{\circ},$$

so that with $\Omega = R(\mathbf{K})$ and $r_j^{\circ} = R(\mathbf{K}_j^{\circ})$, $j = \{1, \dots, m\}$, we have

$$\Omega = \boxplus_{j=1}^{z}\Omega_j^{\circ}$$

with $\boxplus$ indicating orthogonal direct sum of subspaces. We now establish

Lemma 1. *In models with COBS, we have*

$$rank(\boldsymbol{P}_j) = rank(\boldsymbol{A}_j^T\boldsymbol{P}_j) = rank(\boldsymbol{A}_j^T\boldsymbol{P}_j\boldsymbol{A}_j),\ j = \{1, \dots, m\}.$$

Proof. We have

$$\begin{aligned} rank(\mathbf{A}_j^T\mathbf{P}_j\mathbf{A}_j) \leq rank(\mathbf{A}_j^T\mathbf{P}_j) \leq rank(\mathbf{P}_j) &= rank(\mathbf{A}_j\mathbf{A}_j^T\mathbf{P}_j\mathbf{A}_j\mathbf{A}_j^T) \\ &\leq rank(\mathbf{A}_j^T\mathbf{P}_j\mathbf{A}_j),\ j = \{1, \dots, m\}, \end{aligned}$$

which establishes the thesis. ∎

Since the $\mathbf{K}_1, \dots, \mathbf{K}_m$ add up to $\mathbf{I}_n$, we can also establish

Lemma 2. *In models with COBS, we have*

$$\boldsymbol{K}_j^{\circ} = \boldsymbol{K}_j^{\circ\circ}$$

with $\boldsymbol{K}_j^{\circ\circ} = \boldsymbol{A}_j^T\boldsymbol{P}_j\boldsymbol{A}_j$, $j = 1, \dots, z$.

Proof. The $\mathbf{K}_j^{\circ\circ}$ are symmetric and idempotent, so they are OPM. To establish the thesis it is sufficient to show that $R(\mathbf{K}_j^{\circ}) = R(\mathbf{K}_j^{\circ\circ})$, $j = \{1, \dots, z\}$.

Now

$$
\begin{aligned}
R(\mathbf{K}_j^\circ) &= R(\mathbf{K}_j\mathbf{T}) = \mathbf{K}_j R(\mathbf{T}) = \mathbf{A}_j^T\mathbf{A}_j R(\mathbf{X}_\circ) \\
&= \mathbf{A}_j^T R(\mathbf{A}_j\mathbf{X}_\circ) = \mathbf{A}_j^T R(\mathbf{X}_{\circ j}) \\
&= \mathbf{A}_j^T R(\mathbf{P}_j) = R(\mathbf{A}_j^T\mathbf{P}_j) \\
&= R((\mathbf{A}_j^T\mathbf{P}_j)(\mathbf{A}_j^T\mathbf{P}_j)^T) \\
&= R(\mathbf{K}_j^{\circ\circ}),\ j = \{1,\dots,z\}.
\end{aligned} \tag{2}
$$

■

Corollary 1. *In models with COBS,*

$$\boldsymbol{T} = \sum_{j=1}^{z} \boldsymbol{K}_j^{\circ\circ}.$$

Proof. $\mathbf{T} = \sum_{j=1}^z \mathbf{K}_j\mathbf{T} = \sum_{j=1}^z \mathbf{K}_j^\circ = \sum_{j=1}^z \mathbf{K}_j^{\circ\circ}$. ■

Corollary 2. In models with COBS we have, with $k = rank(\mathbf{T})$,

$$k = \sum_{j=1}^{z} p_j.$$

Proof. The thesis follow from Lemma 1 and Corollary 1 of Lemma 2. ■

3. Least Squares Estimators, LSE

We now obtain explicit expressions LSE in models with COBS using the results of preceding section first, and then assuming normality. Namely we establish

