Commutative orthogonal block structure and error orthogonal models

Francisco Carvalho
fpcarvalho@ipt.pt

Joao T. Mexia

Carla Santos

Follow this and additional works at: http://repository.uwyo.edu/ela

Recommended Citation
DOI: https://doi.org/10.13001/1081-3810.1601
COMMUTATIVE ORTHOGONAL BLOCK STRUCTURE AND ERROR ORTHOGONAL MODELS

FRANCISCO CARVALHO†, JOÃO T. MEXIA‡, AND CARLA SANTOS§

Abstract. A model has orthogonal block structure, OBS, if it has variance-covariance matrix that is a linear combination of known pairwise orthogonal orthogonal projection matrices that sum to the identity matrix. These models were introduced by Nelder in 1965, and continue to play an important part in randomized block designs.

Two important types of OBS are related, and necessary and sufficient conditions for model of one type belonging to the other are determined.

The first type, is that of models with commutative orthogonal block structure in which $T$, the orthogonal projection matrix on the space spanned by the mean vector, commutes with the orthogonal projection matrices in the expression of the variance-covariance matrix.

The second type, is that of error orthogonal models.

These results open the possibility of deepening the study of the important class of models with OBS.

Key words. COBS, Error orthogonal models, Commutative Jordan algebras

AMS subject classifications. 15A15, 15F10.

1. Introduction. Models with orthogonal block structure, OBS, have the family

$$\left\{ \gamma^\circ \sum_{j=1}^{m^\circ} \gamma_j^2 Q_j^\circ \right\} = \{V(\gamma^\circ)\}$$

Received by the editors on November 30, 2011. Accepted for publication on March 6, 2013

Handling Editor: Simo Puntanen.

†Unidade Departamental de Matemática e Física, Instituto Politécnico de Tomar, 2300-313 Tomar, Portugal (fpcarvalho@ipt.pt)

CMA - Centro de Matemática e Aplicações, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal.

‡Departamento de Matemática, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal.

CMA - Centro de Matemática e Aplicações, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal.

§Departmento de Matemática e Ciências Físicas, Instituto Politécnico de Beja, Rua Pedro Soares, 7800-295 Beja, Portugal.
of variance-covariance matrices $V(\gamma^o) = \sum_{j=1}^{m^o} \gamma_j^o Q_j^o$. In these linear combinations, $\gamma_1^o, \ldots, \gamma_{m^o}^o$, are nonnegative and the matrices $Q_1^o, \ldots, Q_{m^o}^o$ are known pairwise orthogonal orthogonal projection matrices, POOPM, such that

$$\sum_{j=1}^{m^o} Q_j^o = I_n.$$ 

These models were introduced by [10, 11] and continue to play an important part in the theory of randomized block designs, see e.g. [1, 2].

Our work will be centered on two classes of models:

1. Models with Commutative Orthogonal Block Structure, COBS. These are OBS models where $T$, the orthogonal projection matrix, OPM, on the range space $\Omega$, spanned by the mean vectors, commutes with POOPM $Q_1^o, \ldots, Q_{m^o}^o$. These models were introduced in [7] and further studied in [3]. As we shall see, the least square estimators, LSE, will be uniformly best linear unbiased estimators, UBLUE, i.e., they will be best linear unbiased estimators, BLUE, whatever the variance models;

2. Error orthogonal models, EO. These are models whose LSE are UBLUE, and with $T' = I_n - T$, the submodel $y' = T'y$ is OBS with variance-covariance matrix

$$V'(\gamma) = \sum_{j=1}^{m'} \gamma_j'^o Q_j'^o$$ 

where the matrices $Q_1'^o, \ldots, Q_{m'}'^o$ are known POOPM and moreover, the parameter space, $\Gamma'$, of the variance components $\gamma' = (\gamma_1', \ldots, \gamma_{m'}')$, contains a non-empty open set. These models were introduced in [17, 18]. To distinguish this reinforced version of OBS from the usual one, we will call it robust orthogonal block structure, ROBS.

In our study, we will use commutative Jordan algebras of symmetric matrices, CJAS, which we will consider in the next section. We then have a section on COBS and one in which we relate both classes of models, giving necessary and sufficient conditions for a EO to be COBS and for a COBS to be EO.

