



# MODELOS E FAMÍLIAS DE MODELOS PARA MATRIZES ESTOCÁSTICAS SIMÉTRICAS

*Cristina Paula da Silva Dias*

Tese apresentada à Universidade de Évora  
para obtenção do Grau de Doutor em Matemática  
Especialidade: Estatística

ORIENTADORES: *Prof.<sup>a</sup> Maria Manuela Melo Oliveira*  
*Prof. João Tiago Praça Nunes Mexia*

ÉVORA, JUNHO DE 2013





**Universidade de Évora**

**Modelos e Famílias de Modelos para Matrizes  
Estocásticas Simétricas**

Dissertação apresentada para obtenção do grau de Doutor em Matemática na  
especialidade de Estatística, pela universidade de Évora.

Cristina Paula da Silva Dias

Orientador: Maria Manuela Melo Oliveira

Co-orientador: João Tiago Praça Nunes Mexia

Évora

2013



À memória do meu Pai.

Ao meu filho João.

Ao Miguel

# Agradecimentos

Expresso o meu agradecimento e reconhecimento a todos os que contribuíram e ajudaram na realização deste trabalho:

Primeiramente a Deus, por ser a minha fonte de inspiração e força nos momentos não raros em que, a incerteza caminha lado a lado com as escolhas dos caminhos a seguir.

À minha orientadora Professora Doutora Maria Manuela Oliveira, pelo incentivo e encorajamento ao longo da realização deste trabalho, principalmente na etapa final desta dissertação em que o seu empenho e conselhos foram extremamente valiosos.

Ao Professor Doutor João Tiago Praça Nunes Mexia, pela disponibilidade e encorajamento prestados ao longo destes quatro anos. O ser tão compreensivo com as minhas dificuldades, os seus enormes conhecimentos científicos e a sua bondade, contribuíram para o meu crescimento científico e enriquecimento pessoal.

À minha mãe e irmã e cunhado pelo apoio incondicional, pelo carinho, pela disponibilidade e presença. Sem eles não seria possível percorrer o longo caminho que levou à realização do trabalho apresentado nesta dissertação.

Ao meu filho João, pelas brincadeiras adiadas, pelas descobertas diárias que nesta idade precisam ser partilhadas e, nem sempre o foram comigo, pelos muitos dias e noites, sem a mãe, para ele, um sem fim do meu amor.

Ao Miguel por acreditar em mim, sempre. Por me mostrar um mundo diferente, repleto de rigor e respeito científico, por me inculir um gosto e uma visão alargada pela ciência e sua importância para ajudar na construção de um mundo melhor.

Aos amigos e colegas, principalmente à minha amiga Carla Dias, pelo apoio dado.

Ao meu Pai, que embora já não esteja entre nós, foi o meu maior amigo e protetor.

# Resumo

Nesta dissertação desenvolvem-se modelos da forma  $\mathbf{M} = \boldsymbol{\mu} + \bar{\mathbf{E}}$  de grau  $k$  para matrizes estocásticas simétricas, sendo  $\boldsymbol{\mu}$  a matriz média e  $\bar{\mathbf{E}}$  uma matriz estocástica simétrica, com matriz média nula. Os modelos são desenvolvidos a partir da análise espectral das respectivas matrizes médias  $\boldsymbol{\mu}$ . O ajustamento e validação do modelo requerem a utilização do vector  $\tilde{\boldsymbol{\beta}}_i = \theta_i \boldsymbol{\gamma}_i, i = 1, \dots, k$ , como estimador do vector de estrutura  $\boldsymbol{\beta}_i = \lambda_i \boldsymbol{\alpha}_i, i = 1, \dots, k$  de  $\mathbf{M}$ , com  $\lambda_1, \dots, \lambda_k$  os valores próprios não nulos e  $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_k$  os respetivos vetores próprios. A informação contida na matriz  $\mathbf{M}$  pode ser condensada no par  $(\tilde{\boldsymbol{\beta}}, V)$ , onde  $V$  é a soma de quadrados de resíduos. Os resultados obtidos permitem realizar inferência transversal e longitudinal. Na primeira trabalha-se com as componentes homólogas do vector de estrutura, e na segunda trabalha-se com vetores de contrastes nas componentes desse vector.

O grau do modelo é dado pela característica da matriz média, pelo que o estudo do modelo envolve todos os valores próprios. Para modelos de grau  $k > 1$  considera-se ainda a possibilidade de truncar o modelo quando se têm valores próprios  $\theta_1, \dots, \theta_k$  muito superiores aos restantes

Uma consequência direta da metodologia adotada é a aplicação dos modelos de grau um a matrizes de produtos cruzados e matrizes de produtos escalares de Hilbert-Schmidt. Estas últimas desempenham um papel importante na primeira fase, a inter-estrutura, da metodologia STATIS, enquanto que as matrizes de produtos cruzados destacam-se (em particular as matrizes  $\mathbf{A}\mathbf{A}^t$  e  $\mathbf{A}^t\mathbf{A}$ , que têm os mesmos valores próprios não nulos) pelo seu papel na inferência.

Para além destes modelos, também se consideraram os modelos de famílias estruturadas. Os modelos destas famílias estão associados aos tratamentos de um delineamento base. A ação dos fatores que se consideram no delineamento base, sobre

os vetores de estrutura é também analisada. Nas famílias estruturadas com delineamento base ortogonal, os delineamentos estão associados a partições

$$R^m = \bigoplus_{j=1}^p \varpi_j$$

onde  $\bigoplus$  indica a soma direta de subespaços e as hipóteses formuladas estão associadas aos espaços destas partições. Concluímos com uma aplicação das famílias estruturadas de matrizes de produtos cruzados a dados de melhoramento vegetal, mostrando-se que, existem influências significativas quer do fator ano quer do fator local, relativamente ao comportamento dos cultivares.



# Abstract

## Models and Families of Models for Symmetric Stochastic Matrices

In this work we develop  $k$ -degree models of the form  $\mathbf{M} = \boldsymbol{\mu} + \bar{\mathbf{E}}$  for symmetric stochastic matrices  $\mathbf{M}$ , with mean matrix  $\boldsymbol{\mu}$ , and  $\bar{\mathbf{E}}$  is a symmetric stochastic matrix with null mean. The models are developed using the spectral analysis of the matrices  $\boldsymbol{\mu}$ . The adjustment and validation of the model requires the usage of the vector  $\tilde{\boldsymbol{\beta}}_i = \theta_i \boldsymbol{\gamma}_i, i = 1, \dots, k$ , which is an estimator of the structure vector  $\boldsymbol{\beta}_i = \lambda_i \boldsymbol{\alpha}_i, i = 1, \dots, k$  of  $\mathbf{M}$ . With  $\lambda_1, \dots, \lambda_k$  the non-null eigenvalues and  $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_k$  are the eigenvectors, of the matrix  $\mathbf{M}$ . The information enclosed in matrix  $\mathbf{M}$  can be cast into a pair  $(\tilde{\boldsymbol{\beta}}, V)$ , where  $V$  is the sum of the squared residuals. With the model results one can carry out transversal and longitudinal inferences. In the latter we work with vectors contrasts in the components of the structure vector, while in the former we work with the homologous components of that vector.

As the degree ( $k$ ) of the model is given by the mean matrix characteristic, and thus the model study involves all  $k$  eigenvalues. For the models with a degree  $k > 1$ , we still consider the possibility of truncating the model, when there are eigenvalues,  $\theta_1, \dots, \theta_k$  much greater than the other ones.

A direct consequence of our methodology is that degree-one models can be applied to cross products matrices and Hilbert-Schmidt scalar products matrices. The latter have

an important role in the first stage (inter-structure) of STATIS methodology, while the former matrices (in particular the  $\mathbf{AA}^t$  and  $\mathbf{A}^t\mathbf{A}$  cross products, which have the same non-null eigenvalues) have an important role in inference.

A further development includes structured families models, which are associated to the treatments in a base design. The action of the factors in the base design on the structured vectors of the models in the family is studied. In Structured families with orthogonal base designs, the designs are associated to orthogonal partitions

$$R^m = \bigoplus_{j=1}^p \varpi_j,$$

where  $m$  is the number of treatments and  $\bigoplus$  indicates orthogonal direct sum of subspaces. The hypotheses for these designs are associated to the spaces in this partition. An application of ours results to experiments of plant breeding is made, showing that there are significant influences of year and local factor, concerning the behavior of the cultivars.

# Simbologia e Notações

$\pi$	modelo de efeitos fixos
$\mathbf{Y}$	vetor aleatório (letra maiúscula a negrito)
$\mathbf{v}$	vetor (letra minúscula a negrito)
$\ \cdot\ $	norma euclidiana de um vetor ou de Hilbert-Schmidt para matrizes
$\in$	pertence
$R$	conjunto dos números reais
$\mathbf{A}^t$	transposta de $\mathbf{A}$
$\mathbf{A}^{-1}$	matriz inversa de $\mathbf{A}$
$\mathbf{A} = [a_{ij}]$	matriz com elemento genérico $a_{ij}$
$\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_r]$	matriz com vetores coluna $\mathbf{a}_1, \dots, \mathbf{a}_r$
$\approx$	aproximadamente igual
$\mathbf{D}(\xi)$	matriz diagonal cujos elementos principais são as componentes de $\xi$
$\mathbf{A}^+$	inversa de Moore-Penrose da matriz $\mathbf{A}$
$\mathbf{I}_p$	matriz identidade de ordem $p$
$\text{vec}(\mathbf{B})$	vetor obtido por agrupamento dos vetores coluna da matriz $\mathbf{B}$
$\boldsymbol{\gamma}^k$	vetor com $k$ coordenadas

$\otimes$	produto de Kronecker
$E(\cdot)$	valor médio
$\text{var}(\cdot)$	variância
$\text{cov}(\cdot, \cdot)$	covariância
$\sim$	distribui-se como
$\Sigma(\cdot)$	matriz de covariância
$P(\lambda)$	distribuição de Poisson de parâmetro $\lambda$
$N(\mu, \sigma^2)$	distribuição normal de valor médio $\mu$ e variância $\sigma^2$
$N(\mu, \Sigma)$	distribuição multiforme de vetor médio $\mu$ e matriz de covariância $\Sigma$
$\chi_k^2$	distribuição do qui-quadrado central com $k$ graus de liberdade
$\chi_{k, \delta}^2$	distribuição do qui-quadrado com $k$ graus de liberdade e parâmetro de não centralidade $\delta$
$F(k, g, \delta)$	distribuição F com $k$ e $g$ graus de liberdade e parâmetro de não centralidade $\delta$
i.i.d.	independentes e identicamente distribuídos
$\text{car}(\cdot)$	característica de uma matriz
$\det(\mathbf{A})$	determinante da matriz $\mathbf{A}$
$\mathbf{D}(r_1, \dots, r_k)$	matriz diagonal cujos elementos principais são as componentes de $r_1, \dots, r_k$
$\mathbf{M}^+$	inversa de Moore-Penrose da matriz $\mathbf{A}$
$\mathbf{P}_r$	matriz ortogonal padronizada $r \times r$
$\mathbf{O}_{n,m}$	matriz de ordem $n \times m$ com elementos todos nulos
$\mathbf{O}^k$	vetor com componentes nulas

$\mathbf{1}^k$	vetor cujas componentes são unitárias
$\nabla^\perp$	complemento ortogonal
$\nabla$	subespaço
f.g.m.	função geradora de momentos
$R(\mathbf{A})$	espaço imagem da matriz $\mathbf{A}$
$\xrightarrow[n \rightarrow \infty]{F}$	converge em distribuição para, quando $n$ tende para $+\infty$
$\xrightarrow[n \rightarrow \infty]{p}$	convergência estocástica para, quando $n$ tende para $+\infty$
$\xrightarrow[n \rightarrow \infty]{q.c.}$	converge quase certamente para, quando $n$ tende para $+\infty$
$\oplus$	soma direta ortogonal de subespaços
$ \cdot $	função módulo
$tr(\cdot)$	traço de uma matriz

# Índice

<b>Agradecimentos</b>	<b>ii</b>
<b>Resumo</b>	<b>iii</b>
<b>Abstract</b>	<b>v</b>
<b>Simbologia e Notações</b>	<b>vii</b>
<b>1 Introdução</b>	<b>1</b>
<b>2 Preliminares</b>	<b>6</b>
2.1 Produto de kronecker .....	6
2.2 Matrizes Conjugadas.....	8
2.3 Inversas de Moore-Penrose.....	10
2.4 Modelos Base Ortogonais.....	12
2.4.1 Partições e Inferência .....	12
2.4.2 Cruzamento Equilibrado de Fatores.....	15
2.4.3 Heterocedasticidade Controlada.....	17
2.4.4 Limites Estocásticos e Distribuições Limite .....	18
<b>3 Modelos e Famílias de Modelos para Matrizes Estocásticas Simétricas</b>	<b>24</b>
3.1 Introdução.....	24
3.2 Vetores de Estrutura e Resíduos.....	25
3.3 Grau do Modelo.....	28
3.4 Estimação do Modelo.....	31
3.4.1 Estimação Pontual.....	31
3.4.2 Regiões de Confiança.....	37
3.5 Truncatura do Modelo.....	45

3.6 Relevâncias Relativas.....	49
3.7 Aplicação à Inter-estrutura.....	54
3.8 Aplicação a Matrizes de Produtos Cruzados .....	57
3.9 Famílias Estruturadas.....	58
3.9.1 Caso Geral.....	58
3.9.2 Modelos de Grau um.....	60
3.9.3 Estudo de Simulação.....	63
<b>4. Comparação de Cultivares de Trigo Rijo</b>	<b>65</b>
4.1 Introdução.....	65
4.2 Dados e Objetivos.....	66
4.3 Metodologia.....	69
4.4 Resultados.....	70
<b>5. Considerações Finais</b>	<b>77</b>
<b>A Linearidade Assintótica e Normalidade</b>	<b>79</b>
A.1 Função Assintótica Linear.....	79
A.2 Resultados Assintóticos.....	80
A.3 Aplicação ao Caso Normal.....	83
A.4 Simulações.....	85
<b>B Rotinas Utilizadas</b>	<b>96</b>
B.1 Estimativas do Teste de Shapiro-Wilk para $P_1$ .....	96
B.2 Estimativas do Teste de Kolmogorov-Smirnov para $P_1$ .....	97
B.3 Estimativas do Teste de Kolmogorov-Smirnov para $P_2$ .....	98
B.4 Número de Termos.....	100
<b>Bibliografia</b>	<b>102</b>

# Lista de Tabelas

3.1	Número de termos.....	40
3.2	Valores médios e desvios-padrão para $\mathbf{Z}$ .....	64
4.1	Primeiros valores próprios ajustados e valores da preponderância.....	70
4.2	Primeiros valores próprios ajustados e valores da preponderância.....	70
4.3	Vetores de estrutura ajustados e soma de quadrados de resíduos.....	71
4.4	Vetores de estrutura ajustados e soma de quadrados de resíduos.....	71
4.5	Somas de quadrados.....	72
4.6	Quadros resumo das análises de variância.....	73
4.7	Quadros resumo das análises de variância.....	74
A.1	Estimativas do teste de Kolmogorov-Smirnov para $P_1$ .....	85
A.2	Estimativas do teste de Shapiro-Wilk para $P_1$ .....	86
A.3	p-value do teste Kolmogorov-Smirnov para $P_1$ .....	86
A.4	Estimativas do teste de Kolmogorov-Smirnov para $P_2$ .....	87
A.5	Estimativas do teste de Kolmogorov-Smirnov para $P_2$ .....	88
A.6	Estimativas do teste de Kolmogorov-Smirnov para $P_2$ .....	89
A.7	Estimativas do teste de Kolmogorov-Smirnov para $P_2$ .....	90
A.8	Estimativas do teste Kolmogorov-Smirnov para $P_2$ .....	91
A.9	p-value do teste Kolmogorov-Smirnov para $P_2$ .....	92
A.10	p-value do teste de Kolmogorov-Smirnov para $P_2$ .....	93
A.11	p-value do teste de Kolmogorov-Smirnov para $P_2$ .....	94
A.12	p-value do teste de Kolmogorov-Smirnov para $P_2$ .....	95



# Lista de Figuras

4.1 Mapa das zonas equipotenciais de adaptação .....	68
4.2 Importâncias relativas para o ano das nove cultivares.....	75
4.3 Importâncias relativa para o local das nove cultivares.....	76
4.4 Importâncias relativas e interação das nove cultivares.....	76

# Capítulo 1

## Introdução

Esta dissertação insere-se numa linha de trabalho que se inicia com o estudo de famílias estruturadas dos delineamentos regressivos múltiplos, introduzidos por Mexia (1989). Uma família de modelos diz-se estruturada quando, os modelos que a constituem correspondem aos tratamentos de outro modelo, que se designa por modelo base. As matrizes do modelo das regressões podem diferir de tratamento para tratamento, isto é, um tratamento corresponde a uma combinação de níveis dos fatores envolvidos no delineamento base. Esta linha de desenvolvimento das famílias estruturadas de modelos leva-nos a incluir os trabalhos de investigação desenvolvidos por, Oliveira e Mexia (1998), (1999a), (1999b), agora com aplicação à metodologia STATIS (Structuration de Tableaux à Trois Indices de la Statistique), a qual foi introduzida por Escoufier (1973), sendo posteriormente desenvolvida por Lavit (1988). Nesta metodologia, um estudo é um trio de matrizes  $(\mathbf{X}, \mathbf{D}, \mathbf{D}^{\cdot})$  onde  $\mathbf{X}$  é uma matriz de dados em que as linhas correspondem a unidades estatísticas também designadas por objetos e, as colunas correspondem a variáveis.  $\mathbf{D}$  e  $\mathbf{D}^{\cdot}$  são matrizes diagonais de pesos, uma para objetos e outra para variáveis, respectivamente. A modelação de séries de estudos com aplicação das famílias estruturadas de modelos, agora no âmbito da metodologia STATIS, continuou a ser desenvolvida por Oliveira (2002), com o objetivo de realizar inferência para estas séries. Uma das conclusões resultantes deste estudo, revelou que nos modelos de grau um, o primeiro valor próprio da matriz  $\mathbf{S}$  predominava sobre os restantes, o que evidenciava que as séries de estudo tinham estrutura comum de grau um. O modelo desenvolvido por Oliveira (2002), era da forma  $\mathbf{S} = \lambda \mathbf{a} \mathbf{a}^t + \mathbf{E}$ , o qual mostrou ser bastante manejável, no entanto seria apenas aproximadamente simétrico. Posteriormente, Oliveira e Mexia (2004), obtiveram um modelo matemático mais

correto, tomando  $\mathbf{E} = \frac{1}{2}(\mathbf{E}^0 + \mathbf{E}^{0t})$ , tendo no entanto o inconveniente de se admitir a normalidade dos elementos da matriz  $\mathbf{S}$  e, não nos operadores e nem nas observações finais. Esta modelização foi aprofundada em Ramos (2007), onde esta dificuldade foi ultrapassada ao considerar a normalidade na componente aleatória das observações iniciais. Aqui as séries de estudos correspondem aos tratamentos dum delineamento fatorial de base prima. Em trabalhos posteriores Oliveira e Mexia (2007a e 2007b) e Areia (2009), conseguiram um modelo cuja formulação matemática, apesar de mais complexa, revelou ser mais potente que os modelos anteriores, aquando da realização da inferência. Foram efetuadas aplicações destes modelos a resultados de eleições autárquicas realizadas em Portugal continental entre 1985 e 2001, tendo sido possível com a aplicação destes modelos, estudar a evolução das mesmas no período temporal considerado. Outras aplicações desta teoria podem ser vistas em Moreira (2008), com aplicação das famílias estruturadas de modelos à hidrologia e à remediação ambiental. Na primeira destas aplicações, os modelos considerados eram log-lineares ajustados a tabelas de contingência e, na segunda aplicação, a remediação ambiental, os modelos eram regressões polinomiais. Os resultados obtidos nas aplicações de Moreira (2008), com a utilização das famílias estruturadas de modelos, revelaram ser de grande interesse na área da hidrologia e remediação ambiental, uma vez que podem ser utilizados como métodos alternativos a outros para detetar zonas homogéneas em termos de ocorrência de períodos de seca (severidade e frequência). Além disso os resultados obtidos, mostraram ser similares a outros, obtidos anteriormente, em trabalhos desenvolvidos por Moreira et al. (2005 e 2007). Assim, a utilização desta teoria ganha cada vez mais uma maior credibilidade e abre caminho a desenvolvimentos futuros. Recentemente em Cantarinha (2012) a teoria das famílias estruturadas foi aplicada ao estudo dos fogos florestais em Portugal Continental no período de 1983 a 2007. Aqui os modelos da família eram uma versão multivariada dos modelos coletivos das matemáticas atuariais. Os resultados obtidos permitem concluir que, todos os efeitos e interações, são altamente significativos, isto é, existe uma evolução global, quer do número de fogos quer da área ardida, no período considerado.

No trabalho desenvolvido nesta dissertação além de considerarmos a teoria das famílias estruturadas de modelos e sua aplicação, teremos também por objetivo o desenvolvimento de modelos para matrizes estocásticas simétricas. Estes modelos serão

desenvolvidos a partir da análise espectral das respectivas matrizes médias. Os modelos são da forma

$$\mathbf{M} = \boldsymbol{\mu} + \bar{\mathbf{E}} = \sum_{j=1}^k \lambda_j \boldsymbol{\alpha}_j \boldsymbol{\alpha}_j^t + \bar{\mathbf{E}},$$

onde  $(\lambda_i, \boldsymbol{\alpha}_i), i=1, \dots, k$  são pares de valores próprios e vetores próprios associados a esses valores próprios.  $\bar{\mathbf{E}}$  é uma matriz estocástica simétrica, com matriz média nula. O grau do modelo será dado pela característica da matriz média, isto é  $k = \text{car}(\boldsymbol{\mu})$ . Uma parte importante deste trabalho será a escolha do grau do modelo. Assim, sendo  $(\theta_i, \boldsymbol{\gamma}_i), i=1, \dots, k$  os pares de valores próprios e vetores próprios para os  $k$  maiores valores próprios de  $\mathbf{M}$ , verifica-se que se  $k$  for o grau do modelo a informação contida na matriz  $\mathbf{M}$ , pode ser condensada num vetor de estrutura ajustado

$$\tilde{\boldsymbol{\beta}} = [\tilde{\boldsymbol{\beta}}_1^t, \dots, \tilde{\boldsymbol{\beta}}_k^t]^t,$$

com  $\tilde{\boldsymbol{\beta}}_i = \theta_i \boldsymbol{\gamma}_i, i=1, \dots, k$  e numa soma  $V$  de quadrados de resíduos. A validação do modelo permite-nos utilizar  $\tilde{\boldsymbol{\beta}}$  como estimador do vetor de estrutura de  $\mathbf{M}$  dado por

$$\boldsymbol{\beta} = [\boldsymbol{\beta}_1^t, \dots, \boldsymbol{\beta}_k^t]^t,$$

com  $\boldsymbol{\beta}_i = \lambda_i \boldsymbol{\gamma}_i, i=1, \dots, k$ .

Sempre que  $\theta_1 > \dots > \theta_h \gg \theta_{h+1} > \dots > \theta_k$ , isto é, quando  $\theta_h$  é muito maior que  $\theta_{h+1}$ , obtém-se um modelo truncado com vetor de estrutura ajustado dado por

$$\tilde{\boldsymbol{\beta}}^0 = [\tilde{\boldsymbol{\beta}}_1^t, \dots, \tilde{\boldsymbol{\beta}}_h^t]^t,$$

o qual permite simplificar os cálculos sem grande perda de precisão dos resultados.

Os modelos de grau um têm mostrado ser de grande interesse, em situações em que trabalhamos com matrizes de produtos cruzados. Em particular esses modelos mostraram ser aplicáveis a matrizes de produtos escalares de Hilbert-Schmidt, as quais são utilizadas na primeira etapa, a inter-estrutura, da metodologia STATIS. Observe-se que estas são também matrizes de produtos cruzados. Para além dos modelos isolados, também iremos considerar famílias estruturadas de modelos. Os modelos destas famílias estão associados aos tratamentos de um delineamento base. A ação dos fatores, que se consideram no delineamento base, sobre os vetores de estrutura das matrizes da família será analisada. Utilizamos para isso a ANOVA (Análise de Variância) e técnicas

relacionadas, no estudo da referida ação sobre combinações lineares de  $\mathbf{a}'\boldsymbol{\beta}_i$ ,  $i = 1, \dots, m$ , das componentes dos vetores de estrutura das  $m$  matrizes do modelo. Consideram-se dois tipos de análise:

- Análise Transversal (AT), onde  $\mathbf{a}$  tem uma única componente não nula a  $i$ -ésima que é igual a um. Assim, estuda-se a ação dos fatores do delineamento base sobre as  $i$ -ésimas componentes dos vetores de estrutura. Por exemplo na metodologia STATIS, os modelos estudados aplicam-se a séries de estudos. Em geral os modelos de grau um aplicam-se, neste caso da AT, à  $i$ -ésima componente a qual corresponde ao  $i$ -ésimo estudo da série, isto é, os estudos homólogos correspondem às componentes com o mesmo índice dos vetores de estrutura.

- Análise Longitudinal (AL), na qual  $\mathbf{a}$  tem soma nula para as suas componentes, tratando-se de um vetor de contrastes. Voltando à metodologia STATIS, os estudos em cada série podem corresponder a momentos sucessivos. Desta forma a AL, permite-nos estudar a evolução destas séries no tempo. Por exemplo, em Oliveira (2002), Ramos (2007) e Areia (2009) foi estudada a evolução dos resultados eleitorais nomeadamente, resultados de eleições legislativas e autárquicas realizadas em Portugal Continental, entre 1985 e 2001.

Verifica-se que os modelos de grau um, permitem um tratamento particularmente elegante podendo muitas vezes, ser utilizados para matrizes de produtos cruzados, quer isolados quer integrados em famílias estruturadas.

As famílias estruturadas que consideramos têm, delineamentos base ortogonais, associados a partições ortogonais de  $R^m$ . Com  $m$  o número de tratamentos, temos no caso duma dessas partições

$$R^m = \bigoplus_{j=1}^p \omega_j,$$

onde  $\bigoplus$  indica a soma direta ortogonal de subespaços. Quando este tipo de modelo é aplicado a um vetor aleatório com vetor médio  $\boldsymbol{\mu}$ , as hipóteses de ausência de efeitos e de interação podem ser escritas da seguinte forma:

$$H_{o_j} : \boldsymbol{\mu} = \omega_j = \omega_j^\perp, \quad j = 1, \dots, p,$$

onde  $\perp$  indica complemento ortogonal.

Neste estudo os modelos aplicam-se a matrizes aleatórias simétricas e, em particular a matrizes de produtos cruzados.

Esta dissertação está organizada em 5 Capítulos. Assim, começamos por apresentar alguns resultados preliminares necessários aos desenvolvimentos dos Capítulos seguintes. De seguida consideramos os modelos base ortogonais e, apresentamos esses modelos a partir de partições ortogonais.

O tratamento destes modelos será generalizado de forma a incluir interação de heterocedasticidade controlada em que se trabalha com vetores aleatórios, os quais, possuem matrizes de covariância  $\sigma^2\mathbf{C}$ , com as  $\mathbf{C}_i, i = 1, \dots, m$ , matrizes conhecidas e invertíveis. Segue-se a formulação dos modelos para matrizes estocásticas simétricas. Depois de desenvolvida uma teoria geral destes modelos, esta será aplicada ao caso da metodologia STATIS e a matrizes de produtos cruzados. Neste último caso consideramos ainda os modelos de grau um, que como referido antes conduzem a uma grande simplificação dos cálculos. Em seguida consideramos as famílias estruturadas de matrizes, primeiro no caso geral, depois o caso de famílias com grau um. Como veremos a aplicação dos resultados para modelos ortogonais às famílias estruturadas é imediata. A concluir apresentamos uma aplicação das famílias estruturadas de matrizes de produtos cruzados a dados de melhoramento vegetal, mais concretamente os dados dizem respeito a uma rede de ensaios de produção de trigo rijo. Finalizaremos a dissertação com algumas considerações finais.

# Capítulo 2

## Preliminares

### 2.1 Produto de Kronecker

Nesta secção e seguintes são apresentados alguns resultados que utilizaremos nos Capítulos precedentes.

Recorde-se que dada a matriz  $\mathbf{A} = [a_{ij}]$  do tipo  $r \times s$  o produto de Kronecker de  $\mathbf{A}$  pela matriz  $\mathbf{B}$  é dado por

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & \dots & a_{1s}\mathbf{B} \\ \vdots & & \vdots \\ a_{r1}\mathbf{B} & \dots & a_{rs}\mathbf{B} \end{bmatrix}.$$

Estando definidos os produtos usuais  $\mathbf{A}_1\mathbf{A}_2$  e  $\mathbf{B}_1\mathbf{B}_2$ , tem-se

$$(\mathbf{A}_1 \otimes \mathbf{B}_1)(\mathbf{A}_2 \otimes \mathbf{B}_2) = (\mathbf{A}_1\mathbf{A}_2) \otimes (\mathbf{B}_1\mathbf{B}_2), \quad (2.1)$$

e

$$(\mathbf{A} \otimes \mathbf{B})^t = \mathbf{A}^t \otimes \mathbf{B}^t.$$

O produto de Kronecker de matrizes ortogonais origina matrizes ortogonais. O produto de Kronecker de matrizes estandardizadas origina matrizes estandardizadas. Em particular

$$\mathbf{P} = \mathbf{P}_1 \otimes \mathbf{P}_2$$

será matriz normal estandardizada. Cada linha de  $\mathbf{P}$  é o produto de Kronecker duma linha de  $\mathbf{P}_1$  por uma linha de  $\mathbf{P}_2$ .

Consideremos agora as matrizes diagonais  $\mathbf{D}(\mathbf{r}_i), i=1, 2$ , cujos elementos principais são as componentes dos vetores  $\mathbf{r}_i$ , isto é

$$\mathbf{D}(\mathbf{r}_1) \otimes \mathbf{D}(\mathbf{r}_2) = \mathbf{D}(\mathbf{r}_1 \otimes \mathbf{r}_2). \quad (2.2)$$

Com  $\mathbf{S}_1$  e  $\mathbf{S}_2$  matrizes simétricas e,  $\mathbf{P}_1$  e  $\mathbf{P}_2$  matrizes ortogonais, tem-se

$$\mathbf{P}_i \mathbf{S}_i \mathbf{P}_i^t = \mathbf{D}(\mathbf{r}_i), i=1,2,$$

onde os elementos principais de  $\mathbf{D}(\mathbf{r}_i), i=1, 2$ , são os valores próprios de  $\mathbf{S}_i$  e os correspondentes vetores próprios, serão os vetores linha de  $\mathbf{P}_i, i=1,2$ . Atendendo a (2.1) e (2.2), tem-se

$$\begin{aligned} (\mathbf{P}_1 \otimes \mathbf{P}_2)(\mathbf{S}_1 \otimes \mathbf{S}_2)(\mathbf{P}_1 \otimes \mathbf{P}_2)^t &= (\mathbf{P}_1 \mathbf{S}_1 \mathbf{P}_1^t) \otimes (\mathbf{P}_2 \mathbf{S}_2 \mathbf{P}_2^t) \\ &= \mathbf{D}(\mathbf{r}_1 \otimes \mathbf{r}_2). \end{aligned} \quad (2.3)$$

Assim, os valores próprios de  $\mathbf{S}_1 \otimes \mathbf{S}_2$  serão os produtos dos valores próprios de  $\mathbf{S}_1$  e  $\mathbf{S}_2$  e os vetores próprios são os vetores linha de  $\mathbf{P}_1 \otimes \mathbf{P}_2$ .

Sendo a característica de uma matriz simétrica igual ao número de valores próprios não nulos, tem-se que

$$\text{car}(\mathbf{S}_1 \otimes \mathbf{S}_2) = \text{car}(\mathbf{S}_1) \cdot \text{Car}(\mathbf{S}_2).$$

Sendo o determinante de  $\mathbf{S}_i, i=1,2$ , dado pelo produto dos seus valores próprios, o mesmo sucede para o determinante de  $\mathbf{S}_1 \otimes \mathbf{S}_2$ , i.é.,

$$\det(\mathbf{S}_1 \otimes \mathbf{S}_2) = \det(\mathbf{S}_1)^{k_1} \cdot \det(\mathbf{S}_2)^{k_2},$$

com  $k_1$  e  $k_2$  as ordens de  $\mathbf{S}_1$  e  $\mathbf{S}_2$  (Schott, 1997). Por outro lado (Silvey, 1975), para qualquer matriz  $\mathbf{C}_i$ , tem-se

$$\text{car}(\mathbf{C}_i) = \text{car}(\mathbf{C}_i \mathbf{C}_i^t), i=1,2$$

e

$$(\mathbf{C}_1 \otimes \mathbf{C}_2)(\mathbf{C}_1 \otimes \mathbf{C}_2)^t = (\mathbf{C}_1 \mathbf{C}_1^t) \otimes (\mathbf{C}_2 \mathbf{C}_2^t),$$



resultando que

$$\text{car}(\mathbf{C}_1 \otimes \mathbf{C}_2) = \prod_{i=1}^2 \text{car}(\mathbf{C}_i \mathbf{C}_i^t) = \prod_{i=1}^2 \text{car}(\mathbf{C}_i) .$$

Caso as matrizes  $\mathbf{C}_1$  e  $\mathbf{C}_2$  sejam regulares, tem-se

$$(\mathbf{C}_1 \otimes \mathbf{C}_2)^{-1} = \mathbf{C}_1^{-1} \otimes \mathbf{C}_2^{-1} .$$

## 2.2 Matrizes Conjugadas

Os pares de matrizes conjugadas desempenharão neste trabalho, um papel importante no tratamento de famílias estruturadas de matrizes de produtos cruzados.

As matrizes  $\mathbf{A} = \mathbf{X}'\mathbf{X}$  e  $\mathbf{B} = \mathbf{X}\mathbf{X}'$  são matrizes conjugadas, satisfazendo a proposição:

**Proposição 2.1** As matrizes conjugadas  $\mathbf{A}$  e  $\mathbf{B}$  terão os mesmos valores próprios não nulos  $r_1, \dots, r_l$ , tendo-se  $r_j > 0, j = 1, \dots, l$ . Sendo  $\boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_l$ , os vetores próprios de  $\mathbf{A}$  associados a esses valores próprios, os

$$\boldsymbol{\gamma}_j = \frac{1}{\sqrt{r_j}} \mathbf{X}\boldsymbol{\varepsilon}_j, \quad j = 1, \dots, l,$$

serão os vetores próprios de  $\mathbf{B}$  associados aos valores próprios  $r_1, \dots, r_l$ . Finalmente se os  $\boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_l$  constituem uma base ortonormada para o espaço imagem  $R(\mathbf{A})$  de  $\mathbf{A}$ , os  $\boldsymbol{\gamma}_1, \dots, \boldsymbol{\gamma}_l$ , constituirão uma base ortonormada para  $R(\mathbf{B})$ . Tanto  $R(\mathbf{A})$  como  $R(\mathbf{B})$  terão dimensão  $l$ .

