Orthogonal Models, Structure Crossing, Nesting and Inference.

Acknowledgements

First of all I would like to thank my supervisor Prof. Miguel Fonseca and to Prof. João Tiago Mexia for their friendship, support and trust. Despite much work that they have, my supervisors were always available to clarify my questions. It was a very rich experience. Thank you very much.

I would also like to thank my good friend priest Carlos Jorge Vicente that really helped me in the hard moments. His support was, and is, very important for me. Thank you very much, priest Carlos. I will always remember you.

To all my friends "FIVES", who were, are and will be true friends and companions of a journey, a big thank for your great friendship.

To my friends and colleagues in the Centre of Mathematics and Applications of the Nova University of Lisbon. Their opinions, suggestions and friendship were crucial for this work.

I am very grateful to my in-laws, uncles and my two wonderful sons, Gonçalo and Gustavo.

Last, but certainly not least, I thank my lovely wife Rute for the great patience that she has with me. Thank you very much my dear wife for being part of my life. Without you nothing in my life would be possible, I love you. To God, a discrete but strong presence and support, who will never let me down. My life without you would not be the same. Thank you very much to all.

Abstract

We intend to study the algebraic structure of the simple orthogonal models to use them, through binary operations as building blocks in the construction of more complex orthogonal models.

We start by presenting some matrix results considering Commutative Jordan Algebras of symmetric matrices, CJAs. Next, we use these results to study the algebraic structure of orthogonal models, obtained by crossing and nesting simpler ones.

Then, we study the normal models with OBS, which can also be orthogonal models. We intend to study normal models with OBS (Orthogonal Block Structure), NOBS (Normal Orthogonal Block Structure), obtaining condition for having complete and sufficient statistics, having UMVUE, is unbiased estimators with minimal covariance matrices whatever the variance components.

Lastly, see ([Pereira et al. (2014)]), we study the algebraic structure of orthogonal models, mixed models whose variance covariance matrices are all positive semi definite, linear combinations of known orthogonal pairwise orthogonal projection matrices, OPOPM, 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 expressions for the LSE of these models.

Keywords: Orthogonal Models, Crossing, Nesting, Least Square Estimators.

Sinopse

Pretende-se estudar a estrutura algébrica dos modelos ortogonais de forma a poder-se utilizar os mesmos, através de operações binárias como blocos utilizados na construção de modelos ortogonais mais complexos.

Vamos começar por apresentar alguns resultados matriciais relativos a Álgebras Comtativas de Jordan de matrizes simétricas. A seguir , aplicaremos estes resultados para estudar a estrutura algébrica dos modelos ortogonais, para podermos analisar o cruzamento e o aninhamento de tais modelos.

Em seguida, vamos estudar os modelos normais com OBS, que também incluem os modelos ortogonais. Pretendemos estudar modelos normais com OBS, ou seja, o estudo de problemas de estimação em particular condições para ter estatísticas completas e suficientes, que garantem a existência de estimadores UMVUE (com matriz de covariânia minima na família dos estimadores centrados).

Finalmente, ver [Pereira et al. (2014)], vamos estudar a estrutura algébrica dos modelos ortogonais assim como dos modelos mistos cujas matrizes de covariância são todas as combinações lineares de conhecidas matrizes semi definidas positivas de projeção ortogonal mutuamente ortogonais, OPOPM, cuja soma é $\mathbf{I}_n$, e cujos estimadores de mínimos quadrados dos vetores estimáveis são estimadores lineares centrados com matrizes de covariância mínima, BLUE, quaisquer que sejam as componentes de variância, assim sendo uniformemente BLUE. Dos resultados da estrutura algébrica, obteremos expressões explicitas para esses estimadores de mínimos quadrados.

Palavras-chave: Modelos Ortogonais, Cruzamento, Aninhamento, Estimadores de Mínimos Quadrados.

Notation Index

- $\mathbf{X}^T$: transpose of matrix $\mathbf{X}$;
- $\mathbf{X}^{-1}$: inverse of matrix $\mathbf{X}$;
- $|\mathbf{X}|$: determinant of matrix $\mathbf{X}$;
- $\mathbf{X}^-$:generalized inverse (g-inverse) of matrix $\mathbf{X}$;
- $\mathbf{X}^+$:Moore-Penrose inverse of matrix $\mathbf{X}$;
- $\mathbf{I}_n$: identity matrix of size n;
- $\perp$: orthogonal;
- $\mathbf{O}_{n\times m}$: null matrix of size $n \times m$;
- $\otimes$: kronecker product;
- $\boxplus$: direct sum orthogonal of subspaces;
- $R(\mathbf{X})$: linear space spanned by the column vectors of matrix $\mathbf{X}$;
- $N(\mathbf{X})$: Kernel of matrix $\mathbf{X}$;
- $E[\mathbf{X}]$: expectation of the random vector $\mathbf{X}$;
- $V[\mathbf{X}]$: covariance matrix of the random vector $\mathbf{X}$;
- $N(\boldsymbol{\mu}, \mathbf{V})$: normal random vector with mean vector $\boldsymbol{\mu}$ and covariance matrix $\mathbf{V}$;
- $L(t, \theta)$: loss function of estimate t for θ;

- $R(t,\theta)$: risk function of estimator t for θ;
- $\mathcal{U}(\mathbf{X})$: uniformizing of $\mathbf{X}$;
- pb: principal basis
- CJA: Commutative Jordan Algebras;
- CJAs: Commutative Jordan Algebras constituted by symmetric matrices;
- OPOPM: orthogonal pairwise orthogonal projection matrices;
- OPM: orthogonal projection matrices;
- OBS: orthogonal block structure;
- COBS: commutative orthogonal block structure;
- NOBS: normal orthogonal block structure
- LSE: least square estimators;
- BLUE: best linear unbiased estimator;
- UBLUE: uniformly best linear unbiased estimator;
- ORT: orthogonal models;
- UMVUE: uniformly minimum variance unbiased estimation;
- $Pr(\mathbf{U})$: orthogonal projection matrix on $R(\mathbf{U})$;
- GME: Gauss-Markov estimator;
- GOBS: generalized orthogonal block structure.

Contents

Acknowledgements **v**

Abstract **vii**

Sinopse **ix**

Notation Index **xi**

1 Introduction **1**

2 Algebraic results **5**

2.1 Introduction 5

2.2 Eigenvalues and Eigenvectors 6

2.3 Gram Schmidt 7

2.3.1 Introduction 7

2.3.2 Gram Schmidt Process 7

2.3.3 QR Decomposition 9

2.4 Projection Matrices 10

2.5 The Moore-Penrose Inverse 11

2.6 Kronecker Product 13

2.7 Jordan Algebras 15

2.7.1 Commutative Jordan Algebras 16

2.7.2 Binary Operations on CJA 21

2.8 Important results of Commutative Jordan Algebra 22

3 Estimation and models **27**

3.1 Introduction 27

3.2 Estimators 28

3.3 Exponential families 28

3.4 Sufficient Statistics 29

3.5 Estimation on linear models 35

3.6 Orthogonal Models 42

3.7 Crossing Orthogonal Models 52

3.7.1 First case 53

3.7.2 Second case 55

3.8 Nesting Orthogonal Models 58

3.8.1 First case 58

3.8.2 Second case 60

3.9 Equivalence of models 63

3.9.1 Introduction 63

3.9.2 Equivalence of models 64

3.9.3 OBS Models 66

4 Normal models with OBS **69**

4.1 Introduction 69

4.2 Sufficient statistics and natural parameters 69

4.3 Mixed models 71

5 Orthogonal Models **79**

5.1 Introduction 79

5.2 Algebraic Structure 80

5.2.1 Orthogonal Block Structure, OBS 80

5.2.2 Commutative Orthogonal Block Structure, COBS 82

5.3 Least Squares Estimators, LSE 85

5.4 Uniformly Best Linear Unbiased Estimator, UBLUE 89

6 Final Remarks **95**

bibliography **97**

Chapter 1

Introduction

This work is divided into five chapters, beginning with the introduction in the first chapter. Here, we will review the themes and important results (which will be introduced later in), to be considered in each chapter throughout this work.

In chapter (2) we review some of the basic operations and fundamental properties involved in matrix algebra. Initially we introduce the notions of symmetric and orthogonal matrices, eigenvalues, eigenvectors. We will study in some detail the generalized Moore-Penrose inverses and Jordan Algebras as well as the commutative Jordan Algebras and binary operations defined on them, that will be useful for studying the following chapters. We are interested in commutative Jordan algebras, where matrices commute. These algebras have unique principal basis constituted by orthogonal projection matrices, all of them mutually orthogonal. Finally, we will introduce and demonstrate some concepts which we now state.

1. If $\mathbf{M}_1, \ldots, \mathbf{M}_n$ commute, then $\mathbf{M} = \{\mathbf{M}_1, \ldots, \mathbf{M}_n\}$ generates an algebra $\mathcal{A}(\mathbf{M})$;

2. If $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are orthogonal projection matrices (OPM), then

$$\mathbf{Q}_1\mathbf{Q}_2 = \mathbf{Q}_2\mathbf{Q}_1 \Longleftrightarrow \mathbf{U} = \mathbf{Q}_1\mathbf{Q}_2,$$

   if and only if $\mathbf{U}$ is an orthogonal projection matrix;

3. Let $\mathbf{Q}$ be an orthogonal projection matrix belonging to Commutative Jordan Algebra (CJA) $\mathcal{A}$, with rank 1, thus we will show that it will belong to principal basis (pb) of $\mathcal{A}$ constituted by the matrices $\{\mathbf{Q}_1, \ldots, \mathbf{Q}_n\}$, namely we may take $\mathbf{Q} = \mathbf{Q}_{j'}$.

The definition of Commutative Jordan Algebras, CJA, will be studied in detail in this chapter. See 2.7.1;

4. If $\mathbf{T}_l \in \mathcal{A}$, $\mathbf{T}_l = \sum\limits_{j\in\varphi(l)} \mathbf{Q}_j(l)$, $l = 1, 2$ and

$$\mathcal{A}(1) \otimes \mathcal{A}(2) = \{\mathbf{M}(1) \otimes \mathbf{M}(2);\ \mathbf{M}(l) \in \mathcal{A}(l)\}$$

so, we have the following:

$$\mathbf{T}_1 \otimes \mathbf{T}_2 \in \mathcal{A}(1) \otimes \mathcal{A}(2).$$

Chapter (3) deals with estimation and models. Firstly, are given definitions of estimators and their properties. The result on estimation will refer the use of sufficient and complete statistics to obtain good pontwise estimators. These last results will be useful in the study of the normal case. Then, we will define linear models and their special types are given, like Gauss-Markov Models, Generalized Gauss-Markov Models and definitions of orthogonal block structure, OBS (see definition (33)) and commutative orthogonal block structure, COBS (see definition (34)). We will get generalized result in characterization and expression of the best linear unbiased estimators, BLUE. In section (3.6), we will study a very important class of models that is the class of Orthogonal Models. Orthogonal models are models with COBS. If the matrices $\mathbf{M}$ commute, they generates an Commutative Jordan Algebra $\mathcal{A}(\mathbf{M})$. If the algebra $\mathcal{A}$ contains the matrices $\mathbf{M}_1, \ldots, \mathbf{M}_w$ and $\mathbf{T}$, the orthogonal projection matrix, OPM, on the space spanned by the mean vector the matrices of $\mathbf{M}$ commutes with $\mathbf{T}$, and the least squares estimators, LSE, of estimable vectors are uniformly UBLUE. This section, we will enunciate and demonstrate results that will be used.
In sections (3.7) and (3.8), we will study Crossing and Nesting Orthogonal Models. Orthogonal models , ORT, see [Vanleeuwen et al. (1998)], are mixed models with remarkable properties.
Finally, in section 3.9, we will study the equivalence of models. Two models are equivalent if they have the same family of estimable vectors and covariance matrices.

In chapter (4), we will study the normal models with OBS, and apply our results to orthogonal models. We intend to study normal models with OBS, NOBS (Normal

Orthogonal Block Structure); namely studying estimation problems and obtaining condition for having complete and sufficient statistics. The proposition 23 obtains the condition for having complete and sufficient statistics.

In chapter (5) see [Pereira et al. (2014)], we will study the algebraic structure of the mixed models with OBS and with COBS. Moreover, we will obtain explicit expressions for least squares estimators, LSE, in models with COBS using the results of preceding section, and then assuming normality. Lastly, we will characterize models whose LSE are UBLUE.

The main result we obtained refer to operations with orthogonal models and to the equivalence of such models. Moreover our study of the algebraic structure of mixed models with OBS and COBS is of interest.

Chapter 2

Algebraic results

2.1 Introduction

In this section we present certain preliminary results and definitions that will be useful along our work. In the first part of this work, we will discuss briefly the eigenvalues and eigenvectors, orthogonal projection matrices, Moore-Penrose inverse and Kronecker product. Lastly, we include a section on Jordan Algebras. In that section we will introduce some important results that are used in linear models. In this work, we will use repeatedly symmetric and orthogonal matrices. We now define:

Definition 1. ***(Symmetric Matrix)***

A symmetric matrix is a square matrix that is equal to its transpose . Let $\boldsymbol{A}$ be a symmetric matrix. Then $\boldsymbol{A} = \boldsymbol{A}^T$, with $\boldsymbol{A}^T$ the transpose of $\boldsymbol{A}$.

Definition 2. ***(Orthogonal Matrix)***

An orthogonal matrix is a square matrix with real entries whose columns and rows are pairwise orthogonal unit vectors (orthonormal vectors). An $m \times 1$ vector $\boldsymbol{p}$ is a normalized vector or unit vector if $\boldsymbol{p}^T\boldsymbol{p} = 1$. The $m \times 1$ vectors $\boldsymbol{p}_1, ..., \boldsymbol{p}_n$, with $n \leq m$, are orthogonal if $\boldsymbol{p}_i^T\boldsymbol{p}_j = 0$ for all $i \neq j$. for each $\boldsymbol{p}_i$ is a normalized vector then the vectors are orthonormal.

Equivalently, an matrix $\boldsymbol{A}$ is orthogonal if its transpose is equal to its inverse:

$$\boldsymbol{A}^T = \boldsymbol{A}^{-1}$$

which entails

$$\boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^T = \boldsymbol{I},$$

where $\boldsymbol{I}$ *is the identity matrix.*

2.2 Eigenvalues and Eigenvectors

In this section, we present results on eigenvalues and eigenvectors which will be useful in studying and understanding the next themes.

If $\mathbf{A}$ is an $m \times m$ matrix, then any scalar λ satisfying the equation:

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}, \tag{2.1}$$

will be an eigenvalue associated to the eigenvector $\mathbf{x}$ where $\mathbf{x}$ is an unit vector and the equation (2.1) is called the eigenvalue-eigenvector equation of $\mathbf{A}$. The equation (2.1) can be equivalently expressed as

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = 0 \tag{2.2}$$

If we had $|\mathbf{A} - \lambda\mathbf{I}| \neq 0$, $(\mathbf{A} - \lambda\mathbf{I})^{-1}$ will exist, then premultiplying in equation (2.2) by $(\mathbf{A} - \lambda\mathbf{I})^{-1}$, we would get $\mathbf{x} = \mathbf{0}$ which is impossible, so we must have

$$|\mathbf{A} - \lambda\mathbf{I}| = 0, \tag{2.3}$$

which is called the characteristic equation of $\mathbf{A}$. The characteristic equation can be used to obtain the eigenvalues of matrix $\mathbf{A}$ which can be used in the eigenvalues-eigenvector equation to obtain the corresponding eigenvectors.

Before ending this section, we will enumerate some basic properties of eigenvalues and eigenvectors.

Let $\mathbf{A}$ be an $m \times m$ matrix. Then

1. The eigenvalues of $\mathbf{A}^T$ are the same as the eigenvalues of $\mathbf{A}$;
2. $\mathbf{A}$ is singular if and only if at least one eigenvalue of $\mathbf{A}$ is equal to 0;
3. the diagonal elements of $\mathbf{A}$ are the eigenvalues of $\mathbf{A}$, if $\mathbf{A}$ is a triangular matrix;
4. the eigenvalues of $\mathbf{BAB}^{-1}$ are the same as the eigenvalues of $\mathbf{A}$, if $\mathbf{B}$ is a nonsingular $m \times m$ matrix;

5. if $\mathbf{A}$ is an orthogonal matrix, its eigenvalues are 1 or -1.

2.3 Gram Schmidt

2.3.1 Introduction

The Gram Schmidt process is a method for orthonormalising a set of vectors in an inner product space. The Gram Schmidt process takes a finite, linearly independent set of vectors $V = \{\mathbf{v}_1, \mathbf{v}_2 \ldots, \mathbf{v}_m\}$ and generates an orthonormal set $V^o = \{\mathbf{u}_1, \mathbf{u}_2 \ldots, \mathbf{u}_m\}$ that spans the same m-dimensional subspace of $\mathbb{R}^n$ as V, being $m \leq n$.

2.3.2 Gram Schmidt Process

We have the following subspace V contained in $\mathbb{R}^n$

$$V = \{\mathbf{v}_1, \mathbf{v}_2 \ldots, \mathbf{v}_m\}$$

the subspace V is formed by vectors $\mathbf{v}_1, \mathbf{v}_2 \ldots, \mathbf{v}_m$ linearly independent and it is a basis of V. The goal is to build an orthonormal basis. The Gram Schmidt process has the following steps:

1. We will consider a simple case with a single dimension. We will define a space $V_1 = span(\mathbf{v}_1)$, $\mathbf{v}_1$ is the basis of V_1. How to build an orthonormal basis? We will consider a vector $\mathbf{u}_1$ and we will normalize it. So, we have

$$\mathbf{u}_1 = \frac{1}{||\mathbf{v}_1||}\mathbf{v}_1$$

   So, we can say that $V_1 = span(\mathbf{u}_1)$. It is a orthonormal basis of V_1. What happens in other cases?

2. We have a space $V_2 = span(\mathbf{v}_1, \mathbf{v}_2)$. We know that $\mathbf{v}_1$ is a linear combination of $\mathbf{u}_1$. Thus, $V_2 = span(\mathbf{u}_1, \mathbf{v}_2)$. We will now assume that the vector $\mathbf{v}_2$ is a linear combination of the vectors $\mathbf{x}$ and $\mathbf{y}_2$ and we can write as follow

$$\mathbf{v}_2 = \mathbf{x} + \mathbf{y}_2$$

where $\mathbf{x} \in V_1$ and $\mathbf{y}_2 \in V_1^{\perp}$. The vector $\mathbf{x}$ is the orthogonal projection of $\mathbf{v}_2$ in space V_1. So, the vector $\mathbf{y}_2$ writes as follows

$$\mathbf{y}_2 = \mathbf{v}_2 - proj_{V_1}\mathbf{v}_2.$$

So, we build the vector $\mathbf{v}_2$ from a linear combination of the vectors $\mathbf{u}_1$ and $\mathbf{y}_2$, the vectors are orthogonal. We know previously that V_1 has a orthonormal basis, so we know that

$$proj_{V_1}\mathbf{v}_2 = (\mathbf{v}_2.\mathbf{u}_1)\mathbf{u}_1$$

The vector $\mathbf{y}_2$ is

$$\mathbf{y}_2 = \mathbf{v}_2 - (\mathbf{v}_2|\mathbf{u}_1)\mathbf{u}_1,$$

where $(.|.)$ is the inner product.

We will now normalize the vector $\mathbf{y}_2$

$$\mathbf{u}_2 = \frac{1}{||\mathbf{y}_2||}\mathbf{y}_2.$$

So, $V_2 = span(\mathbf{u}_1, \mathbf{u}_2)$ and it is a orthonormal basis. The vector $\mathbf{u}_2$ is written as follow

$$\mathbf{u}_2 = \frac{1}{||\mathbf{v}_2 - (\mathbf{v}_2|\mathbf{u}_1)\mathbf{u}_1||}(\mathbf{v}_2 - (\mathbf{v}_2|\mathbf{u}_1)\mathbf{u}_1)$$

3. We have the space $V_m = span\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$. Previous following the same reasoning, we have $V_m = span\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{m-1}, \mathbf{v}_m\}$. So, $V_m = span\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{m-1}, \mathbf{u}_m\}$ is a orthonormal basis. The vector $\mathbf{u}_m$ is written as follow

$$\mathbf{u}_m = \frac{1}{\left|\left|\mathbf{v}_m - \sum_{i=1}^{m-1}(\mathbf{v}_m|\mathbf{u}_i)\mathbf{u}_i\right|\right|}\left(\mathbf{v}_m - \sum_{i=1}^{m-1}(\mathbf{v}_m|\mathbf{u}_i)\mathbf{u}_i\right)$$

So, we can say that

1. $||\mathbf{u}_i|| = 1,\ i = \{1, \dots, m\}$, the vectors are unitary, have dimension 1

2. $\mathbf{u}_i|\mathbf{u}_j,\ i \neq j$, the vectors are orthogonal.

2.3.3 QR Decomposition

In linear algebra, a **QR** decomposition (also called a **QR** factorization) of a matrix is a decomposition of a matrix $\mathbf{V}$ into a product $\mathbf{V} = \mathbf{QR}$ of an orthogonal matrix $\mathbf{Q}$ and an upper triangular matrix $\mathbf{R}$. Next, we will study this topic.

From previous section we know that:

1. $\mathbf{y}_1 = \mathbf{v}_1$
2. $\mathbf{y}_2 = \mathbf{v}_2 - (\mathbf{v}_2|\mathbf{u}_1)\mathbf{u}_1$.
3. $\mathbf{y}_3 = \mathbf{v}_3 - [(\mathbf{v}_3|\mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v}_3|\mathbf{u}_2)\mathbf{u}_2]$
4. ...
5. $\mathbf{y}_m = \mathbf{v}_m - \left[\sum_{i=1}^{m-1}(\mathbf{v}_m|\mathbf{u}_i)\mathbf{u}_i\right]$

We then rearrange the equations above so that the $\mathbf{a}_i$s are on the left, using the fact that the $\mathbf{u}_i$ are unit vectors:

1. $\mathbf{v}_1 = (\mathbf{v}_1|\mathbf{u}_1)\mathbf{u}_1$
2. $\mathbf{v}_2 = (\mathbf{v}_2|\mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v}_2|\mathbf{u}_2)\mathbf{u}_2$
3. ...
4. $\mathbf{v}_m = \sum_{i=1}^{m}(\mathbf{v}_m|\mathbf{u}_i)\mathbf{u}_i$

where $(\mathbf{v}_i|\mathbf{u}_i) = ||\mathbf{y}_i||$. This can be written in matrix form:

$$\mathbf{V} = \mathbf{QR}$$

where

$$\mathbf{Q} = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m] \tag{2.4}$$

being $\mathbf{Q}$ a orthogonal matrix, and

$$\mathbf{R} = \begin{bmatrix} \mathbf{v}_1|\mathbf{u}_1 & \mathbf{v}_2|\mathbf{u}_1 & \cdots & \mathbf{v}_m|\mathbf{u}_1 \\ 0 & \mathbf{v}_2|\mathbf{u}_2 & \cdots & \mathbf{v}_m|\mathbf{u}_2 \\ \vdots & 0 & \cdots & \vdots \\ 0 & 0 & \cdots & \mathbf{v}_m|\mathbf{u}_m \end{bmatrix} \tag{2.5}$$

$\mathbf{R}$ is upper triangular matrix.

2.4 Projection Matrices

Let $\mathcal{S}$ be a linear subspace of $\mathbb{R}^n$. Let also $\mathbf{z}_1, ..., \mathbf{z}_n$ be an orthonormal base for $\mathbb{R}^n$ and $\mathbf{z}_1, ..., \mathbf{z}_m$, m $< n$ an orthonormal basis for $\mathcal{S}$. And so, we have the following relationship

$$\mathbf{x} = \sum_{i=1}^{m} \alpha_i \mathbf{z}_i + \sum_{i=m+1}^{n} \alpha_i \mathbf{z}_i = \mathbf{u} + \mathbf{v},$$

where $\mathbf{u}$ is called the orthogonal projection of $\mathbf{x}$ on $\mathcal{S}$ and $\mathbf{v}$ the orthogonal projection of $\mathbf{x}$ on $\mathcal{S}^{\perp}$, and the $\alpha_1, ..., \alpha_n$ are constants.

Let us consider the following matrix

$$\mathbf{Z} = \begin{bmatrix} \mathbf{Z}_1 & \mathbf{Z}_2 \end{bmatrix}$$

where

$$\mathbf{Z}_1 = (\mathbf{z}_1, ..., \mathbf{z}_m) \text{ and } \mathbf{Z}_2 = (\mathbf{z}_{m+1}, ..., \mathbf{z}_n)$$

We also write $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_n)^T$, $\boldsymbol{\alpha_1} = (\alpha_1, ..., \alpha_m)^T$ and $\boldsymbol{\alpha_2} = (\alpha_{m+1}, ..., \alpha_n)^T$. Then the expression for can be written as

$$\mathbf{x} = \mathbf{Z}\boldsymbol{\alpha} = \mathbf{Z}_1\boldsymbol{\alpha_1} + \mathbf{Z}_2\boldsymbol{\alpha_2}$$

that is, $\mathbf{u} = \mathbf{Z}_1\boldsymbol{\alpha_1}$ and $\mathbf{v} = \mathbf{Z}_2\boldsymbol{\alpha_2}$

Due to orthonormality we have $\mathbf{Z}_1^T\mathbf{Z}_1 = \mathbf{I}_n$ and $\mathbf{Z}_1^T\mathbf{Z}_2 = (\mathbf{0})$, and so we have

$$\mathbf{Z}_1\mathbf{Z}_1^T\mathbf{x} = \mathbf{Z}_1\mathbf{Z}_1^T\mathbf{Z}\boldsymbol{\alpha} = \mathbf{Z}_1\mathbf{Z}_1^T \begin{bmatrix} \mathbf{Z}_1 & \mathbf{Z}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha_1} \\ \boldsymbol{\alpha_2} \end{bmatrix} = \begin{bmatrix} \mathbf{Z}_1 & (\mathbf{0}) \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha_1} \\ \boldsymbol{\alpha_2} \end{bmatrix} = \mathbf{Z}_1\boldsymbol{\alpha_1} = \mathbf{u}.$$

Thus, we can define the orthogonal projection of a vector onto $\mathcal{S}$ (see [Schott (1997)], page 53)

Theorem 1. *Suppose column vectors of the $m \times r$ matrix $\boldsymbol{Z}_1$ form an orthonormal basis for the vector space $\mathcal{S}$ which is a subspace of $\mathbb{R}^m$. If $\boldsymbol{x} \in \mathbb{R}^m$, the orthogonal projection of $\boldsymbol{x}$ onto $\mathcal{S}$ is given by $\boldsymbol{Z}_1\boldsymbol{Z}_1^T\boldsymbol{x}$.*

So, the matrix $\mathbf{Z}_1\mathbf{Z}_1^T$ is called the projection matrix on the vector space $\mathcal{S}$ and will be denoted by $\mathbf{U}$.

