Journal of Interdisciplinary Mathematics

ISSN: 0972-0502 (Print) 2169-012X (Online) Journal homepage: https://www.tandfonline.com/loi/tjim20

Inference for *L* orthogonal models

Sandra Saraiva Ferreira , Dário Ferreira , Elsa Moreira & João Tiago Mexia

Inference for L orthogonal models

Sandra Saraiva Ferreira *
Dário Ferreira
Department of Mathematics
University of Beira Interior
Av. Marquês d' Ávila e Bolama
6200 Covilhã
Portugal

Elsa Moreira
João Tiago Mexia
Department of Mathematics
Faculty of Science and Technology
New University of Lisbon
2829 Caparica
Portugal

Abstract

A mixed model $\underline{Y}^o = \sum_{i=1}^m X_i\underline{\beta}_i + \sum_{i=m+1}^w X_i\underline{\widetilde{\beta}}_i + e$ is orthogonal when the matrices $M_i = X_iX_i^T$, $i = 1,\dots,w$, commute. The vectors $\beta_1^{c_1},\dots,\beta_m^{c_m}$ are fixed vectors and the $\beta_{m+1}^{c_{m+1}},\dots,\beta_w^{c_w}$ and e^n are random. For these models we have very interesting results namely we have UMVUE for the relevant parameters when normality is assumed. We now intend to generalize that class of models taking $\underline{Y} = L\big(\sum_{i=1}^m X_i\underline{\beta}_i + \sum_{i=m+1}^w X_i\underline{\widetilde{\beta}}_i\big) + e$ with L a matrix whose column vectors are linearly independent.

Keywords and phrases : *Normal orthogonal models, commutative Jordan algebras, mixed model.*

1. Introduction

An L model is given by

$$\underline{Y} = L\underline{Y}^o + \underline{e}\,, \tag{1.1}$$

* *E-mail*: sandra@mat.ubi.pt

Journal of Interdisciplinary Mathematics
Vol. 12 (2009), No. 6, pp. 815–824

where

$$\underline{Y}^o = \sum_{i=1}^{m} X_i \underline{\beta}_i + \sum_{i=m+1}^{w} X_i \underline{\widetilde{\beta}}_i \tag{1.2}$$

in the core model and $\underline{e}$ is the error vector. The vectors $\beta_1^{c_1}, \ldots, \beta_m^{c_m}$ are fixed vectors and the $\beta_{m+1}^{c_{m+1}}, \ldots, \beta_w^{c_w}$ and e^n are independent, with null mean vectors and covariance matrices $\sigma_i^2 I_{c_i}$, $i = m+1, \ldots, w$ and $\sigma^2 I_n$. Then Y^o has mean vector and covariance matrix

$$u^o = \sum_{i=1}^{m} X_i \underline{\beta}_i \tag{1.3}$$

and

$$V^o = \sum_{i=m+1}^{w} \sigma_i^2 M_i + \sigma^2 L^+ L^{+^\top}. \tag{1.4}$$

We will assume that the column vectors of matrix L are independent and that the $\dot{n}$ components of core model correspond to the treatments. For instance

$$L = D(1_{r_1}, \ldots, 1_{r_{\dot{n}}}) \tag{1.5}$$

is the blockwise diagonal matrix with as principal blocks the vectors whose $r_1, \ldots, r_{\dot{n}}$ components are equal to one. We will use a mixed model with those number of observations for the different treatments.

We will be interested in the case in which the core model is orthogonal associated to a commutative Jordan algebra, CJA, $\mathcal{A}$. Then, see [1], the matrices $M_i = X_i X_i^T$, $i = 1, \ldots, w$, commute and constitute a basis for $\mathcal{A}$.

In the next section we will consider CJA presenting certain results that will be useful later on. Then we study the structure of the core model before presenting inference.