2. Commutative Jordan Algebras. For our purposes, CJAS will be linear spaces of symmetric matrices that commute and are closed under taking squares. These structures were introduced by [8] in a reformulation of Quantum Mechanics. Later on Seely, see [13, 14], rediscovered these structures and used them to carry out linear statistical inference. This rediscovery initiated a very productive line of work, see for instance [15, 16], [4], [17, 18], [5, 6, 7], [3], [9], etc.
Now, the symmetric matrices $M_1, \ldots, M_v$ commute if and only if they are jointly diagonalized by the same orthogonal matrix $P$, see [12]. Thus the set

$$\mathcal{M} = \{M_1, \ldots, M_v\}$$

will be contained in $V(P)$, the family of matrices diagonalized by $P$. Since $V(P)$ is itself a CJAS, we see that a family of symmetric matrices is contained in a CJAS if and only if its matrices commute. Since the intersection of CJAS is a CJAS, if the matrices in $\mathcal{M}$ commute, there will be a minimum CJAS $\mathcal{A}(\mathcal{M})$ containing $\mathcal{M}$, the CJAS generated by $\mathcal{M}$. Let us now establish the following proposition:

**Proposition 2.1.** A symmetric matrix $U$ commutes with the matrices of a family

$$\mathcal{M} = \{M_1, \ldots, M_v\}$$

of symmetric matrices that commute, if and only if it commutes with all the matrices of the CJAS generated by that family, $\mathcal{A}(\mathcal{M})$.

**Proof.** Since $\mathcal{M} \subset \mathcal{A}(\mathcal{M})$, if $U$ commutes with the matrices of $\mathcal{A}(\mathcal{M})$, it will commute with $M_1, \ldots, M_v$. Conversely, if the matrices in $\mathcal{U} = \{U, M_1, \ldots, M_v\}$ commute we will have a CJAS generated by $\mathcal{U}$, $\mathcal{A}(\mathcal{U})$, that, since $\mathcal{M} \subset \mathcal{U}$, contains $\mathcal{A}(\mathcal{M})$ as well as $U$, so $U$ will commute with all the matrices in $\mathcal{A}(\mathcal{M})$. \(\blacksquare\)

Of importance when considering CJAS is the fact, shown in [15], that any CJAS, $\mathcal{A}$, has an unique basis constituted by POOPM. This will be the principal basis of that CJAS, $pb(\mathcal{A})$. From the existence of this principal basis it follows that any CJAS contains the products of its matrices and not only their squares. Moreover, if $Q$ is an OPM belonging to a CJAS $\mathcal{A}$ with principal basis $pb = \{Q_1, \ldots, Q_m\}$, we will have

$$Q = \sum_{j=1}^{m} a_j Q_j$$

and, since $Q$ is idempotent and the $Q_1, \ldots, Q_m$ are idempotent and pairwise orthogonal, we must have $a_j = 0$ or $a_j = 1$, $j = 1, \ldots, m$. Thus the OPM’s belonging to a CJAS are in fact sums of all or part of the matrices in the principal basis of the CJAS.

When $\sum_{j=1}^{m} Q_j = I_n$, the CJAS with principal basis $pb(\mathcal{A}) = \{Q_1, \ldots, Q_m\}$ will be complete. It is now clear that any family of POOPM will be the principal basis of the CJAS constituted by the linear combinations of the matrices in the family. If we consider models with OBS, the $Q_1^\circ, \ldots, Q_m^\circ$ will constitute the principal basis of a complete CJAS, $\mathcal{A}^\circ$. It’s also important to point out that for a CJAS to contain invertible matrices, that CJAS must be complete.
If we consider an \( n \times n \) matrix \( G \) belonging to a complete CJAS \( \mathcal{A} \), it's clear that \( I_n \) and \( G^c = I_n - G \), the complement of \( G \), will also belong to \( \mathcal{A} \) as well as \( GU \) and \( G^c U \), whatever the matrix \( U \) of \( \mathcal{A} \). We may now establish the following proposition.

**Proposition 2.2.** Let \( Q \) be an orthogonal projection matrix that commutes with the matrices of a complete CJAS \( \mathcal{A}^0 \). If \( pb(\mathcal{A}^0) = \{ Q_1^0, \ldots, Q_m^0 \} \), the principal basis of the CJAS generated by \( Q \), the matrices of \( \mathcal{A}^0 \) is constituted by the non-null products \( QQ_j^0 \) and \( Q_c Q_j^0 \), \( j = 1, \ldots, m \).

