



Universidade de Évora - Instituto de Investigação e Formação Avançada

Programa de Doutoramento em Matemática

Área de especialização | Estatística

Tese de Doutoramento

Estimadores Otimais em Modelos Bi-aditivos e Suas Famílias

Armando Alexandre

Orientador(es) | João Tiago Praça Nunes Mexia

Maria Manuela Oliveira

Évora 2022



Universidade de Évora - Instituto de Investigação e Formação Avançada

Programa de Doutoramento em Matemática

Área de especialização | Estatística

Tese de Doutoramento

Estimadores Otimais em Modelos Bi-aditivos e Suas Famílias

Armando Alexandre

Orientador(es) | João Tiago Praça Nunes Mexia
Maria Manuela Oliveira

Évora 2022



A tese de doutoramento foi objeto de apreciação e discussão pública pelo seguinte júri nomeado pelo Diretor do Instituto de Investigação e Formação Avançada:

Presidente | Feliz Manuel Minhós (Universidade de Évora)

Vogais | António Manuel Pacheco Pires (Universidade de Lisboa)
Célia Maria Pinto Nunes (Universidade da Beira Interior)
Jorge Manuel Santos (Universidade de Évora)
João Tiago Praça Nunes Mexia (Universidade Nova de Lisboa) (Orientador)
Russell Alpizar-Jara (Universidade de Évora)
Óscar António Louro Felgueiras (Universidade do Porto - Faculdade de Ciências)

Resumo

Estimadores Otimais em Modelos Bi-aditivos e Suas Famílias

Esta tese tem como fundamental objetivo contribuir para a resolução de algumas dificuldades recorrentes na teoria dos modelos bi-aditivos que os investigadores de diversas áreas enfrentam quando aplicam no tratamento das observações colhidas. Os modelos bi-aditivos são dados pela soma dum termo fixo com seus termos aleatórios independentes. Estes modelos dizem-se bi-aditivos para atender simultaneamente a expressão de modelo e a da sua matriz de covariância, ambas com estrutura aditiva.

A tese inicia-se com apresentação de resultados preliminares seguindo-se um capítulo sobre modelos individuais. Nesse capítulo começamos por obter estimadores centrados dos cumulantes das variáveis aleatórias que figuram nos modelos. Como veremos os estimadores obtidos são de mínimos quadrados. Para estimarmos os cumulantes de quarta ordem consideramos duas possibilidades:

* ter um par de modelos independentes e identicamente distribuído;

** as distribuições das componentes dos vetores aleatórios que intervém no modelo terem parâmetros de localização, dispersão e forma ou só parâmetros de localização e dispersão. Segue-se a utilização de comutação entre matrizes obtidas a partir das matrizes com valores das variáveis controladas para obter estimadores com propriedades otimais: BQUE (Best Quadratic Unbiased Estimators) para as componentes de variância e BLUE (Best Linear Unbiased Estimators) para os coeficientes do vetor fixo. Ainda neste capítulo mostra-se como construir elipsoide de confiança e intervalos de predição. Finalmente segue-se, o estudo das famílias de modelos bi-aditivos. Em particular teremos famílias regressionais em que estarão associadas a vetores de variáveis controladas exógenas, na nossa expressão segue-se as famílias estruturadas. Os modelos dessas famílias correspondem aos tratamentos dum delineamento base tendo as mesmas matrizes e idênticos cumulantes para as variáveis aleatórias intervenientes na parte de efeitos aleatórios dos mesmos.

Palavras-chave: Bi-aditividade; Estimação de Cumulantes; Elipsoide de Confiança; Intervalo de Predição; Famílias de Modelos.

Abstract

Optimal Estimators in Bi-Additive Models and Their Families

The main objective of this thesis is to contribute to the resolution of some recurring difficulties in the theory of bi-additive models that researchers from different areas face when they apply them to the treatment of collected observations. Bi-additive models are given by the sum of a fixed term and its independent random terms. These models are said to be bi-additive to simultaneously meet the model expression and its covariance matrix, both with an additive structure.

The thesis begins with the presentation of preliminary results, followed by a chapter on individual models. In this chapter we start by obtaining centered estimators of the cumulative of the random variables that appear in the models. As we will see, the estimators obtained are least squares. To estimate fourth-order cumulative, we consider two possibilities:

* have a pair of independent and identically distributed models.

** the distributions of the components of the random vectors that intervene in the model have location, dispersion, and shape parameters or only location and dispersion parameters. Next, the use of switching between matrices obtained from the matrices with values of the controlled variables to obtain estimators with optimal properties: BQUE (Best Quadratic Unbiased Estimators) for the variance components and BLUE (Best Linear Unbiased Estimators) for the coefficients of the fixed vector. Also in this chapter, it is shown how to build confidence ellipsoid and prediction intervals. Finally, the study of families of bi-additive models follows. We will have regression families that will be associated with vectors of exogenous controlled variables, in our expression the structured families follow. The models of these families correspond to the treatments of a basic design having the same matrices and identical cumulative for the random variables involved in the random effects part of them.

Keywords: Bi-additivity; Estimation of Cumulants; Confidence Ellipsoid; Prediction Interval; Model Families.

Índice

| | |
|--|------|
| Resumo | II |
| Abstract..... | III |
| Lista de Figuras..... | VII |
| Lista de Tabelas | VIII |
| Dedicatória..... | IX |
| Agradecimentos | X |
| Notações, Símbolos e Siglas..... | XII |
| Capítulo 1..... | 1 |
| 1. Introdução | 1 |
| Capítulo 2..... | 5 |
| 2 - Resultados Preliminares | 5 |
| 2.1 - Matrizes | 5 |
| 2.1.1 - Definições..... | 5 |
| 2.1.2 – Determinante..... | 5 |
| 2.1.3 – Matrizes inversas | 6 |
| 2.1.4 - Potências de Matrizes | 7 |
| 2.1.5 - Matriz Ortogonal | 7 |
| 2.1.6 - Valores e vetores próprios | 7 |
| 2.1.7 - Matrizes inversas de Moore-Penrose | 8 |
| 2.1.8 – Bases de sub-espacos | 8 |
| 2.1.9 - Produto de Kronecker de matrizes | 10 |
| 2.2 - Funções geradoras de momentos e de cumulantes..... | 12 |
| 2.3 - Sub-espacos | 13 |
| 2.4 – Estimadores de Mínimos Quadrados | 14 |
| 2.5 – Álgebras de Jordan Comutativas de Matrizes Simétricas..... | 14 |
| 2.6 - Espaços $G[p]^n$ | 17 |
| 2.7 - Cruzamento e aninhamentos de fatores..... | 21 |
| 2.8 - Expansões de Distribuições..... | 22 |
| Capítulo 3..... | 24 |
| 3 - Modelos Bi-aditivos | 24 |

| | |
|---|----|
| 3.1 - Modelos individuais | 24 |
| 3.1.1 - Formulação | 24 |
| 3.1.2 - Ajustamento..... | 25 |
| 3.1.3 - Cumulantes de quarta ordem para os modelos bi-aditivos..... | 27 |
| 3.1.4 - Casos particulares de modelos bi-aditivos individuais..... | 29 |
| 3.1.4.1 – Modelos Multi-Normais | 30 |
| 3.1.4.2 – Modelos Multi-Normais Generalizados | 31 |
| 3.1.4.3 - Modelos Multi-Student Generalizados..... | 32 |
| 3.1.4.4 – Modelos Multi-Gama..... | 34 |
| 3.1.4.5 – Modelos Multi-Gumbel | 35 |
| 3.1.4.6 - Modelos Multi-Beta | 36 |
| 3.1.4.7 – Modelo Multi-típicos | 37 |
| 3.1.4.8 – Simetria..... | 38 |
| 3.2 – Comutações..... | 38 |
| 3.3 – Elipsoides de Confiança | 43 |
| 3.4 - Intervalos de Predição | 45 |
| 3.5 - Modelos com vetores de erros..... | 46 |
| 3.5.1 - Estimação | 46 |
| 3.5.2 – Testes F..... | 49 |
| 3.5.2.1 – Caso aninhado..... | 49 |
| 3.5.2.2 – Cruzamento..... | 52 |
| Capítulo 4..... | 55 |
| 4 - Famílias de Modelos Bi-aditivos..... | 55 |
| 4.1 - Ajustamento..... | 55 |
| 4.2 - Famílias COBS | 57 |
| 4.3 – Famílias Regressionais | 59 |
| 4.4 - Famílias estruturadas | 61 |
| 4.4.1 – Cruzamento dos fatores | 63 |
| 4.4.2- Cruzamento e aninhamento de fatores | 69 |
| 4.4.2.1 – Um exemplo..... | 71 |
| 4.4.3 – Fatoriais de base prima | 73 |
| 4.4.3.1 – Fatoriais completos | 73 |
| 4.4.3.2 – Exemplo de construção de Blocos | 75 |
| 4.4.3.3 – Replicações Fracionárias | 76 |

| | |
|------------------------------------|----|
| 4.4.4. 1 – Um algoritmo útil | 79 |
| Capítulo 5..... | 83 |
| 5 - Considerações Finais | 83 |
| Referências Bibliográficas..... | 86 |

Lista de Figuras

Figura 1 Elipsoide de confiança -----45

Lista de Tabelas

Tabela 1: Conjunto de fatores, caso de quatro fatores-----65

Tabela 2: Conjunto de fatores, caso de cinco fatores-----66

Dedicatória

Aos meus filhos: Jonas e Liliana.

A minha mãe: Jacinta e aos meus irmãos: Zito, Angelina, Daniel, Rita e Elias.

Agradecimentos

À Deus pai todo-poderoso o criador do céu e da terra, que proporcionou milagres em mim, fazendo que eu continuasse a sonhar no mais alto nível no campo da ciência, e pela saúde ou ajudou-me a superar as dificuldades crescentes ao longo desta formação.

Ao meu Orientador, Professor Doutor João Tiago Praça Nunes Mexia, Professor Catedrático Jubilado, pela disponibilidade, estímulo, dedicação e encorajamento demonstrados durante a realização desta investigação, pois a sua orientação e amizade foram bastante relevantes.

A minha Orientadora, Professora Doutora Maria Manuela Oliveira, Professora Associada C/Agregação, pela sua flexibilidade inspiração e ou nobreza, afeição demonstradas durante a realização deste trabalho, pois a sua orientação e amizade foram deveras importantes.