**Demonstração:** Tem-se  $0 \leq \boldsymbol{\varepsilon}_i' \mathbf{X}' \mathbf{X} \boldsymbol{\varepsilon}_i = \boldsymbol{\varepsilon}_i' (r_i \boldsymbol{\varepsilon}_i) = r_i \|\boldsymbol{\varepsilon}_i\|^2 = r_i, i = 1, \dots, l$ , o que mostra que os valores próprios não nulos de  $\mathbf{A}$  são positivos. Por outro lado

$$\mathbf{B}\boldsymbol{\gamma}_j = \mathbf{X}\mathbf{X}' \left( \frac{1}{\sqrt{r_j}} \mathbf{X}\boldsymbol{\varepsilon}_j \right) = \frac{1}{\sqrt{r_j}} \mathbf{X}(\mathbf{B}\boldsymbol{\varepsilon}_j) = \frac{1}{\sqrt{r_j}} \mathbf{X} r_j \boldsymbol{\varepsilon}_j = r_j \boldsymbol{\gamma}_j, \quad j = 1, \dots, l,$$

o que mostra que os  $\boldsymbol{\gamma}_1, \dots, \boldsymbol{\gamma}_l$  são vetores próprios de  $\mathbf{B}$  associados aos valores próprios  $r_1, \dots, r_l$ . Verifica-se ainda, que as matrizes  $\mathbf{A}$  e  $\mathbf{B}$  têm os mesmos valores próprios não nulos, logo a mesma característica  $l$ . Os  $\boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_l$  constituem uma base ortonormada para  $R(\mathbf{A})$ , sendo mutuamente ortogonais. Ora se,  $\boldsymbol{\varepsilon}_i^t \boldsymbol{\varepsilon}_l = 0, i \neq l$ , tem-se

$$\boldsymbol{\gamma}_i^t \boldsymbol{\gamma}_l = \left( \frac{1}{\sqrt{r_i}} \mathbf{X} \boldsymbol{\varepsilon}_i \right)^t \left( \frac{1}{\sqrt{r_l}} \mathbf{X} \boldsymbol{\varepsilon}_l \right) = \frac{1}{\sqrt{r_i r_l}} \boldsymbol{\varepsilon}_i^t \mathbf{A} \boldsymbol{\varepsilon}_l = \sqrt{\frac{r_l}{r_i}} \boldsymbol{\varepsilon}_i^t \boldsymbol{\varepsilon}_l = 0, i \neq l,$$

o que mostra que os  $\boldsymbol{\gamma}_1, \dots, \boldsymbol{\gamma}_l$ , sendo mutuamente ortogonais, constituem uma base ortonormada para  $R(\mathbf{B})$ , o que completa a demonstração.

Para a matriz quadrada  $k \times k$ ,  $\mathbf{U} = [u_{ll}]$  o respetivo traço é  $tr(\mathbf{U}) = \sum_{l=1}^k u_{ll}$ . Em particular,

se  $\mathbf{U} = v v^t$ , sendo  $v_1, \dots, v_k$  as componentes de  $v$  e,  $\|v\|^2 = \sum_{l=1}^k v_l^2$  o quadrado da respetiva

norma euclidiana, tem-se  $tr(v v^t) = \|v\|^2$ .

Dada a decomposição espectral da matriz simétrica  $\mathbf{M}$  dada por

$$\mathbf{M} = \sum_{j=1}^k \lambda_j \boldsymbol{\alpha}_j \boldsymbol{\alpha}_j^t,$$

tem-se

$$tr(\mathbf{M}) = \sum_{j=1}^k \lambda_j tr(\boldsymbol{\alpha}_j \boldsymbol{\alpha}_j^t) = \sum_{j=1}^k \lambda_j, \quad (2.4)$$

visto os vetores próprios  $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_k$  terem norma euclidiana igual a 1. Assim, da **Proposição 2.1** e de (2.4) resulta ter-se,  $tr(\mathbf{A}) = tr(\mathbf{B})$ . Por outro lado, sendo  $\mathbf{M}$  uma matriz simétrica, tem-se  $\mathbf{M} = \mathbf{P}^t \mathbf{D}(r_1, \dots, r_k) \mathbf{P}$  com  $\mathbf{P}$  matriz ortogonal e  $\mathbf{D}(r_1, \dots, r_k)$  a matriz diagonal cujos elementos principais são os valores próprios  $r_1, \dots, r_k$ , da matriz  $\mathbf{M}$ .

Com  $\mathbf{M} = [m_1, \dots, m_k]$ , sendo  $\|\mathbf{M}\|^2$  o quadrado da norma euclidiana de  $\mathbf{M}$  a soma dos quadrados dos seus elementos, temos

$$\|\mathbf{M}\|^2 = \sum_{j=1}^k \|m_j\|^2,$$

bem como, com  $\mathbf{P}$  matriz ortogonal

$$\|\mathbf{PM}\|^2 = \|\mathbf{P}[\mathbf{m}_1 \dots \mathbf{m}_k]\|^2 = \sum_{j=1}^k \|\mathbf{P}\mathbf{m}_j\|^2 = \sum_{j=1}^k \|\mathbf{m}_j\|^2 = \|\mathbf{M}\|^2,$$

já que

$$\|\mathbf{P}\mathbf{m}_j\|^2 = (\mathbf{P}\mathbf{m}_j)^t (\mathbf{P}\mathbf{m}_j) = \mathbf{m}_j^t \mathbf{P}^t \mathbf{P} \mathbf{m}_j = \mathbf{m}_j^t \mathbf{m}_j = \|\mathbf{m}_j\|^2.$$

Como as matrizes  $\mathbf{W}$  e  $\mathbf{W}^t$  têm os mesmos elementos tem-se  $\|\mathbf{W}\|^2 = \|\mathbf{W}^t\|^2$ ,

vindo

$$\|\mathbf{PMP}^t\|^2 = \|\mathbf{MP}^t\|^2 = \|\mathbf{PM}\|^2 = \|\mathbf{M}\|^2.$$

Dado que  $\mathbf{PMP}^t = \mathbf{D}(r_1, \dots, r_k)$ , tem-se

$$\|\mathbf{M}\|^2 = \|\mathbf{D}(r_1, \dots, r_k)\|^2 = \sum_{j=1}^k r_j^2.$$

Em particular, atendendo à **Proposição 2.1** tem-se

$$\|\mathbf{A}\|^2 = \|\mathbf{B}\|^2,$$

bem como

$$\frac{r_1^2}{\|\mathbf{A}\|^2 - \sum_{j=2}^l r_j^2} = \frac{r_1^2}{\|\mathbf{B}\|^2 - \sum_{j=2}^l r_j^2}. \quad (2.5)$$

com  $\mathbf{A} = \mathbf{X}^t \mathbf{X}$  e  $\mathbf{B} = \mathbf{X} \mathbf{X}^t$ .

Nos denominadores das expressões (2.5) apenas incluímos os  $l$  primeiros valores próprios por os restantes serem nulos.

## 2.3 Inversas de Moore-Penrose

Qualquer que seja a matriz  $\mathbf{B}$  existe uma matriz única  $\mathbf{B}^+$ , que satisfaz as condições

$$\begin{array}{ll}
1) \mathbf{B}\mathbf{B}^+\mathbf{B} = \mathbf{B} & 3) (\mathbf{B}\mathbf{B}^+)^t = \mathbf{B}\mathbf{B}^+ \\
2) \mathbf{B}^+\mathbf{B}\mathbf{B}^+ = \mathbf{B}^+ & 4) (\mathbf{B}^+\mathbf{B})^t = \mathbf{B}^+\mathbf{B}.
\end{array}$$

Facilmente se verifica que, se  $\mathbf{B}$  for invertível, tem-se

$$\mathbf{B}^+ = \mathbf{B}^{-1}$$

e que, em geral,  $(\mathbf{B}^t)^+ = (\mathbf{B}^+)^t$ , (Pollock, 1975; p.55).

As matrizes  $\mathbf{B}\mathbf{B}^+$  e  $\mathbf{B}^+\mathbf{B}$  são simétricas e idempotentes sendo pois matrizes de projeção ortogonal. Ora devido às condições 1) e 2) tem-se

$$\begin{cases} R(\mathbf{B}\mathbf{B}^+) = R(\mathbf{B}) \\ R(\mathbf{B}^+\mathbf{B}) = R(\mathbf{B}^+) \end{cases},$$

pelo que,  $\mathbf{B}\mathbf{B}^+$  e  $\mathbf{B}^+\mathbf{B}$  serão matrizes de projeção ortogonal sobre  $R(\mathbf{B})$  e  $R(\mathbf{B}^+)$ .

Admita-se que  $\mathbf{M}$  é matriz simétrica  $k \times k$ , tendo-se pois

$$\mathbf{M} = \mathbf{P}^t \mathbf{D}(r_1, \dots, r_k) \mathbf{P},$$

em geral, verifica-se ainda

$$\mathbf{M}^+ = \mathbf{P}^t \mathbf{D}(r_1^+, \dots, r_k^+) \mathbf{P},$$

com  $r_j^+ = r_j^{-1}$  caso  $r_j \neq 0$ ,  $j = 1, \dots, k$  e  $r_j^+ = 0$  caso  $r_j = 0$ ,  $j = 1, \dots, k$ .

Assim, temos uma expressão explícita para as inversas de Moore-Penrose de matrizes simétricas. Verifica-se ainda que

$$\mathbf{X}^+ = (\mathbf{X}^t \mathbf{X})^+ \mathbf{X}^t, \quad (2.6)$$

o que permite utilizar a expressão (2.6) para obter a inversa de Moore-Penrose duma matriz qualquer.

## 2.4 Modelos Base Ortogonais

### 2.4.1 Partições e Inferência

Como referimos no primeiro capítulo as matrizes duma família estruturada estarão associadas aos tratamentos de um modelo base. Nesta secção apresentamos os modelos base ortogonais a partir de partições ortogonais. Assim, havendo  $m$  tratamentos consideramos partições ortogonais

$$R^m = \bigoplus_{j=1}^p \varpi_j,$$

onde  $\bigoplus$  indica soma direta ortogonal de subespaços.

A par da partição ortogonal acima apresentada admitimos ter um par  $(\mathbf{Y}, S)$  onde  $\mathbf{Y}$  é normal com vetor médio  $\boldsymbol{\mu}$  e matriz de covariância  $\sigma^2 \mathbf{I}_m$ , pondo-se

$$\mathbf{Y} \sim N(\boldsymbol{\mu}, \sigma^2 \mathbf{I}_m),$$

independente de  $S$  que se distribui como o produto de  $\sigma^2$  por um qui-quadrado central com  $g$  graus de liberdade.

Os subespaços  $\omega_j = \varpi_j^\perp$ ,  $j = 1, \dots, p$ , complementos ortogonais dos  $\varpi_j$ ,  $j = 1, \dots, p$ , desempenham um papel importante na inferência para estes modelos. Assim as hipóteses

$$H_{0j}(\mathbf{b}): \boldsymbol{\mu} - \mathbf{b} \in \omega_j, \quad j = 1, \dots, p,$$

constituem uma família ortogonal de hipóteses. Constituindo os vetores linha de  $\mathbf{A}_j$  uma base ortonormada para  $\varpi_j$ ,  $j = 1, \dots, p$ , temos

$$N(\mathbf{A}_j) = \omega_j, \quad j = 1, \dots, p,$$

com  $N(\mathbf{A})$  o espaço de nulidade da matriz  $\mathbf{A}$ . Assim, com

$$\tilde{\boldsymbol{\eta}}_j = \mathbf{A}_j \mathbf{Y}, \quad \boldsymbol{\eta}_j = \mathbf{A}_j \boldsymbol{\mu} \quad \text{e} \quad \tilde{\boldsymbol{\eta}}_{j0}(\mathbf{b}) = \mathbf{A}_j \mathbf{b}, \quad j = 1, \dots, p$$

as hipóteses da família podem ser reescritas como

$$H_{0j}(\mathbf{b}) : \boldsymbol{\eta}_j = \boldsymbol{\eta}_{j0}(\mathbf{b}), j = 1, \dots, p,$$

ou como

$$H_{0j}(\mathbf{b}) : \delta_j(\mathbf{b}) = 0, j = 1, \dots, p,$$

com

$$\delta_j(\mathbf{b}) = \frac{1}{\sigma^2} \|\mathbf{A}_j \boldsymbol{\mu} - \mathbf{A}_j \mathbf{b}\|^2 = \frac{1}{\sigma^2} \|\boldsymbol{\eta}_j - \boldsymbol{\eta}_{j0}(\mathbf{b})\|^2, j = 1, \dots, p.$$

Ora,

$$S_j(\mathbf{b}) = \|\mathbf{A}_j \mathbf{Y} - \mathbf{A}_j \mathbf{b}\|^2 = \|\tilde{\boldsymbol{\eta}}_j - \boldsymbol{\eta}_{j0}(\mathbf{b})\|^2, j = 1, \dots, p,$$

distribui-se como o produto de  $\sigma^2$  por um qui-quadrado com

$$g_j = \dim(\boldsymbol{\omega}_j) = \text{car}(\mathbf{A}_j), j = 1, \dots, p,$$

graus de liberdade e parâmetro de não centralidade  $\delta_j(\mathbf{b})$ ,  $j = 1, \dots, p$  (Mexia, 1995)

Uma vez que os  $S_j(\mathbf{b})$ ,  $j = 1, \dots, p$ , são independentes de  $S$ , dado que  $\mathbf{Y}$  é independente de  $S$ , as estatísticas

$$\mathcal{F}_j(\mathbf{b}) = \frac{g_j}{g} \frac{S_j(\mathbf{b})}{S}, j = 1, \dots, p,$$

terão distribuição  $F$  com  $g_j$  e  $g$  graus de liberdade e parâmetro de não centralidade  $\delta_j(\mathbf{b})$ ,  $j = 1, \dots, p$ . Como estes parâmetros se anulam quando e só quando as hipóteses testadas se verificam, os testes  $F$  com estas estatísticas serão (Mexia, 1995), estritamente não distorcidos.

Por outro lado, os

$$S_j^o = \|\mathbf{A}_j \mathbf{Y} - \mathbf{A}_j \boldsymbol{\mu}\|^2 = \|\tilde{\boldsymbol{\eta}}_j - \boldsymbol{\eta}_j\|^2, j = 1, \dots, p,$$

distribuem-se como os produtos por  $\sigma^2$  de qui-quadrados centrais com  $g_j$ ,  $j = 1, \dots, p$ , graus de liberdade independentes de  $\mathbf{S}$ . Assim as variáveis pivot

$$\mathcal{F}_j^o = \frac{g}{g_j} \frac{S_j^o}{S}, j=1, \dots, p,$$

terão distribuições  $F$  centrais com  $g_j$  e  $g$  graus de liberdade,  $j=1, \dots, p$ .

Sendo  $f_{p,r,s}$  o  $p$ -ésimo quantil da distribuição  $F$  central com  $r$  e  $s$  graus de liberdade, obtemos as híper esferas de confiança de nível  $1-q$  para os  $\boldsymbol{\eta}_j, j=1, \dots, p$ , dadas por

$$\|\boldsymbol{\eta}_j - \tilde{\boldsymbol{\eta}}_j\|^2 \leq g_j f_{1-q, g_j, g} \frac{S}{g}, j=1, \dots, p.$$

Observe-se que os testes  $F$  atrás introduzidos gozam de dualidade, já que  $H_{0_j}(\mathbf{b})$  é rejeitada pelo teste de nível  $q$  se e só se a correspondente híper esfera de confiança de nível  $1-q$  não contiver  $\boldsymbol{\eta}_{j_0}(\mathbf{b}), j=1, \dots, p$ .

Atendendo ao teorema de Scheffé, obtemos os intervalos de confiança simultâneos, com nível conjunto  $1-q$ , dados por

$$pr \left[ \bigcap_{\mathbf{d}_j} \left( \left| \mathbf{d}_j^t \boldsymbol{\eta}_j - \mathbf{d}_j^t \tilde{\boldsymbol{\eta}}_j \right| \leq \sqrt{g_j \mathcal{F}_{1-q, g_j, g} \frac{S}{g}} \|\mathbf{d}_j\| \right) \right] = 1-q,$$

onde  $\bigcap$  indica que se consideram todos os vetores  $\mathbf{d}_j \in R^{g_j}, j=1, \dots, p$  (Scheffé, 1959). Apresentamos o tratamento para famílias ortogonais de hipóteses em vez do das hipóteses usuais

$$H_{0_j} : \boldsymbol{\eta}_j = 0, j=1, \dots, p,$$

para fazer sobressair as propriedades de dualidade dos testes  $F$  e as regiões de confiança a ele associadas. Estas últimas hipóteses são, em geral, hipóteses de ausência de efeitos ou interações. Este aspeto ressaltará mais na primeira secção em que atenderemos igualmente ao facto de para se puderem aplicar os resultados desta secção se tem de construir as matrizes  $\mathbf{A}_j, j=1, \dots, p$ .

## 2.4.2 Cruzamento Equilibrado de Fatores

De seguida consideram-se os modelos com cruzamento equilibrado os quais constituem um caso importante de modelos ortogonais. Nesta secção basear-nos-emos nos resultados desenvolvidos em Fonseca et al (2003). Admitamos que temos  $u$  fatores com  $a_1, \dots, a_u$  níveis que cruzam, havendo pois

$$m = \prod_{l=1}^u a_l,$$

combinações de níveis a que correspondem outros tantos tratamentos. Dado

$$\Gamma = \{\mathbf{h} : h_l = 0, 1; l = 1, \dots, u\}$$

os vetores de  $\Gamma$  permitem identificar os conjuntos de fatores. Se o  $l$ -ésimo fator pertence ao conjunto associado a  $\mathbf{h}$  ter-se-á  $h_l = 1$ , caso contrário ter-se-á  $h_l = 0$ .

Ao valor médio geral corresponde o conjunto de fatores vazio e,  $\mathbf{h} = \mathbf{0}$ . Aos efeitos do  $l$ -ésimo fator corresponderá  $\mathbf{h} = \boldsymbol{\delta}_l, l = 1, \dots, u$ , sendo  $\boldsymbol{\delta}_l, l = 1, \dots, u$ , o vetor com componentes todas nulas salvo a  $l$ -ésima que é igual a 1. Às interações corresponderão os vetores com componentes não nulas para os índices de fatores participantes, por exemplo com  $m = 6$ , às interações entre o terceiro e o quinto fator corresponderá o vetor

$$\mathbf{h}_{(3 \times 5)} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}.$$

A construção das matrizes  $\mathbf{A}(\cdot)$  pode ser feita numa forma sistemática (Fonseca et al., 2003). Para isso utiliza-se o produto de Kronecker de matrizes, já descrito na secção 2.1.1. Então para o conjunto de fatores associado ao vetor  $\mathbf{h}$ , ter-se-á a matriz



$$\mathbf{A}(\mathbf{h}) = \bigotimes_{l=1}^u \mathbf{A}_l(h_l),$$

onde

$$\mathbf{A}_l(0) = \frac{1}{\sqrt{a_l}} \mathbf{1}_{a_l}^t, \quad l = 1, \dots, u,$$

e  $\mathbf{A}_l(1)$ ,  $l = 1, \dots, u$ , é obtida retirando a primeira linha igual a  $\frac{1}{\sqrt{a_l}} \mathbf{1}_{a_l}^t$ ,  $l = 1, \dots, u$  a uma

matriz ortogonal  $a_l \times a_l$ ,  $l = 1, \dots, u$ . Como

$$\text{car}(k \otimes C) = \text{car}(k) \cdot \text{car}(C)$$

e

$$\begin{cases} \text{car}(\mathbf{A}_l(0)) = 1, \quad l = 1, \dots, u \\ \text{car}(\mathbf{A}_l(1)) = a_l - 1, \quad l = 1, \dots, u \end{cases},$$

tem-se

$$g(\mathbf{h}) = \text{car}(\mathbf{A}(\mathbf{h})) = \prod_{l=1}^u \text{car}(\mathbf{A}_l(h_l)) = \prod_{l=1}^u \text{car}(a_l - 1)^{h_l}$$

já que,  $(a_l - 1)^0 = 1$ ,  $l = 1, \dots, u$  e  $(a_l - 1)^1 = a_l - 1$ ,  $l = 1, \dots, u$ .

A aplicação destas fórmulas não oferece qualquer dificuldade. No entanto, apresentamos o seguinte exemplo: tomemos,  $m = 3$  bem como  $a_1 = 3$ ,  $a_2 = 3$  e  $a_3 = 4$ ,

tem-se

$$\begin{cases} g_1(0) = 1 \\ g_1(1) = 2 \\ g_2(0) = 1 \\ g_2(1) = 2 \\ g_3(0) = 1 \\ g_3(1) = 3 \end{cases},$$

vindo

$$\mathbf{A}_1(0) = \frac{1}{\sqrt{3}} \mathbf{1}_3^t, \quad \mathbf{A}_2(0) = \frac{1}{\sqrt{3}} \mathbf{1}_3^t \quad \text{e} \quad \mathbf{A}_3(0) = \frac{1}{\sqrt{4}} \mathbf{1}_4^t,$$

bem como

$$\mathbf{A}_1(\mathbf{1}) = \begin{bmatrix} -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix}, \mathbf{A}_2(\mathbf{1}) = \begin{bmatrix} -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix},$$

e

$$\mathbf{A}_3(\mathbf{1}) = \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

### 2.4.3 Heterocedasticidade Controlada

Nesta secção generalizamos os resultados da secção 2.4.1, trabalhando no âmbito da heterocedasticidade controlada (Mexia, 1989). Assim, admitimos que  $\mathbf{Y} \sim N(\boldsymbol{\mu}, \sigma^2 \mathbf{C})$ , com  $\mathbf{C}$  matriz conhecida e regular e independente de  $S \sim \sigma^2 \chi_g^2$ . Tal como referimos na secção 2.4.1 passamos agora a admitir que

$$\mathbf{Y} \sim N(\boldsymbol{\mu}, \sigma^2 \mathbf{C}),$$

com  $\mathbf{C}$  matriz conhecida e regular independente da variável aleatória  $S \sim \sigma^2 \chi_g^2$ , vindo

$$\tilde{\boldsymbol{\eta}}_j \sim N(\boldsymbol{\eta}_j, \sigma^2 \mathbf{A}_j \mathbf{C} \mathbf{A}_j^t), j = 1, \dots, p,$$

independente da variável aleatória  $S$ . Logo teremos

$$\tilde{\boldsymbol{\eta}}_j - \boldsymbol{\eta}_{j0}(\mathbf{b}) \sim N(\boldsymbol{\eta}_j - \boldsymbol{\eta}_{j0}(\mathbf{b}), \sigma^2 \mathbf{A}_j \mathbf{C} \mathbf{A}_j^t), j = 1, \dots, p$$

e

$$\tilde{\boldsymbol{\eta}}_j - \boldsymbol{\eta}_j \sim N(\mathbf{0}, \sigma^2 \mathbf{A}_j \mathbf{C} \mathbf{A}_j^t), j = 1, \dots, p,$$

obtendo-se

$$(\tilde{\boldsymbol{\eta}}_j - \boldsymbol{\eta}_{j0}(\mathbf{b}))^t (\mathbf{A}_j \mathbf{C} \mathbf{A}_j^t)^{-1} (\boldsymbol{\eta}_j - \boldsymbol{\eta}_{j0}(\mathbf{b})) \sim \sigma^2 \chi_{g_j, \delta_j(\mathbf{b})}^2, j = 1, \dots, p,$$

com

$$\delta_j(\mathbf{b}) = \frac{1}{\sigma^2} (\boldsymbol{\eta}_j - \boldsymbol{\eta}_{j0}(\mathbf{b}))^t (\mathbf{A}_j \mathbf{C} \mathbf{A}_j^t)^{-1} (\boldsymbol{\eta}_j - \boldsymbol{\eta}_{j0}(\mathbf{b})), \quad j = 1, \dots, p,$$

e

$$(\tilde{\boldsymbol{\eta}}_j - \boldsymbol{\eta}_{j0})^t (\mathbf{A}_j \mathbf{C} \mathbf{A}_j^t)^{-1} (\tilde{\boldsymbol{\eta}}_j - \boldsymbol{\eta}_{j0}) \sim \sigma^2 \chi_{g_j}^2, \quad j = 1, \dots, p.$$

Estes qui-quadrados são independentes de  $S$ , obtendo-se para as hipóteses  $H_{0j}(\mathbf{b})$ ,  $j = 1, \dots, p$ , as estatísticas de teste

$$\mathcal{F}_j(\mathbf{b}) = \frac{g}{g_j} \frac{(\tilde{\boldsymbol{\eta}}_j - \boldsymbol{\eta}_{j0}(\mathbf{b}))^t (\mathbf{A}_j \mathbf{C} \mathbf{A}_j^t)^{-1} (\tilde{\boldsymbol{\eta}}_j - \boldsymbol{\eta}_{j0}(\mathbf{b}))}{S}, \quad j = 1, \dots, p,$$

com  $g_j$  e  $g$  graus de liberdade e parâmetro de não centralidade  $\delta_j(\mathbf{b})$ ,  $j = 1, \dots, p$ .

Teremos igualmente as variáveis pivot

$$\mathcal{F}_j^o = \frac{g}{g_j} \frac{(\tilde{\boldsymbol{\eta}}_j - \boldsymbol{\eta}_j)^t (\mathbf{A}_j \mathbf{C} \mathbf{A}_j^t)^{-1} (\tilde{\boldsymbol{\eta}}_j - \boldsymbol{\eta}_j)}{S}, \quad j = 1, \dots, p,$$

com distribuição  $F$ , com  $g_j$  e  $g$  graus de liberdade. Obtemos agora para os  $\boldsymbol{\eta}_j$ ,  $j = 1, \dots, p$  os elipsoides de confiança de nível  $1 - q$  dados por

$$(\boldsymbol{\eta}_j - \tilde{\boldsymbol{\eta}}_j)^t (\mathbf{A}_j \mathbf{C} \mathbf{A}_j^t)^{-1} (\boldsymbol{\eta}_j - \tilde{\boldsymbol{\eta}}_j) \leq g_j f_{1-q, g_j, g} \frac{S}{g}, \quad j = 1, \dots, p.$$

Raciocinando como atrás mostra-se que os testes  $F$  são estritamente não distorcidos e gozam de dualidade. Por outro lado o teorema de Scheffé, dá-nos para os  $\mathbf{d}^t \boldsymbol{\eta}_j$ ,  $j = 1, \dots, p$ , os intervalos de confiança simultâneos expressos por (Scheffé, 1959),

$$pr \left[ \bigcap_{\mathbf{d}_j} \left( \left| \mathbf{d}_j^t \boldsymbol{\eta}_j - \mathbf{d}_j^t \tilde{\boldsymbol{\eta}}_j \right| \leq \sqrt{g_j \mathcal{F}_{1-q, g_j, g} \frac{S}{g} \mathbf{d}_j^t \mathbf{A} \mathbf{C} \mathbf{A}_j^t \mathbf{d}_j} \right) \right] = 1 - q, \quad j = 1, \dots, p$$

## 2.4.4 Limites Estocásticos e Distribuições Limite

Vamos agora considerar alguns resultados limite aplicáveis, quando:

- o número  $g$  de graus de liberdade de  $S$  é grande;
- se pode trabalhar com amostras cujos vetores médios  $\boldsymbol{\mu}$  têm norma que tende para infinito.

Como a variável aleatória  $S \sim \sigma^2 \chi_g^2$ , tem valor médio  $E(S) = g\sigma^2$  e  $\text{Var}(S) = 2g\sigma^4$ , ter-se-á utilizando a desigualdade de Bienaymé-Tchebycheff

$$pr\left(\left|\frac{S}{g} - \sigma^2\right| \leq \varepsilon\right) \geq 1 - \frac{2\sigma^4}{g\varepsilon^2} \xrightarrow{g \rightarrow \infty} 1,$$

vindo

$$\frac{S}{g} \xrightarrow[p \rightarrow \infty]{g \rightarrow \infty} \sigma^2,$$

onde  $\xrightarrow[p \rightarrow \infty]{g \rightarrow \infty}$  indica convergência estocástica. Resulta daqui que, admitindo a heterocedasticidade controlada,

$$\frac{g_j \mathcal{F}_j(\mathbf{b})}{(\tilde{\boldsymbol{\eta}}_j - \boldsymbol{\eta}_{j0}(\mathbf{b}))' (\mathbf{A}_j \mathbf{C} \mathbf{A}_j^t)^{-1} (\tilde{\boldsymbol{\eta}}_j - \boldsymbol{\eta}_{j0}(\mathbf{b}))} \xrightarrow[p \rightarrow \infty]{g \rightarrow \infty} 1, \quad j = 1, \dots, p,$$

logo, quando  $g$  é suficientemente grande, pode substituir-se  $\mathcal{F}_j(\mathbf{b})$ ,  $j = 1, \dots, p$ , por

$$U_j(\mathbf{b}) = (\tilde{\boldsymbol{\eta}}_j - \boldsymbol{\eta}_{j0}(\mathbf{b}))' (\mathbf{A}_j \mathbf{C} \mathbf{A}_j^t)^{-1} (\tilde{\boldsymbol{\eta}}_j - \boldsymbol{\eta}_{j0}(\mathbf{b})), \quad j = 1, \dots, p.$$

Recorde-se que

$$U_j(\mathbf{b}) \sim \sigma^2 \chi_{g_j, \delta_j(\mathbf{b})}^2, \quad j = 1, \dots, p,$$

e que

$$\delta_j(\mathbf{b}) = \frac{1}{\sigma^2} (\boldsymbol{\eta}_j - \boldsymbol{\eta}_{j0}(\mathbf{b}))' (\mathbf{A}_j \mathbf{C} \mathbf{A}_j^t)^{-1} (\boldsymbol{\eta}_j - \boldsymbol{\eta}_{j0}(\mathbf{b})), \quad j = 1, \dots, p.$$

Para utilizarmos as variáveis aleatórias  $U_j(\mathbf{b})$ ,  $j = 1, \dots, p$ , convém-nos reescreve-las como a soma de quadrados de variáveis normais independentes.

A matriz  $\mathbf{C}$  sendo regular é definida positiva já que  $\sigma^2\mathbf{C}$  é matriz de covariância regular, logo definida positiva. Tem-se então

$$\mathbf{C} = \mathbf{P}^t \mathbf{D}(r_1, \dots, r_m) \mathbf{P},$$

com  $\mathbf{P}$  matriz ortogonal e  $r_1, \dots, r_m$  os valores próprios positivos de  $\mathbf{C}$ . Assim, estará definida

$$\mathbf{C}^{\frac{1}{2}} = \mathbf{P}^t \mathbf{D}\left(r_1^{\frac{1}{2}}, \dots, r_m^{\frac{1}{2}}\right) \mathbf{P},$$

tendo-se

$$\mathbf{C} = \mathbf{C}^{\frac{1}{2}} \mathbf{C}^{\frac{1}{2}},$$

bem como

$$\mathbf{ACA}^t = \left( \mathbf{AC}^{\frac{1}{2}} \right) \left( \mathbf{AC}^{\frac{1}{2}} \right)^t.$$

Ora, (Silvey, 1975)

$$\text{car}(\mathbf{A}_j \mathbf{CA}_j^t) = \text{car}\left(\mathbf{A}_j \mathbf{C}^{\frac{1}{2}}\right) = \text{car}(\mathbf{A}_j) = g_j, \quad j = 1, \dots, p,$$

com  $\mathbf{C}^{\frac{1}{2}}$  matriz regular já que tem valores próprios  $r_j^{\frac{1}{2}} > 0, j = 1, \dots, p$  e sendo  $\mathbf{A}_j \mathbf{CA}_j^t$  uma matriz  $g_j \times g_j$ , é regular logo definida positiva visto,  $\sigma^2 \mathbf{A}_j \mathbf{CA}_j^t$  ser a matriz de covariância de  $\tilde{\boldsymbol{\eta}}_j, j = 1, \dots, p$ .

