

# **Modelos Aditivos – Ajustamento, Elipsóides de Confiança e Intervalos de Predição**

Patrícia Alexandra Robalo Antunes

Tese para obtenção do Grau de Doutor em  
Matemática e Aplicações  
(3.º ciclo de Estudos)

# **Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição**

**Patrícia Alexandra Robalo Antunes**

Tese para obtenção do Grau de Doutor em  
**Matemática e Aplicações**  
(3º ciclo de estudos)

Orientadora: Prof. Doutora Sandra Maria Bargão Saraiva Ferreira  
Co-orientador: Prof. Doutor Dário Jorge da Conceição Ferreira

Provas de Doutoramento:  
28 de janeiro de 2022

## **Júri:**

Presidente: Prof. Doutor Joaquim Mateus Paulo Serra  
1º arguente: Prof. Doutor Carlos Manuel Rebelo Tenreiro da Cruz  
2ª arguente: Prof. Doutora Susana Margarida Ferreira de Sá Faria  
Prof. Doutora Isabel Maria Simões Pereira  
Prof. Doutora Raquel Menezes Mota Leite  
Prof. Doutora Célia Maria Pinto Nunes  
Prof. Doutora Sandra Maria Bargão Saraiva Ferreira

**fevereiro de 2022**



## **Agradecimentos**

É com alegria, entusiasmo e um sentimento profundo de gratidão que escrevo estes agradecimentos. Ao longo deste caminho de aprendizagem e partilha, repleto de desafios, onde a conclusão da elaboração desta tese é, acima de tudo, uma enorme conquista individual, mas a qual seria impossível de acontecer sem o auxílio e a colaboração de diversas pessoas, merecedoras dos meus sinceros agradecimentos.

Começo por expressar o meu mais profundo agradecimento aos meus orientadores. Manifesto aqui o mais genuíno agradecimento à Professora Doutora Sandra Ferreira, a minha orientadora, pela sua inigualável disponibilidade e que de forma sempre muito sensata me orientou neste trabalho. Pela compreensão das minhas dificuldades, pela minha vida profissional complicada e pela falta de tempo, e que em alguns momentos dificultaram a realização desta dissertação. Pela paciência de ouvir-me e ajudar a ultrapassar as muitas angústias deste processo. Por todas as oportunidades de aprendizagem que me proporcionou e que tanto ajudaram na minha evolução, como profissional e investigadora, contribuindo de maneira decisiva para o meu crescimento nesta área. Mas, acima de tudo, agradeço a sua amizade.

Ao meu coorientador, Professor Doutor Dário Ferreira, pelo seu apoio contínuo, amizade e paciência, de ser tão compreensivo com as minhas dificuldades, prontificando-se sempre a ajudar-me, e a orientar-me, e, com as suas ideias, foi possível chegar ao término deste trabalho.

Tenho a sorte de ter duas pessoas maravilhosas como orientadores. Agradeço aos dois todo o tempo que comigo, pensaram, questionaram e trabalharam e por serem um exemplo tanto a nível pessoal como profissional. Para além da orientação na minha investigação, agradeço também o vosso incentivo para a disseminação de resultados em conferências, para a partilha em seminários, para o contacto com outros investigadores. Tem sido uma longa jornada e não teria alcançado a meta sem a vossa ajuda e, foi de facto, um privilégio único caminharmos juntos.

O meu reconhecido agradecimento ao Professor Doutor João Tiago Mexia, pelo apoio, a disponibilidade, a amizade, o encorajamento, a bondade, a partilha do saber e as suas valiosas contribuições, que muito enriqueceram, não apenas para o trabalho, mas que contribuíram para o meu crescimento científico. Foi uma honra trabalhar consigo. Sou muito grata.

Ao Jorge por acreditar em mim, sempre. Que em muitos momentos me apoiou e me ajudou nesta difícil empreitada, incentivando-me sempre para que não desistisse.

À minha família pelo seu incentivo e carinho, em especial à minha avó.

A todos, os meus sinceros agradecimentos, BEM HAJAM!



## Resumo

Modelos aditivos são dados por

$$\underline{Y} = \underline{X}\underline{\beta} + \sum_{i=1}^w \underline{X}_i \underline{Z}_i,$$

onde  $\underline{X}$  e  $\underline{\beta}$  são fixos e os  $\underline{Z}_1, \dots, \underline{Z}_w$  são independentes com  $c_1, \dots, c_w$  componentes *i.i.d.*, com valor médio nulo, e cumulantes  $\chi_{2,i}, \chi_{3,i}$  e  $\chi_{4,i}$  de segunda, terceira e quarta ordens, respectivamente. Numa primeira fase, consideramos o caso geral sem pressupostos extras e consideramos pares de modelos independentes e identicamente distribuídos por forma a obter Estimadores de Mínimos Quadrados para os cumulantes. Usamos também esses pares de modelos para obter Estimadores de Mínimos Quadrados Generalizados para o vetor  $\underline{\beta}$  de coeficientes e para as funções estimáveis  $\underline{c}^\top \underline{\beta}$ . Utilizando as propriedades dos cumulantes, obteremos estimadores para os cumulantes de

$$W(\underline{u}) = \underline{u}^\top (\underline{Y} - \underline{\mu}),$$

onde  $\underline{\mu} = \underline{X}\underline{\beta}$ , e que serão depois utilizados para obter elipsóides de confiança para  $\underline{Y}$ . De seguida obtemos intervalos de predição para o vetor de observações  $\underline{Y}$ , correspondente aos conjuntos de valores da variável que, por sua vez, correspondem às colunas das matrizes  $\underline{X}, \underline{X}_1, \dots, \underline{X}_w$ . A seguir, consideramos modelos múltiplos, nos quais temos um par de modelos para cada tratamento de um delineamento base e consideramos as matrizes  $\underline{X}, \underline{X}_1, \dots, \underline{X}_w$  e os cumulantes  $\chi_{r,i}, r = 1, 2, 3, 4, i = 1, \dots, w$ , iguais para todos os pares de tratamentos. Concluímos o estudo do caso geral evidenciando duas situações:

- o primeiro, em que as componentes dos  $\underline{Z}_1, \dots, \underline{Z}_w$  têm parâmetros de localização e dispersão;
- o segundo, em que as componentes dos  $\underline{Z}_1, \dots, \underline{Z}_w$  têm parâmetros de localização, dispersão e forma.

## Palavras-chave

Cumulantes, Estimadores de Mínimos Quadrados, Estimadores de Mínimos Quadrados Generalizados, Expansões de Edgeworth, Elipsóides de Confiança, Independência, Intervalos de Predição, Modelos Múltiplos.



## Abstract

Bi-additive Models are given by

$$\underline{Y} = \underline{X}\underline{\beta} + \sum_{i=1}^w \underline{X}_i \underline{Z}_i,$$

where  $\underline{X}$  and  $\underline{\beta}$  are fixed and the  $\underline{Z}_1, \dots, \underline{Z}_w$  are independent with  $c_1, \dots, c_w$  independent and identically distributed components with null mean value, and cumulants  $\chi_{2,i}, \chi_{3,i}$  and  $\chi_{4,i}$  of the second, third and fourth orders, respectively. In a first phase, we consider the general case without extra assumptions and consider pairs of independent and identically distributed models in order to obtain Least Square Estimator, *LSE*, for the cumulant and Generalized Least Square Estimator, *GLSE*, for the vector  $\underline{\beta}$  of coefficients and estimable functions  $\underline{c}^T \underline{\beta}$ . Using the cumulants properties we obtain estimators for the cumulants of

$$W(\underline{u}) = \underline{u}^T (\underline{Y} - \underline{\mu}),$$

where  $\underline{\mu} = \underline{X}\underline{\beta}$ , and use them to obtain Confidence Ellipsoids for  $\underline{Y}$ . We also obtain prediction intervals for the value of  $\underline{Y}$  that refer to sets of values of the variable corresponding to the columns of matrices  $\underline{X}, \underline{X}_1, \dots, \underline{X}_w$ . Next we consider Multiple Models in which we have a pair of models for each treatment of a base design. The matrices  $\underline{X}, \underline{X}_1, \dots, \underline{X}_w$  and, the cumulants  $\chi_{r,i}, r = 1, 2, 3, 4, i = 1, \dots, w$ , are the same for all pairs of treatments. We then complete the study of the general case we assume:

- first, that the components of the  $\underline{Z}_1, \dots, \underline{Z}_w$  have parameters of location, and dispersion;
- second, that the components of the  $\underline{Z}_1, \dots, \underline{Z}_w$  have location, dispersion and shape parameters.

## Keywords

Cumulants, Confidence Ellipsoids, Edgeworth Expansions, Generalized Least Square Estimators, Independence, Least Square Estimators, Prediction Intervals and Multiple models.



# Índice

<b>1</b>	<b>Introdução</b>	<b>1</b>
<b>2</b>	<b>Resultados preliminares</b>	<b>9</b>
2.1	Matrizes . . . . .	9
2.1.1	Matrizes Transpostas . . . . .	9
2.1.2	Matrizes Ortogonais . . . . .	10
2.1.3	Matrizes Simétricas . . . . .	11
2.1.4	Matrizes Inversas . . . . .	12
2.1.5	Matrizes Inversas Generalizadas . . . . .	13
2.1.6	Inversas de Moore-Penrose . . . . .	13
2.1.7	Matrizes de covariância e de covariância cruzada . . . . .	15
2.1.8	Matrizes de projeção ortogonal . . . . .	17
2.1.9	Produto de Kronecker de matrizes . . . . .	21
2.2	Funções Geradoras de Cumulantes . . . . .	23
2.3	Expansões de Edgeworth . . . . .	27
2.4	Partições Ortogonais de Sub-Espaços . . . . .	27
2.4.1	Considerações prévias . . . . .	27
2.4.2	Cruzamento e Aninhamento . . . . .	29
2.4.3	Fatoriais de base prima . . . . .	30
<b>3</b>	<b>Caso Geral de Modelos Aditivos</b>	<b>33</b>
3.1	Ajustamentos . . . . .	33
3.2	Elipsóides de Confiança . . . . .	35
3.3	Intervalos de Predição . . . . .	38
3.4	Modelos Múltiplos . . . . .	39
<b>4</b>	<b>Modelos Aditivos com Parâmetros de Localização e Dispersão</b>	<b>43</b>
4.1	Ajustamentos . . . . .	43
4.2	Elipsóides de Confiança . . . . .	44
4.3	Intervalos de Predição . . . . .	45
4.4	Modelos Múltiplos . . . . .	45
4.5	Casos Particulares . . . . .	46
4.5.1	Tipo Normal . . . . .	46
4.5.2	Tipo Exponencial . . . . .	46
4.5.3	Tipo Gumbel . . . . .	47
<b>5</b>	<b>Modelos Aditivos com parâmetros de Localização, Dispersão e Forma</b>	<b>49</b>
5.1	Considerações Preliminares . . . . .	49
5.2	Ajustamentos . . . . .	50
5.3	Elipsóides de Confiança . . . . .	51
5.4	Intervalos de Predição . . . . .	53
5.5	Modelos Múltiplos . . . . .	53
5.6	Casos particulares . . . . .	54
5.6.1	Tipo Gama . . . . .	54

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

5.6.2	Tipo Weibull . . . . .	55
5.6.3	Tipo Fréchet . . . . .	56
<b>6</b>	<b>Duas Aplicações através de simulações</b>	<b>59</b>
6.1	Primeiro estudo de simulação . . . . .	60
6.2	Segundo estudo de simulação . . . . .	61
6.2.1	Primeiro caso . . . . .	61
6.2.2	Segundo caso . . . . .	64
6.3	Exemplo de Aplicação de modelos múltiplos aditivos . . . . .	67
<b>7</b>	<b>Observações Finais</b>	<b>69</b>
7.1	Revisão Global . . . . .	69
7.2	Conclusões e Trabalhos Futuros . . . . .	71
	<b>Bibliografia</b>	<b>73</b>
<b>A</b>	<b>Anexos</b>	<b>79</b>
A.1	Álgebras de Jordan . . . . .	79
A.1.1	Álgebras de Jordan Comutativas . . . . .	80
A.2	Momentos . . . . .	81
A.3	Função Geradora de Momentos . . . . .	82

## Lista de Figuras

- 3.1 Elipsóide que passa pela direção dada por  $\underline{u}$ , à distância  $\underline{u}^\top \tilde{K}_p \underline{u}$  do centro. . . . . 38



## Lista de Tabelas

6.1	Médias obtidas para as estimativas. . . . .	60
6.2	Valores previamente considerados de $\chi_{2,i}$ , $i = 1, 2$ , e médias das estimativas das componentes de variância obtidas. . . . .	62
6.3	Médias obtidas para $\tilde{\beta}$ . . . . .	62
6.4	Número de vezes, em 1000, que $Y_f$ está dentro dos contornos do elipsóide. . . . .	63
6.5	Valores médios obtidos para $[Y_{fi}; Y_{fs}]$ . . . . .	64
6.6	Número de vezes, em 1000, que $Y_f$ estava contida em $[Y_{fi}; Y_{fs}]$ . . . . .	64
6.7	Valores previamente considerados de $\chi_{2,i}$ , $i = 1, 2$ , e médias das estimativas das componentes de variância obtidas. . . . .	65
6.8	Médias obtidas para $\tilde{\beta}$ . . . . .	65
6.9	Número de vezes, em 1000, que $Y_f$ está dentro dos contornos do elipsóide. . . . .	66
6.10	Valores médios obtidos para $[Y_{fi}; Y_{fs}]$ . . . . .	66
6.11	Número de vezes, em 1000, que $Y_f$ estava contida em $[Y_{fi}; Y_{fs}]$ . . . . .	67



## Lista de Acrónimos

$LSE$	Estimadores de Mínimos Quadrados .....	5
$GLSE$	Estimadores de Mínimos Quadrados Generalizados .....	5
$BLUP$	Melhor Preditor Linear Centrado .....	5
$\mathbf{A}^T$	Matriz Transposta de $\mathbf{A}$ .....	10
$\mathbf{A}^{-1}$	Matriz inversa de $\mathbf{A}$ .....	10
$\mathbf{P}$	Matriz ortogonal .....	10
$\mathbf{I}_m$	Matriz identidade de ordem $m$ .....	10
$ \mathbf{A} $	Determinante da matriz $\mathbf{A}$ .....	10
$tr(\mathbf{A})$	Traço da matriz $\mathbf{A}$ .....	12
$car(\mathbf{A})$	Característica da matriz $\mathbf{A}$ .....	13
$\mathbf{A}^+$	Matriz inversa de Moore-Penrose de $\mathbf{A}$ .....	13
$S^\perp$	Complemento ortogonal de $S$ .....	17
$MPO$	Matriz de projeção ortogonal .....	17
$\mathbf{O}_{(m-r) \times r}$	Matriz nula do tipo $(m-r) \times n$ .....	19
$R(\mathbf{A})$	Espaço Imagem da matriz $\mathbf{A}$ .....	20
$\mathcal{N}(\mathbf{P})$	Espaço nulidade da matriz $\mathbf{P}$ .....	21
$\boxplus$	Soma direta de espaços vetoriais .....	21
$\mathbf{A} \otimes \mathbf{B}$	Produto de Kronecker das matrizes $\mathbf{A}$ e $\mathbf{B}$ .....	21
$AJC$	Álgebra de Jordan Comutativa .....	32
$\mathcal{A}$	Álgebra de Jordan .....	79



# Capítulo 1

## Introdução

A Inferência Estatística tem uma natureza dialéctica em que intervêm:

- o modelo, contendo a informação sobre o problema em estudo anterior à recolha de observações;
- as observações.

Os modelos podem ser:

- paramétricos, quando indicam o tipo das distribuições das observações, ver, por exemplo, Lee e Thompson (2008) [38] e Madan (2014) [39]. Os modelos lineares têm sido alvo de grande interesse nos últimos anos e são muito utilizados devido à sua flexibilidade, interpretabilidade e simplicidade. Os modelos lineares, no âmbito da estatística univariada, são talvez os mais utilizados para modelar as relações entre uma variável resposta e uma ou mais variáveis explicativas. Estes modelos assumem, entre outros pressupostos, que o valor esperado da variável resposta é uma combinação linear das variáveis explicativas e que a variável resposta segue a distribuição Normal. São considerados como uma generalização dos testes de hipóteses clássicos mais simples e servem de base para numerosas extensões, algumas populares, como por exemplo os modelos lineares mistos, ver Khuri, A. I., Mathew, T. & Sinha, B. K. (1998) [34]. O nome Modelo Linear Misto provém do facto destes modelos serem lineares nos parâmetros, portanto nas variáveis independentes, e poderem englobar fatores fixos (associados a uma ou mais covariáveis) e fatores aleatórios (associados a um ou mais fatores aleatórios). Eisenhart (1947) [14] introduziu a distinção entre fatores de efeitos fixos e fatores de efeitos aleatórios. A estimação desses parâmetros pode ser feita por vários métodos, sendo um dos mais conhecidos o método da máxima verossimilhança, procedimento usual nos modelos lineares mistos que fornece ferramentas necessárias para estimar os efeitos fixos e aleatórios num mesmo modelo. Caso o interesse seja a estimação das componentes de variância, o artigo de C. R. Henderson, *Estimation of Variance and Covariance Components* (1953) [26], apresenta as primeiras ideias de como estimar componentes da variância de dados desequilibrados. Já em 1950 os seus estudos em relação aos melhores preditores lineares centrados se tinham destacado. Os seus métodos de estimação (alguns com o seu nome) e os seus trabalhos, estimularam vários autores a trabalharem na área dos modelos mistos e nas estimativas das componentes de variância, nomeadamente Hartley, H. O. & Rao, C. R., fazendo estes grandes contribuições nesta área no final dos anos 60 e 70. Em 1970, Rao apresenta um estimador de variâncias heterocedásticas em modelos lineares, e, em 1967 e 1971, foram apresentados os métodos de máxima verossimilhança (também por Hartley & Rao) e máxima verossimilhança restrita, conhecida por REML (por Patterson, H. D. & Thompson, R., [53]), método de estimação das componentes de variância que está, por defeito,

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

presente em todos os packages de *software*. Na área dos intervalos de confiança destaca-se o livro de 1985, de Khuri, A. I. & Sahai, H., [33], que fornece uma pesquisa abrangente sobre intervalos de confiança usando estimativas das componentes de variância. Alguns dos métodos referidos acima são dos mais utilizados, mas um livro que se destaca por apresentar os mais importantes processos de estimação data de 1992, de Searle, S.R., Casella, G. & McCulloch, C.E. (1992) [61], no qual, além dos métodos apresentados, se faz também referência às vantagens e desvantagens dos estimadores obtidos pelos diversos métodos. Por vezes, no contexto dos modelos lineares mistos, deparamo-nos com certos resultados algébricos mais complexos do que nos modelos lineares, mas, com o avanço computacional, esses cálculos tornam-se menos morosos. A título de exemplo, Bryk, A. S. & Raudenbush, S. W. (1992) [7] publicaram o programa HLM, para o cálculo dos modelos lineares hierárquicos, entre outros desenvolvimentos na área destes modelos e, em 1992, o *software* SAS introduziu o package de análise intitulado *Proc Mixed*. Em (2000), Pinheiro, J. C. & Bates, D. M. [54] introduziram a função *lme()* para o *software* R, este utilizado nas simulações apresentadas na tese.

- modelos onde se impõem condições muito gerais a que essas distribuições devem satisfazer, por exemplo, terem momentos definidos até à quarta ordem. Por exemplo, o livro de McCulloch, C. E. & Searle, S. R. (2001, Cap. 11) [42], apresenta modelos não lineares incorporando distribuições não normais. Segundo os autores, as estimativas obtidas em modelos lineares, sob não normalidade, não apresentam resultados tão robustos quanto os do caso normal, mas devido à sua generalidade e eficiência compensam e os procedimentos computacionais são, por vezes, de mais fácil execução, sendo, por isso, uma grande vantagem. Corroborando os autores anteriores, também o artigo de Lee, Y. & Nelder, J. A. (1996) [37], apresenta uma discussão aprofundada sobre os efeitos aleatórios não normais. Na literatura temos os modelos lineares generalizados que são uma generalização dos modelos lineares e que permitem distribuições não normais, sendo a distribuição de probabilidade da variável aleatória qualquer distribuição pertencente a uma classe designada família exponencial de distribuições, sendo que algumas generalizações admitem distribuições além da família exponencial.

Observe-se que, ao se especificar mais o modelo, em princípio se obtêm mais possibilidades para a inferência. Por exemplo, o pressuposto de normalidade permite o uso de algumas das mais elaboradas técnicas de inferência, veja-se a análise de variância (verificados alguns pressupostos adicionais, homocedasticidade, por exemplo). Há, no entanto, que recordar a possibilidade do chamado erro de terceira espécie que resulta do modelo adotado não se ajustar às observações. Por exemplo, assumir a normalidade para observações com densidades com cauda direita predominante. Assim, ao optar por um modelo, tem de se atender a dois aspetos contraditórios:

- possibilidades de inferência baseada no modelo escolhido;
- aplicabilidade do modelo.

Tendo em conta estes aspectos, os objetivos e contribuições deste trabalho são:

- Revisão da literatura sobre modelos lineares e modelos lineares mistos;
- Estimação do vetor de coeficientes fixos do modelo, dado por  $\underline{\beta}$  e dos vetores estimáveis  $\underline{\psi} = \mathbf{G}\underline{\beta}$ ;
- Construção de Elipsóides de Confiança e de Intervalos de Predição;

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

- Realização de inferência para os modelos múltiplos cujos modelos correspondem aos tratamentos dum delineamento base com efeitos fixos;
- Implementação dos métodos utilizados no *software* R;
- Aplicação dos modelos propostos através da utilização de conjuntos de dados reais, um exemplificando o caso em que as componentes dos vetores aleatórios seguem distribuições com dois parâmetros e outro caso em que seguem distribuições com três parâmetros.

Organização da tese:

No capítulo 2 apresentamos uma compilação de conceitos plenamente conhecidos, mas que servem de apoio à obtenção de resultados obtidos nos capítulos posteriores. Estes resultados irão ser referidos sempre que julgarmos oportuno e, embora sejam resultados standard, foram incluídos na secção 2.1 por forma a tornar a tese auto-suficiente, de leitura mais clara e também por forma a canalizar a mesma para leitores do nível de doutoramento, como indica o prefácio do livro seguido neste capítulo (Schott, 2017) [60]. Omitimos as demonstrações dos resultados referenciados neste capítulo dado que estas constam nas referências bibliográficas que apresentamos. Ainda neste capítulo, na secção 2.2, apresentamos a noção de função geradora de cumulantes e apresentamos o relacionamento entre momentos correspondendo a combinações lineares de variáveis aleatórias, e cumulantes, que são dados por vetores de variáveis aleatórias, seguindo de perto os desenvolvimentos apresentados por Balakrishnan, N., Johnson, N. L. & Kotz, S. (1998) [4]. Os cumulantes foram introduzidos pela primeira vez por Thorvald Thiele, em 1889, que os chamou de semi-invariantes, ver Crámer (1946, págs. 166-255) [8], e foram designados cumulantes pela primeira vez num artigo de Fisher, R.A. & John Wishart, J. (1932, págs. 195-208) [20]. Stephen Stigler Neyman (1956, págs. 288-294) [48] mencionou que o nome cumulante foi sugerido a Fisher, através de uma carta de Harold Hotelling. Num artigo de 1929, Fisher (1929, págs. 199-238) [19], chamou os cumulantes de funções cumulativas de momentos. Baseados nas expansões de Taylor, apresentamos na secção 2.3 uma útil aproximação da expansão de Edgeworth utilizando os coeficientes de assimetria e de achatamento, dados por momentos, atendendo a Kolassa, J. E. (1994, págs. 31-61) [35], onde se apresentam resultados em que os cumulantes de ordem mais alta podem ser expressos como funções polinomiais de  $n$  momentos e cumulantes de ordem inferior. Na expansão de Edgeworth apresentamos até ao quarto cumulante, seguindo as orientações de Kendall, M. G., Stuart, A. & Ord, J. K (1987, pág. 50) [32]. Segue-se a secção 2.4, sobre Partições Ortogonais de sub-espacos, cujos resultados são úteis para o estudo dos modelos múltiplos em que se testam hipóteses associadas a partições ortogonais de sub-espacos. Os  $d$  modelos associados a essas famílias corresponderão aos tratamentos dum delineamento base de efeitos fixos. Mostraremos que as hipóteses sobre efeitos e interações dos factores do delineamento base correspondem aos espacos duma partição ortogonal de  $\mathbb{R}^d$  e que as matrizes  $\mathbf{X}, \mathbf{X}_1, \dots, \mathbf{X}_w$ , são as mesmas para todos os modelos da família, bem como os cumulantes  $\chi_{r,i}, r = 2, 3, 4, i = 1, \dots, w$ . Temos então que apenas os vetores  $\underline{\beta}$  variarão de modelo individual para modelo individual, como se segue

$$\begin{aligned} \underline{Y}_1 &= \mathbf{X}\underline{\beta}_1 + \sum_{i=1}^w \mathbf{X}_i \underline{Z}_{1,i} \\ &\vdots \\ \underline{Y}_d &= \mathbf{X}\underline{\beta}_d + \sum_{i=1}^w \mathbf{X}_i \underline{Z}_{d,i}, \end{aligned}$$

com  $\underline{Z}_{1,1}, \dots, \underline{Z}_{d,i}, i = 1, \dots, w$ , *i.i.d.*.

Estudamos, então, a ação dos factores do delineamento base nas funções estimáveis  $\underline{c}^\top \underline{\beta}$ , obtendo testes para as hipóteses de ausência de efeitos e/ou interações. Assim, na secção 2.4 apresentamos os instrumentos algébricos necessários para testar essas hipóteses, permitindo realizar a inferência para modelos múltiplos, clarificando a ligação entre este estudo e o segundo trabalho já publicado, com partes comuns a este capítulo, ver Antunes et al (2020b) [3]. Note-se que os modelos múltiplos são da forma  $\underline{Y} = [\underline{Y}(1)^\top, \dots, \underline{Y}(u)^\top]^\top$ , com  $\underline{Y}(l) = \mathbf{X}\underline{\beta}(l) + \underline{Y}^o(l)$ ,  $l = 1, \dots, u$ , sendo os  $\underline{Y}^o(1), \dots, \underline{Y}^o(u)$  *i.i.d.*, com vetores médios nulos e estes generalizam os modelos múltiplos em que  $\mathcal{L}(\underline{Y}^o(l)) = \sigma^2 \mathbf{I}_n$ ,  $l = 1, \dots, u$ , ou seja, os nossos modelos múltiplos surgiram como generalização dos modelos regressoriais múltiplos, em que se considerava um único termo aleatório da forma  $\mathbf{X}\underline{e}$ , com  $\underline{e} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_n)$ , ver, respetivamente, Mexia, J. T. (1987), Moreira et al. (2005) e Ribeiro, A. & Mexia, J. T. (1997, págs. 257-271) [43], [47] e [57].

No capítulo 3 faz-se o estudo do caso geral dos modelos. Apresentamos modelos com vetores de observações  $\underline{Y}$  com  $n$  componentes, dados por

$$\underline{Y} = \mathbf{X}\underline{\beta} + \sum_{i=1}^w \mathbf{X}_i \underline{Z}_i, \quad (1.0.1)$$

não se introduzindo pressupostos adicionais sobre os mesmos, sendo dadas as matrizes  $\mathbf{X}, \mathbf{X}_1, \dots, \mathbf{X}_w$ , do tipo  $n \times k, n \times c_1, \dots, n \times c_w$  e o vetor de coeficientes de efeitos fixos é dado por  $\underline{\beta} = (\beta_1, \dots, \beta_k)$ . Consideramos os vetores aleatórios  $\underline{Z}_i = (Z_{i,1}, \dots, Z_{i,c_i})$ ,  $i = 1, \dots, w$ , independentes, com componentes *i.i.d.*, com valores médios nulos e cumulantes até à quarta ordem. Assim, nesta fase, estaremos a lidar com modelos não paramétricos, o que apresenta uma inovação traduzida no alargamento da classe de modelos estudada.

Sendo  $\sigma_i^2$ ,  $i = 1, \dots, w$ , a variância das componentes de  $\underline{Z}_i$ ,  $\beta = 1, \dots, w$ , estas terão matrizes de covariância  $\sigma_i^2 \mathbf{M}_i$ , com  $\mathbf{M}_i = \mathbf{X}_i \mathbf{X}_i^\top$ ,  $i = 1, \dots, w$ . Notemos que os segundos cumulantes das componentes dos  $\underline{Z}_i$   $i = 1, \dots, w$ , coincidem com as suas variâncias  $\sigma_i^2$ ,  $i = 1, \dots, w$ . Atendendo à independência e homocedasticidade dos vetores  $\underline{Z}_1, \dots, \underline{Z}_w$ , temos a matriz de covariância de  $\underline{Y}$  dada por

$$\mathcal{L}(\underline{Y}) = \sum_{i=1}^w \sigma_i^2 \mathbf{M}_i, \quad (1.0.2)$$

o que poderia sugerir a designação de modelos bi-aditivos. Assim, talvez fosse mais correto designar os nossos modelos com tal, dada a aditividade da equação do modelo e da matriz da covariância, pois assim conseguia-se distingui-los de outros tipos de modelos que não têm esta

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

estrutura, mas que também são designados como Modelos Aditivos (ver, por exemplo, Friedman, J. H. & Stuetzle, W. (1981, págs. 817-823) e Hoggart, S. & Hawkins, S. (2015, págs. 171-236) ([23], [27]).

Observamos que a aditividade dos modelos nos permitiu centrar o nosso tratamento, inovador, nos cumulantes. No estudo destes modelos pretendemos obter Estimadores de Mínimos Quadrados,  $LSE$ , para os cumulantes das componentes dos vetores  $\underline{Z}_i, i = 1, \dots, w$ . Sendo  $\tilde{\sigma}_i^2, i = 1, \dots, w$ , os estimadores das variâncias das componentes dos  $\underline{Z}_i, i = 1, \dots, w$ , teremos a matriz de covariância estimada

$$\tilde{\Sigma}(\underline{Y}) = \sum_{i=1}^w \tilde{\sigma}_i^2 \mathbf{M}_i, \quad (1.0.3)$$

obtendo-se, para  $\underline{\beta}$ , os Estimadores de Mínimos Quadrados Generalizados,  $GLSE$ , segundo Kariya, T. & Kurata, H. (2004, págs. 25-40) [31]:

$$\tilde{\underline{\beta}} = (\mathbf{X}^T \tilde{\Sigma}(\underline{Y})^{-1} \mathbf{X})^{-1} \mathbf{X}^T \tilde{\Sigma}(\underline{Y})^{-1} \underline{Y}, \quad (1.0.4)$$

desde que as matrizes  $\tilde{\Sigma}(\underline{Y})$  e  $\mathbf{X}^T \tilde{\Sigma}(\underline{Y})^{-1} \mathbf{X}$  sejam invertíveis (se não o forem poder-se-á utilizar a inversa generalizada). Observemos que, se forem dadas as matrizes  $\mathbf{X}_1, \dots, \mathbf{X}_w$ , pode admitir-se que  $\tilde{\Sigma}(\underline{Y})$  e  $\mathbf{X}^T \tilde{\Sigma}(\underline{Y})^{-1} \mathbf{X}$  são invertíveis para poder obter  $\tilde{\underline{\beta}}$ . Na estimação dos cumulantes de ordem  $r$  das componentes dos  $\underline{Z}_i, i = 1, \dots, w$ , designados por  $\chi_{r,i}, i = 1, \dots, w$ , utilizaremos a projecção ortogonal dos vetores das observações  $\underline{Y}$  no complemento ortogonal  $\Omega^\perp$  do espaço imagem  $\Omega = R(\mathbf{X})$ , da matriz  $\mathbf{X}$ . Um ponto importante é que, como veremos, para obter os estimadores dos quartos cumulantes,  $\tilde{\chi}_{4,i}, i = 1, \dots, w$ , teremos de dispor de um par de vetores de observações  $\underline{Y}(1)$  e  $\underline{Y}(2)$  *i.i.d.*.