As Direções do PDM – Programa do Doutoramento em Matemática e do IIFA – Instituto de Investigação e Formação Avançada, que me possibilitaram e/ou asseguraram na continuidade para realização desta Tese, numa altura em que estive em total desespero.

Ao Professor Doutor Russell Gerardo Alpizar-Jara, Professor Associado, pelo encorajamento e contributo ostentado no momento da minha crise neste programa, que quase iria deitar todo esforço abaixo.

Ao Professor Doutor Carlos Aberto dos Santos Braumann, Professor Catedrático Jubilado, pelo acoçoamento e compreensão durante a minha frequência das Unidades Curriculares que foi o momento de grande desafio deste percurso.

A todos Professores Doutores do centro de investigação em Matemática – CIMA, que direta ou indiretamente, contribuíram bastante no ensino e aprendizagem dos conteúdos programáticos de cada unidade curricular do PDM e na amizade consagrada durante todo tempo da minha formação, na Universidade de Évora.

A direção da Universidade Zambeze – Moçambique, instituição a que sou vinculado por me autorizar frequentar o Programa de Doutoramento em Matemática no exterior.

A direção do projeto do Ensino Superior, Ciência e Tecnologia (HEST), junto dos: Ministério da Educação e Desenvolvimento Humano (MEDH); Ministério de Ciências e Tecnologia Ensino Superior e Trabalho Profissional (MCTESTP) & Ministério das Finanças (MF) de Moçambique e ao IBE em particular a direção de gestão de bolsa de estudo para formação de docentes do Ensino Superior. Ao Governo de Moçambique que assegurou as despesas depois do projeto ter encerrado o financiamento antes do fim da vigência da bolsa. A todos vão os meus agradecimentos.

Notações, Símbolos e Siglas

Ao longo desta dissertação utilizar-se-ão as(os) seguintes notações, símbolos e siglas:

| | |
|--|--|
| \mathbf{Y} | Vetor aleatório. |
| $\boldsymbol{\beta}$ | Vetor fixo de coeficientes. |
| \mathbf{Z}_i | Vetor aleatório. |
| X_i | Matriz conhecida. |
| $\mathbf{M}_i = \mathbf{X}_i \mathbf{X}_i^t$ | Matriz conhecida. |
| μ | Valor médio (esperança matemática). |
| σ^2 | Variância. |
| $N(\mu, \sigma^2)$ | Distribuição normal com valor médio μ e variância σ^2 . |
| \mathcal{X}_r | Cumulantes de ordem r . |
| $\mathcal{X}_1 = \mu$ | Valor médio. |
| $\mathcal{X}_2 = \mu_2 = \sigma^2$ | Variância. |
| $a_{i,j}$ | Elemento da linha i e coluna j duma matriz. |
| \mathbf{A}^t | Matriz transposta. |
| $\mathbf{P} \perp$ | Matriz ortogonal. |
| I_k | Matriz identidade de ordem k . |
| I_m | Matriz identidade de ordem m . |
| $\det(\mathbf{A})$ | Determinante da matriz \mathbf{A} . |
| $ \cdot $ | Determinante. |
| $ \mathbf{A} $ | Determinante de \mathbf{A} . |
| \mathbf{M}^{-1} | Inversa da matriz \mathbf{M} . |
| $\mathbf{Q} = \mathbf{Q}^t$ | Matriz simétrica. |
| \mathbf{A}^{-1} | Inversa da matriz \mathbf{A} . |
| \mathbf{X}^+ | Inversa de Moore-Penrose da matriz \mathbf{X} . |
| $\text{Car}(\cdot)$ | Característica. |
| $\text{Car}(\mathbf{W})$ | Característica da matriz \mathbf{W} . |
| $\text{R}(\mathbf{U})$ | Espaço imagem da matriz \mathbf{U} . |
| $\text{N}(\mathbf{U})$ | Espaço de nulidade da matriz \mathbf{U} . |

| | |
|---------------------------------|---|
| $\mathbf{0}_{m,n}$ | Matriz nula do tipo $m \times n$. |
| $\sigma_i^2 = \sigma_{i,i}$ | Variância. |
| $\sigma_{i,j}$ | Covariância, $i \neq j$. |
| $\mathbf{W} \otimes \mathbf{Z}$ | Produto de Kronecker das matrizes \mathbf{W} e \mathbf{Z} . |
| CJA | Álgebra comutativa de Jordan. |
| \boxplus | Soma direta ortogonal de subespaço. |
| ∇ | Subespaço. |
| ∇^\perp | Complemento ortogonal de ∇ . |
| BLUE | Do Inglês (Best Linear Unbiased Estimators). |
| BQUE | Do Inglês (Best Quadratic Unbiased Estimators). |
| CJA | Do inglês (Comutative Jordan Algebra). |
| OBS | Do Inglês (Orthogonal Block Structure). |
| COBS | Do Inglês (Commutative Orthogonal Block Structure). |
| $\text{tr}(\mathbf{A})$ | Traço da matriz \mathbf{A} . |
| BLUP | Do Inglês (Best Linear Unbiased Predictor). |

Capítulo 1

Neste capítulo pretende-se falar de forma resumida sobre o desenvolvimento do trabalho e mencionar a sequência das atividades desenvolvidas durante a investigação. Destacando-se a forma dos modelos bi-aditivos, a definição dos modelos bi-aditivos na introdução, os resultados preliminares que serão úteis tratados no capítulo 2. A obtenção de LSE (Least Square Estimators) para os parâmetros relevantes e a introdução de comutação de matrizes para obter BQUE (Best Quadratic Unbiased Estimators) para componentes de variância e BLUE (Best Linear Unbiased Estimators) para vetor de coeficientes β , obtiveram-se ainda elipsoide de confiança e intervalos de predição no capítulo 3. A seguir trataram-se famílias de modelos no capítulo 4 e por último conclusões e recomendações no capítulo 5.

1 - Introdução

Durante muitos anos os modelos normais lineares foram utilizados na tentativa de descrever uma grande parte dos fenômenos aleatórios. No estudo da relação entre duas variáveis X e Y , procura-se uma função de X que explique Y . Em geral, a relação não é perfeita. O modelo a ser adotado depende de vários fatores, por exemplo, da natureza das variáveis, da relação ser linear ou não, da homogeneidade de variâncias ou não, da independência dos erros, entre outros. Por outro lado, a natureza das variáveis X e Y pode variar, isto é, podem ser fixas (ou controladas) ou aleatórias. Se Y é aleatório, três tipos de modelos podem ser considerados: Modelo em que os X 's são fixos; Modelo em que os X 's são aleatórios; Modelo em que parte dos X 's são fixos e parte, aleatórios.

Além disso, ambas podem ser medidas com ou sem erro. \mathbf{X} pode ser fixa ou aleatória, com erro ou sem erro; \mathbf{Y} pode ser fixa ou aleatória com erro ou sem erro.

Nesta tese pretende-se continuar e desenvolver o trabalho realizado anteriormente em Antunes et al. 2020a e 2020b, onde foi considerado o Modelo Bi-aditivo. Estes modelos são extremamente flexíveis, permitindo a sua aplicação a uma grande família de problemas. São assim designados por, para além da sua formulação aditiva terem matrizes de covariância também com uma estrutura aditiva.

Pode-se considerar que estes modelos representam uma extensão dos modelos regressionais.

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

introduzidos por Gauss em 1809 e 1821. Passou assim a ter-se

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

com os vetores $\mathbf{Z}_1, \dots, \mathbf{Z}_w$ independentes e com componentes independentes e identicamente distribuídos.

Estes modelos além da sua estrutura aditiva têm matriz de covariância da forma:

$$\mathbf{V} = \sum_{i=1}^w \sigma_i^2 \mathbf{M}_i, i = 1, \dots, w$$

com $\mathbf{M}_i = \mathbf{X}_i \mathbf{X}_i^t$, sendo designado como bi-aditivo.

A Inferência Estatística tem uma natureza dialética tendo-se na mesma dois elementos fundamentais: o modelo que deverá conter o que se sabe antes de colher as observações; e as observações.

Observe-se que se se admitir um modelo que não se adequa às observações se está a cometer um erro da terceira espécie.

Vê-se assim que os modelos devem ser flexíveis para se alargar o seu domínio de aplicação.

Os modelos bi-aditivos são da forma

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

com $\boldsymbol{\beta}$ um vetor de coeficientes fixo, sendo independentes os vetores $\mathbf{Z}_i = (z_{1,c_i} \dots z_{i,c_i})$, com componentes *i. i. d* e cumulantes de ordem r , $\mathcal{X}_{r,i}$, $r = 2, 3, 4, \dots$ $i = 1, \dots, w$, [ver Antunes et al. (2020a e 2020b)].

Estes modelos são extremamente flexíveis. Suponhamos que se tinha admitido que as matrizes $\mathbf{M}_i = \mathbf{X}_i \mathbf{X}_i^t$, $i = 1, \dots, w$, comutavam o que, como veremos na secção (3.2),

garante ter-se estimadores BQUE (Best Quadratic Unbiased Estimators) para as componentes de variância

$$\sigma_i^2 = \mathcal{X}_{2,i}, i = 1, \dots, w$$

e que, devido a falhas de observações, que se traduzirão na eliminação das linhas correspondentes nas matrizes $\mathbf{X}, \mathbf{X}_1, \dots, \mathbf{X}_w$, as matrizes $\mathbf{M}_1, \dots, \mathbf{M}_w$ não comutam.

No entanto a flexibilidade destes modelos permite, como veremos, analisar os resultados obtidos.

Os modelos bi-aditivos são assim designados por, para além da sua formulação aditiva, terem matrizes de covariância $\sum_{i=1}^w \sigma_i^2 \mathbf{M}_i$ também com estrutura aditiva.

No capítulo 2 apresentamos resultados preliminares que nos serão úteis para:

- i). estimar os parâmetros relevantes;
- ii). Construção de elipsoides de confiança e intervalo de predição;
- iii). Estudar famílias de modelos associados aos tratamentos de delineamento base.

Segue-se o capítulo 3 em que se estudam modelos individuais. Começa-se por estimar os cumulantes $\mathcal{X}_{r,i}, r = 2, 3, 4, i = 1, \dots, w$ e os vetores dos coeficientes $\boldsymbol{\beta}$. O caso dos cumulantes da quarta ordem é estudado a parte e consideram-se:

- i. Pares de modelos independentes e identicamente distribuídos (*i. i. d.*);
- ii. Distribuições das componentes dos $\mathbf{Z}_i, i = 1, \dots, w$, com parâmetros de localização, dispersão e forma ou só de localização e dispersão. Particulariza-se para os casos em que essas distribuições pertencem a certos tipos: Multi-Normal, Multi-Normal Generalizado, Multi-Student generalizado, Multi-Gama, Multi-Gumbel, Multi-Beta e Multi-Típicos. As densidades poderão pertencer a um ou vários tipos.