Proposition 4. *In a model with COBS, the LSE for the estimable vector*

$$\boldsymbol{\Psi} = \boldsymbol{U}\boldsymbol{\mu}$$

is

$$\tilde{\boldsymbol{\Psi}} = \sum_{j=1}^{z} \boldsymbol{U}_j\boldsymbol{P}_j\boldsymbol{Y}_j$$

with

$$\boldsymbol{U}_j = \boldsymbol{U}\boldsymbol{A}_j^T,\ j = \{1,\dots,m\}$$

and

$$\boldsymbol{Y}_j = \boldsymbol{A}_j\boldsymbol{Y},\ j = \{1,\dots,m\}.$$

Proof. We have only to point out that the LSE for $\boldsymbol{\mu}$ and $\boldsymbol{\Psi}$ are

$$\tilde{\boldsymbol{\mu}} = \mathbf{T}\mathbf{Y}$$

and

$$\tilde{\boldsymbol{\Psi}} = \mathbf{U}\tilde{\boldsymbol{\mu}}$$

and we use the expressions of $\mathbf{T}$ given by Corollary 2 of Lemma 2. ■

We recall that models with COBS have LSE that are UBLUE. Thus, Proposition 4, gives explicit expression for UBLUE in models with COBS.

Let us now assume $\mathbf{Y}$ to be normal with mean vector $\boldsymbol{\mu}$ and variance covariance matrix $V(\boldsymbol{\gamma})$, and so, we write

$$\mathbf{Y} \sim N(\boldsymbol{\mu}, V(\boldsymbol{\gamma})).$$

Then the

$$\mathbf{Y}_j \sim N(\boldsymbol{\mu}_j, \boldsymbol{\gamma}_j \mathbf{I}_{g_j}),$$

with

$$\boldsymbol{\mu}_j = \mathbf{A}_j \boldsymbol{\mu},\ j = \{1, \dots, m\},$$

will be independent, since they have joint normal distribution and null cross covariance matrices.

Let the row vectors of $\mathbf{L}_j$ constituting an orthonormal basis for $R(\mathbf{P}_j)$ and put

$$\boldsymbol{\eta} = \mathbf{L}_j \boldsymbol{\mu}_j$$

as well as

$$\tilde{\boldsymbol{\eta}} = \mathbf{L}_j \mathbf{Y}_j$$

and as $s_j = ||\mathbf{Y}_j||^2$, thus with

$$\mathbf{P}_j^{\perp} = \mathbf{I}_{g_j} - \mathbf{P}_j,\ j = \{1, \dots, z\},$$

we have

$$\begin{aligned} ||\mathbf{Y}_j - \boldsymbol{\mu}_j||^2 &= (\mathbf{Y}_j - \boldsymbol{\mu}_j)^T \mathbf{L}_j^T \mathbf{L}_j (\mathbf{Y}_j - \boldsymbol{\mu}_j) + (\mathbf{Y}_j - \boldsymbol{\mu}_j) \mathbf{P}_j^{\perp} (\mathbf{Y}_j - \boldsymbol{\mu}_j) \\ &= \mathbf{Y}_j^T \mathbf{P}_j \mathbf{Y}_j - 2\boldsymbol{\eta}_j^T \tilde{\boldsymbol{\eta}}_j + \boldsymbol{\eta}_j^T \boldsymbol{\eta}_j + \mathbf{Y}_j^T \mathbf{P}_j^{\perp} \mathbf{Y}_j \\ &= s_j - 2\boldsymbol{\eta}_j^T \boldsymbol{\eta}_j + ||\boldsymbol{\eta}_j||^2,\ j = \{1, \dots, z\} \end{aligned} \tag{3}$$

and so the corresponding densities will be

$$n_j(\mathbf{Y}_j) = \frac{e^{-\dfrac{1}{2\boldsymbol{\gamma}_j}(s_j - 2\boldsymbol{\eta}_j^T\boldsymbol{\eta}_j + ||\boldsymbol{\eta}_j||^2)}}{(2\pi\boldsymbol{\gamma}_j)^{\dfrac{g_j}{2}}}.$$

If $z < m$ we will also have the densities

$$n_j(\mathbf{Y}_j) = \frac{e^{-\dfrac{s_j}{2\boldsymbol{\gamma}_j}}}{(2\pi\boldsymbol{\gamma}_j)^{\dfrac{g_j}{2}}},\ j = \{z+1, \ldots, m\},$$

the joint density of these vectors is

$$n(\mathbf{Y}_j)^{\circ} = \prod_{j=1}^{m} n_j(\mathbf{Y}_j)$$

will also be the density of

$$\mathbf{Y}^{\circ} = \left[\mathbf{Y}_1^T, \ldots, \mathbf{Y}_m^T\right]^T = \mathbf{A}^{\circ}\mathbf{Y}$$