Esta exigência pode ser descartada se apenas se quiser estimar os  $\underline{\beta}$  e os cumulantes até à terceira ordem. No entanto, os  $\tilde{\chi}_{r,i}, i = 1, \dots, w, r = 2, 3, 4$ , serão necessários para construir Elipsóides de Confiança e Intervalos de Predição. Os Elipsóides de Confiança, de nível  $1 - p$ , terão (aproximadamente) probabilidade  $p$  de qualquer realização de um vetor com a distribuição do vetor das observações. Por outro lado, a predição será para um conjunto de vetores adicionais  $\underline{\hat{x}}_1, \dots, \underline{\hat{x}}_w$  nas matrizes  $\mathbf{X}, \mathbf{X}_1, \dots, \mathbf{X}_w$ . Em vez de obter uma predição pontual, solução adotada nas técnicas usuais como melhor preditor linear centrado,  $BLUP$ , podemos reutilizar e complementar a técnica usada geralmente na literatura para construir Elipsóides de Confiança por forma a obter os limites de Predição  $Y_{\frac{p}{2}}$  e  $Y_{1-\frac{p}{2}}$ , tendo assim o Intervalo de Predição  $[Y_{\frac{p}{2}}; Y_{1-\frac{p}{2}}]$ , com nível (aproximado)  $1 - p$ . Observe-se que as componentes dos vetores  $\underline{Z}_1, \dots, \underline{Z}_w$ , podem ter distribuições distintas o que alarga o âmbito destes modelos em que é usual ter um único termo aleatório com distribuição normal. Esta possibilidade é explorada nos capítulos 4 e 5, sendo estes capítulos dedicados ao estudo de casos particulares dos modelos abordados.

No capítulo 4 consideramos o caso em que as distribuições das componentes dos vetores  $\underline{Z}_i, i = 1, \dots, w$ , têm parâmetros de localização  $\lambda_i, i = 1, \dots, w$  e dispersão  $\delta_i, i = 1, \dots, w$ , nomeadamente, os casos em que essas distribuições são Normais, Exponenciais e Gumbel, ver Balakrishnan, N. & Basu, A. P. (2019, págs. 7, 34-35, 55-58 e 74) e Coorey, K. (2010, págs. 171-179) [5] [11].

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

No capítulo 5 consideramos o caso em que as distribuições das componentes dos vetores  $Z_i$ ,  $i = 1, \dots, w$ , têm parâmetros de localização  $\lambda_i$ ,  $i = 1, \dots, w$ , de dispersão  $\delta_i$ ,  $i = 1, \dots, w$ , e forma  $\tau_i$ ,  $i = 1, \dots, w$ . Aquando do ajustamento, integramos os parâmetros de localização  $\lambda_i$ ,  $i = 1, \dots, w$ , no vetor dos coeficientes  $\underline{\beta}$ . Assim, quando estimamos  $\underline{\beta}$ , estimamos  $\lambda_1, \dots, \lambda_w$ , usando técnicas algébricas plenamente conhecidas mas agora aplicadas num contexto diferente. Por outro lado, a partir dos  $(\tilde{\chi}_{2,i}, \tilde{\chi}_{3,i})$ ,  $i = 1, \dots, w$ , obtemos os  $(\tilde{\delta}_i, \tilde{\tau}_i)$ ,  $i = 1, \dots, w$ , e, destes, os  $\tilde{\chi}_{4,i}$ ,  $i = 1, \dots, w$ . Obtemos Elipsóides de Confiança e Intervalos de Predição considerando ainda Modelos Múltiplos. Particularizamos para os casos em que as componentes  $Z_{i,l}$ ,  $l = 1, \dots, c_i$ ,  $i = 1, \dots, w$ , têm distribuições Gama, Weibull e Fréchet, ver Gumbel (1965, págs. 349-363), Lakshmi, R. V. & Vaidyabathan, V. S. (2016) e Rinne, H. (2008 págs. 27-97) [25] [36] [58].

Por fim, no capítulo 6, apresentamos dois estudos de simulação, sendo que o segundo está dividido em dois casos. Todos os cálculos são efetuados no *software* R.

Consideraremos modelos com dois factores aleatórios, que cruzam, em ambos os estudos de simulação. O primeiro fator terá dois níveis e o segundo terá três níveis.

No primeiro estudo de simulação são simulados dois vetores de observações, normalmente distribuídos, com média zero e variância conhecida. São considerados 6 casos e, para cada caso, assume-se um valor para a variância. Parte deste primeiro estudo de simulação foi publicado em Antunes et al. (2020a) [2]. Constatamos que as estimativas médias obtidas estão bastante próximas dos valores considerados.

No segundo estudo simulação a estrutura dos modelos será a mesma que a considerada no primeiro estudo. No entanto, as observações são obtidas a partir da simulação dos vetores de efeitos aleatórios, sendo este procedimento repetido 1000 vezes. Ou seja, o vetor das observações será a soma do vetor de efeitos fixos com os dois vetores de efeitos aleatórios. Cada vetor aleatório será obtido por simulação, tendo em conta que será normalmente distribuído, com vetor médio nulo e uma determinada variância (componente de variância). Para essas variâncias serão usados 6 valores. Assim, tendo em conta todas as combinações, teremos 36 casos diferentes. Apresentamos as estimativas para as componentes de variância, e para o vetor de efeitos fixos, em que consideramos a distribuição normal por forma a ilustrar um modelo com parâmetros de localização e dispersão, respetivamente. Também neste estudo constatamos que as estimativas estão muito perto dos valores previamente considerados. Além disso, exemplificamos a construção de elipsóides de confiança para observações futuras e intervalos de predição para uma futura observação. A fim de aferir a qualidade dos intervalos obtidos, definimos intervalos, com cobertura de 95%, a partir dos quantis estimados, e verificamos que as percentagens de observações incluídas nesses intervalos, de lotes de 1000 novas observações, estão muito perto do valor teórico, logo, o nível de confiança efetivo dos elipsóides de confiança e dos intervalos de predição são exatamente iguais ao nível de confiança fixado à partida. A fim de aferir a qualidade dos intervalos obtidos, apresentamos uma tabela com a contagem do número de vezes, em 1000, em que a simulação de uma nova observação pertence ao intervalo apresentado.

Como parte dos resultados desenvolvidos no âmbito desta dissertação já se encontram publicados, no capítulo 6 apresentamos também um exemplo no contexto dos modelos múltiplos, por

## **Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição**

forma a ilustrar a aplicabilidade destes modelos, sendo que este exemplo integra o artigo Antunes et al. (2000b) [3].

Foi ainda incluído um apêndice sobre álgebras de Jordan Comutativas, dado que essas álgebras são úteis no estudo das partições ortogonais de sub-espacos, como as que consideramos para os modelos múltiplos. Só apresentamos resultados fundamentais para o necessário acompanhamento das definições presentes no capítulo.

Na tese foram colocadas as referências com os nomes dos autores e, seguidamente, com o número segundo o qual a referência aparece na bibliografia. Esta nossa preferência pretende uma identificação mais rápida dos autores dos trabalhos e livros utilizados na tese.

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

# Capítulo 2

## Resultados preliminares

Neste capítulo apresentamos uma compilação de alguns conceitos e resultados importantes na área da estatística. Primeiro, apresentamos alguns resultados algébricos sobre matrizes inversas generalizadas e inversas de Moore-Penrose, matrizes de covariância e de covariância cruzada e matrizes de projeção ortogonal. Posteriormente, apresentamos um breve resumo de resultados conhecidos sobre o produto de Kronecker de matrizes.

Na formulação de modelos lineares é normal usar-se a álgebra matricial, quando se apresentam ou se verificam resultados, que, desta forma, nos permite manipular a complexidade das várias variáveis observadas.

Estes conceitos servirão de apoio à obtenção de resultados dos capítulos seguintes, podendo grande parte deles ser encontrados, por exemplo, em Schott, J. R. (2017) [60]. Com a inclusão destes resultados preliminares pretende-se tornar a tese auto suficiente, para um nível científico com alguma maturidade, uma vez que este livro é indicado, segundo o prefácio do mesmo, para alunos de doutoramento. As demonstrações não incluídas nesta secção poderão ser vistas no mesmo. Uma ou outra referência constarão do texto por forma a clarificar a origem das expressões apresentadas.

Apresentamos de seguida algumas definições e proposições necessárias para uma melhor compreensão de certos conceitos dos capítulos seguintes. Outras se poderiam apresentar mas estas são as mais utilizadas e importantes para o melhor entendimento do conteúdo dos mesmos.

### 2.1 Matrizes

#### 2.1.1 Matrizes Transpostas

**Definição 2.1 (*Matriz transposta*)** Seja  $A$  uma matriz do tipo  $m \times n$ . Então a transposta de  $A$ ,  $A^T$ , define-se como sendo uma matriz do tipo  $n \times m$  cuja entrada  $(i, j)$  é igual à entrada  $(j, i)$  de  $A$ .

Assim, a coluna [linha]  $i$  da  $A^T$  é precisamente a linha [coluna]  $i$  de  $A$ , para  $i = 1, \dots, m$ .

**Proposição 2.1** Seja  $\alpha$  um escalar,  $k$  um número natural e  $A$  e  $B$  matrizes do tipo  $m \times n$  e  $n \times p$ , respetivamente.

A transposição de matrizes goza das seguintes propriedades:

- $(A^T)^T = A$ ;

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

- $(A + B)^T = A^T + B^T$ ;
- $(\alpha A)^T = \alpha A^T$ ;
- $(AB)^T = B^T A^T$ ;
- $(A^k)^T = (A^T)^k$ ;
- Se  $A$  for uma matriz invertível,  $A^T$  também o é, tendo-se  $(A^T)^{-1} = (A^{-1})^T$ .

### 2.1.2 Matrizes Ortogonais

**Definição 2.2 (Matriz Ortogonal)** Uma matriz  $P$ , de ordem  $m$ , é ortogonal se e só se

$$PP^T = I_m, \quad (2.1.1)$$

onde  $I_m$  representa a matriz identidade de ordem  $m$ .

Os vetores  $\underline{p}_1, \dots, \underline{p}_n$ , com  $n \leq m$ , do tipo  $m \times 1$ , são ortogonais dois a dois se  $\underline{p}_i^T \underline{p}_j = 0$ , para todos os  $i \neq j$ . Além disso, uma matriz  $P$ , do tipo  $m \times m$ , é ortogonal se a sua transposta é igual à sua inversa

$$P^T = P^{-1}. \quad (2.1.2)$$

Para mais propriedades sobre matrizes ortogonais ver pág. 15, Teorema 1.12, Schott, J. R. (2017) [60].

Note-se que, a toda a matriz quadrada, de números reais, pode ser associada um número real, que se designa por determinante da matriz. Apresentamos de seguida a sua definição.

**Definição 2.3 (Determinante)** Seja  $A$  uma matriz quadrada de ordem  $m$ . O determinante de  $A$ , designado por  $\det(A)$  ou  $|A|$ , é o número dado por

$$|A| = \sum_{\sigma \in S_m} \text{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \dots a_{n\sigma(m)}, \quad (2.1.3)$$

onde a soma é efectuada considerando todas as permutações  $\sigma$  do conjunto  $\{1, 2, 3, \dots, m\}$ . A permutação é uma função que reordena este conjunto de inteiros e o valor da  $i$ -ésima posição, depois da reordenação  $\sigma$ , é denotado por  $\sigma_i$ . O conjunto de todas as permutações é denotado por  $S_m$ . Para cada permutação  $\sigma$ ,  $\text{sgn}(\sigma)$  denota o sinal de  $\sigma$ , que fornece os valores  $+1$  e  $-1$ , para permutações pares e ímpares, respetivamente.  $\text{sgn}(\sigma)$  é no fundo  $(-1)$  com o expoente que representa o número de vezes que se tem de permutar elementos contíguos para permitir construir a permutação base  $\{1, 2, 3, \dots, m\}$ . Os produtos na expressão acima são obtidos considerando os  $m$  elementos da matriz  $A$ .

**Teorema 2.1** Sejam  $P$  e  $Q$  matrizes ortogonais de ordem  $m$  e  $A$  uma matriz qualquer, também de ordem  $m$ . Então

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

- $|P| = \pm 1$ ;
- $|P^T A P| = |A|$ ;
- $PQ$  é uma matriz ortogonal.

### 2.1.3 Matrizes Simétricas

**Definição 2.4 (Matriz Simétrica)** Uma matriz simétrica é uma matriz quadrada igual à sua transposta, ou seja  $A = A^T$ .

Necessitaremos da definição de forma quadrática que apresentaremos de seguida. Esta definição obtém-se seguindo a definição de forma quadrática da secção 1.11, Schott, J. R. (2017, pág. 16) [60]. Seja  $\underline{x}$  um vetor do tipo  $m \times 1$ . A função  $f(\underline{x}) = \underline{x}^T A \underline{x} = \sum_{i=1}^m \sum_{j=1}^m x_i x_j a_{ij}$ , é chamada de forma quadrática em  $\underline{x}$ . Cada matriz simétrica  $A$  e a respetiva forma quadrática associada são classificadas numa das categorias apresentadas seguidamente.

**Definição 2.5** Seja  $A$  uma matriz quadrada de ordem  $m$ , simétrica, e seja  $\underline{x}$  um vetor do tipo  $m \times 1$ . Essa matriz diz-se

- definida positiva, se se verificar  $\underline{x}^T A \underline{x} > 0, \forall \underline{x} \neq 0$ ;
- definida negativa, se se verificar  $\underline{x}^T A \underline{x} < 0, \forall \underline{x} \neq 0$ ;
- semidefinida positiva, se  $\underline{x}^T A \underline{x} \geq 0, \forall \underline{x}$  and  $\underline{x}^T A \underline{x} = 0$  para algum  $\underline{x} \neq 0$ ;
- semidefinida negativa, se  $\underline{x}^T A \underline{x} \leq 0, \forall \underline{x}$  and  $\underline{x}^T A \underline{x} = 0$  para algum  $\underline{x} \neq 0$  e
- indefinida se  $\underline{x}^T A \underline{x} > 0$ , para algum  $\underline{x}$  e  $\underline{x}^T A \underline{x} < 0$ , para algum  $\underline{x}$ .

Os resultados seguintes seguem os apresentados nas páginas 96-107 de Schott, J. R. (2017) [60].

**Definição 2.6** Seja  $A$  uma matriz quadrada de ordem  $m$ . Um escalar  $\lambda \in \mathbb{R}$  diz-se um valor próprio de  $A$  se e só se existir um vetor não nulo, do tipo  $m \times 1$ , tal que

$$A \underline{x} = \lambda \underline{x}. \quad (2.1.4)$$

Diz-se que  $\underline{x}$  é um vetor próprio, associado ao valor próprio  $\lambda$ , se a equação  $A \underline{x} = \lambda \underline{x}$  se verificar. A definição 2.6 também pode ser expressa, de forma equivalente, por

**Definição 2.7** Seja  $A$  uma matriz quadrada de ordem  $m$ . Então,  $\lambda$  é um valor próprio de  $A$  se e só se  $|A - \lambda I_m| = 0$ .

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

**Corolário 2.1.1** *Seja  $A$  uma matriz quadrada de ordem  $m$ . Então, tem-se  $|A| = 0 \iff \lambda = 0$ , sendo  $\lambda$  um valor próprio de  $A$ .*

**Corolário 2.1.2** *Seja  $A$  uma matriz quadrada de ordem  $m$ . Então,  $A$  é uma matriz invertível se e só se  $A$  não tiver valores próprios nulos.*

**Teorema 2.2** *Seja  $A$  uma matriz simétrica de ordem  $m$ .*

- Os valores próprios  $\lambda_1, \dots, \lambda_m$ , da matriz  $A$  são números reais;
- Os vetores próprios,  $\underline{x}_1, \dots, \underline{x}_k$ , correspondentes a valores próprios distintos,  $\lambda_1, \dots, \lambda_k$ , com  $k \leq m$ , são mutuamente ortogonais.

**Teorema 2.3** *Seja  $A$  uma matriz simétrica com valores próprios  $\lambda_1, \dots, \lambda_m$ , e vetores próprios ortonormados (ortogonais entre si e com norma 1),  $\underline{x}_1, \dots, \underline{x}_m$ , correspondentes a esses valores próprios, então a matriz  $A$  pode ser representada como*

$$A = PD(\lambda_1, \dots, \lambda_m)P^T = \sum_{i=1}^m \lambda_i \underline{x}_i \underline{x}_i^T, \quad (2.1.5)$$

onde  $P$  é uma matriz cujas colunas são os vetores próprios normalizados  $\underline{x}_1, \dots, \underline{x}_m$  e  $D(\lambda_1, \dots, \lambda_m)$  é uma matriz diagonal cujos elementos principais são os valores próprios  $\lambda_1, \dots, \lambda_m$ , da matriz  $A$ . Ao somatório  $\sum_{i=1}^m \lambda_i \underline{x}_i \underline{x}_i^T$ , chama-se a decomposição espectral da matriz  $A$  (Schott, J. R. (2017, pág. 108))[60].

**Corolário 2.1.3** *Se  $A$  é uma matriz simétrica, existe uma matriz diagonalizadora ortogonal da matriz  $A$ , ou seja,*

$$P^T AP = D(\lambda_1, \dots, \lambda_m). \quad (2.1.6)$$

**Proposição 2.2** *Se  $A$  é uma matriz  $m \times m$  com valores próprios  $\lambda_1, \dots, \lambda_m$ , então*

- $|A| = \prod_{i=1}^m \lambda_i$ ;
- $tr(A) = \sum_{i=1}^m \lambda_i$ ,

onde  $tr(A)$  representa o traço da matriz  $A$ . Isto é, o traço de uma matriz é a soma dos elementos da diagonal principal (Schott, J. R. (2017, pág. 101)) [60].

### 2.1.4 Matrizes Inversas

Apresentamos alguns resultados sobre matrizes inversas, seguindo novamente Schott, J. R. (2017, pág 9) [60].

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

**Definição 2.8 (Matriz Inversa)** Uma matriz quadrada com característica completa é dita ser não singular. Uma matriz não singular  $A$ , do tipo  $m \times m$ , tem uma inversa única, representada por  $A^{-1}$ , com

$$AA^{-1} = A^{-1}A = I_m. \quad (2.1.7)$$

**Proposição 2.3** Uma matriz não singular e idempotente é uma matriz identidade  $I_m$ .

**Definição 2.9 (Matriz Invertível)** Se uma matriz  $A$  quadrada de ordem  $m$  tiver inversa, diz-se invertível, (ou regular), se existir uma matriz  $B$ , quadrada, de ordem  $m$ , tal que

$$AB = BA = I_m. \quad (2.1.8)$$

Apesar de só as matrizes quadradas poderem ser regulares, no entanto, nem todas as matrizes quadradas têm inversa. Apresentamos outra condição para que uma matriz seja regular.

**Proposição 2.4** Uma matriz  $A$ , de ordem  $m$ , é regular se e só se é uma matriz quadrada com característica igual à ordem, ou seja,  $A_{m \times m}$  é regular equivale a que  $\text{car}(A) = m$ .

### 2.1.5 Matrizes Inversas Generalizadas

Se uma matriz  $A$  é singular,  $A^{-1}$  não existe, mas existe uma matriz inversa generalizada (Schott, J. R. (2017, pág. 201)) [60].

**Definição 2.10 (Matriz Inversa Generalizada)** Uma matriz inversa generalizada de uma matriz  $A$ , do tipo  $m \times n$ , é uma matriz  $A^{-}$ , do tipo  $n \times m$ , tal que

$$AA^{-}A = A. \quad (2.1.9)$$

Note-se que toda a matriz tem uma inversa generalizada.

### 2.1.6 Inversas de Moore-Penrose

A matriz inversa de Moore-Penrose de uma matriz  $A$ ,  $m \times n$ , é a matriz  $n \times m$ , denotada por  $A^{+}$ , Schott, J. R. (2017, pág 202) [60], que satisfaz as seguintes condições

$$\left\{ \begin{array}{l} AA^{+}A = A \\ A^{+}AA^{+} = A^{+} \\ (AA^{+})^{\top} = AA^{+} \\ (A^{+}A)^{\top} = A^{+}A \end{array} \right. \quad (2.1.10)$$

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

Caso  $A$  seja invertível, tem-se

$$\mathbf{A}^+ = \mathbf{A}^{-1}. \quad (2.1.11)$$

Sendo  $b^+ = b^{-1}$  [ $= 0$ ], quando  $b \neq 0$  [ $= 0$ ], verifica-se que, sendo  $\underline{U}$  um vetor aleatório,

$$\Sigma(\underline{U})^+ = \mathbf{P}^\top \mathbf{D}(r_1^+, \dots, r_n^+) \mathbf{P}. \quad (2.1.12)$$

Seja agora  $A$  uma matriz  $m \times n$  com característica  $s$ , tem-se então, de acordo com Belsey, D. A., Kuh, E. & Welsch, R. E. (1980, pág. 58) [6],

$$\mathbf{A} = \mathbf{P}^\top \mathbf{D}(r_1, \dots, r_s, 0, \dots, 0) \mathbf{P}', \quad (2.1.13)$$

com  $\mathbf{P}$  e  $\mathbf{P}'$  diagonalizadoras ortogonais de  $\mathbf{A}\mathbf{A}^\top$  e de  $\mathbf{A}^\top \mathbf{A}$ , respetivamente, e  $\mathbf{D}(r_1, \dots, r_s, 0, \dots, 0)$  uma matriz  $m \times n$ , cujos únicos elementos não nulos  $r_1, \dots, r_s$ , têm índices de linha e coluna iguais. Estabelecemos a

**Proposição 2.5** *Tem-se  $\mathbf{A}^+ = \mathbf{P}'^\top \mathbf{D}(r_1^{-1}, \dots, r_s^{-1}, 0, \dots, 0) \mathbf{P}$ .*

Ponhamos

$$\begin{cases} \mathbf{P} = [\underline{p}_1, \dots, \underline{p}_s, \dots, \underline{p}_m]^\top \\ \mathbf{P}' = [\underline{p}'_1, \dots, \underline{p}'_s, \dots, \underline{p}'_m] \end{cases} \quad (2.1.14)$$

vindo

$$\mathbf{A}\mathbf{A}^+ = \mathbf{P}^\top \mathbf{D}(1, \dots, 1, 0, \dots, 0) \mathbf{P} = \sum_{j=1}^s \underline{p}_j \underline{p}_j^\top \quad (2.1.15)$$

e

$$\mathbf{A}^+ \mathbf{A} = \mathbf{P}'^\top \mathbf{D}(1, \dots, 1, 0, \dots, 0) \mathbf{P}' = \sum_{j=1}^s \underline{p}'_j \underline{p}'_j^\top, \quad (2.1.16)$$

pelo que  $\mathbf{A}\mathbf{A}^+$  e  $\mathbf{A}^+ \mathbf{A}$  são matrizes de projeção ortogonal com característica  $s$ . Devido à primeira e segunda das condições, (ver 2.1.10), para  $\mathbf{A}^+$  ser inversa de Moore-Penrose de  $\mathbf{A}$  em  $R(\mathbf{W})$ , o espaço imagem de  $\mathbf{W}$ , tem-se

$$\begin{cases} R(\mathbf{A}) = R(\mathbf{A}\mathbf{A}^+) \\ R(\mathbf{A}^+) = R(\mathbf{A}^+ \mathbf{A}) \end{cases}, \quad (2.1.17)$$

pelo que

$$\text{car}(\mathbf{A}) = \text{car}(\mathbf{A}\mathbf{A}^+) = s = \text{car}(\mathbf{A}^+ \mathbf{A}) = \text{car}(\mathbf{A}^+). \quad (2.1.18)$$

Vê-se ainda que

$$\mathbf{A}^\top = \mathbf{P}'^\top \mathbf{D}(r_1, \dots, r_s, 0, \dots, 0) \mathbf{P}, \quad (2.1.19)$$

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

e, como

$$\text{car}(\mathbf{A}^\top) = \text{car}(\mathbf{A}) = s, \quad (2.1.20)$$

tem-se

$$\text{car}(\mathbf{A}^\top) = \text{car}(\mathbf{A}^+). \quad (2.1.21)$$

### 2.1.7 Matrizes de covariância e de covariância cruzada

O vetor aleatório  $\underline{U} = (U_1, \dots, U_w)$ , com vetor médio  $\underline{\mu}(\underline{U}) = (\mu(U_1), \dots, \mu(U_w))$ , tem matriz de covariância, do tipo  $n \times n$ , dada por

$$\underline{\Sigma}(\underline{U}) = E[(U_i - \mu(U_i))(U_j - \mu(U_j))], \quad i, j = 1, \dots, w, \quad (2.1.22)$$

se os valores médios que dão os seus elementos estiverem definidos. Vê-se que, caso  $\underline{\Sigma}(\underline{U})$  esteja definida, se tem

$$\underline{\Sigma}(\underline{U})^\top = \underline{\Sigma}(\underline{U}), \quad (2.1.23)$$

pelo que, se a matriz de covariância  $\underline{\Sigma}(\underline{U})$  estiver definida, é simétrica. Existe pois, como refere Schott, J. R. (2017, pág. 109) [60], uma matriz ortogonal

$$\mathbf{P} = [\underline{p}_1, \dots, \underline{p}_n]^\top, \quad (2.1.24)$$

tal que

$$\mathbf{P} \underline{\Sigma}(\underline{U}) \mathbf{P}^\top = \mathbf{D}(r_1, \dots, r_n), \quad (2.1.25)$$

com  $\mathbf{D}(r_1, \dots, r_n)$  a matriz diagonal com elementos principais  $r_1, \dots, r_n$ , (tal como no Teorema 2.3), sendo os  $(r_i, \underline{p}_i)$ ,  $i = 1, \dots, n$ , os pares de valores e vetores próprios de  $\underline{\Sigma}(\underline{U})$ , respetivamente. Como

$$r_i = \underline{\Sigma}(\underline{p}_i^\top \underline{U}) \geq 0, \quad i = 1, \dots, n, \quad (2.1.26)$$

$\underline{\Sigma}(\underline{U})$  será semi-definida ou definida positiva, sendo definida positiva se for invertível. Vê-se ainda que

$$\underline{\Sigma}(\underline{U}) = \mathbf{P}^\top \mathbf{D}(r_1, \dots, r_n) \mathbf{P}, \quad (2.1.27)$$

e, considerando  $\text{car}(\underline{\Sigma}(\underline{U}))$  a característica de  $\underline{\Sigma}(\underline{U})$ , correspondendo ao número dos seus valores próprios positivos, tem-se ainda que, se  $\underline{\Sigma}(\underline{U})$  for invertível (tendo pois  $n$  valores próprios positivos) se tem

$$\underline{\Sigma}(\underline{U})^{-1} = \mathbf{P}^\top \mathbf{D}(r_1^{-1}, \dots, r_n^{-1}) \mathbf{P}. \quad (2.1.28)$$

Por outro lado, com

$$\underline{\Sigma}(\underline{U})^{\frac{1}{2}} = \mathbf{P}^\top \mathbf{D}(r_1^{\frac{1}{2}}, \dots, r_n^{\frac{1}{2}}) \mathbf{P}, \quad (2.1.29)$$

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

tem-se

$$\Sigma(\underline{U}) = \Sigma(\underline{U})^{\frac{1}{2}} \Sigma(\underline{U})^{\frac{1}{2}}. \quad (2.1.30)$$

Se  $\Sigma(\underline{U})$  for invertível, tem-se  $r_i > 0$  e  $r_i^{\frac{1}{2}} > 0, i = 1, \dots, n$ , sendo pois  $\Sigma(\underline{U})^{\frac{1}{2}}$  invertível. Se  $C$  for uma matriz definida positiva com valores próprios  $r_i > 0, i = 1, \dots, n$ , existe uma matriz ortogonal  $P$ , tal que  $PCP^T = D(r_1, \dots, r_n)$ , sendo  $C$  a matriz de covariância de  $\underline{U}$ , vindo

$$\begin{aligned} D(r_1^{-\frac{1}{2}}, \dots, r_n^{-\frac{1}{2}})P \Sigma(\underline{U})P^T D(r_1^{-\frac{1}{2}}, \dots, r_n^{-\frac{1}{2}}) &= D(r_1^{-\frac{1}{2}}, \dots, r_n^{-\frac{1}{2}})PP^T D(r_1, \dots, r_n)PP^T D(r_1^{-\frac{1}{2}}, \dots, r_n^{-\frac{1}{2}}) \\ &= D(r_1^{-\frac{1}{2}}, \dots, r_n^{-\frac{1}{2}})D(r_1, \dots, r_n)D(r_1^{-\frac{1}{2}}, \dots, r_n^{-\frac{1}{2}}) \\ &= I_n. \end{aligned}$$

Considerando

$$K = D(r_1^{-\frac{1}{2}}, \dots, r_n^{-\frac{1}{2}})P, \quad (2.1.31)$$

tem-se

$$KCK^T = I_n, \quad (2.1.32)$$

bem como,

$$\Sigma(K\underline{U}) = \sigma^2 I_n, \quad (2.1.33)$$

quando

$$\Sigma(\underline{U}) = \sigma^2 C. \quad (2.1.34)$$

Esta situação de heterocedasticidade controlada representa uma generalização da homocedasticidade, ver Mexia, J. T. (1989) [45]. Como vimos, existe então uma matriz  $K$ , conhecida, tal que  $\Sigma(K\underline{U}) = \sigma^2 I_n$  (em geral,  $K$  não é simétrica). Por outro lado, como refere Silvey, S. D. (1980, pág. 12) [65],

$$\text{car}(BB^T) = \text{car}(B) = \text{car}(B^T B), \quad (2.1.35)$$

e, como

$$\Sigma(\underline{AU}) = A \Sigma(\underline{U})A^T = A \Sigma(\underline{U})^{\frac{1}{2}} \Sigma(\underline{U})^{\frac{1}{2}} A^T = (A \Sigma(\underline{U})^{\frac{1}{2}})(A \Sigma(\underline{U})^{\frac{1}{2}})^T, \quad (2.1.36)$$

tem-se

$$\text{car}(\Sigma(\underline{AU})) = \text{car}(A \Sigma(\underline{U})A^T) \quad (2.1.37)$$

e

$$\text{car}(\Sigma(\underline{AU})) = \text{car}(A), \quad (2.1.38)$$

com  $\Sigma(\underline{U})$  invertível, Schott, J. R. (2017, pág. 14) [60].

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

Passando às matrizes de covariância cruzada, o par  $(\underline{U}, \underline{Z})$  de vetores aleatórios, com  $n$  e  $n'$  componentes tem matriz de covariância cruzada

$$\Sigma(\underline{U}, \underline{Z}) = [E((U_i - \mu(U_i))(Z_j - \mu(Z_j)))], \quad (2.1.39)$$

se os valores médios que nela figuram estiverem definidos. Tem-se ainda, ver novamente Schott, J. R. (2017, págs. 23-24) [60],

$$\Sigma(\underline{AU}, \underline{BZ}) = \mathbf{A} \Sigma(\underline{U}, \underline{Z}) \mathbf{B}^\top. \quad (2.1.40)$$

### 2.1.8 Matrizes de projeção ortogonal

Suponhamos que  $S$  é um subespaço linear do espaço vetorial  $E$ , com produto interno definido e que  $\{\underline{z}_1, \dots, \underline{z}_m\}$  é uma base ortonormada de  $E$ , e  $\{\underline{z}_1, \dots, \underline{z}_r\}$  uma base ortonormada de  $S$ , com  $r < m$  e consideremos  $\{a_1, \dots, a_r, a_{r+1}, \dots, a_m\}$  constantes. Dado  $\underline{x} \in E$ , tem-se

$$\underline{x} = \sum_{i=1}^r a_i \underline{z}_i + \sum_{i=r+1}^m a_i \underline{z}_i. \quad (2.1.41)$$

**Definição 2.11 (Complemento ortogonal)** O complemento ortogonal de  $S$ , representado por  $S^\perp$ , é o conjunto de todos os vetores de  $E$  que são ortogonais a cada vetor de  $S$ , em que  $S$  é um subespaço linear de  $E$ , ou seja

$$S^\perp = \{\underline{x} \in E : \underline{x}^\top \underline{y} = 0, \text{ para todo o } \underline{y} \in S\}. \quad (2.1.42)$$

**Teorema 2.4** Se  $S$  é um subespaço linear de  $E$ . Então o seu complemento ortogonal,  $S^\perp$ , também é um subespaço linear de  $E$ . Se  $E$  for um espaço vetorial de dimensão  $m$  e  $S$  um subespaço linear de  $E$ , de dimensão  $r \leq m$ , então o complemento ortogonal,  $S^\perp$ , é um subespaço linear de  $E$ , de dimensão  $m - r$ .

**Proposição 2.6** A condição necessária e suficiente para uma matriz quadrada  $\mathbf{P}$ , de ordem  $m$ , ser uma matriz de projeção ortogonal, MPO, é verificar as duas condições seguintes

- $\mathbf{P}^2 = \mathbf{P}$ , ou seja,  $\mathbf{P}$  é idempotente;
- $\mathbf{P}^\top = \mathbf{P}$ , ou seja,  $\mathbf{P}$  é simétrica.

Note-se que uma MPO é única.

**Definição 2.12** Uma base  $B = \{\underline{z}_1, \dots, \underline{z}_m\}$ , do espaço vetorial  $S$ , é ortogonal se o conjunto dos seus vetores for ortogonal.