According ([Seber (1980)]), we can have the following

$$\begin{cases} R(\mathbf{U}) = \mathcal{S} \\ N(\mathbf{U}) = \mathcal{S}^{\perp} \end{cases},$$

where $R(\mathbf{U})$ is the range subspace and $N(\mathbf{U})$ the kernel of matrix $\mathbf{U}$. We have

$$R(\mathbf{U}^T) = N(\mathbf{U})^{\perp}$$

Similarly, $\mathbf{Z}_2\mathbf{Z}_2^T$ is the orthogonal projection matrix for $\mathcal{S}^{\perp}$ and $\mathbf{Z}\mathbf{Z}^T = \mathbf{I}_n$ is the projection matrix for $\mathbb{R}^n$.

Thus

$$(\mathbf{I}_n - \mathbf{Z}_1\mathbf{Z}_1^T)\mathbf{x} = \mathbf{x} - \mathbf{Z}_1\mathbf{Z}_1^T\mathbf{x} = \mathbf{x} - \mathbf{Z}_1\boldsymbol{\alpha}_{\mathbf{1}} = \mathbf{Z}_2\boldsymbol{\alpha}_{\mathbf{2}} = \mathbf{v}.$$

The next theorem shows that a vector space does not have a unique orthonormal basis, but the projection matrix formed from these orthonormal basis is unique.

Theorem 2. *Suppose the columns of the $m \times r$ matrices $\boldsymbol{Z}_1$ and $\boldsymbol{W}_1$ each form an orthonormal basis for the r-dimensional vector space $\mathcal{S}$. Then $\boldsymbol{Z}_1\boldsymbol{Z}_1^T = \boldsymbol{W}_1\boldsymbol{W}_1^T$.*

Proof: See ([Schott (1997)]), page 53. □

We known that $\mathbf{U} = \mathbf{Z}_1\mathbf{Z}_1^T$, so

$$\mathbf{U}^T = (\mathbf{Z}_1\mathbf{Z}_1^T)^T = \mathbf{Z}_1\mathbf{Z}_1^T = \mathbf{U},$$

thus $\mathbf{U}$ is symmetric and

$$\mathbf{U}\mathbf{U} = \mathbf{Z}_1\mathbf{Z}_1^T\mathbf{Z}_1\mathbf{Z}_1^T = \mathbf{Z}_1\mathbf{I}_n\mathbf{Z}_1^T = \mathbf{Z}_1\mathbf{Z}_1^T = \mathbf{U},$$

is idempotent. So, the next theorem proves the converse.

Theorem 3. *Let $\boldsymbol{P}$ be an $m \times m$ symmetric idempotent matrix of rank r. Then there is an r-dimensional vector space which has $\boldsymbol{P}$ as its projection matrix.*

Proof: See ([Schott (1997)]), page 59 □

2.5 The Moore-Penrose Inverse

Generalized inverses are a very versatile tool, in the study of linear algebra and statistical applications. Generalized inverses can be derived from two different contexts: equations system and linear applications.

Let $\mathbf{A}$ be a general $m \times n$ matrix. Given

$$\mathbf{Ax} = \mathbf{y},\ for\ \mathbf{x} \in \mathbb{R}^m,\ \text{given } \mathbf{y} \in \mathbb{R}^n \tag{2.6}$$

If $\mathbf{A}$ is a square matrix (m = n) and $\mathbf{A}$ is invertible, then (2.6) holds if and only if $\mathbf{x} = \mathbf{A}^{-1}\mathbf{y}$. If $\mathbf{A}$ is not invertible, then equation (2.6) may have no solutions. If the system is undetermined, a possible solution is given by a matrix $\mathbf{A}^{-}$, such that:

Definition 3. *If $\boldsymbol{A}$ is an $m \times n$ matrix, then $\boldsymbol{G}$ is a generalized inverse (g-inverse) of $\boldsymbol{A}$ if $\boldsymbol{G}$ is an $n \times m$ matrix such that*

$$\boldsymbol{AGA} = \boldsymbol{A}, \tag{2.7}$$

and it is represented as $\boldsymbol{A}^{-}$.

The Moore-Penrose inverse is a particular case of the generalized inverses and was studied by [Moore (1920), (1935)] and [Penrose (1955)]. We will now study some basic proprieties of the Moore-Penrose inverse. So, we have the following definition, see ([Schott (1997)]), page 171

Definition 4. *The Moore-Penrose inverse of the $m \times n$ matrix $\mathbf{A}$ is the $n \times m$ matrix, denoted by $\boldsymbol{A}^{+}$, which satisfies the following conditions:*

$$\begin{cases} \boldsymbol{AA}^{+}\boldsymbol{A} = \boldsymbol{A} \\ \boldsymbol{A}^{+}\boldsymbol{AA}^{+} = \boldsymbol{A}^{+} \\ (\boldsymbol{AA}^{+})^T = \boldsymbol{AA}^{+} \\ (\boldsymbol{A}^{+}\boldsymbol{A})^T = \boldsymbol{A}^{+}\boldsymbol{A} \end{cases} \tag{2.8}$$

If $\mathbf{A}$ is regular, then

$$\mathbf{A}^{+} = \mathbf{A}^{-1}. \tag{2.9}$$

The next theorem answers that there is a Moore-Penrose inverse for every matrix.

Theorem 4. *Corresponding to each $m \times n$ matrix $\boldsymbol{A}$, there exists one and only one $n \times m$ matrix $\boldsymbol{A}^{+}$ satisfying conditions in the definition 4.*

Proof: See ([Schott (1997)]), page 171. □

When $\mathbf{A}$ is an symmetric matrix of order k with,

$$l = rank(\mathbf{A}) < k, \tag{2.10}$$

we can always order the row vectors of an orthogonal matrix $\mathbf{P}$, the orthogonal diagnolizer of $\mathbf{A}$, to get

$$\mathbf{PAP}^T = \mathbf{D}(\lambda_1, ..., \lambda_l, 0, \dots, 0)$$

with $\mathbf{D}(\lambda_1, ..., \lambda_l, 0, \dots, 0)$ a diagonal matrix whose main elements $\lambda_1, ..., \lambda_l$ are the non null eigenvalues of $\mathbf{A}$. Thus

$$\mathbf{A} = \mathbf{P}^T\mathbf{D}(\lambda_1, \dots, \lambda_l, 0, \dots, 0)\mathbf{P}$$

and

$$\mathbf{A}^+ = \mathbf{P}^T\mathbf{D}(\lambda_1^{-1}, \dots, \lambda_l^{-1}, 0, \dots, 0)\mathbf{P}.$$

2.6 Kronecker Product

This matrix operation has been widely studied, for instance see [Steeb (1991)] and [Graham (1981)]. We only recall results that will be useful.

Give the $m \times n$ matrix $\mathbf{A} = [a_{ij}]$, its Kronecker product with matrix $\mathbf{B}$ will be

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{m1}\mathbf{B} & \cdots & a_{mn}\mathbf{B} \end{bmatrix}. \tag{2.11}$$

We will now enumerate some properties of Kronecker product:

1. $\mathbf{A} \otimes (\mathbf{B} + \mathbf{C}) = \mathbf{A} \otimes \mathbf{B} + \mathbf{A} \otimes \mathbf{C}$;

2. $(\mathbf{A} + \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes \mathbf{C} + \mathbf{B} \otimes \mathbf{C}$;

3. $(k\mathbf{A}) \otimes \mathbf{B} = \mathbf{A} \otimes (k\mathbf{B}) = k(\mathbf{A} \otimes \mathbf{B})$;

4. $(\mathbf{A} \otimes \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes (\mathbf{B} \otimes \mathbf{C})$.
   Where k is a scalar.

It is easy to see that $(\mathbf{A} \otimes \mathbf{B})^T = \mathbf{A}^T \otimes \mathbf{B}^T$. As for the distributivity with the usual matrix product, we have the following proposition:

Proposition 1. *If the usual matrix products $\boldsymbol{A}_1\boldsymbol{A}_2$ and $\boldsymbol{B}_1\boldsymbol{B}_2$ are defined*

$$(\boldsymbol{A}_1 \otimes \boldsymbol{B}_1)(\boldsymbol{A}_2 \otimes \boldsymbol{B}_2) = (\boldsymbol{A}_1\boldsymbol{A}_2) \otimes (\boldsymbol{B}_1\boldsymbol{B}_2) \tag{2.12}$$

Proof:

$$
\begin{aligned}
(\mathbf{A}_1 \otimes \mathbf{B}_1)(\mathbf{A}_2 \otimes \mathbf{B}_2) &= \begin{bmatrix} a_{11}(1)\mathbf{B}_1 & \cdots & a_{1n}(1)\mathbf{B}_1 \\ \vdots & \ddots & \vdots \\ a_{m1}(1)\mathbf{B}_1 & \cdots & a_{mn}(1)\mathbf{B}_1 \end{bmatrix} \begin{bmatrix} a_{11}(2)\mathbf{B}_2 & \cdots & a_{1n}(2)\mathbf{B}_2 \\ \vdots & \ddots & \vdots \\ a_{m1}(2)\mathbf{B}_2 & \cdots & a_{mn}(2)\mathbf{B}_2 \end{bmatrix} = \\
&= \begin{bmatrix} \left(\sum_{j=1}^{n} a_{1j}(1)a_{j1}(2)\right)\mathbf{B}_1\mathbf{B}_2 & \cdots & \left(\sum_{j=1}^{n} a_{1j}(1)a_{jn}(2)\right)\mathbf{B}_1\mathbf{B}_2 \\ \vdots & \ddots & \vdots \\ \left(\sum_{j=1}^{n} a_{mj}(1)a_{j1}(2)\right)\mathbf{B}_1\mathbf{B}_2 & \cdots & \left(\sum_{j=1}^{n} a_{mj}(1)a_{jn}(2)\right)\mathbf{B}_1\mathbf{B}_2 \end{bmatrix} = \\
&= (\mathbf{A}_1\mathbf{A}_2) \otimes (\mathbf{B}_1\mathbf{B}_2)
\end{aligned}
\tag{2.13}
$$

Thus $(\mathbf{A}_1 \otimes \mathbf{B}_1)(\mathbf{A}_2 \otimes \mathbf{B}_2) = (\mathbf{A}_1\mathbf{A}_2) \otimes (\mathbf{B}_1\mathbf{B}_2)$ □

Some classes of matrices are "closed" for the Kronecker product, with $\mathbf{J}_s = \mathbf{1}_s\mathbf{1}_s^T$, we have

1. $\mathbf{1}_u \otimes \mathbf{1}_v = \mathbf{1}_{uv}$
2. $\mathbf{I}_u \otimes \mathbf{I}_v = \mathbf{I}_{uv}$
3. $\mathbf{J}_u \otimes \mathbf{J}_v = \mathbf{J}_{uv}$

With $\mathbf{J}_s = \mathbf{1}_s\mathbf{1}_s^T$. With $\mathbf{D}(\boldsymbol{\epsilon_1})$ and $\mathbf{D}(\boldsymbol{\epsilon_2})$ the diagonal matrices whose principal elements are the components of vectors $\boldsymbol{\epsilon_1}$ and $\boldsymbol{\epsilon_2}$, we have

$$\mathbf{D}(\boldsymbol{\epsilon_1}) \otimes \mathbf{D}(\boldsymbol{\epsilon_2}) = \mathbf{D}(\boldsymbol{\epsilon_1} \otimes \boldsymbol{\epsilon_2}) \tag{2.14}$$

Let $\mathbf{M}_1$ and $\mathbf{M}_2$ be symmetric matrices, there are orthogonal matrices $\mathbf{P}_1$ and $\mathbf{P}_2$ and diagonal matrices $\mathbf{D}(\boldsymbol{\epsilon_1})$ and $\mathbf{D}(\boldsymbol{\epsilon_2})$ such that

$$\mathbf{P}_j\mathbf{M}_j\mathbf{P}_j^T = \mathbf{D}(\boldsymbol{\epsilon_j});\ j = 1, 2.$$

The principal elements of $\mathbf{D}(\boldsymbol{\epsilon_j})$; $j = 1, 2$ are the eigenvalues of $\mathbf{M}_j$ while the row vectors $\mathbf{P}_j$; $j = 1, 2$ are the eigenvectors. So, from (2.12) and (2.14), we have

$$
\begin{aligned}
(\mathbf{P}_1 \otimes \mathbf{P}_2)(\mathbf{M}_1 \otimes \mathbf{M}_2)(\mathbf{P}_1 \otimes \mathbf{P}_2)^T &= (\mathbf{P}_1\mathbf{M}_1\mathbf{P}_1^T) \otimes (\mathbf{P}_2\mathbf{M}_2\mathbf{P}_2^T) \\
&= \mathbf{D}(\boldsymbol{\epsilon_1} \otimes \boldsymbol{\epsilon_2})
\end{aligned}
\tag{2.15}
$$

So, the eigenvalues of $\mathbf{M}_1 \otimes \mathbf{M}_2$ will be the products of the eigenvalues of $\mathbf{M}_1$ and $\mathbf{M}_2$ and the corresponding eigenvectors are the row vectors of $\mathbf{P}_1 \otimes \mathbf{P}_2$.
The following proposition shows that the Kronecker product of orthogonal projection matrices gives orthogonal projection matrices.

Proposition 2. *If $\boldsymbol{Q}_1$ and $\boldsymbol{Q}_2$ are orthogonal projection matrices, then $\boldsymbol{Q}_1 \otimes \boldsymbol{Q}_2$ is an orthogonal projection matrix.*

Proof: If $\mathbf{Q}_1$ is an orthogonal projection matrix, then we know that it is symmetric and idempotent, and the same for $\mathbf{Q}_2$. Now

1. $\mathbf{Q}_1 \otimes \mathbf{Q}_2$ symmetric:

$$(\mathbf{Q}_1 \otimes \mathbf{Q}_2)^T = \mathbf{Q}_1^T \otimes \mathbf{Q}_2^T = \mathbf{Q}_1 \otimes \mathbf{Q}_2$$

   is symmetric.

2. $\mathbf{Q}_1 \otimes \mathbf{Q}_2$ idempotent:

$$(\mathbf{Q}_1 \otimes \mathbf{Q}_2)(\mathbf{Q}_1 \otimes \mathbf{Q}_2) = (\mathbf{Q}_1\mathbf{Q}_1) \otimes (\mathbf{Q}_2\mathbf{Q}_2) = \mathbf{Q}_1 \otimes \mathbf{Q}_2$$

   is idempotent.

Thus if $\mathbf{Q}_1 \otimes \mathbf{Q}_2$ is symmetric and idempotent, it is an orthogonal projection matrix.

$\square$

2.7 Jordan Algebras

We will devote this section to Jordan algebras.

Jordan algebras were first introduced by [Jordan et al. (1934)], to formalize the notion of an algebra of observables in quantum mechanics. They were originally called "r-number systems", but were renamed "Jordan algebras" by Adrian Albert (1946), who began the systematic study of general Jordan algebras.

Recall now the definition of algebra.

Definition 5. *An algebra $\mathcal{A}$ is a linear space equipped with a binary operation $*$, usually called product, for which the following properties are verified for all $\alpha \in \mathbb{R}$ and $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in \mathcal{A}$:*

1. $\boldsymbol{a} * (\boldsymbol{b} + \boldsymbol{c}) = \boldsymbol{a} * \boldsymbol{b} + \boldsymbol{a} * \boldsymbol{c}$;

2. $(\boldsymbol{a} + \boldsymbol{b}) * \boldsymbol{c} = \boldsymbol{a} * \boldsymbol{c} + \boldsymbol{b} * \boldsymbol{c}$;

3. $\alpha(\boldsymbol{a} * \boldsymbol{b}) = (\alpha \boldsymbol{a}) * \boldsymbol{b} = \boldsymbol{a} * (\alpha \boldsymbol{b})$.

It is important to point out the definitions of associativity and commutativity. The definitions follow:

Definition 6. *An algebra $\mathcal{A}$ is associative if and only if for all $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in \mathcal{A}$, respectively*

$$(\boldsymbol{a} * \boldsymbol{b}) * \boldsymbol{c} = \boldsymbol{a} * (\boldsymbol{b} * \boldsymbol{c})$$

Definition 7. *An algebra $\mathcal{A}$ is commutative if and only if for all $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in \mathcal{A}$,*

$$\boldsymbol{a} * \boldsymbol{b} = \boldsymbol{b} * \boldsymbol{a}.$$

We now define the Jordan product in an algebra.

Definition 8. *A Jordan algebra is an (not necessary associative) algebra whose product (.) satisfies the conditions:*

1. $\boldsymbol{a}.\boldsymbol{b} = \boldsymbol{b}.\boldsymbol{a}$ *(commutative law)*

2. $(\boldsymbol{a}.\boldsymbol{b}).\boldsymbol{a}^{2\cdot} = \boldsymbol{a}.(\boldsymbol{b}.\boldsymbol{a}^{2\cdot})$ *(Jordan identity)*

where $\mathbf{a}^{2\cdot} = \mathbf{a}.\mathbf{a}$. An algebra equipped with this product, is a Jordan algebra.

We will consider an example of Jordan algebra. Let $\mathcal{S}_n$ be the space of symmetric real matrices of size n. It is a linear space equipped the (associative) matrix product. Then, with the product (.) is defined as

$$\mathbf{a}.\mathbf{b} = \frac{1}{2}(\mathbf{ab} + \mathbf{ba}) \tag{2.16}$$

it is a Jordan algebra. It is a very important example of Jordan algebra, because it is the space in which covariance matrices lie.

2.7.1 Commutative Jordan Algebras

Now see ([Schott (1997)]), page 156, we have the theorem

Theorem 5. *Let $\boldsymbol{A}_1, ..., \boldsymbol{A}_m$ be $k \times k$ symmetric matrices. Then there exists an orthogonal matrix $\boldsymbol{P}$ such that, $\boldsymbol{P}^T \boldsymbol{A}_i \boldsymbol{P} = \boldsymbol{D}_i$ is an diagonal matrix whose main elements are the eigenvalues of $\boldsymbol{A}_i$ if and only if $\boldsymbol{A}_i \boldsymbol{A}_j = \boldsymbol{A}_j \boldsymbol{A}_i$ for all pairs (i,j); i,j=1,...,m, i.e., if and only if the matrices $\boldsymbol{A}_1, ..., \boldsymbol{A}_m$ commute.*

Proof: See ([Schott (1997)]), page 156. □

Thus the $\mathbf{A}_1, ..., \mathbf{A}_m$ will belong to the family $\mathcal{D}(\mathbf{P})$ of matrices diagonalized by $\mathbf{P}$ which is a Commutative Jordan Algebra, CJA. Then the matrices of $M = \{\mathbf{M}_1, ..., \mathbf{M}_w\}$ belong to a CJA if and only if they commute.

The Commutative Jordan Algebras, CJA, are linear spaces constituted by symmetric matrices that commute, containing the squares of their matrices. [Seely (1971)] showed that every CJA $\mathcal{A}$ has an unique basis, the principal basis $pb(\mathcal{A})$, constituted by orthogonal pairwise orthogonal projection matrices, OPOPM, $\mathbf{Q}_1, \ldots, \mathbf{Q}_w$

Let $\mathbf{M}$ and $\mathbf{W}$ be matrices, belonging to a CJA with principal basis $Q = \{\mathbf{Q}_1, ..., \mathbf{Q}_m\}$, then

$$\begin{cases} \mathbf{M} = \sum_{j=1}^{m} m_j \mathbf{Q}_j \in \mathcal{A} \\ \\ \mathbf{W} = \sum_{j=1}^{w} w_j \mathbf{Q}_j \in \mathcal{A} \end{cases} \tag{2.17}$$

and so

$$\begin{aligned} \mathbf{MW} &= \left(\sum_{j=1}^{m} m_j \mathbf{Q}_j\right)\left(\sum_{j=1}^{w} w_j \mathbf{Q}_j\right) = \\ &= \sum_{j=1}^{m}\sum_{l=1}^{w} m_j w_l \mathbf{Q}_j \mathbf{Q}_l = \\ &= \sum_{j=1}^{m} m_j w_j \mathbf{Q}_j \in \mathcal{A}. \end{aligned} \tag{2.18}$$

which shows that these CJA contain the products of their matrices. Let $\mathbf{Q}$ be a orthogonal projection matrix, OPM, belonging to a CJA with principal basis $\{\mathbf{Q}_1, ..., \mathbf{Q}_w\}$, so that

$$\mathbf{Q} = \sum_{j=1}^{w} q_j \mathbf{Q}_j.$$

Since $\mathbf{Q}$ is idempotent

$$\sum_{j=1}^{w} q_j \mathbf{Q}_j = \mathbf{Q} = \mathbf{Q}^2 = \sum_{j=1}^{w} q_j^2 \mathbf{Q}_j \tag{2.19}$$

and $q_j = q_j^2,\ j = \{1, \ldots, w\}$. Therefore, $q_j = 0$ or $q_j = 1,\ j = \{1, \ldots, w\}$, and

$$\mathbf{Q} = \sum_{j \in \varphi} \mathbf{Q}_j;\ \varphi = \{j' : q_{j'} = 1\} \tag{2.20}$$

so orthogonal projection matrices of a CJA are sums of matrices of their principal basis. Moreover

$$R(\mathbf{Q}) = \boxplus_{j \in \varphi} R(\mathbf{Q}_j)$$

where $R(\mathbf{Q})$ is the range space of $\mathbf{Q}$ and $\boxplus$ indicates orthogonal direct sum of subspaces. Thus,

$$rank(\mathbf{Q}) = \sum_{j \in \varphi} rank(\mathbf{Q}_j),$$

In general for

$$\mathbf{M} = \sum_{j \in \varphi(\mathbf{M})} m_j \mathbf{Q}_j;\ \text{with}\ \varphi(\mathbf{M}) = \{j : m_j \neq 0\}, \tag{2.21}$$

we have

$$\begin{cases} R(\mathbf{M}) = \boxplus_{j \in \varphi(\mathbf{M})} R(\mathbf{Q}_j) \\ \\ rank(\mathbf{M}) = \sum_{j \in \varphi(\mathbf{M})} g_j \end{cases} \tag{2.22}$$

where $g_j = rank(\mathbf{Q}_j)$. Moreover

$$\mathbf{Q} = \sum_{j \in \varphi} \mathbf{Q}_j,$$

this is the orthogonal projection matrices belonging to $\mathcal{A}$ are sums of all or of a subset of set of matrices of the basis $\mathbf{Q}_1, \ldots, \mathbf{Q}_w$, φ can have only one element. When $\varphi = \{h\}$, we have

$$\sum_{j \in \{h\}} \mathbf{Q}_j = \mathbf{Q}_h.$$

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 $pb(\mathcal{A})$.

So, we have the following proposition:

Proposition 3. *If $\boldsymbol{Q}$ is an orthogonal projection matrix, OPM, with rank 1 belonging to a CJA, $\mathcal{A}$, then $\boldsymbol{Q}$ belongs to the principal basis of $\mathcal{A}$.*

Proof: knowing that $\mathbf{Q} = \sum_{j\in\varphi} \mathbf{Q}_j$ and, given matrices that $\mathbf{Q}_1, \ldots, \mathbf{Q}_m$ are mutually orthogonal, then

$$rank(\mathbf{Q}) = \sum_{j\in\varphi} rank(\mathbf{Q}_j)$$

and

$$rank(\mathbf{Q}_j) = 1$$

If $rank(\mathbf{Q}) = 1$ then $\sharp(\varphi) = 1$. Thus $\varphi = \{j'\}$ and $\mathbf{Q} = \mathbf{Q}_{j'}$ $\square$

As to Moore-Penrose inverse, we have

$$\begin{cases} \mathbf{M}\mathbf{M}^{+}\mathbf{M} = \mathbf{M} \\ \mathbf{M}^{+}\mathbf{M}\mathbf{M}^{+} = \mathbf{M}^{+} \end{cases} \tag{2.23}$$

with $\varphi(\mathbf{M}^{+}) = \varphi(\mathbf{M})$. So

$$\mathbf{M}\mathbf{M}^{+} = \mathbf{M}^{+}\mathbf{M} = \sum_{j\in\varphi(\mathbf{M})} \mathbf{Q}_j = \mathbf{Q} \tag{2.24}$$

since $R(\mathbf{M}) = R(\mathbf{M}^{+})$ given that $\mathbf{M}$ is symmetric.
If $\mathbf{M} = \sum_{j=1}^{m} m_j\mathbf{Q}_j \in \mathcal{A}$ with $\{\mathbf{Q}_1, \ldots, \mathbf{Q}_w\}$ the principal basis of $\mathcal{A}$, then

$$\mathbf{M}^{+} = \left(\sum_{j=1}^{m} m_j\mathbf{Q}_j\right)^{+} = \sum_{j=1}^{m} m_j^{+}\mathbf{Q}_j$$

with

1. $m_j^{+} = 0$ if $m_j = 0$
2. $m_j^{+} = m_j^{-1}$ if $m_j \neq 0$; $j = 1, \ldots, w$, so that any CJA contains the Moore-Penrose inverse of its matrices.

We will now define regular CJA, complete CJA and the completation of a CJA. The definitions follow:

Definition 9. *$\mathcal{A}$ is regular, if $\frac{1}{n}\boldsymbol{J}_n = \frac{1}{n}\boldsymbol{1}_n\boldsymbol{1}_n^T \in \mathcal{A}$, we take $\boldsymbol{Q}_1 = \frac{1}{n}\boldsymbol{J}_n$.*

Definition 10. *If $\sum_{j=1}^{w} g_j = n$, then $\sum_{j=1}^{w} \boldsymbol{Q}_j = \boldsymbol{I}_n$, so $\mathcal{A}$ is complete.*

Definition 11. *If* $\sum_{j=1}^{w} \boldsymbol{Q}_j \neq \boldsymbol{I}_n$ *or* $\sum_{j=1}^{w} g_j < n$, *we can add* $\boldsymbol{Q}_{w+1} = \boldsymbol{I}_n - \sum_{j=1}^{w} \boldsymbol{Q}_j$ *to obtain the principal basis of a complete CJA,* $\bar{\mathcal{A}}$. *We say that* $\bar{\mathcal{A}}$ *is the completation of* $\mathcal{A}$.

Also easily seen that, being M regular, we have

$$\begin{cases} \mathbf{M}^{-1} = \sum_{j=1}^{m} m_j^{-1}\mathbf{Q}_j \\ \\ Det(\mathbf{M}) = \prod_{j=1}^{m} m_j^{g_j} \end{cases} \tag{2.25}$$

since $\mathbf{M}$ has eigenvalues $\{m_1, \ldots, m_m\}$ with multiplicities $\{g_1, \ldots, g_m\}$. where $\varphi(\mathbf{M}) = \bar{\bar{w}} = \{1, \ldots, w\}$, $g_j = rank(\mathbf{Q}_j)$, with

$$\mathbf{M} = \sum_{j=1}^{w} m_j \mathbf{Q}_j$$

and

$$R(\mathbf{M}) = \boxplus_{j=1}^{w} R(\mathbf{Q}_j) = \mathbb{R}^n,$$

this implies that the matrix $\mathbf{M}$ is regular. After these considerations, let us establish the following proposition

Proposition 4. *If* $\mathcal{A}_1 \subset \mathcal{A}_2$ *and* $\mathcal{A}_1$ *is complete, then* $\mathcal{A}_2$ *is complete.*

Proof: If $\mathcal{A}_1$ is complete, then $\sum_{j=1}^{m(1)} \mathbf{Q}_j(1) = \mathbf{I}_{m(1)} \in \mathcal{A}_1$, implies that $\sum_{j=1}^{m(2)} \mathbf{Q}_j(2) = \mathbf{I}_{m(2)} \in \mathcal{A}_2$. So, $\mathcal{A}_2$ is complete. □

Now, see [Schott (1997)], the matrices of a family $\mathbf{M} = \{\mathbf{M}_1, \ldots, \mathbf{M}_n\}$ of symmetric matrices commute if and only if they are diagonalized by the same orthogonal matrix $\mathbf{P}$. Let $\mathcal{D}(\mathbf{P})$ be the family of symmetric matrices diagonalized by $\mathbf{P}$. It is easy to see that $\mathcal{D}(\mathbf{P})$ is a CJA, thus, a family of symmetric matrices is contained in a CJA if and only if they commute.
We will now define a CJA generated by $\mathbf{M}$:

Definition 12. *The intersection of CJA is also a CJA, and if the matrices of* $\boldsymbol{M}$ *commute, intersecting all CJA that contains* $\boldsymbol{M}$, *we obtain the smallest CJA,* $\mathcal{A}(\boldsymbol{M})$, *that contains* $\boldsymbol{M}$. *This CJA will be the CJA generated by* $\boldsymbol{M}$.