Considera-se em seguida a obtenção de elipsoides de confiança com probabilidade dada (aproximadamente) de conter realizações dos modelos.

A concluir o capítulo constroem-se intervalos de predição com probabilidades também dadas (aproximadamente) de conter futuras observações associadas a valores das variáveis controladas que intervêm nos modelos. Observe-se que os conjuntos de valores simultâneos dessas variáveis correspondem aos vetores linha das matrizes $\mathbf{X}, \mathbf{X}_1, \dots, \mathbf{X}_w$.

As famílias de modelos são consideradas no capítulo 4. As matrizes $\mathbf{X}, \mathbf{X}_1, \dots, \mathbf{X}_w$ e os cumulantes $\mathcal{X}_{r,i}, r = 2, 3, 4, i = 1, \dots, w$, serão os mesmos para os modelos duma família. Assim as matrizes $\mathbf{M}_1, \dots, \mathbf{M}_w$ [$\mathbf{M}, \mathbf{M}_1, \dots, \mathbf{M}_w$] comutam para todos ou nenhum dos modelos duma família. Aliás este capítulo inicia-se com a estimação

conjunta dos cumulantes e a estimação dos vetores dos coeficientes. O caso em que as matrizes M, M_1, \dots, M_w comutam é individualizado.

Segue-se o estudo das famílias regressionais cujos modelos estão associados a vetores de variáveis controladas, isto é, variáveis cujos valores são determinados pelos experimentadores. Na formulação inicial dos modelos bi-aditivos esses valores estão nas colunas das matrizes X, X_1, \dots, X_w . No caso dos intervalos de predição esses valores são escolhidos após o ajustamento do modelo. Ou seja, as variáveis controladas são variáveis cujos valores são fixados pelo investigador, esses valores constituem as colunas das matrizes X, X_1, \dots, X_w . Esses valores são também considerados na construção de intervalo de predição.

Teremos também as famílias estruturadas cujos modelos correspondem aos tratamentos dum delineamento de efeitos fixos. Estudar-se-á a ação dos fatores desse delineamento nos vetores de coeficientes dos modelos da família.

Quer nas famílias regressionais quer nas famílias estruturadas o caso em que as matrizes M, M_1, \dots, M_w comutam será considerado. Foi também considerado a expansão de Edgeworth.

Por último seguir-se-á a síntese do estudo na perspectiva das conclusões e recomendações, ou seja, menção do que foi feito no presente trabalho e o que desenvolver para trabalho futuro.

Capítulo 2

Neste presente capítulo descrevem-se os resultados preliminares que nos serão úteis nomeadamente: sobre matrizes (definição, determinante de matriz, matrizes inversas, potência de matrizes, matriz ortogonal, valores e vetores próprios, matrizes inversas de Moore-Penrose), bases de sub-espços, produtos de Kronecker de matrizes, funções geradoras de momentos e de cumulantes, sub-espços, estimadores de mínimos quadrados, álgebras de Jordan comutativas de matrizes simétricas, espaços $G[p]^n$ e cruzamentos e aninhamentos de fatores sendo estes resultados preliminares que nos ajudarão a interpretar alguns tópicos dos capítulos posteriores.

2 - Resultados Preliminares

2.1 - Matrizes

2.1.1 - Definição

Cayley (1858) introduziu as matrizes como uma notação conveniente para sistemas de equações lineares. Essa introdução foi extremamente útil havendo hoje livros dedicados ao estudo das matrizes entre os quais Schott (2017).

2.1.2 – Determinante

O determinante de uma matriz A do tipo $m \times m$ será

$$|A| = \sum (-1)^{f(i_1, \dots, i_m)} a_{1, i_1} a_{2, i_2} \dots a_{m, i_m},$$

onde a soma é efetuada para todas as permutações $(i_1 \dots i_m)$ de $(1, \dots, m)$ e $f(i_1 \dots i_m)$ é o número de permutações necessário para reduzir $(i_1 \dots i_m)$ a $(1, \dots, m)$.

Quando $m = 1$, $|A| = a_{1,1}$ e se $m = 2$, $|A| = a_{1,1} a_{2,2} - a_{1,2} a_{2,1}$ e assim sucessivamente.

Pode mostrar-se, [ver Schott (2017, p.9)] que

$$|A B| = |A| |B|.$$

2.1.3 – Matrizes inversas

A matriz A do tipo $m \times m$ é invertível se $|A| \neq 0$. Existe então uma matriz A^{-1} , a inversa de A tal que

$$A A^{-1} = A^{-1} A = I_m.$$

A matriz inversa é única pois se B satisfizer a equação matricial anterior ter-se-á

$$B = B I_m = B A A^{-1} = I_m A^{-1} = A^{-1}.$$

Tem-se com $\alpha \neq 0$ as seguintes propriedades:

- $(\alpha A)^{-1} = \alpha^{-1} A^{-1}$
- $(A^t)^{-1} = (A^{-1})^t$
- $(A^{-1})^{-1} = A$
- $|A^{-1}| = |A|^{-1}$.

Sendo $diag(a_1 \dots a_m)$ a matriz diagonal com elementos principais $a_1 \dots a_m$, tem-se

$$(diag(a_1 \dots a_m))^{-1} = diag(a_1^{-1} \dots a_m^{-1}), \text{ caso } a_j \neq 0, j = 1, \dots, m,$$

tem-se ainda

$$(A B)^{-1} = B^{-1} A^{-1}.$$

Lema 2.1.3.1 – Dada uma matriz B invertível a inversa dessa matriz é única

Demonstração:

Suponhamos que B e C são matrizes inversas de A . Então:

$$B(A C) = B I_m = B,$$

e

$$(B A)C = I_m C = C$$

uma vez que a multiplicação de matrizes é associativa, tem-se

$$\mathbf{B} = \mathbf{B}(\mathbf{A} \mathbf{C}) = (\mathbf{B} \mathbf{A}) \mathbf{C} = \mathbf{C}.$$

Podemos assim designar por \mathbf{A}^{-1} a (única) matriz inversa de \mathbf{A} . A matriz, \mathbf{A} diz se invertível ou não singular se admitir matriz inversa [ver Paliga (2012, p.15)].

2.1.4 - Potências de Matrizes

Sendo \mathbf{A} uma matriz quadrada de ordem m , define-se a potência de expoente n , de \mathbf{A} , com $n \in \mathbb{N}_0$, de acordo com o seguinte:

- $\mathbf{A}^0 = \mathbf{I}_m$
- $\mathbf{A}^1 = \mathbf{A}$
- $\mathbf{A}^n = \mathbf{A} \mathbf{A}^{n-1} \quad n \geq 2$

ou seja, quando n é um inteiro maior ou igual 2, \mathbf{A}^n o produto de n fatores iguais a \mathbf{A} , $\mathbf{A}^n = \mathbf{A} \dots \mathbf{A}$.

2.1.5 - Matriz Ortogonal

Definição 2.1.5.1 - Uma matriz quadrada \mathbf{P} de ordem k , é ortogonal se e só se

$$\mathbf{P}^t \mathbf{P} = \mathbf{P} \mathbf{P}^t = \mathbf{I}_k$$

onde \mathbf{I}_k representa a matriz identidade de ordem k , isto é, as inversas das matrizes ortogonais são as suas transpostas, [ver Paliga (2012, p.19)].

Teorema 2.1.5.2 - Sejam \mathbf{A} e \mathbf{B} matrizes ortogonais de ordem k , e \mathbf{C} uma matriz qualquer também de ordem k . Tem-se:

- $|\mathbf{A}| = 1$
- $|\mathbf{A}^t \mathbf{C} \mathbf{A}| = |\mathbf{C}|$
- \mathbf{AB} é uma matriz ortogonal

2.1.6 - Valores e vetores próprios

Dada a matriz quadrada \mathbf{B} de ordem m , se

$$\mathbf{B} \mathbf{v} = \lambda \mathbf{v}$$

λ será o valor próprio associado ao vetor próprio \boldsymbol{v} ambos de \mathbf{B} , e se \mathbf{B} for simétrica existe [ver Schott (2017) e Mexia (1995, p.1)] uma matriz ortogonal \mathbf{P} tal que

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

com $r_1 \dots r_m$ os valores próprios de \mathbf{B} .

Tem-se então

$$\prod_{j=1}^w r_j = |\mathbf{P} \mathbf{B} \mathbf{P}^t| = |\mathbf{B}|$$

2.1.7 - Matrizes inversas de Moore-Penrose

Qualquer que seja a matriz \mathbf{X} existe uma e uma só matriz \mathbf{X}^+ , a inversa de Moore-Penrose de \mathbf{X} tal que:

$$\left[\begin{array}{l} \mathbf{X} \mathbf{X}^+ \mathbf{X} = \mathbf{X} \\ \mathbf{X}^+ \mathbf{X} \mathbf{X}^+ = \mathbf{X}^+ \\ (\mathbf{X} \mathbf{X}^+)^t = \mathbf{X} \mathbf{X}^+ \\ (\mathbf{X}^+ \mathbf{X})^t = \mathbf{X}^+ \mathbf{X} \end{array} \right.$$

[ver, Pollock (1979, P.55)].

Caso \mathbf{X} seja invertível tem-se:

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

2.1.8 – Bases de sub-espacos

Os vetores $\boldsymbol{v}_1 \dots \boldsymbol{v}_m$ são linearmente independentes se, sendo $\mathbf{0}$, o vetor nulo

$$\sum_{i=1}^w a_i \boldsymbol{v}_i = \mathbf{0},$$

implicar

$$a_1 = \dots = a_m = 0.$$

Se $\boldsymbol{v}_1 \dots \boldsymbol{v}_m$ forem linearmente independentes

$$\sum_{i=1}^m a_i \boldsymbol{v}_i = \sum_{i=1}^m b_i \boldsymbol{v}_i$$

implicará que

$$a_i = b_i, i = 1, \dots, m$$

pois ter-se-ia

$$\sum_{i=1}^m (b_i - a_i) \boldsymbol{v}_i = \mathbf{0}.$$

Um sub-espço vetorial ∇ é uma família de vetores, fechada para combinações lineares, isto é, se $\boldsymbol{v}_1 \dots \boldsymbol{v}_m \in \nabla$ então $\sum_{i=1}^m a_i \boldsymbol{v}_i \in \nabla$. Se os $\boldsymbol{v}_1 \dots \boldsymbol{v}_m$ forem linearmente independentes e se ∇ for constituído pelas combinações lineares $\sum_{i=1}^m a_i \boldsymbol{v}_i$ diremos que $\{\boldsymbol{v}_1 \dots \boldsymbol{v}_m\}$ é uma base para ∇ . Dado $\boldsymbol{w} \in \nabla$ teremos

$$\boldsymbol{w} = \sum_{i=1}^m c_i \boldsymbol{v}_i.$$

sendo esta representação única.