with

$$\mathbf{A}^{\circ} = \left[\mathbf{A}_1^T, \ldots, \mathbf{A}_m^T\right]^T$$

an orthogonal matrix. Now the statistics derived from $\mathbf{Y}^{\circ}$ are the same as those derived from $\mathbf{Y}$, and

$$\mathbf{Y}_j = \left[\mathbf{O}_{g_j \times g_1}, \ldots, \mathbf{I}_{g_j}, \ldots, \mathbf{O}_{g_j \times g_m}\right]\mathbf{Y}^{\circ},\ j = \{1, \ldots, m\}$$

so we can work with $\mathbf{Y}^{\circ}$ instead of $\mathbf{Y}$. It is easy to see that

$$n(\mathbf{Y}) = \frac{e^{-\dfrac{1}{\boldsymbol{\xi}}\sum_{j=1}^{m}\vartheta_j s_j + \sum_{j=1}^{z}\boldsymbol{\xi}_j^T\tilde{\boldsymbol{\eta}}_j - \dfrac{1}{2}\sum_{j=1}^{z}\vartheta_j||\boldsymbol{\xi}||^2}}{\prod_{j=1}^{m}(2\pi\vartheta_j^{-1})^{\dfrac{g_j}{2}}}$$

with $\vartheta_j = \gamma_j^+$, this is, $\vartheta_j = \gamma_j^{-1}[0]$ when $\gamma_j > 0[= 0]$, $j = \{1, \ldots, m\}$ and

$$\boldsymbol{\xi}_j = \vartheta_j\boldsymbol{\eta}_j,\ j = \{1, \ldots, z\}.$$

Then this density will have the sufficient statistics s_j, $j = \{1, \ldots, m\}$, and $\tilde{\boldsymbol{\eta}}_j$, $j = \{1, \ldots, z\}$.

When we require the variance covariance matrices to be positive definite, we will have $\boldsymbol{\gamma} \in \mathbb{R}^m_{+>}$, with $\nabla_{+>}$ the family of vectors of subspace ∇ with positive components.

Now according to Corollary 2 of Lemma 2

$$\boldsymbol{\eta} = [\boldsymbol{\eta}_1^T, \ldots, \boldsymbol{\eta}_j^T]^T$$

spans $\mathbb{R}^k$. Thus we have the open parameters space $\mathbb{R}^k \times \mathbb{R}^m$, where $\times$ indicates Cartesian product.

Following Lehmann and Casela [9], the statistics s_j, $j = \{1, \ldots, m\}$, and $\tilde{\boldsymbol{\eta}}_j$, $j = \{1, \ldots, z\}$, are not only sufficient but complete.
Since

$$\mathbf{P}_j \mathbf{Y}_j = \mathbf{L}_j^T \tilde{\boldsymbol{\eta}}_j,\ j = \{1, \ldots, z\}$$

the expression for $\tilde{\Psi}$ given in Proposition 4 may be written as

$$\tilde{\Psi} = \sum_{j=1}^{z} \mathbf{U}_j \mathbf{L}_j^T \tilde{\boldsymbol{\eta}}_j.$$

Moreover if

$$p_j = rank(\mathbf{P}_j) < g_j$$

we have also the unbiased estimator

$$\tilde{\gamma}_j = \frac{s_j - ||\tilde{\boldsymbol{\eta}}_{\boldsymbol{j}_j}||^2}{g_j - p_j}$$

of $\boldsymbol{\gamma}_j$, since

$$||\tilde{\boldsymbol{\eta}}_j||^2 = \mathbf{Y}_j^T \mathbf{P}_j \mathbf{Y}_j = ||\mathbf{P}_j \mathbf{Y}_j||^2.$$

If $j > z$, we can rewrite this last estimator as

$$\tilde{\gamma}_j = \frac{s_j}{g_j}.$$

We then established

Proposition 5. *When normality holds, the* $\tilde{\boldsymbol{\Psi}}$ *and* $\tilde{\gamma}_j$, *with* $j \in \mathcal{D} = \{l : p_l < g_l\}$ *are uniformly minimum variance unbiased estimators, UMVUE.*

Proof. It suffices to point out that these estimators are derived from complete and sufficient statistics. ■

Now, we will established the followings corollary

Corollary 3. *If* $\mathcal{D} = \{1, \ldots, m\}$, *then*

$$\tilde{\boldsymbol{\theta}} = (\boldsymbol{B}^T)^{-1}\tilde{\gamma}$$

will be UMVUE.