Temos então

$$\mathbf{A}_j \mathbf{CA}_j^t = \mathbf{P}_j \mathbf{D}(r_{j1}, \dots, r_{jg_j}) \mathbf{P}_j^t, \quad j = 1, \dots, p,$$

com  $\mathbf{P}_j$  matriz ortogonal e os  $(r_{j1}, \dots, r_{jg_j}), j = 1, \dots, p$ , os valores próprios positivos da matriz  $\mathbf{A}_j \mathbf{CA}_j^t, j = 1, \dots, p$ . Estará pois definida

$$\mathbf{G}_j = \mathbf{P}_j \mathbf{D}\left(r_{j1}^{-\frac{1}{2}}, \dots, r_{jg_j}^{-\frac{1}{2}}\right) \mathbf{P}_j^t, \quad j = 1, \dots, p,$$

tendo-se

$$\begin{aligned}\mathbf{G}_j(\mathbf{A}_j\mathbf{C}\mathbf{A}_j^t)\mathbf{G}_j^t &= \mathbf{P}_j\mathbf{D}\left(r_{j1}^{-\frac{1}{2}}, \dots, r_{jg_j}^{-\frac{1}{2}}\right)\mathbf{P}_j^t\mathbf{P}_j\mathbf{D}(r_{j1}, \dots, r_{jg_j})\mathbf{P}_j^t\mathbf{P}_j\mathbf{D}(r_{j1}, \dots, r_{jg_j})\mathbf{P}_j^t \\ &= \mathbf{I}_{g_j}, j = 1, \dots, p,\end{aligned}$$

bem como

$$\begin{aligned}(\mathbf{A}_j\mathbf{C}\mathbf{A}_j^t)^{-1} &= \mathbf{P}_j\mathbf{D}(r_{j1}^{-1}, \dots, r_{jg_j}^{-1})\mathbf{P}_j^t \\ &= \mathbf{P}_j\mathbf{D}\left(r_{j1}^{-\frac{1}{2}}, \dots, r_{jg_j}^{-\frac{1}{2}}\right)\mathbf{P}_j^t\mathbf{P}_j\mathbf{D}\left(r_{j1}^{-\frac{1}{2}}, \dots, r_{jg_j}^{-\frac{1}{2}}\right)\mathbf{P}_j^t \\ &= \mathbf{G}_j\mathbf{G}_j, j = 1, \dots, p.\end{aligned}$$

Logo

$$U_j(\mathbf{b}) = (\tilde{\boldsymbol{\eta}}_j - \boldsymbol{\eta}_{j0}(\mathbf{b}))^t \mathbf{G}_j \mathbf{G}_j (\tilde{\boldsymbol{\eta}}_j - \boldsymbol{\eta}_{j0}(\mathbf{b})),$$

tendo-se

$$\mathbf{G}_j(\tilde{\boldsymbol{\eta}}_j - \boldsymbol{\eta}_{j0}(\mathbf{b})) = \mathbf{X}_j + \boldsymbol{\mu}_j(\mathbf{b}), j = 1, \dots, p,$$

com vetor  $\mathbf{X}_j$ , tal que  $\mathbf{X}_j \sim N(0, \sigma^2 \mathbf{I}_{g_j})$ ,  $j = 1, \dots, p$ ,

e

$$\boldsymbol{\mu}_j(\mathbf{b}) = \mathbf{G}_j(\boldsymbol{\eta}_j - \boldsymbol{\eta}_{j0}(\mathbf{b})), j = 1, \dots, p.$$

Temos pois estatísticas da forma

$$Y = g(\boldsymbol{\mu} + \mathbf{X}),$$

onde  $\mathbf{X} \sim N(0, \sigma^2 \mathbf{I}_m)$ , com  $g(\mathbf{u}) = \sum_{j=1}^m u_j^2$ , tendo-se  $g_1(\mathbf{u}) = 2\mathbf{u}$  e  $g_2(\mathbf{u}) = 2\mathbf{I}_m$  e assim a distribuição limite de

$$Y^o = \frac{Y - g(\boldsymbol{\mu})}{\|g_1(\boldsymbol{\mu})\|} = \frac{Y - \|\boldsymbol{\mu}\|^2}{2\|\boldsymbol{\mu}\|}$$

quando  $\|\boldsymbol{\mu}\| \rightarrow \infty$ , será  $N(0, \sigma^2)$  (Mexia e Oliveira, 2011; Nunes et al., 2013)

Dados os limites  $-z_p^0$  e  $z_p^0$  para as caudas da densidade normal reduzida com probabilidade  $\frac{p}{2}$ ,

tem-se

$$pr\left(-z_q^0 \leq \frac{Y - \|\boldsymbol{\mu}\|^2}{2\|\boldsymbol{\mu}\|} \leq z_q^0\right) \xrightarrow{\|\boldsymbol{\mu}\| \rightarrow \infty} 1 - q.$$

Resolvendo este par de desigualdades em ordem a  $\|\boldsymbol{\mu}\|^2$  obtemos em primeiro lugar

$$Y^2 - 2\|\boldsymbol{\mu}\|^2 Y + \|\boldsymbol{\mu}\|^4 \leq 4z_q^{0^2} \|\boldsymbol{\mu}\|^2,$$

em seguida

$$\|\boldsymbol{\mu}\|^4 - 2(Y + 2z_q^{0^2})\|\boldsymbol{\mu}\|^2 + Y^2 \leq 0,$$

vindo

$$Y + 2z_q^{0^2} - \sqrt{4z_q^{0^2} Y + 4z_q^{0^4}} \leq \|\boldsymbol{\mu}\|^2 \leq Y + 2z_q^{0^2} + \sqrt{4z_q^{0^2} Y + 4z_q^{0^4}}.$$

Este par de desigualdades define um intervalo de confiança bilateral para  $\|\boldsymbol{\mu}\|^2$  de nível limite  $1 - q$ . Para obter os intervalos para os  $\|\boldsymbol{\mu}_j(\mathbf{b})\|^2$ ,  $j = 1, \dots, p$  há que tomar

$$Y = \|\mathbf{G}_j(\tilde{\boldsymbol{\eta}}_j - \boldsymbol{\eta}_{j0}(\mathbf{b}))\|^2, \quad j = 1, \dots, p.$$

Por dualidade podemos testar hipóteses para os  $\boldsymbol{\mu}_j(\mathbf{b})$ ,  $j = 1, \dots, p$ . Observe-se que

$$\boldsymbol{\mu}_j(\mathbf{b}) = 0, \quad j = 1, \dots, p,$$

se e só se

$$\boldsymbol{\eta}_j = \boldsymbol{\eta}_{j0}(\mathbf{b}), \quad j = 1, \dots, p,$$

isto é, quando e apenas quando  $H_{0j}(\mathbf{b}) = 0$ ,  $j = 1, \dots, p$  se verifica.

Observe-se que os resultados de Mexia e Oliveira (2011) utilizados se aplicam a amostras cuja não centralidade cresce. No Apêndice A apresentamos alguns conceitos e

resultados assintóticos da Linearidade Assintótica, referindo o estudo da aplicação ao caso normal bem como um estudo de simulações que confirmam a aplicabilidade desses resultados.



# Capítulo 3

## Modelos e Famílias de Modelos para Matrizes Estocásticas Simétricas

### 3.1 Introdução

Dada uma matriz  $\mathbf{M}=[m_{ij}]$  aleatória simétrica  $k \times k$ , a respetiva matriz média  $\boldsymbol{\mu}$  quando definida, será simétrica tendo a decomposição espectral

$$\boldsymbol{\mu} = \sum_{i=1}^k \lambda_i \boldsymbol{\alpha}_i \boldsymbol{\alpha}_i^t,$$

vindo

$$\mathbf{M} = \sum_{i=1}^k \lambda_i \boldsymbol{\alpha}_i \boldsymbol{\alpha}_i^t + \bar{\mathbf{E}},$$

com  $\bar{\mathbf{E}}$  matriz estocástica simétrica com matriz média nula. O grau de  $\mathbf{M}$  será  $k = \text{car}(\boldsymbol{\mu})$ , que é a característica da matriz média (suposta definida). Nos modelos que se consideram a seguir, admite-se que

$$\mathbf{Z} = \text{vec}(\mathbf{M}) \sim N(\boldsymbol{\eta}, \sigma^2 \mathbf{L}),$$

é (aproximadamente) normal, com vetor médio  $\boldsymbol{\eta}$  e matriz de covariância  $\sigma^2 \mathbf{L}$  (Areia et al., 2008). Na teoria geral começamos por considerar os vetores de estrutura  $\boldsymbol{\beta}_i = \lambda_i \boldsymbol{\alpha}_i$ ,  $i = 1, \dots, k$  e vetores de resíduos. Veremos em seguida como determinar o grau do modelo. Segue-se a estimação dos vetores de estrutura e de parâmetros de não centralidade a eles associados. A partir dos estimadores potenciais obter-se-ão regiões

de confiança para esses parâmetros. Discutimos em seguida o grau de ajustamento conseguido e, para modelos de grau  $k > 1$  considera-se ainda a possibilidade de truncar o modelo quando se têm valores próprios  $\theta_1, \dots, \theta_k$  muito superiores aos restantes. A concluir este capítulo consideramos alguns casos especiais. O primeiro destes casos, diz respeito às matrizes de produtos de Hilbert-Schmidt as quais desempenham um papel importante na primeira etapa, a inter-estrutura, da metodologia STATIS, (ver por exemplo Oliveira, 2002). O produto de Hilbert-Schmidt de duas matrizes é o produto interno dos respetivos  $vec$ , define-se  $vec(\mathbf{A})$  como sendo o agrupamento dos vetores coluna da matriz  $\mathbf{A}$ , isto é, com  $\mathbf{A} = [a_{ij}]_{n \times m}$ ,  $vec(\mathbf{A}) = (a_{11}, \dots, a_{n1}, \dots, a_{1m}, \dots, a_{nm})$ . Assim, as matrizes  $\mathbf{S} = [s_{ij}]$ , com  $s_{ij} = tr(\mathbf{A}_i \mathbf{A}_j^t)$ ,  $i, j = 1, \dots, n$  de produtos de Hilbert-Schmidt, são um caso particular das matrizes de produtos cruzados  $\mathbf{M} = [m_{ij}]$  com

$$m_{ij} = v_i^t v_j, i, j = 1, \dots, k.$$

A segunda aplicação que consideramos referir-se-á a matrizes de produtos cruzados. Ao considerarmos estas matrizes pôr-se-á em evidência a simplificação do tratamento que resulta das propriedades das matrizes conjugadas.

## 3.2 Vetores de Estrutura e Resíduos

Sendo  $\mathbf{M}$  uma matriz  $n \times n$  com valores próprios  $\theta_1, \dots, \theta_n$  a que correspondem os vetores próprios  $\gamma_1, \dots, \gamma_n$ , mutuamente ortogonais, podemos sempre ordenar os valores próprios por forma a ter  $|\theta_1| \geq \dots \geq |\theta_n|$ . Então, seguindo Oliveira (2002) e Areia et al. (2008), podemos utilizar os

$$\tilde{\beta}_i = \theta_i \gamma_i, i = 1, \dots, k,$$

para estimar os vetores de estrutura dados por

$$\beta_i = \lambda_i \alpha_i, i = 1, \dots, k.$$

Ora, tem-se

$$\tilde{\beta}_i = \mathbf{M} \gamma_i, i = 1, \dots, k,$$

pelo que, pondo

$$\mathbf{M} = [m_1 \quad \dots \quad m_n],$$

tem-se

$$\begin{aligned} \tilde{\boldsymbol{\beta}}_i = \mathbf{M}\boldsymbol{\gamma}_i = \mathbf{M}^t\boldsymbol{\gamma}_i &= \begin{bmatrix} m_1^t\boldsymbol{\gamma}_i \\ \vdots \\ m_n^t\boldsymbol{\gamma}_i \end{bmatrix} = \begin{bmatrix} \boldsymbol{\gamma}_i^t m_1 \\ \vdots \\ \boldsymbol{\gamma}_i^t m_n \end{bmatrix} \\ &= (\mathbf{I}_n \otimes \boldsymbol{\gamma}_i^t) \begin{bmatrix} m_1 \\ \vdots \\ m_n \end{bmatrix} \\ &= (\mathbf{I}_n \otimes \boldsymbol{\gamma}_i^t) \mathbf{Z}, \quad i = 1, \dots, k \end{aligned}$$

com  $\mathbf{Z} = \text{vec}(\mathbf{M})$ .

Assim,  $\tilde{\boldsymbol{\beta}}_i$ ,  $i = 1, \dots, k$  terá vetor médio

$$E(\tilde{\boldsymbol{\beta}}_i) = (\mathbf{I}_n \otimes \boldsymbol{\gamma}_i^t) \boldsymbol{\eta}, \quad i = 1, \dots, k$$

e matriz de covariância

$$\Sigma(\tilde{\boldsymbol{\beta}}_i) = \sigma^2 (\mathbf{I}_n \otimes \boldsymbol{\gamma}_i^t) \mathbf{L} (\mathbf{I}_n \otimes \boldsymbol{\gamma}_i), \quad i = 1, \dots, k.$$

Se  $\mathbf{M}$  tem grau  $k$  e os  $\tilde{\boldsymbol{\beta}}_1, \dots, \tilde{\boldsymbol{\beta}}_k$  forem bons estimadores dos  $\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_k$ , e  $\boldsymbol{\gamma}_1, \dots, \boldsymbol{\gamma}_k$  bons estimadores de  $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_k$ , teremos

$$\bar{\mathbf{E}} = \mathbf{M} - \sum_{i=1}^k \boldsymbol{\beta}_i \boldsymbol{\alpha}_i^t \approx \mathbf{M} - \sum_{i=1}^k \tilde{\boldsymbol{\beta}}_i \boldsymbol{\gamma}_i^t = \mathbf{M} - \sum_{i=1}^k (\mathbf{I}_n \otimes \boldsymbol{\gamma}_i^t) \mathbf{Z} \boldsymbol{\gamma}_i^t,$$

vindo

$$\mathbf{R} = \text{vec}(\bar{\mathbf{E}}) \approx \mathbf{Z} - \text{vec} \left( \sum_{i=1}^k ((\mathbf{I}_n \otimes \boldsymbol{\gamma}_i^t) \mathbf{Z}) \boldsymbol{\gamma}_i^t \right).$$

Ora

$$\text{vec} \left( \sum_{i=1}^k (\mathbf{I}_n \otimes \boldsymbol{\gamma}_i^t) \mathbf{Z} \boldsymbol{\gamma}_i^t \right) = \sum_{i=1}^k (\boldsymbol{\gamma}_i \otimes \mathbf{I}_n \otimes \boldsymbol{\gamma}_i^t) \mathbf{Z},$$

pelo que, com

$$\mathbf{W} = \sum_{i=1}^k (\boldsymbol{\gamma}_i \otimes \mathbf{I}_n \otimes \boldsymbol{\gamma}_i^t)$$

teremos,

$$\mathbf{R} = (\mathbf{I}_{n^2} - \mathbf{W})\mathbf{Z},$$

sendo a matriz de covariâncias de  $\mathbf{R}$  será dada por

$$\Sigma(\mathbf{R}) = \sigma^2(\mathbf{I}_{n^2} - \mathbf{W})\mathbf{L}(\mathbf{I}_{n^2} - \mathbf{W}^t).$$

Enquanto  $\mathbf{R}$  pode ser considerado como um vetor de resíduos,

$$\tilde{\boldsymbol{\beta}} = \left[ \tilde{\boldsymbol{\beta}}_1^t, \dots, \tilde{\boldsymbol{\beta}}_k^t \right]^t,$$

será o vetor de estrutura global ajustado.

Os vetores  $\mathbf{R}$  e  $\tilde{\boldsymbol{\beta}}$  têm um papel fundamental na inferência. Como estes vetores não são independentes temos de obter um vetor de resíduos homocedástico e independente do vetor  $\tilde{\boldsymbol{\beta}}$ . Começamos por obter uma matriz  $\mathbf{A}$  tal que  $\mathbf{AR}$  é independente de

$$\tilde{\boldsymbol{\beta}} = \mathbf{B}\mathbf{Z} = \begin{bmatrix} \mathbf{I}_n \otimes \boldsymbol{\gamma}_1^t \\ \vdots \\ \mathbf{I}_n \otimes \boldsymbol{\gamma}_k^t \end{bmatrix} \mathbf{Z}.$$

Dado  $\tilde{\boldsymbol{\beta}}$  e  $\mathbf{AR}$  serem transformados lineares de  $\mathbf{Z} \sim N(\boldsymbol{\eta}, \sigma^2\mathbf{L})$ , serão independentes se e só se

$$\Sigma(\mathbf{AR}, \tilde{\boldsymbol{\beta}}) = 0,$$

sendo

$$r = \text{car}[(\mathbf{I}_{n^2} - \mathbf{W})\mathbf{L}\mathbf{B}^t].$$

Aplicando o processo de ortonormalização de Gram-Schmidt, aos vetores coluna de  $(\mathbf{I}_{n^2} - \mathbf{W})\mathbf{L}\mathbf{B}^t$  obtemos os vetores mutuamente ortogonais com norma um,  $\mathbf{z}_1, \dots, \mathbf{z}_r$ .

Sendo  $\boldsymbol{\delta}_i$ ,  $i = 1, \dots, n^2$ , o vetor com  $n^2$  componentes todas nulas salvo a  $i$ -ésima que é 1, e voltando a aplicar o processo de Gram-Schmidt aos vetores coluna  $\mathbf{z}_1, \dots, \mathbf{z}_r$ ,  $\boldsymbol{\delta}_1, \dots, \boldsymbol{\delta}_{n^2}$ , obtemos, para além dos vetores  $\mathbf{z}_1, \dots, \mathbf{z}_r$  os vetores linha  $\mathbf{a}_1, \dots, \mathbf{a}_{n^2-r}$  da matriz  $\mathbf{A}$ , tal que  $\mathbf{AR}$  é independente de  $\tilde{\boldsymbol{\beta}}$ . A matriz de covariância de  $\mathbf{AR}$  é dada por

$$\Sigma(\mathbf{A}\mathbf{R}) = \sigma^2 \mathbf{A}(\mathbf{I}_{n^2} - \mathbf{W})\mathbf{L}(\mathbf{I}_{n^2} - \mathbf{W}^t)\mathbf{A}^t,$$

com valores próprios não nulos  $v_1, \dots, v_g$  associados aos vetores próprios  $\zeta_1, \dots, \zeta_g$ , uma vez que as matrizes de covariância não têm valores próprios negativos. Assim, a matriz

$$G = D(v_1^{-\frac{1}{2}}, \dots, v_g^{-\frac{1}{2}}) [\zeta_1, \dots, \zeta_g]^t,$$

estará definida. Vê-se facilmente que

$$\Sigma(\mathbf{G}\mathbf{A}\mathbf{R}) = \sigma^2 \mathbf{I}_g,$$

pelo que

$$\mathbf{Y} = \mathbf{G}\mathbf{A}\mathbf{R},$$

será o vetor de resíduos homocedástico e independente de  $\tilde{\boldsymbol{\beta}}$ . A informação contida na matriz é assim condensada no par  $(\tilde{\boldsymbol{\beta}}, V)$  com  $\tilde{\boldsymbol{\beta}} \sim N(\boldsymbol{\beta}\boldsymbol{\eta}, \sigma^2 \mathbf{B}\mathbf{L}\mathbf{B}^t)$  independente de  $V = \|\mathbf{Y}\|^2 \sim \sigma^2 \chi_g^2$ .

### 3.3 Grau do Modelo

Sendo  $k$ , o grau do modelo, a informação é condensada nos vetores de estrutura estimados  $\tilde{\boldsymbol{\beta}}_i$  e na soma de quadrados de resíduos  $V$ . Assim, a existência do vetor homocedástico de resíduos  $\mathbf{Y}$  é uma consequência testável dos seguintes pressupostos:

- o modelo é dado por

$$\mathbf{M} = \sum_{i=1}^k \lambda_i \boldsymbol{\alpha}_i \boldsymbol{\alpha}_i^t + \bar{\mathbf{E}}, \quad |\lambda_i| \geq 0, \quad i = 1, \dots, k,$$

com

$$\mathbf{Z} = \text{vec}(\mathbf{M}) \sim N(\boldsymbol{\eta}, \sigma^2 \mathbf{L}),$$

sendo a matriz  $\mathbf{L}$  conhecida e definida positiva;

- os  $\tilde{\beta}_1, \dots, \tilde{\beta}_k$  são bons estimadores dos  $\beta_1, \dots, \beta_k$ .

Somos assim, levados a testar a hipótese

$$H_{0k} : Y_1, \dots, Y_g \text{ i.i.d. } \sim N(0, \sigma^2),$$

para o que podemos utilizar, por exemplo, o teste de Shapiro-Wilk (Christenses, 1987).

Uma outra possibilidade é utilizar um teste com estatística

$$\mathcal{F}_k = \frac{(g-1) \left( \sum_{i=1}^g Y_i \right)^2}{g \sum_{i=1}^g Y_i^2 - \left( \sum_{i=1}^g Y_i \right)^2},$$

que quando,  $H_{0k}$  se verifica, tem distribuição F central com 1 e  $g-1$  graus de liberdade e escreve-se  $F(. | 1, g-1)$ . Representamos por  $F(. | r, s)$  a distribuição com  $r$  e  $s$  graus de liberdade.

Quando  $H_{0k}$  não se verifica as variáveis aleatórias  $Y_1, \dots, Y_g$  terão valores médios não nulos  $\mu_1, \dots, \mu_g$  e o numerador e denominador de  $\mathcal{F}_k$  terão os parâmetros de não centralidade dados por

$$\delta_{1k} = \frac{1}{\sigma^2} \left( \frac{1}{g} \left( \sum_{i=1}^g \mu_i \right)^2 \right) = g \left( \frac{\bar{\mu}}{\sigma} \right)^2$$

e

$$\delta_{2k} = \frac{1}{\sigma^2} \left( \sum_{i=1}^g \mu_i^2 - \frac{1}{g} \left( \sum_{i=1}^g \mu_i \right)^2 \right) = \frac{1}{\sigma^2} \sum_{i=1}^g (\mu_i - \bar{\mu})^2 = \frac{1}{\sigma^2} (\|\boldsymbol{\mu}\|^2 - g\bar{\mu}^2)$$

com  $\boldsymbol{\mu} = (\mu_1, \dots, \mu_g)$  e  $\bar{\mu} = \frac{1}{g} \sum_{i=1}^g \mu_i$ .

Assim, haverá alternativas à hipótese  $H_{0k}$  em que  $\delta_{1k}$  predomina sobre  $\delta_{2k}$  e, em que  $\mathcal{F}_k$  tende a tomar valores superiores aos que tomaria, caso  $H_{0k}$  se verificasse. Por outro lado, haverá alternativas à hipótese  $H_{0k}$  em que  $\delta_{2k}$  predomina sobre  $\delta_{1k}$  e, em que  $\mathcal{F}_k$  tende a tomar valores inferiores aos que tomaria, caso  $H_{0k}$  se verificasse. Um estudo numérico, mostrou que com  $f_{p,1,g-1}$  o p-ésimo quantil de  $F(. | 1, g-1)$ ,

utilizando-se regiões de rejeição  $\left[ f_{\frac{q}{2}, 1, g-1}; f_{1-\frac{q}{2}, 1, g-1} \right]$  se tem um teste de nível  $q$  cuja potência cresce rapidamente com  $\delta_1$  e  $\delta_2$  (Mexia, 1989).

É importante observar que, a hipótese  $H_{ok}$  verifica-se se

$$\mathbf{M} = \sum_{i=1}^h \lambda_i \mathbf{a}_i \mathbf{a}_i^t + \bar{\mathbf{E}}, \quad |\lambda_i| \geq 0, i = 1, \dots, h,$$

com  $h < k$ . Assim, ao testar-se  $H_{ok}$  está-se a testar um limite superior para o grau de  $\mathbf{M}$ . Na prática deve-se começar por testar  $H_{ok}$  baixando o valor de  $k$  até se obter uma rejeição. O valor de  $\bar{k}$  será sugerido pelos valores dos  $|\theta_1| \geq \dots \geq |\theta_n|$ , em geral tomando-se

$$\bar{k} = \text{Min} \{ m : |\bar{\theta}_m| \leq d \},$$

com " $d$ " fixado á priori. Sendo  $k'$  o valor de  $k$  em que ocorreu, para o modelo de teste escolhido, a primeira rejeição dum  $H_{ok}$ , é-se levado a admitir que  $|\lambda_{k'}| > 0$ , isto é que o modelo tem grau  $k'$ . Para confirmar esta conclusão somos levados a testar

$$H_{0k'} : \lambda_{k'} = 0,$$

sendo a nossa conclusão validada quando esta hipótese é rejeitada. Dada  $H_{ok'+1}$  não ter sido rejeitada, podemos admitir que os  $\tilde{\beta}_1, \dots, \tilde{\beta}_{k'}$  são independentes do vetor homocedástico de resíduos  $\mathbf{Y} \sim N(0, \sigma^2 \mathbf{I}_g)$ .

Com  $\mathbf{U}^+$  a inversa de Moore-Penrose de  $\mathbf{U}$ ,

$$\varphi_{k'} = \tilde{\beta}_{k'}^t [(\mathbf{I}_n \otimes \boldsymbol{\gamma}_{k'}^t) \mathbf{L} (\mathbf{I}_n \otimes \boldsymbol{\gamma}_{k'})]^+ \tilde{\beta}_{k'},$$

será, o produto de  $\sigma^2$  por um qui-quadrado com

$$r_{k'} = \text{car} [(\mathbf{I}_n \otimes \boldsymbol{\gamma}_{k'}^t) \mathbf{L} (\mathbf{I}_n \otimes \boldsymbol{\gamma}_{k'})]^+,$$

graus de liberdade e parâmetro de não centralidade

$$\delta_{k'} = \frac{1}{\sigma^2} \boldsymbol{\beta}_{k'}^t [(\mathbf{I}_n \otimes \boldsymbol{\gamma}_{k'}^t) \mathbf{L} (\mathbf{I}_n \otimes \boldsymbol{\gamma}_{k'}^t)]^+ \boldsymbol{\beta}_{k'}.$$

Este qui-quadrado é independente de  $\|\mathbf{Y}\|^2$  que é o produto por  $\sigma^2$  dum qui-quadrado central com  $g$  graus de liberdade, logo a estatística

$$\mathcal{F}_{k'} = \frac{g}{r_{k'}} \frac{\varphi_{k'}}{\|\mathbf{Y}\|^2},$$

tem distribuição  $F$  com  $r_{k'}$  e  $g$  graus de liberdade e parâmetro de não centralidade  $\delta_{k'}$ . Este último parâmetro anula-se quando a hipótese  $H_{ok'}$  se verifica pelo que o teste  $F$  com estatística  $\mathcal{F}_{k'}$  é não distorcido.

Com efeito,  $\bar{F}(z | t, s, \delta) = F\left(\frac{s}{t} z | t, s, \delta\right)$  sendo a distribuição do quociente  $\frac{\chi_{t,\delta}^2}{\chi_s^2}$  de qui-quadrados independentes função decrescente de  $\delta$ . Assim, a potência do teste  $F$  de nível  $q$ , com  $t$  e  $s$  graus de liberdade é dada por

$$Pot(\delta | t, s) = 1 - F(f_{1-q,t,s} | t, s, \delta) = 1 - \bar{F}\left(\frac{s}{t} f_{1-q,t,s} | t, s, \delta\right),$$

e é crescente com  $\delta$ .

Esta observação é importante para o nosso modelo pois mostra que a relevância dos  $\boldsymbol{\beta}_i, i = 1, \dots, h$  pode ser medida pelos correspondentes parâmetros de não centralidade. Assim, a par dos  $\boldsymbol{\beta}_i, i = 1, \dots, h$  também estamos interessados nesses parâmetros.

## 3.4 Estimação do Modelo

### 3.4.1 Estimação Pontual

Como vimos na secção anterior 3.3, admitindo que  $|\lambda_{k'}| > 0$  o modelo proposto terá grau  $k'$ . Admita-se ainda que a hipótese  $H_{o,k'+1}$  não é rejeitada, neste caso poder-se-á



considerar que a informação é condensada nos vetores de estrutura estimados  $\tilde{\beta}_i$  e na soma de quadrados de resíduos  $V$ .

Interessa-nos pois estimar os  $\beta_i, i = 1, \dots, k'$  com

$$\beta = [\beta_1^t, \dots, \beta_{k'}^t]^t,$$

Para estimar os parâmetros de não centralidade. Para estimar os vetores  $\beta_i, i = 1, \dots, k'$  utilizam-se os estimadores  $\tilde{\beta}_i, i = 1, \dots, k'$  com

$$\tilde{\beta} = [\tilde{\beta}_1^t, \dots, \tilde{\beta}_{k'}^t]^t,$$

pelo que nos basta considerar os parâmetros de não centralidade. Ora, a forma quadrática

$$\varphi_i = \tilde{\beta}_i^t [(\mathbf{I}_n \otimes \gamma_i^t) \mathbf{L}(\mathbf{I}_n \otimes \gamma_i)]^+ \tilde{\beta}_i, \quad i = 1, \dots, k',$$

distribui-se como o produto de  $\sigma^2$  por um qui-quadrado com

$$r_i = \text{car}(\mathbf{I}_n \otimes \gamma_i^t) \mathbf{L}(\mathbf{I}_n \otimes \gamma_i), \quad i = 1, \dots, k'$$

graus de liberdade e parâmetro de não centralidade

$$\delta_i = \tilde{\beta}_i^t [(\mathbf{I}_n \otimes \gamma_i^t) \mathbf{L}(\mathbf{I}_n \otimes \gamma_i)]^+ \tilde{\beta}_i, \quad i = 1, \dots, k',$$

vindo  $\varphi_i \sim \sigma^2 \chi_{r_i, \delta_i}^2, i = 1, \dots, k'$ . Ora, os  $\tilde{\beta}_i$  e consequentemente, os  $\varphi_i, i = 1, \dots, k'$  são independentes de  $\|\mathbf{Y}\|^2$  pelo que

$$\tau_i = \frac{\varphi_i}{\|\mathbf{Y}\|^2}, \quad i = 1, \dots, k',$$

distribui-se como o quociente  $\frac{\chi_{r_i, \delta_i}^2}{\chi_g^2}$  de qui-quadrados independentes, pondo-se

$$\tau_i \sim \bar{F}(r_i, g, \delta_i), \quad i = 1, \dots, k'.$$

Por outro lado, com

$$\mathbf{B} = \begin{bmatrix} \mathbf{I}_n \otimes \gamma_1^t \\ \vdots \\ \mathbf{I}_n \otimes \gamma_{k'}^t \end{bmatrix}$$

temos

$$\tilde{\boldsymbol{\beta}} = \mathbf{B}\mathbf{Z} \sim N(\boldsymbol{\beta}\boldsymbol{\eta}, \sigma^2(\mathbf{B}\mathbf{L}\mathbf{B}^t)),$$

logo

$$\varphi = \tilde{\boldsymbol{\beta}}^t (\mathbf{B}\mathbf{L}\mathbf{B}^t)^+ \tilde{\boldsymbol{\beta}} \sim \sigma^2 \chi_{r,\delta}^2,$$

com

$$r = \text{car}(\mathbf{B}\mathbf{L}\mathbf{B}^t)$$

e

$$\delta = \frac{1}{\sigma^2} \boldsymbol{\beta}^t (\mathbf{B}\mathbf{L}\mathbf{B}^t)^+ \boldsymbol{\beta}$$

independente de  $\|\mathbf{Y}\|^2$ , pelo que

$$\tau = \frac{\varphi}{\|\mathbf{Y}\|^2},$$

terá distribuição  $\bar{F}(\cdot | r, g, \delta)$ . Iremos utilizar as estatísticas  $\tau_i, i=1, \dots, k'$  e  $\tau$  para estimar os  $\delta_i, i=1, \dots, k'$  e  $\delta$ , respetivamente. Para evitar duplicações apresentamos apenas o caso de  $\delta$ . Ora, temos

$$\bar{F}(z | r, g, \delta) = e^{-\frac{\delta}{2}} \sum_{i=0}^{\infty} \frac{\left(\frac{\delta}{2}\right)^i}{i!} \bar{F}(z | r+2i, g, 0),$$

com  $\bar{F}(z | t, s)$  a distribuição do quociente de qui-quadrados centrais independentes com  $t$  e  $s$  graus de liberdade (Mexia, 1995; p. 51), (Robbins, 1948 e 1949).

A distribuição  $\bar{F}(z | t, s)$  tem função densidade

$$\bar{f}(z | t, s) = \begin{cases} 0 & , z < 0 \\ \frac{\Gamma\left(\frac{t+s}{2}\right)}{\Gamma\left(\frac{t}{2}\right)\Gamma\left(\frac{s}{2}\right)} \frac{z^{\frac{t}{2}-1}}{(1+z)^{\frac{t+s}{2}}} & , z > 0 \end{cases},$$

como o integral em  $]-\infty, +\infty[$  duma densidade é 1, tem-se

$$1 = \int_{-\infty}^{+\infty} \bar{f}(z | h, s) dz = \frac{\Gamma\left(\frac{h+s}{2}\right)}{\Gamma\left(\frac{h}{2}\right)\Gamma\left(\frac{s}{2}\right)} \int_0^{+\infty} \frac{z^{\frac{h}{2}-1}}{(1+z)^{\frac{h+s}{2}}} dz,$$

obtendo-se

$$\int_0^{+\infty} \frac{z^{\frac{h}{2}-1}}{(1+z)^{\frac{h+s}{2}}} dz = \frac{\Gamma\left(\frac{h}{2}\right)\Gamma\left(\frac{s}{2}\right)}{\Gamma\left(\frac{h+s}{2}\right)}.$$

Assim, o momento de ordem  $l$  da distribuição  $\bar{F}(\cdot | t, s)$  relativo à origem estará definido quando  $s > 2l$  sendo dado por

$$\begin{aligned} \mu'_l(\cdot | t, s) &= \frac{\Gamma\left(\frac{t+s}{2}\right)}{\Gamma\left(\frac{t}{2}\right)\Gamma\left(\frac{s}{2}\right)} \int_0^{+\infty} \frac{z^{\frac{h+2l}{2}-1}}{(1+z)^{\frac{h+s}{2}}} dz \\ &= \frac{\Gamma\left(\frac{t+s}{2}\right)}{\Gamma\left(\frac{t}{2}\right)\Gamma\left(\frac{s}{2}\right)} \int_0^{+\infty} \frac{z^{\frac{h+2l}{2}-1}}{(1+z)^{\frac{(h+2l)+(s-2l)}{2}}} dz \\ &= \frac{\Gamma\left(\frac{t+s}{2}\right)}{\Gamma\left(\frac{t}{2}\right)\Gamma\left(\frac{s}{2}\right)} \frac{\Gamma\left(\frac{t+2l}{2}\right)\Gamma\left(\frac{s-2l}{2}\right)}{\Gamma\left(\frac{(t+2l)+(s-2l)}{2}\right)} \\ &= \frac{\Gamma\left(\frac{t+2l}{2}\right)}{\Gamma\left(\frac{t}{2}\right)} \frac{\Gamma\left(\frac{s-2l}{2}\right)}{\Gamma\left(\frac{s}{2}\right)}. \end{aligned}$$

Como  $\Gamma(a+1) = a\Gamma(a)$ , teremos para o valor médio (definido caso  $s > 2$ )

$$\mu(\cdot | t, s) = \frac{\frac{t}{2}}{\frac{s-2}{2}} = \frac{t}{s-2},$$

analogamente o segundo momento (definido caso  $s > 4$ ), será dado por

$$\mu'_2(|t, s) = \frac{\Gamma\left(\frac{t+4}{2}\right) \Gamma\left(\frac{s-4}{2}\right)}{\Gamma\left(\frac{t}{2}\right) \Gamma\left(\frac{s}{2}\right)} = \frac{\frac{t+2}{2} \frac{t}{2}}{\frac{s-2}{2} \frac{s-4}{2}} = \frac{(t+2)t}{(s-2)(s-4)},$$

tendo-se pois, a variância (definida caso  $s > 4$ ),

$$\begin{aligned} \text{Var}(|t, s) &= \frac{(t+2)t}{(s-2)(s-4)} - \frac{t^2}{(s-2)^2} = \frac{t}{(s-2)} \left[ \frac{t+2}{s-4} - \frac{t}{s-2} \right] \\ &= \frac{t}{(s-2)} \frac{ts + 2s - 2t - 4 - ts + 4t}{(s-2)(s-4)} = \frac{t}{(s-2)} \frac{2s + 2t + 4}{(s-4)(s-2)}. \end{aligned}$$

Passando ao caso não central, para  $\bar{F}(|r, g, \delta)$  temos, caso  $g > 2$ , o valor médio

$$\begin{aligned} \mu(|r, g, \delta) &= e^{-\frac{\delta}{2}} \sum_{i=0}^{\infty} \frac{\left(\frac{\delta}{2}\right)^i}{i!} \mu(|r+2i, g) \\ &= e^{-\frac{\delta}{2}} \sum_{i=0}^{\infty} \frac{\left(\frac{\delta}{2}\right)^i}{i!} \frac{r+2i}{g-2} \\ &= \frac{r}{g-2} e^{-\frac{\delta}{2}} \sum_{i=0}^{\infty} \frac{\left(\frac{\delta}{2}\right)^i}{i!} + \frac{2}{g-2} e^{-\frac{\delta}{2}} \sum_{i=1}^{\infty} \frac{\left(\frac{\delta}{2}\right)^i}{(i-1)!} \\ &= \frac{r}{g-2} + \frac{2}{g-2} \frac{\delta}{2} e^{-\frac{\delta}{2}} \sum_{i=1}^{\infty} \frac{\left(\frac{\delta}{2}\right)^{i-1}}{(i-1)!} = \frac{r+\delta}{g-2}, \end{aligned}$$

logo

$$\delta = (g-2)\mu(|r, g, \delta) - r,$$

tendo-se pois, o estimador centrado (caso  $g > 2$ )

$$\tilde{\delta} = (g-2)\tau - r.$$

Da mesma forma, o segundo momento (caso  $g > 4$ ), é dado por

$$\begin{aligned} \mu'_2(|r, g, \delta) &= e^{-\frac{\delta}{2}} \sum_{i=0}^{\infty} \frac{\left(\frac{\delta}{2}\right)^i}{i!} \mu'_2(|r+2i, g) \\ &= e^{-\frac{\delta}{2}} \sum_{i=0}^{\infty} \frac{\left(\frac{\delta}{2}\right)^i}{i!} \frac{(r+2i+2)(r+2i)}{(g-2)(g-4)} = e^{-\frac{\delta}{2}} \sum_{i=0}^{\infty} \frac{\left(\frac{\delta}{2}\right)^i}{i!} \frac{r^2 + 4ri + 4i^2 + 2r + 4i}{(g-2)(g-4)} \\ &= \frac{(r^2+2r)}{(g-2)(g-4)} e^{-\frac{\delta}{2}} \sum_{i=0}^{\infty} \frac{\left(\frac{\delta}{2}\right)^i}{i!} + \frac{4(r+1)}{(g-2)(g-4)} e^{-\frac{\delta}{2}} \sum_{i=0}^{\infty} \frac{\left(\frac{\delta}{2}\right)^i}{i!} i + \frac{4}{(g-2)(g-4)} e^{-\frac{\delta}{2}} \sum_{i=0}^{\infty} \frac{\left(\frac{\delta}{2}\right)^i}{i!} i^2 \\ &= \frac{(r^2+2r)}{(g-2)(g-4)} + \frac{\left(\frac{\delta}{2}\right)4(r+1)}{(g-2)(g-4)} e^{-\frac{\delta}{2}} \sum_{i=1}^{\infty} \frac{\left(\frac{\delta}{2}\right)^{i-1}}{(i-1)!} + \frac{4\left(\frac{\delta}{2}\right)^2}{(g-2)(g-4)} e^{-\frac{\delta}{2}} \sum_{i=2}^{\infty} \frac{\left(\frac{\delta}{2}\right)^{i-2}}{(i-2)!} \\ &= \frac{r(r+2) + \delta(r+2+\delta) + \delta^2}{(g-2)(g-4)}, \end{aligned}$$

pelo que

$$\begin{aligned} \text{Var}(|r, g, \delta) &= \frac{r^2 + 2(r+1)\delta + \delta^2}{(g-2)(g-4)} - \frac{(r+\delta)^2}{(g-2)^2} \\ &= \frac{2r^2 - 2\delta(g+2r) + 4\delta - \delta^2(2-g)}{(g-2)^2(g-4)} \\ &= \frac{2(r+\delta)^2 + 2(g-2)(r+1)\delta}{(g-2)^2(g-4)} \end{aligned}$$

Assim, ter-se-á

$$\text{Var}(\tilde{\delta}) = \frac{2(r+\delta)^2 - 2(g-2)(r+1)\delta}{(g-4)},$$

sendo os estimadores centrados dados por

$$\tilde{\delta}_i = (g-2)\tau_i - r_i, \quad i = 1, \dots, k',$$

com as variâncias

$$\text{Var}(\tilde{\delta}_i) = \frac{2(r_i + \delta_i)^2 + 2\delta_i(g-2)(r_i - 1)}{(g-4)}, \quad i=1, \dots, k'.$$

### 3.4.2 Regiões de Confiança

Começamos por obter intervalos de confiança para os parâmetros de não centralidade. Para evitar duplicações centra-se a nossa exposição em  $\delta$ .

