

Normal models with Orthogonal Block Structure

Francisco Carvalho
 CMA - Centro de Matemática
 e Aplicações
 Universidade Nova de Lisboa
 Unidade Departamental de Matemática
 Instituto Politécnico de Tomar
 Portugal
 fpcarvalho@ipt.pt

Roman Zmysłony
 Faculty of Mathematics, Computer
 Science and Econometrics
 University of Zielona Góra
 Poland

J.T. Mexia
 CMA - Centro de Matemática
 e Aplicações
 Universidade Nova de Lisboa
 Portugal

Abstract: A model with Orthogonal Block Structure, OBS, is characterized by its variance-covariance matrices being all positive semi-definite matrices given by $\sum_{j=1}^m \gamma_j K_j$, where K_1, \dots, K_m , are known orthogonal projection matrices that are pairwise orthogonal. We show that when normality is assumed, models with OBS have complete sufficient statistics and we will obtain uniformly minimum variance unbiased estimators both for estimable vectors and variance-covariance matrices. The case of normal mixed models written as $y = \sum_{i=0}^w X_i \beta_i$, where β_0 is fixed and the β_1, \dots, β_w being normal independent, is considered and an example is discussed.

Key-Words: Typing manuscripts, L^AT_EX

1 Introduction

Nelder, see [7, 8], introduced models with Orthogonal Block Structure, OBS, as models where the family $\mathcal{V} = \left\{ \sum_{j=1}^m \gamma_j K_j \right\}$ of variance-covariance matrices is constituted by all positive semi-definite matrices

$$V = \sum_{j=1}^m \gamma_j K_j,$$

where the K_1, \dots, K_m are known orthogonal projection matrices that are pairwise orthogonal and add up to I_n and the $\gamma_1, \dots, \gamma_m$ are the canonic variance components.

We now point out that for all positive semi-definite matrices $\sum_{j=1}^m \gamma_j K_j$ to be variance-covariance matrices, we must have $\gamma = (\gamma_1, \dots, \gamma_m) \in \mathbb{R}_+^m$, with ∇_+ the family of vectors of subspace ∇ with non-negative components on the $\gamma_1, \dots, \gamma_m$. The existence of such restriction would permit, see [5], the model density to be full rank when we assume normality. So we than could not have complete and sufficient statistics from which to derive uniformly minimum variance unbiased estimators (UMVUE).

We will apply our results to normal mixed models

$$y = \sum_{i=0}^w X_i \beta_i$$

where β_0 is fixed and the β_1, \dots, β_w are normal, independent with null mean vectors and variance-covariance matrices $\theta_1 I_{c_1}, \dots, \theta_w I_{c_w}$. Thus y will be normal with mean vector $\mu = X_0 \beta_0$ and variance-covariance matrix $V(\theta) = \sum_{i=1}^w \theta_i M_i$ with $M_i = X_i X_i^T$, $i = 1, \dots, w$. We will show that such model have OBS, when matrices M_1, \dots, M_w commute and so, see [2],

$$M_i = \sum_{j=1}^m b_{i,j} K_j \quad i = 1, \dots, m$$

with K_1, \dots, K_m known orthogonal projection matrices that are pairwise orthogonal, if and only if $w = m$.

This is an important point since considering models such as those given by Example 1 and 2 of [4], do not satisfy this condition. Actually we will present a possible modifications of these models to ensure that the modified models have OBS and, when normality is assumed, complete and sufficient statistics.

Namely this modification holds for models with u random effects factors that cross. Then $w = 2^u$ and

the $X_i\beta_i, i = 1, \dots, w$, corresponds to the effects and interactions of these models.

When normality is assumed we get UMVUE for variance components and estimable vectors in such models.