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

**Definição 2.13** Uma base  $B = \{\underline{z}_1, \dots, \underline{z}_m\}$ , de um espaço vetorial  $E$ , diz-se ortonormada de  $E$ , se  $B$  é uma base ortogonal e todos os seus vetores são unitários, ou seja, tem-se  $\|\underline{z}_i\| = 1$ ,  $i = 1, \dots, m$ , ou seja

$$\underline{z}_i^\top \underline{z}_j = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}. \quad (2.1.43)$$

**Teorema 2.5** Suponhamos que  $\{\underline{z}_1, \dots, \underline{z}_r, \underline{z}_{r+1}, \dots, \underline{z}_m\}$  é uma base ortonormada de  $E$  e  $\{\underline{z}_1, \dots, \underline{z}_r\}$  é uma base ortonormada para o subespaço linear  $S$ , de  $E$ . Temos que  $\{\underline{z}_{r+1}, \dots, \underline{z}_m\}$  é uma base ortonormada para  $S^\perp$ .

A projeção ortogonal de um vetor  $\underline{x}$ , do tipo  $m \times 1$ , num subespaço linear  $S$ , de  $E$ , pode ser representada na forma matricial.

Consideremos,  $\{\underline{z}_1, \dots, \underline{z}_r\}$ , uma base ortonormada para  $S$  e  $\{\underline{z}_1, \dots, \underline{z}_r, \underline{z}_{r+1}, \dots, \underline{z}_m\}$  uma base ortonormada para  $E$ . Consideremos  $a_1, \dots, a_r, a_{r+1}, \dots, a_m$  constantes que satisfazem a seguinte relação

$$\begin{aligned} \underline{x} &= a_1 \underline{z}_1 + \dots + a_r \underline{z}_r + a_{r+1} \underline{z}_{r+1} + \dots + a_m \underline{z}_m & (2.1.44) \\ &= (a_1 \underline{z}_1 + \dots + a_r \underline{z}_r) + (a_{r+1} \underline{z}_{r+1} + \dots + a_m \underline{z}_m) \\ &= \sum_{i=1}^r a_i \underline{z}_i + \sum_{i=r+1}^m a_i \underline{z}_i \\ &= \underline{u} + \underline{v}, \end{aligned}$$

onde  $\underline{x}$  será uma combinação linear dos elementos da base do subespaço linear  $S$  e dos elementos da base do subespaço linear  $S^\perp$ . Assim,  $\underline{x}$  representa-se como um elemento de  $E$ . Desta forma temos

$$\underline{\alpha} = [\alpha_1, \dots, \alpha_r, \alpha_{r+1}, \dots, \alpha_m]^\top = [\underline{\alpha}_1^\top \ \underline{\alpha}_2^\top]^\top,$$

onde  $\underline{\alpha}_1 = [\alpha_1, \dots, \alpha_r]^\top$ ,  $\underline{\alpha}_2 = [\alpha_{r+1}, \dots, \alpha_m]^\top$ , em que  $\underline{\alpha}_1$  representa o vetor coluna com  $r$  coeficientes, sendo  $r$  a dimensão de  $S$  e  $\underline{\alpha}_2$  o vetor coluna com  $m - r$  componentes, coincidente com a dimensão de  $S^\perp$ . Tem-se ainda

$$\mathbf{Z} = [\mathbf{Z}_1 \ \mathbf{Z}_2],$$

com

$$\mathbf{Z}_1 = [\underline{z}_1, \dots, \underline{z}_r], \quad \mathbf{Z}_2 = [\underline{z}_{r+1}, \dots, \underline{z}_m], \quad (2.1.45)$$

em que  $\mathbf{Z}_1$  representa a matriz que tem por colunas os elementos da base de  $S$  e  $\mathbf{Z}_2$  a matriz que tem por colunas os elementos da base de  $S^\perp$ . Obtém-se então

$$\begin{aligned} \underline{x} &= \mathbf{Z} \underline{\alpha} & (2.1.46) \\ &= [\mathbf{Z}_1 \ \mathbf{Z}_2] [\underline{\alpha}_1^\top \ \underline{\alpha}_2^\top]^\top \\ &= \mathbf{Z}_1 \underline{\alpha}_1 + \mathbf{Z}_2 \underline{\alpha}_2 \end{aligned}$$

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

$$\begin{aligned}
 &= \sum_{i=1}^r \alpha_i \mathbf{Z}_i + \sum_{i=r+1}^m \alpha_i \mathbf{Z}_i \\
 &= \underline{u} + \underline{v},
 \end{aligned}$$

onde se considera que  $\underline{u} = \mathbf{Z}_1 \underline{\alpha}_1$  e  $\underline{v} = \mathbf{Z}_2 \underline{\alpha}_2$ .

Devido à ortonormalidade dos vetores  $\underline{z}_i$ ,  $i = 1, \dots, m$ , obtemos

- $\mathbf{Z}_1^\top \mathbf{Z}_1 = \mathbf{I}_r$ ;
- $\mathbf{Z}_2^\top \mathbf{Z}_2 = \mathbf{I}_{m-r}$ ;
- $\mathbf{Z}_1^\top \mathbf{Z}_2 = \mathbf{0}_{r \times (m-r)}$ ;
- $\mathbf{Z}_2^\top \mathbf{Z}_1 = \mathbf{0}_{(m-r) \times r}$ , onde  $\mathbf{0}$  representa a matriz nula.

Considerando agora só a parte relativa a  $\mathbf{Z}_1$ , podemos afirmar que o vetor  $\underline{u}$  representa uma projeção ortogonal de  $\underline{x}$  em  $S$ , pois  $\underline{u} = \mathbf{Z}_1 \mathbf{Z}_1^\top \underline{x}$ , em que a matriz que tem por colunas os elementos da base do subespaço linear  $S$  se designa por  $\mathbf{Z}_1$ .

$$\begin{aligned}
 \mathbf{Z}_1 \mathbf{Z}_1^\top \underline{x} &= \mathbf{Z}_1 \mathbf{Z}_1^\top (\mathbf{Z} \underline{\alpha}) & (2.1.47) \\
 &= \mathbf{Z}_1 \mathbf{Z}_1^\top [\mathbf{Z}_1 \mathbf{Z}_2] [\underline{\alpha}_1^\top \ \underline{\alpha}_2^\top]^\top \\
 &= [\mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{Z}_1 \quad \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{Z}_2] [\underline{\alpha}_1^\top \ \underline{\alpha}_2^\top]^\top \\
 &= \mathbf{Z}_1 \underline{\alpha}_1 \\
 &= \underline{u}.
 \end{aligned}$$

O vetor  $\underline{v}$  representa a projeção ortogonal de  $\underline{x}$  em  $S^\perp$ , uma vez que  $\underline{v} = \mathbf{Z}_2 \mathbf{Z}_2^\top \underline{x}$ , onde  $\mathbf{Z}_2$  é a matriz com colunas dos elementos da base do subespaço linear  $S^\perp$ , tendo-se

$$\begin{aligned}
 \mathbf{Z}_2 \mathbf{Z}_2^\top \underline{x} &= \mathbf{Z}_2 \mathbf{Z}_2^\top (\mathbf{Z} \underline{\alpha}) & (2.1.48) \\
 &= \mathbf{Z}_2 \mathbf{Z}_2^\top [\mathbf{Z}_1 \mathbf{Z}_2] [\underline{\alpha}_1^\top \ \underline{\alpha}_2^\top]^\top \\
 &= [\mathbf{Z}_2 \mathbf{Z}_2^\top \mathbf{Z}_1 \quad \mathbf{Z}_2 \mathbf{Z}_2^\top \mathbf{Z}_2] [\underline{\alpha}_1^\top \ \underline{\alpha}_2^\top]^\top \\
 &= \mathbf{Z}_2 \underline{\alpha}_2 \\
 &= \underline{v}.
 \end{aligned}$$

**Teorema 2.6** *Suponha-se que os vetores coluna de uma matriz  $\mathbf{Z}_1$ , do tipo  $m \times r$ , formam uma base ortonormada para o subespaço linear  $S$  (de dimensão  $r$ ), subespaço linear de  $E$ . Se  $\underline{x} \in E$ , então a projeção ortogonal de  $\underline{x}$  em  $S$  é dada por  $\mathbf{Z}_1 \mathbf{Z}_1^\top \underline{x}$ .*

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

Deste modo, podemos considerar a seguinte proposição.

**Proposição 2.7** *Suponha-se que as colunas da matriz  $Z_2$ , do tipo  $m \times (m - r)$ , formam uma base ortonormada para o subespaço linear  $S^\perp$  (de dimensão  $m - r$ ). Se  $\underline{x} \in E$ , então a projeção ortogonal de  $\underline{x}$  em  $S^\perp$  é dada por  $Z_2 Z_2^\top \underline{x}$ .*

Como a matriz  $Z_1 Z_1^\top$  é a matriz MPO, no subespaço linear  $S$ , e  $Z_2 Z_2^\top$  é a matriz MPO sobre o subespaço linear  $S^\perp$ , podemos dizer que

$$Z Z^\top = [Z_1 Z_2] \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} = Z_1 Z_1^\top + Z_2 Z_2^\top = I_m. \quad (2.1.49)$$

Como

$$Z Z^\top = Z_1 Z_1^\top + Z_2 Z_2^\top, \quad (2.1.50)$$

e  $Z Z^\top = I_m$ , podemos obter a seguinte equação simplificada,

$$Z_2 Z_2^\top = I_m - Z_1 Z_1^\top, \quad (2.1.51)$$

relativamente à MPO sobre o subespaço linear  $S$  e o seu complemento ortogonal,  $S^\perp$ .

Embora um subespaço linear não tenha uma base ortonormada única, a MPO formada a partir desta base ortonormada torna-se única, conforme se pode constatar no seguinte teorema.

**Teorema 2.7** *Suponha-se que as colunas da matriz  $Z_1$ , do tipo  $m \times r$ , formam uma base ortonormada para o subespaço linear  $S$ , de dimensão  $r$  e que as colunas da matriz  $W_1$ , do tipo  $m \times r$ , também formam uma base ortonormada para o mesmo subespaço linear. Então*

$$Z_1 Z_1^\top = W_1 W_1^\top. \quad (2.1.52)$$

**Definição 2.14 (Matriz de Projeção Ortogonal sobre o subespaço linear  $S$ )** *A matriz dada por  $P = Z_1 Z_1^\top$  será a MPO sobre o subespaço linear  $S$ .*

**Definição 2.15 (Espaço Nulo de uma Matriz)** *O espaço nulo de uma matriz  $A$ , representado por  $\mathcal{N}(A)$ , é tal que*

$$\mathcal{N}(A) = \{\underline{x} : A \underline{x} = \underline{0}\}. \quad (2.1.53)$$

**Definição 2.16 (Espaço Imagem de uma Matriz)** *O espaço imagem de uma matriz  $A$  é representado por*

$$R(A) = \{A \underline{x} : \underline{x} \in E\}. \quad (2.1.54)$$

Tem-se que  $R(P) = S$ , uma vez que  $P$  é a MPO em  $S$ , ou seja

$$R(P) = \{P \underline{x} : \underline{x} \in E\} = \{Z_1 Z_1^\top \underline{x} : \underline{x} \in E\} = S, \quad (2.1.55)$$

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

e, como  $\mathcal{N}(\mathbf{P}) = S^\perp$ , logo

$$\mathcal{N}(\mathbf{P}) = \{\underline{x} : \mathbf{P}\underline{x} = \underline{0}\} = \{\underline{x} : \mathbf{Z}_1\mathbf{Z}_1^\top \underline{x} = \underline{0}\} = S^\perp, \quad (2.1.56)$$

uma vez que o espaço imagem de  $\mathbf{P}$  é  $S$  e o espaço nulo de  $\mathbf{P}$  é  $S^\perp$ .

O espaço imagem de  $MPO$  pode ser representado pela soma direta de subespaços lineares, denominado por  $\boxplus$ . Temos assim os teoremas seguintes

**Teorema 2.8** *Sejam as MPO  $\mathbf{P}_1, \dots, \mathbf{P}_k$ , tais que  $\mathbf{P}_i\mathbf{P}_j = \mathbf{0}$ , com  $i \neq j$ ,  $i, j = 1, \dots, k$ . Então*

- $\mathbf{P} = \sum_{i=1}^k \mathbf{P}_i$  é uma MPO;
- $R(\mathbf{P}_i) \cap R(\mathbf{P}_j) = \mathbf{0}$ , com  $i \neq j$ ;
- $R(\mathbf{P}) = \boxplus_{i=1}^k R(\mathbf{P}_i)$ .

**Proposição 2.8** *Se  $\mathbf{P}$  é uma MPO, os seus valores próprios serão iguais a 0 ou 1.*

**Definição 2.17** *Duas MPO,  $\mathbf{P}_1$  e  $\mathbf{P}_2$ , são mutuamente ortogonais, ponhamos  $\mathbf{P}_1 \perp \mathbf{P}_2$ , se*

$$\mathbf{P}_2\mathbf{P}_1 = \mathbf{0}. \quad (2.1.57)$$

**Teorema 2.9** *Sejam  $S_1, \dots, S_k$ , subespaços dum mesmo espaço linear  $E$ , mutuamente ortogonais e  $S = \boxplus_{i=1}^k S_i$ . Seja  $\mathbf{P}_i$  a MPO sobre  $S_i$ ,  $i = 1, \dots, k$ . Então  $\mathbf{P} = \sum_{i=1}^k \mathbf{P}_i$  é uma MPO em  $S$ .*

**Teorema 2.10** *Sejam  $S$  um subespaço linear de  $E$  de dimensão  $r$ . Seja  $\mathbf{X} = [\underline{x}_1, \dots, \underline{x}_r]$  uma matriz, do tipo  $n \times r$ , cujas colunas formam uma base para  $S$ . Então a MPO sobre  $S$  é dada por*

$$\mathbf{P} = \mathbf{X}(\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top.$$

### 2.1.9 Produto de Kronecker de matrizes

O produto de Kronecker, denotado por  $\otimes$ , designa uma operação entre duas matrizes, que resulta numa matriz de blocos. Este tipo de produto tem importantes implicações na Estatística, pois o mesmo permite representações mais compactas e a sua estrutura permite simplificar cálculos de inversas, determinantes e valores próprios. O produto de Kronecker, ao contrário do produto usual de matrizes, está definido para qualquer tipo de matrizes. Esta operação foi estudada, por exemplo por Graham, A. (2018, págs. 21-35) [24] e Steeb, W. H. & Hardy, Y. (2011, págs. 83-146) [67] e apenas enunciaremos aqui os resultados necessários para tornar a leitura o mais auto-suficiente possível.

**Definição 2.18 (Produto de Kronecker)** *Seja a matriz  $\mathbf{A} = [a_{i,j}]$ , de ordem  $m \times n$ , e uma matriz  $\mathbf{B}$ , de ordem  $p \times q$ . O produto de Kronecker  $\mathbf{A} \otimes \mathbf{B}$  é obtido por uma matriz do tipo  $(mp \times nq)$ , definido por*

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

$$A \otimes B = \begin{bmatrix} a_{1,1}B & a_{1,2}B & \dots & a_{1,n}B \\ a_{2,1}B & a_{2,2}B & \dots & a_{2,n}B \\ \vdots & \ddots & \ddots & \vdots \\ a_{m,1}B & a_{m,2}B & \dots & a_{m,n}B \end{bmatrix}. \quad (2.1.58)$$

O produto de Kronecker não verifica a propriedade comutativa, ou seja  $A \otimes B \neq B \otimes A$ . Seguidamente, apresentam-se alguns resultados sobre o Produto de Kronecker, Schott, J. R. (2017, pág. 316) [60].

**Proposição 2.9** *Sejam  $A$ ,  $B$  e  $C$  matrizes quaisquer e  $\underline{a}$  e  $\underline{b}$  dois quaisquer vetores. Obtemos as seguintes propriedades:*

- $\alpha \otimes A = A \otimes \alpha = \alpha A$ , para qualquer escalar  $\alpha$ ;
- $(\alpha A) \otimes (\beta B) = \alpha\beta(A \otimes B)$ , para quaisquer escalares  $\alpha$  e  $\beta$ ;
- $(\alpha A) \otimes B = \alpha(A \otimes B) = A \otimes (\alpha B)$ ;
- $A \otimes (B \otimes C) = (A \otimes B) \otimes C$ ;
- $(A + B) \otimes (C + D) = A \otimes C + A \otimes D + B \otimes C + B \otimes D$ , satisfaz a propriedade distributiva;
- $(A + B) \otimes C = (A \otimes C) + (B \otimes C)$ , se  $A$  e  $B$  são matrizes do mesmo tipo;
- $A \otimes (B + C) = (A \otimes B) + (A \otimes C)$ , se  $B$  e  $C$  são matrizes do mesmo tipo;
- $(A \otimes B)^+ = A^+ \otimes B^+$ ;
- $\underline{a} \underline{b}^T = \underline{a} \otimes \underline{b}^T = \underline{b}^T \otimes \underline{a}$ ;

Assim, consegue-se facilmente estabelecer que

$$\left( \sum_{i=1}^m \alpha_i A_i \right) \otimes \left( \sum_{j=1}^n \beta_j B_j \right) = \sum_{i=1}^m \sum_{j=1}^n \alpha_i \beta_j (A_i \otimes B_j), \quad (2.1.59)$$

se as matrizes  $A_i$  e  $B_j$  forem do mesmo tipo. A proposição seguinte dá um resultado perfeito da combinação entre o produto usual de matrizes e o produto de Kronecker.

**Proposição 2.10** *Sejam as matrizes  $A$ ,  $B$ ,  $C$  e  $D$  matrizes do tipo  $m \times n$ ,  $r \times s$ ,  $n \times p$  e  $s \times t$ , respetivamente. Se o produto usual de matrizes está definido, então tem-se*

$$(A \otimes B)(C \otimes D) = AC \otimes BD. \quad (2.1.60)$$

**Proposição 2.11** *Sejam  $A$  e  $B$  matrizes idempotentes quaisquer. Então, o produto de Kronecker entre as matrizes  $A \otimes B$ , é uma matriz idempotente.*

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

**Proposição 2.12** *Sejam  $A$  e  $B$  matrizes ortogonais quaisquer. Então o produto de Kronecker  $A \otimes B$ , é uma matriz ortogonal.*

**Proposição 2.13** *Supondo que as matrizes  $A$  e  $B$  são matrizes quaisquer, então*

$$(A \otimes B)^{\top} = A^{\top} \otimes B^{\top}. \quad (2.1.61)$$

**Corolário 2.13.1** *Dadas matrizes simétricas quaisquer  $A$  e  $B$ , então*

$$(A \otimes B)^{\top} = A \otimes B. \quad (2.1.62)$$

**Proposição 2.14** *Se  $P_j$  é a matriz diagonalizadora ortogonal de  $A_j$ ,  $j = 1, 2$ ,  $P_1 \otimes P_2$  será diagonalizadora ortogonal de  $A_1 \otimes A_2$  e os valores próprios de  $A_1 \otimes A_2$  serão o produto dos valores próprios de  $A_1$  pelos valores próprios de  $A_2$ .*

Da aplicação desta proposição resulta que o número de valores próprios não nulos de  $B_1 B_1^{\top} \otimes B_2 B_2^{\top}$  é o produto do número de valores próprios relativos a  $B_1 B_1^{\top}$  e  $B_2 B_2^{\top}$ . Como esses números correspondem às características das matrizes  $B_1 B_1^{\top} \otimes B_2 B_2^{\top}$ ,  $B_1 B_1^{\top}$  e  $B_2 B_2^{\top}$ , logo, obtém-se  $\text{car}(B_1 \otimes B_2) = \text{car}(B_1)\text{car}(B_2)$ .

**Proposição 2.15** *O produto de Kronecker de uma MPO é uma MPO.*

## 2.2 Funções Geradoras de Cumulantes

Na secção Anexos apresentamos os principais resultados de momentos e funções geradoras de momentos que iremos necessitar para uma plena compreensão desta secção. Nesta, iremos ainda apresentar o relacionamento entre momentos e cumulantes, como obter os estimadores dos quartos cumulantes e como construir elipsóides de confiança e intervalos de predição utilizando os estimadores desses cumulantes.

Assim, começamos por apresentar os cumulantes, introduzidos por Thiele, T. N. (1931, págs. 165-308) [69], que têm desempenhado um papel importante na caracterização de distribuições, como referem Andrews, D. F. (2001, págs. 7-16) [1], Novak, J. & Sniady, P. (2011, págs. 300-301) [49], Pistone, G. & Wynn, H.P. (2006, págs. 210-221) [55] e Withers, C. S., Nadarajah, S. & Shih, S. H. (2015, págs. 541-564) [71].

Sendo  $\varphi(t|U)$  a função geradora de momentos da variável aleatória  $U$ , a respetiva função geradora de cumulantes será dada pelo logaritmo natural da função geradora de momentos

$$\psi(t|U) = \log \varphi(t|U), \quad (2.2.1)$$

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

sendo os  $r$ -ésimos momentos (relativos à origem) e  $r$ -ésimos cumulantes dados, respetivamente, por

$$\begin{cases} \mu'_r(U) = \varphi^{<r>}(0|U) \\ \mathcal{O}_r(U) = \psi^{<r>}(0|U) \end{cases} . \quad (2.2.2)$$

Aqui  $\mathcal{O}_r(U)$  designa o  $r$ -ésimo cumulante de  $U$ ,  $\psi^{<r>}(0|U)$  designa a  $r$ -ésima derivada de  $U$  na origem, da função geradora de cumulantes, e  $<r>$  indica a ordem de derivação, para evitar alguma possível confusão com potências.

Sendo  $\mu(U)$ ,  $\sigma^2(U)$  e  $\mu_r(U)$  o valor médio, a variância e o  $r$ -ésimo momento central de  $U$ , têm-se os bem conhecidos resultados, DasGupta, A. (2008, pág. 195) [13]).

$$\begin{cases} \mathcal{O}_1(U) = \mu(U) \\ \mathcal{O}_2(U) = \sigma^2(U) \\ \mathcal{O}_3(U) = \mu_3(U) \\ \mathcal{O}_4(U) = \mu_4(U) - 3(\sigma^2(U))^2 \end{cases} . \quad (2.2.3)$$

O facto de nesta última expressão aparecer o termo  $3(\sigma^2(U))^2$  estará na base de, no caso geral, se ter de trabalhar com pares de modelos *i.i.d.*. Assim, cada modelo do par fornece um estimador centrado da variância. Como estes estimadores são independentes, o seu produto será um estimador centrado do quadrado da variância.

Para obter os  $\mathcal{O}_r(U)$ , a partir dos  $\mu_r(U)$ ,  $r > 4$ , observamos que, com

$$\psi'(t|U) = \frac{\varphi'(t|U)}{\varphi(t|U)}, \quad (2.2.4)$$

temos

$$\varphi(t|U) \psi'(t|U) = \varphi'(t|U), \quad (2.2.5)$$

pelo que, utilizando a fórmula de Euler, de acordo com Olver, P. J. (2000, págs. 27-97) [52], para a derivada de ordem  $r$  dum produto de funções, dada por

$$(h(z)l(z))^{<r>} = \sum_{j=0}^r \binom{r}{j} h^{<j>}(z) l^{<r-j>}(z), \quad (2.2.6)$$

obtemos

$$\varphi^{<r+1>}(t|U) = (\varphi(t|U) \psi'(t|U))^{<r>} = \sum_{j=0}^r \binom{r}{j} \varphi^{<j>}(t|U) \psi^{<r+1-j>}(z), \quad (2.2.7)$$

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

vindo

$$\mu'_{r+1}(U) = \sum_{j=0}^r \binom{r}{j} \mu'_j(U) \mathcal{O}_{r+1-j}(U), \quad (2.2.8)$$

bem como

$$\mathcal{O}_{r+1}(U) = \mu'_{r+1}(U) - \sum_{j=1}^r \binom{r}{j} \mu'_j(U) \mathcal{O}_{r+1-j}(U). \quad (2.2.9)$$

Com  $E(U)$  indicando o valor esperado de  $U$ , temos ainda

$$\mu_r(U) = E[(U - \mu(U))^r] = \sum_{j=0}^r \binom{r}{j} (-1)^j \mu(U)^j E(U^{r-j}) = \sum_{j=0}^r \binom{r}{j} (-1)^j \mu(U)^j \mu'_{r-j}(U). \quad (2.2.10)$$

Finalmente, temos

$$\mu'_r(U) = E[(U - \mu(U) + \mu(U))^r] = \sum_{j=0}^r \binom{r}{j} \mu(U)^{r-j} \mu'_j(U). \quad (2.2.11)$$

Vemos assim que os cumulantes, os momentos centrais e os momentos, relativos à origem, se determinam mutuamente, Balakrishnan, N., Johnson, N. L. & Kotz, S. (1998) [4]. Se  $U_1, \dots, U_w$  forem independentes, temos, de acordo com Craig, C. C. (1931) [12],

$$\varphi \left( t \left| \sum_{i=1}^w a_i U_i \right. \right) = E(e^{t \sum_{i=1}^w a_i U_i}) = \prod_{i=1}^w E(e^{a_i t U_i}) = \prod_{i=1}^w \varphi(a_i t | U_i), \quad (2.2.12)$$

vindo

$$\psi \left( t \left| \sum_{i=1}^w a_i U_i \right. \right) = \sum_{i=1}^w \psi(a_i t | U_i), \quad (2.2.13)$$

bem como,

$$\psi^{<r>} \left( t \left| \sum_{i=1}^w a_i U_i \right. \right) = \sum_{i=1}^w a_i^r \psi^{<r>}(a_i t | U_i), \quad (2.2.14)$$

e os cumulantes das funções lineares são dados em função das combinações lineares das parcelas; para  $t = 0$ , temos os  $r$ -ésimos cumulantes das funções lineares na origem

$$\mathcal{O}_r \left( \sum_{i=1}^w a_i U_i \right) = \sum_{i=1}^w a_i^r \mathcal{O}_r(U_i). \quad (2.2.15)$$

Em particular, se as  $U_1, \dots, U_w$ , forem *i.i.d.*, com  $r$ -ésimos cumulantes  $\mathcal{O}_r$ , temos

$$\mathcal{O}_r \left( \sum_{i=1}^w a_i U_i \right) = b(r) \mathcal{O}_r(U), \quad (2.2.16)$$

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

com  $b(r) = \sum_{i=1}^w a_i^r$ . Admitamos agora que a distribuição de  $U$  tem o par de parâmetros de localização e dispersão  $(\lambda, \delta)$ , pondo-se  $U \sim V(|\lambda, \delta)$  e, tendo-se  $U = \lambda + \delta U_0$ , com  $U_0 \sim V(|0, 1)$ , pelo que

$$V(x|\lambda, \delta) = V\left(\frac{x - \lambda}{\delta} | 0, 1\right). \quad (2.2.17)$$

Vê-se ainda que

$$\varphi(t|\lambda, \delta) = E(e^{tU}) = e^{\lambda t} E(e^{\delta t U_0}) = e^{\lambda t} \varphi(\delta t | 0, 1), \quad (2.2.18)$$

e aplicando logaritmo,

$$\psi(t|\lambda, \delta) = \lambda t + \psi(\delta t | 0, 1). \quad (2.2.19)$$

Assim, derivando a expressão anterior, tem-se

$$\begin{cases} \psi'(t|\lambda, \delta) = \lambda + \delta \psi'(t | 0, 1) \\ \psi^{<r>}(t|\lambda, \delta) = \delta^r \psi^{<r>}(t | 0, 1), \quad r > 1 \end{cases}, \quad (2.2.20)$$

vindo, em particular,

$$\delta = \sqrt{\frac{\mathcal{O}_2(|\lambda, \delta)}{\mathcal{O}_2(|0, 1)}}. \quad (2.2.21)$$

Se se admitir que os  $\mathcal{O}_r(|0, 1)$  são conhecidos, não serão precisos pares de modelos para estimar os cumulantes de quarta ordem, visto que, estimando  $\mathcal{O}_2(|\lambda, \delta)$  está estimado  $\delta$  e, conseqüentemente, os cumulantes de qualquer ordem.

Em seguida consideraremos o caso normal, pondo  $U \sim N(|\mu, \sigma^2)$  para indicar que  $U$  segue uma distribuição normal, com valor médio  $\mu$  e variância  $\sigma^2$ . Tem-se então

$$\varphi(t|\mu, \sigma^2) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}, \quad (2.2.22)$$

vindo

$$\psi(t|\mu, \sigma^2) = \mu t + \frac{\sigma^2}{2} t^2, \quad (2.2.23)$$

pelo que

$$\begin{cases} \psi'(t|\mu, \sigma^2) = \mu + \sigma^2 t \\ \psi^{<2>}(t|\mu, \sigma^2) = \sigma^2 \\ \psi^{<r>}(t|\mu, \sigma^2) = 0, \quad r > 2 \end{cases}, \quad (2.2.24)$$

e que

$$\begin{cases} \mathcal{O}_1(\mu, \sigma^2) = \mu \\ \mathcal{O}_2(\mu, \sigma^2) = \sigma^2 \\ \mathcal{O}_r(\mu, \sigma^2) = 0, \quad r > 2 \end{cases}, \quad (2.2.25)$$

ou seja, na distribuição normal todos os cumulantes de ordem superior ao segundo são zero, sendo esta distribuição definida inteiramente pela média e variância.

## 2.3 Expansões de Edgeworth

Nesta secção iremos obter uma útil aproximação à função distribuição  $F_U$ , de  $U$ . Se a variável aleatória  $U$  tiver momentos  $\mu(U)$ ,  $\sigma^2(U)$ ,  $\mu_3(U)$  e  $\mu_4(U)$  definidos, teremos os coeficientes de assimetria e achatamento

$$\begin{cases} \gamma_1(U) = \frac{\mu_3(U)}{(\sigma^2(U))^{\frac{3}{2}}} \\ \gamma_2(U) = \frac{\mu_4(U)}{(\sigma^2(U))^2} - 3 \end{cases}, \quad (2.3.1)$$

tendo-se, DasGupta, A. (2008, pág. 187) [13], Edgeworth, F. (1905, 1906) [15] [16] e Kolassa, J. E. (1994, pág. 65) [35], para a distribuição  $F_{\dot{U}}$  de

$$\dot{U} = \frac{U - \mu(U)}{\sigma(U)} \quad (2.3.2)$$

a Expansão de Edgeworth, Kendall, M. G., Stuart, A. & Ord, J. K. (1987, §3) [32], pode ser escrita em função dos coeficientes da distribuição  $F_{\dot{U}}$

$$Ed_{\dot{U}}(u) = N(u) - \frac{\gamma_1(U)}{3!} n^{<2>}(u) + \frac{\gamma_2(U)}{4!} n^{<3>}(u) + 10 \frac{\gamma_1(U)^2}{6!} n^{<6>}(u), \quad (2.3.3)$$

onde  $N(\cdot)$  e  $n(\cdot)$  representam, respetivamente, a distribuição e a densidade de uma distribuição normal reduzida. Sendo  $\dot{u}_p$  o  $p$ -ésimo quantil de  $Ed_{\dot{U}}(u)$ , que será uma aproximação do  $p$ -ésimo quantil de  $F_{\dot{U}}$ , tem-se

$$u_p = \mu(U) + \sigma(U)\dot{u}_p, \quad (2.3.4)$$

sendo este uma aproximação ao  $p$ -ésimo quantil de  $F_{\dot{U}}$ . No que segue, substituiremos nestas expressões  $\mu(U)$ ,  $\sigma(U)$ ,  $\gamma_1(U)$  e  $\gamma_2(U)$  pelos seus estimadores.

## 2.4 Partições Ortogonais de Sub-Espaços

### 2.4.1 Considerações prévias

O estudo dos modelos aditivos é feito a dois níveis:

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

- modelos individuais;
- modelos múltiplos.

No primeiro nível temos os modelos individuais, sendo  $n$  o número de observações do vetor  $\underline{Y}$ , tem-se o modelo

$$\underline{Y} = \mathbf{X}\underline{\beta} + \sum_{i=1}^w \mathbf{X}_i \underline{Z}_i, \quad (2.4.5)$$

onde apenas admitimos que os vetores aleatórios  $\underline{Z}_1, \dots, \underline{Z}_w$ , são independentes e que as matrizes  $\mathbf{X}, \mathbf{X}_1, \dots, \mathbf{X}_w$ , são dadas. Os vetores  $\underline{Z}_1, \dots, \underline{Z}_w$ , têm vetores médios nulos, com  $c_1, \dots, c_w$  componentes *i.i.d.* e cumulantes  $\chi_{r,i}(\mathbf{0}, \delta_i) = \delta_i^r \chi_{r,i}(\mathbf{0}, 1)$ ,  $i = 1, \dots, w$ , de ordens  $r = 2, 3, 4$ . Recorde-se que o segundo cumulante coincide com a variância. As matrizes  $\mathbf{X}, \mathbf{X}_1, \dots, \mathbf{X}_w$ , são dos tipos  $n \times k$ ,  $n \times c_1, \dots, n \times c_w$ , respetivamente e o vetor dos coeficientes dos efeitos fixos é dado por  $\underline{\beta}$ , com  $k$  componentes. Nos modelos múltiplos tem-se um modelo individual para cada tratamento de um delineamento de efeitos fixos. As matrizes  $\mathbf{X}, \mathbf{X}_1, \dots, \mathbf{X}_w$ , são as mesmas para todos os modelos individuais, bem como as distribuições dos vetores aleatórios homólogos. Assim a inferência centrar-se-á no estudo da ação dos fatores do delineamento base nos vetores individuais dos coeficientes sendo-se levados a testar hipóteses definidas a partir de partições ortogonais de subespaços, de forma semelhante ao que tinha sido feito em Moreira et al. (2005) [47].