A função  $\bar{F}(z | r, g, \delta)$  decresce estritamente com  $\delta$ , pelo que podemos utilizar o método geral para obter intervalos de confiança para  $\delta$ . Ora,

$$\begin{aligned} \frac{\partial \bar{F}(z | r, g, \delta)}{\partial \delta} &= \frac{\partial}{\partial \delta} \left[ e^{-\frac{\delta}{2}} \sum_{i=0}^{\infty} \frac{\left(\frac{\delta}{2}\right)^i}{i!} \bar{F}(z | r+2i, \delta) \right] \\ &= -\frac{1}{2} e^{-\frac{\delta}{2}} \sum_{i=0}^{\infty} \frac{\left(\frac{\delta}{2}\right)^i}{i!} \bar{F}(z | r+2i, \delta) + \frac{e^{-\frac{\delta}{2}}}{2} \sum_{i=1}^{\infty} \frac{\left(\frac{\delta}{2}\right)^{i-1}}{(i-1)!} \bar{F}(z | r+2i, \delta) \\ &= -\frac{1}{2} \bar{F}(z | r, g, \delta) + \frac{e^{-\frac{\delta}{2}}}{2} \sum_{i'=0}^{\infty} \frac{\left(\frac{\delta}{2}\right)^{i'}}{i'!} \bar{F}(z | r+2+2i', \delta) \\ &= \frac{\bar{F}(z | r+2, g, \delta) - \bar{F}(z | r, g, \delta)}{2}. \end{aligned}$$

Sendo  $\chi_{r,\delta}^2$ ,  $\chi_g^2$  e  $\chi_2^2$  qui-quadrados independentes, atendendo á reprodutividade dos qui-quadrados, as variáveis aleatórias

$$U_1 = \frac{(\chi_{r,\delta}^2 + \chi_2^2)}{\chi_g^2}$$

e

$$U_2 = \frac{\chi_{r,\delta}^2}{\chi_g^2},$$

têm distribuição  $\bar{F}(\cdot|r+2, g, \delta)$  e  $\bar{F}(\cdot|r, g, \delta)$ , respetivamente. Como  $pr(U_2 < U_1) = 1$ , tem-se (com  $z > 0$ ),

$$\bar{F}(z|r+2, g, \delta) < \bar{F}(z|r, g, \delta),$$

vindo, também com  $z > 0$

$$\frac{\partial \bar{F}(z|r, g, \delta)}{\partial \delta} < 0,$$

isto é  $\bar{F}(z|r, g, \delta)$  é função decrescente de  $\delta$ , tal como referido atrás. Podendo assim, utilizar-se o resultado geral para obter intervalos de confiança para  $\delta$  (Mood et al. 1967; p. 389-391). Assim, sendo  $\delta_p$  a solução da equação em  $\delta$

$$\bar{F}(z|r, g, \delta) = 1 - p,$$

teremos os intervalos de confiança de nível  $1 - q$  para  $\delta$ :

$$\left[ \frac{\delta_q}{2}; \delta_{1-\frac{q}{2}} \right], [0; \delta_{1-q}] \text{ e } [\delta_{1-q}; +\infty[ ,$$

os quais permitem por dualidade construir testes para

$$H_0(\mathbf{d}) : \delta = \mathbf{d}.$$

O teste bilateral, unilateral esquerdo [direito] de nível  $q$ , rejeita esta hipótese se  $\mathbf{d}$  não estiver contido no primeiro [segundo, terceiro], intervalos de confiança.

Observa-se que se pode utilizar o método da bissecção para resolver as equações em  $\delta$  e obter os  $\delta_p$  (Hanning, 1962; p.382). Como

$$\bar{F}(z|r, g, \delta) = e^{-\frac{\delta}{2}} \sum_{i=0}^{\infty} \frac{\left(\frac{\delta}{2}\right)^i}{i!} \bar{F}(z|r+2i, g),$$

para calcular os valores de  $\bar{F}(z|r, g, \delta)$ , podemos utilizar

$$\bar{F}_N(z|r, g, \delta) = e^{-\frac{\delta}{2}} \sum_{i=0}^N \frac{\left(\frac{\delta}{2}\right)^i}{i!} \bar{F}(z|r+2i, g),$$

sempre que se tenha, com  $z > 0$

$$e^{-\frac{\delta}{2}} \sum_{i=N+1}^{\infty} \frac{\left(\frac{\delta}{2}\right)^i}{i!} < \varepsilon,$$

ter-se-á com,  $z > 0$

$$\bar{F}_N(z|r, g, \delta) < \bar{F}(z|r, g, \delta) < \bar{F}_N(z|r, g, \delta) + \varepsilon,$$

já que

$$0 \leq \bar{F}(z|r+2i, g) \leq 1, i = 0, 1, \dots$$

É possível assim, controlar o erro de truncatura que se tem quando se substitui  $\bar{F}$  por  $\bar{F}_N$ . Considerando  $\varepsilon$  um limite superior para o erro de truncatura, será mais fácil obter o número mínimo de termos necessário para que o erro de truncatura não exceda este (Nunes et al., 2012).

Assim, a função

$$g_N(\delta) = e^{-\frac{\delta}{2}} \sum_{i=N+1}^{\infty} \frac{\left(\frac{\delta}{2}\right)^i}{i!},$$

é crescente, visto ter-se

$$\begin{aligned} g'_N(\delta) &= -\frac{1}{2} e^{-\frac{\delta}{2}} \sum_{i=N+1}^{\infty} \frac{\left(\frac{\delta}{2}\right)^i}{i!} + \frac{1}{2} e^{-\frac{\delta}{2}} \sum_{i=N+1}^{\infty} \frac{\left(\frac{\delta}{2}\right)^{i-1}}{(i-1)!} \\ &= \frac{1}{2} e^{-\frac{\delta}{2}} \frac{\left(\frac{\delta}{2}\right)^N}{N!} > 0 \end{aligned}$$



Logo  $g_N(\delta)$  e por sua vez o número máximo de termos a tomar, cresce com  $\delta$ . Assim, é necessário fixar um limite superior  $\bar{\delta}$  para os valores de  $\delta$  ( $\delta \leq \bar{\delta}$ ) que se consideram, e determinar o número mínimo de termos a tomar de forma a garantir um erro de truncatura quando muito igual a  $\varepsilon$  (com  $\varepsilon$  um limite superior para os valores de  $\delta$ ). Como

$$e^{-\frac{\delta}{2}} \sum_{i=N+1}^{\infty} \frac{\left(\frac{\delta}{2}\right)^i}{i!} = 1 - e^{-\frac{\delta}{2}} \sum_{i=0}^N \frac{\left(\frac{\delta}{2}\right)^i}{i!},$$

podemos por

$$N' = \text{Max} \left\{ N : e^{-\frac{\delta}{2}} \sum_{i=0}^N \frac{\left(\frac{\delta}{2}\right)^i}{i!} > 1 - \varepsilon \right\} \quad (3.1)$$

Na Tabela 3.1 apresentam-se os menores valores de  $N'$  para os erros de truncatura,  $10^{-2}$ ,  $10^{-3}$ ,  $10^{-4}$ ,  $10^{-5}$ ,  $10^{-6}$ ,  $10^{-8}$  e os valores de  $\bar{\delta} : 5, 10, 15, 20, 50$  e  $100$ , que satisfazem a condição 3.1. Estes valores foram obtidos recorrendo a uma rotina elaborada em Fortran 2008 (Apêndice B).

Tabela 3.1: Número de termos.

$\bar{\delta}$	$\varepsilon$					
	$10^{-2}$	$10^{-3}$	$10^{-4}$	$10^{-5}$	$10^{-6}$	$10^{-8}$
5	7	9	10	12	13	16
10	11	13	15	17	19	22
15	15	17	20	22	24	27
20	18	21	24	26	28	32
50	37	42	46	49	52	58
100	67	73	78	83	87	94

Da análise da Tabela 3.1 é possível concluir que é possível fazer o controlar bastante bem o erro de truncatura.

Temos em seguida as regiões de confiança para os  $\beta_j$ ,  $j = 1, \dots, k'$ , e para o vetor  $\beta$ .  
 Mais uma vez para evitar duplicações, centrar-nos-emos no caso do vetor  $\beta$ .

Assim,

$$\varphi' = (\beta - \tilde{\beta})' (\mathbf{B} \mathbf{L} \mathbf{B}' )^+ (\beta - \tilde{\beta}),$$

distribuir-se-á como o produto por  $\sigma^2$  dum qui-quadrado central com  $r$  graus de liberdade independente de  $\|\mathbf{Y}\|^2$ , logo

$$\mathcal{F}' = \frac{g}{r} \frac{\varphi'}{\|\mathbf{Y}\|^2},$$

terá distribuição F central com  $r$  e  $g$  graus de liberdade. Sendo  $f_{p,r,g}$  o p-ésimo quantil dessa distribuição temos, para  $\beta$ , o elipsoide de confiança de nível  $1 - q$ , dado por

$$(\beta - \tilde{\beta})' (\mathbf{B} \mathbf{L} \mathbf{B}' )^+ (\beta - \tilde{\beta}) \leq r f_{1-q,r,g} \frac{\|\mathbf{Y}\|^2}{g}.$$

Por dualidade, pode testar-se a hipótese

$$H_o : \beta = \mathbf{b},$$

rejeitando o teste de nível  $q$  desta hipótese se  $\mathbf{b}$  não estiver contido no elipsoide de confiança de nível  $1 - q$ . Seguindo Scheffé (1959) e recorrendo ao teorema de Scheffé, obtemos os intervalos de confiança simultâneos para os

$$\rho(\mathbf{d}) = \mathbf{d}' \beta,$$

para os quais temos os estimadores centrados dados por

$$\tilde{\rho}(\mathbf{d}) = \mathbf{d}' \tilde{\beta}.$$

Dado o produto de matrizes  $\mathbf{B} \mathbf{L} \mathbf{B}'$  poder não ser invertível, utilizamos a versão deste teorema dada por Mexia (1989). Tem-se então

$$pr \left[ \bigcap_{\mathbf{d}} (|\rho(\mathbf{d}) - \tilde{\rho}(\mathbf{d})| \leq \sqrt{rf_{1-q,r,g}(\mathbf{d}'\mathbf{BLB}'\mathbf{d}) \frac{\|\mathbf{Y}\|^2}{g}}) \right] = 1 - q.$$

Para se considerar  $\beta_i, i = 1, \dots, k'$  em vez de  $\beta$  bastaria substituir  $r$  por  $r_i$  e  $\mathbf{BLB}'$  por  $(\mathbf{I}_n \otimes \gamma_i^t) \mathbf{L} (\mathbf{I}_n \otimes \gamma_i)$ ,  $i = 1, \dots, k'$ . No entanto, do ponto de vista algébrico temos alguns aspetos interessantes. Em primeiro lugar observamos que os vetores linha de  $\mathbf{I}_n \otimes \gamma_i^t, i = 1, \dots, k'$  são linearmente independentes. Temos ainda a Proposição seguinte:

### Proposição 3.1

Quando a matriz  $\mathbf{L}$  é definida positiva, as

$$(\mathbf{I}_n \otimes \gamma_i^t) \mathbf{L} (\mathbf{I}_n \otimes \gamma_i), \quad i = 1, \dots, k'$$

também são definidas positivas, tendo-se

$$\left[ (\mathbf{I}_n \otimes \gamma_i^t) \mathbf{L} (\mathbf{I}_n \otimes \gamma_i) \right]^+ = \left[ (\mathbf{I}_n \otimes \gamma_i^t) \mathbf{L} (\mathbf{I}_n \otimes \gamma_i) \right]^{-1}, \quad i = 1, \dots, k'.$$

**Demonstração:** Caso a matriz  $\mathbf{L}$  seja definida positiva, tem-se

$$\mathbf{L} = \mathbf{P}' \mathbf{D}(s_1, \dots, s_{n^2}) \mathbf{P},$$

com  $\mathbf{P}$  matriz ortogonal e  $\mathbf{D}(s_1, \dots, s_{n^2})$  a matriz diagonal com elementos principais  $s_1, \dots, s_{n^2}$ , os valores próprios positivos de  $\mathbf{L}$ . Assim está definida a matriz

$$\mathbf{L}^{\frac{1}{2}} = \mathbf{P}' \mathbf{D}\left(s_1^{\frac{1}{2}}, \dots, s_{n^2}^{\frac{1}{2}}\right) \mathbf{P},$$

tendo-se

$$\begin{aligned} (\mathbf{I}_n \otimes \gamma_i^t) \mathbf{L} (\mathbf{I}_n \otimes \gamma_i) &= (\mathbf{I}_n \otimes \gamma_i^t) \mathbf{L}^{\frac{1}{2}} \mathbf{L}^{\frac{1}{2}} (\mathbf{I}_n \otimes \gamma_i) \\ &= \left[ (\mathbf{I}_n \otimes \gamma_i^t) \mathbf{L}^{\frac{1}{2}} \right] \left[ (\mathbf{I}_n \otimes \gamma_i^t) \mathbf{L}^{\frac{1}{2}} \right]^t, \quad i = 1, \dots, k'. \end{aligned}$$

Ora,  $\mathbf{L}^{\frac{1}{2}}$  também é definida positiva, logo invertível, tendo-se

$$\text{car}((\mathbf{I}_n \otimes \boldsymbol{\gamma}_i^t) \mathbf{L}^{\frac{1}{2}}) = \text{car}(\mathbf{I}_n \otimes \boldsymbol{\gamma}_i^t) = n, \quad i = 1, \dots, k'$$

e como

$$\text{car}((\mathbf{I}_n \otimes \boldsymbol{\gamma}_i^t) \mathbf{L}^{\frac{1}{2}})((\mathbf{I}_n \otimes \boldsymbol{\gamma}_i^t) \mathbf{L}^{\frac{1}{2}}) = \text{car}((\mathbf{I}_n \otimes \boldsymbol{\gamma}_i^t) \mathbf{L}^{\frac{1}{2}}) = n, \quad i = 1, \dots, k'$$

e

$$(\mathbf{I}_n \otimes \boldsymbol{\gamma}_i^t) \mathbf{L} (\mathbf{I}_n \otimes \boldsymbol{\gamma}_i), \quad i = 1, \dots, k',$$

é uma matriz  $n \times n$ , a mesma é invertível (Silvey, 1975). Observe-se agora que, se  $v \neq 0$

$$v^t \left[ (\mathbf{I}_n \otimes \boldsymbol{\gamma}_i^t) \mathbf{L} (\mathbf{I}_n \otimes \boldsymbol{\gamma}_i) \right] v = \left\| \mathbf{L}^{\frac{1}{2}} (\mathbf{I}_n \otimes \boldsymbol{\gamma}_i) v \right\|^2 > 0, \quad i = 1, \dots, k'.$$

Visto os  $n$  vetores coluna da matriz  $\mathbf{L}^{\frac{1}{2}} (\mathbf{I}_n \otimes \boldsymbol{\gamma}_i)$ ,  $i = 1, \dots, k'$ , serem linearmente independentes, já que

$$\text{car}(\mathbf{L}^{\frac{1}{2}} (\mathbf{I}_n \otimes \boldsymbol{\gamma}_i)) = \text{car}((\mathbf{I}_n \otimes \boldsymbol{\gamma}_i^t) \mathbf{L}) = n, \quad i = 1, \dots, k'.$$

Assim, as matrizes  $(\mathbf{I}_n \otimes \boldsymbol{\gamma}_i^t) \mathbf{L} (\mathbf{I}_n \otimes \boldsymbol{\gamma}_i)$ ,  $i = 1, \dots, k'$ , são definidas positivas, logo invertíveis e, conseqüentemente, as respectivas inversas de Moore-Penrose coincidem com as respectivas matrizes inversas usuais. Vindo

$$\varphi'_i = (\boldsymbol{\beta}_i - \tilde{\boldsymbol{\beta}}_i)' \left[ (\mathbf{I}_n \otimes \boldsymbol{\gamma}_i^t) \mathbf{L} (\mathbf{I}_n \otimes \boldsymbol{\gamma}_i) \right]^{-1} (\boldsymbol{\beta}_i - \tilde{\boldsymbol{\beta}}_i), \quad i = 1, \dots, k',$$

as variáveis pivot

$$\mathcal{F}'_i = \frac{g}{r_i} \frac{\varphi'_i}{\|\mathbf{Y}\|^2}, \quad i = 1, \dots, k,$$

têm distribuição  $F(\cdot | r_i, g)$ , com

$$r_i = \left[ \text{car}(\mathbf{I}_n \otimes \boldsymbol{\gamma}_i^t) \mathbf{L}(\mathbf{I}_n \otimes \boldsymbol{\gamma}_i) \right] = n, \quad i = 1, \dots, k',$$

e  $g$  graus de liberdade. Obtêm-se assim para os  $\boldsymbol{\beta}_i, i = 1, \dots, k'$ , os elipsoides de confiança, de nível  $1 - q$ , dados por

$$(\boldsymbol{\beta}_i - \tilde{\boldsymbol{\beta}}_i)' \left[ (\mathbf{I}_n \otimes \boldsymbol{\gamma}_i^t) \mathbf{L}(\mathbf{I}_n \otimes \boldsymbol{\gamma}_i) \right]^{-1} (\boldsymbol{\beta}_i - \tilde{\boldsymbol{\beta}}_i) \leq n f_{1-q, n, g} \frac{\|\mathbf{Y}\|^2}{g}, \quad i = 1, \dots, k',$$

sempre que  $\mathbf{L}$  é definida positiva. Têm-se também os intervalos de confiança simultâneos dados por

$$pr \left[ \bigcap_{\mathbf{d}_i} |\rho(\mathbf{d}_i) - \tilde{\rho}(\mathbf{d}_i)| \leq \sqrt{n F_{1-q, n, g}(\mathbf{d}_i^t (\mathbf{I}_n \otimes \boldsymbol{\gamma}_i^t) \mathbf{L}(\mathbf{I}_n \otimes \boldsymbol{\gamma}_i) \mathbf{d}_i)} \frac{\|\mathbf{Y}\|^2}{g} \right] = 1 - q, \quad i = 1, \dots, k'.$$

Quando a matriz  $\mathbf{L}$  é definida positiva, temos

$$\Phi_{k'} = \tilde{\boldsymbol{\beta}}_{k'}^t \left[ (\mathbf{I}_n \otimes \boldsymbol{\gamma}_{k'}^t) \mathbf{L}(\mathbf{I}_n \otimes \boldsymbol{\gamma}_{k'}) \right]^{-1} \tilde{\boldsymbol{\beta}}_{k'},$$

e

$$r_{k'} = n$$

com

$$\delta_{k'} = \frac{1}{\sigma^2} \boldsymbol{\beta}_{k'}^t \left[ (\mathbf{I}_n \otimes \boldsymbol{\gamma}_{k'}^t) \mathbf{L}(\mathbf{I}_n \otimes \boldsymbol{\gamma}_{k'}) \right]^{-1} \boldsymbol{\beta}_{k'}.$$

A estatística de teste será dada por

$$\mathcal{F}_{k'}^0 = \frac{g}{n} \frac{\Phi_{k'}}{\|\mathbf{Y}\|^2}.$$

### 3.5 Truncatura do Modelo

Começamos por recordar alguns resultados relativos a valores e vetores próprios, aplicando-os à análise espectral da matriz  $\mathbf{M}$

$$\mathbf{M} = \sum_{j=1}^n \theta_j \gamma_j \gamma_j^t$$

(Schott 1997, p.105-106). Sejam  $\Omega_h$  e  $\Omega_h^\perp$  os espaços gerados pelos  $\gamma_1, \dots, \gamma_h$ , e pelos,  $\gamma_{h+1}, \dots, \gamma_n$ ,  $h=1, \dots, n-1$  respetivamente, então

$$\theta_h = \min_{x \in \Omega_h} \frac{x^t \mathbf{M} x}{x^t x}, \quad h=1, \dots, n-1$$

e

$$\theta_h = \max_{x \in \Omega_h^\perp} \frac{x^t \mathbf{M} x}{x^t x}, \quad h=1, \dots, n-1.$$

Ora, ao escolhermos o grau  $k'$  para o nosso modelo estamos a admitir

$$\mathbf{M} = \sum_{j=1}^{k'} \lambda_j \mathbf{a}_j \mathbf{a}_j^t + \bar{\mathbf{E}},$$

e a termos como modelo ajustado

$$\tilde{\mathbf{M}} = \sum_{j=1}^{k'} \theta_j \gamma_j \gamma_j^t,$$

o que, atendendo aos resultados acima referidos, pode ser interpretado como  $\tilde{\mathbf{M}}$  correspondendo a um ajustamento semelhante ao dos mínimos quadrados. Para aprofundar o que acabámos de referir recordamos que

$$\|\mathbf{M}\|^2 = \sum_{j=1}^n \theta_j^2.$$

A partir de  $\mathbf{M}$  e  $\tilde{\mathbf{M}}$  obtém-se a matriz de resíduos

$$\mathbf{R} = \mathbf{M} - \tilde{\mathbf{M}} = \sum_{j=k'+1}^n \theta_j \boldsymbol{\gamma}_j \boldsymbol{\gamma}_j^t .$$

Raciocinando-se como na secção 3.4, mostra-se que

$$\|\tilde{\mathbf{M}}\|^2 = \sum_{j=1}^{k'} \theta_j^2$$

e

$$\|\mathbf{R}\|^2 = \sum_{j=k'+1}^n \theta_j^2$$

tendo-se

$$\|\mathbf{M}\|^2 = \|\tilde{\mathbf{M}}\|^2 + \|\mathbf{R}\|^2 ,$$

o que justifica medir-se o ajustamento do modelo por

$$v = \frac{\|\tilde{\mathbf{M}}\|^2}{\|\mathbf{M}\|^2} = 1 - \frac{\|\mathbf{R}\|^2}{\|\mathbf{M}\|^2} = \frac{\sum_{j=1}^{k'} \theta_j^2}{\sum_{j=1}^n \theta_j^2} ,$$

com  $0 \leq v \leq 1$ .

De seguida iremos considerar um problema que muitas vezes se nos coloca, assim, pode ter-se um número  $h$  pequeno de valores próprios relevantes, isto é

$$\theta_{h+1} \geq \dots \geq \theta_{k'} > 0 ,$$

mas com  $\theta_{h+1}$  bastante pequeno. Caso

$$v_h = \frac{\sum_{j=1}^h \theta_j^2}{\sum_{j=1}^n \theta_j^2} ,$$

esteja suficientemente próximo de 1, podemos optar pelo modelo aproximado ajustado

$$\tilde{\mathbf{M}}_h = \sum_{j=1}^h \theta_j \boldsymbol{\gamma}_j \boldsymbol{\gamma}_j^t ,$$

Conseguindo-se assim um modelo mais leve. No entanto sendo  $Y_1, \dots, Y_g$  as componentes do vetor dos resíduos homocedástico, as mesmas terão, devido à eliminação incompleta da matriz média, valores médios  $\mu_1, \dots, \mu_g$  pelo que

$$\|\mathbf{Y}\|^2 \sim \sigma^2 \chi_{g,\delta}^2,$$

isto é,  $\|\mathbf{Y}\|^2$  distribuir-se-á como o produto por  $\sigma^2$  de um qui-quadrado com  $g$  graus de liberdade e parâmetro de não centralidade

$$\delta = \frac{1}{\sigma^2} \sum_{i=1}^g \mu_i^2 > 0.$$

Vejamos quais as consequências deste facto. Se quisermos utilizar a estatística

$$\mathcal{F}(\mathbf{b}) = \frac{g}{r} \frac{\varphi(\mathbf{b})}{s},$$

com

$$\varphi(\mathbf{b}) = (\tilde{\boldsymbol{\beta}} - \mathbf{b})^t (\mathbf{B} \mathbf{W} \mathbf{B}^t)^+ (\tilde{\boldsymbol{\beta}} - \mathbf{b}),$$

para testar

$$H_o : \boldsymbol{\beta} = \mathbf{b},$$

teremos uma estatística com distribuição  $F$  duplamente não central com  $r$  e  $g$  graus de liberdade e parâmetro de não centralidade

$$\delta(\mathbf{b}) = \frac{1}{\sigma^2} (\boldsymbol{\beta} - \mathbf{b})^t (\mathbf{B} \mathbf{W} \mathbf{B}^t)^+ (\boldsymbol{\beta} - \mathbf{b}),$$

com  $F(z | r, g, \delta(\mathbf{b}), \delta)$ . De forma a simplificar a escrita consideramos  $\delta_1 = \delta(\mathbf{b})$  e  $\delta_2 = \delta$ .

Ora, a distribuição do quociente de  $\frac{\chi_{r,\delta_1}^2}{\chi_{g,\delta_2}^2}$  é dada por



$$\bar{F}(z | r, g, \delta_1, \delta_2) = e^{-\frac{\delta_1 + \delta_2}{2}} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{\left(\frac{\delta_1}{2}\right)^i}{i!} \frac{\left(\frac{\delta_2}{2}\right)^j}{j!} \bar{F}(z | r + 2i, g + 2j),$$

podendo agrupar-se e reordenar-se os termos desta série dupla visto ter-se

$$0 \leq \bar{F}(z | r + 2i, g + 2j) < 1, \quad i, j = 0, \dots$$

Obtém-se assim

$$\bar{F}(z | r, g, \delta_1, \delta_2) = e^{-\frac{\delta_2}{2}} \sum_{j=0}^{\infty} \frac{\left(\frac{\delta_2}{2}\right)^j}{j!} \bar{F}(z | r, g + 2j, \delta_1),$$

com

$$\bar{F}(z | r, g + 2j, \delta_1) = e^{-\frac{\delta_1}{2}} \sum_{i=0}^{\infty} \frac{\left(\frac{\delta_1}{2}\right)^i}{i!} \bar{F}(z | r + 2i, g + 2j).$$

Temos então

$$\begin{aligned} \frac{\partial \bar{F}(z | r, g, \delta_1, \delta_2)}{\partial \delta_2} &= -\frac{1}{2} e^{-\frac{\delta_2}{2}} \sum_{j=0}^{\infty} \frac{\left(\frac{\delta_2}{2}\right)^j}{j!} \bar{F}(z | r, g + 2j, \delta_2) + \frac{e^{-\frac{\delta_2}{2}}}{2} \sum_{j=0}^{\infty} \frac{\left(\frac{\delta_2}{2}\right)^{j-1}}{(j-1)!} \bar{F}(z | r, g + 2j, \delta_1) \\ &= -\frac{1}{2} \bar{F}(z | r, g, \delta_1, \delta_2) + \frac{e^{-\frac{\delta_2}{2}}}{2} \sum_{j'=0}^{\infty} \frac{\left(\frac{\delta_2}{2}\right)^{j'}}{j'!} \bar{F}(z | r, g + 2 + 2j', \delta_1) \\ &= \frac{\bar{F}(z | r, g + 2, \delta_1, \delta_2) - \bar{F}(z | r, g, \delta_1, \delta_2)}{2}. \end{aligned}$$

Dados os qui-quadrados independentes  $\chi_{r, \delta_1}^2$ ,  $\chi_{g, \delta_2}^2$  e  $\chi_2^2$ , com  $\tau_1 = \frac{\chi_{r, \delta_1}^2}{\chi_{g, \delta_2}^2}$  e

$\tau_2 = \frac{\chi_{r, \delta_1}^2}{\chi_{g, \delta_2}^2 + \chi_2^2}$ , temos  $pr(\tau_2 < \tau_1) = 1$ . Como  $\tau_1$  têm distribuição  $\bar{F}(z | r, g, \delta_1, \delta_2)$  e

$\tau_2$  tem distribuição  $\bar{F}(z | r, g, \delta_1, \delta_2 + 2)$ , temos (para  $z > 0$ )

$$\bar{F}(z|r, g, \delta_1, \delta_2) < \bar{F}(z|r, g, \delta_1, \delta_2),$$

vindo (também para  $z > 0$ )

$$\frac{\partial \bar{F}(z|r, g, \delta_1, \delta_2)}{\partial \delta_2} > 0 .$$

Observe-se que, para  $z < 0$

$$\bar{F}(z|r, g, \delta_1, \delta_2) = \bar{F}(z|r, g, \delta_1, \delta_2) ,$$

vindo

$$\frac{\partial \bar{F}(z|r, g, \delta_1, \delta_2)}{\partial \delta_2} = 0.$$

Como

$$F(z|r, g, \delta_1, \delta_2) = \bar{F}\left(\frac{r}{g}z|r, g, \delta_1, \delta_2\right),$$

o teste  $F$  de nível  $q$  com estatística  $\mathcal{F}(\mathbf{b})$  terá função potência

$$Pot_q(\delta_1, \delta_2) = 1 - \bar{F}\left(\frac{c}{g}f_{1-q, r, g} | r, g, \delta_1, \delta_2\right),$$

decrecente com  $\delta_2$ . Assim, a consequência da truncatura do modelo será perda de potência dos testes  $F$ . No entanto, os resultados detetados como significativos continuam a sê-lo e se os valores próprios “eliminados” do modelo ajustado forem pequenos,  $\delta_2$  será pequeno e a perda de potência não será relevante. Referimos este ponto pois pode interessar ter um modelo compacto e fácil de interpretar.

Por outro lado, pode raciocinar-se como atrás para mostrar que para,  $z > 0$

$$\frac{\partial \bar{F}(z|r, g, \delta_1, \delta_2)}{\partial \delta_1} < 0 ,$$

pelo que,

$$\frac{\partial Pot_q(\delta_1, \delta_2)}{\partial \delta_1} > 0 ,$$

sendo pois os testes  $F$  não distorcidos, já que a potência do teste debaixo das hipóteses alternativas excede o nível do teste.