Dados os vetores $\boldsymbol{v} = (v_1 \dots v_m)$ e $\boldsymbol{u} = (u_1 \dots u_m)$ com componentes $v_1 \dots v_m$ e $u_1 \dots u_m$ o seu produto interno será

$$\boldsymbol{v}^t \boldsymbol{u} = \sum_{i=1}^m v_i u_i$$

dizendo-se que \boldsymbol{v} e \boldsymbol{u} são ortogonais, e escrevendo-se $\boldsymbol{v} \perp \boldsymbol{u}$,

se $\boldsymbol{v}^t \boldsymbol{u} = 0$.

Suponhamos que $\{\boldsymbol{v}_1 \dots \boldsymbol{v}_h\}$ é uma base de ∇ cujos vetores são ortogonais dois a dois, então essa base será ortogonal.

Se ainda se tiver

$$\|\boldsymbol{v}_j\| = \sqrt{\boldsymbol{v}_j^t \boldsymbol{v}_j} = 1, j = 1, \dots, h$$

a base será ortonormada.

Se os vetores linha de \mathbf{B} constituem uma base ortonormada para ∇ , $\mathbf{B}^t \mathbf{B}$ será a matriz de projeção ortogonal sobre ∇ .

Se todo o vetor de ∇ é ortogonal a todo o vetor de Δ diremos que são ortogonais. Sendo $Q(\nabla)$ e $Q(\Delta)$ as respectivas matrizes de projeção ortogonal tem-se

$$Q(\nabla)Q(\Delta) = 0_{m \times m}.$$

com $0_{m \times m}$ a matriz nula $m \times m$. Põe-se então $\nabla \perp \Delta$ bem como $Q(\nabla) \perp Q(\Delta)$, dizendo-se que estas são matrizes de projeção ortogonal mutuamente ortogonais.

Então o sub-espaço formado pelos $\boldsymbol{v} + \boldsymbol{u}$ com $\boldsymbol{v} \in \nabla$ e $\boldsymbol{u} \in \Delta$ será o sub-espaço $\nabla \boxplus \Delta$ e ∇ e Δ constituirão uma partição direta ortogonal de $\nabla \boxplus \Delta$.

Para representar a soma direta ortogonal dos sub-espaços $\nabla_1, \dots, \nabla_d$ poremos $\boxplus_{h=1}^d \nabla_h$.

2.1.9 - Produto de Kronecker de matrizes

O produto de Kronecker foi largamente estudado por: Graham (1981), Steeb (1991), Steeb e Hardy (2011), [ver Fonseca et al. (2006)].

Definição 2.1.9.1 - Seja $\boldsymbol{W} = [w_{i,j}]$ uma matriz do tipo $k \times n$ e \boldsymbol{Z} uma matriz do tipo $r \times c$, então o produto de Kronecker entre estas duas matrizes, $\boldsymbol{W} \otimes \boldsymbol{Z}$, será dado pela matriz por blocos $\boldsymbol{W} \otimes \boldsymbol{Z}$ do tipo $kr \times nc$,

$$\boldsymbol{W} \otimes \boldsymbol{Z} = \begin{bmatrix} w_{1,1}\boldsymbol{Z} & \dots & w_{1,c}\boldsymbol{Z} \\ \vdots & \ddots & \vdots \\ w_{r,1}\boldsymbol{Z} & \dots & w_{r,c}\boldsymbol{Z} \end{bmatrix}$$

com

$$\boldsymbol{Z} = \begin{bmatrix} z_{1,1} & \dots & z_{1,c} \\ \vdots & \ddots & \vdots \\ z_{r,1} & \dots & z_{r,c} \end{bmatrix}$$

Note-se que o produto Kronecker de matrizes não goza da propriedade comutativa.

A seguir apresentam-se alguns resultados sobre o produto de Kronecker, que podem ser encontrados em Schott (2017, pp.316-317).

Teorema 2.1.9.2 - Sejam \boldsymbol{W} , \boldsymbol{Z} e \boldsymbol{Y} matrizes quaisquer e \boldsymbol{w} e \boldsymbol{z} vetores quaisquer. Então

- $\boldsymbol{\alpha} \otimes \boldsymbol{W} = \boldsymbol{W} \otimes \boldsymbol{\alpha} = \boldsymbol{\alpha}\boldsymbol{W}$, para qualquer matriz pseudo escalar $\boldsymbol{\alpha}$.
- $(\boldsymbol{\alpha}\boldsymbol{W}) \otimes (\boldsymbol{\beta}\boldsymbol{Z}) = \boldsymbol{\alpha}\boldsymbol{\beta}(\boldsymbol{W} \otimes \boldsymbol{Z})$, para quaisquer escalares $\boldsymbol{\alpha}$ e $\boldsymbol{\beta}$.
- $(\boldsymbol{W} \otimes \boldsymbol{Z}) \otimes \boldsymbol{Y} = \boldsymbol{W} \otimes (\boldsymbol{Z} \otimes \boldsymbol{Y})$.
- $(\boldsymbol{W} + \boldsymbol{Z}) \otimes \boldsymbol{Y} = (\boldsymbol{W} \otimes \boldsymbol{Y}) + (\boldsymbol{Z} \otimes \boldsymbol{Y})$, se \boldsymbol{W} e \boldsymbol{Z} são do mesmo tipo.

- $W \otimes (Z + Y) = (W \otimes Z) + (W \otimes Y)$, se Z e Y são do mesmo tipo.
- $WZ^t = W \otimes Z^t = Z^t \otimes W$.

Corolário 2.1.9.3 - Sejam W e Z matrizes diagonais quaisquer. Então $W \otimes Z$ é uma matriz diagonal.

Nota: uma matriz pseudo escalar α é a matriz 1×1 cujo único elemento é α .

Tendo em conta as propriedades apresentadas no Teorema 2.1.12.2, facilmente se estabelece que

$$\left(\sum_{i=1}^k w_i \mathbf{G}_i \right) \otimes \left(\sum_{j=1}^n z_j \mathbf{H}_j \right) = \sum_{i=1}^k \sum_{j=1}^n w_i z_j (\mathbf{G}_i \mathbf{H}_j).$$

Teorema 2.1.9.4 - Sejam W, Z, Y e U matrizes dos tipos $n \times h, p \times k, h \times m, e k \times q$, respetivamente. Tem-se

$$(W \otimes Z)(Y \otimes U) = WY \otimes ZU.$$

Demonstração imediata

Corolário 2.1.9.5 - Sejam W e Z matrizes ortogonais quaisquer. Então $W \otimes Z$ é uma matriz ortogonal.

Demonstração:

Basta verificar que sendo W e Z matrizes ortogonais de ordens k , e n , respetivamente, se tem

$$(W \otimes Z)(W \otimes Z)^t = (W \otimes Z)(W^t \otimes Z^t) = WW^t \otimes ZZ^t = I_k \otimes I_n = I_{k,n}$$

analogamente

$$(W \otimes Z)^t(W \otimes Z) = I_{k,n}.$$

Corolário 2.1.9.6 - Sejam W e Z matrizes de projeção ortogonal. Então $W \otimes Z$ é uma matriz de projeção ortogonal.

Teorema 2.1.9.7 - Sejam W e Z matrizes quaisquer. Ter-se-á

$$(W \otimes Z)^t = W^t \otimes Z^t.$$

Corolário 2.1.9.7 - Sejam W e Z matrizes simétricas quaisquer então $W \otimes Z$ é simétrica.

Teorema 2.1.9.8 - Seja W uma matriz quadrada de ordem n , e Z uma matriz quadrada de ordem q . Então

$$\text{car}(W \otimes Z) = \text{car}(W) \cdot \text{car}(Z).$$

Teorema 2.1.9.9 - Seja W uma matriz do tipo $k \times p$, e seja Z uma matriz do tipo $n \times q$. Então

- $(W \otimes Z)^{-1} = W^{-1} \otimes Z^{-1}$, se $k = p$, $n = q$ e W e Z , são matrizes invertíveis.
- $(W \otimes Z)^+ = W^+ \otimes Z^+$.

2.2 - Funções geradoras de momentos e de cumulantes

Embora estas funções estejam também definidas para o caso multivariado só nos interessa o caso univariado.

Dada uma variável aleatória W a respetiva função geradora de momentos é dada por

$$\varphi(d|W) = E(e^{dW})$$

estando definida para os valores d em que o valor médio de e^{dW} está definido.

Quanto á função geradora dos cumulantes de W a mesma é dada por

$$\psi(d|W) = \log(\varphi(d|W))$$

[ver, por exemplo Pistone e Wynn (2006, pp.210-221)].

estando definida para os mesmos valores que $\varphi(d|W)$ [ver Antunes et al. (2020a, 2020b)].

O r -ésimo cumulante de W , quando definido, é dado por

$$C_r(W) = \psi^{<r>}(0|W),$$

com $< r >$ indicando a r -ésima derivada.

Sendo $\mu(W)$ [$\sigma^2(W)$] o valor médio [variância] e $\mu_r(W)$, $r > 1$, o r -ésimo momento central de W tem-se, [ver Balakrishnan et al. (1998)] e [ver Douglas (1980)].

- $C_1(W) = \mu(W)$
- $C_2(W) = \sigma^2(W)$
- $C_3(W) = \mu_3(W)$
- $C_4(W) = \mu_4(W) - 3(\sigma^2(W))^2$

tem-se ainda

$$\begin{cases} C_1(aW + b) = aC_1(W) + b \\ C_r(aW + b) = a^r C_r(W); r > 1 \end{cases}$$

por outro lado, se W_1, \dots, W_m forem independentes temos,

$$C_r \left(\sum_{\ell=1}^m a_{\ell} W_{\ell} \right) = \sum_{\ell=1}^m a_{\ell}^r C_r(W_{\ell})$$

[ver Craig (1931)], [ver Mukherjea et al. (2006, pp.1185-1189)]. [ver McCullagh e Nelder (1989)], [ver Nelder e Wedderburn (1972)] e [ver Turkman e Silva (2000)], para mais detalhes sobre função geradora de momentos. [ver Smith (2012, pp.217-218)].