Assim, estes modelos, em geral, não terão estrutura por blocos ortogonal, para o que seria necessário que as matrizes, ver Mexia, J. T. (1988) [44],

$$\mathbf{M}_i = \mathbf{X}_i \mathbf{X}_i^\top, \quad i = 1, \dots, w, \quad (2.4.6)$$

comutassem. Caso isso se verificasse, o modelo teria matriz de covariância, dada por

$$\mathbf{V}(\underline{\gamma}) = \sum_{j=1}^m \gamma_j \mathbf{Q}_j, \quad (2.4.7)$$

com  $\mathbf{Q}_1, \dots, \mathbf{Q}_m$ , *MPO*, mutuamente ortogonais, e o modelo estaria associado à partição ortogonal

$$\Omega = \boxplus_{j=1}^m \nabla_j, \quad (2.4.8)$$

do espaço imagem  $\Omega = R \left( \sum_{j=1}^m \mathbf{Q}_j \right)$ , nos subespaços  $\nabla_j = R(\mathbf{Q}_j)$ ,  $j = 1, \dots, m$ . No entanto, as partições ortogonais de subespaços serão consideradas nos fatores de modelos múltiplos. Com efeito, os modelos de uma tal família corresponderão aos  $d$  tratamentos de um delineamento de efeitos fixos associados a uma partição ortogonal

$$\mathbb{R}^d = \boxplus_{j=1}^m \nabla_j. \quad (2.4.9)$$

Consideraremos agora dois casos importantes de delineamento de efeitos fixos:

- modelos obtidos cruzando grupos de fatores aninhados;

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

- modelos factoriais de base prima  $p$ ,

apresentando para um, e para outro, as partições relevantes em subespaços.

### 2.4.2 Cruzamento e Aninhamento

Suponhamos que temos  $L$  grupos com  $u_1, \dots, u_L$  fatores. Ponhamos  $c_l(0) = 1$  e  $c_l(-1) = 0$ , por conveniência, e representemos por  $a_l(1)$ ,  $l = 1, \dots, L$ , o número de níveis do primeiro fator do  $l$ -ésimo grupo. Se  $u_l > 1$ , haverá aninhamento para o  $l$ -ésimo grupo de fatores, onde cada nível encaixará  $a_l(2)$  níveis dos fatores seguintes, segundo Ferreira, S. S., Ferreira, D. & Mexia, J.T. (2006) [17] e Fonseca, M., Mexia, J. T. & Zmyslony, R. (2003) [21].

Haverá  $c_l(h) = \prod_{k=1}^h a_l(k)$  combinações de níveis para os  $h$  primeiros fatores do  $l$ -ésimo grupo, cada um dos quais encaixará  $b_l(h) = c_l(u_l)/c_l(h)$ ,  $l = 1, \dots, L$ ,  $h = 1, \dots, u_l$ , combinações de níveis para os restantes fatores do grupo, com  $c_l(u_l) = \prod_{h=1}^{u_l} a_l(h)$ ,  $l = 1, \dots, L$ .

Os efeitos e interações do modelo estarão associados aos vetores  $\underline{h} \in \Gamma$ , com

$$\Gamma = \{\underline{h} : h_l = 0, \dots, u_l, l = 1, \dots, L\}. \quad (2.4.10)$$

Ao vetor nulo  $\underline{0}$  está associado o valor médio geral. Se  $\underline{h}^L \in \Gamma$  tiver uma única componente não nula, digamos  $h_{l'}$ , estão-lhe associados os efeitos dos níveis do fator com esse índice do grupo  $l'$ , e diremos que esse factor está indiciado por  $h_{l'}$ . Se  $\underline{h}^L$  tiver mais de uma componente não nula, corresponder-lhes-ão as interações entre os níveis dos fatores indiciados por essas componentes. Pondo-se  $c(\underline{h}^L) = \prod_{l=1}^L c_l(h_l)$ , ter-se-á  $c(\underline{0}^L) = 1$ , enquanto que para os outros vetores de  $\Gamma$ ,  $c(\underline{h}^L)$  será o número de níveis ou de combinações de níveis dos fatores indiciados por  $\underline{h}^L$ .

O cruzamento e aninhamento de modelos já foi estudado por vários autores, por isso apresentamos aqui apenas as noções necessárias para um maior entendimento de certos desenvolvimentos em secções vindouras; no entanto, para mais pormenores, ver, por exemplo, Ferreira, S. S., Ferreira, D. & Mexia, J.T. (2006) [17] e Khuri, A. I., Mathew, T. & Sinha, B. K. (1998, págs. 19-46, 71-75) [34].

Assim, temos

$$\dim(\nabla(\underline{h}^L)) = \text{car}(\mathbf{A}(\underline{h}^L)^\top) = \prod_{l=1}^L \left( \prod_{h'_l=0}^{k_l} a_{l,h'_l} - \prod_{h'_l=0}^{k_l-1} a_{l,h'_l} \right), \underline{h}^L \in \Gamma. \quad (2.4.11)$$

Poremos

$$g(\underline{h}^L) = \text{car}(\mathbf{A}(\underline{h}^L)), \underline{h}^L \in \Gamma. \quad (2.4.12)$$

Admitamos agora que se tem  $\underline{Y} \sim N(\underline{\mu}, \sigma^2 \mathbf{I}_n)$ . Então,

$$S(\underline{h}^L) = \|\mathbf{A}(\underline{h}^L) \underline{Y}\|^2, \quad (2.4.13)$$

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

onde  $\|\cdot\|$  denota a norma euclidiana, será, como referem Khuri, A. I., Mathew, T. & Sinha, B. K. (1998, págs. 19-46, 71-75) [34], o produto de  $\sigma^2$  por um qui-quadrado com  $g(\underline{h}^L)$  graus de liberdade e parâmetro de não centralidade

$$\delta(\underline{h}^L) = \frac{1}{\sigma^2} \|\mathbf{A}(\underline{h}^L) \underline{\mu}\|^2. \quad (2.4.14)$$

Se admitirmos que existe um conjunto  $D \subset \Gamma$ , tal que, se  $\underline{h}^L \in D$ ,

$$\mathbf{A}(\underline{h}^L) \underline{\mu} = \underline{0}_{g(\underline{h}^L)}, \quad (2.4.15)$$

para testarmos

$$H_0(\underline{h}^L) : \mathbf{A}(\underline{h}^L) \underline{\mu} = \underline{0}_{g(\underline{h}^L)}, \quad \underline{h}^L \in \Gamma - D. \quad (2.4.16)$$

Assim, podemos utilizar a estatística

$$\mathcal{F} = \frac{g}{g(\underline{h}^L)} \frac{S(\underline{h}^L)}{S}, \quad (2.4.17)$$

com  $g(\underline{h}^L)$  e  $g = \sum_{\underline{h}' \in D} g(\underline{h}')$  graus de liberdade, onde

$$S = \sum_{\underline{h}' \in D} S(\underline{h}'). \quad (2.4.18)$$

Observe-se que

$$\mathbf{A}(\underline{h}^L) \mathbf{A}(\underline{h}^L)^\top = \mathbf{I}_{g(\underline{h}^L)}, \quad \underline{h}^L \in \Gamma, \quad (2.4.19)$$

e que a MPO sobre  $\nabla(\underline{h}^L)$  é dada por

$$\mathbf{P}(\underline{h}^L) = \mathbf{A}(\underline{h}^L)^\top \mathbf{A}(\underline{h}^L), \quad \underline{h}^L \in \Gamma, \quad (2.4.20)$$

sendo esta uma MPO. Tem-se ainda

$$\mathbf{A}(\underline{h}^L)^+ = \mathbf{A}(\underline{h}^L)^\top, \quad \underline{h}^L \in \Gamma. \quad (2.4.21)$$

Por outro lado, para obter intervalos de predição, como veremos adiante, é útil considerar

$$\underline{\mu}(\underline{h}^L) = \mathbf{A}(\underline{h}^L) \underline{\mu}, \quad \underline{h}^L \in \Gamma. \quad (2.4.22)$$

### 2.4.3 Fatoriais de base prima

É amplamente aceite que os delineamentos experimentais mais usados são delineamentos fatoriais completos e delineamentos fatoriais fracionários com 2 e 3 níveis, respetivamente. Os delineamentos fatoriais têm uma variável dependente de várias variáveis controláveis, chamadas fatores. Para cada fator, há dois ou mais níveis, ou categorias, dependendo da situação experimental. Uma combinação de níveis é chamada tratamento. Um delineamento fatorial é um delineamento cujos fatores se cruzam e onde o efeito dos fatores individuais, bem como as

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

interações entre eles, são investigados, como por exemplo, Cheng, C. S., Deng, L. Y. & Tang, B. (2002) [9], Oliveira, S. (2005, 2018) [50] [51] e Ye, K. Q. (2003) [72].

Quando todas as combinações de níveis de dois ou mais fatores são consideradas, dizemos que esses fatores se cruzam e temos o caso completo. Quando é impossível considerar todos os tratamentos e, podemos admitir que as interações de ordem mais alta podem ser ignoradas, podemos considerar um subconjunto dos tratamentos e, nesse caso, temos uma replicação fatorial fracionária.

Suponhamos que num destes delineamentos existem  $N > 1$  fatores com um número  $p$ , primo, de níveis. Considerando o caso completo, temos  $p \times \dots \times p = p^N$ , um fatorial de base prima, com  $n = p^N$  tratamentos. Os  $p^N$  tratamentos serão representados pelos vetores  $\underline{x} = \{x_1, \dots, x_N\}$ , com componentes  $x_j = 0, \dots, p-1$ ,  $j = 1, \dots, N$ . Seja  $L(\underline{x}|\underline{a})$  a família de aplicações lineares

$$L(\underline{x}|\underline{a}) = \left( \sum_{j=1}^N a_j x_j \right)_{(p)}, \quad (2.4.23)$$

com  $a_j = 0, \dots, p-1$ ,  $j = 1, \dots, N$ , as componentes do vetor  $\underline{a}$  e onde  $(p)$  indica o uso da aritmética módulo  $p$ . As aplicações lineares  $L(\underline{x}|\underline{a}_l)$ ,  $l = 1, \dots, m$ , são linearmente independentes se e só se  $\underline{a}_1, \dots, \underline{a}_m$  forem linearmente independentes. Haverá assim  $U_N(p) = \frac{p^N - 1}{p - 1}$  aplicações lineares cujo primeiro coeficiente não nulo é 1, Jesus, V. (2008) [28] e Jesus, V., Mexia, J. T., Fonseca, M. & Zmyślony, R. (2009) [29]. Essas aplicações são chamadas de aplicações reduzidas. Podemos ordenar as aplicações reduzidas de 1 a  $U_N(p)$ , e os tratamentos, de acordo com os índices

$$f(\underline{x}) = 1 + \sum_{j=1}^N x_j p^{j-1}. \quad (2.4.24)$$

Sendo  $L_1, \dots, L_u$  aplicações reduzidas linearmente independentes, o conjunto

$$[\underline{L}|\underline{b}] = [L_1, \dots, L_u | b_1, \dots, b_u], \quad (2.4.25)$$

de tratamentos de  $\underline{x}$ , tais que  $L_i(\underline{x})_{(p)} = b_i$ ,  $i = 1, \dots, u$ , será denominado de bloco. Desde que o sistema de equações que define um bloco nos permita expressar as  $u$  componentes do vetor  $\underline{x}$  como combinações lineares das restantes  $N - u$  componentes, em todos os  $[\underline{L}|\underline{b}]$ , haverá  $p^{N-u}$  tratamentos e  $p^u$  blocos.

As aplicações reduzidas podem ser ordenadas respeitando o aumento dos índices, isto é, se  $l(a_1) < l(a_2)$ , então  $U(a_1) < U(a_2)$ ,  $U(a) = 1, \dots, U_N(p)$ . Assim, considera-se que num fatorial  $p^N$ , as aplicações reduzidas de ordem zero dizem respeito aos efeitos dos fatores principais, enquanto que as aplicações reduzidas de ordem superior a zero correspondem às interações fatoriais entre fatores que contêm coeficientes não nulos. A cada aplicação linear  $L_h = L_h(\underline{x})$ ,  $h = 1, \dots, U_N(p)$ , podemos associar a matriz  $C(\ell_h)$ ,  $h = 1, \dots, U_N(p)$ , com elementos

$$C_{i,j}(\ell_h) = \begin{cases} 1, & \ell_h(x_j) = i - 1 \\ 0, & \ell_h(x_j) \neq i - 1 \end{cases}, \quad i = 1, \dots, p, j = 1, \dots, p^N, h = 1, \dots, U_N(p) \quad (2.4.26)$$

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

Temos as matrizes

$$\begin{cases} \mathbf{C}(\ell_h)\mathbf{C}(\ell_h)^\top = p^{N-1}\mathbf{I}_p, & h = 1, \dots, U_N(p) \\ \mathbf{C}(\ell_h)\mathbf{C}(\ell_{h'})^\top = p^{N-2}\mathbf{J}_p, & h \neq h' \end{cases}, \quad (2.4.27)$$

com  $\mathbf{J}_p = \mathbf{1}_p\mathbf{1}_p^\top$ , sendo  $\mathbf{1}_p$  o vetor com todas as  $p$  componentes iguais a 1.

Seja  $\mathbf{K}_p$  a matriz que se obtém eliminando a primeira linha, igual a  $\frac{1}{\sqrt{p}}\mathbf{1}_p^\top$ , a uma matriz ortogonal do tipo  $p \times p$ .

As matrizes para cada aplicação linear reduzida  $\ell_h$ ,  $h = 1, \dots, U_N(p)$ , são dadas por

$$\mathbf{A}(\ell_h) = \frac{1}{\sqrt{p^{N-1}}}\mathbf{K}_p\mathbf{C}^\top(\ell_h), \quad h = 1, \dots, U_N(p), \quad (2.4.28)$$

e são matrizes ortogonais associadas a uma álgebra de Jordan comutativa,  $AJC$ , com base principal  $\{\frac{1}{p^N}\mathbf{J}_{p^N}, \mathbf{Q}(\ell_1), \dots, \mathbf{Q}(\ell_{U_N(p)})\}$ , ver Mexia, J. T. (1995) [46], onde

$$\mathbf{Q}(\ell_h) = \mathbf{A}(\ell_h)^\top \mathbf{A}(\ell_h), \quad h = 1, \dots, U_N(p). \quad (2.4.29)$$

Assim, as matrizes  $\frac{1}{p^N}\mathbf{J}_{p^N}$  e  $\mathbf{Q}(\ell_1), \dots, \mathbf{Q}(\ell_{U_N(p)})$  são matrizes simétricas, idempotentes e mutuamente ortogonais (ver definição 2.4, Proposição 2.4 e definição 2.4 da secção 2). Ver também a secção Anexos para mais detalhes acerca de Álgebras de Jordan Comutativas.

Note-se que podemos usar as matrizes  $\mathbf{A}(\ell_h)$  para obter os vetores  $\tilde{\eta}_h = \mathbf{A}(\ell_h)\underline{y}$ ,  $h = 1, \dots, U_N(p)$  e formular as hipóteses  $H_{0,h} : \eta_h = \underline{0}$ ,  $h = 1, \dots, U_N(p)$ .

## Capítulo 3

### Caso Geral de Modelos Aditivos

#### 3.1 Ajustamentos

Iremos de seguida apresentar o estudo geral dos modelos aditivos. Obteremos os *LSE* dos cumulantes  $\chi_{r,i}, i = 1, \dots, w, r = 2, 3, 4$ , das componentes de  $Z_i, i = 1, \dots, w$ , e dos vetores de coeficientes  $\underline{\beta}$ . Mostraremos que, para obter os estimadores dos quartos cumulantes,  $\tilde{\chi}_{4,i}, i = 1, \dots, w$ , teremos de dispor de um par de vetores de observações  $\underline{Y}(1)$  e  $\underline{Y}(2)$  *i.i.d.*. Esta exigência pode ser descartada se apenas quisermos estimar os vetores  $\underline{\beta}$  e os cumulantes até à terceira ordem. Nesse caso, o modelo a considerar seria

$$\underline{Y} = \mathbf{X}\underline{\beta} + \sum_{i=1}^w \mathbf{X}_i Z_i, \quad (3.1.1)$$

onde  $\underline{Y}$ , o vetor de observações, é do tipo  $n \times 1$ , os vetores  $Z_1, \dots, Z_w$ , são independentes e com componentes *i.i.d.*. Assume-se que estes vetores têm valores médios nulos e cumulantes até à quarta ordem. As matrizes  $\mathbf{X}, \mathbf{X}_1, \dots, \mathbf{X}_w$ , dadas, são dos tipos  $n \times k, n \times c_1, \dots, n \times c_w$  e o vetor  $\underline{\beta}$ , de coeficientes, é do tipo  $k \times 1$ . Note-se que às componentes dos vetores  $Z_1, \dots, Z_w$ , não se associa qualquer distribuição, estamos assim no caso geral, pelo que poderão ter distribuições distintas, alargando-se o âmbito destes modelos em que é usual ter apenas vetores normais. Como admitimos que  $\mathbf{X}$  é do tipo  $n \times k$ , a sua característica será  $k$ ; assim  $\Omega^\perp$ , o complemento ortogonal de  $\Omega = R(\mathbf{X})$ , terá dimensão  $n - k$ . Seja então  $\{\underline{\alpha}_1, \dots, \underline{\alpha}_{n-k}\}$  uma base ortonormada para  $\Omega^\perp$ . Tendo-se os modelos

$$\underline{Y}(h) = \mathbf{X}\underline{\beta}(h) + \sum_{i=1}^w \mathbf{X}_i Z_i(h), \quad h = 1, 2, \quad (3.1.2)$$

cujas matrizes  $\mathbf{X}, \mathbf{X}_1, \dots, \mathbf{X}_w$ , serão as mesmas e tomemos

$$\dot{Y}_l(h) = \underline{\alpha}_l^\top \underline{Y}(h) = \sum_{i=1}^w \underline{a}_{l,i}^\top Z_i(h), \quad l = 1, \dots, \dot{n}, \quad h = 1, 2, \quad (3.1.3)$$

com  $\dot{n} = n - k$ , e

$$\underline{a}_{l,i}^\top = \underline{\alpha}_l^\top \mathbf{X}_i = (a_{l,i,1}, \dots, a_{l,i,c_i}), \quad l = 1, \dots, \dot{n}, \quad i = 1, \dots, w. \quad (3.1.4)$$

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

Como o par de vetores de observações,  $\underline{Y}(1)$  e  $\underline{Y}(2)$ , são *i.i.d.*, o mesmo é verificado para os vetores  $\dot{Y}(1) = (\dot{Y}_1(1), \dots, \dot{Y}_{\dot{n}}(1))$  e  $\dot{Y}(2) = (\dot{Y}_1(2), \dots, \dot{Y}_{\dot{n}}(2))$ .

O  $r$ -ésimo cumulante de  $\dot{Y}_l(h)$ ,  $h = 1, 2$ ,  $l = 1, \dots, \dot{n}$ , será dado por  $\mathcal{O}_{r,l}(h)$ ,

$$\mathcal{O}_{r,l}(h) = \sum_{i=1}^w b_{l,i}(r) \chi_{r,i}, \quad r = 2, 3, 4, \quad l = 1, \dots, \dot{n}, \quad h = 1, 2, \quad (3.1.5)$$

já que  $\dot{Y}_l(1)$  e  $\dot{Y}_l(2)$ ,  $l = 1, \dots, \dot{n}$ , são *i.i.d.* e temos

$$b_{l,i}(r) = \sum_{h=1}^{c_i} a_{l,i,h}^r, \quad l = 1, \dots, \dot{n}, \quad i = 1, \dots, w, \quad r = 2, 3, 4. \quad (3.1.6)$$

Têm-se, ainda, os estimadores

$$\begin{cases} \tilde{\mathcal{O}}_{r,l} = \frac{\dot{Y}_l(1)^r + \dot{Y}_l(2)^r}{2}, \quad r = 2, 3 \\ \tilde{\mathcal{O}}_{4,l} = \frac{\dot{Y}_l(1)^4 + \dot{Y}_l(2)^4}{2} - 3\dot{Y}_l(1)^2 \dot{Y}_l(2)^2 \end{cases}, \quad l = 1, \dots, \dot{n}. \quad (3.1.7)$$

Então, com

$$\mathbf{B}(r) = [b_{l,i}(r)], \quad l = 1, \dots, \dot{n}, \quad i = 1, \dots, w, \quad r = 2, 3, 4, \quad (3.1.8)$$

tem-se

$$E(\tilde{\mathcal{O}}_r) = \mathbf{B}(r) \tilde{\chi}_r, \quad (3.1.9)$$

com  $\tilde{\chi}_r = (\tilde{\chi}_{r,1}, \dots, \tilde{\chi}_{r,w})$  e  $\tilde{\mathcal{O}}_r = (\tilde{\mathcal{O}}_{r,1}, \dots, \tilde{\mathcal{O}}_{r,\dot{n}})$ ,  $r = 2, 3, 4$ . Obtêm-se assim os *LSE*, dos cumulantes  $\chi_{r,i}$ , dados por

$$\tilde{\chi}_r = (\mathbf{B}(r)^\top \mathbf{B}(r))^+ \mathbf{B}(r)^\top \tilde{\mathcal{O}}_r. \quad (3.1.10)$$

Passando à estimação do vetor  $\underline{\beta}$ , dos coeficientes de efeitos fixos, tomando-se

$$\underline{Y}_\bullet = \frac{1}{2}(\underline{Y}(1) + \underline{Y}(2)), \quad (3.1.11)$$

e  $\mathbf{M}_i = \mathbf{X}_i \mathbf{X}_i^\top$ ,  $i = 1, \dots, w$ , tem-se

$$\mathcal{E}(\underline{Y}_\bullet) = \frac{1}{2} \sum_{i=1}^w \chi_{2,i} \mathbf{M}_i, \quad (3.1.12)$$

bem como

$$\tilde{\mathcal{E}}(\underline{Y}_\bullet) = \frac{1}{2} \sum_{i=1}^w \tilde{\chi}_{2,i} \mathbf{M}_i. \quad (3.1.13)$$

Ora

$$E(\underline{Y}_\bullet) = \mathbf{X} \underline{\beta}, \quad (3.1.14)$$

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

tendo-se pois, Kantar, Y. M. (2015) [30] e Kariya, T. & Kurata, H. (2004, págs. 25-40) [31], o  $GLSE$  do vetor de coeficientes,  $\underline{\beta}$ , é dado por

$$\underline{\tilde{\beta}} = (\mathbf{X}^\top \tilde{\Sigma}(\underline{Y}_\bullet)^+ \mathbf{X})^+ \mathbf{X}^\top \tilde{\Sigma}(\underline{Y}_\bullet)^+ \underline{Y}_\bullet. \quad (3.1.15)$$

Observe-se que, se  $\tilde{\Sigma}(\underline{Y}_\bullet)$  for invertível e  $\mathbf{X}$  for  $n \times k$ , com característica  $k$ ,  $\mathbf{X}^\top \tilde{\Sigma}(\underline{Y}_\bullet)^{-1} \mathbf{X}$  é invertível, verificando-se isto sempre que  $\mathbf{X}_w = \mathbf{I}_n$ , vindo  $M_w = \mathbf{I}_n$ , e  $\tilde{\sigma}_w^2 > 0$ . Tem-se então

$$\underline{\tilde{\beta}} = (\mathbf{X}^\top \tilde{\Sigma}(\underline{Y}_\bullet)^{-1} \mathbf{X})^{-1} \mathbf{X}^\top \tilde{\Sigma}(\underline{Y}_\bullet)^{-1} \underline{Y}_\bullet, \quad (3.1.16)$$

bem como,

$$\begin{aligned} \tilde{\Sigma}(\underline{\tilde{\beta}}) &= (\mathbf{X}^\top \tilde{\Sigma}(\underline{Y}_\bullet)^{-1} \mathbf{X})^{-1} \mathbf{X}^\top \tilde{\Sigma}(\underline{Y}_\bullet)^{-1} \tilde{\Sigma}(\underline{Y}_\bullet) \tilde{\Sigma}(\underline{Y}_\bullet)^{-1} \mathbf{X} (\mathbf{X}^\top \tilde{\Sigma}(\underline{Y}_\bullet)^{-1} \mathbf{X}) \\ &= (\mathbf{X}^\top \tilde{\Sigma}(\underline{Y}_\bullet)^{-1} \mathbf{X})^{-1} \mathbf{X}^\top \tilde{\Sigma}(\underline{Y}_\bullet)^{-1} \mathbf{X} (\mathbf{X}^\top \tilde{\Sigma}(\underline{Y}_\bullet) \mathbf{X})^{-1} \\ &= (\mathbf{X}^\top \tilde{\Sigma}(\underline{Y}_\bullet)^{-1} \mathbf{X})^{-1}. \end{aligned}$$

Estabelecemos assim a

**Proposição 3.1** Se  $\mathbf{X}$  for  $n \times k$ , com característica  $k$  e  $\tilde{\Sigma}(\underline{Y}_\bullet)$  for invertível, temos

$$\tilde{\Sigma}(\underline{\tilde{\beta}}) = (\mathbf{X}^\top \tilde{\Sigma}(\underline{Y}_\bullet)^{-1} \mathbf{X})^{-1}.$$

*Nota* : Este enunciado mostra que  $\tilde{\Sigma}(\underline{\tilde{\beta}})$  é invertível, nas condições da tese.

**Corolário 3.1** Verificadas as condições da Proposição 3.1, para o vetor estimável  $\underline{\psi} = \mathbf{G}\underline{\beta}$  tem-se o estimador  $\underline{\tilde{\psi}} = \mathbf{G}\underline{\tilde{\beta}}$ , com matriz de covariância estimada  $\tilde{\Sigma}(\underline{\tilde{\psi}}) = \mathbf{G}(\mathbf{X}^\top \tilde{\Sigma}(\underline{Y}_\bullet)^{-1} \mathbf{X})^{-1} \mathbf{G}^\top$ . Se  $\mathbf{G}$  for uma matriz  $h \times k$ , com característica  $h$ ,  $\tilde{\Sigma}(\underline{\tilde{\psi}})$  será invertível.

**Dem:** A primeira parte da tese é de demonstração direta. Raciocinando-se como na secção 2.1.7, mostramos que existe  $\tilde{\Sigma}(\underline{\tilde{\beta}})^{\frac{1}{2}}$ , pelo que

$$\tilde{\Sigma}(\underline{\tilde{\psi}}) = \mathbf{G}\tilde{\Sigma}(\underline{\tilde{\beta}})\mathbf{G}^\top = (\mathbf{G}\tilde{\Sigma}(\underline{\tilde{\beta}})^{\frac{1}{2}})(\mathbf{G}\tilde{\Sigma}(\underline{\tilde{\beta}})^{\frac{1}{2}})^\top \quad (3.1.17)$$

vindo, ver Silvey (1980) [65],

$$\text{car}(\tilde{\Sigma}(\underline{\tilde{\psi}})) = \text{car}(\mathbf{G}\tilde{\Sigma}(\underline{\tilde{\beta}})^{\frac{1}{2}}) = \text{car}(\mathbf{G}) = h, \quad (3.1.18)$$

atendendo a  $\tilde{\Sigma}(\underline{\tilde{\beta}})^{\frac{1}{2}}$  ser invertível sempre que  $\tilde{\Sigma}(\underline{\tilde{\beta}})$  o é, visto os seus valores próprios serem as raízes quadradas dos de  $\tilde{\Sigma}(\underline{\tilde{\beta}})$  e, portanto, ambas as matrizes terem o mesmo número de valores próprios positivos. Como  $\tilde{\Sigma}(\underline{\tilde{\psi}})$  é  $h \times h$ , com característica  $h$ , logo será invertível. ■

## 3.2 Elipsóides de Confiança

Nesta secção mostraremos a utilização dos  $\tilde{\chi}_{r,i}$ ,  $i = 1, \dots, w$ ,  $r = 2, 3, 4$ , na construção dos Elipsóides de Confiança, de nível  $1 - p$ , que terão (aproximadamente) probabilidade  $p$  de conter

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

qualquer realização de um vetor com a distribuição do vetor das observações. Para isso, dado o vetor  $\underline{u}$ , com  $\|\underline{u}\| = 1$ , tomemos

$$W(\underline{u}) = \underline{u}^\top \left( \left( \frac{Y(1) + Y(2)}{2} \right) - \underline{\mu} \right) = \sum_{i=1}^w \underline{a}_i^\top Z_i, \quad (3.2.19)$$

com

$$\underline{a}_i^\top = \underline{u}^\top \mathbf{X}_i = (a_{i,1}, \dots, a_{i,c_i}), i = 1, \dots, w. \quad (3.2.20)$$

Como

$$E(Z_i) = \underline{0}_{c_i}, i = 1, \dots, w, \quad (3.2.21)$$

temos

$$E(W(\underline{u})) = 0. \quad (3.2.22)$$

Sendo  $v_r(\underline{u})$  o  $r$ -ésimo cumulante de  $W(\underline{u})$ , podemos escrever

$$v_r(\underline{u}) = \frac{1}{2} \sum_{i=1}^w d_i(r, \underline{u}) \chi_{r,i}, r = 2, 3, 4, \quad (3.2.23)$$

com

$$d_i(r, \underline{u}) = \sum_{l=1}^{c_i} a_{i,l}^r, i = 1, \dots, w, r = 2, 3, 4. \quad (3.2.24)$$

Teremos pois os estimadores do  $r$ -ésimo cumulante de  $W(\underline{u})$ , do coeficiente de assimetria e do coeficiente de curtose, dados por

$$\left\{ \begin{array}{l} \tilde{v}_r(\underline{u}) = \sum_{i=1}^w d_i(r, \underline{u}) \tilde{\chi}_{r,i}, r = 2, 3, 4, \\ \tilde{\gamma}_1(\underline{u}) = \frac{\tilde{v}_3(\underline{u})}{\tilde{v}_2(\underline{u})^{\frac{3}{2}}} \\ \tilde{\gamma}_2(\underline{u}) = \frac{\tilde{v}_4(\underline{u})}{\tilde{v}_2(\underline{u})^2} - 3, \end{array} \right. , \quad (3.2.25)$$

pelo que, para a distribuição  $F_{\dot{W}(\underline{u})}$  de  $\dot{W}(\underline{u})$ , temos, de acordo com DasGupta, A. (2008, pág. 187) [13], Edgeworth, F. (1905, 1906) [15] [16] e Kolassa, J. E. (1994, pág. 65) [35],

$$\dot{W}(\underline{u}) = \frac{W(\underline{u})}{\sqrt{\tilde{v}_2(\underline{u})}}, \quad (3.2.26)$$

e a Expansão de Edgeworth, segundo Kendall, M. G., Stuart, A. & Ord, J. K. (1987, §3)

$$Ed_{\dot{W}(\underline{u})}(w) = N(w) - \frac{\tilde{\gamma}_1(\underline{u})}{3!} n^{<2>}(w) + \frac{\tilde{\gamma}_2(\underline{u})}{4!} n^{<3>}(w) + 10 \frac{\tilde{\gamma}_1^2(\underline{u})}{6!} n^{<6>}(w), i = 1, \dots, w, \quad (3.2.27)$$

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

o que nos permite obter o  $p$ -ésimo quantil (estimado),  $\tilde{w}_p(\underline{u})$ , de  $Ed_{\tilde{W}(\underline{u})}(\cdot)$ .