2.7.2 Binary Operations on CJA

In the section we will study binary operations on CJA. These operations will allow us to build complex models from simple models.

The first of these operations is relevant for model crossing. This binary operation on CJA is called Kronecker product of CJA.

Definition 13. (***Kronecker Product of CJA*** $(\otimes)$)
Given the CJA $\mathcal{A}_l$ with $\boldsymbol{Q}_l = \{\boldsymbol{Q}_{1,l}, \ldots, \boldsymbol{Q}_{w_l,l}\}$ the principal basis of $\mathcal{A}_l$, $l = 1, 2$, $\mathcal{A} = \mathcal{A}_1 \otimes \mathcal{A}_2$ will be the CJA with principal basis,

$$\boldsymbol{Q}_{1,l_1} \otimes \boldsymbol{Q}_{2,l_2} = \boldsymbol{Q}_{(l_1-1)w_2+l_2};\ l_1 = 1, \ldots, w_1,\ l_2 = 1, \ldots, w_2$$

If $\mathcal{A}_1$ and $\mathcal{A}_2$ are regular and constituted by $n_1 \times n_1$ and $n_2 \times n_2$ matrices, since

$$\frac{1}{n_1 n_2}\mathbf{J}_{n_1 n_2} = \frac{1}{n_1}\mathbf{J}_{n_1} \otimes \frac{1}{n_2}\mathbf{J}_{n_2} \in \mathcal{A}_1 \otimes \mathcal{A}_2$$

$\mathcal{A}_1 \otimes \mathcal{A}_2$ will be regular.
If $\mathcal{A}_1$ and $\mathcal{A}_2$ are complete, then

$$\sum_{i=1}^{w_1}\sum_{j=1}^{w_2} \mathbf{Q}_{i,1} \otimes \mathbf{Q}_{j,2} = \left(\sum_{i=1}^{w_1} \mathbf{Q}_{i,1}\right) \otimes \left(\sum_{j=1}^{w_2} \mathbf{Q}_{j,2}\right) = \mathbf{I}_{n,1} \otimes \mathbf{I}_{n,2} = \mathbf{I}_{n_1 n_2}$$

And so, $\mathcal{A}_1 \otimes \mathcal{A}_2$ will be complete.

It is shown in [Fonseca et al. (2006)], that if $\mathcal{A}_1$, $\mathcal{A}_2$ and $\mathcal{A}_3$ are CJA, then

$$\mathcal{A}_1 \otimes (\mathcal{A}_2 \otimes \mathcal{A}_3) = (\mathcal{A}_1 \otimes \mathcal{A}_2) \otimes \mathcal{A}_3.$$

The restricted product $*$ of CJA is used in the construction of nested models. A nested model is when each treatment a model nests all treatment of other model, it is said that the first model nests the second. So, the definition of restricted product follow:

Definition 14. (***Restricted Product of CJA***)
Let $\mathcal{A}_1 * \mathcal{A}_2$ be the CJA with principal basis

$$\left\{\boldsymbol{Q}_{1,1} \otimes \frac{1}{n_2}\boldsymbol{J}_{n_2}, \ldots, \boldsymbol{Q}_{1,w_1} \otimes \frac{1}{n_2}\boldsymbol{J}_{n_2}\right\} \cup \left\{\boldsymbol{I}_{n_1} \otimes \boldsymbol{Q}_{2,2}, \ldots, \boldsymbol{I}_{n_1} \otimes \boldsymbol{Q}_{2,w_2}\right\}.$$

When $\mathcal{A}_1$ is regular, $\mathcal{A}_1 * \mathcal{A}_2$ is regular. Moreover, if $\mathcal{A}_1$ is complete and $\mathcal{A}_2$ is regular, we have

$$\sum_{j=1}^{w_1}\left(\mathbf{Q}_{j,1}\otimes\frac{1}{n_2}\mathbf{J}_{n_2}\right)=\left(\sum_{j=1}^{w_1}\mathbf{Q}_{j,1}\right)\otimes\frac{1}{n_2}\mathbf{J}_{n_2}=\mathbf{I}_{n_1}\otimes\frac{1}{n_2}\mathbf{J}_{n_2}=\mathbf{I}_{n_1}\otimes\mathbf{Q}_{1,2}, \quad (2.26)$$

we assume that $\mathbf{Q}_{1,2}=\frac{1}{n_2}\mathbf{J}_{n_2}$.

If $\mathcal{A}_1$ is complete and $\mathcal{A}_2$ is regular and complete, $\mathcal{A}_1 * \mathcal{A}_2$ is complete. So

$$\begin{aligned}\sum_{j=1}^{w_1}\left(\mathbf{Q}_{j,1}\otimes\frac{1}{n_2}\mathbf{J}_{n_2}\right)+\sum_{j=2}^{w_2}\left(\mathbf{I}_{n_1}\otimes\mathbf{Q}_{j,2}\right)&=\mathbf{I}_{n_1}\otimes\mathbf{Q}_{1,2}+\mathbf{I}_{n_1}\otimes\sum_{j=2}^{w_2}\mathbf{Q}_{j,2}\\&=\mathbf{I}_{n_1}\otimes\sum_{j=1}^{w_2}\mathbf{Q}_{j,2}.\end{aligned} \quad (2.27)$$

Clearly that $\mathcal{A}_1 * \mathcal{A}_2$ is complete and regular. Now let $\mathcal{A}_3$ be other CJAS with principal basis $\mathbf{Q}_{3,1},\ldots,\mathbf{Q}_{3,w_3}$ assuming $\mathbf{Q}_{3,1}=\frac{1}{n_3}\mathbf{J}_{n_3}$, we prove that, see [Fonseca et al. (2006)].

$$(\mathcal{A}_1 * \mathcal{A}_2) * \mathcal{A}_3 = \mathcal{A}_1 * (\mathcal{A}_2 * \mathcal{A}_3).$$

2.8 Important results of Commutative Jordan Algebra

In this section, we will develop propositions that will give us important results. This first proposition shows that the product of any matrices in CJA commutes with any orthogonal projection matrix, OPM, in the same CJA.

Proposition 5. *If $\boldsymbol{M}_1,\ldots,\boldsymbol{M}_d$ commute, then $\boldsymbol{M}=\{\boldsymbol{M}_1,\ldots,\boldsymbol{M}_d\}$ generates an algebra $\mathcal{A}=\mathcal{A}(\boldsymbol{M})$. Knowing that*

1. $\boldsymbol{Q}=\{\boldsymbol{Q}_1,\ldots,\boldsymbol{Q}_d\}$ *form the principal basis of the algebra* $\mathcal{A}$;
2. $\boldsymbol{U}=\sum_{j=1}^{h}\boldsymbol{Q}_j$, *with* $h\le d$;
3. $\boldsymbol{M}_i=\sum_{j=1}^{h}b_{i,j}\boldsymbol{Q}_j$.

Given this, $\boldsymbol{U}\boldsymbol{M}_i=\boldsymbol{M}_i\boldsymbol{U}$, in other words, $\boldsymbol{U}$ and $\boldsymbol{M}_i$ commute.

Proof:

$$\mathbf{M}_i\mathbf{U} = \left(\sum_{j=1}^{h} b_{i,j}\mathbf{Q}_j\right)\left(\sum_{j=1}^{h}\mathbf{Q}_j\right) = \sum_{j=1}^{h} b_{i,j}\mathbf{Q}_j = \mathbf{M}_i$$

$$\mathbf{U}\mathbf{M}_i = \left(\sum_{j=1}^{h}\mathbf{Q}_j\right)\left(\sum_{j=1}^{h} b_{i,j}\mathbf{Q}_j\right) = \sum_{j=1}^{h} b_{i,j}\mathbf{Q}_j = \mathbf{M}_i.$$

If $\mathbf{M}_i\mathbf{U} = \mathbf{U}\mathbf{M}_i = \mathbf{M}_i$, we conclude that $\mathbf{U}\mathbf{M}_i = \mathbf{M}_i\mathbf{U}$. □

The following proposition shows us that the product of two orthogonal projection matrices is an orthogonal projection matrix.

Proposition 6. *If $\boldsymbol{Q}_1$ and $\boldsymbol{Q}_2$ are orthogonal projection matrices, then*

$$\boldsymbol{Q}_1\boldsymbol{Q}_2 = \boldsymbol{Q}_2\boldsymbol{Q}_1 \Longleftrightarrow \boldsymbol{U} = \boldsymbol{Q}_1\boldsymbol{Q}_2,$$

if $\boldsymbol{U}$ is an orthogonal projection matrix.

Proof:

1. We admit that $\mathbf{Q}_1\mathbf{Q}_2$ commute, then

   $$\mathbf{U} = \mathbf{Q}_1\mathbf{Q}_2 = \mathbf{Q}_2\mathbf{Q}_1 = \mathbf{Q}_2^T\mathbf{Q}_1^T = (\mathbf{Q}_1\mathbf{Q}_2)^T = \mathbf{U}^T,$$

   and $\mathbf{U}$ is symmetric.

   We saw that:

   $$\mathbf{U}^2 = \mathbf{Q}_1\mathbf{Q}_2\mathbf{Q}_1\mathbf{Q}_2 = \mathbf{Q}_1\mathbf{Q}_1\mathbf{Q}_2\mathbf{Q}_2 = \mathbf{Q}_1\mathbf{Q}_2 = \mathbf{U},$$

   and $\mathbf{U}$ is idempotent. As $\mathbf{U}$ is symmetric and idempotent, $\mathbf{U}$ is an orthogonal projection matrix.

2. If $\mathbf{U} = \mathbf{Q}_1\mathbf{Q}_2$ is orthogonal projection matrix, is symmetric. So, we have:

   If $\mathbf{U} = \mathbf{Q}_1\mathbf{Q}_2$, then $\mathbf{U}^T = (\mathbf{Q}_1\mathbf{Q}_2)^T = \mathbf{Q}_2^T\mathbf{Q}_1^T = \mathbf{Q}_2\mathbf{Q}_1$,

   thus $\mathbf{U} = \mathbf{Q}_1\mathbf{Q}_2 = \mathbf{Q}_2\mathbf{Q}_1$.

   So, we conclude that $\mathbf{Q}_1\mathbf{Q}_2$ commute. □

Now, we have one more proposition.

Proposition 7. *Let $\boldsymbol{Q}$ be an orthogonal projection matrix with characteristic 1 and belonging the principal basis of the algebra $\mathcal{A}$, show that $\boldsymbol{Q} = \boldsymbol{Q}_{j'}$.*

Proof: We know that

$$\mathbf{Q} = \sum_{j\in\varphi} \mathbf{Q}_j$$

and that

$$car(\mathbf{Q}) = \sum_{j\in\varphi} car(\mathbf{Q}_j)$$

If $car(\mathbf{Q}) = 1$, then, we must have $\varphi = \{j\}$, being

$$\mathbf{Q} = \mathbf{Q}_{j'}$$

$\square$

Knowing that:

1. $\mathbf{Q} = \{\mathbf{Q}_1, \ldots, \mathbf{Q}_m\}$ is the principal basis of algebra $\mathcal{A}$;
2. $\varphi \subseteq \overline{\overline{m}} = \{1, \ldots, m\}$, φ constitute a partition of $\overline{\overline{m}} = \{1, \ldots, m\}$;
3. $\mathbf{Q}(\varphi) = \{\mathbf{Q}_j,\, j \in \varphi\} = pb(\mathcal{A}(\varphi))$;
4. being $\varphi^c = \overline{\overline{m}} \setminus \varphi$.

We now have

$$\mathbf{Q}(\varphi^c) = \{\mathbf{Q},\, j \in \varphi^c\} = pb(\mathcal{A}(\varphi^c)).$$

So, we have the following partition

$$\mathbf{M} = \sum_{j=1}^{m} m_j \mathbf{Q}_j = \mathbf{M}(\varphi) + \mathbf{M}(\varphi^c) \tag{2.28}$$

where

$$\mathbf{M}(\varphi) = \sum_{j\in\varphi} m_j \mathbf{Q}_j$$

and

$$\mathbf{M}(\varphi^c) = \sum_{j\in\varphi^c} m_j \mathbf{Q}_j.$$

We now have another partition

$$\mathbf{M} = \mathbf{M}^\circ(\varphi) + \mathbf{M}^\circ(\varphi^c) \tag{2.29}$$

Of (2.28) and (2.29), we want to show that there is a single partition, then we have

$$\mathbf{M}(\varphi) + \mathbf{M}(\varphi^c) = \mathbf{M}^\circ(\varphi) + \mathbf{M}^\circ(\varphi^c).$$

Being $\mathbf{Q}_j\mathbf{Q}_{j'}$ orthogonal projection matrices, $\mathbf{Q}_j\mathbf{Q}_{j'} = \mathbf{0}$, we have

$$\mathbf{M}(\varphi) - \mathbf{M}^\circ(\varphi) = \mathbf{M}^\circ(\varphi^c) - \mathbf{M}(\varphi^c) = \mathbf{0}.$$

So

$$\mathbf{M}(\varphi) - \mathbf{M}^\circ(\varphi) \in \mathcal{A}(\varphi)$$

and

$$\mathbf{M}^\circ(\varphi^c) - \mathbf{M}(\varphi^c) \in \mathcal{A}(\varphi^c)$$

then

$$\mathcal{A}(\varphi) \cap \mathcal{A}(\varphi^c) = \{\mathbf{0}\}.$$

We conclude that

$$\mathcal{A} = \mathcal{A}(\varphi) \boxplus \mathcal{A}(\varphi^c), \tag{2.30}$$

is a single partition with orthogonal matrices. we will now show the last proposition of this section:

Proposition 8. *If $\boldsymbol{T}_l \in \mathcal{A}(l)$, $T_l = \sum_{j \in \varphi(l)} \boldsymbol{Q}_j(l)$, $l = 1, 2$ and*

$$\mathcal{A}(1) \otimes \mathcal{A}(2) = \{\boldsymbol{M}(1) \otimes \boldsymbol{M}(2);\ \boldsymbol{M}(l) \in \mathcal{A}(l)\ \ l = 1, 2\}$$

then, we have the following thesis

$$\boldsymbol{T}_1 \otimes \boldsymbol{T}_2 \in \mathcal{A}(1) \otimes \mathcal{A}(2). \tag{2.31}$$

Proof:

$$\mathbf{T}_1 \otimes \mathbf{T}_2 = \sum_{j \in \varphi(1)} \mathbf{Q}_j(1) \otimes \sum_{j \in \varphi(2)} \mathbf{Q}_j(2) = \sum_{j_1 \in \varphi(1)} \sum_{j_2 \in \varphi(2)} \mathbf{Q}_{j_1}(1) \otimes \mathbf{Q}_{j_2}(2) \in \mathcal{A}(1) \otimes \mathcal{A}(2)$$

$\square$

Chapter 3

Estimation and models

3.1 Introduction

In this section, we will study the estimation and some models like for example, linear models, orthogonal models and nesting models. Firstly, are given definitions of estimators and their properties. The result on estimation will refer the use of sufficient and complete statistics to obtain good pontwise estimators. Then, we will define linear models and their special types are given, like Gauss-Markov Models, Generalized Gauss-Markov Models and definitions of orthogonal block structure, OBS (see definition (33)) and commutative orthogonal block structure, COBS (see definition (34)). We will get generalized results in characterization and expression of the best linear unbiased estimators, BLUE. In section (3.6), we will study a very important class of models that is the class of Orthogonal Models. In this section, we will enunciate and demonstrate results that will be used in the following chapters.

We now study the crossing and nesting of orthogonal models. We start by presenting some matrix results that we will use. Namely we will considerer commutative Jordan algebras of symmetric matrices (see section 2.7.1). Next we use these results to study the algebraic structure of ORT models, in such a way that we can analyse the crossing and nesting of such models.

3.2 Estimators

Initially we assume that

$$P = \{P_\theta, \theta \in \Omega\} \tag{3.1}$$

being P_θ a joint probability distribution, belonging to some known class P. θ is the parameter of distribution taking values in a set Ω, with the aim to specify a plausible value, in other words, to estimate the value θ.

We define an important concept as a way to determine suitably an estimator.

Definition 15. *An estimator is a real-valued function g defined over the sample space. It is used to estimate an estimand, $g(\theta)$, which is a real-valued function of the parameter.*

Quite generally suppose that the consequences of estimating $g(\theta)$ by a valued t are measured by $L(t, \theta)$. So, we have the definition of loss function.

Definition 16. *Let T be an estimator of $g(\theta)$. The real valued function $L(t, \theta)$ is a loss function if it respect the following conditions:*

1. *$L(t, \theta) \geq 0$, for all t, θ;*
2. *$L(g(\theta), \theta) = 0$, for all θ*

When the loss is zero, the estimated is the correct value. So, we can measure the inaccuracy of an estimator T from its risk function, defined as

Definition 17. *Let T be an estimator of $g(\theta)$ and $L(t, \theta)$ its loss function. The risk function is*

$$R(t, \theta) = E_\theta \; L[g(t(X), \theta)]. \tag{3.2}$$

The average loss comes from the use of T. The goal is find a T which minimizes the risk for all values of θ. We will later develop this theme, how to find the best estimator. This topics can be studied in detail in [Lehmann & Casela (1998)].

3.3 Exponential families

The family P_θ of the expression (3.1) of distributions is denominated as a k-dimensional exponential family if the distributions p_θ with

$$p_\theta = h(x)\ exp\left[\sum_{i=1}^{k} \eta_i(\theta) T_i(x) - B(\theta)\right]. \tag{3.3}$$

The η_i and B are real-valued functions of the parameters, the T_i are real-valued sufficient statistics and x is a given observation in the sample space χ, the support of the distribution. Usually, the η_i are known as the natural parameters of the density, and so we can write the density in the canonical form

$$p(x|\eta) = h(x)\ exp\left[\sum_{i=1}^{k} \eta_i T_i(x) - A(\eta)\right] \tag{3.4}$$

The natural parameter space will be denoted by Ξ

For exponential families, the moment and cumulant generating functions can be expressed from the following theorem

Theorem 6. *If $\boldsymbol{X}$ is a random vector with density (3.4), then for any $\eta \in \Xi$, the moment and cumulant generating functions $M_T(u)$ and $K_T(u)$, of the T's exist in some neighborhood of zero and are given by*

$$M_T(u) = \frac{e^{A(\eta+u)}}{e^{A(\eta)}} \tag{3.5}$$

and

$$K_T(u) = A(\eta + u) - A(\eta) \tag{3.6}$$

respectively

Proof: See [Lehmann & Casela (1998)], pages 27-28. □

3.4 Sufficient Statistics

Suppose we have a random sample $\mathbf{X}_1, ..., \mathbf{X}_n$. If $T = T(\mathbf{X}_1, ..., \mathbf{X}_n)$ is a statistic (not necessarily real-valued) and t is a value taken by T, then the conditional joint distribution of $\mathbf{X}_1, ..., \mathbf{X}_n$ given that $T = t$ can be calculated. So, for each value of t, there will be a family of possible conditional distributions of $(\mathbf{X}_1, ..., \mathbf{X}_n)$ corresponding to the different possible values of θ.

Definition 18. *:*

A statistic T is sufficient for θ if $P_\theta((\boldsymbol{X}_1, ..., \boldsymbol{X}_n)|T(\boldsymbol{X}_1, ..., \boldsymbol{X}_n) = T(x_1, ..., x_n))$ is not a function of θ.

We can say that a statistic $T(\mathbf{X}_1, ..., \mathbf{X}_n)$ is said to be sufficient for θ if the distribution of $\mathbf{X}_1, ..., \mathbf{X}_n$, given $T = t$, does not depend on θ for any value of t. This definition is not precise and we shall return to it later in this section. The definition of sufficient statistic is very hard to verify directly. A much easier way to find sufficient statistics is through the factorization theorem that we now state. see [Lehmann & Casela (1998)], page 35.

Theorem 7. *(**Factorization Criterion**) A necessary and sufficient condition for a statistic T to be sufficient for a family $P = \{P_\theta, \theta \in \Omega\}$ of distributions of $\boldsymbol{X}$ there exist non-negative function g_θ and h such that the densities p_θ satisfy*

$$p_\theta(x) = g_\theta[T(x)]h(x)(a.e.\mu) \tag{3.7}$$

Clearly, sufficient statistics are not unique. From theorem (7), it is easy to see that

1. The identity function $T(x_1, ..., x_n) = (x_1, ..., x_n)$ is a sufficient statistic vector and
2. If T is a sufficient statistic for θ then so is any statistic with a 1:1 relation with T.

This leads to the notion of a minimal sufficient statistic.

Definition 19. *A statistic that is sufficient and that is a function of all other sufficient statistics is called a minimal sufficient statistic.*

Before we enunciate the next corollary, we have the following definition:

Definition 20. *A density belonging to the exponential family is of full rank if:*

1. *The T_i, and η_i, $i = 1, \ldots, k$ are linearly independent;*
2. *and Ξ contains a k-dimensional rectangle.*

Now, we have the following proposition:

Proposition 9. *(Exponential Families) Let $\boldsymbol{X}$ be distributed with density (3.4). Then, $T = (T_1, ..., T_s)$ is minimal sufficient provided the family (3.4) satisfies one the following conditions:*

1. It is of full rank.

2. The parameter space contains s+1 points $\eta^{(j)}(j = 0, ..., s)$, which span E_s, in the sense that they do not belong to a proper affine subspace of E_s.

Proof: See [Lehmann & Casela (1998)], page 39. $\square$

Let $\mathbf{X}_1, ..., \mathbf{X}_n$, be i.i.d., with exponential density of full rank. Then, the joint distribution of $\mathbf{X}_1, ..., \mathbf{X}_n$ is a full-rank exponential density with $T = (T_1^*, ..., T_s^*)$ where $T_i^* = \sum_{j=1}^{n} T_i(X_j)$. So, for a sample from the exponential family (3.4), it is possible to reduce the data to an s-dimensional sufficient statistic, regardless of the sample size. The reduction of a sample to a smaller number of sufficient statistics greatly simplifies the statistical analysis. We point out that the sufficient statistics contain all the information in the sample.

Now, we have the definitions of ancillary and complete statistic:

Definition 21. *A statistic $V = V(\boldsymbol{X})$ is ancillary if the distribution of U does not depend on θ for every $P_\theta(\theta \in \Omega)$.*

Definition 22. *A statistic $S = S(\boldsymbol{X})$ is said to be complete if for real-valued function f, $E_\theta[f(S)] = 0$ for real all $\theta \in \Omega$ implies $P_\theta[f(S) = 0] = 1$ for all $\theta \in \Omega$.*

What happens to the ancillary statistics when the minimal sufficient statistic is complete is show by the following result

Theorem 8. ***(Basu's Theorem)*** *If T is a complete sufficient statistic for the family $P = P_\theta, \theta \in \Omega$, then any ancillary statistic V is independent of T.*

Proof: [Lehmann & Casela (1998)], page 42. $\square$

Theorem 9. *If $\boldsymbol{X}$ is distributed according to the exponential family (3.4) and the family is of full rank, then $T = [T_1(\boldsymbol{X}), ..., T_s(\boldsymbol{X})]$ is complete.*

Proof: See [Lehmann et al. (1997)], pages 142-143. $\square$

The property of convexity and of concavity play an important role in point estimation

Definition 23. *A real-valued function ϕ defined over an open interval I = (a,b) with $-\infty \leq a < b \leq \infty$ is convex if for any $a < x < y < b$ and any $0 < y < 1$*

$$\phi[Yx + (1-Y)y] \leq Y\phi(x) + (1-Y)\phi(y). \tag{3.8}$$

The function is said to be strictly convex if the strict inequality holds in (3.8) for all indicated values of x, y and Y. A function ϕ is concave on (a,b) if $-\phi$ is convex.

The notion of convexity is applied to a result that is known as Jensen's Inequality

Theorem 10. ***(Jensen' Inequality)*** *If ϕ is a convex function defined over an open interval I, and $\boldsymbol{X}$ is a random variable with $P(\boldsymbol{X} \in I) = 1$ and finite expectation, then*

$$\phi[E(\boldsymbol{X})] \leq E[\phi(\boldsymbol{X})]. \tag{3.9}$$

If ϕ is strictly convex, the inequality is strict unless $\boldsymbol{X}$ is a constant with probability 1.

Proof: See [Lehmann & Casela (1998)], page 46. □

The statistics are related with the concept of information. An estimator carries a information that will be useful in the interpretation of data. So, we have the following definition of information matrix:

Definition 24. *Let $\boldsymbol{X} = (\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n)^T$ be a sample vector belonging the $\mathcal{P}$ as defined in* (3.1), *with $\theta = (\theta_1, \ldots, \theta_r)^T$ and let T be a sufficient statistic. The element $I_{ij}(\theta)$ of the information matrix $I(\theta)$ is defined of following form:*

$$I_{ij} = E\left[\frac{\partial}{\partial\theta_i} log(P_\theta(\boldsymbol{X})) \frac{\partial}{\partial\theta_j} log(P_\theta(\boldsymbol{X}))\right]$$

The information matrix essentially measures the variation of P_θ with θ. So, it is possible to obtain a bound for the variance of an estimator. Now, we have the following theorem:

Theorem 11. ***(Multiparameter information inequality)***

Let $\boldsymbol{X} = (\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n)^T$ be a sample vector belonging the $\mathcal{P}$ as defined in (3.1), *with $\theta = (\theta_1, \ldots, \theta_r)^T$. Let T be any statistics for which $E[T^2] < \infty$ and either*

1. *For $i = 1, \ldots, r$; $\frac{\partial}{\partial\theta_i} E[T] = E\left[\frac{\partial}{\partial\theta_i} T\right]$ exists, or*

2. There exists functions $b_{\theta,i}(.)$; $i = 1, \ldots, r$, with $E[b^2_{\theta,i}(\boldsymbol{X})] < \infty$ that satisfy

Proof: See [Lehmann & Casela (1998)], page 127. □

Lemma 1. *If T_0 is any unbiased estimator of $g(\theta)$, the totality of unbiased estimators is given by $T = T_0 - U$, where U is any unbiased estimator of zero, that is, it satisfies*

$$E[U] = 0$$

for all $\theta \in \Omega$

Proof: It is simple, because

$$E[T] = E[T_0] - E[U] = E[T_0] = g(\theta)$$

□

The following lemma states that there exists the uniqueness of unbiased estimators given a complete sufficient statistics.