2. Commutative Jordan algebras and core models

Commutative Jordan algebra, CJA, are linear spaces constituted by symmetric matrices that commute and containing the squares of its matrices. These structures were introduced by [3], in a generalized formulation for quantum mechanics. Later on, CJA were undiscovered and used in linear statistical inference by [10], [11], [12] and [14], among others like [13] and by [15] and [16]. Namely, [12] showed that for each CJA, $\mathcal{A}$, there is one and only one basis, the principal basis, $pb(\mathcal{A})$, constituted by pairwise orthogonal projection matrices which are orthogonal. An example of CJA is the family $\mathcal{V}(\mathcal{P})$ of symmetric matrices diagonalized by P.

The symmetric matrices $M_1, \dots, M_u$ commute, see [9], if and only if they are diagonalized by the same orthogonal matrix P thus belonging to $\mathcal{V}(\mathcal{P})$. Thus a family of symmetric matrices is contained in a CJA if and only if its matrices commute.

If $M \in \mathcal{A}$ and $pb(\mathcal{A}) = \{Q_1, \dots, Q_k\}$ we have

$$M = \sum_{j=1}^{k} a_j Q_j = \sum_{j \in \varphi(M)} a_j Q_j \,, \tag{2.1}$$

with $\varphi(M) = \{j : a_j \neq 0\}$. It is easy to see that the orthogonal projection matrix on the range space $R(M)$ of M is

$$Q(M) = \sum_{j \in \varphi(M)} Q_j. \tag{2.2}$$

Namely if Q is an orthogonal projection matrix belonging to $\mathcal{A}$, we saw that $Q = \sum_{j \in \varphi(Q)} Q_j$, since Q is the orthogonal projection matrix on $R(Q)$, thus $\operatorname{rank}(Q) = \sum_{j \in \varphi(Q)} \operatorname{rank}(Q_j)$ and so, if $\operatorname{rank}(Q) = 1$, $Q \in pb(\mathcal{A})$. If $\frac{1}{n} J_n = \frac{1}{n} 1_n 1_n^\top \in Q$ we will have $\frac{1}{n} J_n \in pb(\mathcal{A})$ and we will then put $Q_1 = \frac{1}{n} J_n$ and say that $\mathcal{A}$ is regular. The core model

$$\underline{Y}^o = \sum_{i=1}^{m} X_i \underline{\beta}_i + \sum_{i=m+1}^{w} X_i \widetilde{\beta}_i \tag{2.3}$$

is orthogonal, ORT, if matrices $M_i = X_i X_i^\top$, $i = 1, \dots, w$ commute. Let then $\{Q_1, \dots, Q_k\}$ be the principal basis of a CJA containing $M_1, \dots, M_w$. If we assume that the $\beta_1, \dots, \beta_m$ are fixed and that the $\widetilde{\beta}_{m+1}, \dots, \widetilde{\beta}_w$ have null mean vectors and variance-covariance matrices $\sigma^2_{m+1} I_{c_{m+1}}, \dots, \sigma^2_w I_{c_w}$, $\underline{Y}^o$ will have mean vector

$$\underline{u}^o = \sum_{i=1}^{m} X_i \underline{\beta}_i \tag{2.4}$$

and variance covariance matrices

$$V^o = \sum_{l=m+1}^{w} \sigma_i^2 M_i \quad \text{with} \quad M_i = X_i X_i^\top, \ i = 1, \dots, w \,. \tag{2.5}$$

Now

$$M_i = \sum_{j=1}^{k} b_{i,j} Q_j \,, \quad i = 1, \dots, w, \tag{2.6}$$

so

$$V^o = \sum_{j=1}^{k} \gamma_j Q_j \tag{2.7}$$

with

$$\gamma_j = \sum_{i=m+1}^{w} b_{i,j}\sigma^2, \quad j = 1, \ldots, k. \tag{2.8}$$

Moreover $\underline{u}^o$ will span

$$\Omega^o = R([X_1 \ldots X_m]) = \left(\sum_{i=1}^{m} M_i\right). \tag{2.9}$$

Now $M = \sum_{i=1}^{m} M_i$ and $T = Q(M)$ will belong to $\mathcal{A}$. We can reorder the matrices in $pb(\mathcal{A})$ to have

$$T = \sum_{j=1}^{h} Q_j \,. \tag{2.10}$$

Since $R(M_i) \subseteq R(T)$, $i = 1, \ldots, m$ we will have $b_{i,j} = 0$, $j = h+1, \ldots, k$, $i = 1, \ldots, m$. Then the matrix $B = [b_{i,j}]$ will be

$$B = \begin{bmatrix} B_{1,1} & 0 \\ B_{2,1} & B_{2,2} \end{bmatrix} \tag{2.11}$$

with $B_{1,1}$ a matrix of type $m \times h$, $B_{2,1}$ a matrix of type $(w-m) \times h$ and $B_{2,2}$ a matrix of type $(w-m) \times (k-h)$. 0 is a null matrix of type $m \times (k-h)$.