Assim, teremos os quantis estimados

$$\tilde{w}_p(\underline{u}) = \sqrt{\tilde{v}_2(\underline{u})} \tilde{w}_p(\underline{u}), \quad (3.2.28)$$

para  $W(\underline{u})$ . Vamos agora obter o elipsóide centrado na origem, associado à forma quadrática  $\underline{u}_l^\top \mathbf{K} \underline{u}_l$ , que, dados os vetores  $\underline{u}_1, \dots, \underline{u}_m$ , minimiza

$$\Delta(\mathbf{K}) = \sum_{l=1}^m (\tilde{w}_p(\underline{u}_l) - \underline{u}_l^\top \mathbf{K} \underline{u}_l)^2. \quad (3.2.29)$$

Ora, com

$$\begin{cases} \underline{u}^{[2l]} = (u_1^2, \dots, 2u_1 u_n, u_2^2, \dots, 2u_2 u_n, \dots, u_n^2) \\ \underline{s}(\mathbf{K}) = (k_{1,1}, \dots, k_{1,n}, k_{2,2}, \dots, k_{2,n}, \dots, k_{n,n}) \end{cases}, \quad (3.2.30)$$

temos

$$\underline{u}_l^\top \mathbf{K} \underline{u}_l = \underline{u}_l^{[2l]\top} \underline{s}(\mathbf{K}), \quad l = 1, \dots, m. \quad (3.2.31)$$

Observe-se que  $\underline{u}^{[2l]}$  e  $\underline{s}(\mathbf{K})$  têm  $\frac{n(n+1)}{2}$  componentes. Sendo  $\{\underline{u}_1, \dots, \underline{u}_n\}$  uma base ortonormada para  $\mathbb{R}^n$ , tomemos

$$\begin{cases} \mathbf{U}^{[2l]} = [\underline{u}_1^{[2l]}, \dots, \underline{u}_m^{[2l]}] \\ \tilde{\underline{w}}_p(\underline{u}) = (\tilde{w}_p(\underline{u}_1), \dots, \tilde{w}_p(\underline{u}_m)) \end{cases}, \quad (3.2.32)$$

onde pretendemos minimizar  $\|\tilde{\underline{w}}_p(\underline{u}) - \mathbf{U}^{[2l]} \underline{\lambda}\|$ , considerando  $\underline{\lambda}$  o semi-vec da matriz  $\mathbf{K}$ . Assim,  $\underline{\lambda} = \underline{s}(\mathbf{K})$ , tendo-se  $\underline{s}(\mathbf{K}(\underline{\lambda})) = \underline{\lambda}$ . Dado ter-se

$$E(\tilde{\underline{w}}_p) = \mathbf{U}^{[2l]} \underline{\lambda}_p, \quad (3.2.33)$$

teremos o *LSE*

$$\begin{cases} \tilde{\underline{\lambda}}_p = (\mathbf{U}^{[2l]\top} \mathbf{U}^{[2l]})^{-1} \mathbf{U}^{[2l]\top} \tilde{\underline{w}}_p(\underline{u}) \\ \tilde{\mathbf{K}}_p = \mathbf{K}(\tilde{\underline{\lambda}}_p) \end{cases}, \quad (3.2.34)$$

onde  $\tilde{\mathbf{K}}_p$  é tal que  $\underline{s}(\tilde{\mathbf{K}}_p) = \tilde{\underline{\lambda}}_p$ . Observe-se que, dado um vetor  $\underline{u}$ , com  $\|\underline{u}\| = 1$ , o elipsóide passa pela direção dada por  $\underline{u}$ , à distância  $\underline{u}^\top \tilde{\mathbf{K}}_p \underline{u}$  do centro. Se se quiser centrar o mesmo no "ponto médio"  $\tilde{\underline{\mu}} = \mathbf{X} \tilde{\underline{\beta}}$ , o contorno passará pelo ponto da semi-reta com direção  $\underline{u}$ , à distância  $\underline{u}^\top \tilde{\mathbf{K}}_p \underline{u}$ , de  $\tilde{\underline{\mu}}$ .

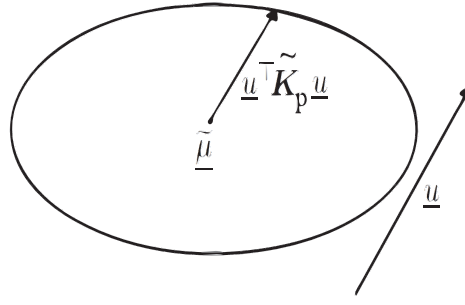


Figura 3.1: Elipsóide que passa pela direção dada por  $\underline{u}$ , à distância  $\underline{u}^T \tilde{K}_p \underline{u}$  do centro.

### 3.3 Intervalos de Predição

A abordagem usual passa por obter uma predição pontual fornecendo um BLUP, com variância mínima para uma futura observação. Complementamos as técnicas usadas geralmente na literatura para apresentar uma alternativa em que se obtém um Intervalo de Predição,  $[Y_{\frac{p}{2}}; Y_{1-\frac{p}{2}}]$ , com uma probabilidade (aproximada),  $1 - p$ , de conter a futura observação.

A predição será para um conjunto de vetores adicionais, digamos  $\dot{\underline{x}}, \dot{\underline{x}}_1, \dots, \dot{\underline{x}}_w$  nas matrizes  $\mathbf{X}, \mathbf{X}_1, \dots, \mathbf{X}_w$ .

Sejam então  $\dot{\underline{x}}, \dot{\underline{x}}_1, \dots, \dot{\underline{x}}_w$  esses vetores, com  $k, c_1, \dots, c_w$ , componentes, que serão os valores das variáveis controladas, para as quais se pretende prever o valor duma futura observação, digamos  $\dot{Y}_f$ . Observe-se que  $E(\dot{Y}_f)$  pode ser estimado por

$$\tilde{E}(\dot{Y}_f) = \dot{\underline{x}}^T \tilde{\underline{\beta}}, \quad (3.3.35)$$

e, segundo Christensen, R. (2011, págs. 33-41) [10], podemos obter o *BLUP* correspondente a

$$\dot{\underline{u}} = [\dot{\underline{x}}^T \dot{\underline{x}}_1^T \dots \dot{\underline{x}}_w^T]^T. \quad (3.3.36)$$

Assim, utilizando os resultados da secção anterior, 3.2, podemos obter para

$$\ddot{\underline{u}} = \frac{1}{\|\dot{\underline{u}}\|} \dot{\underline{u}}, \quad (3.3.37)$$

os estimadores dos  $r$ -ésimos cumulantes de  $W(\underline{u})$ , dados por  $\tilde{v}_r(\ddot{\underline{u}})$ ,  $r = 2, 3, 4$ , e os estimadores dos coeficientes de assimetria e de curtose, dados por  $\tilde{\gamma}_1(\ddot{\underline{u}})$  e  $\tilde{\gamma}_2(\ddot{\underline{u}})$ . Assim, obtemos a Expansão de Edgeworth,  $Ed_{\tilde{W}(\ddot{\underline{u}})}(\cdot)$ , para  $F_{\tilde{W}(\ddot{\underline{u}})}(\cdot)$ , e, a partir daí, obtemos os intervalos de predição seguintes, de nível  $1 - p$ ,

$$[\dot{\underline{x}}^T \tilde{\underline{\beta}} + \tilde{w}_{\frac{p}{2}}(\ddot{\underline{u}}); \dot{\underline{x}}^T \tilde{\underline{\beta}} + \tilde{w}_{1-\frac{p}{2}}(\ddot{\underline{u}})]. \quad (3.3.38)$$

Observe-se que  $1 - p$  é apenas um valor aproximado da probabilidade desse intervalo conter a futura observação.

### 3.4 Modelos Múltiplos

Nesta secção fazemos um estudo sobre modelos múltiplos. Estes são da forma

$$\underline{Y} = [\underline{Y}(1)^\top, \dots, \underline{Y}(u)^\top]^\top, \quad (3.4.39)$$

com  $\underline{Y}(l) = \mathbf{X}\underline{\beta}(l) + \underline{Y}^o(l)$ ,  $l = 1, \dots, u$ , sendo os  $\underline{Y}^o(1), \dots, \underline{Y}^o(u)$  *i.i.d.*, com vetores médios nulos. Estes modelos generalizam os modelos múltiplos em que  $\Sigma(\underline{Y}^o(l)) = \sigma^2 \mathbf{I}_n$ ,  $l = 1, \dots, u$ , ou seja, faremos uma generalização dos modelos regressionais múltiplos, em que se considerava um único termo aleatório da forma  $\mathbf{X}\underline{e}$ , com  $\underline{e} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_n)$ , ver, respetivamente, Mexia, J. T. (1987), Moreira et al. (2005) e Ribeiro, A. & Mexia, J. T. (1997) ([43], [47] e [57]).

Mostraremos que as hipóteses sobre efeitos e interações dos factores de um delineamento base correspondem aos espaços numa partição ortogonal de  $\mathbb{R}^d$ , admitindo que as matrizes  $\mathbf{X}, \mathbf{X}_1, \dots, \mathbf{X}_w$ , são as mesmas para todos os modelos associados aos diferentes tratamentos, bem como os cumulantes  $\chi_{r,i}$ ,  $r = 2, 3, 4$ ,  $i = 1, \dots, w$ . Temos então que apenas os vetores  $\underline{\beta}$  variarão de modelo individual para modelo individual. Estudaremos, então, a ação dos factores do delineamento base nas funções estimáveis  $\underline{e}^\top \underline{\beta}$ , obtendo testes para as hipóteses de ausência de efeitos e/ou interações.

Façamos aqui uma breve referência às funções estimáveis. Quando estamos interessados em estimar funções de  $\underline{\beta}$ , devemos distinguir entre aquelas que são funções apenas de  $\mathbf{X}\underline{\beta}$  e aquelas que não são. Dado que o parâmetro  $\underline{\beta}$  afeta a distribuição de  $\underline{Y}$  apenas através de  $\mathbf{X}\underline{\beta}$ , devemos distinguir entre aquelas que são funções apenas de  $\mathbf{X}\underline{\beta}$ , pois são as que podem ser estimadas satisfatoriamente, McCulloch, C. & Searle, S. (2001, pág. 120) [42]. Assim, apresentaremos os instrumentos algébricos necessários para testar essas hipóteses, permitindo realizar a inferência para modelos múltiplos, clarificando a ligação entre este estudo e um trabalho já publicado, Antunes et al. (2020b) [3].

Consideremos um delineamento base de efeitos fixos com  $d$  tratamentos, associado a uma partição ortogonal  $\mathbb{R}^d = \boxplus_{j=1}^m \nabla_j$ . Sejam  $\mathbf{A}_1, \dots, \mathbf{A}_m$  matrizes cujos  $g_1, \dots, g_m$  vetores linha constituem bases ortonormadas para os sub-espacos  $\nabla_1, \dots, \nabla_m$ .

Por exemplo, se no delineamento base cruzarmos factores com  $f_1, \dots, f_v$  níveis, os sub-espacos  $\nabla_1, \dots, \nabla_m$  corresponderão aos  $m = 2^v$  sub-conjuntos  $\varphi$  de índices dos factores, dado por  $\bar{v} = \{1, \dots, v\}$ . Se  $\sharp(\varphi) = 0$ , estar-lhe-á associado o valor médio geral; se  $\sharp(\varphi) = 1$ , os efeitos dos níveis do único factor com índice em  $\varphi$  e, se  $\sharp(\varphi) > 1$ , as interações entre os factores com índices em  $\varphi$ . Podemos dar aos conjuntos  $\varphi$  e aos sub-espacos correspondentes os índices

$$j(\varphi) = 1 + \sum_{l \in \varphi} 2^{l-1}, \quad (3.4.40)$$

sendo  $j(\varphi)$ ,  $\nabla_j$  e  $\mathbf{A}_j$ ,  $j = 1, \dots, m$ , os conjuntos, sub-espacos e matrizes com os vários índices. Ao conjunto  $\varphi_j$ , com índice  $j$ , está associada a matriz, ver Ferreira et al (2010) [18],

$$\mathbf{A}(\varphi) = \otimes_{l=1}^v \mathbf{A}_l(\varphi), \quad (3.4.41)$$

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

sendo

$$\mathbf{A}_i(\varphi) = \begin{cases} \mathbf{1}_{f_i}^\top; & l \notin \varphi_j \\ \mathbf{L}_{f_i}; & l \in \varphi_j \end{cases}, \quad (3.4.42)$$

onde  $\mathbf{L}_s$  é obtida através da eliminação da primeira linha, igual a  $\frac{1}{\sqrt{f_i}}\mathbf{1}_{f_i}^\top$ , de uma matriz ortogonal. A matriz  $\mathbf{A}(\varphi)$  tem característica  $g_\varphi = \prod_{l \in \varphi} (f_l - 1)$ .

A nossa análise centrar-se-á sobre funções estimáveis.

Sejam  $(\tilde{\psi}_h(1), \tilde{\psi}_h(2)), h = 1, \dots, d$ , os pares de estimadores que se têm para uma dessas funções. Se admitirmos que os  $\underline{Y}_1(1); \underline{Y}_1(2), \dots, \underline{Y}_d(1); \underline{Y}_d(2)$  são independentes, e os vetores de cada par são *i.i.d.*, os  $\tilde{\psi}_1(1); \tilde{\psi}_1(2), \dots, \tilde{\psi}_d(1); \tilde{\psi}_d(2)$  serão independentes e os estimadores de cada par são *i.i.d.*.

Atendendo aos resultados da secção 2.1.7, teremos

$$\mathcal{E}(\tilde{\beta}_h(l)) = \dot{\mathbf{V}}, \quad l = 1, 2, \quad h = 1, \dots, d, \quad (3.4.43)$$

e, tendo-se

$$\tilde{\psi}_h(l) = \underline{c}^\top \tilde{\beta}_h(l), \quad l = 1, 2, \quad h = 1, \dots, d, \quad (3.4.44)$$

ter-se-á

$$\mathcal{E}(\tilde{\psi}_h(l)) = \underline{c}^\top \dot{\mathbf{V}} \underline{c}, \quad l = 1, 2, \quad h = 1, \dots, d, \quad (3.4.45)$$

pondo-se, para aligeirar a escrita,

$$\sigma^2(\underline{c}) = \underline{c}^\top \dot{\mathbf{V}} \underline{c}. \quad (3.4.46)$$

Então, com

$$\begin{cases} \tilde{\underline{\psi}}(l) = (\tilde{\psi}_1(l), \dots, \tilde{\psi}_d(l)) \\ \underline{\psi}(l) = (\psi_1(l), \dots, \psi_d(l)) \end{cases}, \quad l = 1, 2, \quad (3.4.47)$$

os  $\tilde{\underline{\psi}}(1)$  e  $\tilde{\underline{\psi}}(2)$  serão *i.i.d.* com vetor médio  $\underline{\psi}(1) = \underline{\psi}(2) = \underline{\psi}_\bullet$  e matriz de covariância  $\sigma^2(\underline{c})\mathbf{I}_d$ . Assim, o par  $(\tilde{\underline{\psi}}(1), \tilde{\underline{\psi}}(2))$  corresponderá a um modelo homocedástico equilibrado com  $d$  tratamentos e duas observações por tratamento. Atendendo à discussão do capítulo X de Scheffé, H. (1959) [59], concluímos que a Análise de variância, *ANOVA*, é aplicável a este modelo. Tomemos

$$\begin{cases} S = \sum_{l=1}^d (\tilde{\psi}_{l,1} - \tilde{\psi}_{l,2})^2 \\ \tilde{\underline{\psi}}_{l,\bullet} = \frac{1}{2} (\tilde{\psi}_{l,1} + \tilde{\psi}_{l,2}), \quad l = 1, \dots, d, \end{cases}, \quad (3.4.48)$$

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

e

$$\begin{cases} \tilde{\underline{\psi}}_{l,\bullet} = (\tilde{\psi}_{1,\bullet}, \dots, \tilde{\psi}_{d,\bullet}) \\ \underline{\psi}_{l,\bullet} = (\psi_{1,\bullet}, \dots, \psi_{d,\bullet}) \end{cases}, \quad (3.4.49)$$

Então, com

$$\begin{cases} S_j = \|\mathbf{A}_j \tilde{\underline{\psi}}_{l,\bullet}\|^2 \\ g_j = \text{car}(\mathbf{A}_j) \end{cases}, j = 1, \dots, m, \quad (3.4.50)$$

temos a estatística

$$F_j = \frac{2d}{g_j} \frac{S_j}{S}, \quad j = 1, \dots, m, \quad (3.4.51)$$

com  $g_j$  e  $2d$  graus de liberdade e parâmetros de não centralidade

$$\delta_j : \frac{2}{\sigma^2(\underline{c})} \|\mathbf{A}_j \underline{\psi}_{l,\bullet}\|^2, j = 1, \dots, m, \quad (3.4.52)$$

para testar as hipóteses

$$H_{0,j} : \mathbf{A}_j \underline{\psi}_{l,\bullet} = \mathbf{0}_{g_j}, j = 1, \dots, m. \quad (3.4.53)$$

Estes parâmetros de não centralidade só se anulam quando  $H_{0,j}$ ,  $j = 1, \dots, m$ , se verificam. Assim, estes testes são não distorcidos. O factor 2 que aparece nestas últimas expressões, resulta de

$$\Sigma(\tilde{\underline{\psi}}_{l,\bullet}) = \frac{\sigma^2(\underline{c})}{2} \mathbf{I}_d. \quad (3.4.54)$$



## Capítulo 4

### Modelos Aditivos com Parâmetros de Localização e Dispersão

#### 4.1 Ajustamentos

Nesta secção consideramos o caso em que as distribuições das componentes dos vetores  $\underline{Z}_i$ ,  $i = 1, \dots, w$ , têm parâmetros de localização  $\lambda_i$ ,  $i = 1, \dots, w$ , e dispersão  $\delta_i$ ,  $i = 1, \dots, w$ . Aquando do ajustamento, integraremos os parâmetros de localização  $\lambda_i$ ,  $i = 1, \dots, w$ , no vetor dos coeficientes  $\underline{\beta}$ . Assim, quando estimarmos  $\underline{\beta}$  estimamos  $\lambda_1, \dots, \lambda_w$ , usando técnicas algébricas plenamente conhecidas mas agora aplicadas num contexto diferente. Por outro lado, a partir dos  $(\tilde{\chi}_{2,i}, \tilde{\chi}_{3,i})$ ,  $i = 1, \dots, w$ , obtemos os  $\tilde{\delta}_i$ ,  $i = 1, \dots, w$ , e, destes, os estimadores dos cumulantes de quarta ordem, dados por  $\tilde{\chi}_{4,i}$ ,  $i = 1, \dots, w$ . Seguem-se as secções 4.2, 4.3 e 4.4, onde também iremos obter, de forma similar ao capítulo anterior, elipsóides de Confiança, intervalos de predição e modelos múltiplos. Particularizamos para os casos em que as componentes  $Z_{i,l}$ ,  $l = 1, \dots, c_i$ ,  $i = 1, \dots, w$  têm distribuições Normal, Exponencial e Gumbel. Tal como no capítulo 3, consideramos o modelo

$$\underline{\dot{Y}} = \dot{\mathbf{X}} \underline{\dot{\beta}} + \sum_{i=1}^w \dot{\mathbf{X}}_i \underline{\dot{Z}}_i, \quad (4.1.1)$$

onde o vetor de observações  $\underline{\dot{Y}}$  é do tipo  $n \times 1$ , os vetores  $\underline{\dot{Z}}_1, \dots, \underline{\dot{Z}}_w$ , são independentes, com componentes *i.i.d.*, tendo estas componentes pares de localização e dispersão  $(\lambda_i, \delta_i)$ ,  $i = 1, \dots, w$ . O modelo pode agora ser reescrito como

$$\underline{Y} = \mathbf{X} \underline{\beta} + \sum_{i=1}^w \mathbf{X}_i \underline{Z}_i, \quad (4.1.2)$$

com

$$\begin{cases} \mathbf{X} = [\dot{\mathbf{X}} \ \mathbf{X}_1 \mathbf{1}_{c_1} \ \dots \ \mathbf{X}_w \mathbf{1}_{c_w}] \\ \underline{\beta} = [\underline{\dot{\beta}}^\top \ \lambda_1 \ \dots \ \lambda_w]^\top \end{cases}, \quad (4.1.3)$$

e vetores  $\underline{Z}_i$ ,  $i = 1, \dots, w$ , independentes, com componentes *i.i.d.* e com parâmetros de dispersão  $\delta_i$ ,  $i = 1, \dots, w$ . Temos ainda

$$\chi_{r,i}(|0, \delta_i) = \delta_i^r \chi_{r,i}(|0, 1), \quad i = 1, \dots, w, r = 2, 3, 4, \quad (4.1.4)$$

bem como

$$\tilde{\mathcal{O}}_r = (\dot{Y}_1^r, \dots, \dot{Y}_n^r), \quad r = 2, 3, \quad (4.1.5)$$

e com

$$\tilde{\chi}_r = (\mathbf{B}(r)^\top \mathbf{B}(r))^+ \mathbf{B}(r)^\top \tilde{\mathcal{Q}}_r, \quad r = 2, 3. \quad (4.1.6)$$

Tem-se ainda

$$\tilde{\delta}_i = \sqrt{\frac{\chi_{2,i}(|0, \delta_i|)}{\chi_{2,i}(|0, 1|)}}, \quad i = 1, \dots, w, \quad (4.1.7)$$

e

$$\tilde{\chi}_{4,i}(|0, \delta_i|) = \tilde{\delta}_i^4 \chi_{4,i}(|0, 1|), \quad (4.1.8)$$

onde os  $\chi_{2,i}(|0, 1|)$  e os  $\chi_{4,i}(|0, 1|)$  são conhecidos. Por outro lado temos, tal como vimos atrás,

$$\tilde{\Sigma}(\underline{Y}) = \sum_{i=1}^w \tilde{\chi}_{2,i}(|0, \delta_i|) \mathbf{M}_i, \quad (4.1.9)$$

e o *GLSE*,

$$\tilde{\beta} = (\mathbf{X}^\top \tilde{\Sigma}(\underline{Y})^+ \mathbf{X})^+ \mathbf{X}^\top \tilde{\Sigma}(\underline{Y})^+ \underline{Y}. \quad (4.1.10)$$

É ainda fácil de estabelecer, tal como na Proposição 3.1. e respectivo Corolário, do capítulo 3, que, quando  $\tilde{\Sigma}(\underline{Y})$  for invertível e os vetores coluna de  $\mathbf{X}$  forem linearmente independentes, teremos

$$\begin{cases} \tilde{\beta} = (\mathbf{X}^\top \tilde{\Sigma}(\underline{Y})^{-1} \mathbf{X})^{-1} \mathbf{X}^\top \tilde{\Sigma}(\underline{Y})^{-1} \underline{Y} \\ \tilde{\Sigma}(\tilde{\beta}) = (\mathbf{X}^\top \tilde{\Sigma}(\underline{Y})^{-1} \mathbf{X})^{-1} \end{cases}. \quad (4.1.11)$$

## 4.2 Elipsóides de Confiança

Mantendo-se a expressão dos  $d_i(r, \underline{u})$  da secção 3.3 e do  $r$ -ésimo cumulante de  $W(\underline{u})$ ,  $r = 2, 3, 4$ , voltamos a obter

$$\begin{cases} \tilde{v}_r(\underline{u}) = \sum_{i=1}^w d_i(r, \underline{u}) \tilde{\chi}_{r,i}(|0, \delta_i|), \quad r = 2, 3, 4, \\ \tilde{\gamma}_1(\underline{u}) = \frac{\tilde{v}_3(\underline{u})}{\tilde{v}_2(\underline{u})^{\frac{3}{2}}} \\ \tilde{\gamma}_2(\underline{u}) = \frac{\tilde{v}_4(\underline{u})}{\tilde{v}_2(\underline{u})^2} - 3, \end{cases}, \quad (4.2.1)$$

obtendo-se, como vimos atrás, a Expansão de Edgeworth  $Ed_{\tilde{W}(\underline{u})}(w)$  e os quantis estimados  $\tilde{w}_p(\underline{u})$  e

$$\tilde{w}_p(\underline{u}) = \sqrt{\tilde{v}_2(\underline{u})} \quad \tilde{w}_p(\underline{u}). \quad (4.2.2)$$

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

Temos agora, dados os  $\underline{u}_1, \dots, \underline{u}_m$ , que minimizar

$$\Delta(\mathbf{K}) = \sum_{l=1}^m (\tilde{w}_p(\underline{u}_l) - \underline{u}_l^\top \mathbf{K} \underline{u}_l)^2, \quad (4.2.3)$$

obtendo-se para  $\underline{\lambda} = \underline{s}(\mathbf{K})$ , o estimador

$$\tilde{\underline{\lambda}}_p = (U^{12\top} U^{12})^{-1} U^{12\top} \tilde{\underline{w}}_p, \quad (4.2.4)$$

com

$$\tilde{\underline{w}}_p = (\tilde{w}_p(u_1), \dots, \tilde{w}_p(u_m)), \quad (4.2.5)$$

e, a partir daqui, de novo  $\tilde{\mathbf{K}}_p = \mathbf{K}(\tilde{\underline{\lambda}}_p)$ , mantendo-se a interpretação geométrica da Figura 3.1. Tal como na secção 3.3, trabalhamos com um modelo, em vez de um par de modelos.

### 4.3 Intervalos de Predição

Para obter Intervalos de Predição, tal como na secção 3.3, do capítulo 3, temos

$$\underline{\dot{u}} = [\dot{x}^\top \dot{x}_1^\top \dots \dot{x}_w^\top], \quad (4.3.1)$$

e obtemos  $Ed_{\tilde{W}(\underline{u})}(\ddot{w})$  para  $\underline{\ddot{u}} = \frac{1}{\|\underline{\dot{u}}\|} \underline{\dot{u}}$ , obtendo primeiro os os quantis  $\tilde{w}_p(\underline{\ddot{u}})$  e  $\tilde{w}_p(\underline{\ddot{u}})$ . A partir daqui procede-se como vimos atrás, para obter os intervalos de predição de nível  $1 - p$  (aproximado),

$$[\dot{x}^\top \tilde{\underline{\beta}} + \tilde{w}_{\frac{p}{2}}(\underline{\ddot{u}}); \dot{x}^\top \tilde{\underline{\beta}} + \tilde{w}_{1-\frac{p}{2}}(\underline{\ddot{u}})].$$

### 4.4 Modelos Múltiplos

Podemos repetir o tratamento como na secção 3.4, do capítulo 3, para dados

$$\begin{cases} \tilde{\underline{\psi}} = (\tilde{\psi}_1, \dots, \tilde{\psi}_d) \\ \underline{\psi} = (\psi_1, \dots, \psi_d) \end{cases}, \quad (4.4.1)$$

obter primeiro as  $S_j, j = 1, \dots, m$  e  $S = \sum_{j \in D} S_j$ , e, em seguida, as estatísticas

$$\mathcal{F}_j = \frac{g}{g_j} \frac{S_j}{S}, \quad (4.4.2)$$

mantendo-se  $g$  e  $g_j, j \notin D$ , tal como vimos atrás, para testar as hipóteses  $H_{0,j}, j \notin D$ .

## 4.5 Casos Particulares

Seguidamente vamos apresentar a estimação dos parâmetros de localização e dispersão dos casos em que as componentes  $Z_{i,l}$ ,  $l = 1, \dots, c_i$ ,  $i = 1, \dots, w$  têm distribuições Normal, Exponencial e Gumbel.

### 4.5.1 Tipo Normal

Temos agora, sendo  $\mu$  o valor médio e  $\sigma^2$  a variância,

$$\left\{ \begin{array}{l} \chi_1(\mu, \sigma^2) = \mu \\ \chi_2(\mu, \sigma^2) = \sigma^2 \\ \vdots \\ \chi_r(\mu, \sigma^2) = 0, \quad r > 2 \end{array} \right. , \quad (4.5.1)$$

bem como

$$\left\{ \begin{array}{l} \chi_1(0, 1) = \mu = \lambda \\ \chi_2(0, 1) = \sigma^2 \\ \chi_r(0, 1) = 0, \quad r > 2 \end{array} \right. , \quad (4.5.2)$$

não sendo necessário estimar os  $\chi_r(\mu, \sigma^2)$  para  $r > 2$ . Para estimar os  $\chi_1(\mu, \sigma^2)$ , incluímos  $\mu = \lambda$  no vetor dos coeficientes e obtemos *LSE* para  $\chi_1(\mu, \sigma^2)$  e igualmente para  $\chi_2(\mu, \sigma^2)$ , utilizando

$$\tilde{\chi}_2 = (\mathbf{B}(2)^\top \mathbf{B}(2))^{-1} \mathbf{B}(2)^\top \tilde{\mathcal{O}}_2, \quad (4.5.3)$$

com

$$\tilde{\mathcal{O}}_2 = (\dot{Y}_1^2, \dots, \dot{Y}_n^2). \quad (4.5.4)$$

### 4.5.2 Tipo Exponencial

Temos a densidade estandardizada da distribuição exponencial dada por

$$e(x|0, 1) = e^{-x}, \quad x \geq 0, \quad (4.5.5)$$

vindo

$$\varphi(t|0, 1) = \int_0^{+\infty} e^{-(1-t)x} dx = \frac{1}{1-t}, \quad t < 1, \quad (4.5.6)$$

bem como

$$\psi(t|0, 1) = -\log(1-t), \quad (4.5.7)$$

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

e, com

$$\psi^{<r>}(t|0, 1) = \frac{r!}{(1-t)^r}, \quad r \geq 1, \quad (4.5.8)$$

logo

$$\chi_r(|0, 1) = r! \quad r \geq 1. \quad (4.5.9)$$

### 4.5.3 Tipo Gumbel

Se as distribuições das componentes são distribuições de Gumbel, tem-se, no caso estandardizado,

$$F(x|0, 1) = G_b(x|0, 1) = e^{-e^{-x}}, \quad -\infty < x < \infty, \quad (4.5.10)$$

bem como a respetiva função densidade

$$g_b(x|0, 1) = e^{-x} e^{-e^{-x}}, \quad -\infty < x < \infty, \quad (4.5.11)$$

e a função geradora de momentos é dada por

$$\varphi(t|0, 1) = \int_{-\infty}^{+\infty} e^{tx} e^{-x} e^{-e^{-x}} dx = \int_{-\infty}^{+\infty} e^{-(1-t)x} e^{-e^{-x}} dx, \quad (4.5.12)$$

definida para  $t < 1$ . Fazendo-se a transformação  $z = e^{-x}$ , vem

$$\varphi(t|0, 1) = \int_0^1 z^{-t} e^{-z} dz = \Gamma(1-t). \quad (4.5.13)$$

Logo,

$$\psi(t|0, 1) = \log \Gamma(1-t), \quad (4.5.14)$$

donde

$$\psi'(t|0, 1) = -\frac{\Gamma'(1-t)}{\Gamma(1-t)}, \quad (4.5.15)$$

e obtém-se

$$\Gamma'(1-t) = -\psi'(t|0, 1)\Gamma(1-t). \quad (4.5.16)$$

Assim,

$$\begin{aligned} \Gamma^{<r>}(1-t) &= -\psi'(t|0, 1)\Gamma(1-t)^{<r-1>} \\ &= -\sum_{j=0}^{r-1} \binom{r-1}{j} \psi^{<j+1>}(t|0, 1) \Gamma^{<r-1-j>}(1-t), \end{aligned} \quad (4.5.17)$$

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

vindo

$$\Gamma^{<r>}(1-t) = -\psi^{<r>}(t|0,1)\Gamma(1-t) - \sum_{j=0}^{r-2} \binom{r-1}{j} \psi^{<j+1>}(t|0,1)\Gamma^{<r-1-j>}(1-t). \quad (4.5.18)$$

Se  $t = 0$ , então temos

$$\Gamma^{<r>}(1) = -\chi_r(|0,1)\Gamma(1) - \sum_{j=0}^{r-2} \binom{r-1}{j} \chi_{j+1}(|0,1)\Gamma^{<r-1-j>}(1), \quad (4.5.19)$$

onde

$$\chi_r(|0,1) = -\Gamma^{<r>}(1) - \sum_{j=0}^{r-2} \binom{r-1}{j} \chi_{j+1}(|0,1)\Gamma^{<r-1-j>}(1). \quad (4.5.20)$$

Em geral, ter-se-á

$$\chi_r(\delta) = \delta^r \chi_r(1), \quad r > 1, \quad (4.5.21)$$

bem como

$$\delta = \sqrt{\frac{\chi_2(\delta)}{\chi_2(1)}}. \quad (4.5.22)$$

## Capítulo 5

### Modelos Aditivos com parâmetros de Localização, Dispersão e Forma

#### 5.1 Considerações Preliminares

Nesta secção consideramos o caso em que as distribuições das componentes dos vetores  $\underline{Z}_i$ ,  $i = 1, \dots, w$ , têm parâmetros de localização  $\lambda_i$ , dispersão  $\delta_i$  e forma  $\tau_i$ ,  $i = 1, \dots, w$ , respetivamente. De forma análoga à secção anterior, obteremos os estimadores dos cumulantes de quarta ordem, dados por  $\tilde{\chi}_{4,i}$ ,  $i = 1, \dots, w$  bem como os Elipsóides de Confiança, Intervalos de Predição e Modelos Múltiplos, seguindo as secções 3.2, 3.3 e 3.4, em que foram apresentados os casos gerais. Particularizamos para os casos em que as componentes  $Z_{i,l}$ ,  $l = 1, \dots, c_i$ ,  $i = 1, \dots, w$  têm distribuições Gama, Weibull e Fréchet.