Lemma 2. *Let $g(\theta)$ be an estimable function and T a complete sufficient statistic. If f and h are functions, then we have*

$$E[f(T)] = E[h(T)] = g(\theta),$$

and

$$P[f \equiv h] = 1$$

Proof: See [Lehmann & Casela (1998)], page 87-88. □

The estimator of a parameter will be a statistic whose values are "approximation" the true values of the parameter. We define uniformly minimum variance unbiased estimation, UMVUE. The definition follows:

Definition 25. (**UMVUE**) *It says that T is UMVUE of $g(\theta)$ if it is unbiased, that is:*

$$E[T] = g(\theta) \text{ for all } \theta \in \Omega$$

and if for all θ has $V(T) \leq V(U)$, where U is another unbiased estimator of $g(\theta)$.

The next theorem states that the class of all unbiased estimators of zero is important to distinguish the more efficient of less efficient. Thus, we have

Theorem 12. *Let $\boldsymbol{X}$ have distribution P_θ; $\theta \in \Omega$. Let T be an unbiased estimator of $g(\theta)$ such that $E[T^2] < \infty$. Let $\mathcal{U}$ denote the set of all unbiased estimators of zero which such that $E[T^2] < \infty$. Then a necessary and sufficient condition for T to be a UMVUE estimator of $g(\theta)$ is*

$$E[TU] = cov[T, U] = 0,\ U \in \mathcal{U} \text{ and } \theta \in \Omega.$$

Proof: See [Lehmann & Casela (1998)], page 86. □

The Rao-Blackwell theorem can be used to improve unbiased estimators reach the best unbiased estimator, we now establish

Theorem 13. ***(Rao-Blackwell theorem)*** *Let g be a convex function and $\hat{g}$ an unbiased estimator of $g(\theta)$. Given a sufficient statistic T, we have*

1. *$\tilde{g}(t) = E[\hat{g}(\boldsymbol{X}) \mid T{=}t]$ is function of sufficient statistic $T{=}t$ but is not function of the θ;*

2. *$\tilde{g}(T)$ is unbiased estimator of $g(\theta)$;*

3. *$V(\tilde{g}) \leq V(\hat{g})$ for all θ.*

Proof: See [Lehmann & Casela (1998)], page 48. □

Before enunciate the next theorem, we will define the following:

Definition 26. *A sufficient statistic T is complete in the family of measures if*

$$E\left(g\left(T(\boldsymbol{X}_1, ..., \boldsymbol{X}_n)\right)\right) = 0 \Rightarrow P\left(g(T) = 0\right) = 1, \forall_\theta \tag{3.10}$$

Using jointly the concepts of sufficiency and completeness statistics, we get the following result very important.

Theorem 14. ***(Blackwell-Lehman-Sheffé theorem)***

If T is complete and sufficient statistics and there exist an unbiased estimator $\hat{g}$ of $g(\theta)$, $\tilde{g} = [\hat{g}(\boldsymbol{X}) \mid T{=}t]$ is an unbiased estimator of $g(\theta)$ with uniformly mean variance, UMVUE.

Proof: Considering the theorem of Rao-Blackwell, we know that $\tilde{\mathrm{g}}$ is unbiased estimator of $g(\theta)$ and $\mathrm{V}(\hat{\mathrm{g}}) \leq \mathrm{V}(\tilde{\mathrm{g}})$. Now assuming that there exists another unbiased estimator g^* of $g(\theta)$. Again by theorem (13), we know that $\tilde{g}^* = E[g^*(\mathbf{X})|T = t]$ is function of t but is not of $g(\theta)$ and $V(\tilde{g}^*) \leq V(\tilde{g})$. $\square$

The Blackwell-Lehman-Sheffé theorem states that if we find an unbiased estimator that is a function of a complete sufficient statistic, it is the unique UMVUE.

3.5 Estimation on linear models

A random variable described by a linear model is a sum of several terms and it has the following structure

$$\mathbf{Y} = \sum_{i=1}^{w} \mathbf{X}_i \boldsymbol{\beta}_i \tag{3.11}$$

where the parameters $\boldsymbol{\beta}_i$, $i = \{1, \ldots, w\}$ are fixed unknown constants or random variables.

We now define the concepts of fixed effects model, random effects model and mixed effects model.

Definition 27. ***(fixed effects model)***

A linear model

$$\boldsymbol{Y} = \sum_{i=1}^{w} \boldsymbol{X}_i \boldsymbol{\beta}_i + \boldsymbol{\epsilon} \tag{3.12}$$

the vectors $\boldsymbol{\beta}_i$, $i = \{1, \ldots, w\}$, are fixed and unknown and $\boldsymbol{\epsilon}$ is a random vector or random error with null mean and covariance matrix $\theta \boldsymbol{V}$, with $\boldsymbol{V}$ a known matrix and θ unknown, will be a fixed effects model.

This model is important in the application for analysis of variance and linear regression models. This is the most commonly used type of model used in real applications.

We will considerer models with random parameters. We have the following definition.

Definition 28. ***(Random effects model)***

A linear model

$$\boldsymbol{Y} = \boldsymbol{1}\mu + \sum_{i=1}^{w} \boldsymbol{X}_i \boldsymbol{\epsilon}_i, \tag{3.13}$$

where μ is fixed and unknown and $\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_w$ are random independent with null mean vectors and covariance matrices $\theta_1 \boldsymbol{V}_1, \ldots, \theta_w \boldsymbol{V}_w$, with $\boldsymbol{V}_1, \ldots, \boldsymbol{V}_w$ known, it is a random effects model.

This model is applied in analysis of variance of random effects.

After defining these two models, the definition of mixed model arises normally

Definition 29. ***(Mixed models)***

A linear model

$$\boldsymbol{Y} = \sum_{i=1}^{k} \boldsymbol{X}_i \boldsymbol{\beta}_i + \sum_{i=k+1}^{w} \boldsymbol{X}_i \boldsymbol{\epsilon}_i \tag{3.14}$$

with $k < w$, $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta_k}$ fixed and unknown. $\boldsymbol{\epsilon}_{k+1}, \ldots, \boldsymbol{\epsilon}_k$ are random independent vectors (random errors) with null mean vectors and covariance matrices $\theta_i V_i$, $i = \{k+1, \ldots, w\}$. These models are called mixed effects models and have fixed and random effects.

A particular case of the fixed effects model is

Definition 30. ***(Gauss-Markov model)***

A fixed effects model

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

with $\boldsymbol{\epsilon} \sim N(0, \theta \boldsymbol{I}_n)$, is a Gauss-Markov model

However there are situations in which the $\mathbf{Y}_1, \ldots, \mathbf{Y}_w$ have non null covariance, so we consider

Definition 31. ***(Generalized Gauss-Markov model)***

A fixed effects model

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

with $\boldsymbol{\epsilon} \sim N(0, \theta \boldsymbol{V})$, where $\boldsymbol{V}$ is a known positive definitive matrix, is a generalized Gauss-Markov model.

These models are applied in usual regression and heterocedastic regression.

We still have the following definition:

Definition 32. (***BLUE***) *A function $(\boldsymbol{a}^T \boldsymbol{Y})$ is a BLUE of $\mathbb{E}[\boldsymbol{a}^T \boldsymbol{Y}]$ if any other $(\boldsymbol{b}^T \boldsymbol{Y})$ unbiased estimator of $\mathbb{E}[\boldsymbol{a}^T \boldsymbol{Y}]$ has a variance that is not lesser than that of $(\boldsymbol{a}^T \boldsymbol{Y})$.*

Now, for vector estimates we have the following versions of the Gauss-Markov theorems, see [Kariya & Kurata (2004)], page 34:

Theorem 15. (***Gauss-Markov Theorem***) *Let* $\boldsymbol{Y}$ *be a Gauss-Markov model, we have the estimator of the form*

$$\tilde{\boldsymbol{\beta}} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{Y}$$

which we call the Gauss-Markov Estimator, GME. The GME is the BLUE of $\boldsymbol{\beta}$, *that is, the GME is the unique estimator that satisfies*

$$cov(\boldsymbol{\beta}^*) - cov(\tilde{\boldsymbol{\beta}}) \text{ is positive definite},$$

with $\boldsymbol{\beta}^*$ *any other linear estimator. The covariance matrix of* $\tilde{\boldsymbol{\beta}}$ *is given by*

$$cov(\tilde{\boldsymbol{\beta}}) = \boldsymbol{\theta}(\boldsymbol{X}^T\boldsymbol{X})^{-1}$$

Proof: see [Kariya & Kurata (2004)], page 34. □

Theorem 16. (***Generalized Gauss-Markov Theorem***) *Let* $\boldsymbol{Y}$ *be a Generalized Gauss-Markov model, we have the estimator of the form*

$$\tilde{\boldsymbol{\beta}} = (\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{Y}$$

which we call the Gauss-Markov Estimator, GME. GME is the BLUE of $\boldsymbol{\beta}$, *that is, the GME is the unique estimator that satisfies*

$$cov(\boldsymbol{\beta}^*) - cov(\tilde{\boldsymbol{\beta}}) \text{ is positive definite},$$

with $\boldsymbol{\beta}^*$ *any other linear estimator. The covariance matrix of* $\tilde{\boldsymbol{\beta}}$ *is given by*

$$cov(\tilde{\boldsymbol{\beta}}) = \boldsymbol{\theta}(\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}$$

Proof: see [Kariya & Kurata (2004)], page 34. □

Given the lack of homocedasticity and independence of errors, we transform $\mathbf{Y}$, to get the homocedasticity. With $\mathbf{G} \in \mathcal{U}(\mathbf{V})$, we take

$$\begin{cases} \mathbf{Y}^* = \mathbf{G}\mathbf{Y} \\ \mathbf{X}^* = \mathbf{G}\mathbf{X} \end{cases}$$

getting

$$\begin{cases} E[\mathbf{Y}^*] = \mathbf{X}^*\boldsymbol{\beta} \\ Cov[\mathbf{Y}^*] = \mathbf{G}^T Cov[\mathbf{Y}]\mathbf{G} = \theta\mathbf{I}_n \end{cases}$$

So, we now have the homocedastic model

$$\mathbf{Y}^* = \mathbf{X}^*\boldsymbol{\beta} + \boldsymbol{\epsilon}^*$$

We also have the following mixed model,

$$\mathbf{Y} = \sum_{i=0}^{w} \mathbf{X}_i\boldsymbol{\beta}_i \tag{3.15}$$

where $\boldsymbol{\beta}_0$ is fixed and the $\boldsymbol{\beta}_j$, $j = 1, \ldots, w$ are independent random vectors with null mean vectors and variance covariance matrices $\theta_j\mathbf{I}_{c_j}$, $j = 1, \ldots, w$. The matrices $\mathbf{X}_i$, $i = 0, \ldots, w$ are known.

Putting $\mathbf{M}_i = \mathbf{X}_i\mathbf{X}_i^T$, $i = 0, \ldots, w$, $\mathbf{Y}$ will have mean vector and variance covariance matrix

$$\begin{cases} \boldsymbol{\mu} = \mathbf{X}_0\boldsymbol{\beta}_0 \\ \mathbf{V} = \sum\limits_{i=1}^{w} \theta_i\mathbf{M}_i \end{cases} \tag{3.16}$$

The θ_i, $i = 1, \ldots, w$ will be the variance components.

So, we have the mixed model of the expression 3.15, and in the first expression of 3.16, the mean vector which spans the range space $\Omega = R(\mathbf{X}_0)$ of $\mathbf{X}_0$. The orthogonal projection matrices on Ω will be

$$\mathbf{T} = \mathbf{X}_\circ(\mathbf{X}_\circ^T\mathbf{X}_\circ)^+\mathbf{X}_\circ^T = \mathbf{X}_0\mathbf{X}_0^+ \tag{3.17}$$

We now have the

Definition 33. *A linear model has orthogonal block structure (OBS) if the variance-covariance matrix has the form*

$$\boldsymbol{V} = \sum_{j=1}^{m} \gamma_j\boldsymbol{Q}_j \tag{3.18}$$

where $\boldsymbol{Q}_1, ..., \boldsymbol{Q}_m$ are orthogonal pairwise projection matrices, OPOPM, such that

$$\sum_{j=1}^{m} \boldsymbol{Q}_j = \boldsymbol{I}_m$$

These models allow optimal estimation for the canonical variance components $\gamma_1, \ldots, \gamma_m$. These important class of models were introduced by [Nelder (1965b)] and continue to play an important role in the theory of randomized block design, see [Calinski & Kageyama (2000), (2003)]. There are many interesting papers on OBS such as [Houtman & Speed (1983)] and [Mejza (1992)].
We also have the following definition:

Definition 34. *A linear model has COBS (Commutative Orthogonal Block Structure), if it have OBS and,*

$$\boldsymbol{T}\boldsymbol{Q}_j = \boldsymbol{Q}_j\boldsymbol{T},\ j = \ \ldots, m$$

These models were introduced by [Fonseca et al. (2008)] and have been studied in [Carvalho et al. (2008), (2013)] and [Mexia et al. (2010)]. We have the following results whose proofs may be seen in [Fonseca et al. (2008)].

Proposition 10. *When the matrices $\boldsymbol{M}_1, \ldots, \boldsymbol{M}_w$ commute, the model has OBS.*

Proposition 11. *If the matrices $\boldsymbol{M}_1, \ldots, \boldsymbol{M}_w$ and T commute, the model has COBS.*

Moreover, see [Zmyślony (1978)], we have optimal results for estimable vectors in the same models. We have the following mixed model

$$\mathbf{Y} = \sum_{i=1}^{w} \mathbf{X}_i \boldsymbol{\beta}_i,$$

with

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

Let ν be the space spanned by all the covariance matrices $\{\mathbf{V}_1, \ldots, \mathbf{V}_w\}$ and $\mathbf{V}_0$ is a maximal element, this is

$$R(\mathbf{V}_i) \subseteq R(\mathbf{V}_0),\ i = 1, \ldots, w$$

where $R(\mathbf{V})$ is the range of $\mathbf{V}$.

Definition 35. *Let $\boldsymbol{\chi}$ be a vector space and a subset $\boldsymbol{\mathcal{B}}$ of $\boldsymbol{\chi}$. Then $\boldsymbol{\mathcal{B}}$ is an essentially complete class if for $\forall \boldsymbol{a} \notin \boldsymbol{\mathcal{B}},\ \exists \boldsymbol{b} \in \boldsymbol{\mathcal{B}}$, such that*

$$\mathbb{E}_\theta[\boldsymbol{a}^T \boldsymbol{Y}] = \mathbb{E}_\theta[\boldsymbol{b}^T \boldsymbol{Y}]$$

and

$$Var_\theta(\boldsymbol{b}^T\boldsymbol{Y}) \leq Var_\theta(\boldsymbol{a}^T\boldsymbol{Y})$$

for $\forall\theta \in \Omega$.

Moreover, we have the definition:

Definition 36. *Let* $\boldsymbol{X}$ *be a linear operator from a inner product space* $\mathcal{L}[.|.]$ *into* $\boldsymbol{\chi}$ *such that* $R(\boldsymbol{X}) = \epsilon$, *being* ϵ *the space spanned by the mean vector*

Let, the matrix $\mathbf{W}$ be written of the following form

$$\mathbf{W} = \begin{cases} \mathbf{V}_0 + \mathbf{X}\mathbf{X}^T;\ R(\mathbf{X}) \nsubseteq R(\mathbf{V}_0) \\ \mathbf{V}_0;\ R(\mathbf{X}) \subseteq R(\mathbf{V}_0) \end{cases} \tag{3.19}$$

where $\mathbf{X}^T$ is the transposed matrix of $\mathbf{X}$.
Thus, we will prove the following result

Lemma 3. $R(\boldsymbol{W})$ *is an essentially complete class*

Proof: See [Zmyślony (1978)] □

The next lemma shows us the uniqueness of a BLUE

Lemma 4. *If* $(\boldsymbol{a}^T\boldsymbol{Y})$ *and* $(\boldsymbol{b}^T\boldsymbol{Y})$ *are BLUE of a given function* φ *and* $\boldsymbol{a}, \boldsymbol{b} \in R(\boldsymbol{W})$, *then* $\boldsymbol{a} = \boldsymbol{b}$.

Proof: See [Zmyślony (1978)] □

From now on, let us limit ourselves to an essentially complete class of estimators $(\mathbf{a}^T\mathbf{Y})$, with $\mathbf{a} \in R(\mathbf{W})$. In this class the set of all BLUE will characterized.
Now, we have the following result

Lemma 5. *Let* $\boldsymbol{a} \in R(\boldsymbol{W})$. *If* $(\boldsymbol{a}^T\boldsymbol{Y})$ *is the BLUE of* $E(\boldsymbol{a}^T\boldsymbol{Y})$, *then* $\boldsymbol{a} \in \boldsymbol{W}^+$.

Proof: See [Zmyślony (1978)] □

Let $\mathbf{P} = \mathbf{X}\mathbf{X}^+$ be the orthogonal projection matrix on $R(\mathbf{X})$, and put $\mathbf{M} = \mathbf{I} - \mathbf{P}$, where $\mathbf{I}$ is the identity matrix. We have

$$\mathbf{U} = \mathbf{X}^T\mathbf{W}^-\left(\sum_{i=1}^{k}\mathbf{V}_i\mathbf{M}\mathbf{V}_i\right)\mathbf{W}^-\mathbf{X}$$

and

$$\mathbf{Z} = \mathbf{I} - \mathbf{U}^{-}\mathbf{U}$$

where $\mathbf{X}^{-}$ is a general inverse matrix.

Theorem 17. *Let $\boldsymbol{a} \in R(\boldsymbol{W})$. The function $(\boldsymbol{a}^T\boldsymbol{Y})$ is a BLUE for $E[\boldsymbol{a}^T\boldsymbol{Y}]$ if and only if*

$$\exists \boldsymbol{z} \in N(\boldsymbol{U}) : \boldsymbol{W}^{+}\boldsymbol{X}\boldsymbol{z} = \boldsymbol{a}$$

Proof: See [Zmyślony (1978)]. □

The next theorem shows us the existence of optimal estimators

Theorem 18. *There exists a BLUE for $(\boldsymbol{a}^T\boldsymbol{\beta})$ if and only if*

$$\boldsymbol{a} \in R\left(\boldsymbol{X}^T\boldsymbol{W}^{-}\boldsymbol{X}\boldsymbol{Z}\right)$$

Proof. See [Zmyślony (1978)]. □

The expression for these estimators is given by the following theorem

Theorem 19. *If there exists a BLUE for $(\boldsymbol{a}^T\boldsymbol{\beta})$, then $(\boldsymbol{a}^T\tilde{\boldsymbol{\beta}})$, where $\tilde{\boldsymbol{\beta}}$ is a solution of the equation*

$$\tilde{\boldsymbol{\beta}} = \left(\boldsymbol{Z}^T\boldsymbol{X}\boldsymbol{W}^{-}\boldsymbol{X}\right)^{-}\boldsymbol{Z}^T\boldsymbol{X}^T\boldsymbol{W}^{-}\boldsymbol{y},$$

is a BLUE of $(\boldsymbol{a}^T\boldsymbol{\beta})$.

Proof. See [Zmyślony (1978)]. □

3.6 Orthogonal Models

In this section, we will study a very important class of models that is the class of orthogonal models. Orthogonal models, ORT, see [Vanleeuwen et al. (1998)] and [Vanleeuwen et al. (1999)], are mixed models with remarkable properties.
From the definition (33), we have,

$$V(\boldsymbol{\gamma}) = \sum_{j=1}^{m} \gamma_j \mathbf{Q}_j \tag{3.20}$$

is OBS case $\mathbf{Q}_1, \ldots \mathbf{Q}_m$ are orthogonal pairwise orthogonal projection matrices, OPOPM, and $\sum_{j=1}^{m} \mathbf{Q}_j = \mathbf{I}_n$. We define that the model has COBS if it has OBS and $\mathbf{T}\mathbf{Q}_j = \mathbf{Q}_j\mathbf{T}$ with

$$\mathbf{T} = \mathbf{X}_0\mathbf{X}_0^+$$

(see definition (34)).
We now have the following definitions:

Definition 37. ***(LSE)***
The LSE for

$$\boldsymbol{\Psi} = \boldsymbol{G}\boldsymbol{\beta}$$

is

$$\tilde{\boldsymbol{\Psi}} = \boldsymbol{G}\tilde{\boldsymbol{\beta}}$$

when $\tilde{\boldsymbol{\beta}}$ *is obtained minimizing*

$$||\,\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta}||^2$$

Definition 38. ***(UBLUE)***
When the least square estimative (LSE) for estimable vectors (see [Mexia et al. (2011)], page 14), are BLUE whatever the variance components $\{\gamma_1, \ldots, \gamma_m\}$*, we say, following [Vanleeuwen et al. (1998)], that they are Uniformly BLUE, UBLUE.*

A necessary and sufficient condition for its LSE to be UBLUE is, see [Zmyślony (1978)], that $\mathbf{T}$ commutes with the matrices in $V(\gamma)$. The orthogonal models have least square estimators, LSE that are UBLUE. The definition of ROBS follow:

Definition 39. ***(ROBS)***
When a model has a covariance structure like in expression (3.20) with $\boldsymbol{\gamma} \in \Gamma$ *where* Γ *contains non-empty open set, we say that the model has ROBS (Robust OBS).*

Let $\mathbf{Q} = \{\mathbf{Q}_1, \ldots, \mathbf{Q}_m\}$ be the principal basis of an CJA $\mathcal{A}(\mathbf{M})$, we have

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

If the matrices $\{\mathbf{M}_1, \ldots, \mathbf{M}_w\}$ commute, they generates an CJA $\mathcal{A}(\mathbf{M})$. If the algebra $\mathcal{A}$ contains the matrices $\{\mathbf{M}_1, \ldots, \mathbf{M}_w\}$ and $\mathbf{T}$, then the matrices $\{\mathbf{M}_1, \ldots, \mathbf{M}_w\}$ commute with $\mathbf{T}$. A model associated to this CJA has UBLUE.
Now, with $\mathbf{B} = [b_{ij}]$, the transition matrix, we have

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

with

$$\boldsymbol{\gamma}_j = \sum_{i=1}^{w} b_{ij}\boldsymbol{\theta}_i$$

this is

$$\boldsymbol{\gamma} = \mathbf{B}^T\boldsymbol{\theta}\ \in\ R(\mathbf{B}^T)_+$$

The γ are linear combinations of θ and otherwise, because $\mathbf{B}^T$ is invertible. ∇_+ represents the family of vectors of subspace ∇ with non-negative components. So,

$$\Gamma \subseteq R(\mathbf{B}^T)_+,$$

we say that the parameters are positive.
If the matrices $\mathbf{M}_1, \ldots, \mathbf{M}_w$ are linearly independent, the same happens with the row vectors of $\mathbf{B}$ being

$$car(\mathbf{B}) = w$$

and hence

$$dim(R(\mathbf{B}^T)) = w$$

In the case of orthogonal models, $\mathbf{M}$ is basis of a CJA, $\mathcal{A}(\mathbf{M})$. When $w = m$, $\mathbf{M}$ will be a perfect family. So, we have the following definition:

Definition 40. ***Perfect Family***

A family $\boldsymbol{M}$ *of symmetric matrices that commutes is perfect when it is a basis for the CJA that it generates. In this case* $\Gamma = \mathbb{R}^m_+$ *contains non-empty open. If* $w < m$*,* Γ *it does not contain non-empty open.*

If $\mathbf{B}$ is a transition matrix and $\mathbf{M}$ is a basis for $\mathcal{A}(\mathbf{M})$ it will be, see [Fonseca et al. (2007)], a perfect family and $\mathbf{B}$ will be invertible.
Let $\mathbf{M}_1, \ldots, \mathbf{M}_w$ be $b(\mathcal{A}(\mathbf{M}))$ and let $\mathbf{Q}_1, \ldots, \mathbf{Q}_w$ be the $pb(\mathcal{A}(\mathbf{M}))$.
$\mathbf{M}_1, \ldots, \mathbf{M}_w$ and $\mathbf{Q}_1, \ldots, \mathbf{Q}_w$ are a perfect family, because they are paired and have the same number of elements.

Let $\{\mathbf{0}\} \subset R(\mathbf{Q}^\circ_j) \subset R(\mathbf{T})$ be; $j = \{1, \ldots, m^\circ\}$, if

$$\mathbf{Q}_j = \mathbf{Q}^\circ_j$$

and

$$\mathbf{Q}_{j+m^\circ} = \mathbf{Q}^\circ_j(\mathbf{I}_n - \mathbf{T}) = \mathbf{Q}^\circ_j\mathbf{T}^c$$

then $\mathbf{Q}_j$ and $\mathbf{Q}_{j+m^\circ}$ are paired.

If $\mathbf{Y}$ has normal distribution and the model is orthogonal, the density belongs to the exponential family and the parametric space contains open non-empty sets. The model has sufficient and complete statistics from which we obtain the UMVUE. Now, we will consider the following propositions.

Proposition 12. *We have*

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

Proof:

1. we will prove that

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

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

$$\begin{aligned} V(\theta_1) = V(\theta_2) \Longleftrightarrow 0 = V(\theta_2) - V(\theta_1) = \\ = \sum_{i=1}^{w} \theta_i(2)\mathbf{M}_i - \sum_{i=1}^{w} \theta_i(1)\mathbf{M}_i = \\ = \sum_{i=1}^{w} (\theta_i(2) - \theta_i(1))\mathbf{M}_i \end{aligned} \tag{3.21}$$

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

$$\theta_1 = \theta_2,$$

□

the proof of this proposition is also in [Pereira et al. (2014)].