### 3.6. Relevâncias Relativas

Na secção anterior considerámos a truncatura do modelo ajustado pensando em situações em que havia valores próprios de  $\mathbf{M}$ ,  $\theta_1, \dots, \theta_h$  muito maiores que os restantes. Relativamente aos vetores  $\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_h$ , interessará considerar a relevância relativa dos mesmos, as quais podem ser medidas pelos

$$\mathbf{v}_j = \frac{\delta_j}{\sum_{j'=1}^h \delta_{j'}} = \frac{\boldsymbol{\beta}_j^t \mathbf{W}_j^{-1} \boldsymbol{\beta}_j}{\sum_{j'=1}^h \boldsymbol{\beta}_{j'}^t \mathbf{W}_{j'}^{-1} \boldsymbol{\beta}_{j'}}, \quad j = 1, \dots, h,$$

com

$$\mathbf{W}_j = (\mathbf{I}_n \otimes \boldsymbol{\gamma}_j^t) \mathbf{L} (\mathbf{I}_n \otimes \boldsymbol{\gamma}_j), \quad j = 1, \dots, h.$$

Como vimos as matrizes  $\mathbf{W}_j$ ,  $j = 1, \dots, h$ , são definidas positivas, existindo pois matrizes ortogonais  $\mathbf{P}_j$ ,  $j = 1, \dots, h$ , tais que

$$\mathbf{P}_j \mathbf{W}_j \mathbf{P}_j^t = \mathbf{D}(s_{j1}, \dots, s_{jn}), \quad j = 1, \dots, h,$$

onde  $\mathbf{D}(s_{j1}, \dots, s_{jn})$ ,  $j = 1, \dots, h$ , é a matriz diagonal cujos elementos principais são os valores próprios positivos  $s_{j1}, \dots, s_{jn}$  da matriz  $\mathbf{W}_j$ ,  $j = 1, \dots, h$ . Assim, estão definidas as matrizes

$$\mathbf{G}_j = \mathbf{D} \left( s_{j1}^{-\frac{1}{2}}, \dots, s_{jn}^{-\frac{1}{2}} \right) \mathbf{P}_j, \quad j = 1, \dots, h,$$

tendo-se,

$$\mathbf{G}_j \mathbf{W}_j \mathbf{P}_j^t = \mathbf{I}_n, \quad j = 1, \dots, h,$$

bem como

$$\mathbf{W}_j^{-1} = \mathbf{G}_j \mathbf{G}_j^t, \quad j = 1, \dots, h.$$

Considerando  $\dot{\boldsymbol{\beta}}_j = \mathbf{G}_j \boldsymbol{\beta}_j$ ,  $j = 1, \dots, h$  e  $\tilde{\boldsymbol{\beta}}_j = \mathbf{G}_j \tilde{\boldsymbol{\beta}}_j$ ,  $j = 1, \dots, h$ , vem

$$\boldsymbol{\beta}_j^t \mathbf{W}_j^{-1} \boldsymbol{\beta}_j = \dot{\boldsymbol{\beta}}_j^t \dot{\boldsymbol{\beta}}_j = \|\dot{\boldsymbol{\beta}}_j\|^2, \quad j = 1, \dots, h.$$

Ora, com  $\boldsymbol{\mu}_j = \dot{\boldsymbol{\beta}}_j$ ,  $j = 1, \dots, h$  e  $\mathbf{X}_j = \tilde{\boldsymbol{\beta}}_j - \dot{\boldsymbol{\beta}}_j$ ,  $j = 1, \dots, h$ , temos  $\tilde{\boldsymbol{\beta}}_j = \boldsymbol{\mu}_j + \mathbf{X}_j$ ,  $j = 1, \dots, h$  com  $\mathbf{X}_j \sim N(0, \sigma^2 \mathbf{I}_n)$ ,  $j = 1, \dots, h$ , bem como

$$\|\tilde{\boldsymbol{\beta}}_j\|^2 = g(\boldsymbol{\mu}_j + \mathbf{X}_j), \quad j = 1, \dots, h,$$

com  $g(\mathbf{x}) = \|\mathbf{x}\|^2$ .

Esta função tem gradiente e matriz Hessiana dadas por  $g_1(\mathbf{x}) = 2\mathbf{x}$  e  $g_2(\mathbf{x}) = 2\mathbf{I}$ , logo é assintoticamente linear já que  $\forall d > 0$ , se tem

$$\frac{1}{2\|\boldsymbol{\mu}\|} \text{Sup}\{ \|g_2(\boldsymbol{\mu} + \mathbf{u})\|; \|\mathbf{u}\| \leq d \} = \frac{1}{\|\boldsymbol{\mu}\|} \xrightarrow{\|\boldsymbol{\mu}\| \rightarrow +\infty} 0.$$

Logo a distribuição limite do vetor

$$\mathbf{y}_j^0 = \frac{\tilde{\boldsymbol{\beta}}_j^t \mathbf{W}_j^{-1} \tilde{\boldsymbol{\beta}}_j - \boldsymbol{\beta}_j^t \mathbf{W}_j^{-1} \boldsymbol{\beta}_j}{\sqrt{\boldsymbol{\beta}_j^t \mathbf{W}_j^{-1} \boldsymbol{\beta}_j}} = \frac{\|\tilde{\boldsymbol{\beta}}_j\|^2 - \|\boldsymbol{\beta}_j\|^2}{\|\boldsymbol{\beta}_j\|}, \quad j = 1, \dots, h,$$

é  $N(\cdot | 0, \sigma^2)$  (Mexia e Oliveira, 2011).

Sendo  $z_q^o$  o limite das caudas com probabilidade  $\frac{q}{2}$  da densidade normal estandardizada, teremos

$$\text{pr}(-z_q^o \sigma \leq \frac{\tilde{\boldsymbol{\beta}}_j^t \mathbf{W}_j^{-1} \tilde{\boldsymbol{\beta}}_j - \boldsymbol{\beta}_j^t \mathbf{W}_j^{-1} \boldsymbol{\beta}_j}{\sqrt{\boldsymbol{\beta}_j^t \mathbf{W}_j^{-1} \boldsymbol{\beta}_j}} \leq z_q^o \sigma) \xrightarrow{\|\boldsymbol{\mu}_j\| \rightarrow +\infty} 1 - q, \quad j = 1, \dots, h,$$

o que é equivalente à

$$pr(\boldsymbol{\beta}_j^t \mathbf{W}_j^{-1} \boldsymbol{\beta}_j - z_q^o \sigma \sqrt{\boldsymbol{\beta}_j^t \mathbf{W}_j^{-1} \boldsymbol{\beta}_j} \leq \tilde{\boldsymbol{\beta}}_j^t \mathbf{W}_j^{-1} \tilde{\boldsymbol{\beta}}_j \leq \boldsymbol{\beta}_j^t \mathbf{W}_j^{-1} \boldsymbol{\beta}_j + z_q^o \sigma \sqrt{\boldsymbol{\beta}_j^t \mathbf{W}_j^{-1} \boldsymbol{\beta}_j}) \xrightarrow{\|\boldsymbol{\mu}_j\| \rightarrow +\infty} 1 - q, \\ j = 1, \dots, h,$$

e, a

$$pr\left(1 - \frac{z_q^o \sigma}{\sqrt{\boldsymbol{\beta}_j^t \mathbf{W}_j^{-1} \boldsymbol{\beta}_j}} \leq \frac{\tilde{\boldsymbol{\beta}}_j^t \mathbf{W}_j^{-1} \tilde{\boldsymbol{\beta}}_j}{\boldsymbol{\beta}_j^t \mathbf{W}_j^{-1} \boldsymbol{\beta}_j} \leq 1 + \frac{z_q^o \sigma}{\sqrt{\boldsymbol{\beta}_j^t \mathbf{W}_j^{-1} \boldsymbol{\beta}_j}}\right) \xrightarrow{\|\boldsymbol{\mu}_j\| \rightarrow +\infty} 1 - q, j = 1, \dots, h,$$

bem como

$$pr\left(\frac{1}{1 + \frac{z_q^o \sigma}{\sqrt{\boldsymbol{\beta}_j^t \mathbf{W}_j^{-1} \boldsymbol{\beta}_j}}} \leq \frac{\boldsymbol{\beta}_j^t \mathbf{W}_j^{-1} \boldsymbol{\beta}_j}{\tilde{\boldsymbol{\beta}}_j^t \mathbf{W}_j^{-1} \tilde{\boldsymbol{\beta}}_j} \leq \frac{1}{1 - \frac{z_q^o \sigma}{\sqrt{\boldsymbol{\beta}_j^t \mathbf{W}_j^{-1} \boldsymbol{\beta}_j}}}\right) \xrightarrow{\|\boldsymbol{\mu}_j\| \rightarrow +\infty} 1 - q, j = 1, \dots, h.$$

Observemos agora que

$$1 - \frac{1}{1 + \frac{z_q^o \sigma}{\sqrt{\boldsymbol{\beta}_j^t \mathbf{W}_j^{-1} \boldsymbol{\beta}_j}}} = \frac{\frac{z_q^o \sigma}{\sqrt{\boldsymbol{\beta}_j^t \mathbf{W}_j^{-1} \boldsymbol{\beta}_j}}}{1 + \frac{z_q^o \sigma}{\sqrt{\boldsymbol{\beta}_j^t \mathbf{W}_j^{-1} \boldsymbol{\beta}_j}}} = \frac{z_q^o \sigma}{\sqrt{\boldsymbol{\beta}_j^t \mathbf{W}_j^{-1} \boldsymbol{\beta}_j} + z_q^o \sigma}, j = 1, \dots, h,$$

e

$$\frac{1}{1 - \frac{z_q^o \sigma}{\sqrt{\boldsymbol{\beta}_j^t \mathbf{W}_j^{-1} \boldsymbol{\beta}_j}}} - 1 = \frac{\frac{z_q^o \sigma}{\sqrt{\boldsymbol{\beta}_j^t \mathbf{W}_j^{-1} \boldsymbol{\beta}_j}}}{1 - \frac{z_q^o \sigma}{\sqrt{\boldsymbol{\beta}_j^t \mathbf{W}_j^{-1} \boldsymbol{\beta}_j}}} = \frac{z_q^o \sigma}{\sqrt{\boldsymbol{\beta}_j^t \mathbf{W}_j^{-1} \boldsymbol{\beta}_j} - z_q^o \sigma}, j = 1, \dots, h,$$

vindo a

$$pr\left(1 - \frac{z_q^o \sigma}{\sqrt{\boldsymbol{\beta}_j^t \mathbf{W}_j^{-1} \boldsymbol{\beta}_j} + z_q^o \sigma} \leq \frac{\boldsymbol{\beta}_j^t \mathbf{W}_j^{-1} \boldsymbol{\beta}_j}{\tilde{\boldsymbol{\beta}}_j^t \mathbf{W}_j^{-1} \tilde{\boldsymbol{\beta}}_j} \leq 1 + \frac{z_q^o \sigma}{\sqrt{\boldsymbol{\beta}_j^t \mathbf{W}_j^{-1} \boldsymbol{\beta}_j} - z_q^o \sigma}\right) \xrightarrow{\|\boldsymbol{\mu}_j\| \rightarrow +\infty} 1 - q, j = 1, \dots, h.$$

Ora

$$\|\boldsymbol{\mu}_j\| = \|\dot{\boldsymbol{\beta}}_j\| = \sqrt{\boldsymbol{\beta}_j^t \mathbf{W}_j^{-1} \boldsymbol{\beta}_j}, j = 1, \dots, h,$$

pelo que

$\|\hat{\boldsymbol{\mu}}_j\| \rightarrow \infty$ ,  $j=1, \dots, h$ , é equivalente a  $\sqrt{\boldsymbol{\beta}_j^t \mathbf{W}_j^{-1} \boldsymbol{\beta}_j} \rightarrow +\infty$ ,  $j=1, \dots, h$ .

Tem-se ainda que

$$\frac{\frac{z_q^o \sigma}{\sqrt{\boldsymbol{\beta}_j^t \mathbf{W}_j^{-1} \boldsymbol{\beta}_j \pm z_q^o \sigma}}}{\frac{1}{\sqrt{\boldsymbol{\beta}_j^t \mathbf{W}_j^{-1} \boldsymbol{\beta}_j}}} \xrightarrow{\|\hat{\boldsymbol{\mu}}_j\| \rightarrow +\infty} z_q^o \sigma, \quad j=1, \dots, h,$$

podendo dizer-se que o intervalo de confiança, com nível limite  $1-q$  quando

$$\|\hat{\boldsymbol{\mu}}_j\| = (\boldsymbol{\beta}_j^t \mathbf{W}_j^{-1} \boldsymbol{\beta}_j)^{\frac{1}{2}} \rightarrow +\infty, \quad j=1, \dots, h,$$

tem semi-amplitude proporcional a  $\|\hat{\boldsymbol{\mu}}_j\|$ ,  $j=1, \dots, h$ , logo

$$\frac{\tilde{\boldsymbol{\beta}}_j^t \mathbf{W}_j^{-1} \tilde{\boldsymbol{\beta}}_j}{\boldsymbol{\beta}_j^t \mathbf{W}_j^{-1} \boldsymbol{\beta}_j} \xrightarrow[\|\hat{\boldsymbol{\mu}}_j\| \rightarrow \infty]{p} 1, \quad j=1, \dots, h,$$

onde  $\xrightarrow{p}$  indica convergência estocástica e, conseqüentemente

$$\tilde{\boldsymbol{\beta}}_j^t \mathbf{W}_j^{-1} \tilde{\boldsymbol{\beta}}_j \xrightarrow[\|\hat{\boldsymbol{\mu}}_j\| \rightarrow \infty]{p} \boldsymbol{\beta}_j^t \mathbf{W}_j^{-1} \boldsymbol{\beta}_j, \quad j=1, \dots, h.$$

Com

$$\|\hat{\boldsymbol{\mu}}\| = \text{Min}\{\|\hat{\boldsymbol{\mu}}_j\|\}, \quad j=1, \dots, h,$$

e,  $\tilde{\boldsymbol{v}}$  e  $\boldsymbol{v}$  os vetores com componentes  $\tilde{\boldsymbol{\beta}}_j^t \mathbf{W}_j^{-1} \tilde{\boldsymbol{\beta}}_j$  e  $\boldsymbol{\beta}_j^t \mathbf{W}_j^{-1} \boldsymbol{\beta}_j$ ,  $j=1, \dots, h$ , temos  $\tilde{\boldsymbol{v}} \xrightarrow[\|\hat{\boldsymbol{\mu}}\| \rightarrow \infty]{p} \boldsymbol{v}$ .

Com efeito, quaisquer que sejam  $q < 1$  e  $\varepsilon > 0$  existe  $k$  tal que, para  $\|\hat{\boldsymbol{\mu}}\| > k$ , se tem

$$pr(|\tilde{v}_j - v_j| \leq \varepsilon) \geq 1 - q, \quad j=1, \dots, h,$$

vindo, quando  $\|\hat{\boldsymbol{\mu}}\| > k$

$$pr(\|\tilde{\boldsymbol{v}} - \boldsymbol{v}\| \leq \sqrt{h} \varepsilon) \geq 1 - hq,$$

uma vez que atendendo às desigualdades de BOOLE

$$pr \left[ \bigcap_{j=1}^h (|\tilde{v}_j - v_j| \leq \varepsilon) \right] \geq \sum_{j=1}^h pr(|\tilde{v}_j - v_j| \leq \varepsilon) - (h-1) \geq 1 - hq ,$$

e que

$$\bigcap_{j=1}^h (|\tilde{v}_j - v_j| \leq \varepsilon) \Rightarrow \|\tilde{v} - v\| \leq \sqrt{h} \varepsilon .$$

Ora  $v_{j'}$  é dada por uma função contínua  $l(v)$  de  $v$  logo,

$$\forall \delta > 0, \exists \varepsilon > 0: \|v' - v\| \leq \sqrt{h} \varepsilon \Rightarrow (|l(v') - l(v)| \leq \delta) .$$

Assim, quando  $\|\hat{\mu}\| > k$ , temos  $pr(|\tilde{v}_{j'} - v_{j'}| \leq \delta) \geq 1 - hq$ , com

$$\tilde{v}_{j'} = \frac{\tilde{\beta}_{j'}^t \mathbf{W}_{j'}^{-1} \tilde{\beta}_{j'}}{\sum_{j=1}^h \tilde{\beta}_j^t \mathbf{W}_j^{-1} \tilde{\beta}_j} \xrightarrow[\|\hat{\mu}\| \rightarrow \infty]{p} v_{j'} , \quad j' = 1, \dots, h$$

e, portanto  $\tilde{v} \xrightarrow[\|\hat{\mu}\| \rightarrow \infty]{p} v$ .

Nas secções anteriores apresentamos modelos para matrizes estocásticas simétricas. Apresentaremos de seguida duas aplicações a casos especiais destas matrizes.

### 3.7 Aplicação à Inter-estrutura

Na Análise Conjunta de Tabelas (ACT), é habitual dispor-se de várias tabelas de dados de dupla entrada, os dados que compõem estas tabelas são de natureza quantitativa, podendo estas ser obtidas em diferentes situações. Cada tabela é composta por  $n$  linhas (objetos) e  $p$  colunas (variáveis). As várias técnicas desenvolvidas para a análise deste tipo de tabelas remontam às décadas de 70 e 80, nomeadamente encontram-se estudos destas técnicas em Escoufier e L'Hermier des Plantes (1976), Foucart (1981) entre outros. As várias técnicas desenvolvidas permitem efetuar uma comparação global de todas as tabelas, bem como o estudo de uma eventual existência de uma estrutura comum, a que se dá o nome de inter-estrutura. A inter-estrutura é a primeira etapa da metodologia STATIS (método STATIS e STATIS Dual). Esta metodologia foi introduzida por Hermier des Plantes (1976), e foi desenvolvida em particular por Lavit

(1988). A metodologia STATIS utiliza distâncias euclidianas entre configurações observadas em  $k$  situações diferentes a que chamamos estudos. Sempre que, as variáveis diferem ao longo dos  $k$  estudos e o número de objetos é fixo, existindo uma proximidade entre os indivíduos (objetos) dizemos que estamos no âmbito do método STATIS. Sempre que ocorre o inverso, isto é, o número de variáveis é fixo e os objetos podem mudar ao longo dos estudos, dizemos que estamos no âmbito do método STATIS Dual. Nestes dois métodos estudam-se séries de estudos em que se consideram trios de matrizes  $(\mathbf{X}, \mathbf{D}, \dot{\mathbf{D}})$ , constituídos por uma matriz de dados  $\mathbf{X}$  e duas matrizes de pesos, uma para objetos e outra para variáveis,  $\mathbf{D}$  e  $\dot{\mathbf{D}}$ , respetivamente. Quando os objetos são os mesmos para todos os estudos numa série a mesma diz-se do primeiro tipo. Quando as variáveis são as mesmas para todos os estudos numa série a mesma diz-se do segundo tipo. Escoufier (1978) definiu os operadores  $\mathbf{A}_i = \mathbf{X}_i \dot{\mathbf{D}}_i \mathbf{X}_i^t \mathbf{D}_i$ ,  $i = 1, \dots, n$  e  $\mathbf{B}_i = \mathbf{X}_i^t \mathbf{D}_i \mathbf{X}_i \dot{\mathbf{D}}_i$ ,  $i = 1, \dots, n$  para representar os estudos  $(\mathbf{X}_i, \mathbf{D}_i, \dot{\mathbf{D}}_i)$ ,  $i = 1, \dots, n$ , no caso de séries de primeiro e do segundo tipo, respetivamente.

No resto da secção tomaremos  $\mathbf{U} = \mathbf{A}$ , no caso de séries de estudo do primeiro tipo e  $\mathbf{U} = \mathbf{B}$ , no caso de séries do segundo tipo. Para evitar duplicações vamos considerar o estudo das séries do primeiro tipo. Temos então as matrizes  $\mathbf{S} = [s_{ij}]$  com  $s_{ij} = tr(\mathbf{U}_i \mathbf{U}_j^t)$ ,  $i, j = 1, \dots, n$  o produto de Hilbert-Schmidt dos operadores  $\mathbf{U}_i$  e  $\mathbf{U}_j$ ,  $i, j = 1, \dots, n$ .

Sendo  $\theta_1 \geq \dots \geq \theta_n$  os valores próprios e  $\gamma_1, \dots, \gamma_n$  os correspondentes vetores próprios, Escoufier (1978) considerou uma representação geométrica para os estudos. O  $j$ -ésimo estudo era representado pelo ponto cujas coordenadas eram as  $j$ -ésimas componentes dos vetores  $\frac{1}{\theta_1^2} \gamma_1, \dots, \frac{1}{\theta_n^2} \gamma_n$ . Quando estes pontos estiverem ao longo do primeiro eixo com a direção de  $\gamma_1$ ,  $\theta_1$  predominará sobre os restantes valores próprios e a série tem estrutura de grau 1 (Lavit, 1988). Nesse caso pode admitir-se ter

$$\mathbf{S} = \lambda \boldsymbol{\alpha} \boldsymbol{\alpha}^t + \bar{\mathbf{E}},$$

tendo-se pois uma matriz estocástica simétrica com grau 1. Este modelo é um caso particular do modelo

$$\mathbf{M} = \mathbf{S} = \sum_{j=1}^k \lambda_j \boldsymbol{\alpha}_j \boldsymbol{\alpha}_j^t + \bar{\mathbf{E}},$$



de grau  $k$ , estudado na secção anterior. Admite-se então que

$$\mathbf{U}_i = \mathbf{U}_i^o + \mathbf{E}_i, \quad i = 1, \dots, n,$$

com  $\mathbf{U}_i^o = [u_{ilh}]$  e  $\mathbf{E}_i = [e_{ilh}]$  com  $\text{vec}(\mathbf{E}_i) \sim N(0, \sigma^2 \mathbf{I}_{m^2})$ ,  $i = 1, \dots, n$  (Areia, 2009).

Admite-se ainda que os  $\text{vec}(\mathbf{E}_i)$ ,  $i = 1, \dots, n$ , são independentes. Então

$$\begin{aligned} S_{ij} &= \text{tr}(\mathbf{U}_i \mathbf{U}_j^t) = \sum_{l=1}^m \sum_{h=1}^m (u_{ilh} + e_{ilh})(u_{jlh} + e_{jlh}) \\ &= \sum_{l=1}^m \sum_{h=1}^m (u_{ilh} u_{jlh} + u_{ilh} e_{jlh} + u_{jlh} e_{ilh} + e_{ilh} e_{jlh}), \\ &= S_{ij}^o + \sum_{l=1}^m \sum_{h=1}^m (e_{ilh} e_{jlh}), \quad i, j = 1, \dots, n \end{aligned}$$

com  $S_{ij}^o = \sum_{l=1}^m \sum_{h=1}^m (u_{ilh} u_{jlh} + u_{ilh} e_{jlh} + u_{jlh} e_{ilh})$ ,  $i, j = 1, \dots, n$ .

Ora, com  $\mathbf{S}^o = [s_{ij}^o]$ ,  $\text{vec}(\mathbf{S}^o)$  estará normalmente distribuído. Podemos pois admitir que  $\mathbf{Z} = \text{vec}(\mathbf{S})$ , estará (aproximadamente) normalmente distribuído desde que se possam desprezar  $\sum_{l=1}^m \sum_{h=1}^m (e_{ilh} e_{jlh})$ ,  $i, j = 1, \dots, n$ . Simulações efetuadas (Apêndice A,

p.85-95) mostram que quando os coeficientes de variação  $\frac{\sigma_j}{\mu_j}$ , dum par de variáveis

normais  $\mathbf{X}_j$ ,  $j = 1, 2$  são inferiores a  $\frac{1}{17}$ , o produto  $\mathbf{X}_1 \mathbf{X}_2$  pode ser tratado como

estando normalmente distribuído. Assim quando os elementos das matrizes  $\mathbf{U}_1, \dots, \mathbf{U}_n$

tiverem coeficientes de variação inferior a  $\frac{1}{17}$  podemos tratar  $\mathbf{Z}$  como normal.

Com

$$\mathbf{K}_{ij} = \text{tr}(\mathbf{U}_i^o \mathbf{U}_j^{o t}) = \sum_{l=1}^m \sum_{h=1}^m (u_{ilh} + u_{jlh}), \quad i, j = 1, \dots, n,$$

dada a independência e homocedasticidade dos  $\text{vec}(\mathbf{E}_i)$ ,  $i = 1, \dots, n$ , temos

$$\begin{aligned} \text{Var}(S_{ij}^o) &= \text{Var}\left(\sum_{l=1}^m \sum_{h=1}^m u_{ilh} + e_{jlh} + u_{jlh} + e_{ilh}\right) = \sigma^2 \sum_{l=1}^m \sum_{h=1}^m (u_{ilh}^2 + u_{jlh}^2), \\ &= \sigma^2 (K_{ii} + K_{jj}), \quad i \neq j \end{aligned}$$

bem como

$$\begin{aligned} \text{Var}(S_{ii}^o) &= \text{Var}\left(2\sum_{l=1}^m\sum_{h=1}^m u_{ilh}e_{ilh}\right) = 4\sigma^2\sum_{l=1}^m\sum_{h=1}^m u_{ilh}^2 \\ &= 4\sigma^2 K_{ii}, \quad i = 1, \dots, n. \end{aligned}$$

Os restantes elementos da matriz de variâncias-covariâncias  $\mathbf{L}$  são (aproximadamente) dadas por (Aníbal, 2009).

$$\begin{cases} \text{Cov}(S_{ii}, S_{ij}) = \text{Cov}(S_{ii}, S_{ji}) = 2\sigma^2 K_{ij}, & i \neq j \\ \text{Cov}(S_{ii}, S_{ji}) = 0 \\ \text{Cov}(S_{ij}, S_{ji}) = \text{Var}(S_{ij}) = \sigma^2(K_{ii} + K_{jj}), & i \neq j. \\ \text{Cov}(S_{ij}, S_{il}) = \text{Cov}(S_{ij}, S_{li}) \approx \sigma^2 K_{jl}, & i \neq j \neq l \\ \text{Cov}(S_{ij}, S_{lu}) = 0, & i \neq j \neq l \neq u \end{cases}$$

Podemos então admitir (com a devida aproximação) que

$$\mathbf{Z} \sim N(\boldsymbol{\eta}, \sigma^2 \mathbf{L})$$

com  $\boldsymbol{\eta} = \text{vec}([\mathbf{K}])$  e  $\mathbf{K} = [k_{ij}]$ , já que as componentes de  $\text{vec}(\mathbf{S}^o)$  sendo funções lineares dum vetor normal terão distribuição conjunta normal.

### 3.8 Aplicação a Matrizes de Produtos Cruzados

Dada a matriz de dados  $\mathbf{X}$  do tipo  $v \times d$  as matrizes  $\mathbf{X}\mathbf{X}^t$  e  $\mathbf{X}^t\mathbf{X}$  são as matrizes de produtos cruzados para vetor linha e coluna de  $\mathbf{X}^t\mathbf{X}$ . Em particular se os vetores coluna de  $\mathbf{X}$  forem os  $\text{vec}(\mathbf{U}_i)$ ,  $i = 1, \dots, d$ ,  $\mathbf{X}^t\mathbf{X}$  será a matriz dos produtos de Hilbert-Schmidt das matrizes  $\mathbf{U}_1, \dots, \mathbf{U}_d$ . Assim estes modelos generalizam diretamente os da alínea anterior, sendo fácil de verificar que o tratamento apresentado se estende diretamente.

No caso da matriz  $\mathbf{X}\mathbf{X}^t$  devemos ter

$$k_{ij} = \sum_{l=1}^d x_{il}x_{jl}, \quad i, j = 1, \dots, m,$$

com  $\mathbf{X} = [x_{ih}]$ . Para a matriz  $\mathbf{X}^t\mathbf{X}$  teremos

$$k_{ij} = \sum_{l=1}^m x_{li}x_{lj}, \quad i, j = 1, \dots, d .$$

Observe-se agora que as matrizes  $\mathbf{X}\mathbf{X}'$  e  $\mathbf{X}'\mathbf{X}$  têm os mesmos valores próprios não nulos  $\theta_1 \geq \dots \geq \theta_r > 0$ , com  $r = \text{car}(\mathbf{X}) = \text{car}(\mathbf{X}\mathbf{X}') = \text{car}(\mathbf{X}'\mathbf{X})$ . Assim, os modelos para  $\mathbf{X}\mathbf{X}'$  e  $\mathbf{X}'\mathbf{X}$  terão o mesmo grau  $k'$ . Admitindo que  $d < v$ , deve-se realizar o ajustamento do modelo para  $\mathbf{X}'\mathbf{X}$  que será uma matriz  $d \times d$  em vez de  $\mathbf{X}\mathbf{X}'$  que é  $v \times v$ . Aos  $\dot{\gamma}_1, \dots, \dot{\gamma}_{k'}$  primeiros  $k'$  vetores próprios para  $\mathbf{X}'\mathbf{X}$  corresponderão, como vetores próprios de  $\mathbf{X}\mathbf{X}'$ , os

$$\gamma_i = \frac{1}{\sqrt{\theta_i}} \mathbf{X} \dot{\gamma}_i, \quad i = 1, \dots, k',$$

que poderão ser utilizados na inferência para os modelos  $\mathbf{X}\mathbf{X}'$  e suas famílias.

Como  $V = \|\mathbf{Y}\|^2$  é independente dos

$$\dot{\beta}_i = (\mathbf{X}'\mathbf{X})\dot{\gamma}_i, \quad i = 1, \dots, k',$$

será independente dos

$$\beta_i = \mathbf{X}\dot{\beta}_i, \quad i = 1, \dots, k',$$

pelo que os resultados para o modelo correspondente a  $\mathbf{X}'\mathbf{X}$  se transportam diretamente para o modelo correspondente a  $\mathbf{X}\mathbf{X}'$ .

Em muitos casos podem utilizar-se modelos de grau um para estas matrizes (Oliveira 2001, Ramos, 2007 e Areia, 2009).

## 3.9 Famílias Estruturadas

### 3.9.1 Caso Geral

Recordamos que no caso geral a informação contida numa matriz estocástica simétrica é condensada no par  $(\tilde{\beta}, V)$ , podendo admitir-se que  $\tilde{\beta} \sim N(\beta, \sigma^2 \mathbf{H})$ ,

(aproximadamente) com  $\boldsymbol{\beta} = \mathbf{B}\boldsymbol{\eta}$  e  $\mathbf{H} = \mathbf{B}\mathbf{L}\mathbf{B}'$  e a variável aleatória  $V \sim \sigma^2\chi_g^2$  sendo  $\tilde{\boldsymbol{\beta}}$  independente de  $V$ .

No caso de uma família estruturada de matrizes temos os pares  $(\tilde{\boldsymbol{\beta}}_i, V_i), i = 1, \dots, m$ . Sendo as matrizes que compõem a família independentes, também serão independentes os pares  $(\tilde{\boldsymbol{\beta}}_i, V_i), i = 1, \dots, m$ . Devido á reprodutividade dos qui-quadrados, caso se tenham as

variáveis aleatórias,  $V_i \sim \sigma^2\chi_{g_i}^2, i = 1, \dots, m$ , ter-se-á  $V \sim \sigma^2\chi_g^2$ , com  $g = \sum_{i=1}^m g_i$ ,

independentes de

$$\tilde{\boldsymbol{\beta}} = [\tilde{\boldsymbol{\beta}}_1^t, \dots, \tilde{\boldsymbol{\beta}}_m^t]^t,$$

podendo admitir-se que,  $\tilde{\boldsymbol{\beta}} \sim N(\boldsymbol{\beta}, \sigma^2\mathbf{H})$  com

$$\boldsymbol{\beta} = [\boldsymbol{\beta}_1^t, \dots, \boldsymbol{\beta}_m^t]^t,$$

e

$$\mathbf{H} = \mathbf{D}(\mathbf{H}_1, \dots, \mathbf{H}_m).$$

Como referimos anteriormente o estudo conjunto dos modelos numa família estruturada analisa a ação dos fatores do delineamento base sobre os parâmetros dos modelos. Neste caso interessa-nos estudar a ação sobre os vetores de estrutura  $\tilde{\boldsymbol{\beta}}_i, i = 1, \dots, m$  tendo-se  $w = hn'$  componentes com  $n'$  a ordem das matrizes que constituem as famílias.

Para realizar inferência, escolhe-se um vetor  $\mathbf{a}$  com  $w$  componentes, trabalhando-se com o par  $(\mathbf{z}(\mathbf{a}), V)$ . As componentes de  $\mathbf{z}(\mathbf{a})$  serão

$$\mathbf{z}_i(\mathbf{a}) = \mathbf{a}^t \tilde{\boldsymbol{\beta}}_i \sim N(\mathbf{a}^t \boldsymbol{\beta}_i, \sigma^2 \mathbf{a}^t \mathbf{H}_i \mathbf{a}), i = 1, \dots, m,$$

independentes entre si e de  $V$ . Assim, temos  $\mathbf{z}(\mathbf{a}) \sim N(\boldsymbol{\mu}, \sigma^2\mathbf{C})$ , as componentes de  $\boldsymbol{\mu}$  serão

$$\boldsymbol{\mu}_i = \mathbf{a}^t \boldsymbol{\beta}_i, i = 1, \dots, m,$$

tendo-se ainda

$$\mathbf{C} = \mathbf{D}(\mathbf{a}^t \mathbf{H}_1 \mathbf{a}, \dots, \mathbf{a}^t \mathbf{H}_m \mathbf{a}).$$

Observe-se que se  $\mathbf{a} = \boldsymbol{\delta}_i, i = 1, \dots, n$ , com  $\boldsymbol{\delta}_i, i = 1, \dots, n'$  o vetor com componentes todas nulas salvo a  $i$ -ésima que é 1, a análise incidirá sobre as  $i$ -ésimas componentes dos vetores de estrutura dizendo-se que se trata duma AT. Além da análise transversal pode interessar analisar uma AL. Neste caso,  $\mathbf{a}$  será um vetor de contrastes sendo nula a soma das suas componentes.

Quando o grau do modelo é um podemos estar interessados nos respetivos primeiros vetores de estrutura. Nesse caso, no estudo da AT restringimo-nos aos  $\boldsymbol{\delta}_i, i = 1, \dots, n'$ , e no caso da AL, as componentes não nulas de  $\mathbf{a}$  têm índices que não excedem  $n'$ .

Se as componentes dos primeiros vetores de estrutura correspondem a instantes sucessivos a AL incidirá sobre a evolução no tempo. Por exemplo, no caso particular das séries de estudos consideradas na metodologia STATIS, os estudos podem suceder-se no tempo. Em Oliveira (2002), Ramos (2007) e Areia (2009), foram analisados resultados de sucessivas eleições autárquicas em Portugal Continental, nos períodos de 1976 a 1997 e 1998 a 2001 respetivamente. Nestes casos, foi possível estudar a evolução dos resultados eleitorais, através da Análise Longitudinal.

A concluir, observamos que dado a matriz  $\mathbf{C}$  ser diagonal os resultados atrás apresentados para a heterocedasticidade controlada terão a sua aplicabilidade facilitada.

### 3.9.2 Modelos de Grau um

Quando a predominância do primeiro valor próprio é muito grande, podemos adotar um modelo de grau um da forma

$$\mathbf{M} = \lambda \boldsymbol{\alpha} \boldsymbol{\alpha}' + \bar{\mathbf{E}}$$

onde  $\bar{\mathbf{E}}$  tem matriz média nula.

No mais simples destes modelos tem-se  $\bar{\mathbf{E}} = \frac{1}{2}(\mathbf{E} + \mathbf{E}^t)$  com  $\text{vec}(\mathbf{E}) \sim N(0, \sigma^2 \mathbf{I}_{n^2})$ .

