Discussiones Mathematicae
Probability and Statistics 28 (2008) 171–181

CANONIC INFERENCE AND COMMUTATIVE ORTHOGONAL BLOCK STRUCTURE

Francisco P. Carvalho,

Department of Mathematics, Management School
Polytechnic Institute of Tomar, Portugal

e-mail: fpcarvalho@ipt.pt

João Tiago Mexia

Mathematics Department, Faculty of Science and Technology
New University of Lisbon
Monte da Caparica 2829–516 Caparica, Portugal

e-mail: jtm@fct.unl.pt

and

M. Manuela Oliveira

Department of Mathematics, Évora University, Portugal

e-mail: mmo@uevora.pt

Abstract

It is shown how to define the canonic formulation for orthogonal models associated to commutative Jordan algebras. This canonic formulation is then used to carry out inference. The case of models with commutative orthogonal block structures is stressed out.

Keywords: COBS, canonical inference, commutative Jordan algebras.

2000 Mathematics Subject Classification: 62J05.

1. Introduction

A model has Orthogonal Block Structure (OBS), if its observation vector has covariance matrix given by

$$\mathscr{V}(y) = \sum_{j=1}^{w} \gamma_j P_j \tag{1}$$

with $P_1, \dots, P_w$ known mutually orthogonal orthogonal projection matrices. The mean vector will be

$$\mu = X\beta. \tag{2}$$

This mean vector spans the range space $\Omega = R(X)$ of X. The orthogonal projection matrix over this space image is given by (see Mexia, 1999),

$$T = X(X^T X)^+ X^T = X(X^T X)^- X^T, \tag{3}$$

where A^+, A^- stands for the Moore-Penrose inverse matrix, and for general inverse, respectively.

In the case that T commutes with the $P_1, \dots, P_w$, the model will have Commutative Orthogonal Block Structures (COBS).

In this work we'll present, for the normal case, the canonic formulation of such models and use it to preform inference.

2. Commutative Jordan algebras

Commutative Jordan algebras are vectorial spaces constituted by symmetric matrices that commute and contains their squares.

Seely (1971), shows that all of these algebras have one and only one base made of mutually orthogonal orthogonal projection matrices. This will be the principal base of the algebra.

In our case we have two commutative Jordan algebras, $\mathscr{A}(\mathbf{P})$ and $\mathscr{A}(\mathbf{T})$ with principal bases $\{P_1, \dots, P_w\}$ and $\{T, I - T\}$.

Besides this, see Schott (1997, pag. 157), the symmetric matrices $M_1, \dots, M_u$ commute if and only if they are diagonalized by the same orthogonal matrix K. Therefore they will belong to the commutative Jordan algebra $\mathscr{A}(K)$ of the matrices diagonalized by K. So, $P_1, \dots, P_w$, T and $I - T$ belongs to a commutative Jordan algebra.

If we intercept commutative Jordan algebras, we'll have a commutative Jordan algebra, so there is a minimal commutative Jordan algebra, $\mathscr{A}^\circ$, that contains $P_1, \ldots, P_w$, T and $I - T$. The principal base of $\mathscr{A}^\circ$, (see Fonseca *et al.* 2006), is composed by the non-null matrices

$$Q_j = P_j T, \qquad j = 1, \ldots, w, \tag{4}$$

and

$$Q_{j+w} = P_j(I - T), \qquad j = 1, \ldots, w. \tag{5}$$

3. Canonic formulation

Let's represent by $Q_1^\circ, \ldots, Q_u^\circ$ the non-null matrices Q_j, $j = 1, \ldots, w$, and by $Q_{u+1}^\circ, \ldots, Q_v^\circ$ the non-null matrices Q_{j+w}, $j = 1, \ldots, w$. Therefore,

$$T = \sum_{j=1}^{u} Q_j^\circ \tag{6}$$

so that

$$\mu = T\mu = \sum_{j=1}^{u} Q_j^\circ \mu \tag{7}$$

and

$$I = \sum_{j=1}^{v} Q_j^\circ. \tag{8}$$

Now

$$Q_j^\circ = A_j^T A_j, \qquad j = 1, \ldots, v, \tag{9}$$

with A_j a matrix whose row vectors constitute an orthonormal base for the range image of Q_j°, $j = 1, \ldots, v$.