We still have the following result:

Proposition 13. *The matrices in $M = \{\boldsymbol{M}_1, \ldots, \boldsymbol{M}_w\}$ are linearly independent if and only if $\boldsymbol{B}^0$ is horizontally free.*

Proof: Assume that $\mathbf{M}_i = \sum_{j=1}^{w} b_{ij}\mathbf{Q}_j$, with $\mathbf{b}_1 \ldots, \mathbf{b}_w$ the row vectors of $\mathbf{B}$, in other words

$$\mathbf{B}^\circ = \begin{bmatrix} \mathbf{b}_1^T \\ \vdots \\ \mathbf{b}_w^T \end{bmatrix}$$

As $\mathbf{Q}_1, \ldots, \mathbf{Q}_m$ are mutually orthogonal, are linearly independent. Since

$$\sum_{i=1}^{u} c_i\mathbf{M}_i = \sum_{i=1}^{u} c_i \sum_{j=1}^{w} b_{ij}\mathbf{Q}_j = \sum_{j=1}^{w} \left(\sum_{i=1}^{u} c_i b_{ij} \right) \mathbf{Q}_j$$

then

$$\sum_{i=1}^{u} c_i\mathbf{M}_i = \mathbf{0}_{n\times n} \Longleftrightarrow \sum_{i=1}^{u} c_i b_{ij} = 0$$

So, $\sum_{i=1}^{u} c_i\mathbf{M}_i = \mathbf{0}$ implies $c_1 = \cdots = c_u = 0$. Also $\sum_{i=1}^{u} c_i\mathbf{b}_{ij} = 0$ implies $c_1 = \cdots = c_u = 0$, that is, implies the linear independence of the row vectors of $\mathbf{B}$, so the matrix $\mathbf{B}$ is horizontally free, it has linearly independent row vectors $\square$

We now have

$$\mathbf{M}_i = \sum_{j=1}^{m} b_{ij}\mathbf{Q}_j$$

so

$$\begin{aligned} V(\boldsymbol{\theta}) &= \sum_{i=1}^{w} c_i\mathbf{M}_i = \sum_{i=1}^{w} c_i \sum_{j=1}^{m} b_{ij}\mathbf{Q}_j = \\ &= \sum_{i=1}^{w}\sum_{j=1}^{m} c_i b_{ij}\mathbf{Q}_j = \sum_{j=1}^{m}\left(\sum_{i=1}^{w} c_i b_{ij}\right)\mathbf{Q}_j = \\ &= \sum_{j=1}^{m} \boldsymbol{\gamma}_j\mathbf{Q}_j \end{aligned} \tag{3.22}$$

with $\boldsymbol{\gamma}_j = \sum_{i=1}^{w} c_i b_{i,j}$; $j = 1, \ldots, m$, i.e., $\boldsymbol{\gamma} = \mathbf{B}^T\mathbf{c}$, with $\mathbf{c} \geq 0$, so that $\boldsymbol{\gamma} \in \Gamma = \mathbf{B}^T\mathbb{R}_+^w \subset R(\mathbf{B}^T)$. We point out that

$$dim\left(R(\mathbf{B}^T)\right) = car(\mathbf{B}^T) \leq w,$$

since $\mathbf{B}$ is a $m \times w$ matrix.

If $w < m$, the m components of γ have to satisfy linear solutions, therefore $\Gamma = \mathbf{B}^T\mathbb{R}_+^w$ not containing any non-empty open set.

So, we have the following propositions:

Proposition 14.

$$\boldsymbol{TM}_i = \boldsymbol{M}_i\boldsymbol{T} \Longleftrightarrow \textit{the LSE are UBLUE.}$$

Proof: From [Zmyślony (1978)], the least square estimators, LSE, (see [Mexia et al. (2011)], page 14) are uniformly best linear unbiased estimator, UBLUE, if and only if $\forall_{\boldsymbol{\theta}}$, $TV(\boldsymbol{\theta}) = V(\boldsymbol{\theta})T$, $\boldsymbol{\theta} \geq 0$. Let

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

and

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

we then have

$$\mathbf{M}_i = V(\boldsymbol{\theta} + \boldsymbol{\delta}_i) - V(\boldsymbol{\theta}) = \sum_{i=1}^{w} (\boldsymbol{\theta} + \boldsymbol{\delta}_i)\mathbf{M}_i - \sum_{i=1}^{w} \boldsymbol{\theta}\mathbf{M}_i = \boldsymbol{\delta}_i^T \mathbf{M}_i = \mathbf{M}_i.$$

So

$$\mathbf{T}\mathbf{M}_i = \mathbf{T}\left(V(\boldsymbol{\theta} + \boldsymbol{\delta}_i) - V(\boldsymbol{\theta})\right) = (V(\boldsymbol{\theta} + \boldsymbol{\delta}_i) - V(\boldsymbol{\theta}))\mathbf{T} = \mathbf{M}_i\mathbf{T},$$

because the LSE are UBLUE. □

Proposition 15. *We have:*

$$\boldsymbol{T}\boldsymbol{M}_i = \boldsymbol{M}_i\boldsymbol{T} \Longrightarrow \boldsymbol{M}_i = \boldsymbol{M}_i^{\circ} + \boldsymbol{M}_i^{c},$$

with $\boldsymbol{M}_i^{\circ} = \boldsymbol{T}\boldsymbol{M}_i\boldsymbol{T}$ *and* $\boldsymbol{M}_i^{c} = \boldsymbol{T}^c\boldsymbol{M}_i\boldsymbol{T}^c$

Proof: We know that $\mathbf{T}^c = \mathbf{I}_n - \mathbf{T} \Longleftrightarrow \mathbf{I}_n = \mathbf{T} + \mathbf{T}^c$ and $\mathbf{T}\mathbf{T}^c = \mathbf{T}(\mathbf{I}_n - \mathbf{T}) = \mathbf{T} - \mathbf{T} = 0$. So, we have

$$\begin{aligned} \mathbf{M}_i &= \mathbf{I}_n\mathbf{M}_i\mathbf{I}_n = (\mathbf{T} + \mathbf{T}^c)\mathbf{M}_i(\mathbf{T} + \mathbf{T}^c) = \\ &= \mathbf{T}\mathbf{M}_i\mathbf{T} + \mathbf{T}\mathbf{M}_i\mathbf{T}^c + \mathbf{T}^c\mathbf{M}_i\mathbf{T} + \mathbf{T}^c\mathbf{M}_i\mathbf{T}^c = \\ &= \mathbf{M}_i^{\circ} + \mathbf{M}_i\mathbf{T}\mathbf{T}^c + \mathbf{M}_i\mathbf{T}\mathbf{T}^c + \mathbf{M}_i^c = \\ &= \mathbf{M}_i^{\circ} + \mathbf{M}_i^c \end{aligned} \tag{3.23}$$

We conclude that $\mathbf{M}_i = \mathbf{M}_i^{\circ} + \mathbf{M}_i^c$ and the thesis is established. □

Proposition 16. *We have:*

$$\boldsymbol{M}_i\boldsymbol{M}_l = \boldsymbol{M}_l\boldsymbol{M}_i \Longleftrightarrow \boldsymbol{M}_i^{\circ}\boldsymbol{M}_l^{\circ} = \boldsymbol{M}_l^{\circ}\boldsymbol{M}_i^{\circ} \;\wedge\; \boldsymbol{M}_i^{c}\boldsymbol{M}_l^{c} = \boldsymbol{M}_l^{c}\boldsymbol{M}_i^{c}$$

Proof: Knowing that

$$\begin{cases} \mathbf{M}_i = \mathbf{M}_i^{\circ} + \mathbf{M}_i^c \\ \mathbf{M}_l = \mathbf{M}_l^{\circ} + \mathbf{M}_l^c \end{cases} \tag{3.24}$$

and that

$$\begin{cases} \mathbf{M}_i^\circ = \mathbf{T}\mathbf{M}_i\mathbf{T} \\ \mathbf{M}_i^c = \mathbf{T}^c\mathbf{M}_i\mathbf{T}^c \end{cases} \tag{3.25}$$

we get

$$\begin{aligned} \mathbf{M}_i\mathbf{M}_l &= (\mathbf{M}_i^\circ + \mathbf{M}_i^c)(\mathbf{M}_l^\circ + \mathbf{M}_l^c) = \\ &= \mathbf{M}_i^\circ\mathbf{M}_l^\circ + \mathbf{M}_i^\circ\mathbf{M}_l^c + \mathbf{M}_i^c\mathbf{M}_l^\circ + \mathbf{M}_i^c\mathbf{M}_l^c = \\ &= \mathbf{M}_i^\circ\mathbf{M}_l^\circ + \mathbf{T}\mathbf{M}_i\mathbf{T}\mathbf{T}^c\mathbf{M}_l\mathbf{T}^c + \mathbf{T}^c\mathbf{M}_i\mathbf{T}^c\mathbf{T}\mathbf{M}_l\mathbf{T} + \mathbf{M}_i^c\mathbf{M}_l^c = \\ &= \mathbf{M}_i^\circ\mathbf{M}_l^\circ + \mathbf{M}_i^c\mathbf{M}_l^c \end{aligned} \tag{3.26}$$

We now will get $\mathbf{M}_l\mathbf{M}_i$

$$\begin{aligned} \mathbf{M}_l\mathbf{M}_i &= (\mathbf{M}_l^\circ + \mathbf{M}_l^c)(\mathbf{M}_i^\circ + \mathbf{M}_i^c) = \\ &= \mathbf{M}_l^\circ\mathbf{M}_i^\circ + \mathbf{M}_l^\circ\mathbf{M}_i^c + \mathbf{M}_l^c\mathbf{M}_i^\circ + \mathbf{M}_l^c\mathbf{M}_i^c = \\ &= \mathbf{M}_l^\circ\mathbf{M}_i^\circ + \mathbf{M}_l^c\mathbf{M}_i^c \end{aligned} \tag{3.27}$$

Therefore $\mathbf{M}_i\mathbf{M}_l = \mathbf{M}_l\mathbf{M}_i$, $\mathbf{M}_i$ e $\mathbf{M}_l$ commute.

We will now demonstrate that $\mathbf{M}_i^\circ\mathbf{M}_l^\circ + \mathbf{M}_i^c\mathbf{M}_l^c = \mathbf{M}_l^\circ\mathbf{M}_i^\circ + \mathbf{M}_l^c\mathbf{M}_i^c$. Knowing that:

$$\begin{cases} \mathbf{T}\mathbf{M}_i^\circ = \mathbf{M}_i^\circ \ \wedge \ \mathbf{T}\mathbf{M}_i^c = 0 \\ \mathbf{T}^c\mathbf{M}_i^c = \mathbf{M}_i^c \ \wedge \ \mathbf{T}^c\mathbf{M}_i^\circ = 0 \end{cases} \tag{3.28}$$

we have

$$\mathbf{T}\mathbf{M}_i\mathbf{M}_l = \mathbf{T}\mathbf{M}_i^\circ\mathbf{M}_l^\circ + \mathbf{T}\mathbf{M}_i^c\mathbf{M}_l^c = \mathbf{T}\mathbf{M}_i^\circ\mathbf{M}_l^\circ = \mathbf{M}_i^\circ\mathbf{M}_l^\circ$$

and

$$\mathbf{T}\mathbf{M}_l\mathbf{M}_i = \mathbf{T}\mathbf{M}_l^\circ\mathbf{M}_i^\circ + \mathbf{T}\mathbf{M}_l^c\mathbf{M}_i^c = \mathbf{T}\mathbf{M}_l^\circ\mathbf{M}_i^\circ = \mathbf{M}_l^\circ\mathbf{M}_i^\circ$$

so

$$\mathbf{T}\mathbf{M}_i\mathbf{M}_l = \mathbf{T}\mathbf{M}_l\mathbf{M}_i$$

and the matrices $\mathbf{M}_1^\circ, \ldots, \mathbf{M}_m^\circ$ commute.
We now have

$$\mathbf{T}^c\mathbf{M}_i\mathbf{M}_l = \mathbf{T}^c\mathbf{M}_i^\circ\mathbf{M}_l^\circ + \mathbf{T}^c\mathbf{M}_i^c\mathbf{M}_l^c = \mathbf{T}^c\mathbf{M}_i^c\mathbf{M}_l^c = \mathbf{M}_i^c\mathbf{M}_l^c$$

and

$$\mathbf{T}^c\mathbf{M}_l\mathbf{M}_i = \mathbf{T}^c\mathbf{M}_l^\circ\mathbf{M}_i^\circ + \mathbf{T}^c\mathbf{M}_l^c\mathbf{M}_i^c = \mathbf{T}^c\mathbf{M}_l^c\mathbf{M}_i^c = \mathbf{M}_l^c\mathbf{M}_i^c$$

so

$$\mathbf{T}^c\mathbf{M}_i\mathbf{M}_l = \mathbf{T}^c\mathbf{M}_l\mathbf{M}_i$$

and the matrices $\mathbf{M}_1^c, \ldots, \mathbf{M}_m^c$ commutes. So, the thesis is established. □

Proposition 17. *Models with COBS are models with UBLUE whose matrices $\boldsymbol{M}_1, \ldots, \boldsymbol{M}_w$ commute.*

Proof: Let $V(\boldsymbol{\theta}) = \sum_{i=1}^{w} \boldsymbol{\theta}_i\mathbf{M}_i$. The model has COBS if $\mathbf{T}, \mathbf{Q}_1, \ldots, \mathbf{Q}_m$ commute and it has UBLUE if the matrices $\mathbf{T}$ and $V(\boldsymbol{\theta})$ commute, this is,

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

If M= $\{\mathbf{M}_1, ..., \mathbf{M}_w\}$ commute, they generate a Commutative Jordan Algebra, CJA, with principal basis, pb $= \{\mathbf{Q}_1, \ldots, \mathbf{Q}_m\}$. Then, if the matrix $\mathbf{T}$ and the matrices in the CJA commute, the matrices $\mathbf{T}$ and $\mathbf{M}$ commute.
If

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

then

$$\mathbf{T}\left(\sum_{i=1}^{w} \boldsymbol{\theta}_i\mathbf{M}_i\right) = \left(\sum_{i=1}^{w} \boldsymbol{\theta}_i\mathbf{M}_i\right)\mathbf{T}.$$

From proposition 14, we get

$$\mathbf{T}\mathbf{M}_i = \mathbf{M}_i\mathbf{T}$$

So, if $\mathbf{T}, \mathbf{M}_1, \ldots, \mathbf{M}_w$ commute, then the model has COBS. And so, the thesis is established. $\square$

Now, we again have the following mixed model, (see expression 3.15 in section 3.5)

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

will have mean vector

$$\boldsymbol{\mu} = \mathbf{X}_0 \boldsymbol{\beta}_0$$

which spans the range space $\Omega = R(\mathbf{X}_0)$ of $\mathbf{X}_0$. The orthogonal projection matrices on Ω will be

$$\mathbf{T} = \mathbf{X}_0 \mathbf{X}_0^+$$

The family of variance covariance matrices for this model will be

$$\mathbf{V} = \sum_{i=1}^{w} \boldsymbol{\theta}_i \mathbf{M}_i; \ \boldsymbol{\theta} \geq 0$$

with $\mathbf{M}_i = \mathbf{X}_i \mathbf{X}_i^T, \ i = \{1, \ldots, w\}$. For the matrices of $\mathbf{V}$ to be invertible we have only to take $\boldsymbol{\theta}_w > 0$ and $\mathbf{M}_w = \mathbf{X}_w = \mathbf{I}_n$.
We now establish

Lemma 6. *The model has UBLUE if and only if $\boldsymbol{T}$ commutes with $\boldsymbol{M}_1, \ldots, \boldsymbol{M}_w$.*

Proof: If $\mathbf{T}$ commutes with $\mathbf{M}_1, \ldots, \mathbf{M}_w$ it commutes with the matrices $V(\boldsymbol{\theta}) = \sum_{i=1}^{w} \boldsymbol{\theta}_i \mathbf{M}_i$ and the model has UBLUE, see [Pereira et al. (2014)]. Inversely if the model has UBLUE, with $\boldsymbol{\delta_i}$ the vector with null components but the $i-th$ which is 1, $\mathbf{T}$ will commute with $V(\boldsymbol{\theta})$, $V(\boldsymbol{\theta}+\boldsymbol{\delta}_i)$ and $\mathbf{M}_i = V(\boldsymbol{\theta}+\boldsymbol{\delta_i}) - V(\boldsymbol{\delta}_i)$, $i = \{1, \ldots, w\}$, which completes the proof. $\square$

If the matrices $\mathbf{M}_1, \ldots, \mathbf{M}_w$ commute they generate $\mathcal{A}^\circ = \mathcal{A}(\mathbf{M})$ with principal basis $\mathbf{Q}^\circ = \{\mathbf{Q}_1^\circ, \ldots, \mathbf{Q}_{m^\circ}^\circ\}$. Let $\mathbf{B}^\circ = [b_{ij}^\circ]$ be the transition matrix, then

$$V(\boldsymbol{\theta}) = \sum_{i=1}^{w} \boldsymbol{\theta}_i \mathbf{M}_i = \sum_{i=1}^{w} \boldsymbol{\theta}_i \left(\sum_{j=1}^{m^\circ} b_{ij}^\circ \mathbf{Q}_j^\circ \right) = \sum_{j=1}^{m^\circ} \boldsymbol{\gamma}_j^\circ \mathbf{Q}_j^\circ$$

with

$$\boldsymbol{\gamma}_j^\circ = \sum_{i=1}^{w} b_{ij}^\circ \boldsymbol{\theta}_i, \ j = \{1, \ldots, m^\circ\}$$

so the matrices $\mathbf{V}$ will be linear combinations of known orthogonal pairwise orthogonal projections matrices, OPOPM, $\mathbf{Q}_1^\circ, \ldots, \mathbf{Q}_{m^\circ}^\circ$, . If the $V(\boldsymbol{\theta})$ are invertible, $\mathcal{A}$ must be complete, so

$$\sum_{j=1}^{m^\circ} \mathbf{Q}_j^\circ = \mathbf{I}_n.$$

Now we get

Lemma 7. *When the matrices of $\boldsymbol{M}_1, \ldots, \boldsymbol{M}_w$ commute the model will have BLUE if and only if $\boldsymbol{T}$ commutes with the matrices of $\mathcal{A}^\circ$.*

Proof: If $\mathbf{T}$ commutes with the matrices of $\mathcal{A}^\circ$, then it commutes with $\mathbf{M}_1, \ldots, \mathbf{M}_w$ and, from (6), the model will have UBLUE. If the model will have UBLUE, also from (6), $\mathbf{T}$ will commute with the matrices $\mathbf{M}_1, \ldots, \mathbf{M}_w$. assuming now the matrices $\mathbf{M}_1, \ldots, \mathbf{M}_w$ commute with all the matrices of

$$\mathbf{M}' = \{\mathbf{T}, \mathbf{M}_1, \ldots, \mathbf{M}_w\},$$

then generating a algebra $\mathcal{A}' = \mathcal{A}(\mathbf{M'})$, which contains $\mathcal{A}^\circ$, so $\mathbf{T}$ will commute with the matrices of $\mathcal{A}'$ namely with these of $\mathcal{A}^\circ$ and the proof is complete. $\square$

From now on we assume that $\mathbf{M}_1, \ldots, \mathbf{M}_w$ commute.
For the model to be well defined, we have

$$V(\boldsymbol{\theta}_1) = V(\boldsymbol{\theta}_2)$$

must imply

$$\boldsymbol{\theta}_1 = \boldsymbol{\theta}_2,$$

so the matrices $\mathbf{M}_1, \ldots, \mathbf{M}_w$ must be linearly independent which, since

$$\mathbf{M}_i = \sum_{j=1}^{m^\circ} b_{ij}^\circ \mathbf{Q}_j^\circ,\ i = 1, \ldots, w$$

is equivalent to the row vectors of $\mathbf{B}^\circ$ being linearly independent and so

$$rank(\mathbf{B}^{\circ T}) = rank(\mathbf{B}^\circ) = w$$

since

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

thus there will be no linear restrictions on the m° components of $\boldsymbol{\gamma}^\circ$ when and only when $\mathbf{B}^{\circ^T}$ is invertible. Then $\mathbf{B}^\circ$ will be invertible and $\mathbf{M}$ will be perfect being a basis for $\mathcal{A}^\circ$. We now have the following lemma:

Lemma 8. *When $\boldsymbol{M}_1, \ldots, \boldsymbol{M}_w$ commute and $\boldsymbol{B}^{\circ^T}$ is invertible then there are no linear restrictions on the components of $\boldsymbol{\gamma}^\circ$ and the corresponding parameter space contains non void open sets, if and only if $\boldsymbol{M}$ is perfect.*

Proof: To complete the proof we have only to point out that

$$\{\mathbf{v};\ \mathbf{v} = \mathbf{B}^{\circ^T}\mathbf{u},\ \mathbf{u} > \mathbf{0}\} \subseteq \Gamma^\circ$$

and that when $\mathbf{M}$ is perfect $\mathbf{B}^{\circ^T}$ is invertible. □

Now orthogonal models, ORT, as defend in [Vanleeuwen et al. (1998)] are UBLUE models with families of variance covariance matrices

$$\mathbf{V} = \{V(\boldsymbol{\gamma}^\circ) = \sum_{j=1}^{m} \boldsymbol{\gamma}_j^\circ \mathbf{Q}_j;\ \boldsymbol{\gamma}^\circ \in \Gamma\}.$$

When $\mathbf{Q}$ is a family of OPOPM and Γ contains non void open sets, then we now have the proposition:

Proposition 18. *A mixed model*

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

*is orthogonal if and only if the matrices **T**, $\boldsymbol{M}_1, \ldots, \boldsymbol{M}_w$ commute and the family $\boldsymbol{M} = \{\boldsymbol{M}_1, \ldots, \boldsymbol{M}_w\}$ is perfect*

The properties of ORT models was studied in [Vanleeuwen et al. (1998)] and [Vanleeuwen et al. (1999)]. Our goal in this section was to study their algebraic structure as expressed in proposition above to obtain the basis for binary operations on these models.
In the next sections, we will study the crossing and nesting ORT models. This topics was also studied in [Santos C. (2012)].

3.7 Crossing Orthogonal Models

We will represent by $\otimes$ the Kronecker product, and take

$$\{U_1, \ldots, U_d\} \otimes \{G_1, \ldots, G_t\} = \{U_l G_j,\ i = 1, \ldots, d\ e\ j = 1, \ldots, t\}$$

given two CJA $\mathcal{A}(l),\ l = 1, 2$ with principal basis $\mathbf{Q}(l) = \{\mathbf{Q}_1(l), \ldots, \mathbf{Q}_{m(l)}(l)\},\ l = 1, 2$, and two families of symmetric matrices

$$\mathbf{M}(l) = \{\mathbf{M}_1(l), \ldots, \mathbf{M}_{w(l)}(l)\},\ l = 1, 2,$$

we have

$$\mathbf{M}(1) \otimes \mathbf{M}(2) \subset \mathcal{A}(1) \otimes \mathcal{A}(2),$$

that

$$\mathbf{Q} = pb(\mathcal{A}(1) \otimes \mathcal{A}(2)) = pb(\mathcal{A}(1)) \otimes pb(\mathcal{A}(2)) = \mathbf{Q}(1) \otimes \mathbf{Q}(2)$$

and that the transition matrix $\mathbf{M}(1) \otimes \mathbf{M}(2) \setminus \mathbf{Q}(1) \otimes \mathbf{Q}(2)$ is

$$\mathbf{B} = \mathbf{B}(1) \otimes \mathbf{B}(2)$$

with $\mathbf{B}(l)$ the transition matrix $\mathbf{M}(l) \setminus \mathbf{Q}(l),\ l = 1, 2$, the kronecker product $\otimes$ of CJA will be useful in studying model crossing.
We will now study two cases

3.7.1 First case

As stated above we will have the Commutative Jordan Algebra, CJA

$$\mathcal{A}^{\circ}(1) \otimes \mathcal{A}^{\circ}(2)$$

and

$$\mathcal{A}(1) \otimes \mathcal{A}(2)$$

in studying the crossing of ORT models. We will consider the first case where the two models have fixed effects. When we cross two ORT models with mean vector

$$\boldsymbol{\mu}(l) = \mathbf{X}_0(l)\boldsymbol{\beta}_0(l),\ l = 1, 2,$$

we obtain a model with mean vector

$$\boldsymbol{\mu} = \mathbf{X}_0\boldsymbol{\beta}_0 \tag{3.29}$$

where $\mathbf{X}_0 = \mathbf{X}_0(1) \otimes \mathbf{X}_0(2)$. Since

$$\mathbf{X}_0^+ = \mathbf{X}_0^+(1) \otimes \mathbf{X}_0^+(2)$$

the orthogonal projection matrices in the space spanned by $\boldsymbol{\mu}$ will be

$$\begin{aligned}\mathbf{T} &= \mathbf{X}_\circ \mathbf{X}_\circ^+ = \\ &= (\mathbf{X}_0(1) \otimes \mathbf{X}_0(2))(\mathbf{X}_0^+(1) \otimes \mathbf{X}_0^+(2)) = \\ &= (\mathbf{X}_0(1)\mathbf{X}_0^+(1)) \otimes (\mathbf{X}_0(2)\mathbf{X}_0^+(2)) = \\ &= \mathbf{T}(1) \otimes \mathbf{T}(2)\end{aligned} \tag{3.30}$$

where $\mathbf{T}(l) = \mathbf{X}_0(l)\mathbf{X}_o^+(l),\ l = 1,2$ is the orthogonal projection matrix on the space spanned by

$$\boldsymbol{\mu}(l),\ l = 1,2.$$

Now $\mathbf{T}(l) \in \mathcal{A}(l)$, so we can consider the matrices $\{\mathbf{Q}_1(l), \ldots, \mathbf{Q}_{m(l)}(l)\}$ in

$$Q(l) = bp(\mathcal{A}(l)),\ l = 1,2,$$

so, we have

$$\mathbf{T}(l) = \sum_{j=1}^{z(l)} \mathbf{Q}_j(l),\ l = 1,2$$

as well as

$$\mathbf{T} = \sum_{j'=1}^{z(1)} \sum_{j''=1}^{z(2)} \mathbf{Q}_{j'}(1) \otimes \mathbf{Q}_{j''}(2) \in \mathcal{A}(1) \otimes \mathcal{A}(2)$$

Thus $\mathbf{T}$ will commute with the matrices of $\mathcal{A}^\circ(1) \otimes \mathcal{A}^\circ(2)$ and so the new model will be UBLUE.

As pointed above the transition matrix

$$\mathbf{M}(1) \otimes \mathbf{M}(2) \setminus \mathbf{Q}^\circ(1) \otimes \mathbf{Q}^\circ(2)$$

is

$$\mathbf{B}^\circ = \mathbf{B}^\circ(1) \otimes \mathbf{B}^\circ(2)$$

and since $\mathbf{B}^\circ(1)$ and $\mathbf{B}^\circ(2)$ are invertible, $\mathbf{B}^\circ$ is invertible and $\mathbf{M}(1) \otimes \mathbf{M}(2)$ is perfect. Thus, we get the following proposition:

Proposition 19. *Crossing ORT model gives ORT models*

3.7.2 Second case

When the matrices $\{\mathbf{M}_1, \ldots, \mathbf{M}_w\}$ commute, they generate a CJA $\mathcal{A}^\circ$, with principal basis $\{\mathbf{Q}_1^\circ, \ldots, \mathbf{Q}_{m^\circ}^\circ\}$.

If $\mathbf{T}$ and $\{\mathbf{M}_1, \ldots, \mathbf{M}_w\}$ commute, they generate a CJA $\mathcal{A}$. The matrices $\{\mathbf{Q}_1^\circ, \ldots, \mathbf{Q}_{m^\circ}^\circ\}$. and $\mathbf{T}$ will belong the $\mathcal{A}$, with principal basis constituted by matrices $\{\mathbf{Q}_1, \ldots, \mathbf{Q}_m\}$.

Finally, we have the matrices $\{\mathbf{M}_0,\ \mathbf{M}_1, \ldots, \mathbf{M}_w\}$, with $\mathbf{M}_0 = \mathbf{X}_0\mathbf{X}_0^T$ that if commute, they generate a CJA $\overset{\circ}{\mathcal{A}}$ constituting the matrices $\{\overset{\circ}{\mathbf{Q}}_1, \ldots, \overset{\circ}{\mathbf{Q}}_{\overset{\circ}{m}}\}$ a principal basis of $\overset{\circ}{\mathcal{A}}$.