2.3 - Sub-espços

Dado um sub-espço ∇ de R^n o conjunto de todos vetores de R^n ortogonais a todos vetores de ∇ será o complemento ortogonal ∇^\perp de ∇ . Observe-se que, se $\mathbf{u}, \mathbf{v} \in \nabla^\perp$, e $\mathbf{w} \in \nabla$ se tem

$$(a\mathbf{u} + b\mathbf{v})^t \mathbf{w} = a(\mathbf{u}^t \mathbf{w}) + b(\mathbf{v}^t \mathbf{w}) = 0$$

Pelo que ∇^\perp será sub-espço de R^n . Aliás é fácil de mostrar que todo o vetor de ∇ é ortogonal a todo vetor de ∇^\perp , vindo $\nabla \perp \nabla^\perp$.

Se $\mathbf{v} \in \Delta \cap \nabla$ com $\Delta \perp \nabla$ tem-se $\mathbf{v}^t \mathbf{v} = 0$ o que implica $\mathbf{v} = \mathbf{0}$.

Dada uma matriz X o seu espaço de nulidade, $N(X)$, é formado pelos vetores \mathbf{u} tais que $X\mathbf{u} = \mathbf{0}$. Por outro lado, o espaço imagem, $R(U)$, da matriz U é formado pelos vetores

$$\mathbf{z} = U\mathbf{u}.$$

Tem-se, ver Seber (1980, p.2)

$$N(\mathbf{B}) = R(\mathbf{B}^t)^\perp.$$

O sub-espço ∇ é soma direta ortogonal dos sub-espço $\Delta_1, \dots, \Delta_w$, mutuamente ortogonais, tendo-se

$$\nabla = \boxplus_i^w \Delta_i,$$

se e só se todo o vetor $\mathbf{x} \in \nabla$ for dado por

$$\mathbf{x} = \sum_{i=1}^w \mathbf{x}_i,$$

com $\mathbf{x}_i \in \Delta_i$.

Mostra-se que esta representação é única, sendo $\mathbf{Q}, \mathbf{Q}_1 \dots \mathbf{Q}_w$ as matrizes de projeção ortogonal sobre $\nabla, \Delta_1, \dots, \Delta_w$, tem-se

$$Q = \sum_{i=1}^w Q_i.$$

2.4 – Estimadores de Mínimos Quadrados

Dado o vetor aleatório Y com vetor médio

$$\mu = X \beta,$$

o LSE (Least Squares Estimators) de β é

$$\tilde{\beta} = (X^t X)^+ X^t Y.$$

Estes estimadores minimizam

$$S(B) = \|Y - X \beta\|^2,$$

dando o nome de estimador de mínimos quadrados.

Quando Y tem matriz de covariância $\Sigma(Y)$ para a qual se tem um estimador centrado $\tilde{\Sigma}(Y)$ tem-se o GLSE do Inglês (Generalized Least Squares Estimators), ver Kuriya & Kurata (2004)

$$\tilde{\beta} = (X^t \tilde{\Sigma}(Y) X)^+ X^t \Sigma(Y) Y.$$

2.5 – Álgebras de Jordan Comutativas de Matrizes Simétricas

Estas álgebras são espaços vetoriais, isto é, contêm, as combinações lineares de suas matrizes que comutam, bem como os quadrados das mesmas [ver (Leon (2010))].

Assim se M_1, \dots, M_w pertencem á álgebra de Jordan comutativa, AJC, \mathcal{A} temos

$$\sum_{i=1}^m a_i M_i \in \mathcal{A},$$

bem como

$$M_i^2 \in \mathcal{A}, i = 1, \dots, m.$$

Ora, ver Schott (2017, p.184), as matrizes simétricas W_1, \dots, W_u comutam se e só forem diagonalizadas pela mesma matriz ortogonal P , tendo-se

$$P^t W_j P = D_j, j = 1, \dots, u,$$

com D_j matriz diagonal, $j = 1, \dots, u$.

Assim existe uma matriz ortogonal \dot{P} tal que diagonaliza as matrizes da AJC.

Dadas $M_1, M_2 \in \mathcal{A}$ ter-se-á

$$M_\ell = \dot{P}^t D_\ell \dot{P}, \ell = 1, 2,$$

com D_ℓ matriz diagonal cujos elementos principais são os valores próprios $\mathcal{E}_{\ell,1}, \dots, \mathcal{E}_{\ell,n}$ (admitindo-se que as matrizes de \mathcal{A} são do tipo $n \times n$) de $M_\ell, \ell = 1, 2$.

Tem-se então

$$M_1 M_2 = \dot{P}^t D_1 \dot{P} \dot{P}^t D_2 \dot{P} = \dot{P}^t D_1 D_2 \dot{P},$$

com $D_1 D_2$ a matriz diagonal cujos elementos principais são os produtos dos elementos principais homólogos de M_1 e M_2 . Como estes elementos principais são os valores próprios de M_1, M_2 e $M_1 M_2$ vê-se que os valores próprios de $M_1 M_2$ são os produtos dos valores próprios homólogos de M_1 e M_2 .

Por outro lado, ver Seely (1971), toda a AJC tem uma base única, a base principal, $\mathcal{b}p$, constituída por matrizes de projeção ortogonal mutuamente ortogonais. Sendo

$$\{Q_1, \dots, Q_m\} = \mathcal{b}p(\mathcal{A}),$$

dado $M \in \mathcal{A}$ tem-se

$$M = \sum_{j=1}^m a_j Q_j,$$

e se

$$M_\ell = \sum_{j=1}^m a_{j,\ell} Q_j, \ell = 1, 2,$$

tem-se

$$M_1 M_2 = \sum_{j=1}^m (a_{j,1} a_{j,2}) Q_j$$

visto as Q_1, \dots, Q_m serem idempotentes e mutuamente ortogonais.

Sendo Q uma matriz de projeção ortogonal pertencente a \mathcal{A} tem-se

$$Q = \sum_{j=1}^m q_j Q_j,$$

e como Q é idempotente tem-se

$$\mathbf{Q} = \mathbf{Q}^2 = \left(\sum_{j=1}^m q_j \mathbf{Q}_j \right) \left(\sum_{j'=1}^m q_{j'} \mathbf{Q}_{j'} \right) = \sum_{j=1}^m q_j^2 \mathbf{Q}_j,$$

vindo

$$q_j = q_j^2, j = 1, \dots, m.$$

Logo $q_j \in \{0, 1\}$, $j = 1, \dots, m$ vindo

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

com $\varphi = \{j'; q_j = 1\}$, vindo-se ainda que

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

e que

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

Resulta ainda que se $\mathbf{Q} \in \mathcal{A}$ e $\text{car}(\mathbf{Q}) = 1$, \mathbf{Q} terá de pertencer á $\mathcal{L}p(\mathcal{A})$.

Em particular se $\frac{1}{n}J_n = \frac{1}{n}1_n 1_n^t \in \mathcal{A}$, tomamos $\mathbf{Q}_j = \frac{1}{n}J_n$ e a AJC é regular.

Por outro lado, se

$$\mathbf{U} = \sum_{j=1}^m \mathbf{Q}_j,$$

com $\mathcal{L}p(\mathcal{A}) = \{\mathbf{Q}_1, \dots, \mathbf{Q}_m\}$, \mathbf{U} será a unidade de \mathcal{A} já que com $M \in \mathcal{A}$ tendo-se

$$\mathbf{M} = \sum_{j=1}^m a_j \mathbf{Q}_j,$$

tem-se

$$\mathbf{U} \mathbf{M} = \mathbf{M} \mathbf{U} = \mathbf{M}.$$

Vê-se ainda que, se

$$\mathbf{M} = \sum_{j=1}^m a_j \mathbf{Q}_j \in \mathcal{A},$$

com $a_j^+ = a_j^{-1}$ se $a_j \neq 0$ e $a_j^+ = 0$ se $a_j = 0$, se tem

$$\mathbf{M}^+ = \sum_{j=1}^m a_j^+ \mathbf{Q}_j \in \mathcal{A},$$

já que

$$\left\{ \begin{array}{l} \mathbf{M} \mathbf{M}^+ \mathbf{M} = \mathbf{M} \\ \mathbf{M}^+ \mathbf{M} \mathbf{M}^+ = \mathbf{M}^+ \\ (\mathbf{M} \mathbf{M}^+)^t = \mathbf{M} \mathbf{M}^+ \\ (\mathbf{M}^+ \mathbf{M})^t = \mathbf{M}^+ \mathbf{M} \end{array} \right.$$

Assim as AJC conterão as inversas de Moore-Penrose das suas matrizes.

Vê-se ainda que \mathbf{U} é simétrica e idempotente pelo que, é matriz de projeção ortogonal, [ver Schott (2017, p.64)].

Dada $\mathbf{M} \in \mathcal{A}$ tem-se,

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

com $\varphi = \{j; a_j \neq 0\}$, vindo

$$\text{car}(\mathbf{M}) = \sum_{j \in \varphi} \text{car}(\mathbf{Q}_j) \leq \sum_{j=1}^m \text{car}(\mathbf{Q}_j) = \text{car}(\mathbf{U})$$

se $\text{car}(\mathbf{U}) = n$, $\mathbf{U} = I_n$ e \mathcal{A} conterá matrizes invertíveis (em particular \mathbf{U}). se $\text{car}(\mathbf{U}) < n$ todas as matrizes de \mathcal{A} terão características inferiores a n , não podendo existir matrizes invertíveis em \mathcal{A} . Assim para \mathcal{A} conter matrizes invertíveis é necessário e suficiente que, $\text{car}(\mathbf{U}) = n$ dizendo-se que, quando \mathbf{U} é invertível, a AJC é completa.

Pondo-se

$$\Theta = R(\mathbf{U}),$$

tem-se

$$\Theta = \boxplus_{j=1}^m R(\mathbf{Q}_j),$$

o que mostra a relação entre partições ortogonais de subespaços e bases principais de AJC.

2.6 - Espaços $G[p]^n$

Nesta secção utilizaremos, com p primo, a aritmética modulo p em que se substituem os resultados da adição e multiplicação usuais pelos restos da sua divisão por p .