We now apply our results to normal mixed models,

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

where β_0 is fixed and the β_1, \dots, β_w are normal, independent with null mean vectors and variance-covariance matrices $\theta_1 I_{c_1}, \dots, \theta_w I_{c_w}$. In what follows we put $z \sim \mathcal{N}(\eta; W)$ to indicate that z is normal with mean vector η and variance-covariance matrix W . Namely for these normal mixed models, we will have $y \sim \mathcal{N}\left(X_0\beta_0; \sum_{i=1}^w \theta_i M_i\right)$, with $M_i = X_i X_i^T, i = 1, \dots, w$. We will characterize these models that are NOBS or have OBS when normality is discarded and, in the case of NOBS, our results.

2 Sufficient statistics and natural parameters

Let the row vectors of a matrix A_j constitute an orthonormal basis for the range space $\nabla_j = \mathcal{R}(K_j)$ of $K_j, j = 1, \dots, m$. Then, with $g_j = \text{rank}(K_j) = \text{rank}(A_j), j = 1, \dots, m$, we have

$$\begin{cases} A_j A_j^T = I_{g_j} & j = 1, \dots, m \\ A_j^T A_j = K_j & j = 1, \dots, m. \end{cases}$$

If the initial NOBS model y has, besides variance-covariance matrix V , mean vector $\mu = X\beta$, we get the normal homocedastic sub-models

$$y_j = A_j y \sim \mathcal{N}(\mu_j; \gamma_j I_{g_j}), \quad j = 1, \dots, m.$$

The y_1, \dots, y_m are independent since they have joint normal density and null cross-covariance matrices.

Let P_j and Q_j be the orthogonal projection matrices on the space Ω_j spanned by the mean vector μ_j and its orthogonal complement, $j = 1, \dots, m$. We put $p_j = \text{rank}(P_j)$ and $q_j = \text{rank}(Q_j)$ and consider the sets

$$\begin{cases} \mathcal{C} = \{j : p_j > 0\} \\ \mathcal{D} = \{j : q_j > 0\}. \end{cases}$$

Let also the row vectors of $W_j, j \in \mathcal{C}$, constitute an orthonormal basis for Ω_j so that $P_j = W_j^T W_j, j \in \mathcal{C}$. Then it is easy to show that

$$\begin{aligned} \|y_j - \mu_j\|^2 &= S_j - 2\mu_j^T y_j + \|\mu_j\|^2 \\ &= S_j - 2\eta_j^T \tilde{\eta}_j + \|\eta_j\|^2, \quad j \in \mathcal{C}, \end{aligned}$$

with $\eta_j = W_j \mu_j, \tilde{\eta}_j = W_j y_j$ and $S_j = \|y_j\|^2, j \in \mathcal{C}$, and that

$$\|y_j - \mu_j\|^2 = S_j, \quad j \notin \mathcal{C}.$$

Thus, the sub-models will have densities

$$\begin{cases} n_j(y_j) = \frac{e^{-\frac{1}{2\gamma_j}(S_j - 2\eta_j^T \tilde{\eta}_j + \|\eta_j\|^2)}}{(2\pi\gamma_j)^{g_j/2}}, & j \in \mathcal{C}, \\ n_j(y_j) = \frac{e^{-\frac{S_j}{2\gamma_j}}}{(2\pi\gamma_j)^{g_j/2}}, & j \notin \mathcal{C}. \end{cases}$$

If we take $\mathcal{C} = \{1, \dots, z\}$, the joint density of the sub-models will be, since they are independent,

$$n(y) = \frac{e^{-\frac{1}{2}\left(\sum_{j=1}^z (\mu_j S_j - 2\xi_j^T \tilde{\eta}_j + \nu_j \|\xi_j\|^2) + \sum_{j=z+1}^m \nu_j S_j\right)}}{\prod_{j=1}^m \left(\frac{2\pi}{\nu_j}\right)^{\frac{g_j}{2}}},$$

with sufficient statistics $S_j, j = 1, \dots, m$ and $\tilde{\eta}_j, j = 1, \dots, z$, and natural parameters ν with components $\nu_j = \gamma_j^{-1}, j = 1, \dots, m$, and

$$\xi = [\xi_1^T \dots \xi_m^T]^T,$$

where $\xi_j = \frac{1}{\gamma_j} \eta_j, j = 1, \dots, m$.