If one of the families

$$\{\mathbf{M}_1, \ldots, \mathbf{M}_w\},\ \{\mathbf{T}, \mathbf{M}_1, \ldots, \mathbf{M}_w\}\ and\ \{\mathbf{M}_0,\ \mathbf{M}_1, \ldots, \mathbf{M}_w\}$$

is perfect, the corresponding transition matrix

$$\mathbf{B}^\circ = [b_{ij}^\circ],\ \mathbf{B} = [b_{ij}]\ or\ \overset{\circ}{\mathbf{B}} = [\overset{\circ}{b}_{ij}]$$

will be invertible.

We have the following mixed models:

$$\mathbf{Y}(1) = \sum_{i=0}^{w(1)} \mathbf{X}_i(1)\boldsymbol{\beta}_i(1)$$

and

$$\mathbf{Y}(2) = \sum_{j=0}^{w(2)} \mathbf{X}_j(2)\boldsymbol{\beta}_j(2).$$

Crossing the models, using the definition in [Fonseca et al. (2008)], we obtain the following model:

$$\begin{aligned}\mathbf{Y} &= \sum_{i=0}^{w(1)} \sum_{j=0}^{w(2)} \left(\mathbf{X}_i(1) \otimes \mathbf{X}_j(2)\right) \boldsymbol{\beta}_{ij} = \\ &= \sum_{l=0}^{w} \mathbf{X}_l \boldsymbol{\beta}_l\end{aligned} \tag{3.31}$$

with

$$\begin{cases} w = (w(1)+1)(w(2)+1)-1 \\ l = iw(2)+j \end{cases}$$

and $\mathbf{X}_l = \mathbf{X}_i(1) \otimes \mathbf{X}_j(2)$, if $\mathbf{M}_l = \mathbf{X}_l \mathbf{X}_l^T$,

then

$$\begin{aligned} \mathbf{M}_l &= (\mathbf{X}_i(1) \otimes \mathbf{X}_j(2))\,(\mathbf{X}_i(1) \otimes \mathbf{X}_j(2))^T = \\ &= (\mathbf{X}_i(1) \otimes \mathbf{X}_j(2))\left(\mathbf{X}_i^T(1) \otimes \mathbf{X}_j^T(2)\right) = \\ &= \left(\mathbf{X}_i(1)\mathbf{X}_i^T(1)\right) \otimes \left(\mathbf{X}_j(2)\mathbf{X}_j^T(2)\right) = \\ &= \mathbf{M}_i(1) \otimes \mathbf{M}_j(2) \end{aligned} \tag{3.32}$$

Knowing that

$$\mathbf{M}_l = \sum_{k=1}^{\mathring{m}} \mathring{b}_{lk} \mathring{\mathbf{Q}}_k,\ \ l = \{1, \dots, w\}$$

from expression (3.32), we have

$$\begin{aligned} \mathbf{M}_i(1) \otimes \mathbf{M}_j(2) &= \left(\sum_{i=1}^{\mathring{m}(1)} \mathring{b}_{li}(1)\mathring{\mathbf{Q}}_i(1)\right) \otimes \left(\sum_{j=1}^{\mathring{m}(2)} \mathring{b}_{lj}(2)\mathring{\mathbf{Q}}_j(2)\right) = \\ &= \sum_{i=1}^{\mathring{m}(1)} \sum_{j=1}^{\mathring{m}(2)} \left(\mathring{b}_{li}(1)\mathring{\mathbf{Q}}_i(1)\right) \otimes \left(\mathring{b}_{lj}(2)\mathring{\mathbf{Q}}_j(2)\right) = \\ &= \sum_{k=1}^{\mathring{m}} \mathring{b}_{lk} \mathring{\mathbf{Q}}_k,\ \ l = \{1, \dots, w\} \end{aligned} \tag{3.33}$$

The transition matrix for the crossed model is written as follows

$$\mathring{\mathbf{B}} = \mathring{\mathbf{B}}(1) \otimes \mathring{\mathbf{B}}(2)$$

having

$$car(\mathring{\mathbf{B}}) = car(\mathring{\mathbf{B}}(1))car(\mathring{\mathbf{B}}(2))$$

with

$$\begin{cases} nl = nl(1)nl(2) \\ nc = nc(1)nc(2) \\ n = n(1)n(2) \end{cases}$$

being n, n(1) and n(2) the orders of the matrices $\overset{\circ}{\mathbf{B}}$, $\overset{\circ}{\mathbf{B}}(1)$ *and* $\overset{\circ}{\mathbf{B}}(2)$ respectively. So, we can say that if the matrices $\overset{\circ}{\mathbf{B}}(1)$ *and* $\overset{\circ}{\mathbf{B}}(2)$ are invertible, $\overset{\circ}{\mathbf{B}}$ is invertible and $\mathbf{M}(1) \otimes \mathbf{M}(2)$ is perfect when $\mathbf{M}(1)$ and $\mathbf{M}(2)$ are perfect. So,

Proposition 20. *The model (*3.31*) is an orthogonal model if the initial models are orthogonal.*

Proof:

1. $\mathbf{M}_1, \ldots, \mathbf{M}_w$ forms the basis of the algebra $\mathcal{A}(\mathbf{M})$;
2. $\mathbf{Q}_1, \ldots, \mathbf{Q}_w$ forms the principal basis the same algebra $\mathcal{A}(\mathbf{M})$;
3. $\mathbf{T}\mathbf{Q}_i = \mathbf{Q}_i\mathbf{T}$; $\mathbf{Q}_i$ and $\mathbf{T}$ commute.

We have

$$\mathbf{Q}_{ii'} = \mathbf{Q}_i(1) \otimes \mathbf{Q}_{i'}(2)$$

and

$$\mathbf{T} = \mathbf{T}(1) \otimes \mathbf{T}(2)$$

with

$$\mathbf{Q}(1) \otimes \mathbf{Q}(2)$$

the principal basis of

$$\mathcal{A}(1) \otimes \mathcal{A}(2)$$

So,

$$\begin{aligned} \mathbf{T}\mathbf{Q}_{ii'} &= \mathbf{T}(1)\mathbf{Q}_i(1) \otimes \mathbf{T}(2)\mathbf{Q}_{i'}(2) = \\ &= \mathbf{Q}_1(1)\mathbf{T}(1) \otimes \mathbf{Q}_{i'}(2)\mathbf{T}(2) = \\ &= \mathbf{Q}_{ii'}\mathbf{T} \end{aligned} \tag{3.34}$$

So, the models (3.29) and (3.31) are orthogonal. $\square$

Thus, we get again the following proposition:

Proposition 21. *Crossing ORT model gives ORT models*

3.8 Nesting Orthogonal Models

In considering model nesting, see [Fonseca et al. (2006)], we will use the restricted kronecker product $*$ of CJA. So, with $\mathcal{A}(1)$ complete and regular and if the matrices of $\mathcal{A}(l)$ are $n(l) \otimes n(l), l = 1, 2$, we will have principal basis

$$pb(\mathcal{A}) = pb(\mathcal{A}(1) * \mathcal{A}(2)) = \{\mathbf{Q}_1(1) \otimes \frac{1}{n(2)}\mathbf{J}_{n(2)}, \dots, \mathbf{Q}_{m(1)}(1) \otimes \frac{1}{n(2)}\mathbf{J}_{n(2)}\} \cup$$
$$\cup \{\mathbf{I}_{n(1)} \otimes \mathbf{Q}_2(2), \dots, \mathbf{I}_{n(1)} \otimes \mathbf{Q}_{m(2)}(2)\} \tag{3.35}$$

with $\mathbf{J}_n = \mathbf{1}_n \mathbf{1}_n^T$.

3.8.1 First case

The first model will be mixed and given by

$$\mathbf{Y}(1) = \sum_{i=0}^{w(1)} \mathbf{X}_i(1)\boldsymbol{\beta}_i(1) \tag{3.36}$$

$\boldsymbol{\beta}_0(1)$ is fixed and $\boldsymbol{\beta}_1(1), \dots, \boldsymbol{\beta}_{w(1)}(1)$ are independent random vectors with $\boldsymbol{\beta}_i(1) \sim N(0, \sigma_i^2 \mathbf{I}_{c_i})$; $i = 1, \dots, w(1)$ and $\mathbf{X}_0, \dots, \mathbf{X}_{w(1)}$ known. The second model has random effects and is given by

$$\mathbf{Y}(2) = \sum_{i=0}^{w(2)} \mathbf{X}_i(2)\boldsymbol{\beta}_i(2) \tag{3.37}$$

where $\mathbf{X}_0(2) = \mathbf{1}_{n(2)}$. Through nesting, we get the following model

$$\mathbf{Y} = \sum_{i=0}^{w(1)} (\mathbf{X}_i(1) \otimes \mathbf{1}_{n(2)})\boldsymbol{\beta}_i + \sum_{i=w(1)+1}^{w} (\mathbf{I}_{n(1)} \otimes \mathbf{X}_{i-w(1)}(2))\boldsymbol{\beta}_i, \tag{3.38}$$

with w = w(1) + w(2).

Let $\mathbf{M}_i = (\mathbf{X}_i(1) \otimes \mathbf{1}_{n(2)})(\mathbf{X}_i(1) \otimes \mathbf{1}_{n(2)})^T$ be, with $i = 0, \ldots, w(1)$. Making calculations, we have

$$\begin{aligned}
\mathbf{M}_i &= (\mathbf{X}_i(1) \otimes \mathbf{1}_{n(2)})(\mathbf{X}_i(1) \otimes \mathbf{1}_{n(2)})^T = \\
&= (\mathbf{X}_i(1) \otimes \mathbf{1}_{n(2)})(\mathbf{X}_i(1)^T \otimes \mathbf{1}_{n(2)}^T) = \\
&= (\mathbf{X}_i(1)\mathbf{X}_i(1)^T) \otimes \mathbf{J}_{n(2)} = \\
&= \mathbf{M}_i(1) \otimes \mathbf{J}_{n(2)}
\end{aligned} \tag{3.39}$$

From expression (3.38) where $i = w(1)+1, \ldots, w$, we have

$$\begin{aligned}
\mathbf{M}_i &= (\mathbf{I}_{n(1)} \otimes \mathbf{X}_{i-w(1)}(2))(\mathbf{I}_{n(1)} \otimes \mathbf{X}_{i-w(1)}(2))^T = \\
&= (\mathbf{I}_{n(1)} \otimes \mathbf{X}_{i-w(1)}(2))(\mathbf{I}_{n(1)}^T \otimes \mathbf{X}_{i-w(1)}^T(2)) = \\
&= \mathbf{I}_{n(1)} \otimes \mathbf{X}_{i-w(1)}(2)\mathbf{X}_{i-w(1)}^T(2) = \\
&= \mathbf{I}_{n(1)} \otimes \mathbf{M}_{i-w(1)}(2)
\end{aligned} \tag{3.40}$$

From (3.39), we have

$$\begin{aligned}
\mathbf{M}_i = \mathbf{M}_i(1) \otimes \mathbf{J}_{n(2)} &= \left(\sum_{j=1}^{m(1)} b_{ij}(1)\mathbf{Q}_j(1) \right) \otimes \mathbf{J}_{n(2)} = \\
&= n(2) \sum_{j=1}^{m(1)} b_{ij}(1)\mathbf{Q}_j(1) \otimes \frac{1}{n(2)}\mathbf{J}_{n(2)} = \\
&= \sum_{j=1}^{m(1)} b_{ij}\mathbf{Q}_j
\end{aligned} \tag{3.41}$$

being $b_{ij} = n(2)b_{ij}(1)$ and $\sum_{j=1}^{m(1)} \mathbf{Q}_j = \mathbf{I}_{n(1)} \otimes \frac{1}{n(2)}\mathbf{J}_{n(2)}$ with $i = \{0, \ldots, w(1)\}$ and $j = \{1, \ldots, m(1)\}$.

From (3.40), we have

$$\begin{aligned}
\mathbf{M}_i &= \mathbf{I}_{n(1)} \otimes \mathbf{M}_{i-w(1)}(2) = \mathbf{I}_{n(1)} \otimes \left(\sum_{j=1}^{m(2)} b_{i-w(1),j}(2)\mathbf{Q}_j(2) \right) = \\
&= b_{i-w(1),1}(2)(\mathbf{I}_{n(1)} \otimes \frac{1}{n(2)}\mathbf{J}_{n(2)}) + \sum_{j=m(1)+1}^{m} b_{i-w(1)}(2)(\mathbf{I}_{n(1)} \otimes \mathbf{Q}_j(2)) = \\
&= b_{i-w(1),1}(2) \sum_{j=1}^{m(1)} \mathbf{Q}_j + \sum_{j=m(1)+1}^{m} b_{i-w(1),j}(2)\mathbf{Q}_j = \sum_{j=1}^{m} b_{i,j}\mathbf{Q}_j
\end{aligned} \tag{3.42}$$

with $i = w(1)+1,\ldots,w$ and w = w(1) + w(2).

Thus, the expressions 3.41 and 3.42 the transition matrix $\mathbf{B}$ is written as follows:

$$\mathbf{B} = \begin{bmatrix} n(2)\mathbf{B}(1) & 0 \\ \mathbf{b}_1(2)\mathbf{1}^T & \mathbf{B}^\circ(2) \end{bmatrix} \tag{3.43}$$

where

$$\mathbf{B}(2) = \begin{bmatrix} b_1(2) & \mathbf{B}^0(2) \end{bmatrix}$$

If the matrices $\{\mathbf{M}_1,\ldots,\mathbf{M}_w\}$ form a basis of algebra $\mathcal{A}(\mathbf{M})$ and $\{\mathbf{Q}_1,\ldots,\mathbf{Q}_w\}$ form the principal basis of algebra $\mathcal{A}(\mathbf{M})$, M $= \{\mathbf{M}_1,\ldots,\mathbf{M}_w\}$ will be a perfect family, and the transition matrix $\mathbf{B}$ will be invertible, this because the transition matrices $\mathbf{B}(1)$ and $\mathbf{B}^\circ(2)$ are invertible. Thus, we can say that the model written in (3.38) is orthogonal.

3.8.2 Second case

We will now consider the second case where the first model has fixed effects with w(1)=1 and $\mathbf{X}_1(1) = \mathbf{I}_{n(1)}$, thus the first model is written as follows

$$\mathbf{Y}(1) = \mathbf{X}_0(1)\beta_0(1) + \mathbf{X}_1(1)\beta_1(1) \tag{3.44}$$

Knowing that w(1)=1 and $\mathbf{X}_1(1) = \mathbf{I}_{n(1)}$, we have

$$\mathbf{M}_1(1) = \mathbf{X}_1(1)\mathbf{X}_1^T(1) = \mathbf{I}_{n(1)}$$

We consider that the matrix $\mathbf{Q}^\circ(1) = \{\mathbf{I}_{n(1)}\}$ is the principal basis of algebra $\mathcal{A}^\circ(1)$, then

$$\mathbf{Q}_1^\circ(1) = \mathbf{I}_{n(1)} \text{ and } m^\circ(1) = 1,$$

then

$$\mathbf{M}_1(1) = \mathbf{I}_{n(1)} = \mathbf{Q}_1^\circ(1),$$

thus we obtain the following transition matrix

$$\mathbf{B}^\circ(1) = \mathbf{I}_1$$

We now have

$$\mathbf{T}(1) = \mathbf{X}_0(1)\mathbf{X}_0^+(1) = \mathbf{X}_0(1)(\mathbf{X}_0^T(1)\mathbf{X}_0(1))^{-1}\mathbf{X}_0^T(1) = \frac{1}{n(1)}\mathbf{X}_0(1)\mathbf{X}_0^T(1) = \frac{1}{n(1)}\mathbf{J}_{n(1)},$$

and

$$\mathbf{T}^c(1) = \mathbf{I}_{n(1)} - \mathbf{T}(1) = \mathbf{I}_{n(1)} - \frac{1}{n(1)}\mathbf{J}_{n(1)},$$

being

1. $\mathbf{X}_0(1) = \mathbf{1}_{n(1)}$

2. $\mathbf{X}_0^T(1) = \mathbf{1}_{n(1)}^T$

3. $\mathbf{1}_{n(1)}\mathbf{1}_{n(1)}^T = n(1)$

4. $\mathbf{1}_{n(1)}^T\mathbf{1}_{n(1)} = \mathbf{J}_{n(1)}$

We considerer that $\mathbf{Q}(1) = \{\mathbf{T}(1); \mathbf{T}^c(1)\}$ is the principal basis of algebra $\mathcal{A}(1)$, thus we have

$$\mathbf{M}_1(1) = \mathbf{T}(1) + \mathbf{T}^c(1),$$

thus we obtain the following transition matrix

$$\mathbf{B}(1) = \begin{bmatrix} 1 & 1 \end{bmatrix}$$

The second model being mixed is

$$\mathbf{Y}(2) = \sum_{i=0}^{w(2)} \mathbf{X}_i(2)\boldsymbol{\beta}_i(2). \tag{3.45}$$

We assume that

$$\mathbf{X}_0(2) = \begin{bmatrix} \mathbf{1}_{n(1)} & \mathbf{X}_0(2) \end{bmatrix},$$

with $\mathbf{1}_{n(2)}^T\mathbf{X}_0(2) = \mathbf{0}^T$. Through nesting, using the definition [Fonseca et al. (2008)], we get the following model

$$\mathbf{Y} = \sum_{i=0}^{w} \mathbf{X}_i\boldsymbol{\beta}_i \tag{3.46}$$

with w=1+w(2) = w(1)+w(2) and

$$\begin{cases} \mathbf{X}_0 = \begin{bmatrix} \mathbf{X}_0(1) \otimes \mathbf{1}_{n(2)} & \mathbf{I}_{n(1)} \otimes \mathbf{X}_0(2) \end{bmatrix} \\ \mathbf{X}_1 = \mathbf{X}_1(1) \otimes \mathbf{1}_{n(2)} = \mathbf{I}_{n(1)} \otimes \mathbf{1}_{n(2)} \\ \mathbf{X}_i = \mathbf{I}_{n(1)} \otimes \mathbf{X}_{i-1}(2) \qquad\qquad j = 2, ..., w. \end{cases} \tag{3.47}$$

From the expression (3.47), we have

$$\begin{aligned}\mathbf{M}_1 &= \mathbf{X}_1\mathbf{X}_1^T = (\mathbf{I}_{n(1)} \otimes \mathbf{1}_{n(2)})(\mathbf{I}_{n(1)} \otimes \mathbf{1}_{n(2)})^T = \\ &= (\mathbf{I}_{n(1)} \otimes \mathbf{1}_{n(2)})(\mathbf{I}_{n(1)} \otimes \mathbf{1}_{n(2)}^T) = \mathbf{I}_{n(1)} \otimes J_{n(2)} = \\ &= n(2)\sum_{j=1}^{2} \mathbf{Q}_j(1) \otimes \frac{1}{n(2)} J_{n(2)} \text{ with } m(1) = 2.\end{aligned} \tag{3.48}$$

and

$$\begin{aligned}\mathbf{M}_i &= \mathbf{X}_i\mathbf{X}_i^T = (\mathbf{I}_{n(1)} \otimes \mathbf{X}_{i-1}(2))(\mathbf{I}_{n(1)} \otimes \mathbf{X}_{i-1}(2))^T = \\ &= (\mathbf{I}_{n(1)} \otimes \mathbf{X}_{i-1}(2))(\mathbf{I}_{n(1)} \otimes \mathbf{X}_{i-1}^T) = \mathbf{I}_{n(1)} \otimes \mathbf{M}_{i-1}(2) = \\ &= \mathbf{I}_{n(1)} \otimes \left(\sum_{j=1}^{m(2)} b_{i-1,j}(2)\mathbf{Q}_j(2)\right).\end{aligned} \tag{3.49}$$

$\mathbf{Q}_j = \mathbf{Q}_j(1) \otimes \frac{1}{n(2)} J_{n(2)}$; $j = 1, 2$ is the principal basis of algebra $\mathcal{A}(1)$ then of expression (3.48), we have

$$\mathbf{M}_1 = n(2)\sum_{j=1}^{2} \mathbf{Q}_j = n(2)(\mathbf{Q}_1 + \mathbf{Q}_2) \tag{3.50}$$

thus we obtain the following transition matrix

$$\mathbf{B}(1) = \begin{bmatrix} n(2) & n(2) \end{bmatrix}. \tag{3.51}$$

knowing that

$$\mathbf{M}_{i-1}(2) = \sum_{j=1}^{m(2)} b_{i-1,j}(2)\mathbf{Q}_j(2) = b_{i-1,j}(2)\mathbf{Q}_1(2) + \sum_{j=2}^{m(2)} b_{i-1,j}(2)\mathbf{Q}_j(2)$$

From expression (3.49), we have

$$\mathbf{M}_i = \mathbf{I}_{n(1)} \otimes \mathbf{M}_{i-1}(2) = b_{i-1,1}(2)\mathbf{I}_{n(1)} \otimes \frac{1}{n(2)}\mathbf{J}_{n(2)} + \sum_{j=2}^{m(2)} b_{i-1,j}(2)\mathbf{I}_{n(1)} \otimes \mathbf{Q}_j(2)$$

Of expression $b_{i-1,1}(2)\mathbf{I}_{n(1)} \otimes \frac{1}{n(2)}\mathbf{J}_{n(2)}$, we have

$$b_{i-1,1}(2)\mathbf{I}_{n(1)} \otimes \frac{1}{n(2)}\mathbf{J}_{n(2)} = b_{i-1,1}(2)\sum_{j=1}^{2} \mathbf{Q}_j(1) \otimes \frac{1}{n(2)}\mathbf{J}_{n(2)} = b_{i-1,1}(2)\sum_{j=1}^{2} \mathbf{Q}_j,$$

and of expression $b_{i-1,j}(2)\mathbf{I}_{n(1)} \otimes \mathbf{Q}_j(2)$, we have

$$b_{i-1,j}(2)\mathbf{I}_{n(1)} \otimes \mathbf{Q}_j(2) = b_{i-1,j}(2)\mathbf{Q}_{j+1}$$

with $j = \{2, \ldots, m(2)\}$. Thus, the transition matrix is written as follows

$$\mathbf{B}(2) = \begin{bmatrix} \mathbf{b}_1(2) & \mathbf{b}_2(2) & \ldots & \mathbf{b}_m(2) \end{bmatrix} \tag{3.52}$$

Then the matrix $\mathbf{B}^{\circ}(2)$ is

$$\mathbf{B}^{\circ}(2) = \begin{bmatrix} \mathbf{b}_2(2) & \ldots & \mathbf{b}_m(2) \end{bmatrix} \tag{3.53}$$

with m = m(2) + 1.

So, the matrix $\mathbf{M}_i$ is written as follows

$$\mathbf{M}_i = b_{i-1,1}(2)\mathbf{Q}_1 + b_{i-1,1}(2)\mathbf{Q}_2 + \sum_{j=2}^{m(2)} b_{i-1,j}(2)\mathbf{Q}_{j+1}. \tag{3.54}$$

From the expressions (3.50) and (3.54) we obtain the transition matrix

$$\mathbf{B} = \begin{bmatrix} n(2) & n(2) & 0 \ldots & 0 \\ \mathbf{b}_1(2) & \mathbf{b}_1(2) & \mathbf{B}^{\circ}(2) & \end{bmatrix} \tag{3.55}$$

3.9 Equivalence of models

3.9.1 Introduction

In this study, we introduce the notion of equivalence between models. So, we have the following definition:

Definition 41. *Two models are equivalent if they have the same family of:*

1. *estimable vectors;*
2. *covariance matrices*

We show that all the mixed model

$$\mathbf{Y} = \mathbf{X}_{\circ}\boldsymbol{\beta}_{\circ} + \sum_{i=1}^{w} \mathbf{X}_i\boldsymbol{\beta}_i \tag{3.56}$$

where $\boldsymbol{\beta}_{\circ}$ is fixed and the $\boldsymbol{\beta}_i$, $i = \{1, \ldots, w\}$ have null mean vectors and cross-covariances matrices and covariance matrices $\theta_j\mathbf{I}_{c_j}$, $j = 1, \ldots, w$, is equivalent to a model centered

$$\overset{\circ}{\mathbf{Y}} = \overset{\circ}{\mathbf{X}}_{\circ}\overset{\circ}{\boldsymbol{\beta}}_{\circ} + \sum_{i=1}^{w} \mathbf{X}_i\boldsymbol{\beta}_i \tag{3.57}$$

with $\mathbf{T} = \overset{\circ}{\mathbf{X}}_\circ \overset{\circ}{\mathbf{X}}_\circ^+$, the orthogonal projection matrix on the space spanned by the mean vector. This notion allows to introduce the role played by the part of the fixed effects model.

Being $\mathbf{U} = \sum_{j \in \varphi} \mathbf{U}_j$ an orthogonal projection matrix in a CJA with principal basis $\{\mathbf{U}_1, \dots, \mathbf{U}_m\}$, we know that $R(\mathbf{U}) = \boxplus_{j \in \varphi(\mathbf{M})} R(\mathbf{U}_j)$, as mentioned earlier in section 2.7.1. If $\overset{\circ}{\mathbf{U}} = GS(\mathbf{U})$, with column vectors obtained by applying the Gram Schmidt process to the columns vectors of $\mathbf{U}$, we have:

$$Pr(\mathbf{U}) = \overset{\circ}{\mathbf{U}} \overset{\circ}{\mathbf{U}}^T$$

with $Pr(\mathbf{U})$ the orthogonal projection matrix on $R(\mathbf{U})$

3.9.2 Equivalence of models

The mixed models 3.56 and 3.57, where $\overset{\circ}{\mathbf{X}}_\circ = GS(\mathbf{X}_\circ)$, are equivalent, because they clearly have the same covariance matrices and $R(\mathbf{X}_\circ) = R(\overset{\circ}{\mathbf{X}}_\circ)$.

We now have the following application. The mixed model 3.56 has COBS if the matrix $\mathbf{T}$ commutes with the matrices $\mathrm{M} = \{\mathbf{M}_1, \dots, \mathbf{M}_w\}$. The model 3.57 is orthogonal and centered if the matrix $\overset{\circ}{\mathbf{M}}_\circ = \mathbf{T}$ commutes with the matrices $\mathrm{M} = \{\mathbf{M}_1, \dots, \mathbf{M}_w\}$. So, it is easy to establish the

Proposition 22. *Every model with COBS is equivalent to an ORT centered model.*

Relatively to the crossing of models, we have the following mixed models:

$$\mathcal{M}_l = \mathbf{X}_\circ(l)\boldsymbol{\beta}_\circ(l) + \sum_{j=1}^{w_l} \mathbf{X}_j(l)\boldsymbol{\beta}_j(l), \; l = 1, 2 \tag{3.58}$$

and

$$\overset{\circ}{\mathcal{M}}_l = \overset{\circ}{\mathbf{X}}_\circ(l)\overset{\circ}{\boldsymbol{\beta}}_\circ(l) + \sum_{j=1}^{w_l} \mathbf{X}_j(l)\boldsymbol{\beta}_j(l), \; l = 1, 2 \tag{3.59}$$

Knowing that:

$$\mathcal{M}_1 \equiv \overset{\circ}{\mathcal{M}}_1$$

and

$$\mathcal{M}_2 \equiv \overset{\circ}{\mathcal{M}}_2.$$

We say that $\mathcal{M}_1$ is equivalent the $\overset{\circ}{\mathcal{M}}_1$, because

$$\begin{cases} \mathbf{T}_1 = \overset{\circ}{\mathbf{T}}_1 \\ E(\mathcal{M}_1) = E(\overset{\circ}{\mathcal{M}}_1) \\ V(\mathcal{M}_1) = V(\overset{\circ}{\mathcal{M}}_1) \end{cases}$$

being $E(\mathcal{M}) = \mathbf{U}\mu$ where

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

and

$$\begin{cases} \boldsymbol{\beta}_\circ = \mathbf{X}_\circ^+ \mu = \mathbf{X}_\circ^+ \mathbf{X}_\circ \boldsymbol{\beta}_\circ \\ \overset{\circ}{\boldsymbol{\beta}}_\circ = \overset{\circ}{\mathbf{X}}_\circ^{\;+} \mu = \overset{\circ}{\mathbf{X}}_\circ^{\;+} \overset{\circ}{\mathbf{X}}_\circ \overset{\circ}{\boldsymbol{\beta}}_\circ \end{cases}$$

the same happening for $\mathcal{M}_2$ to be equivalent the $\overset{\circ}{\mathcal{M}}_2$.