Estes modelos foram utilizados com êxito em várias aplicações, nomeadamente á primeira fase da metodologia STATIS, a inter-estrutura. Dado que as matrizes de Hilbert- Schmidt são matrizes de produtos cruzados, podemos utiliza-los para estas últimas, passando a modelos de maior grau se os de grau um não se ajustarem.

Este tipo de modelos encontra-se estudado mais detalhadamente em Oliveira (2002). No que segue basear-nos-emos em apresentar os principais resultados que se têm para os mesmos.

Sejam  $\theta_1 \geq \dots \geq \theta_n$  os valores próprios e  $\gamma_1, \dots, \gamma_n$  os vetores próprios de  $\mathbf{M}$ . Suponhamos que  $\theta_1 \gg \theta_2$ , passando-se a admitir que  $\theta_j \approx 0, j = 2, \dots, n$ . A validação do modelo pode ser decomposta em duas fases. Na primeira mostra-se que:

- o grau do modelo não excede um;
- $vec(\mathbf{E}) \sim N(0, \sigma^2 \mathbf{I}_{n^2})$ .

Para verificação dos pressupostos anteriores, considera-se uma matriz

$$\mathbf{W} = \mathbf{P}^o \mathbf{M} \mathbf{P}^{o^t} = [w_{ln}],$$

com  $\mathbf{P}^o$  matriz ortogonal com primeiro vetor linha  $\gamma_1$ . Mostra-se que, quando o grau do modelo não excede 1 e  $vec(\mathbf{E}) \sim N(0, \sigma^2 \mathbf{I}_{n^2})$ , também se verifica

$$H_0: Y_1, \dots, Y_g \text{ i.i.d. } \sim N(0, \sigma^2),$$

com as variáveis aleatórias  $Y_1 = w_{11}, \dots, Y_k = w_{kk}, Y_{k+1} = \sqrt{2}w_{2n} \dots Y_g = \sqrt{2} w_{(n-1)n}$ , sendo  $g = \frac{n(n+1)}{2}$ .

Para testar a hipótese  $H_0$  pode utilizar-se um teste com estatística

$$\mathcal{F} = \frac{(g-1) \left( \sum_{i=1}^g Y_i \right)^2}{g \sum_{i=1}^g Y_i^2 - \left( \sum_{i=1}^g Y_i \right)^2},$$

introduzido por Mexia (1959, p.42). Esta estatística servirá para testar a hipótese  $H_0$  contra hipóteses alternativas em que as  $Y_1, \dots, Y_g$  são variáveis normais independentes

com variância  $\sigma^2$  e valores médios  $\mu_1, \dots, \mu_g$ . Com  $\bar{\mu} = \frac{1}{g} \sum_{i=1}^g \mu_i$ ,  $\delta_1 = g \frac{\bar{\mu}^2}{\sigma^2}$  e

$$\delta_2 = \frac{1}{\sigma^2} \sum_{i=1}^g (\mu_i - \bar{\mu})^2.$$

A estatística  $\mathcal{F}$  distribui-se como o quociente  $\frac{\chi_{1,\delta_2}^2}{\chi_{g-1,\delta_1,\delta_2}^2}$  de qui-quadrados independentes, pondo-se  $\mathcal{F} \sim \bar{F}( |1, g-1, \delta_1, \delta_2)$ .

Ora, existem alternativas à hipótese  $H_0$  em que  $\delta_1$  predomina sobre  $\delta_2$ , e em que  $\mathcal{F}$  tende a tomar valores superiores aos que tomaria, caso a hipótese  $H_0$  se verificasse. Por outro lado, haverá alternativas à hipótese  $H_0$  em que  $\delta_2$  predomina sobre  $\delta_1$ , e em que  $\mathcal{F}$  tende a tomar valores inferiores aos que tomaria, caso  $H_0$  se verificasse, podendo aliás reescrever-se a hipótese nula como

$$H_0 : \delta_1 = \delta_2 = 0.$$

Assim, sendo  $\bar{f}_{\frac{q}{2}} \left[ \bar{f}_{1-\frac{q}{2}} \right]$  o quantil para a probabilidade  $\frac{q}{2} \left[ 1 - \frac{q}{2} \right]$  de  $\bar{F}(.|1, g-1, \delta_1, \delta_2)$ ,

para um teste de nível  $q$  pode tomar-se a região de rejeição  $\left] 0, \bar{f}_{\frac{q}{2}} \left[ \cup \right] \bar{f}_{1-\frac{q}{2}} ; +\infty \left[ .$

Assim, tem-se um teste de nível  $q$  cuja potência decresce com  $\delta_1$  e  $\delta_2$  (Mexia, 1989).

Esta primeira fase da validação considera-se satisfeita quando  $H_0$  não é rejeitada.

Na segunda fase da validação do modelo testa-se a hipótese nula

$$H_0^0 : \theta_1 = 0,$$

pode então utilizar-se o teste com estatística

$$F = \frac{k-1}{2} \frac{\theta_1^2}{V},$$

com  $V = \sum_{i=2}^k \theta_i^2$ . Quando a hipótese  $H_0^0$  se verifica, esta estatística tem distribuição

$F(.|k, l)$  com  $k$  e  $l = \frac{k(k-1)}{2}$  graus de liberdade (Oliveira, 2002). A segunda fase da

validação considera-se satisfeita quando a hipótese  $H_0^0$  não é rejeitada.

### 3.9.3 Estudo de Simulação

Nestes modelos temos o mesmo número  $g'$  de graus de liberdade para o erro para todas as matrizes. Estamos pois no caso equilibrado em que a Análise de Variância e as técnicas a ela associadas são robustas, ver Scheffé (1959) e Ito (1980), para a heterocedasticidade e ainda mais para a não normalidade. Nos modelos de grau um tem-se a matriz de covariâncias dada por

$$\Sigma(\tilde{\boldsymbol{\beta}}_i) = \sigma^2 (\mathbf{I}_{n'} + \gamma_{1i} \boldsymbol{\gamma}_{1i}^t), \quad i = 1, \dots, m,$$

como  $\|\boldsymbol{\gamma}_{1i}\| = 1$  a heterocedasticidade dos  $\tilde{\boldsymbol{\beta}}_i$  é pouco acentuada. Assim, atendendo á robustez da Análise de Variância nos modelos equilibrados admite-se que os  $\tilde{\boldsymbol{\beta}}_1, \dots, \tilde{\boldsymbol{\beta}}_m$  têm matrizes de covariância  $\sigma^2 \mathbf{I}_{n'}$  independentes das somas de quadrados de resíduos  $V_1, \dots, V_m$ , tendo-se  $V_i \sim \sigma^2 \chi_l^2, i = 1, \dots, m$ . Assim, ter-se-á

$$\tilde{\boldsymbol{\beta}}_i = [\tilde{\boldsymbol{\beta}}_1^t, \dots, \tilde{\boldsymbol{\beta}}_m^t] \sim N(\boldsymbol{\beta}, \sigma^2 \mathbf{I}_{mn'}),$$

que é independente de

$$V = \sum_{i=1}^m V_i \sim \sigma^2 \chi_g^2,$$

com  $\boldsymbol{\beta} = [\boldsymbol{\beta}_1^t, \dots, \boldsymbol{\beta}_m^t]^t$  e  $g = ml$ . Um estudo de simulação pode ser utilizado para conferir uma validação adicional a estes modelos. Em particular garante obter condições em que  $\theta_1$  e  $\gamma_1$  podem ser utilizados como estimadores de  $\lambda$  e de  $\alpha$  respetivamente. Com

$\tau^0 = \frac{\theta_1^2}{\sum_{i=2}^k \theta_i^2}$ , obtemos os valores médios e os desvios padrão estimados por

$$\mathbf{Z} = \|\boldsymbol{\alpha} - \boldsymbol{\gamma}\|^2.$$



Assim, na Tabela 3.2 apresentamos resultados relativos a lotes de 1000 simulações e, os respectivos valores médios e desvio padrão para  $Z = \|\alpha - \gamma\|^2$ . Os valores de  $k$  escolhidos foram: 6, 8, 10, 12 e os valores de  $\tau^0$  foram 12.5, 50, 200 e 400.

Tabela 3.2 Valores médios e desvios padrão para Z.

$\tau^0$	$k = 6$		$k = 8$		$k = 10$		$k = 12$	
	Média	D.P	Média	D.P	Média	D.P	Média	D.P
12.5	0.957	0.574	1.101	0.571	1.253	0.536	1.392	0.484
50	0.238	0.259	0.329	0.345	0.409	0.390	0.583	0.470
200	0.047	0.012	0.057	0.027	0.072	0.031	0.100	0.038
400	0.011	0.005	0.014	0.007	0.017	0.007	0.024	0.008

Da análise da Tabela conclui-se que, quando  $\tau^0 > 200$  se pode utilizar  $\theta_1$  e  $\gamma_1$  como estimadores de  $\lambda$  e de  $\alpha$ .

# Capítulo 4

## Aplicação: comparação de cultivares de trigo rijo

### 4.1 Introdução

A utilização de plantas na nossa alimentação é um ato comum que faz parte do cotidiano humano. Além disso as plantas também são utilizadas no vestuário, na energia, em medicamentos, na alimentação de animais (os quais por sua vez nos servirão de alimento), etc. Podemos afirmar que a sobrevivência humana depende em grande parte das plantas e do seu manuseamento.

O conhecimento das qualidades tecnológicas do trigo são bem conhecidas, principalmente para a indústria alimentar, onde é usado basicamente para a produção de massas. As suas características nutricionais estão relacionadas com a quantidade de amido, proteínas, lípidos, vitaminas e fibras e pela sua capacidade antioxidante, as quais podem ser utilizadas em benefício de doenças degenerativas, graças ao seu teor em componentes bioativos. Todas estas qualidades do trigo rijo são de grande utilidade não só para a indústria alimentar, que detém assim a possibilidade de produzir produtos alimentares de alta qualidade, a qual, cada vez mais é exigida pelos consumidores. Na agricultura, ao fazer-se a seleção das variedades com maior qualidade tecnológica, verá aumentadas as condições de comercialização da sua produção. Embora a qualidade tecnológica dos trigos dependa em grande parte da variedade, outros aspetos como as condições ambientais de cada local e de cada ano e, as técnicas de cultivo, a fertilização, também influem para aumentar a qualidade desejada.

No caso do trigo rijo, dentre as culturas anuais cultivadas em Portugal, é uma das mais sensíveis às influências ambientais. Essa característica resulta em recomendações regionalizadas, onde determinados cultivares devem ser alocados em condições específicas, tirando o máximo aproveitamento dos recursos do ambiente. Um cultivar é uma variedade de uma planta, criada ou escolhida intencionalmente e mantida através do cultivo e à qual corresponde um determinado genótipo e fenótipo que recebe um nome único devidamente registado, (Wilhelm Johannsen, 1911). O trigo rijo (*Triticum durum*) tem uma expressão maior sobretudo nos países da orla mediterrânica (entre os quais se inclui Portugal). Os principais produtores a nível mundial de trigo rijo são: União Europeia, China, Índia, Estados Unidos, Rússia, Canadá, Austrália, Paquistão, Ucrânia e Turquia (FAO - Organização das Nações Unidas para a Agricultura e Alimentação, 2009). Em Portugal a cultura do trigo rijo ocupa atualmente uma superfície de 35.000ha, (INE, I.P., Lisboa -Portugal, 2012. [www.ine.pt](http://www.ine.pt)).

## 4.2 Dados e Objetivos

A aplicação prática a que se refere este Capítulo tem por base um trabalho anterior desenvolvido na tese de mestrado de Dias (2000), onde se tinha por objetivo estudar as interações específicas, as quais se verificam quando uma dada cultivar (variedade cultivada) se comporta anormalmente bem ou anormalmente mal num dado ensaio Gusmão, Mexia & Baeta; (1992). A técnica então utilizada para análise e interpretação dos dados foi a *Análise de Regressão Conjunta (Joint Regression Analysis - J.R.A.)*. Na base do método J.R.A. está a ideia de utilizar a média das produções num dado local, para medir a capacidade de produção correspondente. Esta produção média é designada por índice ambiental, trata-se de um fator único responsável pelo padrão de resposta das cultivares às situações ambientais.

Historicamente a análise de regressões tem sido utilizada para interpretar redes de ensaios destinados à comparação de cultivares. O nome da técnica provinha de se ajustar uma regressão linear por cultivar. A técnica foi originalmente descrita por Mooers (1921) tendo sido introduzida por Yates & Cochran (1938), que calcularam a média total de todos os cultivares em determinado ambiente e designaram essa medida por índice ambiental. Os autores estudavam o comportamento dum cultivar através da regressão linear das produções do mesmo nos índices ambientais. Na década de sessenta

sucesivas investigações em ACR conduzem a grandes avanços da técnica, permitindo a sua utilização na comparação do desempenho de cultivares quando sujeitos a um vasto leque de capacidades produtivas dos pares (local, ano). Salientam-se os contributos dos trabalhos de Finlay & Wilkinson (1963), de Eberhart & Russel (1966) e de Perkins & Jinks (1966, 1968a, 1968b). Finlay & Wilkinson (1963) propuseram o ajustamento de regressões lineares, para as diferentes cultivares, das produções respetivas nesses índices ambientais. Em 1997, Ng & Grunwal desenvolveram métodos iterativos para o ajustamento conjunto de regressões. Recentemente, Pinto (2006) mostrou como utilizar a ACR na condução de planos de melhoramento. Aplicou os seus resultados ao plano português de melhoramento de trigo mole (1986-2000).

É ainda possível utilizar este método para definir regiões tão vastas quanto possível, as zonas equipotenciais de adaptação, e podendo-se aplicar a técnica a redes de ensaios implementados nas mesmas. O que se pretende com a aplicação prática que se apresenta neste trabalho é o de contemplar outros aspetos para comparação de cultivares, esperando-se em trabalhos futuros fazer a comparação dos resultados obtidos aquando da utilização das duas técnicas, quer separadamente quer em conjunto.

Os nove ensaios que se consideraram neste trabalho encontravam-se implementados na zona “ Sul de Montejunto-Estrela” (ZEA II), que abrange a mais vasta área cerealífera do País, como mostra a figura



Figura 4.1 Mapa das zonas equipotenciais de adaptação

Os dados foram cedidos pela estação de melhoramento de plantas de Elvas, a quem se agradece a disponibilização dos mesmos. Os ensaios considerados referem-se aos anos de 1992/1993 e 1993/1994 e foram realizados em quatro locais distintos em cada ano,

Beja, Benavila, Elvas e Tavira. Em ambos os anos foram colocadas em ensaio nove variedades de trigo rijo (*CELTA, HELVIO, TE 9006, TE 9007, TE 9008, TE 9110, TE 9115, TE 9204, TROVADOR*).

Os ensaios foram realizados segundo um dispositivo experimental em blocos casualizados completos com quatro repetições.

Consideramos os fatores: ano com dois níveis (1991/1993 e 1993/1994) e, local com quatro níveis (Beja, Benavila, Elvas e Tavira)

O objetivo deste estudo será o de avaliar e verificar os efeitos e interações dos respetivos fatores, ano e local, no comportamento dos cultivares nos diferentes locais e para os diferentes anos.

### 4.3 Metodologia

Para cada par (local, ano) tem-se um ensaio em blocos casualizados com 4 repetições. As produções foram colocadas em matrizes  $\mathbf{X}_{lh}$  com  $l=1,2$  o índice do ano e  $h = 1, 2, 3, 4$  o índice do local.

Nestas matrizes tem-se uma linha por variedade e uma coluna por bloco. Estamos interessados nas variedades o que nos leva a considerar as matrizes

$$\mathbf{M}_{lh} = \mathbf{X}_{lh} \mathbf{X}_{lh}^t, \quad l = 1, 2; \quad h = 1, 2, 3, 4$$

Para simplificar os cálculos utilizamos as propriedades das matrizes conjugadas trabalhando com as

$$\dot{\mathbf{M}}_{lh} = \mathbf{X}_{lh}^t \mathbf{X}_{lh}, \quad l = 1, 2; \quad h = 1, 2, 3, 4$$

Estes pares de matrizes terão um máximo de quatro valores próprios não nulos.

Para verificar quais os efeitos e interações dos respetivos fatores no comportamento das cultivares, nos diferentes locais e para os diferentes anos, analisa-se uma família estruturada dessas matrizes, correspondendo os tratamentos do modelo base aos pares (local, ano). Para realizar o ajustamento dos modelos individuais, utilizamos a técnica das matrizes conjugadas, (secção 2.2).

As componentes dos vetores de estrutura correspondem aos cultivares. Assim para estudar a ação dos fatores, ano e local, para cada cultivar, podemos realizar uma Análise

transversal (secção 3.9.1). Esta análise incide sobre as componentes homólogas dos vetores de estrutura, realizando-se a análise de variância correspondente.

## 4.4 Resultados

Utilizando as propriedades das matrizes conjugadas

$$\dot{\mathbf{M}}_{lh} = \mathbf{X}_{lh}^t \mathbf{X}_{lh}, \quad l = 1, 2; \quad h = 1, 2, 3, 4$$

verificamos que estes pares de matrizes têm quatro valores próprios não nulos. Esses valores próprios, bem como os valores das preponderâncias (secção 3.9.3), são apresentados na Tabela 4.1 e na Tabela 4.2.

Tabela 4.1: Valores próprios ajustados ( $\theta_i, i = 1, \dots, 4$ ) e valores da preponderância ( $\tau$ )

<b>Ano 1992/1993</b>				
	Beja	Benavila	Elvas	Tavira
$\theta_1$	1594579365.1	265876447.50	984951776	1493579365.1
$\theta_2$	7246605.7	5396645.36	7965263	6946605.7
$\theta_3$	3408466.9	2084587.51	2085464	3338466.9
$\theta_4$	655574.1	75917.05	660160	66557.10
$\tau$	39384.37	2111.73	14218.45	37079.84

Tabela 4.2: Valores próprios ajustados ( $\theta_i, i = 1, \dots, 4$ ) e valores da preponderância ( $\tau$ )

<b>Ano 1993/1994</b>				
	Beja	Benavila	Elvas	Tavira
$\theta_1$	96943968.5	476006696	539018343.1	10788358
$\theta_2$	3259516.1	2159445	3510632.0	19725327
$\theta_3$	1274789.2	1417157	2301972.9	8335463
$\theta_4$	116186.5	477976	935209.9	4524393
$\tau$	766.38	32838.04	15706.41	2425.34

Da sua análise verifica-se que o primeiro valor próprio é muito superior em várias ordens de grandeza aos restantes, para todos as cultivares; o valor da preponderância ( $\tau$ ) em todos os casos é superior a 766, indicando que se pode utilizar modelos de grau um para todos os casos. As Tabela 4.3 e Tabela 4.4 mostram os vetores de estrutura para as várias matrizes estocásticas e as somas de quadrados de resíduos, para as respetivas variedades e para os fatores ano e local.

Tabela 4.3: Vetores de estrutura ajustados e soma de quadrados de resíduos

	Ano de 1992/1993			
	Beja	Benavila	Elvas	Tavira
CELTA	14540.54	7430.448	13346.89	22596.21
HELVIO	8964.56	10441.59	5285.928	4169.252
TE 9006	13386.49	1522.099	7777.199	22730.64
TE 9007	20278.34	7643.398	10168.73	6322.33
TE 9008	3514.226	9680.511	2264.447	38063.24
TE 9110	1712.533	2367.907	1629.267	3747.177
TE 9115	15541.44	6520.88	12376.82	19898.47
TE 9204	4634.124	9596.588	7654.076	34462.95
TROVAD	10089.29	3447.883	9809.606	4472.869
V	11310646.7	7557149.2	10710887	10710887

Tabela 4.4: Vetores de estrutura ajustados e soma de quadrados de resíduos

	Ano de 1993/1994			
	Beja	Benavila	Elvas	Tavira
CELTA	11767.08	8399.241	9890.214	11529.6
HELVIO	21619.88	12764.94	1911.708	3590.561
TE 9006	3553.035	11186.26	3312.333	12371.35
TE 9007	4760.576	11018.63	10660.63	10856.13
TE 9008	13593.19	1972.86	1654.533	27398.45
TE 9110	1277.349	1920.22	13447.06	16256.56
TE 9115	12063.3	9840.968	10536.18	2862.637
TE 9204	4844.458	4153.382	13729.72	25687.27
TROVAD	4917.375	7502.148	12995.64	7538.134
V	4650491.8	4054578	6747814.8	32585183



As componentes homogêneas correspondem aos diferentes cultivares, por exemplo, as primeiras, correspondem ao *CELTA*.

Assim, a AT incidirá sobre as diferentes cultivares (secção, 3.9.2).

Para cada uma destas testam-se as seguintes hipóteses

$H_{01}$  : ausência de efeitos dos anos;

$H_{02}$  : ausência de efeitos dos locais;

$H_{03}$  : ausência de interação entre o fator ano e o fator local.

Realizou-se, para cada hipótese, a ANOVA correspondente. Para isso utilizamos a matriz ortogonal estandardizada (secção, 2.4.2). O produto da segunda e terceira matriz corresponde ao fator local que tem quatro níveis.

$$P_8 = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \otimes \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \otimes \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

Dado  $\mathbf{Z} = \mathbf{PY}$  com componentes  $\mathbf{Z}_i$ ,  $i = 1, \dots, 8$ , temos as somas de quadrados para os efeitos e interação, indicando-se na Tabela 4.7 como estas podem ser calculadas.

Tabela 4.5: Somas de quadrados

1ºFator	$S_1 = \mathbf{Z}_5^2$
2ºFator	$S_2 = \mathbf{Z}_2^2 + \mathbf{Z}_3^2 + \mathbf{Z}_4^2$
Interação	$S_{1 \times 2} = \mathbf{Z}_6^2 + \mathbf{Z}_7^2 + \mathbf{Z}_8^2$

Na Tabela 4.6 e na Tabela 4.7, encontra-se o resumo das análises de variância efetuadas aos cultivares. Foram testadas as 3 hipóteses  $H_{01}$ ,  $H_{02}$ ,  $H_{03}$ , usando-se um nível de significância de 5%, verificando-se que existe interação significativa para os cultivares, bem como influências significativas quer do fator ano quer do fator local para os cultivares.

Tabela 4.6: Quadros resumo das análises de variância

CELTA				
	SQ	GL	MQ	<i>F</i>
ANO	504590427	1	504590427	272.36 ***
LOCAL	385773291	3	128591097	69.409 ***
INTERAÇÃO	342523172	3	114174390.7	61.63***
RESÍDUO(MQE)	1852654.12			
HELVIO				
	SQ	GL	MQ	<i>F</i>
ANO	5348841	1	5348841	2.887123
LOCAL	120982674	3	40327558	21.76745***
INTERAÇÃO	197085839	3	65695280	35.46***
RESÍDUO(MQE)	1852654			
TE 9006				
	SQ	GL	MQ	<i>F</i>
ANO	48389639	1	48389639	26.1191***
LOCAL	125557773	3	41852591	22.592***
INTERAÇÃO	320953707	3	106984569	57.7467***
RESÍDUO(MQE)	1852654			
TE 9007				
	SQ	GL	MQ	<i>F</i>
ANO	177637688	1	177637688	95.88281***
LOCAL	2620035227	3	873345076	471.4021***
INTERAÇÃO	914476443	3	304825481	164.5345***
RESÍDUO(MQE)	1852654,12			
TE 9008				
	SQ	GL	MQ	<i>F</i>
ANO	218005888	1	218005888	117.6722***
LOCAL	1552502738	3	517500913	279.3295***
INTERAÇÃO	1827569546	3	609189849	328.8201***
RESÍDUO(MQE)	1852654,12			

Tabela 4.7: Quadros resumo das análises de variância

TE 9110				
	SQ	GL	MQ	F
ANO	15067409	1	15067409	8.132878**
LOCAL	235722803	3	78574268	42.41173***
INTERAÇÃO	318336316	3	106112105	57.27572***
RESÍDUO(MQE)	1852654,12			
TE 9115				
	SQ	GL	MQ	F
ANO	838805.2	1	838805.2	0.4527587
LOCAL	33046621	3	11015540	5.945816**
INTERAÇÃO	407358444	3	135786148	73.29277***
RESÍDUO(MQE)	1852654,12			
TE 9204				
	SQ	GL	MQ	F
ANO	194371842	1	194371842	104.9153***
LOCAL	218729891	3	72909964	39.35433***
INTERAÇÃO	193429437	3	64476479	34.80222***
RESÍDUO(MQE)	1852654,12			
TROVADOR				
	SQ	GL	MQ	F
ANO	1751986	1	1751986	2.456628
LOCAL	686301950	3	228767317	123.4809***
INTERAÇÃO	1728724052	3	576241351	311.0356***
RESÍDUO(MQE)	1852654,12			

onde \*- significa ao nível de 5%; \*\*- significa ao nível de 1% e \*\*\*- significa ao nível de 0,1%

conclui-se assim que

- Há efeito significativo do fator ano para as cultivares:  
*CELTA, TE9006, TE9007, TE9008, TE9110 e TE9204.*
- Há efeito significativo do fator local para as cultivares:  
*CELTA, HELVIO, TE9006, TE9007, TE9008, TE9110, TE9115, TE9204 e TROVADOR.*
- Há interação significativa para as cultivares:

*CELTA, HELVIO, TE9006, TE9007, TE9008, TE9110, TE9115, TE9204 e TROVADOR.*

A partir daqui podemos ainda discutir a importância relativa dos fatores e da interação. Quanto maior for a média das estatísticas F para um efeito ou interação tanto maior será o peso da correspondente causa de variação. Assim somos levados a medir as importâncias relativas utilizando as médias  $\bar{F}_A$ ,  $\bar{F}_L$  e  $\bar{F}_I$ , das estatísticas respectivas, tomando como medida

$$\begin{cases} \tau_A = \frac{\bar{F}_A}{\bar{F}_A + \bar{F}_L + \bar{F}_I} \\ \tau_L = \frac{\bar{F}_L}{\bar{F}_A + \bar{F}_L + \bar{F}_I} \\ \tau_I = \frac{\bar{F}_{A \times L}}{\bar{F}_A + \bar{F}_L + \bar{F}_I} \end{cases}$$

Os resultados obtidos foram  $\tau_A = 0.22$ ,  $\tau_L = 0.38$  e  $\tau_I = 0.40$ , o que mostra a grande relevância do fator local.

No entanto essas relevâncias variam muito de cultivar para cultivar como se pode observar pela análise das Figuras seguintes.

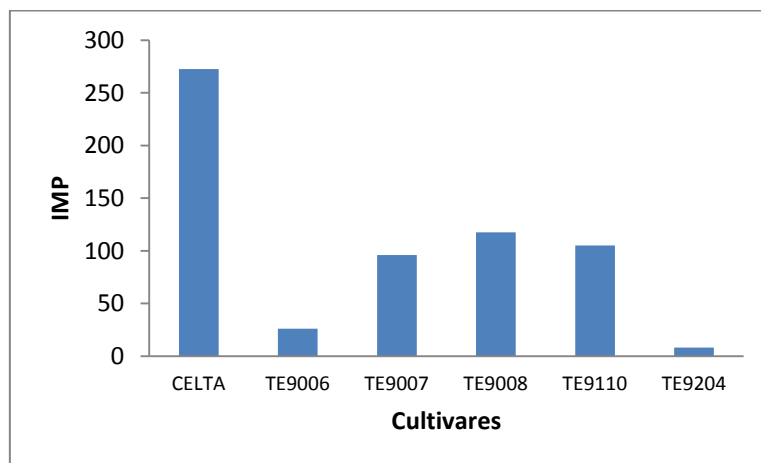


Figura 4.1: Importância relativa do ano para as nove cultivares

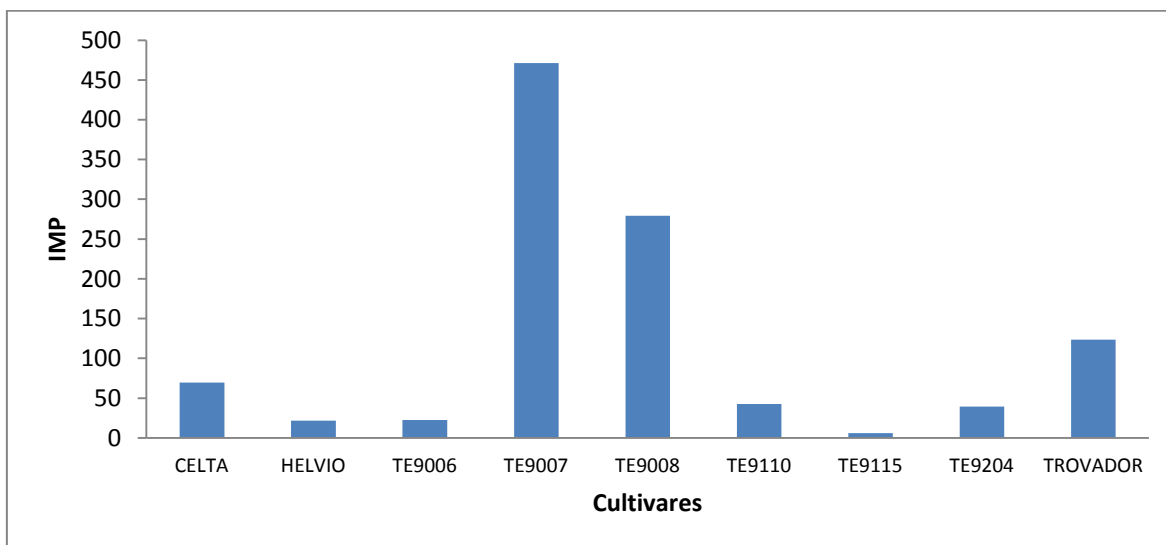


Figura 4.2: Importância relativa do local para as nove cultivares

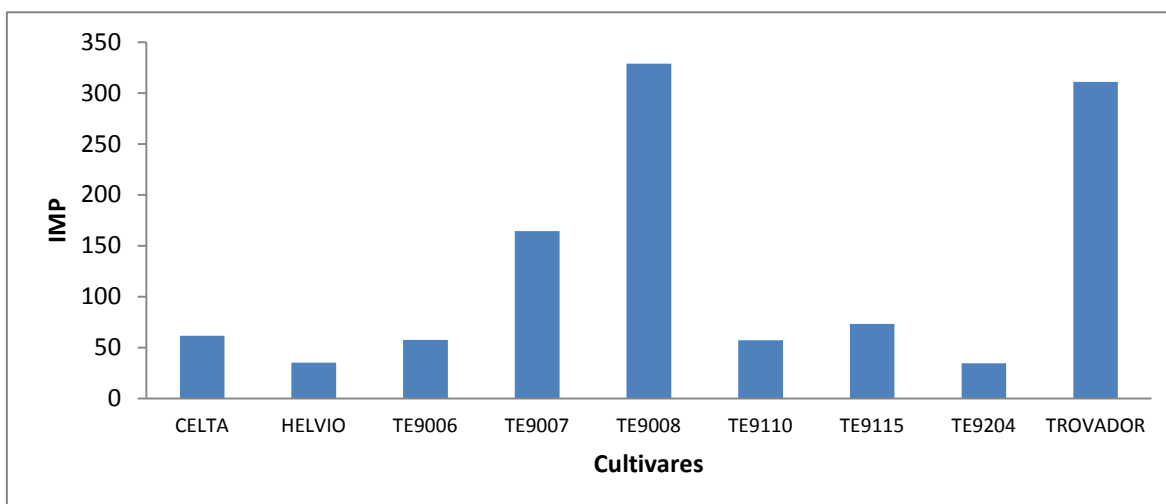


Figura 4.3: Importância relativa da interação para as nove cultivares

# Capítulo 5

## Considerações Finais

Nesta dissertação vimos como ajustar e validar modelos assentes na análise espectral de matrizes médias de matrizes estocásticas simétricas. Consegue-se assim uma formulação que permite condensar a informação contida numa matriz estocástica simétrica num par constituído por um vetor  $\tilde{\beta}$  de estrutura ajustado e numa soma  $V$  de quadrados de resíduos. Esta condensação é paralela á que se tem para regressões lineares sendo então  $\tilde{\beta}$  o vetor dos coeficientes ajustados e, continuando  $V$  a ser uma soma de quadrados de resíduos.

Esta possibilidade de condensação da informação contida numa matriz estocástica simétrica torna estes modelos adequados para o estudo de famílias estruturadas. Nestas famílias os modelos correspondem aos tratamentos de um delineamento base. Assim, ao analisar-se uma tal família pode-se ir mais fundo do que quando se trabalha apenas com um modelo, já que, se pode estudar a ação dos fatores do delineamento base sobre os parâmetros dos modelos da família. No estudo apresentado considerou-se o caso de famílias estruturadas de modelos com delineamento base ortogonal, o que corresponde a estar-se numa situação de equilíbrio. Realizam-se análises de variância para estudar a ação dos fatores do delineamento base sobre combinações lineares  $\mathbf{a}'\beta_1, \dots, \mathbf{a}'\beta_m$  das componentes dos vetores de estrutura das matrizes da família. Quando  $\mathbf{a} = \delta_i$  a análise incidirá sobre as  $i$ -ésimas componentes dos vetores de estrutura. Por exemplo na primeira fase da metodologia STATIS, a inter-estrutura, incidirá sobre as componentes associadas aos  $i$ -ésimos estudos. Temos então, como vimos, uma AT. Se as somas das componentes for nula,  $\mathbf{a}$  será um vetor de contraste e temos uma, AL. Voltando à inter-

estrutura quando os estudos que constituem as várias séries correspondem a instantes sucessivos, a AL incidirá sobre a evolução das mesmas. Aliás na representação de Escoufier dos operadores os pontos representativos dos estudos, no caso de modelos de grau um, dispõem-se ao longo do eixo paralelo ao vetor de estrutura. Compreende-se a designação de análise longitudinal uma vez que a mesma incidirá sobre a evolução ao longo desse eixo dos pontos representativos dos estudos. Referimo-nos à metodologia STATIS pois foi para a respetiva primeira etapa desta metodologia, a inter-estrutura, que se iniciou a modelização (Oliveira 2002), de matrizes estocásticas simétricas. Posteriormente verificou-se que o mesmo tipo de modelos eram aplicáveis a matrizes estocásticas simétricas em geral podendo, à semelhança do que já se tinha feito antes para a inter-estrutura, utilizar esses modelos para realizar AT e AL para famílias estruturadas de modelos. Aliás pode-se considerar este trabalho como um alargamento do tratamento que tinha sido feito para a inter-estrutura. Nesse tratamento tinha-se verificado que na generalidade dos casos, se podem utilizar os modelos de grau um, o mesmo facto verifica-se agora com a aplicação das matrizes de produtos cruzados a dados de melhoramento vegetal, mais concretamente a variedades de trigo rijo. Aqui para o modelo base considerou-se o cruzamento de dois fatores com interação (local e ano), o primeiro fator (local), com quatro níveis (Beja, Benavila, Elvas, Tavira), e o segundo fator (ano) com dois níveis (1992/1993 e 1993/1994). Os resultados obtidos indicaram que existem influências significativas quer do fator ano quer do fator local, relativamente ao comportamento das cultivares, indicando ainda a existência de interação significativa. Ainda, relativamente aos fatores e interação, foi possível discutir a importância relativa dos mesmos, considerando as médias das estatísticas F, pelo que estas medidas representam um primeiro passo para colmatar as dificuldades que se colocam, quando o utilizador comum é confrontado com um volume excessivo de informação, tendo agora a possibilidade de conseguir interpretar os resultados mais claramente. Ora, como vimos, a adoção de modelos de grau um permite uma grande simplificação dos cálculos. Surge assim, a questão de analisar mais completamente as situações em que tais modelos podem ser utilizados. Observe-se que os casos em que isso se verifica diziam respeito a matrizes de produtos cruzados, o que sugere que se comece por verificar quando é que modelos de grau um se aplicam a tais matrizes. Este será o ponto de partida para um trabalho futuro, onde pensamos aplicar a técnica a um conjunto de dados mais recentes.