Assim com $p = 3$ tem-se as tabelas para as respetivas adição e multiplicação

$+(3)$

| | 0 | 1 | 2 |
|---|---|---|---|
| 0 | 0 | 1 | 2 |
| 1 | 1 | 2 | 0 |
| 2 | 2 | 0 | 1 |

$x_{(3)}$

| | 0 | 1 | 2 |
|---|---|---|---|
| 0 | 0 | 0 | 0 |
| 1 | 0 | 1 | 2 |
| 2 | 0 | 2 | 1 |

com $p = 5$ as tabelas seriam

$+_{(5)}$

| | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| 0 | 0 | 1 | 2 | 3 | 4 |
| 1 | 1 | 2 | 3 | 4 | 0 |
| 2 | 2 | 3 | 4 | 0 | 1 |
| 3 | 3 | 4 | 0 | 1 | 2 |
| 4 | 4 | 0 | 1 | 2 | 3 |

$x_{(5)}$

| | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 1 | 2 | 3 | 4 |
| 2 | 0 | 2 | 4 | 1 | 3 |
| 3 | 0 | 3 | 1 | 4 | 2 |
| 4 | 0 | 4 | 3 | 2 | 1 |

onde apenas se consideram os $\{1, \dots, p - 1\}$ os quais constituem o suporte do corpo $G[p]$ de Galois.

Observe-se que

$$G(p)^n = \{(x_1 \dots x_n); x_j \in G[p], j = 1, \dots, n\},$$

sendo fechado para combinação linear é um espaço vetorial (sobre o corpo $G[p]$).

Dado $\mathbf{a} \in G[p]^n$ temos a aplicação linear

$$L(\mathbf{x}|\mathbf{a}) = \left(\sum_{j=1}^n a_j x_j \right)_{(p)},$$

onde (p) indica a utilização de aritmética modulo p . Com $c_1, \dots, c_m \in G[p]$ seja

$$\sum_{j=1}^m c_j L(\mathbf{x}|\mathbf{a}_j) = L\left(\mathbf{x} \left| \left(\sum_{j=1}^m c_j \mathbf{a}_j \right)_{(p)} \right.\right)$$

vê-se assim que $\mathcal{L}_r([p]^n)$, a família destas aplicações lineares, também é fechada para combinações lineares sendo, pois, um espaço vetorial (em terminologia algébrica será o dual de $G[p]^n$).

As aplicações lineares reduzidas têm o seu primeiro coeficiente não nulo igual a 1.

Enquanto há p^n aplicações lineares o número de aplicações lineares reduzida será

$$N_r = \frac{p^n - 1}{p - 1}.$$

Por exemplo com $p = 3$ e $n = 3$, $\mathcal{L}([p]^3)$ terá 27 aplicações ordenadas pelos índices

$$i(\mathbf{a}) = 1 + \sum_{h=1}^3 (\mathbf{a}_h - 1) 3^{h-1},$$

no caso geral este índice será

$$i(\mathbf{a}) = 1 + \sum_{h=1}^n (\mathbf{a}_h - 1) p^{h-1}.$$

Neste exemplo teremos, com L_0 a aplicação nula correspondente a $\mathbf{a} = \mathbf{0}$

| | | | | | | |
|--------|---------------|---------------------|----------------------|--------------------|---------------------|-----|
| L_0 | x_2 | $2x_2$ | x_3 | $x_2 + x_3$ | $2x_2 + x_3$ | ... |
| x_1 | $x_1 + x_2$ | $x_1 + 2x_2$ | $x_1 + x_3$ | $x_1 + x_2 + x_3$ | $x_1 + 2x_2 + x_3$ | ... |
| $2x_1$ | $2x_1 + x_2$ | $2x_1 + 2x_2$ | $2x_1 + x_3$ | $2x_1 + x_2 + x_3$ | $2x_1 + 2x_2 + x_3$ | ... |
| | $2x_3$ | $x_2 + 2x_3$ | $2x_2 + 2x_3$ | | | |
| | $x_1 + 2x_3$ | $x_1 + x_2 + 2x_3$ | $x_1 + 2x_2 + 2x_3$ | | | |
| | $2x_1 + 2x_3$ | $2x_1 + x_2 + 2x_3$ | $2x_1 + 2x_2 + 2x_3$ | | | |

sendo coloridas (azul) as 13 aplicações lineares reduzidas.

As aplicações lineares $L(\mathbf{x}|\mathbf{a}_1), \dots, L(\mathbf{x}|\mathbf{a}_m)$ são linearmente independentes se

$$\sum_{j=1}^m c_j L(\mathbf{x}|\mathbf{a}_j) = L_0,$$

implicar $c_1 = \dots = c_m = 0$ ou o que o mesmo

$$\sum_{j=1}^m c_j \mathbf{a}_j = \mathbf{0}.$$

Assim as $L(\mathbf{x}|\mathbf{a}_1) \dots L(\mathbf{x}|\mathbf{a}_m)$ são linearmente independentes se e só se os seus vetores de coeficientes o forem.

Dada $\mathbf{L} = (L(\mathbf{a}_1) \dots L(\mathbf{a}_m))$ pomos

$$A(\mathbf{L}) = [\mathbf{a}_1 \dots \mathbf{a}_m]^t.$$

Observe-se que, dado $\mathcal{L}([p]^n)$ ter dimensão n se tem de ter $m \leq n$

e que, se as $(L(\mathbf{a}_1) \dots L(\mathbf{a}_m))$ poderem ser linearmente independentes, tendo-se então

$$\text{car}(A(\mathbf{L})) = m,$$

e tendo o sistema $A(\mathbf{L})\mathbf{x} = \boldsymbol{\theta}$, com $x, \boldsymbol{\theta} \in G[p]^n$, p^{n-m} soluções que constituem o bloco $[\mathbf{L}|\boldsymbol{\theta}]$.

Assim os p^n vetores, de $G[p]^n$ distribuem-se por p^m blocos.

Estes vetores podem ser ordenados pelos índices

$$j(\mathbf{x}) = 1 + \sum_{h=1}^n (x_h - 1)p^{h-1}.$$

Se os $\mathbf{x} \in G[p]^n$ tiverem como componentes $x_1 \dots x_n$ os níveis dos fatores dum delineamento, estes índices afetarão as observações, $Y_1 \dots Y_{p^n}$.

A $L(\cdot) \in \mathcal{L}_r([p]^n)$ podemos associar uma matriz

$$B(\mathbf{L}) = [\boldsymbol{\theta}_{i,j}(\mathbf{L})],$$

com

$$\begin{cases} \boldsymbol{\theta}_{i,j}(\mathbf{L}) = 0 & ; L(\mathbf{x}_j) \neq i - 1 \\ \boldsymbol{\theta}_{i,j}(\mathbf{L}) = 1 & ; L(\mathbf{x}_j) = i - 1 \end{cases}$$

Sendo \mathbf{T}_p obtida eliminando a primeira linha igual a $\frac{1}{\sqrt{p^n}} \mathbf{1}_p^t$ duma matriz ortogonal de ordem p e tomando-se

$$A(\mathbf{L}) = \mathbf{T}_p B(\mathbf{L}),$$

tem-se a matriz de projeção ortogonal

$$\mathbf{Q}(L) = A(L)^t A(L); L \in \mathcal{L}_r([p]^n),$$

associada a L , com

$$g(L) = \text{car}(\mathbf{Q}[L]) = p - 1.$$

Estas matrizes de projeção ortogonal são mutuamente ortogonais e

$$\sum_{L \in \mathcal{L}_r([p]^n)} \mathbf{Q}(L) = I_{p^n} - \frac{1}{p^n} \mathbf{1}_{p^n} \mathbf{1}_{p^n}^t,$$

tendo-se ainda

$$\sum_{L \in \mathcal{L}_r([p]^n)} g(L) = p^n - 1.$$

Logo pondo-se

$$\mathbf{Q}_0 = \frac{1}{p^n} \mathbf{1}_{p^n} \mathbf{1}_{p^n}^t$$

tem-se

$$\mathbf{Q}_0 + \sum_{L \in \mathcal{L}_r([p]^n)} \mathbf{Q}(L) = I_{p^n},$$

pelo que $\mathcal{A}(p^n)$ com

$$\mathcal{A}(p^n) = \{\mathbf{Q}_0; \mathbf{Q}(L), L \in \mathcal{L}_r([p]^n)\},$$

será regular e completa.

2.7 - Cruzamento e aninhamentos de fatores

Suponhamos que L grupos com u_1, \dots, u_L fatores que aninham, cruzam tendo-se um modelo com cruzamento e aninhamento. As hipóteses de ausência de efeitos e interações nesses modelos estão associadas aos vetores de

$$\Gamma = \{\mathbf{h}^L; \mathbf{h}_\ell = 0, \dots, \mathbf{h}_\ell = 1, \dots, L\}.$$

O valor médio geral do modelo estará associado a $\mathbf{0}_L$; os efeitos dos fatores do ℓ – ésimo grupo aos vetores de Γ cuja única componente não nula é a ℓ – ésima que será o índice desses fatores.

Se \mathbf{h} tiver mais de que uma componente não nula, estará associado às interações dos fatores de grupos distintos indicados por essas componentes não nulas.

As hipóteses $H_0(\mathbf{h})$ de ausência dos efeitos ou interações associadas as $\mathbf{h} \in \Gamma$ podem ser [ver Fonseca et al. (2003)] escritas como

$$H_0(\mathbf{h}): \mu \in \Delta(\mathbf{h}),$$

sendo a matriz de projeção ortogonal sobre $\Delta(\mathbf{h})$ dada por

$$Q(\mathbf{h}) = \mathbf{B}(\mathbf{h})^t \mathbf{B}(\mathbf{h}), \mathbf{h} \in \Gamma.$$

Com, [ver, Fonseca et al. (2003)]

$$\mathbf{B}(\mathbf{h}) = \otimes_{\ell} \mathbf{B}_{\ell}(\mathbf{h}_{\ell}); \mathbf{h} \in \Gamma$$

onde

$$\mathbf{B}_{\ell}(\mathbf{h}_{\ell}) = \left(\otimes_{\mathbf{h}'=1}^{\mathbf{h}_{\ell}-1} I_{a_{\ell}}(\mathbf{h}') \right) \otimes \mathbf{T}_{a_{\ell}}(\mathbf{h}_{\ell}) \otimes \left(\otimes_{\mathbf{h}''=\mathbf{h}_{\ell}+1}^{u_{\ell}} \frac{1}{\sqrt{a_{\ell}(\mathbf{h}'')}} \cdot \mathbf{1}_{a_{\ell}}^t(\mathbf{h}'') \right)$$

sendo $a_{\ell}(\mathbf{h})$ o número de níveis de \mathbf{h}_{ℓ} – éssimo fator do grupo ℓ e \mathbf{T}_r a matriz obtida retirando a primeira linha igual a $\frac{1}{\sqrt{r}} \mathbf{1}_r^t$ duma matriz ortogonal $r \times r$.