So, if we cross the models, we have

$$\mathcal{M}_1 \otimes \mathcal{M}_2 = \overset{\circ}{\mathcal{M}}_1 \otimes \overset{\circ}{\mathcal{M}}_2$$

with

$$\begin{aligned} \mathcal{M}_1 \otimes \mathcal{M}_2 = (\mathbf{X}_\circ(1) \otimes \mathbf{X}_\circ(2))\boldsymbol{\beta}_\circ + \sum_{i=1}^{w(2)} (\mathbf{X}_\circ(1) \otimes \mathbf{X}_i(2))\boldsymbol{\beta}_i + \\ + \sum_{i=w(2)+1}^{w(2)+w(1)} (\mathbf{X}_i(1) \otimes \mathbf{X}_\circ(2))\boldsymbol{\beta}_{i-w(1)} + \sum_{i=1}^{w_1} \sum_{j=1}^{w_2} (\mathbf{X}_i(1) \otimes \mathbf{X}_j(2))\boldsymbol{\beta}_{l(ij)} \end{aligned} \tag{3.60}$$

with l(ij) = w(1) + w(2) + (i - 1) w(2) + j + 1.

This happens because, knowing that

$$\begin{cases} \mathbf{T}(l) = \mathbf{X}_\circ(l)\mathbf{X}_\circ^+(l), \;\; l = 1, 2 \\ \overset{\circ}{\mathbf{T}}(l) = \overset{\circ}{\mathbf{X}}_\circ(l)\overset{\circ}{\mathbf{X}}_\circ^+(l), \;\; l = 1, 2 \end{cases}$$

we have

$$\begin{aligned} \mathbf{T}(1) \otimes \mathbf{T}(2) &= (\mathbf{X}_\circ(1)\mathbf{X}_\circ^+(1)) \otimes (\mathbf{X}_\circ(2)\mathbf{X}_\circ^+(2)) = \\ &= (\mathbf{X}_\circ(1) \otimes \mathbf{X}_\circ(2))(\mathbf{X}_\circ^+(l) \otimes \mathbf{X}_\circ^+(2)) \\ &= \mathbf{X}_\circ \mathbf{X}_\circ^+ = \mathbf{T} \end{aligned} \tag{3.61}$$

and

$$
\begin{aligned}
\overset{\circ}{\mathbf{T}}(1) \otimes \overset{\circ}{\mathbf{T}}(2) &= (\overset{\circ}{\mathbf{X}}_\circ(1)\overset{\circ}{\mathbf{X}}{}_\circ^+(1)) \otimes (\overset{\circ}{\mathbf{X}}_\circ(2)\overset{\circ}{\mathbf{X}}{}_\circ^+(2)) = \\
&= (\overset{\circ}{\mathbf{X}}_\circ(1) \otimes \overset{\circ}{\mathbf{X}}_\circ(2))(\overset{\circ}{\mathbf{X}}{}_\circ^+(1) \otimes \overset{\circ}{\mathbf{X}}{}_\circ^+(2)) \\
&= \overset{\circ}{\mathbf{X}}_\circ \overset{\circ}{\mathbf{X}}{}_\circ^+ = \overset{\circ}{\mathbf{T}}
\end{aligned}
\tag{3.62}
$$

So, knowing that

$$\mathbf{T}(l) = \overset{\circ}{\mathbf{T}}(l),\ l = 1, 2$$

we have

$$\mathbf{T} = \overset{\circ}{\mathbf{T}}$$

and, we can say that

$$\mathbf{T}_1 \otimes \mathbf{T}_2 = \overset{\circ}{\mathbf{T}}_1 \otimes \overset{\circ}{\mathbf{T}}_2$$

since that

$$Pr(\mathbf{X}_\circ(1) \otimes \mathbf{X}_\circ(2)) = Pr(\mathbf{X}_\circ(1)) \otimes Pr(\mathbf{X}_\circ(2))$$

3.9.3 OBS Models

Let Q $= \{\mathbf{Q}_1, \ldots, \mathbf{Q}_m\}$ be the principal basis of the Jordan Algebra $\mathcal{A}(\mathbf{M})$, with M $= \{\mathbf{M}_1, \ldots, \mathbf{M}_w\}$, and $\mathbf{M}_1, \ldots, \mathbf{M}_w$ linearly independent. If m = w, $\mathbf{M}$ will be a perfect family and will be a basis for the algebra $\mathcal{A}(\mathbf{M})$. with principal basis $\mathcal{A}(\mathbf{M}) = \{\mathbf{Q}_1, \ldots, \mathbf{Q}_m\}$. If the $\mathbf{g}_j$ row vectors of $\mathbf{A}_j$ constitute an orthonormal basis for the range space $R(\mathbf{Q}_j)$ of $\mathbf{Q}_j$ $j = \{1, \ldots, m\}$. We then have

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

will have the mean vector

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

with

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

and covariance matrix

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

Let $\mathbf{P}_j$ and $\mathbf{P}_j^{\perp}$ be the orthogonal projection matrices on $R(\mathbf{X}_{\circ j})$ and on its orthogonal complement. The matrices are written as follows:

$$\begin{cases} \mathbf{P}_j = Pr(\mathbf{X}_{\circ j}) \\ \mathbf{P}_j^{\perp} = \mathbf{I}_{g_j} - \mathbf{P}_j \end{cases},$$

so, we have

$$\begin{cases} \overset{\circ}{\mathbf{Q}}_{2j-1} = \mathbf{A}_j^T \mathbf{P}_j \mathbf{A}_j = \overset{\circ}{\mathbf{A}}{}_{2j-1}^T \overset{\circ}{\mathbf{A}}_{2j-1} \\ \overset{\circ}{\mathbf{Q}}_{2j} = \mathbf{A}_j^T \mathbf{P}_j^{\perp} \mathbf{A}_j = \overset{\circ}{\mathbf{A}}{}_{2j}^T \overset{\circ}{\mathbf{A}}_{2j} \end{cases},$$

$$\mathbf{Q}_j = \overset{\circ}{\mathbf{Q}}_{2j-1} + \overset{\circ}{\mathbf{Q}}_{2j}$$

with

$$\mathbf{A}_j = \begin{bmatrix} \overset{\circ}{\mathbf{A}}_{2j-1} \\ \overset{\circ}{\mathbf{A}}_{2j} \end{bmatrix} \tag{3.63}$$

Knowing that

$$\mathbf{T} = \sum_{j=1}^{m} \overset{\circ}{\mathbf{Q}}_{2j-1}$$

with $\mathbf{T} = \overset{\circ}{\mathbf{X}}_{\circ}$, $\mathbf{X}_{\circ} = \sum_{j'=1}^{m} \overset{\circ}{\mathbf{A}}{}_{2j'-1}^T \overset{\circ}{\mathbf{A}}_{2j'-1}$ and $\mathbf{X}_{\circ j}$ a $(g_j \times n)$ matrix, we have:

$$\begin{aligned} \mathbf{X}_{\circ j} &= \begin{bmatrix} \overset{\circ}{\mathbf{A}}_{2j-1} \\ \overset{\circ}{\mathbf{A}}_{2j} \end{bmatrix} \left(\sum_{j'=1}^{m} \overset{\circ}{\mathbf{A}}{}_{2j'-1}^T \overset{\circ}{\mathbf{A}}_{2j'-1} \right) = \\ &= \begin{bmatrix} \mathbf{I}_{\overset{\circ}{g}_{2j-1}} \\ \mathbf{O}_{\overset{\circ}{g}_{2j}, \overset{\circ}{g}_{2j-1}} \end{bmatrix} \end{aligned} \tag{3.64}$$

So, the matrices $\mathbf{P}_j$ and $\mathbf{P}_j^{\perp}$ are written as follows:

$$\mathbf{P}_j = \begin{bmatrix} \mathbf{I}_{\overset{\circ}{g}_{2j-1}} & \mathbf{O} \\ \mathbf{O} & \mathbf{O}_{\overset{\circ}{g}_{2j} \times \overset{\circ}{g}_{2j}} \end{bmatrix} = Pr(R(\mathbf{X}_{0j}))$$

and

$$\mathbf{P}_j^{\perp} = \begin{bmatrix} \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \mathbf{I}_{\overset{\circ}{g}_{2j}} \end{bmatrix} = Pr(R(\mathbf{X}_{0j}^{\perp}))$$

Chapter 4

Normal models with OBS

4.1 Introduction

As stated previously in the introduction, [Nelder (1965a)] introduced models with ortoghonal block structure, OBS, models with variance-covariance matrices

$$\mathbf{V} = \sum_{j=1}^{m} \gamma_j \mathbf{Q}_j \tag{4.1}$$

where $\mathbf{Q}_1, ..., \mathbf{Q}_m$ are know OPOPM. We now intend to study normal models with OBS, NOBS (Normal Orthogonal Block Structure); namely studying estimation problems and obtaining condition for having complete, sufficient statistics.

4.2 Sufficient statistics and natural parameters

Let the row vectors of matrix $\mathbf{A}_j$ constitute an orthonormal basis for the range space $\nabla_j = R(\mathbf{Q}_j)\ of\ \mathbf{Q}_j,\ j = \{1, ..., m\}$. Then, with $g_j = rank(\mathbf{Q}_j) = rank(\mathbf{A}_j)$ we have

$$\begin{cases} \mathbf{A}_j\mathbf{A}_j^T = \mathbf{I}_{g_j} & ;\ \text{j=\{1,...,m\}} \\ \mathbf{A}_j^T\mathbf{A}_j = \mathbf{Q}_j & ;\ \text{j=\{1,...,m\}} \end{cases}$$

If the initial NOBS model $\mathbf{Y}$ has besides variance-covariance matrix $\mathbf{V}$, mean vector $\boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta}$, we get the homocedastic sub-models

$$\mathbf{Y}_j = \mathbf{A}_j\mathbf{Y},\ j = \{1, ..., m\} \tag{4.2}$$

with mean vectors $\boldsymbol{\mu}_j = \mathbf{X}_j\boldsymbol{\beta}$ where $\mathbf{X}_j = \mathbf{A}_j\mathbf{X}$ and variance-covariance matrices $\boldsymbol{\gamma}_j\mathbf{I}_{g_j},\ j = \{1, ..., m\}$. These sub-vectors will have joint normal distribution and are independent since as it is easy to see, their cross-covariance matrices are null. Let $\mathbf{P}_j$ be the orthogonal projection matrix on $\Omega_j = R(\mathbf{X}_j)$ and $\mathbf{L}_j$ a matrix where row vectors constitute an orthonormal basis for Ω_j, so $\mathbf{P}_j = \mathbf{L}_j^T\mathbf{L}_j,\ j = \{1, ..., m\}$. And so

$$\begin{aligned}(\mathbf{Y}-\boldsymbol{\mu})^T\mathbf{V}^{-1}(\mathbf{Y}-\boldsymbol{\mu}) &= (\mathbf{Y}-\boldsymbol{\mu})^T\sum_{j=1}^{m}\boldsymbol{\gamma}_j^{-1}\mathbf{Q}_j(\mathbf{Y}-\boldsymbol{\mu}) = \\ &= \sum_{j=1}^{m}\boldsymbol{\gamma}_j^{-1}(\mathbf{Y}-\boldsymbol{\mu})^T\mathbf{A}_j^T\mathbf{A}_j(\mathbf{Y}-\boldsymbol{\mu}) = \\ &= \sum_{j=1}^{m}\frac{||\mathbf{Y}_j-\boldsymbol{\mu}_j||^2}{\boldsymbol{\gamma}_j}\end{aligned} \tag{4.3}$$

Thus

$$\begin{aligned}||\mathbf{Y}_j-\boldsymbol{\mu}_j||^2 &= (\mathbf{Y}_j-\boldsymbol{\mu}_j)^T\mathbf{L}_j^T\mathbf{L}_j(\mathbf{Y}_j-\boldsymbol{\mu}_j) + (\mathbf{Y}_j-\boldsymbol{\mu}_j)^T\mathbf{P}_j^{\perp}(\mathbf{Y}_j-\boldsymbol{\mu}_j) = \\ &= \mathbf{Y}_j^T\mathbf{P}_j\mathbf{Y}_j - 2\boldsymbol{\eta}_j^T\tilde{\boldsymbol{\eta}}_j + \boldsymbol{\eta}_j^T\boldsymbol{\eta}_j + \mathbf{Y}_j^T\mathbf{P}_j^{\perp}\mathbf{Y}_j = \\ &= s_j - 2\boldsymbol{\eta}_j^T\tilde{\boldsymbol{\eta}}_j + ||\boldsymbol{\eta}_j||^2,\ j = \{1, \ldots, z\}\end{aligned} \tag{4.4}$$

with $\boldsymbol{\eta}_j = \mathbf{L}_j\boldsymbol{\mu}_j, \tilde{\boldsymbol{\eta}}_j = \mathbf{L}_j\mathbf{Y}_j$ and $\mathbf{s}_j = \|\mathbf{Y}_j\|^2,\ j = \{1, ..., m\}$, then the densities for the sub-models will be

$$\begin{aligned}n_j(\mathbf{y}_j) &= \frac{e^{-\frac{1}{2\boldsymbol{\gamma}_j}(\mathbf{s}_j - 2\boldsymbol{\eta}_j^T\tilde{\boldsymbol{\eta}}_j + \|\boldsymbol{\eta}_j\|^2)}}{(2\pi\boldsymbol{\gamma}_j)^{\frac{g_j}{2}}} = \\ &= \frac{e^{-\frac{1}{2}(\boldsymbol{\theta}_j\mathbf{s}_j - 2\boldsymbol{\eta}_j^T\tilde{\boldsymbol{\eta}}_j + \boldsymbol{\theta}_j^{-1}\|\boldsymbol{\xi}_j\|)}}{(2\pi\boldsymbol{\gamma}_j)^{\frac{g_j}{2}}}\end{aligned} \tag{4.5}$$

So we have the natural parameters $\boldsymbol{\theta}_j = \dfrac{1}{\boldsymbol{\gamma}_j}$ and $\boldsymbol{\xi}_j = \dfrac{1}{\boldsymbol{\gamma}_j}\boldsymbol{\eta}_j,\ j = \{1, ..., m\}$. We point out that with $p_j = rank(\mathbf{P}_j) = rank(\mathbf{L}_j)$ we have $0 \neq p_j \leq g_j$ with $z = z_1 > 0[z_1 \leq z_2 \leq m]$, we may assume that

$$\begin{cases} p_j = g_j & ; j = \{1, ..., z_1\} \\ 0 < p_j < g_j & ; j = \{z_{1+1}, ..., z_2\} \\ p_j = 0 & ; j = \{z_{2+1}, ..., m\} \end{cases}$$

where as stated above we may have $z_1 = [z_1 = z_2; z_2 = m]$. If $z_2 < m$ we take $\tilde{\boldsymbol{\eta}} = \boldsymbol{\xi}_j = \mathbf{0}_{g_j},\ j = \{z_{2+1}, ..., m\}$. The joint density of the sub-models will then be

$$n(\mathbf{y}) = \frac{e^{-\frac{1}{2}\sum_{j=1}^{m}(\boldsymbol{\theta}_j \mathbf{s}_j - 2\boldsymbol{\xi}_j^T \boldsymbol{\eta}_j + \boldsymbol{\theta}_j \|\boldsymbol{\xi}_j\|^2)}}{\prod_{j=1}^{m}(2\pi\boldsymbol{\gamma}_j)^{\frac{g_j}{2}}} \tag{4.6}$$

with sufficient statistics $(\mathbf{s}_j, \tilde{\boldsymbol{\eta}}_j),\ j = \{1, ..., m\}$, and natural parameters $\boldsymbol{\theta}_j$, with components $\{\boldsymbol{\theta}_1, ..., \boldsymbol{\theta}_m\}$, and

$$\boldsymbol{\xi} = [\boldsymbol{\xi}_1^T, ..., \boldsymbol{\xi}_m^T]^T$$

The completeness of the statistics will only hold, see [Lehmann & Casela (1998)], if the parameter space contains the Cartesian product of non degenerate intervals one for each of the natural parameter. This ceases to happen when restrictions on the parameters are introduced.

4.3 Mixed models

We now assume that

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

where $\boldsymbol{\beta}_0$ is fixed. For $\mathbf{Y}$ to have NOBS it suffices to assume that the $\{\boldsymbol{\beta}_1, ..., \boldsymbol{\beta}_w\}$ are normal independent, with null mean vectors and variance-covariance matrices $\{\boldsymbol{\theta}_1 \mathbf{I}_{c_1}, ..., \boldsymbol{\theta}_w \mathbf{I}_{c_w}\}$ see [Fonseca (2007)], that the matrices $\mathbf{M}_i = \mathbf{X}_i \mathbf{X}_i^T,\ i = \{1, ..., w\}$ commute. Thus we will have, see [Fonseca (2007)]

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

where the $\{\mathbf{Q}_1, ..., \mathbf{Q}_m\}$ are OPOPM that could lead to in wherever we assume the variance-covariance matrix

$$\mathbf{V} = \sum_{i=1}^{w} \boldsymbol{\theta}_i \mathbf{M}_i = \sum_{j=1}^{m} \gamma_j \mathbf{Q}_j,$$

with

$$\gamma_j = \sum_{i=1}^{w} b_{ij} \boldsymbol{\theta}_i, \; j = \{1, ..., m\}$$

Thus with

$$\boldsymbol{\gamma} = \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_m \end{bmatrix}; \boldsymbol{\theta} = \begin{bmatrix} \boldsymbol{\theta}_1 \\ \vdots \\ \boldsymbol{\theta}_w \end{bmatrix}; \mathbf{B} = \begin{bmatrix} b_{ij} \end{bmatrix}$$

We have $\boldsymbol{\gamma} = \mathbf{B}^T \boldsymbol{\theta} \in R(\mathbf{B}^T)$. Now, see again [Fonseca (2007)], we have $w \leq m$ and so, when $w < m$, there are linear restrictions on the components of $\boldsymbol{\gamma}$ and so the parameter space cannot contain a non degenerate interval for each of the $\{\gamma_1, ..., \gamma_m\}$. This is equivalent to not having non degenerate interval for the $\{\theta_1, ..., \theta_m\}$ and so, when $w < m$, we cannot have complete sufficient statistics. When $w = m$, the $\{\mathbf{M}_1, ..., \mathbf{M}_m\}$ constitute a basis for the smallest commutative Jordan Algebra that contains these matrices. Thus, see [Ferreira et al. (2007)], $\mathbf{M} = \{\mathbf{M}_1, ..., \mathbf{M}_m\}$ will be a perfect family of symmetric matrices. Observe that

$$\boldsymbol{\Psi} = G\boldsymbol{\beta}$$

is estimable if and only if

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

Thus we may replace $\mathbf{X}_o$ by $\overset{\circ}{\mathbf{X}}_o$ another $n \times k$ matrix with the same range space Ω thus keeping the family of estimable vectors. namely we may choice $\overset{\circ}{\mathbf{X}}_o$ to have column vectors that constitute an orthonormal basis for Ω thus having, besides $\boldsymbol{\mu} = \overset{\circ}{\mathbf{X}}_o \boldsymbol{\beta}$, the Moore-Penrose Inverse

$$\overset{\circ}{\mathbf{X}}_o^{+} = \overset{\circ}{\mathbf{X}}_o^{T}$$

so that

$$\boldsymbol{\beta} = \overset{\circ}{\mathbf{X}}_o^{T} \boldsymbol{\mu}$$

we also have

$$\boldsymbol{\mu} = \mathbf{A}^T \begin{bmatrix} \boldsymbol{\mu}_1 \\ \vdots \\ \boldsymbol{\mu}_m \end{bmatrix}$$

when $\mathbf{A} = [\mathbf{A}_1^T ... \mathbf{A}_m^T]^T$ is an orthogonal matrix. Since $\boldsymbol{\mu}_j = \mathbf{L}_j^T \boldsymbol{\eta}_j$, $j = \{1, ..., m\}$, we also will have

$$\boldsymbol{\mu} = \mathbf{A}^T D\left(\mathbf{L}_1^T, ..., \mathbf{L}_m^T\right) \boldsymbol{\eta}$$

where $D\left(\mathbf{L}_1^T, ..., \mathbf{L}_m^T\right)$ is the blockwise diagonal matrix with principal blocks $\mathbf{L}_1^T, ..., \mathbf{L}_m^T$ and $\boldsymbol{\eta} = [\boldsymbol{\eta}_1^T, ..., \boldsymbol{\eta}_m^T]^T$, then

$$\boldsymbol{\beta} = \overset{\circ}{\mathbf{X}}_o^{T} \mathbf{A}^T D\left(\mathbf{L}_1^T ... \mathbf{L}_m^T\right) \boldsymbol{\eta}$$

now $\boldsymbol{\beta}$ may be any vector in R^k so

$$rank(\overset{\circ}{\mathbf{X}}_o^{T} \mathbf{A}^T D(\mathbf{L}_1^T ... \mathbf{L}_m^T)) = k$$

and, since $\overset{\circ}{\mathbf{X}}_o^{T} \mathbf{A}^T D(\mathbf{L}_1^T ... \mathbf{L}_m^T)$ is a $k \times p$ with

$$p = \sum_{j=1}^{m} rank(X_j)$$

When $k < p$ there will be linear restrictions on the components of η thus on the components of

$$\boldsymbol{\xi} = [\boldsymbol{\xi}_1^T ... \boldsymbol{\xi}_m^T]^T$$

and we cannot have complete sufficient statistics. We thus get the following proposition

Proposition 23. *To have complete and sufficient statistics we must have $w = m$ and $k = p$*

Proof: The previous discussion establishes the necessary condition. For sufficiency we start by pointing out that $\boldsymbol{\theta}$ may be any vector with non negative components and so, when $w = m$, and $\mathbf{B}^T$ is invertible there may be no restrictions the components of $\boldsymbol{\gamma}$ since we now have $\boldsymbol{\theta} = \left(\mathbf{B}^T\right)^{-1} \boldsymbol{\gamma}$. Likewise when $p = k$ matrix $\mathbf{Z} = \overset{\circ}{\mathbf{X}}{}_o^T \mathbf{A}^T D(\mathbf{L}_1^T ... \mathbf{L}_m^T)$ will be invertible and we can reason as above to show that there may be no restrictions on the components of $\boldsymbol{\eta}$ and on those of $\boldsymbol{\xi}$. Since the existence of restrictions on the components of $\boldsymbol{\gamma}$ and $\boldsymbol{\eta}$ was the sole reason for there not being non degenerate interval natural parameter the proof is complete. $\square$

Above we assumed that, with $z_1 \geq 0$, we have

$$\begin{cases} p_j = g_j & ; j = \{1, ..., z_1\} \\ p_j < g_j & ; j = \{z_{1+1}, ..., m\} \end{cases}$$

Assuming that $z_1 > 0$ it is easy to see that for $\mathbf{Y}_j,\ j = \{1, ..., z_1\}$, we cannot obtain any unbiased estimator for the corresponding $\boldsymbol{\eta}_j,\ j = \{1, ..., z\}$. Thus, we also have the

$$\tilde{\boldsymbol{\eta}}_j = \frac{\mathbf{Y}_j^T \mathbf{P}_j^{\perp} \mathbf{Y}_j}{f_j},\ j = \{z + 1, ..., w\}$$

with $\mathbf{P}_j^{\perp} = \mathbf{I}_{g_j} - \mathbf{P}_j$ and $f_j = g_j - p_j,\ j = \{z+1, ..., m\}$. Moreover with

$$\boldsymbol{\gamma}(1) = \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_{z_1} \end{bmatrix};\ \boldsymbol{\gamma}(2) = \begin{bmatrix} \gamma_{z_1+1} \\ \vdots \\ \gamma_m \end{bmatrix}$$

and taking the matrix partition

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}(1) & \mathbf{B}(2) \end{bmatrix}$$

where $\mathbf{B}(1)$ has z_1 columns we will have $\boldsymbol{\gamma}(l) = \mathbf{B}(l)^T\boldsymbol{\theta}; l = 1, 2$ and, moreover when the row vectors of $\mathbf{B}(2)$ are linearly independent,

$$\boldsymbol{\theta} = \mathbf{B}(2)^{T^-}\boldsymbol{\gamma}(2)$$

with $\mathbf{B}(2)^{T^-}$a generalized inverse of $\mathbf{B}(2)^T$. We then also have

$$\boldsymbol{\gamma}(1) = \mathbf{B}(1)^T\mathbf{B}(2)^{T^-}\boldsymbol{\gamma}(2)$$

replacing in these expression $\boldsymbol{\gamma}(2)$ by it's estimator with components $\{\tilde{\gamma}_{z_1+1}, ..., \tilde{\gamma}_m\}$. So we havean estimate for $\boldsymbol{\theta}$ and $\boldsymbol{\gamma}(2)$. We then say that we have segregation, see for instance [Ferreira S. (2006)], since the relevant parameters, $\boldsymbol{\theta}$ and $\boldsymbol{\gamma}(2)$, for the random effects parts of the model determine each other so this part segregate random effects part as a sub-model. We point out that for having segregation we must have

$$w \leq m - z$$

so we cannot have, then complete and sufficient statistics. Thus in this case we have a trade-off between the existence of UMVUE and the possibility of completing the estimation of variance components. We may also considerer the other conditions for having UMVUE which is more related to estimable vectors. Now $\mathbf{\Psi}$ is an estimable vectors if and only if

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

So the family of estimable vectors will be determined by the space Ω spanned by $\boldsymbol{\mu}$. Then we may, given Ω, choose a convenient matrix with range space Ω. Namely we can work with the orthogonal projection matrix $\mathbf{T}$ on Ω. Now the condition we are considering is

$$rank(\overset{\circ}{\mathbf{X}}{}_o^T \mathbf{A}^T D(\mathbf{L}_1^T, ..., \mathbf{L}_m^T)) = k$$

when $\overset{\circ}{\mathbf{X}}_o = \mathbf{T}$, and so we work with

$$\mathbf{T}\mathbf{A}^T D\left(\mathbf{L}_1^T, ..., \mathbf{L}_m^T\right) = [\mathbf{T}\mathbf{A}_1^T, ..., \mathbf{T}\mathbf{A}_m^T] D\left(\mathbf{L}_1^T, ..., \mathbf{L}_m^T\right) = [\mathbf{T}\mathbf{A}_1^T\mathbf{L}_1^T, ..., \mathbf{T}\mathbf{A}_m^T\mathbf{L}_m^T]$$