Therefore with

$$\eta_j = A_j \mu, \qquad j = 1, \ldots, u, \tag{10}$$

we'll have

$$\mu = \sum_{j=1}^{u} A_j^T \eta_j, \tag{11}$$

since

$$\mu = T\mu = \sum_{l=1}^{u} Q_l^{\circ}\mu = \sum_{l=1}^{u} A_l^T A_l \mu = \sum_{l=1}^{u} A_l \eta_l.$$

Besides this

$$P_j = Q_j + Q_{j+w}, \qquad j = 1, \dots, w, \tag{12}$$

so that at least one of the matrices of the second member will be non-null.

Now from (8) and (9) we get

$$\mathbf{y} = I\mathbf{y} = \sum_{l=1}^{v} Q_l^{\circ}\mathbf{y} = \sum_{l=1}^{v} A_l^T A_l \mathbf{y} = \sum_{l=1}^{v} A_l^T \widetilde{\eta}_l. \tag{13}$$

Let's now consider

$$\mathscr{D} = \{j; Q_j \neq 0\}$$

as well as

$$\mathscr{D}_1 = \mathscr{D} \cap \{1, \dots, w\} \quad \text{and} \quad \mathscr{D}_2 = \mathscr{D} \cap \{w+1, \dots, 2w\}.$$

We can now rewrite (11) as

$$\begin{cases} \mu = \sum_{l \in \mathscr{D}_1} A_l^T \eta_l \\ \\ \mathbf{y} = \sum_{h=1}^{2} \sum_{l \in \mathscr{D}_h} A_l^T \widetilde{\eta}_l \end{cases} . \tag{14}$$

The covariance matrix given by (1), according to these considerations, can be written as

$$\begin{aligned} \mathscr{V} &= \sum_{j=1}^{w} \gamma_j P_j \\ &= \sum_{l \in \mathscr{D}_1} \gamma_l Q_l + \sum_{l \in \mathscr{D}_2} \gamma_{l-w} Q_l, \end{aligned} \tag{15}$$

where the matrices in the second and third terms are mutually orthogonal orthogonal projection matrices.

Given the expression of the mean vector and covariance matrix, we get

$$\mathbf{y} \sim \mathcal{N}\left(\sum_{l\in\mathscr{D}_1} A_l^T\eta_l; \sum_{l\in\mathscr{D}_1}\gamma_l Q_l + \sum_{l\in\mathscr{D}_2}\gamma_{l-w}Q_l\right). \tag{16}$$

From (15), we obtain the inverse of the covariance matrix as

$$\mathscr{V}^{-1} = \sum_{l\in\mathscr{D}_1}\frac{1}{\gamma_l}Q_l + \sum_{l\in\mathscr{D}_2}\frac{1}{\gamma_{l-w}}Q_l, \tag{17}$$

since the Q_l, $l \in \mathscr{D}_1 \cup \mathscr{D}_2$, are mutually orthogonal orthogonal projection matrices.

Given (17), we obtain

$$\begin{aligned}\mathbf{U} &= (\mathbf{y}-\mu)\,\mathscr{V}^{-1}\,(\mathbf{y}-\mu)\\ &= \sum_{l\in\mathscr{D}_1}\frac{1}{\gamma_l}(\mathbf{y}-\mu)^T Q_l(\mathbf{y}-\mu) + \sum_{l\in\mathscr{D}_2}\frac{1}{\gamma_{l-w}}(\mathbf{y}-\mu)^T Q_l(\mathbf{y}-\mu)\ .\end{aligned} \tag{18}$$

Considering both parts of the model (fixed effects and random effects), let's study what can be said for each one.