**Proof.** Let \( Q_1, \ldots, Q_m \) be the non-null products \( QQ_j^0 \) and \( Q_c Q_j^0 \), \( j = 1, \ldots, m \). It is easy to see that \( Q_1, \ldots, Q_m \) are POOPM and therefore constitute the principal basis of a CJAS \( \mathcal{A} \) that contains \( Q \) and the matrices of \( \mathcal{A}^0 \), since it contains the \( Q_1^0, \ldots, Q_m^0 \). To complete the proof we have only to point out that any CJAS containing both \( Q \) and the matrices of \( \mathcal{A}^0 \) contains the \( Q_1, \ldots, Q_m \), so it will contain \( \mathcal{A} \). \( \square \)

3. Models with commutative orthogonal block structure. As stated in the introduction, this section is focussed on COBS and related models.

We will consider a mixed model with \( n \) observations written in its usual form,

\[
y = \sum_{i=0}^{w} X_i \beta_i,
\]

where \( \beta_0 \) is fixed and the \( \beta_1, \ldots, \beta_w \) are random, independent, with null mean vectors and variance-covariance matrices \( \sigma_1^2 I_{c_1}, \ldots, \sigma_w^2 I_{c_w} \), having mean vector and variance-covariance matrix

\[
\begin{cases}
\mu_0 = X_0 \beta_0 \\
V = \sum_{i=1}^{w} \sigma_i^2 M_i,
\end{cases}
\]

where \( M_i = X_i X_i^\top, i = 1, \ldots, w \). The \( \sigma_1^2, \ldots, \sigma_w^2 \) will be the variance components, while the OPM on \( \Omega \) will be given by

\[
T = X_0 \left( X_0^\top X_0 \right)^\dagger X_0^\top = X_0 A^\dagger,
\]

where \( A^\dagger \) denotes the Moore-Penrose inverse of matrix \( A \).

The range space of \( U \) is denoted by \( R(U) \)

**Proposition 3.1.** If the matrices \( M_1, \ldots, M_w \) commute and \( R \left( \begin{bmatrix} X_1 & \cdots & X_w \end{bmatrix} \right) = \mathbb{R}^n \), the model is OBS.
Proof. Let \( \mathbb{P}^o = \{ Q_1^o, \ldots, Q_{m^o}^o \} \) be the principal basis of the CJAS, \( \mathcal{A}^o \), generated by the \( M_1, \ldots, M_w \), then \( M_i = \sum_{j=1}^{m^o} b_{i,j}^o Q_j^o \), \( i = 1, \ldots, w \), and so the variance-covariance matrix can be expressed as the linear combination

\[
V = \sum_{i=1}^{w} \sigma_i^2 \left( \sum_{j=1}^{m^o} b_{i,j}^o Q_j^o \right) = \sum_{j=1}^{m^o} \gamma_j^o Q_j^o,
\]

with \( \gamma_j^o = \sum_{i=1}^{w} b_{i,j}^o \sigma_i^2 \), \( j = 1, \ldots, m^o \). Moreover, since

\[
R \left( \sum_{i=1}^{w} M_i \right) = R \left( \begin{bmatrix} X_1 & \cdots & X_w \end{bmatrix} \right) = \mathbb{R}^n,
\]

\( \sum_{i=1}^{w} M_i \) is an invertible matrix belonging to \( \mathcal{A}^o \) which will be complete. \( \square \)

We are now able to establish the following proposition.

**Proposition 3.2.** If the matrices \( T, M_1, \ldots, M_w \) commute and

\[
R \left( \begin{bmatrix} X_1 & \cdots & X_w \end{bmatrix} \right) = \mathbb{R}^n,
\]

the model is COBS.

Proof. According to Proposition 3.1, the model is OBS. Moreover, from Proposition 2.1 it follows that \( T \) will commute with the \( Q_1^o, \ldots, Q_{m^o}^o \), so the model will be COBS. \( \square \)

The principal basis of the CJAS \( \mathcal{A}^o \) generated by \( T \) and the matrices \( Q_1^o, \ldots, Q_{m^o}^o \) will be constituted by the \( \{ Q_1^o, \ldots, Q_{m^o}^o \} \), these are the non-null products \( T Q_j^o \) and \( T^o Q_j^o \), \( j = 1, \ldots, m^o \). Moreover, \( T \) will be the sum of matrices in \( pb(\mathcal{A}) \) which can be reordered in such a way that

\[
T = \sum_{j=1}^{z} Q_j.
\]