$Q(\mathbf{h}_{\ell})$ tem característica

$$g(\mathbf{h}^{\ell}) = \prod_{\ell=1}^L \left(\prod_{\mathbf{h}'=0}^{\mathbf{h}_{\ell}} a_{\ell, \mathbf{h}'} - \prod_{\mathbf{h}''=0}^{\mathbf{h}_{\ell}-1} a_{\ell, \mathbf{h}''} \right)$$

tomando-se $a_{\ell,0} = 1, \ell = 1, \dots, u_{\ell}$.

2.8 - Expansões de Distribuições

Expansões em série de distribuições podem ser úteis por exemplo no cálculo de quantis.

Muitas vezes parte-se de expansões em serie de densidades para integrando termo a termo se obtém a correspondente expansão da distribuição.

As expansões usuais para densidades são/ou densidades estandardizadas com recurso por exemplo, [ver Stuart & Ord (2000, p.227)] a polinómios de Chelychw-Hermite dados por

$$H_r(x) = x^r - \frac{r^{[2]}}{2 \cdot 1!} x^{r-2} + \frac{r^{[4]}}{2^2 \cdot 2!} x^{r-4} - \frac{r^{[6]}}{2^3 \cdot 3!} x^{r-6}$$

com $r^{[s]} = r(r-1)\dots(r-s+1)$

Assim sendo $f(x)$ uma densidade estandardizada obtém-se [ver Stuart & Ord (2000, p.229)]

$$f(x) = n(x) \left(1 + \frac{1}{2}(\mu_2 - 1)H_2(x) + \frac{1}{6}\mu_3H_3(x) + \frac{1}{24}(\mu_4 - 6\mu_2 + 3)H_4(x) + \dots \right)$$

com $n(x)$ a densidade normal estandardizada.

Utilizando as transformações de Fourier dos termos da última expansão de Edgeworth obtém-se [ver Stuart & Ord (2000, p.229)] respectivas expansões uma das quais será utilizada na obtenção de elipsoide de confiança.

Capítulo 3

No presente capítulo descrevem-se os modelos individuais destacando-se: a formulação, ajustamento, determinação dos cumulantes da 4ª ordem, apresentam-se os casos particulares de modelos nomeadamente: Multi-Normal, Multi-Normal Generalizados, Multi-Student Generalizados, estes três tipos são simétricos, seguem-se os três tipos não simétricos: Multi-Gama, Multi-Gumbel e Multi-Beta e por fim Multi-Típico. Em seguida introduzem-se comutação entre matrizes para obter o BQUE (Best Quadratic Unbiased Estimators) para componentes de variância e BLUE (Best Linear Unbiased Estimators) para o vetor β dos coeficientes segue-se a construção de elipsoides de confiança e de intervalos de predição, modelos com vetores de erro, testes F em particular casos aninhados e cruzamentos.

3 - Modelos Bi-aditivos

3.1 - Modelos individuais

Estimamos os parâmetros relevantes dos modelos considerando-se para além do caso geral casos particulares com interesse.

3.1.1 - Formulação

Os modelos bi-aditivos são de forma

$$Y = X\beta + \sum_{i=1}^w X_i Z_i$$

onde β é vetor fixo e os Z_1, \dots, Z_w aleatórios e independentes. Admite-se ainda ter-se $Z_i = (Z_{i,1}, \dots, Z_{i,c_i})$, com $Z_{i,1}, \dots, Z_{i,c_i}$ i.i.d. tendo cumulantes de ordem r , $\mathcal{X}_{r,i}$, $i =$

$1, \dots, w$. Em particular ter-se-á $\mathcal{X}_{1,i} = 0, i = 1, \dots, w$, quando os valores médios das componentes forem nulos e põe-se

$$\sigma_i^2 = \mathcal{X}_{2,i}, i = 1, \dots, w.$$

A matriz de covariância de \mathbf{Y} , será

$$\Sigma(\mathbf{Y}) = \sum_{i=1}^w \sigma_i^2 \mathbf{M}_i$$

com $\mathbf{M}_i = \mathbf{X}_i \mathbf{X}_i^t, i = 1, \dots, w$, dada a independência dos $\mathbf{Z}_1 \dots \mathbf{Z}_w$

3.1.2 - Ajustamento

Seja $\Omega = R(\mathbf{X})$ o espaço imagem de \mathbf{X} , com complemento ortogonal Ω^\perp para o qual se terá uma base ortonormada $\{\alpha_1, \dots, \alpha_n\}$. Seja o vetor

$$\dot{\mathbf{Y}}_j = \alpha_j^t \mathbf{Y} = \sum_{i=1}^w \alpha_j^t \mathbf{X}_i \mathbf{Z}_i, j = 1, \dots, \dot{n}$$

ponhamos

$$\alpha_j^t \mathbf{X}_i = (a_{\{j,i,1\}} \cdot a_{\{j,i,\ell\}} \cdot a_{\{j,i,c_i\}}), j = 1, \dots, \dot{n}, i = 1, \dots, w$$

então, representando por $\Theta_r(\dot{\mathbf{Y}}_j)$ o r -ésimo cumulante de $\dot{\mathbf{Y}}_j, r = 2, 3, j = 1, \dots, \dot{n}$, teremos.

$$\Theta_r(\dot{\mathbf{Y}}_j) = \sum_{i=1}^w \mathcal{L}_{i,j}(r) \mathcal{X}_{r,j}$$

com

$$\mathcal{L}_{j,i}(r) = \sum_{\ell=1}^{c_i} a_{j,i,\ell}^r, j = 1, \dots, \dot{n}, i = 1, \dots, w$$

pondo-se, com $r = 2, 3$

$$\mathbf{B}(r) = [\mathcal{L}_{i,j}(r)]$$

$$\Theta(r) = (\Theta_r(\dot{\mathbf{Y}}_1), \dots, \Theta_r(\dot{\mathbf{Y}}_{\dot{n}}))$$

e

$$\mathcal{X}_r = (\mathcal{X}_{r,1} \dots \mathcal{X}_{r,w})$$

tem-se então

$$\Theta_r = \mathbf{B}(r) \mathcal{X}_r, r = 2, 3.$$

Para Θ_r tem-se o estimador

$$\tilde{\Theta}_r = (\dot{Y}_1^r \dots \dot{Y}_n^r), r = 2, 3, \quad (3.1)$$

o que origina o estimador de mínimos quadrados

$$\tilde{\mathcal{X}}_r = (\mathbf{B}(r)^t \mathbf{B}(r))^+ \mathbf{B}(r)^t \tilde{\Theta}_r, r = 2, 3 \quad (3.2)$$

onde, tal como atrás, + indica inversa de Moore-Penrose.

Em particular tem-se

$$\tilde{\mathcal{X}}_2 = (\tilde{\sigma}_1^2 \dots \tilde{\sigma}_w^2),$$

com σ_i^2 componentes de variância de \mathbf{Z}_i , $i = 1, \dots, w$.

Pondo-se $\tilde{\sigma}_i^2 = \tilde{\mathcal{X}}_{2,i}$, $i = 1, \dots, w$, tem-se, para \mathbf{Y} a matriz de covariância estimada

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

o que, [ver Kariya & Kurata (2004)] origina os estimadores de mínimos quadrados generalizados

$$\tilde{\beta} = (\mathbf{X}^t \tilde{\Sigma}(\mathbf{Y})^{-1} \mathbf{X})^{-1} \mathbf{X}^t \tilde{\Sigma}(\mathbf{Y})^{-1} \mathbf{Y},$$

quando os $\mathbf{Z}_1 \dots \mathbf{Z}_w$ têm vetores médios nulos.

Admitamos agora ter o modelo bi-aditivo

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

em que as componentes de \mathbf{Z}_i têm parâmetro de localização λ_i , $i = 1, \dots, w$

Então as componentes de

$$\mathbf{Z}_i^0 = \mathbf{Z}_i - \mathbf{1}_{c_i} \lambda_i, i = 1, \dots, w,$$

terão parâmetro de localização nulo, tendo-se

$$\mathbf{X}_i \mathbf{Z}_i = \mathbf{X}_i \mathbf{Z}_i^0 + \mathbf{X}_i \mathbf{1}_{c_i}, i = 1, \dots, w.$$

Tomando-se

$$\begin{cases} \mathbf{X}^0 = [\mathbf{X} \mathbf{X}_1 \mathbf{1}_{c_1} \dots \mathbf{X}_w \mathbf{1}_{c_w}] \\ \boldsymbol{\beta}^0 = [\boldsymbol{\beta}^t \lambda_1, \dots, \lambda_w]^t \end{cases},$$

teremos

$$Y = X^o \beta^o + \sum_{i=1}^w X_i Z_i^o, i = 1, \dots, w.$$

Observe-se que as componentes de $Z_i^o, i = 1, \dots, w$ têm os mesmos cumulantes de ordem $r > 1$ que $Z_i, i = 1, \dots, w$, podendo estimá-los a partir dos resultados anteriores (expressões (3.1) e (3.2)).

Por outro dado teremos

$$\tilde{\beta}^o = (X^{o^t} \tilde{\Sigma}(Y)^{-1} X^o)^+ X^{o^t} \tilde{\Sigma}(Y) Y.$$

3.1.3 - Cumulantes de quarta ordem para os modelos bi-aditivos

Dada uma variável aleatória U com variância $\sigma^2(U)$ e quarto momento central $\mu_4(U)$, o respetivo cumulante de quarta ordem é, [ver Balakrishnan et al. (1998, p.49-54) e Douglas (1980)]

$$\mathcal{X}_4(U) = \mu_4(U) - 3(\sigma^2(U))^2.$$

Assim para se estimar $\mathcal{X}_4(U)$ tem de se estimar $\mu_4(U)$ e $(\sigma^2(U))^2$. Surge assim o problema de estimar $(\sigma^2(U))^2$. Uma primeira solução é trabalharmos com pares de vetores $(Y(1), Y(2))$ *i. i. d.* A partir destes pares obtêm-se os estimadores $\tilde{\sigma}^2(U, 1)$ e $\tilde{\sigma}^2(U, 2)$ os quais serão *i. i. d.* e dar-nos-ão

$$(\tilde{\sigma}^2(U))^2 = \tilde{\sigma}^2(U, 1) \tilde{\sigma}^2(U, 2).$$

Por outro lado, teremos os estimadores *i. i. d.* $(\tilde{\mu}'_r(1), \tilde{\mu}'_r(2))$ dos momentos de ordem r relativos á origem que serão *i. i. d.* $r = 2, 3, 4$.