For $l \in \mathscr{D}_1$, we get

$$\begin{aligned}(\mathbf{y}-\mu)^T Q_l(\mathbf{y}-\mu) &= (\mathbf{y}-\mu)^T A_l^T A_l(\mathbf{y}-\mu)\\ &= \|A_l(\mathbf{y}-\mu)\|^2 = \|\widetilde{\eta}_l-\eta_l\|^2\end{aligned} \tag{19}$$

and for $l \in \mathscr{D}_2$, we get, with $\widetilde{\eta}_l = A_l\mathbf{y}$,

$$(\mathbf{y}-\mu)^T Q_l(\mathbf{y}-\mu) = \|\widetilde{\eta}_l-\eta_l\|^2 = \|\widetilde{\eta}_l\|^2 = S_l \tag{20}$$

since $\eta_l = \mathbf{0}$ thus replacing in (18), we obtain

$$U = \sum_{l\in\mathscr{D}_1}\frac{\|\widetilde{\eta}_l-\eta_l\|^2}{\gamma_l} + \sum_{l\in\mathscr{D}_2}\frac{S_l}{\gamma_{l-w}}\ . \tag{21}$$

Moreover, the determinant of the covariance matrix is

$$det(\mathscr{V}) = \prod_{l \in \mathscr{D}_1} (\gamma_l)^{g_l} \prod_{l \in \mathscr{D}_2} (\gamma_{l-w})^{g_l}, \tag{22}$$

where $g_l = car(Q_l)$. We can then write the density function as

$$n(\mathbf{y}) = \frac{e^{-\frac{1}{2}\sum\limits_{l \in \mathscr{D}_1} \frac{\|\tilde{\eta}_l - \eta_l\|^2}{\gamma_l} - \frac{1}{2}\sum\limits_{l \in \mathscr{D}_2} \frac{S_l}{\gamma_{l-w}}}}{2\pi^{\frac{n}{2}} \sqrt{\prod\limits_{l \in \mathscr{D}_1} (\gamma_l)^{gl} \prod\limits_{l \in \mathscr{D}_2} (\gamma_{l-w})^{gl}}}. \tag{23}$$

We point out that from the expression of the $\boldsymbol{\eta}_j$, $j = 1, \ldots, u$, and the fact that $\gamma_j \geq 0$, $j = 1, \ldots, w$, it clearly follows that the parameter space of this density contains open sets, thus we have sufficient and complete statistics and according to the Blackwell-Lehmann-Scheffé theorem, the following uniform minimal variance unbiased estimators - UMVUE.

$$\widetilde{\eta}_l \sim \mathcal{N}\left(\eta_l; \gamma_l I_{gl}\right), \qquad l \in \mathscr{D}_1, \tag{24}$$

and

$$\widetilde{\gamma}_l = \frac{S_l}{g_l} \sim \chi^2_{g_l}, \qquad l \in \mathscr{D}_2 . \tag{25}$$

4. Inference

Starting with the variance components, if $l \in \mathscr{D}_2$, we can easily construct confidence intervals for γ_l and test hypothesis

$$H_{0,l} : \gamma_{l-w} = 0, \qquad l \in \mathscr{D}_2,$$

or more generally

$$H_{0,l} : \gamma_{l-w} = c_l, \qquad l \in \mathscr{D}_2.$$

In this last case we can derive bilateral tests or right and left unilateral tests.

If $l \in \mathscr{D}_1$ and $l + w \notin \mathscr{D}_2$, we have to admit that

$$\gamma_l = \sum_{h \in \mathscr{D}_2} c_h \gamma_{h-w}, \tag{26}$$

which would give us the UMVUE for γ_l,

$$\widetilde{\gamma}_l = \sum_{h \in \mathscr{D}_2} c_h \widetilde{\gamma}_{h-w}. \tag{27}$$