Putting $\underline{\sigma}^2 = (\sigma^2_{m+1}, \ldots, \sigma^2_w)$, $\underline{\gamma}_1 = (\gamma_1, \ldots, \gamma_h)$ and $\underline{\gamma}_2 = (\gamma_{h+1}, \ldots, \gamma_k)$ we have

$$\begin{cases} \underline{\gamma}_{(1)} = B^\top_{2,1}\underline{\sigma}^2 \\ \underline{\gamma}_{(2)} = B^\top_{2,2}\underline{\sigma}^2 \,. \end{cases} \tag{2.12}$$

Then, if the row vectors of $B_{2,2}$ are linearly independent the column vectors of $B^\top_{2,2}$ will be linearly independent and we have, see [7],

$$\begin{cases} \underline{\sigma}^2 = (B^\top_{2,2})^\top \underline{\gamma}(2) \\ \underline{\gamma}_{(1)} = B^\top_{2,1}(B^\top_{2,2})^\top \underline{\gamma}(2) \end{cases} \tag{2.13}$$

with C^+ the Moore-Penrose inverse of C, see [8]. Thus the parameters of $\underline{\gamma}(2)$ and $\underline{\sigma}^2$ of random effects part of the core model will determine each other. Then the random effects part of the core model segregates itself as a separate sub-model. We thus say that these core models are segregated.

3. Estimation

The orthogonal projection matrices on $\Omega = R(L)$ will be $Q(\Omega) = LL^+$. Let now the row vectors of K constitute an orthonormal basis for the orthogonal complement $\Omega^{\perp}$ of Ω, then, with n the number of components of $\underline{Y}$ we will have

$$Q(\Omega^{\perp}) = I_n - Q(\Omega) = K^{\top}K . \tag{3.1}$$

The orthogonal projections $\underline{Y}_{\Omega}$ and $\underline{Y}_{\Omega}^{\perp}$ of $\underline{Y}$ in Ω and $\Omega^{\perp}$ will be

$$\underline{Y}_{\Omega} = L\underline{Y}^{o} + \underline{e}_{\Omega} \tag{3.2}$$

and

$$\underline{Y}_{\Omega}^{\perp} = \underline{e}_{\Omega}^{\perp} , \tag{3.3}$$

thus, if we assume that cross-variance matrices

$$\not{\Sigma}(\widetilde{\beta}_i, \underline{e}), \quad i = m+1, \ldots, w, \tag{3.4}$$

we will have

$$\not{\Sigma}(\underline{Y}_{\Omega}, \underline{Y}_{\Omega}^{\perp}) = 0_{n\times n} \tag{3.5}$$

since

$$\not{\Sigma}(\underline{e}_{\Omega}, \underline{e}_{\Omega}^{\perp}) = 0_{n\times n} . \tag{3.6}$$

Thus, taking $S = \|\underline{Y}_{\Omega}^{\perp}\|^2 = \|\underline{e}_{\Omega}^{\perp}\|^2$ we will have $E(S) = g\sigma^2$ with $g = n - n^{o}$ where $n^{o} = \text{rank}(L)$. Moreover, since the column vectors of L are linearly independent we have $L^{+}L = I_{n},$ so that

$$\begin{aligned} \underline{Z} \;&=\; L^{+}\underline{Y} = L^{+}LL^{+}\underline{Y} = L^{+}\underline{Y}_{\Omega} = \underline{Y}^{o} + L^{+}\underline{e}_{\Omega} && (3.7)\\ &=\; \underline{Y}^{o} + L^{+}LL^{+}\underline{e} = \underline{Y}^{o} + L^{+}\underline{e} , && (3.8) \end{aligned}$$