Assim, admitamos que as componentes dos vetores  $\underline{Z}_i$ ,  $i = 1, \dots, w$ , têm cumulantes de ordem  $r$

$$\begin{cases} \chi_{1,i}(\lambda_i, \delta_i, \tau_i) = \lambda_i + \delta_i g_{1,i}(\tau_i), & i = 1, \dots, w, \\ \chi_{r,i}(\lambda_i, \delta_i, \tau_i) = \delta_i^r g_{r,i}(\tau_i), & i = 1, \dots, w, \quad r = 2, \dots \end{cases}, \quad (5.1.1)$$

sendo conhecidas as funções  $g_{r,i}(\tau_i)$ ,  $[g_{r,1}(\tau_i) = \chi_{r,i}(0, 1, \tau_i)$ ,  $i = 1, \dots, w$ ,  $r = 1, 2, \dots]$ . Observe-se que

$$\dot{\underline{Z}}_i = \lambda_i \underline{1}_{c_i} + \underline{Z}_i, \quad i = 1, \dots, w, \quad (5.1.2)$$

tendo as componentes de  $\underline{Z}_i$  parâmetros  $(0, \delta_i, \tau_i)$ ,  $i = 1, \dots, w$ . O modelo

$$\underline{Y} = \dot{\underline{X}} \underline{\dot{\beta}} + \sum_{i=1}^w \dot{\underline{X}}_i \dot{\underline{Z}}_i, \quad (5.1.3)$$

pode agora ser reescrito como

$$\underline{Y} = \underline{X} \underline{\beta} + \sum_{i=1}^w \underline{X}_i \underline{Z}_i, \quad (5.1.4)$$

com

$$\begin{cases} \underline{X} = [\dot{\underline{X}} \quad \underline{X}_1 \underline{1}_{c_1} \quad \dots \quad \underline{X}_w \underline{1}_{c_w}] \\ \underline{\beta} = [\dot{\underline{\beta}}^\top \quad \lambda_1 \quad \dots \quad \lambda_w]^\top \end{cases} \quad (5.1.5)$$

Observe-se que o  $r$ -ésimo cumulante pode ser escrito como

$$\chi_{r,i}(0, \delta_i, \tau_i) = \delta_i^r g_{r,i}(\tau_i), \quad i = 1, \dots, w, \quad r = 2, \dots \quad (5.1.6)$$

## 5.2 Ajustamentos

Procedendo-se como em (3.1.6), obtêm-se as matrizes  $\mathbf{B}(r)$ ,  $r = 2, 3, 4, \dots$ , já que, como vimos, apenas temos de estimar os cumulantes  $\chi_{r,i}$ ,  $i = 1, \dots, w$ ,  $r = 2, 3$ , necessitando, portanto, apenas dum modelo. Considerando

$$\tilde{\mathcal{Q}}_r = (\dot{Y}_1^r, \dots, \dot{Y}_w^r), \quad r = 2, 3, \quad (5.2.1)$$

e

$$\tilde{\chi}_r = (\tilde{\chi}_{r,1}, \dots, \tilde{\chi}_{r,w}) \quad (5.2.2)$$

obtém-se

$$\tilde{\chi}_r = (\mathbf{B}(r)^\top \mathbf{B}(r))^+ \mathbf{B}(r)^\top \tilde{\mathcal{Q}}_r, \quad r = 2, 3. \quad (5.2.3)$$

Ora, com

$$h_i(\cdot) = \frac{g_{3,i}^2(\cdot)}{g_{2,i}^3(\cdot)}, \quad (5.2.4)$$

temos

$$\frac{\chi_{3,i}^2}{\chi_{2,i}^3} = \frac{g_{3,i}^2(\tau_i)}{g_{2,i}^3(\tau_i)} = h_i(\tau_i), \quad i = 1, \dots, w, \quad (5.2.5)$$

vindo

$$\tau_i = h_i^{-1} \left( \frac{\chi_{3,i}^2}{\chi_{2,i}^3} \right), \quad i = 1, \dots, w. \quad (5.2.6)$$

Obtemos assim o estimador

$$\tilde{\tau}_i = h_i^{-1} \left( \frac{\tilde{\chi}_{3,i}^2}{\tilde{\chi}_{2,i}^3} \right), \quad i = 1, \dots, w, \quad (5.2.7)$$

bem como

$$\tilde{\delta}_i = \sqrt{\frac{\tilde{\chi}_{2,i}^2}{g_{2,i}(\tilde{\tau}_i)}}, \quad i = 1, \dots, w, \quad (5.2.8)$$

onde se tem

$$\tilde{\chi}_{r,i} = \tilde{\delta}_i^r g_{r,i}(\tilde{\tau}_i), \quad i = 1, \dots, w, \quad r \geq 4. \quad (5.2.9)$$

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

Assim, tem-se

$$\tilde{Q}_r = \mathbf{B}(r)\tilde{\chi}_r, \quad r = 2, 3, \dots \quad (5.2.10)$$

Para estimar o vetor das componentes, trabalhamos com o vetor  $\underline{Y}$  das observações do modelo dado em (5.1.3), tendo-se

$$\underline{\Sigma}(\underline{Y}) = \sum_{i=1}^w \chi_{2,i} \mathbf{M}_i, \quad (5.2.11)$$

bem como

$$\tilde{\Sigma}(\underline{Y}) = \sum_{i=1}^w \tilde{\chi}_{2,i} \mathbf{M}_i. \quad (5.2.12)$$

De acordo com Kariya & Kurata (2004) [31], obtemos o GLSE

$$\underline{\tilde{\beta}} = (\mathbf{X}^\top \tilde{\Sigma}(\underline{Y}) + \mathbf{X})^+ \mathbf{X}^\top \tilde{\Sigma}(\underline{Y})^+ \underline{Y}. \quad (5.2.13)$$

Tal como foi visto anteriormente, se  $\tilde{\Sigma}(\underline{Y})$  for invertível e os vectores coluna de  $\mathbf{X}$  forem linearmente independentes,  $\mathbf{X}^\top \tilde{\Sigma}(\underline{Y})^{-1} \mathbf{X}$  é invertível, vindo

$$\underline{\tilde{\beta}} = (\mathbf{X}^\top \tilde{\Sigma}(\underline{Y})^{-1} \mathbf{X})^{-1} \mathbf{X}^\top \tilde{\Sigma}(\underline{Y})^{-1} \underline{Y}, \quad (5.2.14)$$

bem como

$$\tilde{\Sigma}(\underline{\tilde{\beta}}) = (\mathbf{X}^\top \tilde{\Sigma}(\underline{Y})^{-1} \mathbf{X})^{-1}. \quad (5.2.15)$$

Obtemos assim a

**Proposição 5.1** *Se  $\mathbf{X}$  tiver vectores coluna linearmente independentes e  $\tilde{\Sigma}(\underline{Y})$  for invertível, então*

$$\tilde{\Sigma}(\underline{\tilde{\beta}}) = (\mathbf{X}^\top \tilde{\Sigma}(\underline{Y})^{-1} \mathbf{X})^{-1}. \quad (5.2.16)$$

**Corolário 5.1** *Verificadas as condições da tese da Proposição 5.1., para o vetor estimável  $\underline{\psi} = \mathbf{G}\underline{\beta}$ , tem-se o estimador  $\underline{\tilde{\psi}} = \mathbf{G}\underline{\tilde{\beta}}$ , com  $\tilde{\Sigma}(\underline{\tilde{\psi}}) = \mathbf{G}(\mathbf{X}^\top \tilde{\Sigma}(\underline{Y})^{-1} \mathbf{X})^{-1} \mathbf{G}^\top$ . Se  $\mathbf{G}$  for uma matriz  $h \times k$ , com característica  $h$ ,  $\tilde{\Sigma}(\underline{\tilde{\psi}})$  será invertível.*

### 5.3 Elipsóides de Confiança

Dado ter-se um único modelo, temos

$$W(\underline{u}) = \underline{u}^\top (\underline{Y} - \underline{\mu}) = \sum_{i=1}^w a_i^\top(\underline{u}) Z_i, \quad (5.3.1)$$

onde, como atrás,

$$\underline{a}_i^\top(\underline{u}) = \underline{u}^\top \mathbf{X}_i = (a_{i,1}(\underline{u}), \dots, a_{i,c_i}(\underline{u})), \quad i = 1, \dots, w. \quad (5.3.2)$$

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

Temos igualmente

$$E(W(\underline{u})) = 0, \quad (5.3.3)$$

bem como, tal como definido atrás, sendo  $v_r(\underline{u})$  o  $r$ -ésimo cumulante de  $W(\underline{u})$ , podemos escrever

$$v_r(\underline{u}) = \sum_{i=1}^w d_i(r, \underline{u}) \chi_{r,i}, \quad r = 2, 3, 4, \quad (5.3.4)$$

com

$$d_i(r, \underline{u}) = \sum_{l=1}^{c_i} a_{i,l}^r(\underline{u}), \quad i = 1, \dots, w, \quad r = 2, 3, 4. \quad (5.3.5)$$

Teremos pois os estimadores

$$\left\{ \begin{array}{l} \tilde{v}_r(\underline{u}) = \sum_{i=1}^w d_i(r, \underline{u}) \tilde{\chi}_{r,i}, \quad r = 2, 3, 4, \\ \tilde{\gamma}_1(\underline{u}) = \frac{\tilde{v}_3(\underline{u})}{\tilde{v}_2(\underline{u})^{\frac{3}{2}}} \\ \tilde{\gamma}_2(\underline{u}) = \frac{\tilde{v}_4(\underline{u})}{\tilde{v}_2(\underline{u})^2} - 3, \end{array} \right. , \quad (5.3.6)$$

obtendo-se, para a distribuição  $F_{\dot{W}(\underline{u})}$  de

$$\dot{W}(\underline{u}) = \frac{W(\underline{u})}{\sqrt{\tilde{v}_2(\underline{u})}}, \quad (5.3.7)$$

a expansão de Edgeworth

$$Ed_{\dot{W}(\underline{u})}(w) = N(w) - \frac{\tilde{\gamma}_1(\underline{u})}{3!} n^{<2>}(w) + \frac{\tilde{\gamma}_2(\underline{u})}{4!} n^{<3>}(w) + 10 \frac{\tilde{\gamma}_1^2(\underline{u})}{6!} n^{<6>}(w), \quad i = 1, \dots, w. \quad (5.3.8)$$

Dado o  $p$ -ésimo quantil (estimado)  $\tilde{w}_p(\underline{u})$  de  $\dot{W}(\underline{u})$ , para  $W(\underline{u})$ , teremos os quantis estimados

$$\tilde{w}_p(\underline{u}) = \sqrt{\tilde{v}_2(\underline{u})} \tilde{w}_p(\underline{u}). \quad (5.3.9)$$

A partir daqui, obtém-se o elipsóide de confiança de "nível  $p$ ," minimizando  $\Delta(\mathbf{K})$ , tal como em (3.2.29).

Mantendo as definições de (3.2.30) e tomando igualmente  $\underline{\lambda} = \underline{s}(\mathbf{K})$ , temos os *LSE*

$$\left\{ \begin{array}{l} \tilde{\underline{\lambda}}_p = (U]^{2[^\top U]^{2[}]^{-1} U]^{2[^\top} \tilde{w}_p \\ \tilde{\mathbf{K}}_p = \mathbf{K}(\tilde{\underline{\lambda}}_p) \end{array} \right. , \quad (5.3.10)$$

com

$$\tilde{w}_p = (\tilde{w}_p(\underline{u}_1), \dots, \tilde{w}_p(\underline{u}_m)), \quad (5.3.11)$$

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

e  $\tilde{K}_p = \mathbf{K}(\tilde{\lambda}_p)$ . A interpretação geométrica da Figura 3.1. mantém-se. Esta secção é paralela à secção 3.2, do capítulo 3. No entanto, é de salientar que agora trabalhamos com um modelo, em vez de um par de modelos.

### 5.4 Intervalos de Predição

Tal como na secção 4.3, teremos

$$\underline{\dot{u}} = [\dot{x}^\top \dot{x}_1^\top \dots \dot{x}_w^\top]^\top, \quad (5.4.1)$$

e

$$\underline{\ddot{u}} = \frac{1}{\|\underline{\dot{u}}\|} \underline{\dot{u}}, \quad (5.4.2)$$

obtendo-se  $Ed_{\hat{W}(\underline{\ddot{u}})}(\cdot)$ , e, a partir daí os  $\hat{w}_p(\underline{\ddot{u}})$  e os  $\tilde{w}_p(\underline{\ddot{u}})$ , por forma a construir os Intervalos de Predição seguintes, de nível  $1 - p$ ,

$$[\underline{\dot{x}}^\top \tilde{\beta} + \tilde{w}_{\frac{p}{2}}(\underline{\ddot{u}}); \underline{\dot{x}}^\top \tilde{\beta} + \tilde{w}_{1-\frac{p}{2}}(\underline{\ddot{u}})]. \quad (5.4.3)$$

Tal como na secção 4.3,  $1 - p$  é apenas um valor aproximado deste intervalo conter a futura observação.

### 5.5 Modelos Múltiplos

Ao contrário do que admitimos na secção 3.4, do capítulo 3, temos agora de considerar um modelo por tratamento do delineamento base, logo, um único estimador, com valores  $\tilde{\psi}_1, \dots, \tilde{\psi}_d$ , para os valores duma função estimável, para os vários tratamentos. Temos agora

$$\begin{cases} \underline{\psi} = (\psi_1, \dots, \psi_d) \\ \underline{\tilde{\psi}} = (\tilde{\psi}_1, \dots, \tilde{\psi}_d) \end{cases}, \quad (5.5.1)$$

bem como

$$\begin{cases} S_j = \|\mathbf{A}_j \underline{\tilde{\psi}}\|^2 \\ g_j = \text{car}(\mathbf{A}_j) \end{cases}, \quad j = 1, \dots, m. \quad (5.5.2)$$

Dado ter-se um único modelo por tratamento, temos de admitir existir um conjunto  $\varphi$ , de índices, tal que

$$\|\mathbf{A}_j \underline{\psi}\|^2 = 0_{g_j}, \quad j \in \varphi. \quad (5.5.3)$$

Para testar as hipóteses

$$H_{0,j} : \mathbf{A}_j \underline{\psi} = 0_{g_j}, \quad j \notin \varphi, \quad (5.5.4)$$

temos as estatísticas

$$F_j = \frac{g}{g_j} \frac{S_j}{S}, \quad j \notin \varphi, \quad (5.5.5)$$

com  $S = \sum_{j \in \varphi} S_j$  e onde  $g_j$  e  $g = \sum_{j' \in \varphi} g_{j'}$  representam os graus de liberdade; as hipóteses só se anulam quando  $H_{0,j}$ ,  $j \notin \varphi$ , se verifica.

## 5.6 Casos particulares

A teoria anterior aplica-se desde que se conheçam as funções  $g_r(p)$ , que dão os  $\chi_r(0, 1, p)$ ,  $r = 2, 3$ . Vamos obtê-las para três casos particulares:

### 5.6.1 Tipo Gama

Seja  $x$  uma *v.a.* com distribuição Gama, cuja densidade estandardizada é dada por

$$g(x|0, 1, p) = \frac{1}{\Gamma(p)} x^{p-1} e^{-x}, \quad x > 0, \quad (5.6.1)$$

vindo

$$\begin{aligned} \varphi(t|0, 1, p) &= \int_0^{+\infty} e^{tx} \frac{x^{p-1}}{\Gamma(p)} e^{-x} dx = \\ &= \frac{1}{\Gamma(p)} \int_0^{+\infty} e^{-(1-t)x} x^{p-1} dx, \end{aligned} \quad (5.6.2)$$

e, fazendo-se a transformação  $z = (1-t)x$ , temos

$$\begin{aligned} \varphi(t|0, 1, p) &= \frac{1}{\Gamma(p)} \int_0^{+\infty} \left( \frac{z}{1-t} \right)^{p-1} e^{-z} \frac{1}{1-t} dz = \\ &= \frac{1}{(1-t)^p}, \quad t < 1. \end{aligned} \quad (5.6.3)$$

Temos também

$$\psi(t|0, 1, p) = -p \log(1-t), \quad t < 1, \quad (5.6.4)$$

vindo

$$\left\{ \begin{array}{l} \psi^{<1>}(t|0, 1, p) = \frac{p}{1-t} \\ \vdots \\ \psi^{<r>}(t|0, 1, p) = \frac{p(r-1)!}{(1-t)^r}; \quad r = 2, 3, \dots \end{array} \right. , \quad (5.6.5)$$

bem como

$$\left\{ \begin{array}{l} \chi_1(|0, 1, p) = p = g_1(p) \\ \vdots \\ \chi_r(|0, 1, p) = p(r-1)! = g_r(p) \end{array} \right. . \quad (5.6.6)$$

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

Tem-se então ( $p$  é o parâmetro de forma),

$$h(p) = \frac{4p^2}{p^3} = \frac{4}{p}, \quad (5.6.7)$$

bem como

$$p = \frac{\chi_2^2( |0, 1, p)}{4\chi_3^2( |0, 1, p)}. \quad (5.6.8)$$

Vê-se ainda que se tem

$$\delta = \sqrt{\frac{\chi_2( |0, 1, p)}{\chi_2( |0, 1, p)}}. \quad (5.6.9)$$

### 5.6.2 Tipo Weibull

Temos a densidade estandardizada,

$$f(x|0, 1, \xi) = \xi x^{\xi-1} e^{-x^\xi}, \quad \text{se } x > 0 \quad (5.6.10)$$

a que correspondem os momentos relativos à origem

$$\mu'_r( |0, 1, \xi) = \int_0^{+\infty} \xi x^{r+\xi-1} e^{-x^\xi} dx, \quad (5.6.11)$$

Fazendo a transformação  $z = x^\xi$ , bem como  $\frac{dx}{dz} = \frac{1}{\xi} z^{\frac{1}{\xi}-1}$ , temos

$$\begin{aligned} \mu'_r( |0, 1, \xi) &= \int_0^{+\infty} \xi z^{\frac{r+\xi-1}{\xi}} e^{-z} \frac{1}{\xi} z^{\frac{1}{\xi}-1} dz = \int_0^{+\infty} z^{\frac{r}{\xi}} e^{-z} dz \\ &= \Gamma(1 + \frac{r}{\xi}). \end{aligned} \quad (5.6.12)$$

Os quatro momentos centrados serão:

$$\left\{ \begin{array}{l} \mu_1( |0, 1, \xi) = 0 \\ \mu_2( |0, 1, \xi) = \Gamma(1 + \frac{2}{\xi}) - \Gamma^2(1 + \frac{1}{\xi}) \\ \mu_3( |0, 1, \xi) = \Gamma(1 + \frac{3}{\xi}) - 3\Gamma(1 + \frac{1}{\xi})\Gamma(1 + \frac{2}{\xi}) + 2\Gamma(1 + \frac{1}{\xi})^3 \\ \mu_4( |0, 1, \xi) = \Gamma(1 + \frac{4}{\xi}) - 4\Gamma(1 + \frac{1}{\xi})\Gamma(1 + \frac{3}{\xi}) + 6\Gamma(1 + \frac{1}{\xi})^2\Gamma(1 + \frac{2}{\xi}) - 3\Gamma(1 + \frac{1}{\xi})^4 \end{array} \right. \quad (5.6.13)$$

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

Para os cumulantes, temos

$$\begin{cases} \chi_1(|\lambda, \delta, \xi) = \lambda + \delta\chi_1(|0, 1, \xi) \\ \vdots \\ \chi_r(|\lambda, \delta, \xi) = \delta^r\chi_r(|0, 1, \xi), \quad r > 1 \end{cases} \quad (5.6.14)$$

Então, com

$$\chi_1(|0, 1, \xi) = \mu'(|0, 1, \xi) = \Gamma\left(1 + \frac{1}{\xi}\right) \quad (5.6.15)$$

e

$$\chi_r(|0, 1, \xi) = \mu_r(|0, 1, \xi), \quad r = 2, 3, \quad (5.6.16)$$

temos

$$\chi_4(|\lambda, \delta, \xi) = \delta^4\chi_4(|0, 1, \xi) = \delta^4[\mu_4(|0, 1, \xi) - 3\mu_2(|0, 1, \xi)^2]. \quad (5.6.17)$$

### 5.6.3 Tipo Fréchet

Sabendo que a distribuição de Fréchet, Stuart, A. & Ord, K. A. (1994, pág. 485) [68], tem função densidade estandardizada dada por

$$f(x|0, 1, \xi) = \xi x^{-\xi-1} e^{-x^{-\xi}}, \quad x > 0, \quad (5.6.18)$$

os momentos da distribuição de Fréchet são dados por

$$\mu'_r(|0, 1, \xi) = \int_0^{+\infty} \xi x^{r-(\xi+1)} e^{-x^{-\xi}} dx = \xi \int_0^{+\infty} x^{r-(\xi+1)} e^{-x^{-\xi}} dx. \quad (5.6.19)$$

Fazendo agora a transformação  $z = x^{-\xi}$ , vem  $x = z^{-\frac{1}{\xi}}$  e  $\frac{dx}{dz} = -\frac{1}{\xi} z^{-\frac{1}{\xi}-1}$ , logo

$$\begin{aligned} \mu'_r(|0, 1, \xi) &= \int_0^{+\infty} \xi z^{\frac{\xi+1-r}{\xi}} e^{-z} \frac{1}{\xi} z^{-\frac{1}{\xi}-1} dz \\ &= \int_0^{+\infty} z^{-\frac{r}{\xi}} e^{-z} dz \\ &= \Gamma\left(1 - \frac{r}{\xi}\right), \end{aligned} \quad (5.6.20)$$

não tendo portanto elementos de ordem superior a  $\xi$ .

Em particular, tem-se

$$\left\{ \begin{array}{l} \mu'_1(0, 1, \xi) = \Gamma(1 - \frac{1}{\xi}), \quad \xi < 1 \\ \mu'_2(0, 1, \xi) = \Gamma(1 - \frac{2}{\xi}) \\ \mu'_3(0, 1, \xi) = \Gamma(1 - \frac{3}{\xi}) \\ \mu'_4(0, 1, \xi) = \Gamma(1 - \frac{4}{\xi}) \end{array} \right. , \quad (5.6.21)$$

vindo, de (5.6.21)

$$\left\{ \begin{array}{l} \mu_2(0, 1, \xi) = \Gamma(1 - \frac{2}{\xi}) - \Gamma^2(1 - \frac{1}{\xi}) \\ \mu_3(0, 1, \xi) = \Gamma(1 - \frac{3}{\xi}) - 3\Gamma(1 - \frac{1}{\xi})\Gamma(1 - \frac{2}{\xi}) + 2\Gamma(1 - \frac{1}{\xi})^3 \\ \mu_4(0, 1, \xi) = \Gamma(1 - \frac{4}{\xi}) - 4\Gamma(1 - \frac{1}{\xi})\Gamma(1 - \frac{3}{\xi}) + 6\Gamma(1 - \frac{1}{\xi})^2\Gamma(1 - \frac{2}{\xi}) - 3\Gamma(1 - \frac{1}{\xi})^4 \end{array} \right. . \quad (5.6.22)$$

Atendendo a (5.6.20), obtemos os cumulantes

$$\left\{ \begin{array}{l} \chi_2(0, 1, \xi) = \mu_2(0, 1, \xi) \\ \chi_3(0, 1, \xi) = \mu_3(0, 1, \xi) \\ \chi_4(0, 1, \xi) = \Gamma(1 - \frac{4}{\xi}) - 4\Gamma(1 - \frac{1}{\xi})\Gamma(1 - \frac{3}{\xi}) + 6\Gamma^2(1 - \frac{1}{\xi})\Gamma(1 - \frac{2}{\xi}) - 3\Gamma(1 - \frac{1}{\xi})^4 + 3\Gamma(1 - \frac{1}{\xi})^4 \\ \quad + 3\Gamma(1 - \frac{1}{\xi})^4 - 3(\Gamma(1 - \frac{2}{\xi}) - \Gamma^2(1 - \frac{1}{\xi}))^2 \end{array} \right. \quad (5.6.23)$$

Note-se que se pode, perfeitamente, admitir que as componentes dos  $\underline{Z}_1, \dots, \underline{Z}_w$ , pertencem a tipos distintos.



## Capítulo 6

### Duas Aplicações através de simulações

De forma a ilustrar a aplicabilidade dos resultados obtidos iremos agora primeiro apresentar dois estudos de simulação e, de seguida, um exemplo de aplicação dos modelos múltiplos aditivos. Consideraremos modelos com dois factores aleatórios, que cruzam, em ambos os estudos de simulação, diferindo, no entanto, na forma como as observações são simuladas.

Assim, consideraremos um caso com dois factores aleatórios, que cruzam ( $A$  com dois níveis e  $B$  com três níveis).

Geramos dois vetores de observações, de tamanho 6, independentes e identicamente distribuídos. Todos os cálculos foram feitos no software R. Os modelos considerados podem ser escritos como se apresenta em (3.1.2), com  $w = 2$ , vindo

$$\underline{Y}(h) = \mathbf{X}\underline{\beta}(h) + \mathbf{X}_1\underline{Z}_1(h) + \mathbf{X}_2\underline{Z}_2(h), \quad h = 1, 2, \quad (6.0.1)$$

onde

$$\mathbf{X} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}^T,$$

$$\mathbf{X}_1 = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}^T,$$

$$\mathbf{X}_2 = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}^T,$$

$\underline{\beta}$  é fixo e os vetores  $\underline{Z}_1(h)$  e  $\underline{Z}_2(h)$ ,  $h = 1, 2$ , são aleatórios e não correlacionados, com vetores médios nulos e matrizes de variância e covariância  $\chi_{2,1}(h)\mathbf{I}_2$  e  $\chi_{2,2}(h)\mathbf{I}_3$ .

De forma a realizar os ajustamentos, necessitamos considerar os vetores  $\alpha_1, \dots, \alpha_5$ , apresentados em (6.2.4), que constituem uma base ortonormada para  $\Omega^\perp$ , o complemento ortogonal do espaço  $\Omega$ . Assim, consideremos

$$\mathbf{A} = \begin{bmatrix} -\frac{1}{2} & 0 & \frac{1}{2} & -\frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{\sqrt{12}} & -\frac{2}{\sqrt{12}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{12}} & -\frac{2}{\sqrt{12}} & \frac{1}{\sqrt{12}} \\ -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{1}{2} & 0 & -\frac{1}{2} & -\frac{1}{2} & 0 & \frac{1}{2} \\ -\frac{1}{\sqrt{12}} & -\frac{2}{\sqrt{12}} & -\frac{1}{\sqrt{12}} & \frac{1}{\sqrt{12}} & -\frac{2}{\sqrt{12}} & \frac{1}{\sqrt{12}} \end{bmatrix}.$$

sendo  $\alpha_l^\top$ , a  $l$ -ésima linha da matriz  $\mathbf{A}$ ,  $l = 1, \dots, 5$ .

## 6.1 Primeiro estudo de simulação

No primeiro estudo de simulação, geramos 1000 vezes os seguintes vetores de observações,

$$\underline{Y}(h) = [Y_1(h), \dots, Y_6(h)]^\top, h = 1, 2,$$

com

$$\underline{Y}(1), \underline{Y}(2) \sim N(\underline{0}, \sigma_s^2 \mathbf{I}_6), \quad s = 1, \dots, 6, \quad (6.1.2)$$

e  $\sigma_1^2 = 0.3$ ,  $\sigma_2^2 = 0.6$ ,  $\sigma_3^2 = 1.2$ ,  $\sigma_4^2 = 1.8$ ,  $\sigma_5^2 = 2.4$  e  $\sigma_6^2 = 4.8$ . Assim, serão considerados 6 casos, sendo que em cada caso se terão dois vetores de observações independentes e identicamente distribuídos.

De (3.1.3) temos  $\dot{Y}_l(1) = \mathbf{A}\underline{Y}(1)$  e  $\dot{Y}_l(2) = \mathbf{A}\underline{Y}(2)$ ,  $l = 1, \dots, 5$ , obtendo-se, de (3.1.7),

$$\begin{cases} \tilde{\mathcal{O}}_{r,l} = \frac{\dot{Y}_l(1)^r + \dot{Y}_l(2)^r}{2}, \quad r = 2, 3 \\ \tilde{\mathcal{O}}_{4,l} = \frac{\dot{Y}_l(1)^4 + \dot{Y}_l(2)^4}{2} - 3\dot{Y}_l(1)^2\dot{Y}_l(2)^2 \end{cases}, l = 1, \dots, 5. \quad (6.1.3)$$

Como já mencionado na teoria, enquanto que para obter um estimador centrado para  $\mathcal{O}_{4,l}$ ,  $l = 1, \dots, 5$ , necessitamos de dois modelos independentes e identicamente distribuídos, para calcular os estimadores centrados de  $\mathcal{O}_{2,l}$  e  $\mathcal{O}_{3,l}$  necessitamos apenas de um modelo. Contudo, como agora temos dois modelos, vamos utilizá-los para melhorar os estimadores obtendo  $\tilde{\mathcal{O}}_{r,l}$ ,  $r = 2, 3$  como se indica em (6.1.3). Os  $\sigma_s^2$ ,  $s = 1, \dots, 6$ , previamente considerados e os valores médios das estimativas de  $\mathcal{O}_{r,l}$ , dadas por  $\tilde{\mathcal{O}}_{r,l}$ ,  $r = 2, 3, 4$ ,  $l = 1, \dots, 5$ , são apresentados na Tabela 6.1.

Tabela 6.1: Médias obtidas para as estimativas.

$\sigma_i^2, i = 1 \dots 6:$		0.3	0.6	1.2	1.8	2.4	4.8
$\tilde{\mathcal{O}}_{2,l}$	$l = 1$	0.3018	0.6082	1.2062	1.7123	2.3856	4.4692
	$l = 2$	0.2915	0.5872	1.2177	1.8806	2.4989	4.8668
	$l = 3$	0.2997	0.6079	1.2614	1.8481	2.3242	4.8386
	$l = 4$	0.2961	0.6047	1.1682	1.9162	2.3206	4.8771
	$l = 5$	0.3009	0.5839	1.1124	1.7218	2.2960	4.7057
$\tilde{\mathcal{O}}_{3,l}$	$l = 1$	0.0419	0.0364	0.1851	0.0331	-0.3278	0.4252
	$l = 2$	0.0193	-0.0214	-0.1413	0.1595	-0.2323	0.1794
	$l = 3$	-0.0225	0.0288	-0.0571	-0.3290	-0.0814	0.2556
	$l = 4$	0.0073	0.0130	-0.0776	0.2581	0.3230	0.6529
	$l = 5$	0.0143	-0.0288	0.0709	-0.1193	0.2897	0.8125
$\tilde{\mathcal{O}}_{4,l}$	$l = 1$	-0.0067	0.0477	-0.2182	-1.0895	-0.7089	0.4231
	$l = 2$	-0.0029	-0.0774	0.5716	-0.1009	0.2468	0.9922
	$l = 3$	-0.0339	-0.0159	-0.0176	-0.9297	0.0770	-0.0499
	$l = 4$	0.0060	-0.0086	0.1360	-0.2694	0.0155	-1.2685
	$l = 5$	0.0058	-0.1303	0.2589	0.0432	0.5745	0.7388

Podemos constatar que  $\tilde{\mathcal{O}}_{2,l}$ ,  $l = 1, \dots, 5$ , apresenta excelentes estimativas dos  $\sigma_s^2$ ,  $s = 1, \dots, 6$ , em todos os casos considerados. Além disso, também estimamos o segundo, terceiro e quarto momentos centrais de  $\dot{Y}_l(h)$ ,  $h = 1, 2$ ,  $l = 1, \dots, 5$ , usando o package "moments" implementado no software R (R Core Team, 2013) [56], tendo obtido valores iguais aos obtidos com a técnica

aqui apresentada.

Este primeiro estudo de simulação está publicado em Antunes et al. (2020a) [2].

## 6.2 Segundo estudo de simulação

Uma das diferenças entre este estudo e o anterior consiste na forma como os dados foram simulados. Neste, voltámos a considerar o modelo descrito em (6.0.1). No entanto, neste estudo não são necessários dois vetores de observações *i.i.d.*. Assim, em cada simulação, obtemos apenas um vetor de observações, dado por

$$\underline{Y} = \mathbf{X}\underline{\beta} + \mathbf{X}_1\underline{Z}_1 + \mathbf{X}_2\underline{Z}_2, \quad (6.2.4)$$

que basicamente é o modelo considerado em (6.0.1), com  $h = 1$ . Por outro lado, em vez de simular as observações, considerando a condição em (6.1.2), estas são obtidas simulando os vetores  $\underline{Z}_1$  e  $\underline{Z}_2$ . Mais uma vez, o procedimento foi repetido 1000 vezes.