Proof. Since

$$\boldsymbol{\gamma} = \mathbf{B}^T\boldsymbol{\theta},$$

we have

$$\boldsymbol{\theta} = (\mathbf{B}^T)^{-1}\boldsymbol{\gamma}$$

and the thesis follows from Proposition 5. ∎

Corollary 4. *If the sub-matrix* $\boldsymbol{B}(\mathcal{D})$ *of* $\boldsymbol{B}$ *constituted by the elements with both index in* $\mathcal{D}$ *is invertible,*

$$\tilde{\gamma}(\mathcal{D}) = (\tilde{\gamma}_j,\ \gamma \in \mathcal{D})$$

and

$$\tilde{\boldsymbol{\theta}}(\mathcal{D}) = [\boldsymbol{B}^T(\mathcal{D})]^{-1}\tilde{\gamma}(\mathcal{D})$$

are UMVUE.

The results in this section complete those in often works on COBS such as Areia and Carvalho [1], M. Fonseca *et al.* [7], Carvalho *et al.* [4] and Mexia *et. al* [11].

4. Uniformly Best Linear Unbiased Estimator, UBLUE

We now characterize models whose LSE are UBLUE. The estimable vectors of a model with mean vector

$$\boldsymbol{\mu} = \mathbf{X}_\circ\boldsymbol{\beta}_\circ$$

are the

$$\boldsymbol{\Psi} = \mathbf{U}\boldsymbol{\mu}$$

the corresponding linear unbiased estimators are the

$$\boldsymbol{\Psi}^* = \mathbf{AY}$$

with

$$\mathbf{A} \in [\boldsymbol{\Psi}] = \{\mathbf{A} : \xi(\mathbf{AY}) = \boldsymbol{\Psi}\}$$

where $\xi(.)$ is the mean vector. We now establish the following lemma:

Lemma 3. $\xi(\boldsymbol{A}_1, \boldsymbol{Y}) = \xi(\boldsymbol{A}_2, \boldsymbol{Y})$ *if and only if* $\boldsymbol{A}_1\boldsymbol{T} = \boldsymbol{A}_2\boldsymbol{T}$.

Proof. Since

$$\xi(\mathbf{A}_l\mathbf{Y}) = \mathbf{A}_l\boldsymbol{\mu} = \mathbf{A}_l\mathbf{T}\boldsymbol{\mu},\ l = 1, 2$$

the sufficient condition is established. Inversely, if

$$\xi(\mathbf{A}_1, \mathbf{Y}) = \xi(\mathbf{A}_2, \mathbf{Y})$$

we will have, whatever $\boldsymbol{\beta}_\circ$,

$$\mathbf{A}_1\mathbf{T}\mathbf{X}_\circ\boldsymbol{\beta}_\circ = \mathbf{A}_2\mathbf{T}\mathbf{X}_\circ\boldsymbol{\beta}_\circ$$

so that

$$\mathbf{A}_1\mathbf{T}\mathbf{X}_\circ = \mathbf{A}_2\mathbf{T}\mathbf{X}_\circ$$

and that

$$(\mathbf{A}_1\mathbf{T} - \mathbf{A}_2\mathbf{T})\mathbf{X}_\circ = \mathbf{0}$$

being $\mathbf{0}$ a null matrix.

Thus the row vectors of

$$\mathbf{W} = \mathbf{A}_1\mathbf{T} - \mathbf{A}_2\mathbf{T} = (\mathbf{A}_1 - \mathbf{A}_2)\mathbf{T}$$

have to be orthogonal to $\Omega = R(\mathbf{X}_\circ)$, but these vectors also belong to Ω, so they are null, which gives

$$\mathbf{A}_1\mathbf{T} - \mathbf{A}_2\mathbf{T} = \mathbf{0}$$

and so

$$\mathbf{A}_1\mathbf{T} = \mathbf{A}_2\mathbf{T}$$

as we wanted to establish. ■

Now, the LSE for

$$\boldsymbol{\Psi} = \mathbf{U}\boldsymbol{\mu}$$

is

$$\tilde{\boldsymbol{\Psi}} = \mathbf{A}(\boldsymbol{\Psi})\mathbf{y}$$

with

$$\mathbf{A}(\boldsymbol{\Psi}) = \mathbf{U}\mathbf{T}$$

and moreover

$$\tilde{\boldsymbol{\mu}} = \mathbf{T}\mathbf{Y}.$$

We see that $\mathbf{A}(\boldsymbol{\Psi}) \in [\boldsymbol{\Psi}]$, since

$$\xi(\tilde{\boldsymbol{\Psi}}) = \mathbf{A}(\boldsymbol{\Psi})\boldsymbol{\mu} = \mathbf{U}\mathbf{T}\mathbf{X}_\circ\boldsymbol{\beta}_\circ = \mathbf{U}\boldsymbol{\mu} = \boldsymbol{\Psi}$$