# Apêndice A

## Linearidade Assintótica e Normalidade

De seguida apresentam-se alguns conceitos sobre Linearidade Assintótica (Mexia e Oliveira, 2011) e sua aplicação a estudos de simulações (utilizados no Capítulo 3).

### A.1 Função Assintótica Linear

Dada uma função suficientemente regular  $g : R^k \rightarrow R$ . Seja  $g_1$  o gradiente e  $g_2$  a matriz Hessiana de  $g$  e  $r_d(\mathbf{X})$  o supremo do raio espectral  $r(\mathbf{Y})$  de  $g_2(\mathbf{Y})$ , quando  $\|\mathbf{Y} - \mathbf{X}\| \leq d$ , a função  $g$  é assintoticamente linear, se qualquer que seja  $d > 0$ , se tem  $k_d(u) \xrightarrow{u \rightarrow \infty} 0$ , com

$$k_d(u) = \sup \left\{ \frac{r_d(\mathbf{X})}{\|g_1(\mathbf{X})\|}; \|\mathbf{X}\| \geq u \right\}.$$

Sempre que se deriva um polinómio o grau diminui numa unidade, e os polinómios serão, salvo raras exceções, assintoticamente lineares.

Seja  $\mathbf{X}$  um vetor aleatório e  $g(\cdot)$  uma função assintoticamente linear, o método delta pode ser utilizado para aproximar as estatísticas

$$Y = g(\boldsymbol{\mu} + \mathbf{X}),$$

e

$$Y^0 = \frac{(Y - g(\boldsymbol{\mu}))}{\|g_1(\boldsymbol{\mu})\|} = \frac{g(\boldsymbol{\mu} + \mathbf{X}) - g(\boldsymbol{\mu})}{\|g_1(\boldsymbol{\mu})\|},$$

por, respetivamente,

$$Z = g(\boldsymbol{\mu}) + g_1(\boldsymbol{\mu})' \mathbf{X},$$



e

$$Z^o = \frac{(Z - g(\boldsymbol{\mu}))}{\|g_1(\boldsymbol{\mu})\|} = \frac{g_1(\boldsymbol{\mu})^t \mathbf{X}}{\|g_1(\boldsymbol{\mu})\|},$$

quando  $\|\boldsymbol{\mu}\| \rightarrow +\infty$ , (Oliveira e Mexia, 2011).

Sejam  $y_p^o$  e  $z_p^o$  os quantis para a probabilidade  $p$  de  $\mathbf{Y}^o$  e  $\mathbf{Z}^o$ . Quando

$$\frac{1}{\|g_1(\boldsymbol{\mu})\|} g_1(\boldsymbol{\mu}) \xrightarrow{\|\boldsymbol{\mu}\| \rightarrow \infty} \mathbf{b},$$

tem-se que  $F_{Z^o}$  converge para  $F_{Z_o^o}$  com  $Z_o^o = \mathbf{b}^t \mathbf{X}$ .

Estes resultados podem aplicar-se sempre que  $\mathbf{X}$  é normal e, nesse caso também  $Z_o^o$  é normal e ter-se-á um teorema do limite central que não estará vinculado ao crescimento da dimensão da amostra, mas sim ao da não centralidade.

Antes de entrarmos no estudo da aplicação ao caso normal, apresentamos alguns resultados assintóticos.

## A.2 Resultados Assintóticos

Se  $g : R^k \rightarrow R$  é suficientemente regular temos (Khuri, 2003),

$$Y = g(\boldsymbol{\mu} + \mathbf{X}) = g(\boldsymbol{\mu}) + g_1(\boldsymbol{\mu})^t \mathbf{X} + \int_0^1 (1-h) (\mathbf{X}^t g_2(\boldsymbol{\mu} + h\mathbf{X}) \mathbf{X}) dh.$$

Quando  $\|\mathbf{X}\| \leq d$ , vem

$$|Y - Z| \leq \left| \int_0^1 (1-h) (\mathbf{X}^t g_2(\boldsymbol{\mu} + h\mathbf{X}) \mathbf{X}) dh \right| \leq r_d(\boldsymbol{\mu}) d^2,$$

e

$$|Y^o - Z^o| \leq k_d(\|\boldsymbol{\mu}\|) d^2,$$

Estabelecendo-se a

**Proposição A.1** Se  $g(\cdot)$  é assintoticamente linear, então  $|Y^o - Z^o| \xrightarrow[\|\boldsymbol{\mu}\| \rightarrow \infty]{q.c.} 0$ , (onde q.c.

indica convergência quase certa).

**Demonstração:** Se  $g(\cdot)$  é assintoticamente linear, então para qualquer  $\varepsilon > 0$  e  $d > 0$  existe  $u(\varepsilon)$  tal que, para  $u > u(\varepsilon)$ ,  $k_d(u) < \varepsilon$ . Assim quando  $\|\boldsymbol{\mu}\| \geq u(\varepsilon)$ ,  $\|\mathbf{X}\| \leq d$  implica  $|Y^o - Z^o| \leq \varepsilon$ . Para qualquer  $\delta > 0$ , pode-se escolher  $d > 0$ , para assegurar que  $\Pr(\|\mathbf{X}\| < d) \geq 1 - \delta$ , assim quando  $\|\boldsymbol{\mu}\| \geq u(\varepsilon)$ , também temos  $\Pr(|Y^o - Z^o| \leq \varepsilon) \geq 1 - \delta$ . Ficando a tese estabelecida.

Sabendo que  $k_d(u)$  decresce com  $u$  existe

$$u(d, \varepsilon) = \text{Min} \{u : k_d(u)d^2 \leq \varepsilon\},$$

Considere-se a distribuição  $\mathbf{L} = \|\mathbf{X}\|^2$  e, seja  $l_q$  o quantil de ordem  $q$  dessa distribuição quando  $\mathbf{L} \leq l_q$ , com  $\|\boldsymbol{\mu}\| \geq u_q(\varepsilon) = u(l_q, \varepsilon)$ , ter-se-á

$$|Y^o - Z^o| \leq \varepsilon.$$

Designemos por  $q(\varepsilon)$  a probabilidade do evento  $A(\varepsilon)$  com

$$A(\varepsilon) = \{|Y^o - Z^o| \leq \varepsilon\},$$

tem-se que,  $q(\varepsilon) \geq q$ , quando  $\|\boldsymbol{\mu}\| > u_q(\varepsilon)$ . Represente-se ainda por  $A^c(\varepsilon)$  o complementar do evento  $A(\varepsilon)$  e  $q^c(\varepsilon)$  o complementar do evento  $q(\varepsilon)$  com,  $q^c(\varepsilon) = 1 - q(\varepsilon)$ .

Assim temos a

**Proposição A.2** Qualquer que seja  $z$  e  $\varepsilon > 0$ ,

$$F_{Z^o}(z - \varepsilon) - q^c(\varepsilon) \leq F_{Y^o}(z) \leq F_{Z^o}(z + \varepsilon) + q^c(\varepsilon).$$

**Demonstração:** tem-se,

$$\begin{aligned} F_{Y^o}(z) &= \Pr(Y^o \leq z) = \\ &= q(\varepsilon)\Pr(Y^o \leq z | A(\varepsilon)) + q^c(\varepsilon)\Pr(Y^o \leq z | A^c(\varepsilon)) \leq q(\varepsilon)\Pr(Z^o \leq z + \varepsilon | A(\varepsilon)) + q^c(\varepsilon) \end{aligned}$$

pelo que

$$F_{Z^o}(z - \varepsilon) - q^c(\varepsilon) \leq q(\varepsilon) \Pr(Z^o \leq z - \varepsilon | A(\varepsilon)) \leq q(\varepsilon) \Pr(Y^o \leq z | A(\varepsilon)) \leq \Pr(Y^o \leq z) = F_{Y^o}(z).$$

Ficando a tese estabelecida.

**Corolário A.1** Quando  $F_{Z^o}(z)$  tem densidade  $f_{Z^o}(z)$  limitada por  $C$ , tem-se

$$\text{Sup}\{|F_Y - F_Z|\} = \text{Sup}\{|F_{Y^o} - F_{Z^o}|\} \leq d(\varepsilon),$$

com  $d(\varepsilon) = 2(C\varepsilon + q^c(\varepsilon))$ .

**Demonstração:** sendo  $F_Y(x)$  e  $F_Z(x)$  as respectivas funções de distribuição das variáveis aleatória  $Y$  e  $Z$ , tem-se

$$F_Y(x) = F_{Y^o}\left(\frac{x - g(\boldsymbol{\mu})}{\|g_1(\boldsymbol{\mu})\|}\right) \quad \text{e} \quad F_Z(x) = F_{Z^o}\left(\frac{x - g(\boldsymbol{\mu})}{\|g_1(\boldsymbol{\mu})\|}\right),$$

com  $\text{Sup}\{|F_Y - F_Z|\} = \text{Sup}\{|F_{Y^o} - F_{Z^o}|\}$ . Para a demonstração ficar completa basta apenas aplicar a proposição A.1, sabendo que  $f_{Z^o}$  é limitada por  $C$ , então  $F_{Z^o}(z + \varepsilon) - F_{Z^o}(z - \varepsilon) \leq 2C\varepsilon$ . Ficando a tese estabelecida.

Tem-se agora o

**Corolário A.2** Quando  $f_{Z^o}$  é limitada por  $C$  e  $g(\cdot)$  é assintoticamente linear,

$$\text{Sup}\{|F_Y - F_Z|\}_{\|\boldsymbol{\mu}\| \rightarrow +\infty} \rightarrow 0.$$

Ficamos assim em condições de garantir que  $F_Z$  converge uniformemente para  $F_Y$ , quando  $\|\boldsymbol{\mu}\| \rightarrow \infty$ .

Para as densidades tem-se a:

**Proposição A.3** Se  $f_{Z^o}$  é limitada por  $C$  e se  $f_{Z^o}(z) \geq m(\delta) > 0$ , sempre que

$$z_{\delta-d(\varepsilon)}^o \leq z \leq z_{1-\delta+d(\varepsilon)}^o, \text{ ter-se-á } \text{Sup}\{|y_p^o - z_p^o|; \delta \leq p \leq 1-\delta\} \leq 2\left(\varepsilon + \frac{d(\varepsilon)}{m(\delta)}\right).$$

**Demonstração:** de acordo com o corolário A.1 da proposição A.2, tem-se

$$F_{Z^o}(y_p^o - \varepsilon) - d(\varepsilon) \leq F_{Y^o}(y_p^o) = p \leq F_{Z^o}(y_p^o + \varepsilon) + d(\varepsilon),$$

logo

$$F_{z^o}(y_p^o - \varepsilon) \leq p + d(\varepsilon) \text{ e } p - d(\varepsilon) \leq F_{z^o}(y_p^o + \varepsilon),$$

com

$$y_p^o - \varepsilon \leq z_{p+d(\varepsilon)} \quad \text{e} \quad z_{p-d(\varepsilon)}^o \leq y_p^o + \varepsilon,$$

resulta

$$z_{p-d(\varepsilon)}^o - \varepsilon \leq y_p^o \leq z_{p+d(\varepsilon)}^o + \varepsilon.$$

Para completar a demonstração basta lembrar que

$$z_{p+d(\varepsilon)}^o - z_{p-d(\varepsilon)}^o \leq \frac{2d(\varepsilon)}{m(\delta)}.$$

**Corolário A.3** Se  $\|\boldsymbol{\mu}\| \geq u_q(\varepsilon)$ ,  $f_{z^o}$  é limitada por  $C$  e  $f_{z^o}(z)_{z \geq m(\delta) > 0}$ , então sempre que  $z_{\delta-d(\varepsilon)}^o \leq z \leq z_{1-\delta+d(\varepsilon)}^o$ , tem-se  $\text{Sup} \left\{ |y_p^o - z_p^o|, \delta \leq p \leq 1 - \delta \right\} \leq 2(\varepsilon + d_q(\varepsilon))/m(\delta)$ , com  $d_q(\varepsilon) = 2(C\varepsilon + q^c)$  com  $q^c = 1 - q$ .

**Demonstração:** da proposição A.3 vem que  $u_q(\varepsilon) > q$ , sempre que  $\|\boldsymbol{\mu}\| \geq u_q(\varepsilon)$ . Ficando a tese estabelecida.

**Proposição A.4** Se  $\frac{1}{\|g_1(\boldsymbol{\mu})\|} g_1(\boldsymbol{\mu}) \xrightarrow{\|\boldsymbol{\mu}\| \rightarrow \infty} \mathbf{b}$ ,  $F_{z^o} \xrightarrow{\|\boldsymbol{\mu}\| \rightarrow \infty} F_{z^o}$ , esta convergência será uniforme quando  $f_{z^o}$  é limitada for  $C$  e  $g(\cdot)$  é assintoticamente linear.

**Demonstração:** A primeira parte da tese segue de  $\left( \frac{1}{\|g_1(\boldsymbol{\mu})\|} g_1(\boldsymbol{\mu}) - \mathbf{b} \right)^t \mathbf{X} \xrightarrow{\|\boldsymbol{\mu}\| \rightarrow \infty} 0$ , sempre

que  $\frac{1}{\|g_1(\boldsymbol{\mu})\|} g_1(\boldsymbol{\mu}) \xrightarrow{\|\boldsymbol{\mu}\| \rightarrow \infty} \mathbf{b}$ .

Recorrendo à proposição A.3, temos que  $u_q(\varepsilon) > q$ , sempre que  $\|\boldsymbol{\mu}\| > u_q(\varepsilon)$ , pelo que a tese fica estabelecida.

### A.3 Aplicação ao Caso Normal.

Seja  $\mathbf{X}$  é um vetor normal, com vetor médio nulo e matriz de covariância  $\mathbf{M}$ . Então

$$f_{z^o}(z) = \frac{e^{-\|g_1(\boldsymbol{\mu})\|^2(z-g(\boldsymbol{\mu}))^2 / g_1(\boldsymbol{\mu})^t \mathbf{M} g_1(\boldsymbol{\mu})}}{\sqrt{2\pi g_1(\boldsymbol{\mu})^t \mathbf{M} g_1(\boldsymbol{\mu})}} \|g_1(\boldsymbol{\mu})\|$$

é limitada por

$$C(\boldsymbol{\mu}) = \frac{\|g_1(\boldsymbol{\mu})\|}{\sqrt{2\pi g_1(\boldsymbol{\mu})^t \mathbf{M} g_1(\boldsymbol{\mu})}}.$$

Se  $\theta_k$  é o menor dos valores próprios de  $\mathbf{M}$  tem-se

$$C(\boldsymbol{\mu}) \leq C = \frac{1}{\sqrt{2\pi\theta_k}}.$$

Sendo  $x_p$  o quantil para a probabilidade  $p$  da densidade normal estandardizada, os quantis de  $f_{z^o}$  vêm dados por

$$z_p^o(\boldsymbol{\mu}) = g_1(\boldsymbol{\mu}) + x_p \frac{\sqrt{g_1(\boldsymbol{\mu})^t \mathbf{M} g_1(\boldsymbol{\mu})}}{\|g_1(\boldsymbol{\mu})\|},$$

enquanto o mínimo de  $f_{z^o}$  em  $[z_{\delta-d(\varepsilon)}^o(\boldsymbol{\mu}); z_{1-\delta+d(\varepsilon)}^o(\boldsymbol{\mu})]$ , será

$$m(\delta | \boldsymbol{\mu}) = \frac{e^{-x_{\delta-d(\varepsilon)}^2/2}}{\sqrt{2\pi g_1(\boldsymbol{\mu})^t \mathbf{M} g_1(\boldsymbol{\mu})}} \|g_1(\boldsymbol{\mu})\|,$$

vindo

$$\frac{C(\boldsymbol{\mu})}{m(\delta | \boldsymbol{\mu})} = e^{x_{\delta-d(\varepsilon)}^2/2},$$

e

$$\frac{d(\varepsilon | \boldsymbol{\mu})}{m(\delta | \boldsymbol{\mu})} = \frac{2(C(\boldsymbol{\mu})\varepsilon + q^c(\varepsilon))}{m(\delta | \boldsymbol{\mu})} = 2e^{x_{\delta-d(\varepsilon)}^2/2} \left( \varepsilon + \frac{q^c(\varepsilon)\|g_1(\boldsymbol{\mu})\|}{\sqrt{2\pi g_1(\boldsymbol{\mu})^t \mathbf{M} g_1(\boldsymbol{\mu})}} \right) \leq 2e^{x_{\delta-d(\varepsilon)}^2/2} \left( \varepsilon + \frac{q^c(\varepsilon)}{\sqrt{2\pi\theta_k}} \right)$$

Obtemos limites que não dependem de  $\boldsymbol{\mu}$ , para  $f_{z^o}(z)$  e para  $\frac{d(\varepsilon | \boldsymbol{\mu})}{m(\delta | \boldsymbol{\mu})}$ , sendo assim,

possível aplicar os resultados anteriores a este caso. Além disso se  $\mathbf{X}$  é normal, com vetor médio nulo e matriz de covariância  $\mathbf{M}$  e,  $\theta_1 \geq \dots \geq \theta_h$  são os valores próprios de  $\mathbf{M}$  com multiplicidades  $g_1, \dots, g_h$ , tem-se  $\mathbf{L} = \|\mathbf{X}\|^2$ , que pode escrever como

combinação linear de  $\sum_{j=1}^h \theta_j \chi_{gj}^2$ , independentes da variação do qui-quadrado (Imhof, 1961). Quando  $\mathbf{M}$  é conhecida torna-se fácil calcular os quantis  $l_p$  (Fonseca et al., 2007).

## A.4 Simulações

Através do estudo de simulações mostra-se numericamente como polinômios de baixo grau em variáveis normais estandardizadas  $X$ ,  $X_1$  e  $X_2$  independentes com baixos coeficientes de variação têm distribuição aproximadamente normal. Para realizar as simulações utilizaram-se os polinômios  $P_1 = (c\mu + X)^r - (c\mu)^r$  e  $P_2 = (c_1\mu_1 + X_1)^r \cdot (c_2\mu_2 + X_2)^s$ . Consideram-se os testes de Kolmogorov-Smirnov (K-S) e Shapiro-Wilk para testar a aproximação à normal de  $P_1$  e  $P_2$ . Foram elaborados programas de simulação utilizando o Software R, foram geradas inicialmente 1000 observações pseudo-aleatórias para cada uma das variáveis  $X$ ,  $X_1$  e  $X_2$ . Às amostras obtidas aplicou-se o teste de Kolmogorov-Smirnov, tal como descrito acima. Assim para um nível de significância de 0.05, cujo valor crítico é de 0.0608. Os valores são apresentados nas tabelas que se seguem considerando arredondamentos à quarta casa decimal.

O polinômio inicialmente considerado é  $P_1 = (c\mu + X)^r - (c\mu)^r$ , onde  $r \in \{1, 2, 3, 4\}$ , para o qual obtivemos as seguintes estatísticas do teste de Kolmogorov-Smirnov, para a amostra de observações de  $P_1$ .

Tabela A.1: Estimativas do teste de Kolmogorov-Smirnov para  $P_1$

		$n = 1000$							
$r \setminus c$	1	5	10	20	30	40	50	100	200
1	0.0171	0.0171	0.0171	0.0171	0.0171	0.0171	0.0171	0.0171	0.0171
2	0.2095	0.0461	0.0286	0.0208	0.0194	0.0188	0.0185	0.0178	0.0174
3	0.2527	0.0835	0.0469	0.029	0.0236	0.0209	0.02	0.0185	0.0178
4	0.3501	0.1186	0.0663	0.0377	0.0291	0.025	0.0225	0.0192	0.0181

Para reforçar a nossa convicção nos resultados expressos na tabela 2.1, aplicou-se também o teste de Shapiro-Wilk a  $P_1 = (c\mu + X)^r - (c\mu)^r$  com  $r \in \{1, 2, 3, 4\}$ , as estatísticas de teste obtidas encontram-se na Tabela 2.2.

Tabela A.2: Estimativas do teste de Shapiro-Wilk para  $P_1$

$n = 1000$				
$c \setminus r$	1	2	3	4
1	0.048	1.000	1.000	1.000
5	0.060	1.000	1.000	1.000
10	0.050	0.940	1.000	1.000
20	0.050	0.390	0.931	1.000
30	0.054	0.214	0.626	0.945
40	0.045	0.115	0.384	0.736
50	0.056	0.106	0.255	0.542
100	0.060	0.066	0.105	0.159
200	0.052	0.055	0.070	0.080

Para uma melhor análise dos resultados obtidos basta ter em conta a Tabela A.1 e a Tabela A.3 e fazer a relação entre os valores observados das estatísticas e, os p-values do teste de Kolmogorov-Smirnov, para amostras de dimensão 1000.

Tabela A.3: P-value do teste de Kolmogorov-Smirnov para  $P_1$

$n = 1000$									
$r \setminus c$	1	5	10	20	30	40	50	100	200
1	0.9315	0.9315	0.9315	0.9315	0.9315	0.9315	0.9315	0.9315	0.9315
2	-	<u>0.0285</u>	0.3859	0.7795	0.8454	0.8704	0.8844	0.9097	0.9211
3	-	-	<u>0.0245</u>	0.3701	0.6353	0.7747	0.8203	0.8836	0.9095
4	-	-	<u>3e-04</u>	0.1161	0.3645	0.5604	0.6896	0.8532	0.8969

Em seguida consideremos o polinómio  $P_2 = (c_1\mu_1 + X_1)^r \cdot (c_2\mu_2 + X_2)^s$ . As estatísticas do teste de Kolmogorov-Smirnov, são apresentados nas tabelas, A.4, A.5 e A.6 nas quais consideramos todas as combinações de expoentes inteiros  $r, s$  tais que  $r + s \in \{2, 3, 4\}$ .

Tabela A.4: Estimativas do teste de Kolmogorov-Smirnov para  $P_2$

	$n = 1000 \quad r = 1 \quad s = 1$								
$c_1 \setminus c_2$	1	5	10	20	30	40	50	100	200
1	0.2085	0.0735	<b>0.0432</b>	<b>0.0256</b>	<b>0.0194</b>	<b>0.0162</b>	<b>0.0144</b>	<b>0.0139</b>	<b>0.0141</b>
5	0.0735	<b>0.0469</b>	<b>0.0334</b>	<b>0.0225</b>	<b>0.0179</b>	<b>0.0154</b>	<b>0.0141</b>	<b>0.0139</b>	<b>0.0141</b>
10	<b>0.0432</b>	<b>0.0334</b>	<b>0.0266</b>	<b>0.0198</b>	<b>0.0164</b>	<b>0.0145</b>	<b>0.0138</b>	<b>0.0139</b>	<b>0.0141</b>
20	<b>0.0256</b>	<b>0.0225</b>	<b>0.0198</b>	<b>0.0164</b>	<b>0.0145</b>	<b>0.0138</b>	<b>0.0137</b>	<b>0.014</b>	<b>0.0141</b>
30	<b>0.0194</b>	<b>0.0179</b>	<b>0.0164</b>	<b>0.0145</b>	<b>0.0138</b>	<b>0.0137</b>	<b>0.0138</b>	<b>0.014</b>	<b>0.0141</b>
40	<b>0.0162</b>	<b>0.0154</b>	<b>0.0145</b>	<b>0.0138</b>	<b>0.0137</b>	<b>0.0138</b>	<b>0.0139</b>	<b>0.014</b>	<b>0.0141</b>
50	<b>0.0144</b>	<b>0.0141</b>	<b>0.0138</b>	<b>0.0137</b>	<b>0.0138</b>	<b>0.0139</b>	<b>0.0139</b>	<b>0.014</b>	<b>0.0142</b>
100	<b>0.0139</b>	<b>0.0139</b>	<b>0.0139</b>	<b>0.014</b>	<b>0.014</b>	<b>0.014</b>	<b>0.014</b>	<b>0.0141</b>	<b>0.0142</b>
200	<b>0.0141</b>	<b>0.0141</b>	<b>0.0141</b>	<b>0.0141</b>	<b>0.0141</b>	<b>0.0141</b>	<b>0.0142</b>	<b>0.0142</b>	<b>0.0142</b>



Tabela A.5: Estimativas do teste de Kolmogorov-Smirnov para  $P_2$

		$n = 1000$		$r = 2$		$s = 1$			
$c_1 \setminus c_2$	1	5	10	20	30	40	50	100	200
1	0.2427	0.1375	0.0749	<b>0.0438</b>	<b>0.032</b>	<b>0.0258</b>	<b>0.022</b>	<b>0.0144</b>	<b>0.0139</b>
5	0.2462	0.0842	<b>0.056</b>	<b>0.0365</b>	<b>0.0282</b>	<b>0.0235</b>	<b>0.0205</b>	<b>0.0141</b>	<b>0.0139</b>
10	0.2294	0.0691	<b>0.0463</b>	<b>0.0315</b>	<b>0.0252</b>	<b>0.0215</b>	<b>0.019</b>	<b>0.0139</b>	<b>0.0139</b>
20	0.2195	<b>0.0594</b>	<b>0.0385</b>	<b>0.0265</b>	<b>0.0217</b>	<b>0.0189</b>	<b>0.0171</b>	<b>0.0137</b>	<b>0.014</b>
30	0.216	<b>0.0556</b>	<b>0.0351</b>	<b>0.024</b>	<b>0.0198</b>	<b>0.0174</b>	<b>0.0159</b>	<b>0.0138</b>	<b>0.014</b>
40	0.2142	<b>0.0536</b>	<b>0.0332</b>	<b>0.0225</b>	<b>0.0185</b>	<b>0.0164</b>	<b>0.0151</b>	<b>0.0138</b>	<b>0.014</b>
50	0.2131	<b>0.0523</b>	<b>0.032</b>	<b>0.0215</b>	<b>0.0177</b>	<b>0.0158</b>	<b>0.0146</b>	<b>0.0138</b>	<b>0.014</b>
100	0.2108	<b>0.0497</b>	<b>0.0294</b>	<b>0.0192</b>	<b>0.0158</b>	<b>0.0141</b>	<b>0.0137</b>	<b>0.0139</b>	<b>0.0141</b>
200	0.2097	<b>0.0483</b>	<b>0.028</b>	<b>0.0179</b>	<b>0.0146</b>	<b>0.0137</b>	<b>0.0138</b>	<b>0.014</b>	<b>0.0141</b>

		$n = 1000$		$r = 3$		$s = 1$			
$c_1 \setminus c_2$	1	5	10	20	30	40	50	100	200
1	0.3492	0.1978	0.104	0.0605	<b>0.044</b>	<b>0.0351</b>	<b>0.0296</b>	<b>0.0182</b>	<b>0.0137</b>
5	0.2761	0.1051	0.0655	<b>0.0424</b>	<b>0.0331</b>	<b>0.0279</b>	<b>0.0245</b>	<b>0.0166</b>	<b>0.0138</b>
10	0.2597	0.1051	0.0655	<b>0.0424</b>	<b>0.0331</b>	<b>0.0279</b>	<b>0.0245</b>	<b>0.0166</b>	<b>0.0138</b>
20	0.251	0.0958	<b>0.0572</b>	<b>0.0364</b>	<b>0.0287</b>	<b>0.0245</b>	<b>0.0217</b>	<b>0.0155</b>	<b>0.0138</b>
30	0.248	0.0922	<b>0.0539</b>	<b>0.0337</b>	<b>0.0265</b>	<b>0.0226</b>	<b>0.0201</b>	<b>0.0148</b>	<b>0.0138</b>
40	0.2466	0.0903	<b>0.0522</b>	<b>0.0322</b>	<b>0.0251</b>	<b>0.0214</b>	<b>0.0191</b>	<b>0.0142</b>	<b>0.0139</b>
50	0.2458	0.0891	<b>0.0511</b>	<b>0.0312</b>	<b>0.0242</b>	<b>0.0206</b>	<b>0.0184</b>	<b>0.014</b>	<b>0.0139</b>
100	0.2443	0.0867	<b>0.0488</b>	<b>0.029</b>	<b>0.0222</b>	<b>0.0188</b>	<b>0.0167</b>	<b>0.0138</b>	<b>0.014</b>
200	0.2435	0.0855	<b>0.0476</b>	<b>0.0278</b>	<b>0.021</b>	<b>0.0177</b>	<b>0.0157</b>	<b>0.0138</b>	<b>0.014</b>

Tabela A.6: Estimativas do teste de Kolmogorov-Smirnov para  $P_2$

		$n = 1000$			$r = 2$	$s = 2$			
$c_1 \setminus c_2$	1	5	10	20	30	40	50	100	200
1	0.3492	0.2781	0.2492	0.2304	0.2234	0.2198	0.2176	0.2176	0.2108
5	0.2781	0.0908	0.0655	<b>0.0495</b>	<b>0.043</b>	<b>0.0395</b>	<b>0.0372</b>	<b>0.0322</b>	<b>0.0295</b>
10	0.2492	0.0908	0.0655	<b>0.0495</b>	<b>0.043</b>	<b>0.0395</b>	<b>0.0372</b>	<b>0.0322</b>	<b>0.0295</b>
20	0.2304	0.0712	<b>0.0495</b>	<b>0.0364</b>	<b>0.0311</b>	<b>0.0281</b>	<b>0.0262</b>	<b>0.0218</b>	<b>0.0193</b>
30	0.2234	0.0637	<b>0.043</b>	<b>0.0311</b>	<b>0.0265</b>	<b>0.0238</b>	<b>0.0221</b>	<b>0.0182</b>	<b>0.0159</b>
40	0.2198	<b>0.0599</b>	<b>0.0395</b>	<b>0.0281</b>	<b>0.0238</b>	<b>0.0214</b>	<b>0.0199</b>	<b>0.0163</b>	<b>0.0142</b>
50	0.2176	<b>0.0575</b>	<b>0.0372</b>	<b>0.0262</b>	<b>0.0221</b>	<b>0.0199</b>	<b>0.0184</b>	<b>0.0152</b>	<b>0.0137</b>
100	0.2131	<b>0.0524</b>	<b>0.0322</b>	<b>0.0218</b>	<b>0.0182</b>	<b>0.0163</b>	<b>0.0152</b>	<b>0.0138</b>	<b>0.0139</b>
200	0.2108	<b>0.0497</b>	<b>0.0295</b>	<b>0.0193</b>	<b>0.0159</b>	<b>0.0142</b>	<b>0.0137</b>	<b>0.0139</b>	<b>0.014</b>

		$n = 1000$			$r = 3$	$s = 2$			
$c_1 \setminus c_2$	1	5	10	20	30	40	50	100	200
1	0.3618	0.3057	0.2676	0.2409	0.2307	0.2254	0.2221	0.2154	0.212
5	0.3157	0.1588	0.1105	0.0831	0.072	0.066	0.0624	<b>0.0551</b>	<b>0.0511</b>
10	0.279	0.1237	0.085	0.0601	<b>0.0505</b>	<b>0.0454</b>	<b>0.0421</b>	<b>0.0349</b>	<b>0.0309</b>
20	0.2607	0.1066	0.0682	<b>0.0462</b>	<b>0.0381</b>	<b>0.0336</b>	<b>0.0307</b>	<b>0.0243</b>	<b>0.0206</b>
30	0.2544	0.0998	0.0615	<b>0.0407</b>	<b>0.0331</b>	<b>0.029</b>	<b>0.0264</b>	<b>0.0206</b>	<b>0.0171</b>
40	0.2513	0.0962	<b>0.0579</b>	<b>0.0377</b>	<b>0.0304</b>	<b>0.0265</b>	<b>0.024</b>	<b>0.0185</b>	<b>0.0154</b>
50	0.2494	0.0939	<b>0.0557</b>	<b>0.0357</b>	<b>0.0286</b>	<b>0.0248</b>	<b>0.0224</b>	<b>0.0173</b>	<b>0.0144</b>
100	0.2459	0.0892	<b>0.0512</b>	<b>0.0314</b>	<b>0.0246</b>	<b>0.0211</b>	<b>0.0189</b>	<b>0.0146</b>	<b>0.0138</b>
200	0.2443	0.0868	<b>0.0488</b>	<b>0.0291</b>	<b>0.0223</b>	<b>0.0189</b>	<b>0.0169</b>	<b>0.0137</b>	<b>0.0139</b>

Tabela A.7: Estimativas do teste de Kolmogorov-Smirnov para  $P_2$

		$n = 1000$		$r = 3$		$s = 3$				
		1	5	10	20	30	40	50	100	200
1		0.4168	0.3538	0.2983	0.2712	0.2611	0.2561	0.2533	0.2476	0.2451
5		0.3538	0.2092	0.142	0.1167	0.1071	0.1019	0.0986	0.0917	0.088
10		0.2983	0.142	0.1036	0.0789	0.0691	0.0638	0.0604	<b>0.0536</b>	<b>0.0501</b>
20		0.2712	0.1167	0.0789	<b>0.056</b>	<b>0.0475</b>	<b>0.043</b>	<b>0.0401</b>	<b>0.0338</b>	<b>0.0303</b>
30		0.2611	0.1071	0.0691	<b>0.0475</b>	<b>0.0397</b>	<b>0.0355</b>	<b>0.0328</b>	<b>0.0269</b>	<b>0.0235</b>
40		0.2561	0.1019	0.0638	<b>0.043</b>	<b>0.0355</b>	<b>0.0315</b>	<b>0.0289</b>	<b>0.0233</b>	<b>0.0201</b>
50		0.2533	0.0986	0.0604	<b>0.0401</b>	<b>0.0328</b>	<b>0.0289</b>	<b>0.0265</b>	<b>0.0211</b>	<b>0.018</b>
100		0.2476	0.0917	<b>0.0536</b>	<b>0.0338</b>	<b>0.0269</b>	<b>0.0233</b>	<b>0.0211</b>	<b>0.0165</b>	<b>0.014</b>
200		0.2451	0.088	<b>0.0501</b>	<b>0.0303</b>	<b>0.0235</b>	<b>0.0201</b>	<b>0.018</b>	<b>0.014</b>	<b>0.0138</b>

		$n = 1000$		$r = 4$		$s = 4$				
$c_1 \setminus c_2$		1	5	10	20	30	40	50	100	200
1		0.4499	0.4112	0.3897	0.3728	0.3658	0.362	0.3596	0.3546	0.3519
5		0.4112	0.2904	0.2189	0.1666	0.1465	0.1395	0.1355	0.1272	0.1233
10		0.3897	0.2189	0.1386	0.1069	0.0948	0.0882	0.084	0.0752	0.0705
20		0.3728	0.1666	0.1069	0.076	0.0641	<b>0.0577</b>	<b>0.0537</b>	<b>0.0456</b>	<b>0.0412</b>
30		0.3658	0.1465	0.0948	0.0641	<b>0.0527</b>	<b>0.0469</b>	<b>0.0433</b>	<b>0.0355</b>	<b>0.0312</b>
40		0.362	0.1395	0.0882	<b>0.0577</b>	<b>0.0469</b>	<b>0.0413</b>	<b>0.0378</b>	<b>0.0303</b>	<b>0.0261</b>
50		0.3596	0.1355	0.084	<b>0.0537</b>	<b>0.0433</b>	<b>0.0378</b>	<b>0.0344</b>	<b>0.0271</b>	<b>0.023</b>
100		0.3546	0.1272	0.0752	<b>0.0456</b>	<b>0.0355</b>	<b>0.0303</b>	<b>0.0271</b>	<b>0.0204</b>	<b>0.0168</b>
200		0.3519	0.1233	0.0705	<b>0.0412</b>	<b>0.0312</b>	<b>0.0261</b>	<b>0.023</b>	<b>0.0168</b>	<b>0.0138</b>

Tabela A.8: Estimativas do teste de Kolmogorov-Smirnov para  $P_2$

	$n = 1000 \quad r = 3 \quad s = 4$								
$c_1 \setminus c_2$	1	5	10	20	30	40	50	100	200
1	0.4237	0.3989	0.3807	0.3672	0.3618	0.3589	0.3571	0.3532	0.3512
5	0.3766	0.253	0.192	0.151	0.1392	0.1344	0.1313	0.1253	0.1223
10	0.3183	0.1685	0.1209	0.0973	0.0878	0.0827	0.0796	0.0728	0.0692
20	0.281	0.1263	0.0892	0.066	<b>0.0571</b>	<b>0.0524</b>	<b>0.0495</b>	<b>0.0433</b>	<b>0.0401</b>
30	0.2683	0.1141	0.0764	<b>0.0542</b>	<b>0.0462</b>	<b>0.0419</b>	<b>0.0392</b>	<b>0.0333</b>	<b>0.03</b>
40	0.2613	0.1074	0.0695	<b>0.0481</b>	<b>0.0405</b>	<b>0.0364</b>	<b>0.0338</b>	<b>0.0282</b>	<b>0.025</b>
50	0.2572	0.1031	0.0652	<b>0.0444</b>	<b>0.0369</b>	<b>0.033</b>	<b>0.0305</b>	<b>0.025</b>	<b>0.0219</b>
100	0.2495	0.0941	<b>0.0559</b>	<b>0.0361</b>	<b>0.0291</b>	<b>0.0255</b>	<b>0.0232</b>	<b>0.0184</b>	<b>0.0158</b>
200	0.2459	0.0893	<b>0.0513</b>	<b>0.0315</b>	<b>0.0247</b>	<b>0.0213</b>	<b>0.0192</b>	<b>0.0149</b>	<b>0.0138</b>

Com base nas Tabelas, A.9 a A.12 dos p-value e fazendo a relação com as tabelas das estatísticas do teste de Kolmogorov-Smirnov para amostras de dimensão 1000, conseguimos obter uma melhor análise dos resultados. Assim para um nível de significância de 0.05, a hipótese da normalidade das distribuições não será rejeitada no teste, sempre que o valor observado da estatística seja inferior ou igual a 0.061.