Este estudo de simulação está dividido em dois casos. No primeiro (Subsecção 6.2.1), consideramos modelos com parâmetros de localização e dispersão, ilustrando os casos apresentados no capítulo 4. No segundo (Subsecção 6.2.2) consideramos modelos com parâmetros de localização, dispersão e forma, ilustrando os casos apresentados no capítulo 5.

### 6.2.1 Primeiro caso

Como mencionado anteriormente, nesta subsecção ilustra-se a teoria apresentada no capítulo 4, considerando a distribuição normal. Para tal, considerou-se o modelo descrito em (6.2.4), com  $\underline{\beta} = \underline{0}$ ,  $\underline{Z}_1 \sim N(\underline{0}, \chi_{2,1}\mathbf{I}_2)$  e  $\underline{Z}_2 \sim N(\underline{0}, \chi_{2,2}\mathbf{I}_3)$ , com os cumulantes de  $\underline{Z}_i, i = 1, 2$  também denominados componentes de variância, dados por  $\chi_{2,i}, i = 1, 2$ , tomando valores em  $\{0.3, 0.6, 1.2, 1.8, 2.4, 4.8\}$ . Assim, tendo em conta todos os diferentes valores para  $\chi_{2,i}, i = 1, 2$ , são considerados 36 vetores de observações  $\underline{Y}$ . Note-se que os vetores de observações terão, em todos os casos, vetor médio nulo e variância dependendo de  $\chi_{2,i}, i = 1, 2$ .

Note-se também que, enquanto no estudo anterior eram conhecidos os segundos momentos centrais dos vetores de observações, aqui conhecem-se as componentes de variância,  $\chi_{2,i}, i = 1, 2$ . Assim, estas podem ser estimadas, a partir de (3.1.10), com  $r = 2$ , e as suas estimativas comparadas com os valores assumidos acima. Para tal, obtemos a partir de (3.1.4),

$$\begin{array}{ll} \underline{\alpha}_1^\top \mathbf{X}_1 = [0 \ 0]; & \underline{\alpha}_1^\top \mathbf{X}_2 = [-1 \ 0 \ 1] \\ \underline{\alpha}_2^\top \mathbf{X}_1 = [0 \ 0]; & \underline{\alpha}_2^\top \mathbf{X}_2 = [0.577 \ -1.155 \ 0.577] \\ \underline{\alpha}_3^\top \mathbf{X}_1 = [-1.225 \ 1.225]; & \underline{\alpha}_3^\top \mathbf{X}_2 = [0 \ 0 \ 0] \\ \underline{\alpha}_4^\top \mathbf{X}_1 = [0 \ 0]; & \underline{\alpha}_4^\top \mathbf{X}_2 = [0 \ 0 \ 0] \\ \underline{\alpha}_5^\top \mathbf{X}_1 = [0 \ 0]; & \underline{\alpha}_5^\top \mathbf{X}_2 = [0 \ 0 \ 0] \end{array}$$

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

e de (3.1.8), com  $r = 2$ ,

$$B(2) = \begin{bmatrix} 0 & 2 \\ 0 & 2 \\ 3 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Então, atendendo a (3.1.10) e considerando  $r = 2$ , obtém-se

$$\tilde{\chi}_2 = (B(2)^\top B(2))^+ B(2)^\top \tilde{O}_2.$$

Depois de repetido este procedimento 1000 vezes, obteve-se a Tabela 6.2, onde constam as médias das estimativas obtidas, para cada um dos 36 casos considerados, tendo em conta as várias combinações de valores previamente considerados para as componentes de variância. Pode verificar-se que as estimativas médias estão muito perto dos valores previamente considerados.

Tabela 6.2: Valores previamente considerados de  $\chi_{2,i}$ ,  $i = 1, 2$ , e médias das estimativas das componentes de variância obtidas.

		$\chi_{2,1}$					
$\chi_{2,2}$		0.3	0.6	1.2	1.8	2.4	4.8
0.3	$\tilde{\chi}_{2,1}$	0.2936	0.6087	1.2315	1.7948	2.3762	4.6244
	$\tilde{\chi}_{2,2}$	0.3008	0.3013	0.2997	0.3033	0.3009	0.3012
0.6	$\tilde{\chi}_{2,1}$	0.3029	0.6128	1.1622	1.8165	2.4335	4.8770
	$\tilde{\chi}_{2,2}$	0.6096	0.6090	0.5940	0.5879	0.5926	0.5909
1.2	$\tilde{\chi}_{2,1}$	0.2979	0.6024	1.2016	1.8117	2.3822	4.7587
	$\tilde{\chi}_{2,2}$	1.1870	1.2099	1.1997	1.2153	1.2066	1.1958
1.8	$\tilde{\chi}_{2,1}$	0.3065	0.6114	1.2113	1.8233	2.3465	4.6282
	$\tilde{\chi}_{2,2}$	1.7555	1.7818	1.7874	1.7975	1.8271	1.8108
2.4	$\tilde{\chi}_{2,1}$	0.3018	0.5905	1.1618	1.7717	2.4373	4.7646
	$\tilde{\chi}_{2,2}$	2.4230	2.3864	2.4469	2.3731	2.4113	2.3937
4.8	$\tilde{\chi}_{2,1}$	0.2937	0.6003	1.1804	1.7969	2.3998	4.8816
	$\tilde{\chi}_{2,2}$	4.7961	4.8335	4.8328	4.7178	4.8049	4.8615

A Tabela 6.3 apresenta os valores médios obtidos para  $\tilde{\beta}$ , descrito em (4.1.10), para cada um dos 36 casos considerados. Lembramos que ao se simularem as observações se considerou  $\beta = \underline{0}$  e neste caso também se verifica que as estimativas médias estão muito perto dos valores previamente considerados.

Tabela 6.3: Médias obtidas para  $\tilde{\beta}$ .

		$\chi_{2,1}$					
$\chi_{2,2}$		0.3	0.6	1.2	1.8	2.4	4.8
0.3		-0.0043	-0.0008	-0.0027	-0.0032	-0.0009	-0.0006
0.6		-0.0070	-0.0085	-0.0090	-0.0043	-0.0027	-0.0010
1.2		-0.0051	-0.0025	-0.0022	-0.0002	-0.0008	0.0003
1.8		0.0020	0.0036	0.0020	0.0015	0.0011	0.0001
2.4		0.0004	0.0003	-0.0005	-0.0006	-0.0008	-0.0002
4.8		-0.0002	0.0008	0.0008	0.0007	0.0005	0.0004

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

Com vista à obtenção de elipsóides de confiança para observações futuras que, como mencionado no capítulo 3, passam pela direção dada por  $\underline{u}$ , à distância  $\underline{u}^\top \mathbf{K} \underline{u}$  do centro, consideraram-se vetores  $\underline{g}$ , com 6 componentes (tantas quantas o vetor de observações), podendo cada componente tomar os valores  $-0.9, 0.1, \text{ ou } 1.1$ , por exemplo e consideramos

$$\underline{u} = \frac{1}{\|\underline{g}\|} \times \underline{g}. \quad (6.2.5)$$

Desta forma foram considerados  $3^6 = 729$  vetores diferentes, todos com norma 1 e diferentes direções. Um desses vetores é, por exemplo,  $\underline{u} = \frac{1}{2.0149}[-0.9 \ 0.1 \ 1.1 \ -0.9 \ 0.1 \ 1.1]^\top$ .

Para cada vetor  $\underline{u}$ , obtiveram-se os valores do estimador do segundo cumulante de  $W(\underline{u})$ , dado por  $\tilde{v}_2(\underline{u})$ , ( $r=2$ ), do coeficiente de assimetria,  $\tilde{\gamma}_1(\underline{u})$ , e do coeficiente de curtose,  $\tilde{\gamma}_2(\underline{u})$ , apresentados em (3.2.25). De seguida, obteve-se o valor correspondente para  $Ed_{\tilde{W}(\underline{u})}(w)$ , segundo (3.2.27), o que nos permitiu obter os quantis estimados  $\tilde{w}_{0.025}(\underline{u})$  e  $\tilde{w}_{0.975}(\underline{u})$ , indicados em (3.2.28).

Considerou-se, então,  $U^{[2]}$ , tal como descrita em (3.2.32), por forma a obtermos  $\tilde{\lambda}_p$  tal como em (3.2.34) e, conseqüentemente,  $\tilde{\mathbf{K}}_{0.975}$ , com  $p = 0.975$ . Por exemplo, para  $\chi_{2,1} = \chi_{2,2} = 4.8$ , obteve-se

$$\tilde{\mathbf{K}}_{0.975} = \begin{bmatrix} 9.0122 & 5.5376 & 5.5866 & 6.9825 & -1.2271 & -1.2985 \\ 5.5376 & 8.9601 & 5.6214 & -1.2688 & 7.0123 & -1.2847 \\ 5.5866 & 5.6214 & 8.9450 & -1.2360 & -1.3083 & 7.0276 \\ 6.9825 & -1.2688 & -1.2360 & 8.9317 & 5.5936 & 5.6336 \\ -1.2271 & 7.0123 & -1.3083 & 5.5936 & 8.9460 & 5.5824 \\ -1.2985 & -1.2847 & 7.0276 & 5.6336 & 5.5824 & 8.9612 \end{bmatrix}.$$

Assim, novamente, por exemplo, para  $\chi_{2,1} = \chi_{2,2} = 4.8$ , a distância  $\underline{u}^\top \tilde{\mathbf{K}}_{0.975} \underline{u}$ , indicada na Figura 3.1, foi de 20.7402.

A fim de se verificar se os elipsóides continham 97.5% futuras observações, digamos  $Y_f$ , geraram-se 1000 novas observações, para cada caso, e contámos as vezes que  $Y_f$  estavam dentro dos contornos do elipsóide. Ou seja, para o caso em que  $\chi_{2,1} = \chi_{2,2} = 4.8$ , contámos o número de vezes que  $Y_f^\top \tilde{\mathbf{K}}_{0.975} Y_f \leq 20.7402$ . Os resultados obtidos estão contidos na Tabela 6.4. Pode verificar-se que, de uma maneira geral, as contagens não estão muito longe de 975.

Tabela 6.4: Número de vezes, em 1000, que  $Y_f$  está dentro dos contornos do elipsóide.

	$\chi_{2,1}$					
$\chi_{2,2}$	0.3	0.6	1.2	1.8	2.4	4.8
0.3	980	975	976	973	974	981
0.6	976	978	975	971	973	976
1.2	975	974	974	978	975	977
1.8	980	974	980	972	977	978
2.4	972	970	975	976	977	978
4.8	971	979	980	979	972	976

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

De forma a construir-se um intervalo de predição  $[Y_{fi}; Y_{fs}]$ , sendo  $i$  o limite inferior e  $s$  o limite superior, com uma probabilidade (aproximada) de 0.95 de conter uma futura observação,  $Y_f$ , tal como descrito em (3.3.38), considerou-se  $\hat{x} = [1]$ , ou seja, uma matriz com um único elemento igual a 1. Para  $\tilde{\beta}$  e para os quantis estimados  $\tilde{w}_{0.025}(\underline{u})$  e  $\tilde{w}_{0.975}(\underline{u})$ , indicados em (3.2.28), utilizaram-se as estimativas obtidas anteriormente. A Tabela 6.5 contém os limites inferiores médios e os limites superiores médios dos intervalos de predição, obtidos para cada caso, ao fim de 1000 simulações.

Tabela 6.5: Valores médios obtidos para  $[Y_{fi}; Y_{fs}]$ .

$\chi_{2,2}$	$\chi_{2,1}$					
	0.3	0.6	1.2	1.8	2.4	4.8
0.3	[-1.16; 1.19]	[-1.45; 1.44]	[-2.23; 2.29]	[-2.70; 2.36]	[-3.18; 3.11]	[-4.20; 4.20]
0.6	[-1.55; 1.51]	[-1.81; 1.87]	[-2.42; 2.50]	[-3.07; 3.06]	[-3.47; 3.46]	[-4.47; 4.47]
1.2	[-2.39; 2.35]	[-2.79; 3.11]	[-3.43; 3.43]	[-3.85; 3.85]	[-3.99; 3.99]	[-5.07; 5.07]
1.8	[-3.62; 3.62]	[-3.48; 3.36]	[-4.11; 4.12]	[-4.30; 4.30]	[-4.50; 4.51]	[-5.30; 5.30]
2.4	[-4.27; 4.27]	[-4.46; 4.46]	[-4.52; 4.52]	[-4.81; 4.81]	[-4.99; 4.99]	[-5.78; 5.78]
4.8	[-6.00; 6.00]	[-6.03; 6.03]	[-6.00; 6.00]	[-6.13; 6.13]	[-6.73; 6.73]	[-7.11; 7.11]

Os intervalos obtidos são, aparentemente, os que se esperaríamos obter, uma vez que têm todos o valor zero, aproximadamente, ao centro. Note-se que nas simulações se considerou  $\beta = \underline{0}$ , tendo os vetores aleatórios também vetor médio nulo. No entanto, para verificar se os intervalos construídos continham 95% de futuras observações, geraram-se 1000 novas observações  $Y_f$ , para cada caso, e, novamente, contámos as vezes que cada observação estava contida no respetivo intervalo  $[Y_{fi}; Y_{fs}]$ . Os resultados obtidos estão contidos na Tabela 6.6. Pode verificar-se que, de uma maneira geral, os valores não estão muito longe do valor 950.

Tabela 6.6: Número de vezes, em 1000, que  $Y_f$  estava contida em  $[Y_{fi}; Y_{fs}]$ .

$\chi_{2,2}$	$\chi_{2,1}$					
	0.3	0.6	1.2	1.8	2.4	4.8
0.3	872	854	945	916	942	945
0.6	907	918	947	950	951	946
1.2	952	978	966	980	962	954
1.8	989	970	989	975	975	968
2.4	990	990	979	980	993	971
4.8	991	990	983	987	979	979

### 6.2.2 Segundo caso

Esta subsecção pretende ilustrar a teoria apresentada na secção 5.1. Para tal, considerou-se o modelo descrito em (6.2.4), com  $\underline{\beta} = \underline{0}$  e  $\underline{Z}_i \sim \Gamma(a, b) - \frac{a}{b}$ ,  $i = 1, 2$ , com  $a$  tomando valores em  $\{0.3, 0.6, 1.2, 1.8, 2.4, 4.8\}$  e  $b = 1$ .

Tem-se assim que os vetores  $\underline{Z}_i$ ,  $i = 1, 2$ , são vetores aleatórios, com distribuição Gama, vetor médio nulo e variância  $\chi_{2,i}$ ,  $i = 1, 2$ , tomando valores em  $\{0.3, 0.6, 1.2, 1.8, 2.4, 4.8\}$ . Tal como no primeiro caso, tendo em conta as várias combinações para os  $\chi_{2,i}$ ,  $i = 1, 2$ , são considerados 36 vetores de observações,  $\underline{Y}$ .

Neste segundo caso tudo é simulado e estimado da mesma forma que no primeiro caso. Portanto, também aqui o procedimento foi repetido 1000 vezes. A única diferença é que, como já foi referido, agora os vetores de observações seguem distribuição gama, em vez de distribuição

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

normal. Logo, também neste segundo caso se conhecem as componentes da variância,  $\chi_{2,i}$ ,  $i = 1, 2$ . Assim, estas podem ser estimadas e as suas estimativas podem ser comparadas com os valores predefinidos. Neste caso, as estimativas foram obtidas a partir de (5.2.3). As médias dessas estimativas, para cada um dos 36 casos considerados, podem ser vistos na Tabela 6.7.

Tabela 6.7: Valores previamente considerados de  $\chi_{2,i}$ ,  $i = 1, 2$ , e médias das estimativas das componentes de variância obtidas.

		$\chi_{2,1}$					
$\chi_{2,2}$		0.3	0.6	1.2	1.8	2.4	4.8
0.3	$\tilde{\chi}_{1,2}$	0.3171	0.3087	0.3429	0.3033	0.2861	0.2952
	$\tilde{\chi}_{2,2}$	0.3008	0.3013	0.2997	0.3033	0.3009	0.3012
0.6	$\tilde{\chi}_{1,2}$	0.5953	0.6430	0.6028	0.6048	0.6601	0.5708
	$\tilde{\chi}_{2,2}$	0.6096	0.6090	0.5940	0.5879	0.5926	0.5909
1.2	$\tilde{\chi}_{1,2}$	1.0838	1.1289	1.1606	1.2008	1.1517	1.28057
	$\tilde{\chi}_{2,2}$	1.1870	1.2099	1.1997	1.2153	1.2066	1.1958
1.8	$\tilde{\chi}_{1,2}$	1.8517	1.7960	1.9114	1.8255	1.8932	1.9034
	$\tilde{\chi}_{2,2}$	1.7555	1.7818	1.7874	1.7975	1.8271	1.8108
2.4	$\tilde{\chi}_{1,2}$	2.2450	2.4245	2.4676	2.3385	2.7129	2.4745
	$\tilde{\chi}_{2,2}$	2.4230	2.3864	2.4469	2.3731	2.4113	2.3937
4.8	$\tilde{\chi}_{1,2}$	5.0440	5.0892	4.9497	4.6422	4.9647	4.9958
	$\tilde{\chi}_{2,2}$	4.7961	4.8335	4.8328	4.7178	4.8049	4.8615

Pode verificar-se que as estimativas médias estão muito perto dos valores previamente considerados.

A Tabela 6.8 apresenta os valores médios obtidos para  $\tilde{\beta}$ , descrito em (4.1.10), para cada um dos 36 casos considerados. Note-se que ao se simularem as observações se considerou  $\beta = 0$ .

Tabela 6.8: Médias obtidas para  $\tilde{\beta}$ .

		$\chi_{2,1}$					
$\chi_{2,2}$		0.3	0.6	1.2	1.8	2.4	4.8
0.3		-0.0039	-0.0060	-0.0009	-0.0001	-0.0013	-0.0005
0.6		-0.0074	-0.0035	-0.0037	-0.0021	-0.0010	-0.0004
1.2		-0.0063	-0.0038	-0.0031	-0.0016	-0.0017	-0.0008
1.8		-0.0033	-0.0031	-0.0016	-0.0012	-0.0006	-0.0002
2.4		-0.0020	-0.0019	-0.0017	-0.0014	-0.0009	-0.0005
4.8		-0.0007	-0.0004	-0.0004	-0.0006	-0.0006	-0.0005

Também neste caso se pode verificar que as estimativas médias estão muito perto dos valores previamente considerados.

Com vista à obtenção de elipsóides de confiança para observações futuras, consideraram-se os vetores já considerados na subsecção anterior. Assim, consideraram-se os vetores  $\underline{g}$ , com 6 componentes, podendo cada componente tomar os valores  $-0.9$ ,  $0.1$ , ou  $1.1$ , tomando-se

$$\underline{u} = \frac{1}{\|\underline{g}\|} \times \underline{g}. \quad (6.2.6)$$

Mais uma vez, para cada vetor  $\underline{u}$ , obtiveram-se os valores de  $\tilde{v}_2(\underline{u})$ ,  $\tilde{\gamma}_1(\underline{u})$  e  $\tilde{\gamma}_2(\underline{u})$ , apresentados em (3.2.25) e, de seguida, o valor correspondente para  $Ed_{\tilde{W}(\underline{u})}(w)$ , em (3.2.27), permitindo-nos, assim, obter os quantis estimados  $\tilde{w}_{0.025}(\underline{u})$  e  $\tilde{w}_{0.975}(\underline{u})$ , indicados em (3.2.28).

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

As matrizes  $\widetilde{\mathbf{K}}_p$  foram obtidas da forma anteriormente descrita. Neste caso, novamente, por exemplo, para  $\chi_{2,1} = \chi_{2,2} = 4.8$ , obteve-se

$$\widetilde{\mathbf{K}}_{0.975} = \begin{bmatrix} 8.9533 & 5.6927 & 5.7477 & 6.7337 & -1.2480 & -1.1888 \\ 5.6927 & 8.9630 & 5.6494 & -1.2460 & 6.6855 & -1.1204 \\ 5.7477 & 5.6494 & 8.9782 & -1.1740 & -1.1577 & 6.5915 \\ 6.7337 & -1.2460 & -1.1740 & 8.8557 & 5.7419 & 5.7256 \\ -1.2480 & 6.6855 & -1.1577 & 5.7419 & 8.9649 & 5.6917 \\ -1.1889 & -1.1204 & 6.5915 & 5.7256 & 5.6917 & 9.0261 \end{bmatrix}.$$

Para  $\chi_{2,1} = \chi_{2,2} = 4.8$ , a distância  $\underline{u}^T \widetilde{\mathbf{K}}_{0.975} \underline{u}$ , indicada na Figura 3.1, neste caso, foi de 20.7838.

A fim de se verificar se os elipsóides continham 97.5% de futuras observações,  $Y_f$ , geraram-se 1000 novas observações, para cada caso, e contou-se as vezes que  $Y_f$  estava dentro dos contornos do elipsóide. Ou seja, para o caso em que  $\chi_{2,1} = \chi_{2,2} = 4.8$ , contámos o número de vezes que  $Y_f^T \widetilde{\mathbf{K}}_{0.975} Y_f \leq 20.7838$ . Os resultados obtidos estão contidos na Tabela 6.9. Pode verificar-se que, de uma maneira geral, as contagens não estão muito longe de 975.

Tabela 6.9: Número de vezes, em 1000, que  $Y_f$  está dentro dos contornos do elipsóide.

	$\chi_{2,1}$					
$\chi_{2,2}$	0.3	0.6	1.2	1.8	2.4	4.8
0.3	978	978	974	973	971	977
0.6	974	978	976	979	972	978
1.2	977	972	978	973	972	972
1.8	971	978	979	976	977	979
2.4	977	975	975	979	978	979
4.8	970	974	979	979	979	973

A fim de se construir um intervalo de predição  $[Y_{fi}; Y_{fs}]$ , com uma probabilidade (aproximada) de 0.95 de conter uma futura observação,  $Y_f$ , descrito em (3.3.38), considerou-se, mais uma vez,  $\hat{x} = [1]$ . Para  $\hat{\beta}$  e para os quantis estimados  $\tilde{w}_{0.025}(\underline{u})$  e  $\tilde{w}_{0.975}(\underline{u})$ , indicados em (3.2.28), utilizaram-se as estimativas obtidas anteriormente. A Tabela 6.10 contém os limites inferiores médios e os limites superiores médios dos intervalos de predição, obtidos para cada caso, depois de se fazerem 1000 simulações.

Tabela 6.10: Valores médios obtidos para  $[Y_{fi}; Y_{fs}]$ .

	$\chi_{2,1}$					
$\chi_{2,2}$	0.3	0.6	1.2	1.8	2.4	4.8
0.3	[-1.62; 1.62]	[-1.82; 1.81]	[-2.30; 2.30]	[-2.55; 2.55]	[-2.80; 2.80]	[-3.79; 3.78]
0.6	[-2.08; 2.07]	[-2.31; 2.31]	[-2.53; 2.52]	[-2.89; 2.89]	[-3.20; 3.20]	[-4.03; 4.03]
1.2	[-2.65; 2.64]	[-2.89; 2.89]	[-3.16; 3.15]	[-3.44; 3.44]	[-3.60; 3.60]	[-4.52; 4.52]
1.8	[-3.38; 3.37]	[-3.46; 3.45]	[-3.77; 3.77]	[-3.95; 3.95]	[-4.17; 4.17]	[-4.85; 4.85]
2.4	[-3.70; 3.69]	[-3.96; 3.96]	[-4.15; 4.15]	[-4.27; 4.27]	[-4.70; 4.70]	[-5.24; 5.24]
4.8	[-5.47; 5.47]	[-5.57; 5.57]	[-5.63; 5.63]	[-5.61; 5.61]	[-5.92; 5.91]	[-6.51; 6.51]

Também neste caso os intervalos obtidos são, aparentemente, os que se esperariam obter, uma vez que, mais uma vez, têm todos o valor zero, aproximadamente, ao centro. Note-se que tam-

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

bém neste caso, aquando das simulações dos vetores de observações, se considerou  $\underline{\beta} = 0$ , tendo os vetores aleatórios também vetor médio nulo. A fim de se verificar se os intervalos construídos continham 95% de futuras observações, geraram-se 1000 novas observações  $Y_f$ , para cada caso e contámos as vezes que cada observação estava contida no respetivo intervalo  $[Y_{fi}; Y_{fs}]$ . Os resultados obtidos estão contidos na Tabela 6.11. Pode verificar-se que, também neste caso, de uma maneira geral, os valores não estão muito longe do valor 950, valor correspondente aos 95%.

Tabela 6.11: Número de vezes, em 1000, que  $Y_f$  estava contida em  $[Y_{fi}; Y_{fs}]$ .

	$\chi_{2,1}$					
$\chi_{2,2}$	0.3	0.6	1.2	1.8	2.4	4.8
0.3	964	939	934	919	916	908
0.6	969	959	948	940	923	915
1.2	969	970	947	950	933	924
1.8	981	974	969	958	959	925
2.4	978	975	970	958	973	945
4.8	992	985	978	980	971	962

### 6.3 Exemplo de Aplicação de modelos múltiplos aditivos

Modelos aditivos são considerados homólogos quando têm as mesmas matrizes  $X, X_1, \dots, X_w$ . Nesse caso, constituem um modelo múltiplo aditivo, se corresponderem a tratamentos de um delineamento base. Assim, os modelos múltiplos aditivos podem ser utilizados para ter uma análise conjunta de matrizes de modelos homólogos. Estes podem, por exemplo, ter como delineamento base blocos casualizados. Esses delineamento base são utilizados no melhoramento de plantas onde o objetivo é selecionar culturas (variedades cultivadas), com bom comportamento em grandes áreas.

Vamos supor que estamos a comparar  $k$ , culturas, em  $m$  locais. Assim, temos os estimadores  $\tilde{\mu}_{j,l}, j = 1, \dots, k, l = 1, \dots, m$ , da produção média da  $j$ -ésima cultura no  $l$ -ésimo local, obtido a partir das produções do  $l$ -ésimo ensaio. Se considerarmos que as variâncias do erro para os diferentes ensaios não diferem significativamente, os locais irão corresponder aos tratamentos de um delineamento base. Podemos, então, estudar a ação dos fatores do delineamento base nas  $\mu_j = E(\tilde{\mu}_j), j = 1, \dots, k$ , com  $\tilde{\mu}_j = (\tilde{\mu}_{j,1}, \dots, \tilde{\mu}_{j,m}), j = 1, \dots, k$ . Se os  $u$  fatores do delineamento base cruzarem, podemos realizar uma análise conjunta, ver secção 3.4. .

Realçamos que os locais podem pertencer a diferentes regiões e que os ensaios podem ser realizados em anos sucessivos. Então, teríamos dois fatores que cruzam, ou seja, região e ano, respetivamente.

Note-se que poderemos realizar uma abordagem por duas etapas, começando com ensaios individuais e, em seguida, realizar o estudo de regiões e anos. A razão por trás desta opção, é que os blocos não eram os mesmos para os diferentes ensaios. Realçamos ainda que os ensaios de comparação entre culturas, são realizados em locais representativos de regiões agrícolas para que possamos relacionar as suas localidades com as regiões a que pertencem. Então, teríamos os fatores  $L$ , de localização, e  $T$ , de tempo.

Como ambos os fatores têm efeitos fixos podemos realizar uma ANOVA simples de dois fatores.

## **Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição**

Os nossos desenvolvimentos permitem-nos considerar ensaios individuais mais complexos com fatores de efeitos aleatórios. No entanto, o exemplo simples que apresentamos destaca a abordagem em duas etapas:

- na primeira fase estudam-se os ensaios individuais;
- na segunda fase estudamos a ação dos fatores do delineamento base dessas componentes usando a técnica bem conhecida denominada por ANOVA.

# Capítulo 7

## Observações Finais

### 7.1 Revisão Global

Começámos pelo caso geral com modelos da forma

$$\underline{Y} = \underline{X}\underline{\beta} + \sum_{i=1}^w \underline{X}_i \underline{Z}_i, \quad (7.1.1)$$

onde o vetor de observações  $\underline{Y}$  é do tipo  $n \times 1$ , com matrizes  $\underline{X}, \underline{X}_1, \dots, \underline{X}_w$ , dadas, sendo  $\underline{\beta}$  o vetor dos coeficientes a estimar e as componentes  $\underline{Z}_1, \dots, \underline{Z}_w$ , são independentes com componentes *i.i.d.*, com:

- valores médios nulos;
- cumulantes  $\chi_{r,i}, r = 2, 3, 4, i = 1, \dots, w$ .

Para estimar os cumulantes de quarta ordem, admitimos ter um par de modelos  $(\underline{Y}(1), \underline{Y}(2))$  *i.i.d.*. A partir dos estimadores dos cumulantes  $\chi_{r,i}, r = 2, 3, 4, i = 1, \dots, w$ , obtivemos estimadores para os cumulantes de

$$W(\underline{u}) = \underline{u}^\top (\underline{Y} - \underline{\mu}), \quad (7.1.2)$$

com  $\underline{u} = \underline{X}\underline{\beta}$  e para os cumulantes de

$$\dot{W}(\underline{u}) = \frac{W(\underline{u})}{\sqrt{\tilde{v}_2(\underline{u})}}. \quad (7.1.3)$$

(Recordemos que  $\tilde{v}_2(\underline{u})$  representa o segundo cumulante de  $W(\underline{u})$ ).

Passámos à obtenção da Expansão de Edgeworth para a distribuição  $F_{\dot{W}(\underline{u})}$ , de  $\dot{W}(\underline{u})$ , e, a partir daí, às aproximações dos quantis  $\tilde{w}_p(\underline{u})$  da distribuição  $F_{\dot{W}(\underline{u})}$  e  $\tilde{w}_p(\underline{u})$ , da distribuição  $F_{W(\underline{u})}$ . Dada uma matriz  $\mathbf{K} = [k_{l,h}]$ , considerámos o seu semi-vec, dado por

$$\underline{\lambda} = (k_{1,1} \dots 2k_{1,n}, k_{2,2} \dots 2k_{2,n}, \dots, k_{n,n}), \quad (7.1.4)$$

e representámos por  $\mathbf{K}(\underline{\lambda})$  a matriz com semi-vec  $\underline{\lambda}$ .

Depois, considerámos um vetor  $\underline{u} = (u_1, \dots, u_n)$ , com  $\underline{u}^{[2]} = (u_1^2, \dots, 2u_1 u_n, u_2^2, \dots, 2u_2 u_n, \dots, u_n^2)$ , e

$$\underline{u}^\top \mathbf{K}(\underline{\lambda}) \underline{u} = \underline{u}^{[2]\top} \underline{\lambda}, \quad (7.1.5)$$

o que nos permitiu reescrever as formas quadráticas como produtos internos. Assim, ajustámos os Elipsóides de Confiança. De seguida, considerámos as componentes  $w_p(\underline{u}_1), \dots, w_p(\underline{u}_m)$ , e minimizámos

$$\Delta(\mathbf{K}) = \sum_{l=1}^m (w_p(\underline{u}_l) - \underline{u}_l^\top \mathbf{K} \underline{u}_l)^2, \quad (7.1.6)$$

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

pelo que, com

$$\underline{w}_p(\underline{u}) = (w_p(\underline{u}_1), \dots, w_p(\underline{u}_n)), \quad (7.1.7)$$

$$\mathbf{U}^{[2]} = [\underline{u}_1^{[2]}, \dots, \underline{u}_n^{[2]}], \quad (7.1.8)$$

obtivémos o *LSE*

$$\tilde{\lambda}_p = [\mathbf{U}^\top \mathbf{U}]^+ \mathbf{U}^\top \tilde{w}_p(\underline{u}). \quad (7.1.9)$$

Seguiu-se a construção de Intervalos de Predição. Estes intervalos foram construídos para conjuntos de valores das variáveis controladas correspondentes às colunas das matrizes  $\mathbf{X}, \mathbf{X}_1, \dots, \mathbf{X}_w$ , constituindo vetores linha  $\hat{x}, \hat{x}_1, \dots, \hat{x}_w$ , adicionais para as mesmas. Considerámos  $\hat{u} = [\hat{x}^\top \hat{x}_1^\top \dots \hat{x}_w^\top]^\top$  e  $\underline{u} = \frac{1}{\|\hat{u}\|} \hat{u}$ , e obtivémos a Expansão de Edgeworth,  $Ed_{\hat{W}(\underline{u})}$ , para a distribuição  $F_{\hat{W}(\underline{u})}$  e, a partir daí, amostras com uma determinada distribuição. A partir dos quantis empíricos dessa distribuição, obtivémos os intervalos de predição.