If matrix $\mathbf{T}$ commutes with $\mathbf{Q}_j,\ j = \{1, ..., m\}$, the model will have commutative orthogonal block structure, COBS. Now matrices

$$\overset{\circ}{\mathbf{Q}}_j = \mathbf{T}\mathbf{Q}_j = \mathbf{Q}_j\mathbf{T},\ j = \{1, ..., m\},$$

will be symmetric and idempotent so they will be orthogonal projection matrices. Hence it is easy to see that they are OPOPM. We also have

$$\overset{\circ}{\mathbf{Q}}_j = \mathbf{T}\mathbf{Q}_j\mathbf{T} = \mathbf{T}\mathbf{A}_j^T\mathbf{A}_j\mathbf{T} = \overset{\circ}{\mathbf{A}}_j\mathbf{A}_j^T,\ j = \{1, ..., m\}$$

with

$$\overset{\circ}{\mathbf{A}}_j = \mathbf{A}_j\mathbf{T},\ j = \{1, ..., m\},$$

so

$$\mathbf{T}\mathbf{A}^T D(\mathbf{L}_1^T, ..., \mathbf{L}_m^T) = [\overset{\circ}{\mathbf{A}}{}_1^T \mathbf{L}_1^T, ..., \overset{\circ}{\mathbf{A}}{}_m^T \mathbf{L}_m^T]$$

Now, since $\mathbf{T}$ and $\mathbf{Q}_j$ commute, $j = \{1, ..., m\}$,

$$\begin{aligned} R\left(\overset{\circ}{\mathbf{A}}{}_j^T \mathbf{L}_j^T\right) \subseteq R\left(\overset{\circ}{\mathbf{A}}{}_j^T\right) = R\left(\overset{\circ}{\mathbf{A}}{}_j^T \overset{\circ}{\mathbf{A}}_j\right) = R\left(\overset{\circ}{\mathbf{Q}}_j\right) = \\ = R\left(\mathbf{T}\mathbf{Q}_j\right) = R\left(\mathbf{Q}_j\mathbf{T}\right) \subseteq R\left(\mathbf{Q}_j\right),\ j = \{1, ..., m\} \end{aligned} \tag{4.7}$$

The $R\left(\overset{\circ}{\mathbf{A}}{}_j^T \mathbf{L}_j^T\right)$, $j = \{1, ..., m\}$ being subspaces of orthogonal spaces are orthogonal. Thus

$$\begin{aligned} k &= rank\left(\mathbf{T}\mathbf{A}^T D(\mathbf{L}_1^T ... \mathbf{L}_m^T)\right) = \\ &= dim\left(R(\mathbf{T}\mathbf{A}^T D(\mathbf{L}_1^T ... \mathbf{L}_m^T)\right) = \\ &= \sum_{j=1}^{m} dim\left(R\left(\overset{\circ}{\mathbf{A}}{}_j^T \mathbf{L}_j^T\right)\right) = \\ &= \sum_{j=1}^{m} p_j = \\ &= p \end{aligned} \tag{4.8}$$

since

$$rank(\overset{\circ}{\mathbf{A}}{}_j^T \mathbf{L}_j^T) \leq rank(\mathbf{L}_j^T) = p_j = rank(\overset{\circ}{\mathbf{A}}_j \overset{\circ}{\mathbf{A}}{}_j^T \mathbf{L}_j^T) \leq rank(\overset{\circ}{\mathbf{A}}{}_j^T \mathbf{L}_j^T),\ j = \{1, ..., m\}$$

We thus establishes

Proposition 24. *For models with COBS the condition $p = k$ holds.*

Chapter 5

Orthogonal Models

5.1 Introduction

A mixed model has OBS if its variance covariance matrices are all positive semi definite linear combinations, so, it is written in the following form:

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

where $\mathbf{Q}_1, \ldots, \mathbf{Q}_m$ are known ortoghonal parwise projection matrices, OPOPM, that add up to $\mathbf{I}_n$, thus,

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

with ∇_+ the family of vectors of subspace ∇ with non negative components. We are interested in mixed models, see expression 3.15 in section 3.5 and [Pereira et al. (2014)]

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

that have:

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

Following, [Vanleeuwen et al. (1998)], 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.

5.2 Algebraic Structure

5.2.1 Orthogonal Block Structure, OBS

In the study of the mixed models, we will use Commutative Jordan Algebras of symmetric matrices, CJA, see [Pereira et al. (2014)] and section 2.7.1. 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

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

If the $\mathbf{M}_i,\ i = \{1, ..., w\}$, commute, they generate a CJA $\mathcal{A}$ with principal basis Q $= \{\mathbf{Q}_1, \ldots, \mathbf{Q}_m\}$.

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

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

Now $\mathcal{A}$ is a linear space constituted by symmetric matrices that, given the $\mathbf{Q}_1, \ldots, \mathbf{Q}_m$ are OPOPM, commute. The squares of the matrices of $\mathcal{A}$ will, since the $\mathbf{Q}_1, \ldots, \mathbf{Q}_m$ are OPOPM, belong to $\mathcal{A}$. So $\mathcal{A}$ will be a commutative Jordan algebra, CJA. Now, see [Seely (1971)], every CJA $\mathcal{A}^{\circ}$ has an unique basis, this is, the principal basis $pb(\mathcal{A}^{\circ})$, constituted by OPOPM. It is easy to see that

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

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

$$Q = pb(\mathcal{A}),\ is\ OBS(\mathcal{A}).$$

Since,

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

we also have

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

and so

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

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 [Pereira et al. (2014)] and proposition 12 in section 3.6 of this thesis, 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 $\mathbf{M} = \{\mathbf{M}_1, \dots, \mathbf{M}_m\}$ will be a basis for $\mathcal{A}$, $\mathbf{M} = b(\mathcal{A})$. We now have the following proposition:

Proposition 25. *A mixed model*

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

has 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 $\boldsymbol{\gamma} = \mathbf{B}^T\boldsymbol{\theta} \in \mathbb{R}(\mathbf{B}^T)$ □

This proof is in [Pereira et al. (2014)].

5.2.2 Commutative Orthogonal Block Structure, COBS

Now, see [Zmyślony (1978)], 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

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

following [Fonseca et al. (2008)], we say that these models have commutative orthogonal block structure, COBS, associated with $\mathcal{A}$, $COBS(\mathcal{A})$.

Combining proposition (25) with the lemma (6), we get, see [Pereira et al. (2014)].

Proposition 26. *The orthogonal models are models with COBS* $(\mathcal{A})$.

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

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

when both $\mathcal{M}_1$ and $\mathcal{M}_2$ 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 $R(\mathbf{Q}_j)$ of $\mathbf{Q}_j,\ j = 1, \ldots, m$, we have

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

and

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

will have mean vector

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

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

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

Since $\mathbf{T}$ and $\mathbf{Q}_j$ commute, $\mathbf{Q}_j^\circ = \mathbf{T}\mathbf{Q}_j = \mathbf{Q}_j\mathbf{T}$ will be symmetric and idempotent, so they are OPM. We point out that when a product matrix is null it does not belong to the basis. Given that

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

the $\mathbf{Q}_j^\circ$ will be pairwise orthogonal orthogonal projection matrices, OPOPM. Reordering, if necessary, assume that the $\mathbf{Q}_1^\circ, \ldots, \mathbf{Q}_z^\circ$ are the non null product matrices, $z \leq m$.

Now

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

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

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

with $\boxplus$ indicating orthogonal direct sum of subspaces. We now establish, see [Pereira et al. (2014)].

Lemma 9. *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

$$\begin{aligned} rank(\mathbf{A}_j^T\mathbf{P}_j\mathbf{A}_j) &\le rank(\mathbf{A}_j^T\mathbf{P}_j) \le rank(\mathbf{P}_j) = \\ &= rank(\mathbf{A}_j\mathbf{A}_j^T\mathbf{P}_j\mathbf{A}_j\mathbf{A}_j^T) \le rank(\mathbf{A}_j^T\mathbf{P}_j\mathbf{A}_j),\ j = \{1, \dots, m\}, \end{aligned} \tag{5.1}$$

which establishes the thesis. □

Since the $\mathbf{Q}_1, \dots, \mathbf{Q}_m$ add up to $\mathbf{I}_n$, we can also establish, see [Pereira et al. (2014)].

Lemma 10. *In models with COBS, we have*

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

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

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

Now

$$\begin{aligned} R(\mathbf{Q}_j^{\circ}) &= R(\mathbf{Q}_j\mathbf{T}) = \mathbf{Q}_jR(\mathbf{T}) = \mathbf{A}_j^T\mathbf{A}_jR(\mathbf{X}_{\circ}) = \\ &= \mathbf{A}_j^TR(\mathbf{A}_j\mathbf{X}_{\circ}) = \mathbf{A}_j^TR(\mathbf{X}_{\circ j}) = \\ &= \mathbf{A}_j^TR(\mathbf{P}_j) = R(\mathbf{A}_j^T\mathbf{P}_j) = \\ &= R((\mathbf{A}_j^T\mathbf{P}_j)(\mathbf{A}_j^T\mathbf{P}_j)^T) = \\ &= R(\mathbf{Q}_j^{\circ\circ}),\ j = \{1, \dots, z\}. \end{aligned} \tag{5.2}$$

□

The following corollaries and the respective proofs are in [Pereira et al. (2014)].

Corollary 1. *In models with COBS,*

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

Proof: $\mathbf{T} = \sum_{j=1}^{z} \mathbf{Q}_j \mathbf{T} = \sum_{j=1}^{z} \mathbf{Q}_j^{\circ} = \sum_{j=1}^{z} \mathbf{Q}_j^{\circ\circ}$ □

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

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

Proof: The thesis follow from lemma 9 and corollary 1 of lemma 10.

5.3 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, see [Pereira et al. (2014)].

Proposition 27. *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 1 of lemma 10. □

We recall that models with COBS have LSE that are UBLUE. Thus, proposition 27, 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

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

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

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

as well as

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

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

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

we have

$$\begin{aligned} ||\mathbf{Y}_j - \boldsymbol{\mu}_j||^2 &= (\mathbf{Y}_j - \boldsymbol{\mu}_j)^T \mathbf{L}_j^T \mathbf{L}_j (\mathbf{Y}_j - \boldsymbol{\mu}_j) + (\mathbf{Y}_j - \boldsymbol{\mu}_j)^{\perp} \mathbf{P}_j^{\perp} (\mathbf{Y}_j - \boldsymbol{\mu}_j) = \\ &= \mathbf{Y}_j^T \mathbf{P}_j \mathbf{Y}_j - 2\boldsymbol{\eta}_j^T \tilde{\boldsymbol{\eta}}_j + \boldsymbol{\eta}_j^T \boldsymbol{\eta}_j + \mathbf{Y}_j^T \mathbf{P}_j^{\perp} \mathbf{Y}_j = \\ &= s_j - 2\boldsymbol{\eta}_j^T \tilde{\boldsymbol{\eta}}_j + ||\boldsymbol{\eta}_j||^2,\ j = \{1, \ldots, z\} \end{aligned} \tag{5.3}$$

and so the corresponding densities will be

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

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

$$n(\mathbf{y}^\circ) = \frac{e^{-\dfrac{1}{\boldsymbol{\xi}} \sum\limits_{j=1}^{m} \vartheta_j s_j + \sum\limits_{j=1}^{z} \boldsymbol{\xi}_j^T \tilde{\boldsymbol{\eta}}_j - \dfrac{1}{2} \sum\limits_{j=1}^{z} \vartheta_j ||\boldsymbol{\xi}||^2}}{\prod\limits_{j=1}^{m} (2\pi\vartheta_j^{-1})^{\dfrac{g_j}{2}}}.$$

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 10

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

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 & Casela (1998)], the statistics $s_j,\ j = \{1, \dots, m\}$, and $\tilde{\boldsymbol{\eta}}_j,\ j = \{1, \dots, z\}$, are not only sufficient but complete.

Since

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

the expression for $\tilde{\Psi}$ given in proposition 27 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

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

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, [Pereira et al. (2014)].

Proposition 28. *When normality holds, the* $\tilde{\boldsymbol{\Psi}}$ *and* $\tilde{\boldsymbol{\gamma}}_j$*, 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 following corollaries, see [Pereira et al. (2014)].

Corollary 3. *If* $\mathcal{D} = \{1, \dots, m\}$*, 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 28. □

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 & Carvalho (2013)], [Fonseca et al. (2008)], [Carvalho et al. (2008), (2013)] and [Mexia et al. (2010)].

5.4 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

$$\boldsymbol{\Psi}^* = \mathbf{A}\mathbf{Y}$$

with

$$\mathbf{A} \in [\boldsymbol{\Psi}] = \{\mathbf{A} : E(\mathbf{A}\mathbf{Y}) = \boldsymbol{\Psi}\}$$

where $E(.)$ is the mean vector. We now establish the following lemma, see [Pereira et al. (2014)].

Lemma 11. *$E(\boldsymbol{A}_1\boldsymbol{Y}) = E(\boldsymbol{A}_2\boldsymbol{Y})$ if and only if $\boldsymbol{A}_1\boldsymbol{T} = \boldsymbol{A}_2\boldsymbol{T}$*

Proof: Since

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

the sufficient condition is established. Inversely, if

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

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 $\square$

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

$$\tilde{\boldsymbol{\mu}} = \mathbf{TY}.$$

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

$$E(\tilde{\boldsymbol{\Psi}}) = \mathbf{A}(\boldsymbol{\Psi})\boldsymbol{\mu} = \mathbf{UTX}_\circ\boldsymbol{\beta}_\circ = \mathbf{U}\boldsymbol{\mu} = \boldsymbol{\Psi}$$

and that, according to lemma 11, $\mathbf{A} \in [\boldsymbol{\Psi}]$ if and only if

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

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

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

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

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

Thus,

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

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

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

if and only if

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

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

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

for every estimable vector, we must have

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

which gives

$$\begin{aligned}\mathbf{T}V(\boldsymbol{\theta})(\mathbf{I}_n - \mathbf{T}) = 0 \Leftrightarrow \mathbf{T}V(\boldsymbol{\theta}) - \mathbf{T}V(\boldsymbol{\theta})\mathbf{T} = 0 \Leftrightarrow \\ \Leftrightarrow \mathbf{T}V(\boldsymbol{\theta}) = \mathbf{T}V(\boldsymbol{\theta})\mathbf{T}\end{aligned} \tag{5.4}$$

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, see [Pereira et al. (2014)].

Theorem 20. *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

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

Definition 42. ***(Generalized Orthogonal Block Structure, GOBS)***

We have the mixed model, see expression 3.15 in section 3.5

$$\boldsymbol{Y} = \sum_{i=1}^{w} \boldsymbol{X}_i \boldsymbol{\beta}_i$$

then $\boldsymbol{Y}$ has a variance covariance matrix

$$\boldsymbol{V}(\gamma) = \sum_{j=1}^{m} \gamma_j \boldsymbol{Q}_j$$

where the $\{\boldsymbol{Q}_1, \dots, \boldsymbol{Q}_m\}$ are OPOPM. Let $\boldsymbol{Q} = \{\boldsymbol{Q}_1, \dots, \boldsymbol{Q}_m\}$ be the principal basis of an CJA $\mathcal{A}(\boldsymbol{M})$, we have

$$\boldsymbol{M}_i = \sum_{j=1}^{m} \boldsymbol{b}_{ij} \boldsymbol{Q}_j; \; i = 1, \dots, w$$

In section 3.6, we saw that if the matrices $\{\boldsymbol{M}_1, \ldots, \boldsymbol{M}_m\}$ commute, they generate a CJA, $\mathcal{A}(\boldsymbol{M})$, but $\boldsymbol{M}$ is not necessarily a basis of CJA, $\mathcal{A}$. Now, with $\boldsymbol{B} = [b_{ij}]$, we have

$$\boldsymbol{M}_i = \sum_{j=1}^{m} \boldsymbol{b}_{ij}\boldsymbol{Q}_j;\ i = 1, \ldots, w$$

as well as

$$\boldsymbol{\gamma} = \boldsymbol{B}^T\boldsymbol{\theta}\ \in\ R(\boldsymbol{B}^T)_+.$$

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 20, 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 20 continues to hold whatever the dimension of ∇.

Chapter 6

Final Remarks

We studied mixed models
$$\mathbf{Y} = \sum_{i=0}^{w} \mathbf{X}_i \boldsymbol{\beta}_i$$
where $\boldsymbol{\beta}_\circ$ was fixed and the $\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_w$ were independent with null mean vectors and variance covariance matrices, such that the matrices

$$\mathbf{M}_i = \mathbf{X}_i \mathbf{X}_i^T, \ i = 1, \dots, w$$

commutes. Then $\mathbf{Y}$ has a variance covariance matrix
$$V(\boldsymbol{\gamma}) = \sum_{j=1}^{m} \gamma_j \mathbf{Q}_j$$
where the $\mathbf{Q}_1, \dots, \mathbf{Q}_m$ are OPOPM, moreover if

$$R([\mathbf{X}_1, \dots, \mathbf{X}_w]) = \mathcal{R}^n$$

we must have
$$\sum_{j=1}^{m} \mathbf{Q}_j = \mathbf{I}_n$$

These models are said to have OBS. Besides this if $\mathbf{T}$, the OPM on the space Ω spanned by the mean vector, commutes with the $\mathbf{M}_1, \dots, \mathbf{M}_w$, it commutes with $V(\gamma)$ and the LSE of estimable vectors are UBLUE (see [Zmyślony (1978)]). We showed that for models with OBS to have LSE for estimable vectors that are UBLUE it is necessary and sufficient that $\mathbf{T}$ commutes with the $\mathbf{Q}_1, \dots, \mathbf{Q}_m$.

In the study of the orthogonal models, we consider the crossing and the nesting of such models which are building blocks for more complex orthogonal models.

We thus contributed for the study of an important class of models, considering in more detail the case of normal models. namely we showed that for having complete and sufficient statistics we must have w=m and k=p. Lastly, we showed that for models with COBS the condition p=k holds. These two results are very important.

Lastly, see [Pereira et al. (2014)], the theorem 20 uses the algebraic structure of the orthogonal models to characterize those with LSE that are UBLUE.

These results studied and established in this thesis will be useful in the study of structured families of orthogonal models. In these families each model is associated to a treatment of a base design. These designs have usually fixed effects factors and we may study the action of these factors on the estimable vectors.

Bibliography

[Areia & Carvalho (2013)] Areia, A; Carvalho, F. (2013). Perfect Families: an appliacation to Orthogonal and Error Orthogonal Models. *AIP Conf. Proc. Numerical Analysis and Applied Mathematics: International Conferences of Numerical Analysis and pplied Mathematics,* 1558, 841-846.

[Calinski & Kageyama (2000), (2003)] Calinski, T.; Kageyama, S. (2000). Block Designs: A Randomization Approach. Vol. I: Analysis. *Lecture Notes in Statistics, Springer*, 150

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

[Carvalho et al. (2008), (2013)] Carvalho, F.; Mexia, J. T.; Oliveira, M. M. (2008). Canonic inference and commutative orthogonal block structure. *Discussiones Mathematicae* **28**(2), 171-181.

[Carvalho et al. (2013)] Carvalho, Francisco; Mexia, JoÃ£o T.; Santos, Carla (2013). commutative orthogonal block structure and error orthogonal models. *Electronic Journal of Linear Algebra* **25**, 119-128.

[Covas et al. (2010)] Covas, R.; Mexia, J. T.; Zmyślony, R. (2010). Lattices of Jordan Algebras. *Linear Algebra and its Applications*, 2679-2690.

[Ferreira S. (2006)] Ferreira, S. (2006). Inferência para modelos ortogonais com segregação. Dissertação de Doutoramento em Matemática Estatística. Univesidade da Beira Interior.

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

[Fonseca et al. (2003)] Fonseca, M; Mexia, J. T.; Zmyślony, R. (2003). Estimators and tests for variance components in cross nested orthogonal designs. *Discussiones Mathematicae: Prob. Stat.*, **23**(2), 175-201

[Fonseca et al. (2006)] Fonseca , M; Mexia, J.T.; Zmyślony, R. (2006). Binary Operation on Jordan algebras and orthogonal normal models. *Linear Algebra and Its Applications*, **417**, 75-86.

[Fonseca et al. (2007)] Fonseca, M; Mexia, J.T.; Zmyślony, R. (2007). Jordan algebras. Generating pivot variables and orthogonal normal models. *Journal of Interdisciplinary Mathematics*, **1**, 305-326. [Fon

[Fonseca (2007)] Fonseca, M. (2007) Estimation and Hypothesis Testing in Mixed Linear Models. *Dissertação de Doutoramento em Matemática, especialização em Estatística. FCT, UNL.*

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

[Fonseca et al. (2009)] Fonseca, M; Jesus, V.; Mexia, J.T.; Zmyślony, R. (2009). Binary operations and canonical forms for factorial and related models. *Linear Algebra and its Applications*, **430**, 2781-2797.

[Graham (1981)] Graham, A. (1981). Kronecker Product and Matrix Calculus with Applications. *Ellis Horwood Ltd.*

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

[Jesus et al. (2009)] Jesus, V. M.; Mexia, J. T.; Fonseca, M.; Zmyślony, R. (2009). Binary Operations and Cannonical Forms for Factorial and Related Models. *Linear Algebra and its Applications*, **430**, 2781-2797.

[Jordan et al. (1934)] Jordan, P.; Von Newman, J.; Wigner, E. (1934). On the algebraic generalization of the quantum mechanical formalism. *Ann. of Math.*, **36**, 26-64.

[Kariya & Kurata (2004)] Kariya, T.; Kurata, H. (2004). Generalized Least Squares. *John Wiley and Sons, Ltd.*

[Kendall (2004)] Kendall, M. G. (2004). A course in the Geometry of n Dimensions. *Dover Publications.*

[Khury & Gosh (1990)] Khury, A. I.; Gosh, M. (1990). Minimal Sufficient Statistics for the Unbiased Two-Fold Nested Design. *Stat. Prob Letters*, **10**, 351-353.

[Khury et al. (1997)] Khury, A. L.; Mathew, T.; Sinha, B. K. (1997). Statistical Tests for Mixed Linear Models, *John Wiley and Sons, Ltd*

[Lehmann et al. (1997)] Lehmann, E.L. (1997); Testing statistical hypotheses. *Reprint of the 2nd ed. publ. by Wiley 1986. New York, NY: Springer.*

[Lehmann & Casela (1998)] Lehmann, E. L.; Casela, G. (1998). Theory of Point estimation. *2nd ed. Springer Tests Statistics. New York, NY: Springer.*

[Malley (1994)] Malley, J. D. (1994). Statistical Applications of Jordan Algebras. *Lecture Notes in Statistics*, **91**, springer-Verlag, 1994.

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

[Mexia (1995)] Mexia, J. T. (1995). Introdução à Inferência Estatística. *Edições Lusøfonas.*

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

[Mexia et al. (2011)] Mexia, J. T.; Carvalho, F.; Covas, R.; (...) (2011). Error Orthogonal Models: Structure, Binary Operations, Extensions and Normality.

[Montgomery (1997)] Montgomery, D. C. (1997). Design and Analysis of Experiments - 5th Edition. *John Willey and Sons. New York.*

[Moore (1920), (1935)] Moore E.H. (1920). On the reciprocal of the general algebraic matrix (abstract). *Bulletin of the American Mathematical Society.* **26**, 394-395.

[Moore (1935)] Moore E.H. (1935). General analysis. *Memoirs of the American Philosophical Society.* **1**, 147-209.

[Moreira (2008)] Moreira, E. (2008). Famílias Estruturadas de Modelos com Modelo Base Ortogonal. Teoria e Aplicações. *Dissertação de Doutoramento em Matemática, especialização em Estatística. FCT, UNL.*

[Nelder (1965a)] Nelder, J. A. (1965a). 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.

[Nelder (1965b)] 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), 163-178.

[Penrose (1955)] Penrose R. (1955). A generalized inverse for matrices. *Procedings of the Cambridge Philosophical Society.* **51**, 406-413.

[Pereira et al. (2014)] . Orthogonal Models: Algebraic Structure and Explicit Estimators for Estimable Vectors. *DM-PS.*

[Pollock (1973)] Pollock, D. S. G. (1973). The Algebra of Econometric . *Wiley Series in Probability and Mathematical Statistics.* Wiley, New York.

[Rao & Rao (1998)] Rao, C. R.; Rao, M. B. (1998). Matrix Algebra and its Applications to Statistics and Econometrics, *World Scientific Publishing Company*, Singapore.

[Santos C. (2012)] 2012 Santos, E. (2012). Error Orthogonal Models: Structure, Operations and Inference. *Dissertação de Doutoramento em Matemática, especialização em Estatística. UBI.*

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

[Seber (1980)] Seber, G.A.F. (1980). The Linear Hypothesis: *A General Theory 2nd edition, Griffin.*

[Seely (1970a)] Seely, J. (1970a). Linear spaces and unbiased estimators. *The Annals of Mathematical Statistics*, **41**, 1735-1745.

[Seely (1970b)] Seely, J. (1970b). Linear spaces and unbiased estimators. pplication to a mixed linear model *The Annals of Mathematical Statistics*, **41**, 1735-1745.

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

[Seely et al. (1971)] Seely, J.; Zyskind (1971). Linear Spaces and Minimum Variance estimators. *The Annals of Mathematical Statistics*, **42**, 691-703.

[Seely (1977)] Seely, J. (1977). Minimal Sufficient Statistic and Completeness for Multivariate Normal Families. *Sankhya: The Indian Journal of Statistics, Series A*, **39** (2), 170-185.

[Silvey (1977)] Silvey, S. D. (1977). Statistical Inference. *Chapman and Hall / CRC Monographs on Statistics and Applied Probability.*

[Steeb (1991)] Steeb, W. H. (1991). Kronecker products of Matrices and Applications. *Bibliographisces Institut, Mannhein, Germany.*

[Vanleeuwen et al. (1998)] Vanleeuwen, Dawn M.; Birks, David s.; Seely, Justus F.; J. Mills, J.; Greenwood, j. A.; Jones, C. W. (1998). Sufficient Conditions for Orthogonal Designs in mixed Linear Models. *Journal of statistics Planning and Inference*, **73**, 373-389.

[Vanleeuwen et al. (1999)] Vanleeuwen, Dawn M.; Birks, David s.; Seely, Justus F. (1999). Balances and Orthogonality in Designs for Mixed Classification Models. *Journal of statistics Planning and Inference*, **27**, 1927-1947.

[Williams (1991)] Williams, D. (1991). Probability with martingales. *Cambridge etc.: Cambridge University Press.*

[Zmyślony (1978)] Zmyślony, R. (1978). A Characterization of the Best Linear Unbiased Estimators in the General Linear Model. *Lecture Notes Statistics*, **2**, 365-373.