Later on we will find an example of such situation. To construct confidence intervals and test the hypothesis $H_{0,l}(c_l)$, we can use, (see Fonseca, 2005), induced pivot variables. Then, with the $X_{h,i}$ central, independent chi-squares, with g_h degree of freedom, $h \in \mathscr{D}_2$, $i = 1, \ldots, N$, we can construct samples constituted by the

$$Z_l = \sum_{h \in \mathscr{D}_2} c_l \frac{S_h}{X_{h,i}}, \qquad l = 1, \ldots \tag{28}$$

obtaining the correspondent empirical quantiles $z_{N,q}$ for probabilities q. We can prove (see Ferreira, 2005), that

$$\begin{cases} pr\left(z_{n,q_1} \leq \gamma_l \leq z_{n,q_2}\right) \longrightarrow q_2 - q_1 \\ pr\left(z_{n,q_1} \leq \gamma_l\right) \longrightarrow 1 - q_1 \\ pr\left(\gamma_l \leq z_{n,q_2}\right) \longrightarrow q_2 \end{cases}. \tag{29}$$

Thus if we are testing $H_0 : \gamma_l = \gamma_{0,l}$ against $H_1 : \gamma_l \neq \gamma_{0,l}$, we may reject H_0, with a two sided test, when $\gamma_{0,l} \notin \left[z_{n,\frac{q}{2}}; z_{n,1-\frac{q}{2}}\right]$. This two sided test will have limit power $1 - q$. In the same way one sided tests are derived.

Considering the vectors η_l with $l \in \mathscr{D}_1$, if $l+w \in \mathscr{D}_2$, γ_l can be estimated, since $\gamma_l = \gamma_{l+w}$. Then we can construct confidence hyperspheres for η_l since

$$\|\widetilde{\eta}_l - \eta_l\|^2 \sim \gamma_l \chi^2_{gl}, \qquad l \in \mathscr{D}_1, \tag{30}$$

independent of $S_{l+w} \sim \gamma_{l+w} \chi^2_{gl+w}$. Since $\gamma_l = \gamma_{l+w}$ we'll get

$$pr\left(\|\eta_l - \widetilde{\eta}_l\|^2 \leq g_l F_{1-q, g_l, g_{l+w}} \frac{S_{l+w}}{g_{l+w}}\right) = 1 - q, \qquad l \in \mathscr{D}_1, \tag{31}$$

which allows us to use duality to test

$$H_{0,l} : \eta_l = \mathbf{b}_l.$$

When $l+w \notin \mathscr{D}_2$ we use (26) to derive the test, rejecting the hypothesis when $\mathbf{b}_l$ is not covered by the confidence hypershere. We can generate samples constituted by the variables $W_i = Z_i U_i$, $i = 1, \ldots$, with U_i, $i = 1, \ldots, N$, cental chi-squared with g_l degrees of freedom independent of each other, and the Z_i, $i = 1, \ldots$ defined as earlier. With $W_{N,q}$ the empirical quantile for probability q, we have (see Ferreira, 2006),

$$pr\left(\|\eta_l - \widetilde{\eta}_l\|^2 \leq w_{N,1-q}\right) \xrightarrow[N\to\infty]{} 1-q, \tag{32}$$

which allows us to construct confidence spheres and to test $H_{0,l}$ through duality.

5. Associated orthogonal models

Many times, mixed models are written as

$$\mathbf{Y} = \sum_{i=1}^{v} X_i\beta_i + \sum_{i=v+1}^{w} X_i\widetilde{\beta}_i \tag{33}$$

with β_i, $i = 1, \ldots, v$, fixed and the $\widetilde{\beta}_i \sim \mathcal{N}\left(0; \sigma_i^2 I_i\right)$, $i = v+1, \ldots, w$.

If the matrices of $M = \{M_1, \ldots, M_w\}$ commute they, will generate a commutative Jordan algebra $\mathscr{A}(M)$. When M is a basis for $\mathscr{A}(M)$, with $Q = \{Q_1, \ldots, Q_w\}$, the principal basis we will have

$$M_i = \sum_{j=1}^{w} b_{i,j} Q_j, \tag{34}$$

where $B = [b_{i,j}]$ is the transition matrix which will be regular.