with $L^{+}\underline{e}$ independent from $\underline{Y}^{o}$, so $\underline{Z}$ will have mean vector $\underline{u}^{o}$ and variance covariance matrix

$$V^{o} + \sigma^2 L^{+}L^{+^{\top}} . \tag{3.9}$$

Let now the g_j row vectors of A_j, $j = 1, \ldots, k$ constitute an orthonormal basis for $R(Q_j)$, $j = 1, \ldots, k$. Taking

$$\begin{cases} \underline{\eta}_j = A_j \underline{u}^{o} , \\ \widetilde{\underline{\eta}}_j = A_j \underline{Z} , \end{cases} \quad j = 1, \ldots, k, \tag{3.10}$$

since $\underline{u}$ spans ω and $T = \sum_{j=1}^{h} Q_j$ is the orthogonal projection matrix on ω, we will have $\underline{\eta}_j = \underline{0}, j = h+1, \ldots, k$. Moreover since the $Q_1, \ldots, Q_k$ are pairwise orthogonal, the variance-covariance matrix of $\widetilde{\underline{\eta}}_j$ will be $\gamma_j I_{g_j} + \sigma^2 A_j L^{+}L^{+^{\top}} A_j^{\top}$, thus, taking

$$S_j = \|\widetilde{\underline{\eta}}_j\|^2 , \quad j = 1, \ldots, k, \tag{3.11}$$

we have

$$E(S_j) = g_j\gamma_j + t_j\sigma^2\,, \quad j = m+1, \dots, w, \tag{3.12}$$

see [9], with $t_j = \mathrm{trace}(A_j L^+ L^{+^\top} A_j^\top)$, $j = 1, \dots, k$ and $g_j = \mathrm{rank}(Q_j)$, $j = 1, \dots, k$. Thus we have the unbiased estimators

$$\widetilde{\gamma}_j = \frac{S_j}{g_j} - \frac{t_j}{g_j}\sigma^2, \quad j = m+1, \dots, k \tag{3.13}$$

for the components of $\underline{\gamma}(2)$. If the core model are segregated we also have the unbiased estimators

$$\begin{cases} \underline{\widetilde{\sigma}}^2 = (B_{2,2}^\top)^\top \underline{\widetilde{\gamma}}(2) \\ \underline{\widetilde{\gamma}}_{(1)} = B_{2,1}^\top (B_{2,2}^\top)^\top \underline{\widetilde{\gamma}}(2)\,. \end{cases} \tag{3.14}$$

As to estimable vectors, putting

$$\begin{cases} X = [X_1 \dots X_m] \\ \underline{\beta} = [\underline{\beta_1}^\top \dots \underline{\beta_m}^\top]^\top \end{cases} \tag{3.15}$$

we have

$$\underline{u}^o = X\underline{\beta} \tag{3.16}$$

and so the mean vector of $\underline{Y}$ will be

$$\underline{u} = L\underline{u}^o = LX\underline{\beta}\,. \tag{3.17}$$

Given a matrix G, $G\underline{Y}$ and $G\underline{Y}_\omega$ will have the same mean vector since both $\underline{Y}$ and $\underline{Y}_\omega$ have mean vector $\underline{u}$. Now

$$\underline{\psi} = G\underline{\beta} \tag{3.18}$$

is estimable if there is a linear unbiased estimator $\underline{\psi}^o$ for $\underline{\psi}$.

The mean vector of $U\underline{Y}$ is $ULX\underline{\beta}$, and $U\underline{Y}$ is an unbiased estimator of $\underline{\psi}$ if

$$ULX = G, \tag{3.19}$$

this is if

$$\underline{\psi} = G\underline{\beta} = UL\underline{u}^o. \tag{3.20}$$

Since

$$\underline{u}^o = T\underline{u}^o = \sum_{j=1}^{h} Q_j \underline{u}^o = \sum_{j=1}^{h} A_j^\top A_j \underline{u}^o = \sum_{j=1}^{h} A_j^\top \underline{\eta}_j \tag{3.21}$$

we have

$$\underline{\psi} = UL \sum_{j=1}^{h} A_j^\top \underline{\eta}_j \,, \tag{3.22}$$

thus

$$\underline{\psi}^* = UL \sum_{j=1}^{h} A_j^\top \underline{\widetilde{\eta}}_j \tag{3.23}$$

will be an unbiased estimator of $\underline{\psi}$.