We also will have

\[
M_i = \sum_{j=1}^{m} b_{i,j} Q_j \quad j = 1, \ldots, w,
\]

so the variance-covariance matrix is written as

\[
V = \sum_{j=1}^{m} \gamma_j Q_j,
\]

with \( \gamma_j = \sum_{i=1}^{w} b_{i,j} \sigma_i^2 \), \( j = 1, \ldots, m \). These will be the canonical variance components.
According to the last expressions let us take the vectors

\[
\gamma(1) = \begin{bmatrix}
\gamma_1 \\
\vdots \\
\gamma_z \\
\end{bmatrix}; \quad \gamma(2) = \begin{bmatrix}
\gamma_{z+1} \\
\vdots \\
\gamma_m \\
\end{bmatrix}; \quad \sigma^2 = \begin{bmatrix}
\sigma^2_1 \\
\vdots \\
\sigma^2_w \\
\end{bmatrix}
\]

and the matrix \( B = [b_{i,j}] \) that can be partitioned as

\[
B = \begin{bmatrix}
B(1) & B(2)
\end{bmatrix},
\]

where the sub-matrix \( B(1) \) has \( z \) columns and the sub-matrix \( B(2) \) has \( m - z \) columns, and we can express the canonical variance components vectors as

\[
\gamma(\ell) = B(\ell)^\top \sigma^2, \quad \ell \in \{1, 2\}.
\]

We now point out that, due to \( R(Q_j) \subset \Omega, \ j = 1, \ldots, z \), only the \( \gamma_{z+1}, \ldots, \gamma_m \) are directly estimable. But if the row vectors of \( B(2), \) which are the columns vectors of \( B(2)^\top, \) are linearly independent, we will have

\[
\begin{cases}
\sigma^2 = (B(2)^\top)^+ \gamma(2) \\
\gamma(1) = B(1)^\top (B(2)^\top)^+ \gamma(2),
\end{cases}
\]

so all variance components, either usual or canonic, will be estimable. This condition is not a very restrictive condition in most situations. In this case the relevant parameters for the random effects part determine each other. Thus that part segregates as a sub-model and we say there is segregation.

Another interesting case is when \( B(1) \) is a sub-matrix of \( B(2). \) Thus \( \gamma(1) \) is a sub-vector of \( \gamma(2) \) and its components are directly estimable. This case is called matching since it is based on the matching of columns of \( B(1) \) and \( B(2). \)

Now, in COBS, matrices \( T \) and \( V \) commute whatever the variance components, which, see Theorem 2 in [19], is a necessary and sufficient condition for LSE to be UBLUE. A similar result can also be found in [20].

4. Relating the models. To lighten the writing, we name as UBLUE the models whose LSE are UBLUE. Since EO and COBS are UBLUE, we start by considering the wider class of models. We put \( \Theta = \{\sigma^2; \sigma^2 \geq 0\} \) and \( \mathcal{V} = \{V(\sigma^2); \sigma^2 \in \Theta\}, \) with

\[
V(\sigma^2) = \sum_{i=1}^w \sigma^2_i M_i.
\]
With $\delta_i$ the vectors whose components are zero but the $i$-th, which is 1, if $\sigma^2 \in \Theta$, we have $\sigma^2 + \delta_i \in \Theta$, $i = 1, \ldots, w$. Now, if the models is UBLUE, we have, see again Theorem 2 in [19],

$$TV = \{TV (\sigma^2); \sigma^2 \in \Theta\} = \{V (\sigma^2) T; \sigma^2 \in \Theta\} = \mathcal{V} T,$$

so, since $M_i = V (\sigma^2 + \delta_i) - V (\sigma^2)$, $i = 1, \ldots, w$, we have

$$TM_i = T \left(V (\sigma^2 + \delta_i) - V (\sigma^2)\right) = \left(V (\sigma^2 + \delta_i) - V (\sigma^2)\right) T = M_i T$$

for $i = 1, \ldots, w$.

Moreover, with $T^c = I_n - T$, we also have

$$T^c M_i = M_i T^c, \quad i = 1, \ldots, w,$$

so, with $M_i^c = TM_i T$ and $M_i' = T^c M_i T^c$, $i = 1, \ldots, w$, it is easy to see that

$$M_i = M_i^c + M_i', \quad i = 1, \ldots, w,$$

as well as

$$M_i M_\ell = M_i^c M_\ell^c + M_i' M_\ell', \quad i = 1, \ldots, w, \quad \ell = 1, \ldots, w,$$

since $M_i^c M_\ell' = M_\ell' M_i^c = 0_{n \times n}$, $i = 1, \ldots, w, \quad \ell = 1, \ldots, w$.