We now establish the following theorem.

Theorem 1

For NOBS the statistics $S_j, j = 1, \dots, m$, and $\tilde{\eta}_j, j = 1, \dots, z$, are complete and sufficient.

Proof: Let Γ be the parameter space for the joint density. To show that the mentioned statistics are complete and sufficient, we have only to show that Γ contains the cartesian product of non degenerate intervals, one for each component of $\lambda = [\nu; \xi]$, see [5] which follows from

- $\nu_j = \gamma_j^{-1}$ where γ_j may take non negative values, $j = 1, \dots, m$;
- ξ_j spanning \mathbb{R}^{p_j} , since $\mu_j = W^T \xi_j$ spans $\Omega_j, j = 1, \dots, z$.

□

A first interesting application of this result is to estimable vectors. It is easy to see that Ψ will be estimable if it may be written as

$$\Psi = U\mu.$$

Now

$$\mu = \left(\sum_{j=1}^z K_j\right) \mu = \sum_{j=1}^z A_j^T \mu_j^T = \sum_{j=1}^z A_j^T W_j \eta_j,$$

so

$$\tilde{\mu} = \sum_{j=1}^z A_j^\top W_j \tilde{\eta}_j,$$

and

$$\tilde{\Psi} = U\tilde{\mu},$$

being unbiased and function of the complete and sufficient statistics, is UMVUE.

Since

$$y_j^\top Q_j y_j = y_j^\top y_j - y_j^\top P_j y_j = S_j - \|\tilde{\eta}_j\|^2, \quad j \in \mathcal{D},$$

the estimators

$$\tilde{\gamma}_j = \frac{S_j - \|\eta_j\|^2}{q_j}, \quad j \in \mathcal{D},$$

will be unbiased and derived from the complete sufficient statistics, so they will be UMVUE.

3 Mixed models

As stated in the Introduction, we now consider normal mixed models $y \sim \mathcal{N}\left(X_0\beta_0; \sum_{i=1}^w \theta_i M_i\right)$, where the matrices of $\mathcal{M} = \{M_1, \dots, M_w\}$ commute.

The matrices M_1, \dots, M_w are symmetric so they commute if and only if they are diagonalized by the same orthogonal projection matrix, see [9]. Thus they belong to the family of matrices diagonalized by P of symmetric matrices that commute. Moreover $\mathcal{A}(P)$ will contain the squares of its matrices, thus being a commutative Jordan algebra (CJA) of symmetric matrices. Each such algebra \mathcal{A} has, see [10], an unique basis constituted by orthogonal projections matrices, that are pairwise orthogonal, the principal basis of \mathcal{A} , $pb(\mathcal{A})$. With $pb(\mathcal{A}) = \{K_1, \dots, K_m\}$, we will have

$$M_i = \sum_{j=1}^m b_{i,j} K_j \quad i = 1, \dots, w$$

thus, with V the variance-covariance matrix of the model, we have

$$V = \sum_{i=1}^w \theta_i M_i = \sum_{j=1}^m \gamma_j K_j$$

with

$$\gamma_j = \sum_{i=1}^w b_{i,j} \theta_i$$

so that $\gamma \in \mathcal{R}(B^\top)_+$, for the model to have OBS. Since w is the number of random vectors and m the number of sub-models, or strata, used to carry out the

inference, we see that this equality between both numbers is necessary for having OBS. Again, considering the Example 1 and 2 in [4], for better comparison, we see that they do not have OBS, unless we discard the requirement that all positive definite matrices $\sum_{j=1}^m \gamma_j K_j$ may be variance-covariance matrices.