4. The orthogonal case

Let us assume that $L^+ = L^\top$, then matrices

$$\overline{Q}_j = LQ_jL^\top, \quad j = 1, \ldots, w, \tag{4.1}$$

will be pairwise orthogonal projection matrices which are orthogonal. Moreover

$$\sum_{j=1}^{w} \overline{Q}_j = LL^\top = Q(\Omega) \tag{4.2}$$

with $\Omega = R(L)$. Thus adding to the $\overline{Q}_j$, $j = 1, \ldots, w$ the matrix

$$Q^\perp = I_n - LL^\top \tag{4.3}$$

we get the principal basis of a complete CJA, $\overline{\mathcal{A}}$.

If the row vectors of A_j constitute an orthogonal basis for $R(Q_j)$, $j = 1, \ldots, w$ we have $A_j^\top A_j = Q_j$ and $A_jA_j^\top = I_{g_j}$, with $g_j = \text{rank}(Q_j)$ $= \text{rank}(\overline{Q}_j)$, $j = 1, \ldots, w$. Then, with

$$\overline{A}_j = A_jL^\top, \quad j = 1, \ldots, w \tag{4.4}$$

we will have

$$\begin{cases} \overline{Q}_j = \overline{A}_j^\top \overline{A}_j \\ I_{g_j} = \overline{A}_j \overline{A}_j^\top \,, \end{cases} \quad j = 1, \ldots, w, \tag{4.5}$$

and the row vectors of $\overline{A}_j$ will constitute an orthogonal basis for $R(\overline{Q}_j)$. Thus, since $L^+L = I_n$,

$$\begin{cases} \overline{A}_j\underline{u} = A_jL^\top L\underline{u}^o = A_j\underline{u}^o = \underline{\eta}_j \\ \overline{A}_j\underline{Y} = A_jL^\top L\underline{Z} = A_j\underline{Z} = \underline{\widetilde{\eta}}_j \,, \end{cases} \quad j = 1, \ldots, w. \tag{4.6}$$

Besides this the variance-covariance matrix of $\underline{Y}$ will be

$$V = L\Big(\sum_{j=1}^{w}\gamma_j Q_j\Big)L^\top + \sigma^2 I_n = \sum_{j=1}^{w}\gamma_j \overline{Q}_j + \sigma^2 I_n$$
$$= \sum_{j=1}^{w}(\gamma_j+\sigma^2)\overline{Q}_j + \sigma^2 Q^\perp, \qquad (4.7)$$

so

$$\begin{cases} V^{-1} = \sum_{j=1}^{w}(\gamma_j+\sigma^2)^{-1}\overline{Q}_j + \frac{1}{\sigma^2}Q^\perp \\ \det(V) = \prod_{j=1}^{w}(\gamma_j+\sigma^2)^{g_j}(\sigma^2)^{g}, \end{cases} \qquad (4.8)$$

with $g = n - n^o$. Thus

$$(\underline{y}-\underline{u})^\top V^{-1}(\underline{y}-\underline{u}) = \sum_{j=1}^{w}\frac{(\underline{y}-\underline{u})^\top \overline{A}^\top \overline{A}(\underline{y}-\underline{u})}{(\gamma_j+\sigma^2)} + \frac{\underline{y}^\top Q^\perp \underline{y}}{\sigma^2}$$
$$= \sum_{j=1}^{w}\frac{\|\widetilde{\underline{\eta}}_j - \underline{\eta}_j\|^2}{\gamma_j+\sigma^2} + \sum_{j=m+1}^{w}\frac{S_j}{\gamma_j+\sigma^2} + \frac{S}{\sigma^2}, \qquad (4.9)$$

where, as before, $S_j = \|\widetilde{\underline{\eta}}_j\|^2$, $j = m+1,\dots,w$, and

$$S = \underline{y}^\top Q^\perp \underline{y} = \|\underline{y}_{\Omega^\perp}\|^2. \qquad (4.10)$$