Besides, these matrices $M_1, \ldots, M_w$ commute if and only if the matrices

$$M_1^c, \ldots, M_w^c \quad [M_1', \ldots, M_w']$$

commute. Thus, $\mathcal{M} = \{M_1, \ldots, M_w\}$ will be commutative when and only when $\mathcal{M}^c = \{M_1^c, \ldots, M_w^c\}$ and $\mathcal{M}' = \{M_1', \ldots, M_w'\}$ are commutative. As we saw, when $\mathcal{M}$ is commutative, it generates a CJAS, and likewise, when $\mathcal{M}^c [\mathcal{M}']$ are commutative, they generate a CJAS $\mathcal{D}^c [\mathcal{D}']$, with principal basis $\mathcal{Q} = \{Q_1^c, \ldots, Q_m^c\}$ $\mathcal{Q}' = \{Q_1', \ldots, Q_m'\}$.

Now let the model be EO and put $\mathcal{V}' = \{V' (\sigma^2); \sigma^2 \in \Theta\}$, with

$$V' (\sigma^2) = \sum_{i=1}^w \sigma_i^2 M_i'.$$

As we saw, the variance-covariance matrices $V' (\sigma^2)$ belong, when the model is EO, to a CJAS, $\mathcal{A}'$, with principal basis $\{Q_1', \ldots, Q_m'\}$. Thus, with $\sigma^2 \in \Theta$, we have,

$$M_i' = V' (\sigma^2 + \delta_i) - V' (\sigma^2) \in \mathcal{A}', \quad i = 1, \ldots, w,$$
and so
\[ M'_i = \sum_{j=1}^{m'} b'_{i,j} Q'_j, \quad i = 1, \ldots, w. \]
Thus
\[ V'(\sigma'^2) = \sum_{j=1}^{m'} \gamma'_j Q'_j, \]
with
\[ \gamma'_j = B'^\top \sigma^2, \]
where \( B' = [b'_{i,j}] \). Now \( \Gamma' = \{ \gamma' : \sigma^2 \in \Theta \} \) contains non-empty open sets if and only if there are no linear restrictions on the \( \gamma'_1, \ldots, \gamma'_{m'} \), which is equivalent to \( B'^\top \) and, consequently, \( B' \) being invertible. Thus a necessary and sufficient condition for the model \( y' = T^c y \) enjoying ROBS is that matrix \( B' \) is invertible.

We now have the following proposition.

**Proposition 4.1.** A UBLUE is EO if and only if its family \( \mathcal{M}' \) is commutative and its matrix \( B' \) is invertible.

**Proof.** As we saw, if a model is EO, it is UMVUE and has a commutative \( \mathcal{M}' \) matrix family and the corresponding matrix \( B' \) is invertible. Besides this, if the model is UMVUE and has a commutative matrix family \( \mathcal{M}' \), we will have
\[ M'_i = \sum_{i=1}^{m'} b'_{i,j} Q'_j, \]
with \( \{ Q'_1, \ldots, Q'_{m'} \} \), the principal basis of \( \mathcal{A}' = \mathcal{A}(\mathcal{M}') \), so that we will have
\[ V(\gamma') = \sum_{j=1}^{m'} \gamma'_j Q'_j, \]
with \( \gamma'_j = B'^\top \sigma^2. \)

To complete the proof, we have only to point out that if \( B' \) is invertible, there will be no linear restrictions on the components of \( \gamma' \) and so the corresponding parameter space will contain nonempty open sets. \( \Box \)

**Corollary 4.2.** A COBS is EO if and only if its matrix \( B(2) \) is invertible.

**Proof.** A COBS is EO and its family \( \mathcal{M} \) is commutative, so its family \( \mathcal{M}' \) will be commutative. The rest of the proof follows from Proposition 4.1, since we now have \( B' = B(2) \). \( \Box \)
We now establish the following proposition,

**Proposition 4.3.** An EO is COBS, if and only if its family $\mathcal{M}'$ is commutative.

**Proof.** Since an EO is UBLUE with commutative family $\mathcal{M}'$, it will be UBLUE with commutative family $\mathcal{M}$ if and only if its family $\mathcal{M}'$ is commutative. The rest of the proof is straightforward. □

**Acknowledgment.** We deeply thank the referee for insightful comments, namely those on the distinction between the usual OBS concept and the one used by Van-Leeuwen et al., which led to significant enhancement of this work.

This work was partially supported by CMA/FCT/UNL, under the project PEst-OE/MAT/UI0297/2011.

**REFERENCES**