With $\mathscr{C} = \{j : b_{i,j} \neq 0\}$, we will have

$$R(M_i) = \boxplus_{j\in\mathscr{D}_i} R(Q_j), \tag{35}$$

where $\boxplus$ represents the orthogonal direct sum of subspaces.
The model $\mathbf{Y}$ has mean vector

$$\boldsymbol{\mu} = \sum_{i=1}^{v} \boldsymbol{X_i\beta_i} \tag{36}$$

so μ spans the range space of

$$X = [X_1, \ldots, X_v] .$$

With

$$M = XX^T = \sum_{i=1}^{v} X_i X_i^T = \sum_{i=1}^{v} M_i, \tag{37}$$

we have

$$R(M) = \boxplus_{j \in \bigcup_{i=1}^{v} \mathscr{C}_i} R(Q_j) \tag{38}$$

as well as

$$T = \sum_{j \in \bigcup_{i=1}^{v} \mathscr{C}_i} Q_j. \tag{39}$$

Besides this, the covariance matrix of $\mathbf{Y}$ is

$$\begin{aligned} \mathscr{V} &= \sum_{i=1}^{w} \sigma_i^2 M_i = \sum_{i=1}^{w} \sigma_i^2 \sum_{j=1}^{w} b_{i,j} Q_j \\ &= \sum_{j=1}^{w} \gamma_j Q_j \end{aligned} , \tag{40}$$

with

$$\begin{cases} \gamma_j = \sum_{i=1}^{w} b_{i,j} \sigma_i^2 \\ \gamma = B^T \sigma^2 \\ \sigma^2 = \left(B^T\right)^{-1} \gamma \end{cases} . \tag{41}$$

It is now obvious that these models have COBS, since T commutes with $Q_1, \ldots, Q_w$.

It is of interest for us to consider a specific class of these models, those whith segregation. In these model we have

$$B = \begin{bmatrix} B_{1,1} & 0 \\ B_{2,1} & B_{2,2} \end{bmatrix} \quad \text{and} \quad B^T = \begin{bmatrix} B_{1,1}^T & B_{2,1}^T \\ 0 & B_{2,2}^T \end{bmatrix}$$

with $B_{1,1}$, a $v \times v$ matrix. Let's define

(42)
$$\begin{cases} \sigma_{(2)}^2 = \left(\sigma_{v+1}^2, \dots, \sigma_w^2\right) \\ \gamma_{(1)} = (\gamma_1, \dots, \gamma_v) \\ \gamma_{(2)} = (\gamma_{v+1}, \dots, \gamma_w) \end{cases},$$

then we'll have

(43)
$$\begin{cases} \sigma_{(2)}^2 = \left(B_{2,2}^T\right)^{-1} \gamma_{(2)} \\ \gamma_{(1)} = B_{2,1}^T \left(B_{2,2}^T\right)^{-1} \gamma_{(1)} \end{cases},$$

which constitutes precisely an example of the relation beetwen variance components previously mentioned.

References

[1] D.J. Ferreira, *Inducing pivot variables and variance components in orthogonal normal models*, Ph.D Thesis 2005.

[2] S.M. Ferreira, *Inferęncia para Modelos Ortogonais com Segregação*, Tese de Doutoramento 2006.

[3] M. Fonseca, J.T. Mexia and R. Zmyślony, *Binary Operation on Jordan algebras and orthogonal normal models*, Linear Algebra And Its Applications (2006), 75–86.

[4] A.I. Khuri, *Advanced Calculus with Applications in Statistics*, 2nd ed., Wiley 2003.

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

[6] J.T. Mexia, *Introdução à Inferȩncia Estatística Linear*, Edições Universitárias Lusófonas 1995.

[7] J.R. Schott, *Matrix Analysis for Statistics*, Wiley Series in Probability and Statistics 1997.

[8] J. Seely, *Linear spaces and unbiased estimators. Application to a mixed linear model*, The Annals of Mathenatical Statistics **41** (5) (1970), 1735–1745.

[9] R. Zmyślony, *Completeness for a family of normal distributions*, Mathematic Statistics, Banach Center Publications **6** (1980), 355–357.

Received 26 October 2007
Revised 5 April 2008