Moreover, when $w < m$, the condition $\gamma \in \mathcal{R}(B^\top)_+$ implies the existence of linear restrictions between the components of γ so, the density cannot have full rank, see [5]. Thus the requirement of B being invertible is highly connected with the model having OBS and, if normality is assumed, with having complete and sufficient statistics. Actually this condition ($w = m$) follows, as we showed, from the requirement of the model having OBS as defined in [7], [8].

Thus for there being no restriction on the $\gamma_1, \dots, \gamma_m$, $B = [b_{i,j}]$ must be invertible. Reasoning as for establishing theorem 1, we get the following theorem.

Theorem 2

A normal mixed model such that its matrices M_1, \dots, M_w commute, has complete and sufficient statistics if and only if matrix B is invertible.

We also may establish the following theorem.

Theorem 3

A normal mixed model such that its matrices M_1, \dots, M_w commute is NOBS, if and only if matrix B is invertible.

Proof: The variance-covariance matrices may be all positive semi-definite matrices, $\sum_{j=1}^m \gamma_j K_j$, if and only if there are no linear restrictions on the $\gamma_1, \dots, \gamma_m$ and, since the vector $\gamma \in \mathcal{R}(B^\top)$, this happens if and only if matrix B is invertible. \square

Now $\Psi = G\beta$ is estimable if and only if $\Psi = U\mu = UX_0\beta_0$, this happening if

$$G = UX_0,$$

and we have, when B is invertible and thus are complete sufficient statistics, the UMVUE

$$\tilde{\Psi} = U\tilde{\mu} = U \sum_{j=1}^z A_j^\top W_j \tilde{\eta}_j.$$

This result is very similar to the ones in the preceding section. Where there is something new is on the estimation of the variance-covariance components

$\theta_1, \dots, \theta_m$, since if B is invertible, we must have $w = m$. Putting

$$C = [B^T]^{-1} = [c_{\ell,h}], \quad \ell = 1, \dots, m; h = 1, \dots, m,$$

from

$$\gamma = B^T \theta,$$

we get

$$\theta = C\gamma.$$

If

$$\theta_\ell = \sum_{h \in \mathcal{D}} c_{\ell,h} \gamma_h, \quad \ell = 1, \dots, m,$$

we get the UMVUE,

$$\tilde{\theta}_\ell = \sum_{h \in \mathcal{D}} c_{\ell,h} \tilde{\gamma}_h, \quad \ell = 1, \dots, m.$$

4 An example

We now assume the model to have a random effect factor that cross with a_1, \dots, a_u levels, so there will be $n = \prod_{h=1}^u a_h$ observations, $m = w = 2^u$ terms in the random effects part associated to the sets \mathcal{X} of factors. If $\#(\mathcal{X}) = 1 [> 1]$ the term will be associated to the factor [factors] with index [indexes] in \mathcal{X} . These sets may be ordered according to the indices

$$i(\mathcal{X}) = 1 + \sum_{h \in \mathcal{X}} 2^{h-1} = 1, \dots, w,$$

so the model can be written as

$$y = \sum_{i=0}^w X_i \beta_i$$

where β_0 is fixed and the β_1, \dots, β_w are independent, with null mean vectors and variance-covariance matrices $\theta_1 I_{c_1}, \dots, \theta_w I_{c_w}$.

We point out that the fixed effects part $X_0 \beta_0$, where X_0 is a $n \times k$ matrix, is quite fluid since the only restriction to be applied to it is that the column vectors of X_0 are linearly independent.

This case has been studied thus, see for instance [2],

$$X_i = \bigotimes_{h=1}^u X_{i,h} \quad i = 1, \dots, w,$$

where \otimes indicates the Kronecker matrix product, and, with \mathcal{X}_i the set with index $i, i = 1, \dots, w$, we have

$$X_{i,h} = \begin{cases} 1_{a_h} & h \notin \mathcal{X}_i \\ I_{a_h} & h \in \mathcal{X}_i \end{cases}$$

for $h = 1, \dots, u; i = 1, \dots, m$.