Thus when normality is assumed $\underline{Y}$ will have the density

$$n(\underline{y}) = \frac{e^{-\frac{1}{2}\left[\sum_{j=1}^{m}\frac{\|\widetilde{\underline{\eta}}_j-\underline{\eta}_j\|^2}{\gamma_j+\sigma^2} + \sum_{j=m+1}^{w}\frac{S_j}{\gamma_j+\sigma^2} + \frac{S}{\sigma^2}\right]}}{(2\pi)^{\frac{n}{2}}\prod_{j=1}^{w}(\gamma_j+\sigma^2)^{\frac{g_j}{2}}\sigma^{g}}. \qquad (4.11)$$

We now have

Proposition 1. *The* $\widetilde{\underline{\eta}}_j$, $j = 1,\dots,m$, *the* S_j, $j = m+1,\dots,w$ *and* S *will be sufficient complete statistics.*

Proof. Applying the Factorization theorem we see that the statistics are sufficient. Moreover they will be complete, see [5], since the normal distribution belongs to the exponential family and, for these models, the parameter space contains the cartesian product of non degenerate intervals. □

Corollary 2. *The* $\widetilde{\underline{\psi}}_j$, *the* $\widetilde{\underline{\gamma}}(1)$, $\widetilde{\underline{\gamma}}(2)$ *and* $\widetilde{\underline{\sigma}}^2$ *will be UMVUE.*

Proof. The thesis follows from the proposition and the Blackwell-Lehman-Scheffé theorem. □

References

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

[2] M. Fonseca, J.T. Mexia and R. Zmyślony, Inference in normal models with commutative orthogonal block structure, accepted for publication in *Acta et Commentationes Universitatis Tartuences de Mathematica* (in press) (2008).

[3] P. Jordan, J. Von Neumann and E. Wigner, On a algebraic generalization of the quantum mechanical formulation, *Ann. of Math.*, Vol. 35 (1934), pp. 29–64.

[4] A. I. Khuri, T. Matthew and B. K. Sinha, Statistical tests for mixed linear models, *Wiley Series in Probability and Statistics*, Jonh Wiley & Sons, New York, 1997.

[5] E. L. Lehman, *Testing Statistical Hypothesis*, reprint of 2nd edition, Wiley, 1997.

[6] J. T. Mexia, Best linear unbiased estimates duality of F tests and the Scheffé multiple comparison method in presence of controlled heterodecedasticity, *Comp. Stat. & Data Analysis*, Vol. 10 (1990), p. 3.

[7] J. T. Mexia, *Introdução à inferência estatística linear*, Centro de Estudos de Matemática Aplicada. Edições Universitárias Lusófonas - Lisboa, 1995.

[8] D. S. G. Pollock, *The Algebra of Econometrics*, Jonh Wiley & Sons, New York, 1979.

[9] J. R. Schott, *Matrix Analysis for Statistics*, Jonh Wiley & Sons, New York, 1997.

[10] J. Seely, Linear spaces and unbiased estimators, *Ann. Math. Statist.*, Vol. 41 (1970), pp. 1725–1745.

[11] J. Seely, Linear spaces and unbiased estimators. Application to a mixed linear model, *Ann. Math. Statist.*, Vol. 41 (1970), pp. 1735–1748.

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

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

[14] J. Seely, Minimal sufficient statistics and completeness for multivariate normal families, *Sankhya*, Vol. 39 (1977), pp. 170–185.

[15] D. Van Leeuwen, J. Seely and D. Birkes, Sufficient conditions for orthogonal designs in mixed linear models, *J. Statist. Plann. Inference*, Vol. 73 (1998), pp. 373–389.

[16] D. Van Leeuwen, D. Birkes and J. Seely, Balance and orthogonality in designs for mixed classification models, *Ann. Statist.*, Vol. 27 (1999) , pp. 1927–1947.

Received April, 2009