and that, according to Lemma 3, $\mathbf{A} \in [\mathbf{\Psi}]$ if and only if

$$\mathbf{AT} = \mathbf{A}(\mathbf{\Psi})\mathbf{T} = \mathbf{UTT} = \mathbf{UT} = \mathbf{A}(\mathbf{\Psi}).$$

Putting $\mathbf{T}^c = \mathbf{I}_n - \mathbf{T}$, we have, with $\mathbf{A} \in [\mathbf{\Psi}]$,

$$\mathbf{A} = \mathbf{AT} + \mathbf{AT}^c = \mathbf{A}(\mathbf{\Psi}) + r\mathbf{B}$$

with $-\infty < r < +\infty$, this is, $r \in \mathbb{R}$ and

$$\mathbf{B} = \frac{1}{r}\mathbf{AT}^c.$$

Thus,

$$Cov_\theta(\mathbf{AY}) = cov_\theta(\mathbf{A}[\mathbf{\Psi}]\mathbf{Y}) + 2r cov_\theta(\mathbf{A}[\mathbf{\Psi}]\mathbf{Y}, \mathbf{BY}) + r^2 cov_\theta(\mathbf{BY}),$$

it is easy to see that we have, whatever $r \in \mathbb{R}$

$$cov_\theta(\tilde{\mathbf{\Psi}}) = cov_\theta(\mathbf{A}(\mathbf{\Psi})\mathbf{Y}) \leq cov_\theta(\mathbf{AY})$$

if and only if

$$cov_\theta(\mathbf{A}(\mathbf{\Psi})\mathbf{Y}, \mathbf{BY}) = \mathbf{0}.$$

Since

$$\mathbf{B} = \frac{1}{r}\mathbf{AT}^c,$$

we get

$$cov_\theta(\mathbf{ATY}, \mathbf{AT}^c\mathbf{Y}) = \mathbf{0}.$$

Now

$$cov_\theta(\mathbf{ATY}, \mathbf{AT}^c\mathbf{Y}) = \mathbf{AT}V(\boldsymbol{\theta})\mathbf{T}^c\mathbf{A}^T$$

so that to have

$$cov_\theta(\tilde{\mathbf{\Psi}}) \leq cov_\theta(\mathbf{AY})$$

for every estimable vector, we must have

$$\mathbf{T}V(\boldsymbol{\theta})\mathbf{T}^c = 0$$

which gives

$$\begin{aligned} \mathbf{T}V(\boldsymbol{\theta})(\mathbf{I}_n - \mathbf{T}) = 0 &\Leftrightarrow \mathbf{T}V(\boldsymbol{\theta}) - \mathbf{T}V(\boldsymbol{\theta})\mathbf{T} = 0 \\ &\Leftrightarrow \mathbf{T}V(\boldsymbol{\theta}) = \mathbf{T}V(\boldsymbol{\theta})\mathbf{T} \end{aligned} \tag{4}$$

and

$$V(\boldsymbol{\theta})\mathbf{T} = (\mathbf{T}V(\boldsymbol{\theta}))^T = (\mathbf{T}V(\boldsymbol{\theta})\mathbf{T})^T = \mathbf{T}V(\boldsymbol{\theta})\mathbf{T} = \mathbf{T}V(\boldsymbol{\theta})$$

which implies that $V(\boldsymbol{\theta})\mathbf{T} = \mathbf{T}V(\boldsymbol{\theta})$ and so, the matrices $\mathbf{T}$ and $V(\boldsymbol{\theta})$ commutes.