It is easy to show that the matrices $M_i = X_i X_i^T, i = 1, \dots, u$, commute, moreover, see [6], we have the transition matrix

$$B = \begin{bmatrix} a_u & 0 \\ 1 & 1 \end{bmatrix} \otimes \dots \otimes \begin{bmatrix} a_1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} n & 0^T \\ \mathbf{b} & B^\circ \end{bmatrix}$$

and the orthogonal projection matrices that are pairwise orthogonal, for this model, are

$$K_j = \bigotimes_{h=1}^u K_{j,h} \quad j = 1, \dots, m$$

where

$$K_{j,h} = \begin{cases} \frac{1}{a_h} J_{a_h} & h \notin \mathcal{X}_j \\ I_{a_h} - \frac{1}{a_h} J_{a_h} & h \in \mathcal{X}_j \end{cases}$$

for $h = 1, \dots, u; j = 1, \dots, m$, with $J_a = 1_a 1_a^T$.

Since the Kronecker matrix product of invertible matrices results in invertible matrices, the transition matrix B will be invertible, as well as B°, B^T and $B^{\circ T}$, so this models will be NOBS.

It is important to point out that

$$K_1 = \bigotimes_{i=1}^u \left(\frac{1}{a_1} J_{a_1} \right)$$

so $p_1 = g_1 = 1$ and $q_1 = 0$, so $1 \notin \mathcal{D}$. Thus despite there being complete and sufficient statistics, we do not have unbiased estimators for γ_1 . Putting $\gamma^\circ = (\gamma_2, \dots, \gamma_m)$ and $\theta^\circ = (\theta_2, \dots, \theta_m)$, we have

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

and so, when $\{2, \dots, m\} = \mathcal{D}$, we will have the UMVUE

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

where the components of $\tilde{\gamma}^\circ$ are obtained as before.

Moreover, see [6],

$$A_j = \bigotimes_{h=1}^u A_{j,h} \quad j = 1, \dots, m$$

with

$$A_{j,h} = \begin{cases} \frac{1}{\sqrt{a_h}} 1_{a_h}^T & h \notin \mathcal{X}_j \\ F_{a_h} & h \in \mathcal{X}_j \end{cases}$$

for $h = 1, \dots, u; j = 1, \dots, m$, where F_a is obtained deleting the first row equal to $\frac{1}{\sqrt{a_h}}, \dots, \frac{1}{\sqrt{a_h}}$ of a $a_h \times a_h$ orthogonal matrix P_{a_h} , so that

$$\begin{aligned} I_{a_h} &= P_{a_h}^\top P_{a_h} \\ &= \begin{bmatrix} \frac{1}{\sqrt{a_h}} \mathbf{1}_{a_h} & F_{a_h}^\top \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{a_h}} \mathbf{1}_{a_h} & F_{a_h}^\top \end{bmatrix}^\top \\ &= \frac{1}{a_h} \mathbf{J}_{a_h} + F_{a_h}^\top F_{a_h} \end{aligned}$$

and so

$$F_{a_h}^\top F_{a_h} = I_{a_h} - \frac{1}{a_h} \mathbf{J}_{a_h},$$

so that

$$g_j = \prod_{h \in \mathcal{X}_j} (a_h - 1).$$

We can now reorder the terms to have $p_j > 0$ if and only if $j \neq z$, in order to obtain the

$$\begin{cases} y_j = A_j y & j = 1, \dots, z \\ X_{0,j} = A_j X_0 & j = 1, \dots, z \end{cases}$$

and the orthogonal projection matrices

$$P_j = W_j^\top W_j \quad j = 1, \dots, z$$

on the $\mathcal{R}(X_{0,j}), j = 1, \dots, z$, as well as the

$$\tilde{\eta}_j = W_j y_j \quad j = 1, \dots, z$$

in order to get the

$$\tilde{\Psi} = U \sum_{j=1}^z A_j^\top W_j \tilde{\eta}_j.$$

We point out that the fact that X_0 may be any matrix with linearly independent k column vectors renders difficult the obtainment of more detailed general results. Thus we opted for showing how to obtain the matrices specific for this analysis, namely the matrices K_j and $A_j, j = 1, \dots, m$.