Tabela A.9: p-value do teste de Kolmogorov-Smirnov para  $P_2$

$n = 1000 \quad r = 1 \quad s = 1$									
$c_1 \setminus c_2$	1	5	10	20	30	40	50	100	200
1	-	-	<b>0.0476</b>	0.5279	0.8473	0.956	0.9858	0.9904	0.9886
5	-	<b>0.0245</b>	0.2154	0.6912	0.9067	0.9723	0.9889	0.9903	0.9885
10	<b>0.0476</b>	0.2154	0.4796	0.8276	0.9502	0.9914	0.9917	0.9898	0.9885
20	0.5279	0.6912	0.8276	0.9502	0.9842	0.9914	0.9917	0.9898	0.9884
30	0.8473	0.9067	0.9502	0.9842	0.9914	0.9917	0.9912	0.9896	0.9883
40	0.956	0.9723	0.9842	0.9914	0.9917	0.9912	0.9908	0.9894	0.9882
50	0.9858	0.9889	0.9914	0.9917	0.9912	0.9908	0.9904	0.9892	0.9882
100	0.9904	0.9903	0.9901	0.9898	0.9896	0.9894	0.9892	0.9886	0.9879
200	0.9886	0.9885	0.9885	0.9884	0.9883	0.9882	0.9882	0.9879	0.9875
$n = 1000 \quad r = 2 \quad s = 1$									
$c_1 \setminus c_2$	1	5	10	20	30	40	50	100	200
1	-	-	-	<b>0.0013</b>	<b>0.0417</b>	0.1696	0.3447	0.8952	0.9916
5	-	-	-	<b>0.0162</b>	0.1186	0.29	0.4692	0.9234	0.9915
10	-	-	<b>4e-04</b>	0.0552	0.2218	0.417	0.5872	0.9459	0.9913
20	-	-	<b>0.0029</b>	0.1403	0.3823	0.5875	0.7329	0.9699	0.991
30	-	-	<b>0.0059</b>	0.2049	0.4851	0.6864	0.8116	0.9814	0.9908
40	-	-	<b>0.0086</b>	0.2512	0.5526	0.7473	0.8576	0.9876	0.9906
50	-	-	<b>0.0108</b>	0.285	0.5994	0.7874	0.8865	0.9899	0.9904
100	-	-	<b>0.017</b>	0.37	0.7073	0.8724	0.9424	0.9915	0.9899
200	-	-	<b>0.0215</b>	0.4231	0.7674	0.9138	0.9658	0.991	0.9894

Tabela A.10: p-value do teste de Kolmogorov-Smirnov para  $P_2$

$n = 1000 \quad r = 3 \quad s = 1$									
$c_1 \setminus c_2$	1	5	10	20	30	40	50	100	200
1	-	-	-	<b>0.0013</b>	<b>0.0417</b>	0.1696	0.3447	0.8952	0.9916
5	-	-	-	<b>0.0162</b>	0.1186	0.29	0.4692	0.9234	0.9915
10	-	-	<b>4e-04</b>	0.0552	0.2218	0.417	0.5872	0.9459	0.9913
20	-	-	<b>0.0029</b>	0.1403	0.3823	0.5875	0.7329	0.9699	0.991
30	-	-	<b>0.0059</b>	0.2049	0.4851	0.6864	0.8116	0.9814	0.9908
40	-	-	<b>0.0086</b>	0.2512	0.5526	0.7473	0.8576	0.9876	0.9906
50	-	-	<b>0.0108</b>	0.285	0.5994	0.7874	0.8865	0.9899	0.9904
100	-	-	<b>0.017</b>	0.37	0.7073	0.8724	0.9424	0.9915	0.9899
200	-	-	<b>0.0215</b>	0.4231	0.7674	0.9138	0.9658	0.9910	0.9894
$n = 1000 \quad r = 2 \quad s = 2$									
$c_1 \setminus c_2$	1	5	10	20	30	40	50	100	200
1	-	-	-	-	-	-	-	-	-
5	-	-	-	<b>1e-04</b>	<b>6e-04</b>	<b>0.0015</b>	<b>0.0027</b>	<b>0.0082</b>	<b>0.0142</b>
10	-	-	<b>4e-04</b>	<b>0.0149</b>	<b>0.0491</b>	0.0887	0.1261	0.251	0.3497
20	-	<b>1e-04</b>	<b>0.0149</b>	0.1403	0.2865	0.4065	0.4989	0.7293	0.852
30	-	<b>6e-04</b>	<b>0.0491</b>	0.2865	0.4851	0.6206	0.7123	0.8962	0.9626
40	-	<b>0.0015</b>	0.0887	0.4065	0.6206	0.7473	0.8244	0.9528	0.9873
50	-	<b>0.0027</b>	0.1261	0.4989	0.7123	0.8244	0.8865	0.9752	0.9919
100	-	<b>0.0082</b>	0.251	0.7293	0.8962	0.9528	0.9752	0.9915	0.9905
200	-	<b>0.0142</b>	0.3497	0.852	0.9626	0.9873	0.9919	0.9905	0.9894

Tabela A.11: p-value do teste de Kolmogorov-Smirnov para  $P_2$

		$n = 1000$		$r = 3$	$s = 2$				
$c_1 \setminus c_2$	1	5	10	20	30	40	50	100	200
1	-	-	-	-	-	-	-	-	-
5	-	-	-	-	<b>1e-04</b>	<b>3e-04</b>	<b>8e-04</b>	<b>0.0046</b>	<b>0.0108</b>
10	-	-	-	<b>0.0015</b>	<b>0.0121</b>	<b>0.0325</b>	0.0578	0.1753	0.2955
20	-	-	<b>2e-04</b>	<b>0.0279</b>	0.11	0.2088	0.3017	0.5954	0.7885
30	-	-	<b>0.001</b>	0.0725	0.2225	0.3689	0.4888	0.7921	0.9312
40	-	-	<b>0.0025</b>	0.117	0.3154	0.4855	0.6121	0.8817	0.9712
50	-	-	<b>0.004</b>	0.156	0.3882	0.5696	0.6948	0.9265	0.9858
100	-	-	<b>0.0105</b>	0.2769	0.5818	0.7663	0.8664	0.9838	0.9912
200	-	-	<b>0.0169</b>	0.3674	0.7025	0.8675	0.9386	0.9916	0.9902

		$n = 1000$		$r = 3$	$s = 3$				
$c_1 \setminus c_2$	1	5	10	20	30	40	50	100	200
1	-	-	-	-	-	-	-	-	-
5	-	-	-	-	-	-	-	-	-
10	-	-	-	-	<b>1e-04</b>	<b>6e-04</b>	<b>0.0013</b>	<b>0.0064</b>	<b>0.0133</b>
20	-	-	-	<b>0.0038</b>	<b>0.022</b>	<b>0.0498</b>	0.0804	0.2034	0.3174
30	-	-	<b>1e-04</b>	<b>0.022</b>	0.0854	0.1612	0.2326	0.4654	0.6371
40	-	-	<b>6e-04</b>	<b>0.0498</b>	0.1612	0.2756	0.3735	0.649	0.8139
50	-	-	<b>0.0013</b>	0.0804	0.2326	0.3735	0.4856	0.7647	0.9017
100	-	-	<b>0.0064</b>	0.2034	0.4654	0.649	0.7647	0.9492	0.9895
200	-	-	<b>0.0133</b>	0.3174	0.6371	0.8139	0.9017	0.9895	0.9909

Tabela A.12 p-value do teste de Kolmogorov-Smirnov para  $P_2$ 

$n = 1000 \quad r = 4 \quad s = 4$									
$c_1 \setminus c_2$	1	5	10	20	30	40	50	100	200
1	-	-	-	-	-	-	-	-	-
5	-	-	-	-	-	-	-	-	-
10	-	-	-	-	-	-	-	-	1e-04
20	-	-	-	-	<b>5e-04</b>	<b>0.0026</b>	<b>0.0062</b>	<b>0.0315</b>	0.0673
30	-	-	-	<b>5e-04</b>	<b>0.0077</b>	<b>0.0247</b>	<b>0.0472</b>	0.1606	0.2845
40	-	-	-	<b>0.0026</b>	<b>0.0247</b>	0.0657	0.1141	0.3169	0.5021
50	-	-	-	<b>0.0062</b>	<b>0.0472</b>	0.1141	0.1865	0.4535	0.6629
100	-	-	-	<b>0.0315</b>	0.1606	0.3169	0.4535	0.7981	0.9408
200	-	-	1e-04	0.0673	0.2845	0.5021	0.6629	0.9408	0.9909

$n = 1000 \quad r = 3 \quad s = 4$									
$c_1 \setminus c_2$	1	5	10	20	30	40	50	100	200
1	-	-	-	-	-	-	-	-	-
5	-	-	-	-	-	-	-	-	-
10	-	-	-	-	-	-	-	-	<b>1e-04</b>
20	-	-	-	<b>3e-04</b>	<b>0.0029</b>	<b>0.0083</b>	<b>0.015</b>	<b>0.0467</b>	0.0814
30	-	-	-	<b>0.0056</b>	<b>0.028</b>	0.0595	0.0924	0.2171	0.3275
40	-	-	<b>1e-04</b>	<b>0.0195</b>	0.075	0.1409	0.2029	0.4062	0.5606
50	-	-	<b>4e-04</b>	<b>0.0391</b>	0.1307	0.2272	0.3116	0.5595	0.7228
100	-	-	<b>0.0038</b>	0.147	0.364	0.5337	0.6522	0.8867	0.9647
200	-	-	<b>0.0104</b>	0.2729	0.5731	0.7556	0.8558	0.9788	0.9914

Os resultados destas tabelas mostram claramente que, quando o valor de  $c_1$  e  $c_2$  cresce, o coeficiente de variação é suficientemente baixo  $cv = 1/c$  é próximo de zero, pelo que  $P_1$  e  $P_2$  têm distribuição aproximadamente normal.



# Apêndice B

## Rotinas Utilizadas

As rotinas, utilizadas para obtenção dos resultados descritos nas tabelas A.1 a A.11 e 3.1 (Capítulos 3), foram elaboradas respectivamente em R (<http://www.r-project.org>) e em Fortran 2008 (vide, e.g., Metcalf et al. 2011).

### B.1 Estimativas do teste de Shapiro-Wilk para $P_1$

(Tabela A.2, rotina em R)

```
nsim<-1000
nv<-1000
c1<-c(1,5,10,20,30,40,50,100,200)
r1<-1:4
nc<-length(c1)
nr<-length(r1)
alfa<-0.05
{P<-function(x,cc,r)
return((cc+x)^r-cc^r) }
vn<-rnorm(nv)
pvm<-array(0,c(nc,nr))
for(ic in 1:nc)
{for(ir in 1:nr)
{for(i in 1:nsim) }
```

```

{xn<-P(rnorm(nv),c1[ic],r1[ir])
test<-shapiro.test(xn)
if(test$p.value<alfa)
{pvm[ic,ir]<-pvm[ic,ir]+1
hist(xn)
}
}
}
}
pvm<-pvm/nsim
pvm

```

## B.2 Estimativas do teste de Kolmogorov-Smirnov para

$P_1$

(tabelas A.1 e A.3; rotina em R)

```

sigma<-1
n<-1000
c1<-c(1,5,10,20,30,40,50,100,200)
# Geração dos NPAs de uma variável aleatória normal
x<-rnorm(1000,0,sigma)
res<-array(10,dim=c(4,9))
resp<-array(10,dim=c(4,9))
for (r in 1:4){
for(j in 1:9){
cc1<-c1[j]
#cálculo dos valores de  $P1=(cc1+x)^r-(cc1)^r$ 
P1<-1:n
for(i in 1:n){
P1[i]<-((cc1+x[i])^r-(cc1)^r)

```

```

}
hist(P1)
# A matriz resguarda os resultados do teste de K-S para P1
# A matriz resguarda os resultados dos p-values do teste de K-S para P1
ks<-ks.test(P1,"pnorm",mean(P1),sd(P1))
res[r-1,j]<-round(ks$statistic,4)
resp[r-1,j]<-round(ks$p.value,4)
}
}
latex<-rbind(c1,res)
latex<-cbind(c("r",1,2,3,4),latex)
In cbind(c("r", 1,2, 3, 4), latex)
# Escrita em ficheiro dos valores obtidos na simulação
write("Valores das estatísticas do teste de Kolmogorov-Smirnov para uma
amostra",file="Teste-statisticsP1.txt")
write(t(latex),file="Teste-statisticsX.txt",append=TRUE,ncol=10,sep="&")
latex<-rbind(c1,resp)
latex<-cbind(c("r",1,2,3,4),latex)
# Escrita em ficheiro dos valores obtidos na simulação
write("Valores dos p-values do teste de Kolmogorov-Smirnov para uma
amostra",file="Teste-p-valueP1.txt")
write(t(latex),file="Teste-p-valueX.txt",append=TRUE,ncol=10,sep="&")

```

## B.3 Estimativas do teste de Kolmogorov-Smirnov para

$P_2$ .

(Tabelas A.4 a A.11, rotina em R)

```

sigma<-1
n<-1000
## Geração de NPAs para variáveis normais independentes
x<-rnorm(n,0,sigma)

```

```

y<-rnorm(n,0,sigma)
# Vectors dos valores dos parametros cc1 e ccc2
c1<-c(1,5,10,20,30,40,50,100,200)
c2<-c(1,5,10,20,30,40,50,100,200)
res<-array(10,dim=c(4,4,9,9))
cont<-0
for (s in 1:4){
  for (r in s:3){
    for(i in 1:9){
      for(j in 1:9){
        cc1<-c1[i]
        ccc2<-c2[j]
        #cálculo dos valores de P2=(cc1+x)^r*(ccc2+y)^s
        P2<-1:n
        for(k in 1:n){
          P2[k]<-((cc1+x[k])^r)*(ccc2+y[k])^s
        }
        #hist(P2)
        #res[r,s,i,j] guarda os valores da estatística
        # do teste de ks para r,s, cc1=c[i] e ccc2=c[j]
        ks<-ks.test(P2,"pnorm",mean(P2),sd(P2))
        #res[r,s,i,j]<-ks$statistic
        res[r,s,i,j]<-round(ks$p.value,4)
      }
    }
  }
  latex<-rbind(c2,res[r,s,,])
  latex<-cbind(c("ccc2",c1),latex)
  # Escrita em ficheiro das matrizes res[r,s,,]
  write(paste("r=",r,"s=",s),file="D-statisticsP2.txt",append=TRUE)
  write("ccc2/cc1",file="D-statisticsP2.txt",append=TRUE)
  write(latex,file="D-statisticsP2.txt",append=TRUE,ncol=10,sep="&")
}
}

```

## B.4 Número de Termos

(Tabela 3.1; rotina em Fortran 2008)

```
module kinds
  implicit none
  private
  public :: quad
  integer, parameter :: quad = selected_real_kind(18) ! Precisão quad
end module kinds
```

```
program numero_termos
  use kinds
  implicit none
  integer :: i,n,kdelta,kepsilon, istart
  real(quad)::prodi
  real(quad):: delta2,sum,erro
  character(len=17):: fileout
  integer,dimension(1:6)::delta
  real(quad),dimension(1:6)::epsilon
  integer,dimension(1:6,1:6)::ntermos
  data delta/5, 10, 15,20,50,100/
  data epsilon/1e-2,1e-3,1e-4,1e-5,1e-6,1e-8/
  ntermos=0
  write(*,*) 'Indique o indice inicial do somatorio?'
  read (*,*) istart
  write(fileout,'(a11,i2.2,a4)') 'Termos-i_0=',istart,'.dat'
  open(10,file=fileout,status='unknown')
  L2: do kepsilon=1,6
    erro=(1.0-epsilon(kepsilon))
    L1: do kdelta=1,6
```

```

delta2=0.5*delta(kdelta)
prodi=1.0
sum=0.0
write(10,*) 'kdelta',kdelta
do i=istart,10000
  if(i==0) then
    prodi=1.0 ! Fatorial de i
  else
    prodi=prodi*i
  endif
  sum=sum+delta2**i/prodi ! somatorio
  if(prodi>=1e+900_quad) then
    ntermos(kdelta,kepsilon)=i
    ! numero de termos a partir
    ! dos quais o fatorial e' > 1e900
    cycle L1
  endif
  write(10,'(i3.3,2x,2(es16.9e3,2x),f12.10,2x,f10.8))' ) &
    i,prodi,sum,sum*exp(-delta2),erro
  if(sum*exp(-delta2)>erro) then
    ntermos(kdelta,kepsilon)=i
    cycle L1
  endif
enddo
enddo L1
enddo L2
! Escreve os termos em forma matricial (delta vs epsilon)
write(10,'(20x,6(es7.1,4x))' ) (epsilon(kepsilon),kepsilon=1,6)
do kdelta=1,6
  write(10,*) delta(kdelta),(ntermos(kdelta,kepsilon),kepsilon=1,6)
enddo
end program numero_termos

```

# Bibliografia

- Areia, A., 2009, *Séries Emparelhadas de Estudos: Estrutura Comum e Inferência*, Tese de Doutoramento, Universidade de Évora.
- Areia, A., Oliveira, M. M & Mexia, J. T., 2008, *Models for a Series of Studies Based on Geometrical Representation*, *Statistical Methodology*, 5, 277-288.
- Areia, A., Oliveira, M. M & Mexia, J. T., 2011, *Modelling the Compromise Matrix in STATIS*, *Methodology*, 5, 277-288.
- Azzalini, A., 1996, *Statistical Inference*, Monographs on Statistics and Applied Probability 68, Chapman & Hall, New York.
- Bilingsley, P., 1968, *Convergence of Probability Measure*, John Wiley & Sons, New York.
- Calinski, T. & Lejeune, M., 1998, *Dimensionality in Manova Tested by a Closed Testing Procedure*, *Journal of Multivariate Analysis* 65, 181-194.
- Cantarinha, A., 2012, *Resultados Assintóticos para Famílias Estruturadas de Modelos Colectivos. Aplicação aos Fogos Florestais em Portugal Continental*, Tese de Doutoramento, Universidade de Évora.
- Carvalho, L. & Diamantino, F., 1993, *Epidemia de SIDA em Portugal, Projeções pelo Método de Back - Calculation*, in *A Estatística e o Futuro da Estatística*, D. Pestana (Ed), Lisboa: Sociedade Portuguesa de Estatística e Edições Salamandra, 19-29.
- Christensen, R., 1987, *Plane Answers to Complex Questions*, The Theory of Linear Models, Springer-Verlag.
- Courant, R. & Hilbert, D., 1953, *Methods of Mathematical Physics*, Interscience Publishers.
- Dias, C., 2000, *Análise Conjunta Pesada de Regressões*, Tese de Mestrado, Universidade de Évora.
- Dias, C., Oliveira, M. M. & Mexia, J. T., 2012, *Models of Symmetric Stochastic Matrices*, in *AIP Conf. Proc.* 1479, 1666-1669.
- Eberhart & Russel, 1966, *Stability Parameters for Comparing Varieties*, *Crop Science*, Society of America, 6, 3640.
- Escoufier, Y., 1973, *Le Traitement des Variables Vectorielles*, *Biometrics*, 29(4), 751-760.
- Escoufier, Y. & L'Hermier H., 1978, *A Propos de la Comparaison Graphique des Matrices de Variance*, *Biom. J.*, 20(5), 477-483.

- Finlay, R. & Wilkinson, G. N., 1963, *The analysis of adaptation in a plant-breeding programme*, Aust. J. Agric. Res., 14: 742 - 754.
- Fisher, R. & Yates, F., 1963, *Statistical Tables for Biological Agricultural and Medical Research*, Sixth Edition, Olivier and Boyd.
- Fonseca, M., Mexia, J. T. & Zmyslosny, R., 2003, *Estimators and tests for variance components in cross-nested orthogonal models*, *Discussiones Mathematicae, Probability and Statistics*, 23(2), 175-201.
- Fonseca, M., Mexia, J. T. & Zmyslosny, R., 2007, *Jordan Algebras, Generating Pivot Variables and Orthogonal Normal Models*, *Journal of Interdisciplinary Mathematics* 1, 305-326.
- Fonseca, M., Mexia, J. T. & Zmyslosny, R., 2008, *Inference in normal models with commutative orthogonal block structure*, *Acta et Commentationes Universitatis Tartuensis de Mathematica*, 3-16.
- Fonseca, M., Jesus, V., Mexia, J. T. & Zmyslosny, R., 2009, *Binary operations and canonical forms for factorial and related models*, *Linear Algebra and its Applications*, 430, 2781-2797.
- Foucart, T., 1981, *Suites de Tableaux et de Sous-Tableaux*, *Revue de Statistique Appliquée*, XXIX 2, 31-42.
- Gusmão, L., Mexia, J. T. & Baeta, J., 1992, *Trimmed Joint Regression: a new approach to the Joint Regression Analysis for cultivar relative performance evaluation*, *Theoretical and Applied Genetics*, 84: 735 - 738.
- Hanning, R. W., 1962, *Numerical Methods for Scientists and Engineers*, McGraw-Hill, New York.
- Foucart, T., 1981, *Suites de Tableaux et de Sous-Tableaux*, *Revue de Statistique Appliquée*, XXIX 2, 31-42.
- Harville, A., 1998, *Matrix Algebra from a Statistician's Perspective*, Springer-Verlag.
- Henry, R. J., 2007, *Genomics as a Tool for Cereal Chemistry*, *Cereal Chemistry Journal*, 84, 365-369.
- Horn, R. & Johnson C., 1985, *Matrix Analysis*, Cambridge University Press.
- Imhof, J., P, 1961, *Computing the Distribution of Quadratic Forms in Normal Variables*. *Biometrics* 48 (3-4), 419-426.
- Ito, K., 1980, *Robustness of Anova and Macanova Test Procedures*, *Handbook for Statistics*, P.R. Krishnaiah (Ed), North Holland.
- Johannsen, W. L., 1911, *The genotype conception of heredity*, *The American Naturalist*, 45, N° 531, 129-159, Chicago.
- Kearseyn, I. J. & Jinks, J. L., 1968, *A general method of detecting additive, dominance and epistatic variation for metrical traits: Theory*, *Heredity* I., 23, 403-409.



- Khuri, A. I., Mathew, T. & Sinha., B.K., 1998, *Statistical Tests for Mixed Linear Models*, John Willey & Sons, New York.
- Khuri, A. I., 2003, *Advanced Calculus with Applications in Statistics*, second ed. Wiley, New York.
- Lavit, C., 1988, *Analyse Conjointe de Tableaux Quantitatifs*, Collection Methods Programmes, Masson, Paris.
- Layard, J., 1973, *Robust Large-Sample Tests for Homogeneity of Variances*, Journal of the American Statistical Association, 341, 195-198.
- Lehmann, E. L., 1959, *Testing Statistical Hypothesis*, John Willey & Sons, New York.
- L'Hermier des Plantes H., 1976, *Structuration des Tableaux à Trois Indices de la Statistique*, Thèse de 3ème cycle, Université de Montpellier II.
- Metcalf, M., Reid, J., & Cohen, M., 2012, *Modern Fortran Explained*, 7ª Edição, Oxford University Press, New York
- Mexia, J. T., 1987, *Multitreatment regression designs*, Trabalhos de Investigação Nº1, Departamento de Matemática, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa
- Mexia, J. T., 1988a, *Estimable Functions, Duality of F Tests and Scheffé's Multiple Comparison Method*, Trabalhos de Investigação Nº1, Departamento de Matemática, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa.
- Mexia, J. T., 1988b, *Standardized Orthogonal Matrices and the Decomposition of the Sum of Squares for Treatments*, Trabalhos de Investigação Nº2, Departamento de Matemática, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa.
- Mexia, J. T., 1989, *Simultaneous Confidence Intervals: Generalization of the Scheffé Theorem*, Trabalhos de Investigação Nº2, Departamento de Matemática, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa.
- Mexia, J. T., 1990, *Variance free models*, Trabalhos de Investigação Nº2, Departamento de Matemática, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa.
- Mexia, J. T., 1990, *Best Linear Unbiased Estimates, Duality of F Tests and the Scheffé Multiple Comparison Method in Presence of Controlled Heterocedasticity*, Comp. Stat & Data Analysis, 10(3), 271-281.
- Mexia, J. T., 1992, *Asymptotic chi-squared tests, design and log-linear models*, Trabalhos de Investigação, Nº1, Departamento de Matemática, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa.
- Mexia, J. T., 1995, *Introdução à Inferência Estatística Linear*, Centro de Estudos de Matemática Aplicada, Edições Universitárias Lusófonas, Lisboa.
- Mexia, J. T., 2005, *Modelos Normais Ortogonais: dos Anos 70 ao Presente*. Memorial of the Portuguese Statistical Society, pp. 179-198.

- Mexia, J. T. & Oliveira, M. M., 2011, *Asymptotic linearity and limit distributions, approximations*, Journal of Statistical Planning and Inference, 140(2), 353-357.
- Montgomery, Douglas, C., 2005, *Design and Analysis of Experiments*, 6ª Edição, John Wiley & Sons, New York.
- Mood, A., Graybill R. & Boes D., 1987, *Introduction to the Theory of Statistics*, 3ª Edição, MacGraw-Hill.
- Mooers, C. A., 1921, *The agronomic placement of varieties*. Journal of the American Society of Agronomy 13, 337-352.
- Moreira, E., E., Ribeiro, A., B., Mateus, E., P., Mexia, J., T., Ottosen, L., M., 2005, *Regression modeling of electro-dialytic removal of Cu, Cr and As from CCA timber waste: application to sawdust*. Wood Sci Technol 39(4):291–309.
- Moreira, E., E., Ribeiro, A., B., Mateus, E., P., Mexia, J., T., Ottosen, L., M., 2005, *Regression modeling of electro-dialytic removal of Cu, Cr and As from CCA timber waste: application to wood chips*. Biometri Lett 4(1):11–25.
- Moreira, E., E., Mexia, J., T., 2007, *Multiple regression models with cross nested orthogonal base model*. In: Proceedings of the 56th session of the ISI 2007—International Statistical Institute, Lisboa
- Moreira, E. E., 2008, *Famílias Estruturadas de Modelos com Modelo Base Ortogonal*, Tese de Doutorado, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa.
- Ng, M. P. & Grunwald, G. K., 1997, *Nonlinear Regression Analysis of the Joint Regression Model*, Biometrics, 53, 1366-1372.
- Nunes, C., Oliveira, M., M., & Mexia, J., T., 2013, *Applications Domains for the Delta Method*. Statistics, 47(2), 317-328.
- Nunes, C., Ferreira, D., Ferreira S., S. & Mexia, J., T., 2012, *Controlo f the Truncation Errors for Generalized F Distribution*, vol 82(2), 165-171.
- Oliveira, T., 1980, *Statistical Choice of Univariate Extreme Models in Statistical Distribution Scientific Work*, vol. C. Taillie et al. (eds.), 367-382, Reichel Dordrécht.
- Oliveira, M. M. & Mexia, J., 1998, *Tests for the Rank of Hilbert-Schmidt Product Matrices, in Advances Data Science and Classifications*, Rizzi, Vichi & Bock, (Ed.), 619-625, Springer.
- Oliveira, M. M. & Mexia J., 1999a, *F Tests for Hypothesis on the Structure Vectors of Series*, Discussiones Mathematicae, 19(2), 345-353.
- Oliveira, M. M. & Mexia J., 1999b, *Multiple Comparisons for Rank one Common Structures*, Biometrical Letters, 36(2), 159-167.
- Oliveira, M. M., 2002, *Modelação de Series Emparelhadas com Estrutura Comum*, Tese de Doutorado, Universidade de Évora.
- Oliveira, M.M., Mexia, T.J., 2004, *AIDS in Portugal: endemic versus epidemic forecasting scenarios for mortality*, International Journal Forecasting 20, 131–137.

- Oliveira, M. M. & Mexia, J. T., 2007a, *ANOVA like analysis of matched series of studies with a common structure*, Journal of Statistical Planning and Inference, 137, 1862-1870.
- Oliveira, M. M. & Mexia, J. T., 2007b, *Modelling series of studies with a common structure*, Computational Statistics & Data Analysis, 51, 5876-5885.
- Perkins, Jean M, and Jinks, J L. 1968a, *Environmental and genotype-environmental components of variability*, III. Multiple lines and crosses, Heredity, 23, 339-356.
- Perkins, Jean M, and Jinks, J L. 1968b, *Environmental and genotype-environmental components of variability*, IV. Non-linear interactions for multiple inbred lines, Heredity, 23, 525-535.
- Perkins, J.M. & Jinks, J.L., 1966, *Environmental and genotype-environmental components of variability*, 111 Multiple lines and crosses: Heredity, 23; 339-356.
- Pinto, I., 2006, *Análise Conjunta de Regressões e Planos de Melhoramento*, Tese de Doutorado, Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa.
- Pollock, D. S. G., 1975, *The Algebra of Econometrics*, John Wiley & Sons, New York.
- Ramos, L., 2007, *Quase Normalidade e Inferência para Séries de Estudos Emparelhados*, Tese de Doutorado, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa.
- Robbins, H., 1948, *Mixture of Distributions*. The Annals of Mathematics Statistics, 19, 360-369.
- Robbins, H. & Pitman, E., J., G., 1949, *Application of the method of mixtures to quadratic forms in normal variates*, The Annals of mathematical Statistics, 20, 552-560.
- Rao, C. R., 1973, *Linear Statistical Inference and Its Applications*, 2ª Edição, John Wiley & Sons, New York.
- Rohatgi, V. K., 1976, *An Introduction to Probability and Mathematical Statistics*, John Wiley & Sons, New York.
- Scheffé, H., 1959, *The Analysis of Variance*, John Wiley & Sons, New York.
- Schott, J. R., 1997, *Matrix Analysis for Statistics*, Wiley Series in Probability and Statistics, New York.
- Seber, G. A. F., 1980, *The Linear Hypothesis, A General Theory*, 2ª Edição, Charles Griffin & Co, London.
- Silvey, S. D., 1975, *Statistical Inference*, Chapman & Hall, London.
- Steeb, W. H. 1991, *Kronecker product of Matrices and Applications*, Berdichsbibl Mathematik. Dortmund.
- Vaquinhas, R. & Mexia, J., 1992, *Convergence of Matrices and Subspaces with Statistical Applications*, Distância 92-Rennes.
- Wilhelm, Johannsen., L., 1911, “The Genotype Conception of [Heredity](#).” *The American Naturalist* 45, 129–59.
- Wilks, S., 1961, *Mathematical Statistics*, John Wiley & Sons, New York.

Yates, F. & Cochran, W. G., 1938, *The analysis of groups of experiments*, The Journal of Agricultural Science, 28, 556-580.