Em seguida, considerámos Modelos Múltiplos. Os pares de modelos duma dessas famílias têm as mesmas matrizes  $\mathbf{X}, \mathbf{X}_1, \dots, \mathbf{X}_w$  e os cumulantes  $\chi_{r,i}, r = 2, 3, 4, i = 1, \dots, w$ , são os mesmos. A partir desses pares de modelos, obtivémos, para os  $d$  tratamentos, pares  $(\tilde{\psi}_h(1), \tilde{\psi}_h(2)), h = 1, \dots, d$ , de estimadores para a função estimável  $\psi$ .

Considerámos  $\tilde{\psi}(l) = (\tilde{\psi}_1(l), \dots, \tilde{\psi}_d(l)), l = 1, 2$ , bem como  $\tilde{\psi}_{l,\bullet} = \frac{1}{2}(\tilde{\psi}_{l,1} + \tilde{\psi}_{l,2}), l = 1, \dots, d$ .

Verificámos que, dado as hipóteses a testar corresponderem aos sub-espacos  $\nabla_1, \dots, \nabla_m$ , duma partição ortogonal de  $\mathbb{R}^d$ , existem matrizes  $\mathbf{A}_1, \dots, \mathbf{A}_m$ , cujos vetores linha constituem bases ortonormadas para esses sub-espacos. Com  $g_j = \text{car}(\mathbf{A}_j)$ , verificámos que temos

$$\begin{cases} S_j = \|\mathbf{A}_j \tilde{\psi}_{l,\bullet}\|^2 \\ S = \|\tilde{\psi}(1) - \tilde{\psi}(2)\|^2 \end{cases}, j = 1, \dots, m, l = 1, \dots, d. \quad (7.1.10)$$

Obtivémos os testes F com estatísticas

$$F_j = \frac{2d S_j}{g_j S}, j = 1, \dots, m, \quad (7.1.11)$$

com  $g_j$  e  $d$  graus de liberdade, tal como no capítulo 4.

Seguidamente admitimos que as componentes dos  $\underline{z}_i, i = 1, \dots, w$ , tinham parâmetros de localização,  $\lambda_i$ , dispersão,  $\delta_i$  e forma,  $\tau_i, i = 1, \dots, w$ .

Os seus cumulantes de ordem  $r$  são dados por:

$$\chi_{r,i}(\lambda_i, \delta_i, \tau_i) = \delta_i^r g_{r,i}(\tau_i), i = 1, \dots, w, r \geq 2. \quad (7.1.12)$$

Por outro lado, integrámos, como vimos, os  $\lambda_1, \dots, \lambda_w$ , no vetor dos coeficientes. Assim, a partir dos cumulantes, estimámos os  $(\delta_i, \tau_i), i = 1, \dots, w$ , e, a partir da estimação de

$$\underline{\beta} = [\underline{\beta}^\top \lambda_1 \dots \lambda_w]^\top, \quad (7.1.13)$$

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

estimámos não só esse vetor  $\underline{\beta}$ , mas também os  $\lambda_1, \dots, \lambda_w$ .

Obtivémos, então, os

$$\tilde{\chi}_{4,i} = \chi_{4,i}(\tilde{\lambda}_i, \tilde{\delta}_i, \tilde{\tau}_i), \quad (7.1.14)$$

não sendo necessário ter pares de modelos.

Passando aos Modelos Múltiplos (como não tínhamos pares de modelos) admitimos, à priori, que existia um conjunto

$$\begin{cases} g = \sum_{j \in \varphi} g_j \\ S = \sum_{j \in \varphi} S_j \end{cases}, \quad (7.1.15)$$

e, para testar  $H_{0,j}$ ,  $j \notin \varphi$ , considerámos as estatísticas

$$F_j = \frac{g}{g_j} \frac{S_j}{S}, \quad j \notin \varphi, \quad (7.1.16)$$

com  $g_j$  e  $g$  graus de liberdade.

Destacamos os sub-casos em que considerámos que as distribuições das componentes dos vetores  $\underline{Z}_i, i = 1, \dots, w$ , tinham parâmetros de localização e dispersão. Particularizámos para os casos das distribuições Normal, Exponencial e Gumbel. Por último, admitimos que as distribuições das componentes dos vetores  $\underline{Z}_i, i = 1, \dots, w$ , tinham distribuições Gamma, Weibull e Fréchet, apresentando os cumulantes  $\chi_{r,i}(0, 1, \tau_i), r = 2, 3, 4$  e  $i = 1, \dots, w$ .

Na fase inicial apenas estimámos o  $\chi_{2,i}, i = 1, \dots, w$ , e o resto do tratamento foi paralelo ao do caso anterior.

## 7.2 Conclusões e Trabalhos Futuros

Chegados ao fim da nossa apresentação dos modelos aditivos, verificámos que estes, apesar da grande generalidade de pressupostos em que assentam, permitiram obter resultados, nomeadamente, para

- Elipsóides de Confiança;
- Intervalos de Predição;
- Testes de hipóteses para Famílias Estruturadas de modelos.

Observe-se ainda que, à medida que se reforçam os pressupostos relativos às distribuições das componentes dos vetores  $\underline{Z}_1, \dots, \underline{Z}_w$ , se aprofunda o tratamento. Assim, de pares *i.i.d.* de modelos, passamos a modelos isolados e, a ordem dos cumulantes das distribuições dos componentes  $\underline{Z}_1, \dots, \underline{Z}_w$ , que estimamos diretamente, vai diminuindo. Resumindo, os modelos aditivos representam um ponto de equilíbrio entre a generalidade dos mesmos e o interesse dos resultados que permitem obter.

Como continuação natural do nosso trabalho pretendemos desenvolver o estudo de casos mistos em que as distribuições das componentes dos  $\underline{Z}_1, \dots, \underline{Z}_w$ , pertencem a tipos distintos. Em particular, as de  $\underline{Z}_w$  poderão ser normais. Pretende-se-á então aproveitar a bem conhecida ligação das distribuições normais aos erros de medida para aprofundar o efeito perturbador de  $\underline{Z}_w$  em modelos aditivos.



## Bibliografia

- [1] Andrews, D. F. (2001). Asymptotic expansions of moments and cumulants. *Statistics Computing*, 11(1), 7-16.
- [2] Antunes, P., Ferreira, S. S., Ferreira, D., Nunes, C. & Mexia, J. T. (2020a). Estimation in additive models and ANOVA-like applications. *Journal of Applied Statistics*, 47, 1- 10. <https://doi.org/10.1080/02664763.2020.1723501>
- [3] Antunes, P., Ferreira, S. S., Ferreira, D. & Mexia, J. T. (2020b). Multiple Additive Models. *Communications in Statistics - Theory and Methods*, 49. <https://doi.org/10.1080/03610926.2020.1723636>
- [4] Balakrishnan, N., Johnson, N. L. & Kotz, S. (1998). A note on relationships between moments, central moments and cumulants from multivariate distributions. *Statistics & Probability Letters*, 39(1), 49 - 54. [https://doi.org/10.1016/S0167-7152\(98\)00027-3](https://doi.org/10.1016/S0167-7152(98)00027-3)
- [5] Balakrishnan, N. & Basu, A. P. (2019). *The Exponential distribution: Theory, Methods and Applications*. CRC Press, Taylor & Francis Group.
- [6] Belsey, D. A., Kuh, E. & Welsch, R. E. (1980). *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. John Wiley & Sons, Inc.
- [7] Bryk, A. S. & Raudenbush, S. W. (1992). *Hierarchical linear models in social and behavioral research: Applications and data analysis methods* (1st ed.). Newbury Park, CA: Sage.
- [8] Cramér, H. (1946). In *Mathematical Methods of Statistics*. Princeton University Press.
- [9] Cheng, C. S., Deng, L. Y. & Tang, B. (2002). Generalized minimum aberration and design efficiency for nonregular fractional factorial designs. *Statistica Sinica*, 12(4), 991-1000. <https://www.jstor.org/stable/24307009>
- [10] Christensen, R. (2011). *Plane answers to complex questions: the theory of linear models*. Springer. <https://doi.org/10.1007/978-1-4419-9816-3>
- [11] Coorey, K. (2010). Generalized Gumbel distribution. *Journal of Applied Statistics*, 37(1), 171-179. <https://doi.org/10.1080/02664760802698995>
- [12] Craig, C. C. (1931). On A Property of the Semi-Invariants of Thiele. *Annals of Mathematical Statistics*, 2(2), 154-164. <https://doi.org/10.1214/aoms/1177732994>
- [13] DasGupta, A. (2008) *Asymptotic Theory of Statistics and Probability*, Springer.
- [14] Eisenhart, C. (1947). The assumptions underlying the analysis of variance, *Biometrics* (3),

1â21. <https://doi.org/10.2307/3001534>

- [15] Edgeworth, F. (1905). The law of error, *Cambridge philosophical Society Transactions*, 20(1), 37-65.
- [16] Edgeworth, F. (1906). The Generalised Law of Error, or Law of Great Numbers, *Journal of the Royal Statistical Society*, 69(3), 497-539. <https://doi.org/10.2307/2339343>
- [17] Ferreira, S. S., Ferreira, D. & Mexia, J.T. (2006). Cross Additivity-An Application. *Discussiones Mathematicae - Probability and Statistics*, 26, 207-219. <http://www.discuss.wmie.uz.zgora.pl/ps/index.php?doi=10.7151/dmps.1083>
- [18] Ferreira, S., Ferreira, D., Moreira, E. & Mexia, J.T. (2010). Exact estimators for normal linear mixed models. *Far East Journal of Theoretical Statistics*, 31(1), 33-48. <http://hdl.handle.net/10400.6/9124>
- [19] Fisher, R. A. (1929). Moments and Product Moments of Sampling Distributions. *Proceedings of the London Mathematical Society*, 30(1), 199-238. <https://doi.org/10.1112/plms/s2-30.1.199>
- [20] Fisher, R.A. & John Wishart, J. (1932). The derivation of the pattern formulae of two-way partitions from those of simpler patterns. *Proceedings of the London Mathematical Society*, s2-33 (1), 195-208 <https://doi.org/10.1112/plms/s2-33.1.195>
- [21] Fonseca, M., Mexia, J. T. & Zmyślony, R. (2003). Estimators and tests for variance components in cross-nested orthogonal models. *Discussiones Mathematicae, Probability and Statistics*, 23(2), 175-201. <https://doi.org/10.7151/DMPS.1043>
- [22] Fonseca, M., Mexia, J. T. & Zmyślony, R. (2009). Binary Operations and canonical forms for factorial and related models. *Linear Algebra and its Applications*, 430(10), 2781-2797. <https://doi.org/10.1016/j.laa.2009.01.013>
- [23] Friedman, J. H. & Stuetzle, W. (1981). Projection pursuit regression. *Journal of the American Statistical Association*, 76(376), 817-823. <https://doi.org/10.2307/2287576>
- [24] Graham, A. (2018) *Kronecker Products and Matrix Calculus with Applications*. Dover Publications, Inc.
- [25] Gumbel, E. J. (1965). A quick estimation of the parameters in Fréchet's distribution. *Review of the International Statistical Institute*, 33 349-363. <https://doi.org/10.2307/1401700>
- [26] Henderson, C. (1953). *Estimation of Variance and Covariance Components*. *Biometrics*, 9 (2), 226-252. <https://doi:10.2307/3001853>
- [27] Hoggart, S. & Hawkins, S. (2015). Ecological Approaches to Coastal Risk Mitigation. In S., Hoggart, R., Thompson, B., Zanuttigh, R., Nicholls, J.P., Vanderlinden, & H., Burcharth (Eds), *Coastal Risk Management in a Changing Climate* (pp.171-236). Elsevier.

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

- [28] Jesus, V. (2008). *Jordan algebras and crossing of factorial and fractional replicates* (Dissertação de Doutoramento em Matemática com especialização em Estatística, não publicada]. Faculdade Ciências e Tecnologia - Universidade Nova de Lisboa.
- [29] Jesus, V., Mexia, J. T., Fonseca, M. & Zmyślony, R. (2009). Binary Operations and Canonical Forms for Factorial and Related Models. *Linear Algebra and Its Applications*, 430(10), 2781-2797. <https://doi.org/10.1016/j.laa.2009.01.013>
- [30] Kantar, Y. M. (2015). Generalized Least Squares and Weighted Least Squares Estimation Methods for Distributional Parameters. *Revstat*, 13(3), 263-282. <https://www.ine.pt/revstat/pdf/rs150306.pdf>
- [31] Kariya, T. & Kurata, H. (2004). *Generalized Least Squares*. John Wiley & Sons, Ltd.
- [32] Kendall, M. G., Stuart, A. & Ord, J. K. (1987). *Kendall's Advanced Theory of Statistics*. London: Griffin.
- [33] Khuri, A. & Sahai, H. (1985). Variance components analysis: a selective literature survey. *International Statistical Review*. (53) 279-300.
- [34] Khuri, A. I., Mathew, T. & Sinha, B. K. (1998). *Statistical Tests for Mixed Linear Models*. Wiley Interscience Publication, John Wiley & Sons, Inc.
- [35] Kolassa, J. E. (1994). *Series Expansion Methods in Statistics*. Springer-Verlag.
- [36] Lakshmi, R. V. & Vaidyathan, V. S. (2016). Three-parameter gamma distribution: Estimation using likelihood, spacings and least squares approach. *Journal of Statistics and Management Systems*, 19(1), 37-53. <https://doi.org/10.1080/09720510.2014.986927>
- [37] Lee, Y. & Nelder, J. A. (1996). Hierarchical generalized linear models *Journal of the Royal Statistical Society, Series B.* , 58 (4), 619-678. <http://www.jstor.org/stable/2346105>
- [38] Lee, K. J. & Thompson, S. G. (2008). Flexible parametric models for random-effects distributions. *Statistics in Medicine*, 27(3), 418-434. <https://doi.org/10.1002/sim.2897>
- [39] Madan, D. B. (2014). Estimating Parametric Models of Probability distributions. *Methodology and Computing in Applied Probability*, 17, 823-831. <http://doi.org/10.2139/ssrn.2635613>
- [40] Malley, J. D. (1994). *Statistical Applications of Jordan Algebras*. Springer-Verlag.
- [41] Martins, J. P. O. (2008). *Feira dos momentos: Planeamento e experimental e Investigação de localização e Escala em Populações não Gaussianas* [Dissertação de Doutoramento em Estatística e Investigação Operacional, Universidade de Lisboa-Faculdade de Ciências]. Repositório Científico da Universidade de Lisboa. <http://hdl.handle.net/10451/1591>
- [42] McCulloch, C. & Searle, S. (2001). *Generalized, Linear and Mixed Models*. New York: John Wiley & Sons, Inc.

- [43] Mexia, J. T. (1987). *Multi-treatment regression designs*, Research Studies, 1, Mathematics Department, Faculty of Sciences and Technology, Universidade Nova de Lisboa.
- [44] Mexia, J. T. (1988). *Standardized Orthogonal Matrices and the Decomposition of the sum of Squares for Treatments*, Research Studies, 2, Mathematics Department, Faculty of Sciences and Technology. Universidade Nova de Lisboa.
- [45] Mexia, J.T. (1989). *Controlled Heteroscedasticity, Quocient Vector Spaces and F Tests for Hypothesis on Mean Vectors*, Research Studies, 1, Mathematics Department, Faculty of Sciences and Technology. Universidade Nova de Lisboa.
- [46] Mexia, J.T. (1995). *Introdução à Inferência Estatística Linear*. Centro de Estudos de Matemática Aplicada. Edições Universitárias Lusófonas.
- [47] Moreira, E., Ribeiro, A., Mateus, E., Mexia, J. T. & Ottosen, L. (2005). Regressional modeling of electrolytic removal of Cu, Cr and As from CCA treated timber waste: application to sawdust. *Wood Science and Technology*, 39(4), 291-305. <https://doi.org/10.1007/s00226-004-0267-z>
- [48] Neyman, J. (1956). Note on an Article by Sir Ronald Fisher. *Journal of the Royal Statistical Society, Series B (Methodological)*, 18(2), 288-294. <https://www.jstor.org/stable/2983716>
- [49] Novak, J. & Sniady, P. (2011). What Is a Free Cumulant? *Notices of the American Mathematical Society*, 58(2), 300-301. <https://pdfs.semanticscholar.org/a49d/0c1c8263b35f05a5b6877575df2f5dffe561.pdf>
- [50] Oliveira, S. (2005). *Delineamentos regressionais múltiplos com base factorial prima*. [Dissertação de mestrado em Estatística e Optimização, Universidade Nova de Lisboa-Faculdade de Ciências e Tecnologia]. Repositório Científico da Universidade Nova de Lisboa. <http://hdl.handle.net/10362/45217>
- [51] Oliveira, S. (2018). *Delineamento regressional múltiplo para um factorial de base prima estritamente associado a uma Álgebra de Jordan Comutativa*. [Dissertação de doutoramento em Estatística e Gestão do Risco, Universidade Nova de Lisboa-Faculdade de Ciências e Tecnologia]. Repositório Científico da Universidade Nova de Lisboa. <http://hdl.handle.net/10362/45217>
- [52] Olver, P. J. (2000). *Applications of Lie Groups to Differential Equations*. Springer Science & Business Media.
- [53] Patterson, H.D. & Thompson, R. (1971). Recovery of inter-block information when block sizes are unequal, *Biometrika* (58) 545-554.
- [54] Pinheiro, J. C. & Bates, D. M. (2000), *Mixed-Effects Models in S and S-PLUS*, Springer, New York.
- [55] Istone, G. & Wynn, H.P. (2006). Cumulant varieties. *Journal of Symbolic Computation*, 41(2),

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

210-221. <https://doi.org/10.1016/j.jsc.2005.04.008>

- [56] R Core Team (2013). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. <http://www.R-project.org/>
- [57] Ribeiro, A. & Mexia, J. T. (1997). A dynamic model for the electrokinetic removal of copper from a polluted soil. *Journal of Hazardous Materials*, 56(3), 257-271. [https://doi.org/10.1016/S0304-3894\(97\)00060-5](https://doi.org/10.1016/S0304-3894(97)00060-5)
- [58] Rinne, H. (2009). *The Weibull distribution, A Handbook*. CRC Press, Taylor & Francis Group.
- [59] Scheffé, H. (1959). *The analysis of variance*. Wiley Classics Series, John Wiley & Sons.
- [60] Schott, J.R. (2017). *Matrix Analysis for Statistics-Third Edition*. John Wiley & Sons.
- [61] Searle, S.R., Casella, G. & McCulloch, C.E. (1992). *Variance Components*. John Wiley & Sons.
- [62] Seely, J. (1970a). Linear spaces and unbiased estimation. *The Annals of Mathematical Statistics*, 41(5), 1725-1734. <https://projecteuclid.org/euclid.aoms/1177696817>
- [63] Seely, J. (1970b). Linear spaces and unbiased estimation- An application to the mixed linear model. *The Annals of Mathematical Statistics*, 41(5), 1735-1748. <https://projecteuclid.org/euclid.aoms/1177696818>
- [64] Seely, J. (1971). Quadratic subspaces and completeness. *The Annals of Mathematical Statistics*, 42(2), 710-721. <https://projecteuclid.org/euclid.aoms/1177693420>
- [65] Silvey, S.D. (1980). *Optimal Design: An Introduction to the Theory for Parameter Estimation*. Chapman and Hall.
- [66] Smith, P. J. (1995). A Recursive Formulation of the Old Problem of Obtaining Moments from Cumulants and Vice Versa. *The American Statistician*, 49(2), 217-218.
- [67] Steeb, W. H. & Hardy, Y. (2011). *Matrix calculus and Kronecker Product: a practical approach to linear and multilinear algebra*. World Scientific Publishing Company.
- [68] Stuart, A. & Ord, K. A. (1994). *Kendall's Advanced Theory of Statistics, Volume 1: Distribution Theory (6<sup>th</sup> Edition)*. Wiley-Blackwell.
- [69] Thiele, T. N. (1931). Theory of Observations. *The Annals of Mathematical Statistics*, 2, 165-308.
- [70] Walpole, R. E., Myers, R. H., Myers, S. L. & Ye, K. (2012). *Probability & Statistics for Engineers & Scientists*. Prentice Hall.
- [71] Withers, C. S., Nadarajah, S. & Shih, S. H. (2015). Moments and Cumulants

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

of a Mixture. *Methodology and Computing in Applied Probability*, 17(3), 541-564.  
<https://doi.org/10.1007/s11009-013-9379-y>

- [72] Ye, K. Q. (2003). Indicator function and its application in two-level factorial designs. *The Annals of Statistics*, 31(3), 984-994. <https://projecteuclid.org/euclid.aos/1056562470>
- [73] Zheng, Q. (2002). Computing relations between moments and cumulants, *Computational Statistics*, 17(4), 507-515. <https://doi.org/10.1007/s001800200123>

## Apêndice A

### Anexos

#### A.1 Álgebras de Jordan

As álgebras de Jordan são importantes e foram introduzidas, em 1932, pelo físico alemão Pascual Jordan, numa tentativa de formular a base mecânica quântica, em termos do produto de Jordan em vez do produto associado. Os primeiros registos de álgebras de Jordan datam de 1934, quando Pascual Jordan, John von Neumann e Eugene Wigner, publicaram o artigo *On an algebraic generalization of the quantum mechanical formalism*, onde deduziram as propriedades das matrizes no contexto da mecânica quântica, sem o auxílio da álgebra matricial subjacente, mas não observável. De referir que Seely (1970a, 1970b) [62] [63], designou as álgebras de Jordan Comutativas como espaços vectoriais quadráticos. Desde então, as álgebras de Jordan têm sido aplicadas em diversos ramos da Matemática, por diversos autores, citando apenas Malley, J. D. (1994) [40], a título de exemplo.

De seguida serão apresentadas algumas definições importantes sobre álgebras de Jordan.

**Definição A.1** *Um conjunto  $\mathcal{A}$  é chamado de uma álgebra sobre o corpo  $k$ , se  $\mathcal{A}$  possui uma estrutura de  $k$ -espaço vectorial e uma aplicação bilinear,  $\otimes$ , geralmente designada por multiplicação, que verifica as seguintes condições, para todo  $x, y, z \in \mathcal{A}$  e  $\alpha \in \mathbb{R}$ :*

$$a) \quad x \otimes (y + z) = x \otimes y + x \otimes z;$$

$$b) \quad (x + y) \otimes z = x \otimes z + y \otimes z;$$

$$c) \quad \alpha(x \otimes y) = (\alpha x) \otimes y = x \otimes (\alpha y).$$

**Definição A.2** *Seja  $\mathcal{A}$  uma álgebra de Jordan e  $S \subseteq \mathcal{A}$ .  $S$  é uma sub-álgebra de Jordan de  $\mathcal{A}$  se:*

a) *é uma sub-álgebra de  $\mathcal{A}$ ;*

b) *para todo o  $x, y \in \mathcal{A}$  :  $x \otimes y \in S$ .*

**Definição A.3** *Uma álgebra de Jordan é uma álgebra munida com o produto de Jordan, que satisfaz as condições seguintes:*

a)  $x.y = y.x$  *Lei comutativa*

b)  $x^2.(y.x) = (x^2.y).x$ , *Identidade de Jordan,*

com  $x^2 = x.x$ , para todo  $x, y \in \mathcal{A}$ .

### A.1.1 Álgebras de Jordan Comutativas

Temos a seguinte proposição, de acordo com Schott (2017, pág.156) [60]:

**Proposição A.1** *Sejam  $A_1, \dots, A_m$  matrizes simétricas  $k \times k$ . Então existe uma matriz ortogonal  $P$  tal que,  $P^\top A_i P = D_i$  é uma matriz diagonal, com os valores próprios de  $A_i$  como elementos principais, se, e somente se  $A_i A_j = A_j A_i$ , para todos os pares  $(i; j)$ ;  $i, j = 1, \dots, m$ .*

Desta forma,  $A_1, \dots, A_m$ , pertencem à família  $\mathcal{D}(P)$ , de matrizes diagonalizadas por  $P$ , que é também uma *AJC*. Logo, as matrizes de  $\mathcal{A} = \{A_1, \dots, A_m\}$  pertencem a uma *AJC* se e somente se elas comutam.

Todas as *AJC* de matrizes simétricas têm uma só base, a base principal,  $bp(\mathcal{A})$ , constituída por *MPO*, que são mutuamente ortogonais, ver Seely, J. (1971) [64].

**Definição A.4 (Álgebra de Jordan Comutativa)** *Um espaço linear de matrizes reais simétricas,  $\mathcal{A}$ , diz-se uma *AJC*, se além de ser uma álgebra Jordan, dadas quaisquer matrizes  $A$  e  $B$  de  $\mathcal{A}$ , se tiver  $AB = BA$ .*

**Definição A.5** *Uma *AJC* é uma matriz de Jordan composta por matrizes simétricas que comutam e contém os quadrados das suas matrizes.*

**Proposição A.2** *Uma condição necessária e suficiente para que um subespaço  $S \subseteq S_n$  seja uma *AJC* parte da existência de uma base,  $Q = \{Q_1, \dots, Q_m\}$ , para  $S$ , constituída por matrizes idempotentes, com  $Q_i Q_j = 0$ , para  $i \neq j, i, j = 1, \dots, m$ , sendo essa base considerada única.*

Se  $Q$  for uma *MPO* pertencente à *AJC*, com base principal  $Q = \{Q_1, \dots, Q_m\}$ , teremos:

$$Q = \sum_{j=1}^m b_j Q_j.$$

Como  $Q$  é idempotente, tem-se  $b_j = 1, j = 1, \dots, m$ .

Assim, temos

$$Q = \sum_{j \in \mathcal{D}} Q_j; \mathcal{D} = \{j : b_j \neq 0\},$$

isto é, as *MPO* duma *AJC* são a soma das matrizes da sua base principal onde a *MPO* de uma *AJC* é a soma direta de matrizes das suas bases principais.

Seguem-se as definições de *AJC* regular, *AJC* completa e a *AJC* complementar.

**Definição A.6 (*AJC* regular)**

$\mathcal{A}$  é regular se  $\frac{1}{n} J_n = \frac{1}{n} \mathbf{1}_n \mathbf{1}_n^\top \in \mathcal{A}$  (considerando  $Q_1 = \frac{1}{n} J_n$ ).

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

### Definição A.7 (AJC completa)

Álgebra completa é aquela que contém a matriz identidade,  $I_n$ ; então, a matriz identidade será a soma das matrizes da base principal.

### Definição A.8 (AJC complementar)

Se  $\sum_{j=1}^w g_j < n$ , ou  $\sum_{j=1}^w Q_j \neq I_n$ , podemos adicionar  $Q_{w+1} = I_n - \sum_{j=1}^w Q_j$  obtendo-se então a base principal de uma AJC completa,  $\bar{\mathcal{A}}$ . Dizemos que  $\bar{\mathcal{A}}$  é a complementada de  $\mathcal{A}$ .

**Proposição A.3** Se  $\mathcal{A}_1 \subset \mathcal{A}_2$  e  $\mathcal{A}_1$  é completa, então  $\mathcal{A}_2$  também é completa.

## A.2 Momentos

Os momentos de uma variável aleatória reproduzem o seu comportamento probabilístico, ver Martins (2008, págs. 6-7) [41]. Temos:

**Definição A.9 (Momentos):** Se  $X$  tem distribuição  $F_X$ , o seu  $r$ -ésimo momento relativo à origem é dado por:

$$\mu'_r(X) = E[X^r] = \int_{-\infty}^{\infty} x^r dF_X, \quad (\text{A.2.1})$$

quando este integral converge. Em particular, caso esteja definido,  $\mu(X) = \mu'_1(X)$  será o valor médio de  $X$ .

Segue-se a

**Definição A.10 (Momentos centrais):** Caso  $\mu = \mu(X)$  esteja definido, o  $r$ -ésimo momento central de  $X$  é dado por:

$$\mu_r(X) = E[(X - \mu)^r] = \int_{-\infty}^{\infty} (X - \mu)^r dF_X. \quad (\text{A.2.2})$$

O momento central de ordem 0 está sempre definido e é igual a 0. Se  $\mu$  estiver determinado,  $\mu_1$  também está e é igual a 1. O segundo momento central, quando definido, é a variância, dada por  $\sigma^2(X) = \mu_2(X)$ . Observa-se que se  $\mu'_r(X)$  está definido e  $r' < r$ ,  $\mu'_{r'}(X)$  também está definido. Existem dois coeficientes importantes definidos em termos de momentos centrais.

**Definição A.11 (Coeficiente de Assimetria):** Se  $\mu_3(X)$  estiver definido, o seu coeficiente de assimetria é dado por

$$\gamma_1(X) = \frac{\mu_3(X)}{\sigma(X)^3}. \quad (\text{A.2.3})$$

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

com  $\sigma(X) = \sqrt{\sigma^2(X)}$ .

Note-se que  $\gamma_1(X)$  pode assumir qualquer valor real. Se a variável aleatória  $X$  for simétrica, temos  $\gamma_1(X) = 0$  e todos os momentos centrados de ordem ímpar serão nulos. Mas se  $\gamma_1(X) < 0$ , a variável  $X$  é assimétrica à esquerda. Caso  $\gamma_1(X) > 0$ , a variável  $X$  é assimétrica à direita.

Os coeficientes de assimetria e curtose são aplicados para verificar se um conjunto de dados estão de acordo com a distribuição normal. Caso uma variável aleatória siga uma distribuição normal, então  $\gamma_1(X) = 0$  e  $\gamma_2(X) = 3$ . Contudo, para facilitar a interpretação do coeficiente de curtose ver, para mais pormenores, por exemplo, Martins (2008, pág. 8) [41].

Segue-se a

**Definição A.12 (Curtose):** Se  $\mu_4(X)$  estiver definida, tem-se o coeficiente de achatamento dado por:

$$\gamma_2(X) = \frac{\mu_4(X)}{\sigma^4(X)} - 3. \quad (\text{A.2.4})$$

Temos:

- (a) Se  $\gamma_2(X) < 0$ , a distribuição  $F_X$  tem caudas mais pesadas do que a normal (platicúrtica);
- (b) Se  $\gamma_2(X) = 0$ , a distribuição  $F_X$  tem caudas com o mesmo peso das de uma normal (mesocúrtica);
- (b) Se  $\gamma_2(X) > 0$ , a distribuição  $F_X$  tem caudas mais leves do que a normal (leptocúrtica).

Sendo a Curtose uma medida que caracteriza o achatamento da curva da função de distribuição, esta permite tornar o valor da expressão anterior nulo para uma variável aleatória Gaussiana padrão. Das definições apresentadas anteriormente, torna-se possível estabelecer algumas noções importantes, nomeadamente para as Funções Geradoras de Momentos e Funções Geradoras de Cumulantes.

### A.3 Função Geradora de Momentos

A função geradora de momentos está associada a uma função de densidade de probabilidade, podendo ser usada para gerar momentos, Walpole et al. (2012, págs. 218-222) [70]

**Definição A.13** A variável aleatória  $X$ , com função de densidade  $f_X$ , tem função geradora de momentos dada por

$$\varphi(t | X) = \int_{-\infty}^{\infty} e^{tx} f_X dx, \quad (\text{A.3.5})$$

## Modelos Aditivos - Ajustamento, Elipsóides de Confiança e Intervalos de Predição

definida para todos os valores de  $t$  para os quais este integral converge. Prova-se que

**Teorema A.1** Se  $\varphi(t|X)$  estiver definida para  $|t| < t_0$ , com  $t_0 > 0$ , temos

$$\varphi^{<r>}(X) = \varphi^{<r>}(0|X), \quad (\text{A.3.6})$$

onde  $<r>$  indica a ordem de derivação.

Temos também:

**Teorema A.2 (Teorema da Unicidade)** Se  $\varphi(t|X) = \varphi(t|Y)$  para todos os valores de  $t$  em que estão definidos, então  $f_X = f_Y$ .

Por outro lado, é fácil mostrar que

$$\begin{cases} \varphi(t|X+a) = e^{at}\varphi(t|X) \\ \varphi(t|aX) = \varphi(at|X) \end{cases}. \quad (\text{A.3.7})$$

Temos ainda o seguinte teorema:

**Teorema A.3** Se  $X_1, X_2, \dots, X_n$  são variáveis aleatórias independentes, com funções geradoras de momentos  $\varphi(t|X_1), \varphi(t|X_2), \dots, \varphi(t|X_n)$  todas definidas para  $t'$ , então:

$$\varphi(t'|\sum_{i=1}^n X_i) = \prod_{i=1}^n \varphi(t'|X_i). \quad (\text{A.3.8})$$

Os Momentos proporcionam um conjunto de parâmetros úteis para aferir as propriedades de uma função distribuição, como referem Smith, J. P. (1995) [66] e Zheng, Q. (2002) [73], entre outros. No entanto, existe outro conjunto de parâmetros, designados de Cumulantes ou momentos cumulativos, importantes na descrição de uma variável aleatória, estando intimamente ligados com os momentos da população e que serão utilizados ao longo da tese.