5 Final comments

We have shown how assuming normality leads to optimal results for models with OBS. Thus NOBS emerges as an important class of models with OBS. Our approach is quite distinct from other that rest, as is the case of OBS, on the algebraic structure of the models. This clearly can be seen in connection with estimable vectors. Thus see [12], when T , the orthogonal projection matrix on the space spanned by the

mean vector, commutes with V , the least squares estimators of estimable vectors are best linear unbiased estimators. Now, assuming normality, leads to an interesting result without no requirement on matrix T .

The requirement of all positive semi-definite matrices given by $\sum_{j=1}^m \gamma_j K_j$ being possible variance-covariance matrices is also considered for orthogonal models, see [11]. These are OBS models in which the orthogonal projection matrix T commutes with the known orthogonal projection matrices K_1, \dots, K_m that are pairwise orthogonal. Thus if we assume normality for an orthogonal model it will have complete sufficient statistics as well as UMVUE for the relevant parameters. Thus the normal orthogonal models constitute in themselves an interesting class of models which may deserve to be studied.

Similar studies can be extended to an also interesting class of models with OBS, the ones with Commutative Orthogonal Block Structures see *e.g.* [3], and their relation with error orthogonal models, see [1].

Acknowledgements: This work was partially supported by CMA / FCT / UNL, under the project PEst-OE/MAT/UI0297/2014.

References:

- [1] Carvalho, Francisco; Mexia, João T.; Santos, Carla (2013). Commutative orthogonal block structure and error orthogonal models. *Electronic Journal of Linear Algebra*. **25**, 119–128.
- [2] Fonseca, M.; Mexia, J. T. and Zmysłony, R. (2006). Binary operations on Jordan algebras and orthogonal normal models. *Linear Algebra and its Applications* **417**, pp. 75–86.
- [3] Fonseca, M.; Mexia, J. T.; Zmysłony, R. (2008). Inference in normal models with commutative orthogonal block structure. *Acta et Commentationes Universitatis Tartunensis de Mathematica* **12**, pp. 3–16.
- [4] Houtman, A.M. and Speed, T.P. (1983). Balance in Designed Experiments with Orthogonal Block Structure. *Annals of Statistics*. **11** 4, pp. 1069–1085.
- [5] Lehmann, E.L. and Casella, G. (1998). *Theory of point estimation*. Springer.
- [6] Mexia, J. T.; Vaquinhas, R.; Fonseca, M. and Zmysłony, R. (2010). COBS: segregation, matching, crossing and nesting, *Latest Trends and Applied Mathematics, simulation, Modelling, 4-th International Conference on Applied Mathematics, Simulation, Modelling (ASM'10)*, pp. 249–255.

- [7] Nelder, J.A. (1965a). The Analysis of Randomized Experiments with Orthogonal Block Structure. I - Block Structure and the Null Analysis of Variance. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, **283** (1393), pp. 147–162.
- [8] Nelder, J.A. (1965b). 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), pp. 163–17.
- [9] Schott, J. R. (1997). *Matrix Analysis for Statistics*, Wiley Series in Probability and Statistics. John Wiley.
- [10] Seely, J. (1971). Quadratic subspaces and completeness. *The Annals of Mathematical Statistics* **42** (2), pp. 710–721.
- [11] VanLeeuwen, Dawn M.; Birks, David S.; Seely, Justus F.; J. Mills, J.; Greenwood, J. A. and Jones, C. W. (1998). Sufficient conditions for orthogonal designs in mixed linear models. *Journal of Statistical Planning and Inference* **73**, pp. 373–389.
- [12] Zmysłony, R. (1980). A characterization of Best Linear Unbiased Estimators in the general linear model. *Lecture Notes in Statistics* **2**, pp. 365–373.