We now establishes

Theorem 4. *The least square estimators, LSE are uniformly best linear unbiased estimator, UBLUE if and only if, for every $\boldsymbol{\theta}$, $\boldsymbol{T}$ commute with $V(\boldsymbol{\theta})$.*

Proof. The preceding discussion establishes the necessary condition to complete the proof. We point out that, we have

$$cov_\theta(\mathbf{AY}) = cov_\theta(\mathbf{A(\Psi)Y}) + r^2 cov_\theta(\mathbf{BY}) \geq cov_\theta(\mathbf{A}[\mathbf{\Psi}]\mathbf{Y}) = cov_\theta(\tilde{\mathbf{\Psi}}).\blacksquare$$

Now, the models with GOBS where $\mathbf{T}$ commutes with $\{\mathbf{M}_1, \ldots, \mathbf{M}_r\}$ and so with $V(\boldsymbol{\theta})$, whatever $\boldsymbol{\theta}$, are those models with COBS whose LSE are UBLUE.

In establishing Theorem 4, we have

$$V(\boldsymbol{\theta}) = \sum_{i=1}^{w} \boldsymbol{\theta}_i \mathbf{M}_i$$

in order to widen the class of models to which our result applies.

Moreover if we extend the models considered requiring only that $\boldsymbol{\gamma} \in \nabla_+$ a subspace of $\mathbb{R}^m$, theorem 4 continues to hold whatever the dimension of ∇.

Acknowledgements

This work was funded by project PEst-OE/MAT/UI0297/2014 (Fundação para a Ciência e Tecnologia).

References

[1] A. Areia and F. Carvalho, *Perfect families: an appliacation to orthogonal and error orthogonal models*, AIP Conf. Proc. Numerical Analysis and Applied Mathematics: International Conferences of Numerical Analysis and Applied Mathematics **1558** (2013), 841–846.

[2] T. Calinski and S. Kageyama, Block Designs: A Randomization Approach. Vol. I: Analysis (Lecture Notes in Statistics, Springer, 2000) 150.

[3] T. Calinski and S. Kageyama, Block Designs: A Randomization Approach. Vol. II: Design (Lecture Notes in Statistics, Springer, 2003) 170.

[4] F. Carvalho, J.T. Mexia and M.M. Oliveira, *Canonic inference and commutative orthogonal block structure*, Discuss. Math. Probab. and Stat. **28** (2) (2008), 171–181.

[5] F. Carvalho, J.T. Mexia and C. Santos, *Commutative orthogonal block structure and error orthogonal models*, Electronic Journal of Linear Algebra **25** (2013), 119–128.

[6] S.S. Ferreira, D. Ferreira, C. Fernandesa and J.T. Mexia, *Orthogonal models and perfect families of symmetric matrices*, Bulletin of the International Statistical Institute. Proc. ISI 2007, Lisboa 22–28 August (2007), 3252–3254.

[7] M. Fonseca, J.T. Mexia and R. Zmyślony, *Inference in normal models with commutative orthogonal block structure*, Acta et Commentationes Universitatis Tartunesis de Mathematica (2008), 3–16.

[8] A. Houtman and T.P. Speed, *Balance in designed experiments with orthogonal block structure*, The annals of Statistics **11** (4) (1983), 1069–1085.

[9] E.L. Lehmann and G. Casela, Theory of Point estimation, 2nd ed. Springer Tests Statistics (New York, Springer, 1998).

[10] S. Mejza, *On some aspects of general balance in designed experiments*, Statistica **52** (1992), 263–278.

[11] J.T. Mexia, R. Vaquinhas, M. Fonseca and R. Zmyślony, *COBS: segregation, matching, crossing and nesting*, Latest Trends on Applied Mathematics, Simulation, Modelling (2010), 249–255.

[12] J.A. Nelder, *The analysis of randomized experiments with orthogonal block structure and the null analysis of variance*, Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences **283** (1393), 147–162.

[13] J.A. Nelder, *The analysis of randomized experiments with orthogonal block structure II. Treatment structure and the general analysis of variance*, Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences **283** (1393), 163–178.

[14] J.R. Schott, Matrix Analysis for Statistics (Wiley Series in Probability and Statistics, 1997).

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

[16] D.M. Vanleeuwen, D.S. Birks, J.F. Seely, J. Mills, J.A. Greenwood and C.W. Jones, *Sufficient conditions for orthogonal designs in mixed linear models*, Journal of Statistics Planning and Inference **73** (1998), 373–389.

[17] R. Zmyślony, *A characterization of the best linear unbiased estimators in the general linear model*, Lecture Notes Statistics **2** (1978), 365–373.

Received 20 November 2014