Tomando-se

$$\tilde{\mu}_4 = \tilde{\mu}'_4 - 4 \tilde{\mu}'_3 \tilde{\mu}'_1 + 6 \tilde{\mu}'_2{}^2 - 3 \tilde{\mu}'_1{}^4$$

teremos

$$\tilde{\mathcal{X}}_4 = \tilde{\mu}_4 - 3(\tilde{\sigma}^2)^2.$$

Se se dispuser de parâmetros de localização (λ), dispersão (δ) e forma (τ), teremos os cumulantes

$$\mathcal{X}_r(\lambda, \delta, \tau) = \mathcal{X}_r(0, \delta, \tau) = \delta^r \mathcal{X}_r(0, 1, \tau); r = 2, 3$$

vindo para $r = 2$

$$\delta = \sqrt{\frac{\mathcal{X}_2(\lambda, \delta, \tau)}{\mathcal{X}_2(0, 1, \tau)}} = \sqrt{\frac{\mathcal{X}_2(0, \delta, \tau)}{\mathcal{X}_2(0, 1, \tau)}}. \quad (3.3)$$

Para estimar τ temos

$$\frac{\mathcal{X}_2(0, \delta, \tau)^3}{\mathcal{X}_3(0, 1, \tau)^2} = \frac{\mathcal{X}_2(0, 1, \tau)^3}{\mathcal{X}_3(0, 1, \tau)^2} = \mathfrak{h}(\tau)$$

Os cumulantes são dados, para cada tipo, por funções conhecidas o que, neste caso, nos permite obter a expressão da função $\mathfrak{h}(\tau)$ e, conseqüentemente de $\tilde{\mathfrak{h}}(\tau)$. Como o valor de $\tilde{\mathfrak{h}}(\tau)$ pode ser calculado obtém-se uma equação em τ cuja solução é $\tilde{\tau}$. O que nos permite, a partir dos $\tilde{\mathcal{X}}_r(0, \delta, \tau)$ obtidos no ajustamento obter o estimador

$$\tilde{\mathfrak{h}}(\tau) = \frac{\tilde{\mathcal{X}}_2(0, \delta, \tau)^3}{\tilde{\mathcal{X}}_3(0, \delta, \tau)^2}$$

resolvendo-se esta equação em ordem a τ obtém-se $\tilde{\tau}$.

Em seguida obtém-se

$$\tilde{\delta} = \sqrt{\frac{\tilde{\mathcal{X}}_2(0, \delta, \tau)}{\mathcal{X}_2(0, 1, \tilde{\tau})}}$$

, sendo $\mathcal{X}_2(0, 1, \tilde{\tau})$ obtido substituindo τ por $\tilde{\tau}$ na expressão do segundo cumulante com $\delta = 1$.

Estimados δ e τ teremos além disso os estimadores $\tilde{\mathcal{X}}_r(0, \delta, \tau) = \tilde{\delta}^r \mathcal{X}_r(0, 1, \tilde{\tau})$, para os $\mathcal{X}_r(0, \delta, \tau)$, $r = 4, \dots$

Se a distribuição não tiver parâmetro de forma a expressão (3.3) aligeira-se para

$$\delta = \sqrt{\frac{\mathcal{X}_2(0, \delta)}{\mathcal{X}_2(0, 1)}} \quad (3.4)$$

tendo-se o estimador

$$\tilde{\delta} = \sqrt{\frac{\tilde{\mathcal{X}}_2(\lambda, \delta)}{\mathcal{X}_2(0, 1)}}.$$

Observe-se que, como vimos se podem incorporar os parâmetros de localização $\lambda_1, \dots, \lambda_w$ correspondentes às componentes de $\mathbf{Z}_1, \dots, \mathbf{Z}_w$ no vetor dos coeficientes.

Assim as componentes de

$$\mathbf{Z}_i^0 = \mathbf{Z}_i - \lambda_i \mathbf{1}_{c_i}, \quad i = 1, \dots, w,$$

terão parâmetros $(0, \delta_i, \tau_i)$ ou $\delta_i, i = 1, \dots, w$ consoante tenhamos ou nos parâmetros de forma.

Por outro lado, com

$$\begin{cases} \mathbf{X}^0 = [\mathbf{X} \mathbf{X}_1 \mathbf{1}_{c_1} \dots \mathbf{X}_w \mathbf{1}_{c_w}] \\ \boldsymbol{\beta}^0 = [\boldsymbol{\beta}^t \lambda_1 \dots \lambda_w]^t \end{cases}$$

tem-se

$$\mathbf{Y} = \mathbf{X} \boldsymbol{\beta} + \sum_{i=1}^w \mathbf{X}_i \mathbf{Z}_i = \mathbf{X}^0 \boldsymbol{\beta}^0 + \sum_{i=1}^w \mathbf{X}_i \mathbf{Z}_i^0.$$

Podendo-se utilizar a segunda formulação do modelo em que os $\lambda_1 \dots \lambda_w$ são estimados ao estimar-se $\boldsymbol{\beta}^0$.

3.1.4 - Casos particulares de modelos bi-aditivos individuais

Vamos considerar vários tipos para as distribuições das componentes de $\mathbf{Z}_1, \dots, \mathbf{Z}_w$ com vista a facilitar o ajustamento dos modelos correspondentes.

Os três primeiros tipos são:

- ✓ Multi-Normal
- ✓ Multi-Normal generalizado
- ✓ Multi-Student generalizado,

no primeiro caso trabalharemos com densidades normais, no segundo com densidades normais generalizadas.

$$f_n(x|\lambda, \delta, r) = \frac{1}{\delta I_{2r}} \left(\frac{x - \lambda}{\delta} \right)^{2r} e^{-\frac{1}{2} \left(\frac{x - \lambda}{\delta} \right)^2}, -\infty < x < +\infty$$

onde

$$I_s = \int_{-\infty}^{+\infty} u^s e^{-\frac{u^2}{2}} du.$$

Segue-se o terceiro caso em que se consideram densidades f generalizadas

$$f(x|\lambda, \delta, m) = \frac{1}{\delta} \frac{\Gamma(m)}{\Gamma\left(\frac{1}{2}\right) \Gamma\left(m - \frac{1}{2}\right)} \frac{1}{\left[1 + \left(\frac{x - \lambda}{\delta}\right)^2\right]^m}$$

Seguem-se três tipos com densidades não simétricas:

- ✓ Multi-Gama

- ✓ Multi-Gumbel
- ✓ Multi-Beta.

No primeiro caso as densidades são de forma

$$f(x|\lambda, \delta, \ell) = \frac{1}{\delta \Gamma(\ell)} \left(\frac{x - \lambda}{\delta} \right)^{\ell-1} e^{-\left(\frac{x-\lambda}{\delta}\right)}, x > \lambda$$

enquanto no segundo temos

$$f(x|\lambda, \delta) = \frac{1}{\delta} e^{-\left(\frac{x-\lambda}{\delta}\right)} e^{-e^{-\left(\frac{x-\lambda}{\delta}\right)}}; -\infty < x < +\infty.$$

Em ambos estes tipos existem parâmetros de localização e dispersão. O terceiro tipo não simétrico tem densidades beta

$$f(x|p, q) = \frac{1}{B(p, q)} x^{p-1} (1-x)^{q-1}; 0 < x < 1,$$

Com, [ver Taylor (1955, p.680)],

$$B(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}.$$

Neste tipo não teremos parâmetros de localização e dispersão, mas assim pares (p, q) de parâmetros de forma.

Por último consideramos o caso de modelos em que as distribuições das componentes dos vetores $\mathbf{Z}_1, \dots, \mathbf{Z}_w$ podem pertencer a tipos distintos.

O nosso objetivo ao considerarmos estes tipos foi pôr em evidencia a flexibilidade dos modelos bi-aditivos.

3.1.4.1 – Modelos Multi-Normais

Nestes modelos as componentes dos vetores $\mathbf{Z}_1, \dots, \mathbf{Z}_w$, têm densidades normais. As densidades normais estandardizadas são

$$n(y|0, 1) = \frac{e^{-y^2/2}}{\sqrt{2\pi}}; -\infty < y < +\infty.$$

Obtendo-se a função geradora de momento

$$\varphi(t|0,1) = \int_{-\infty}^{+\infty} \frac{e^{ty - y^2/2}}{\sqrt{2\pi}} dy = e^{-t^2/2} \int_{-\infty}^{+\infty} \frac{e^{-t^2/2 + ty - y^2/2}}{\sqrt{2\pi}} dy =$$

$$= e^{t^2/2} \int_{-\infty}^{+\infty} \frac{e^{-(y-t)^2/2}}{\sqrt{2\pi}} dy.$$

Fazendo-se a transformação $y = u + t$ vem

$$\varphi(t|0,1) = e^{t^2/2} \int_{-\infty}^{+\infty} \frac{e^{-u^2/2}}{\sqrt{2\pi}} du = e^{t^2/2}$$

logo a função geradora dos cumulantes será

$$\psi(t|0,1) = t^2/2,$$

tendo-se

$$\begin{cases} \mathcal{X}_1(|0,1) = \psi'(0|0,1) = 0 \\ \mathcal{X}_2(|0,1) = \psi''(0|0,1) = 1 \\ \mathcal{X}_r(|0,1) = \psi^{<r>}(0|0,1) = 0, \quad r > 2. \end{cases}$$

3.1.4.2 – Modelos Multi-Normais Generalizados

Nestes modelos as componentes dos vetores $\mathbf{Z}_1, \dots, \mathbf{Z}_w$ têm densidade da forma

$$f(x|\lambda, \delta, r) = \frac{1}{\delta I_{2r}} \left(\frac{x - \lambda}{\delta} \right)^{2r} e^{-\frac{1}{2} \left(\frac{x-\lambda}{\delta} \right)^2}; \quad -\infty < x < +\infty,$$

com

$$I_s = \int_{-\infty}^{+\infty} u^s e^{-\frac{u^2}{2}} du.$$

Integrando por partes obtém-se a fórmula de recorrência

$$I_s = (s - 1)I_{s-2}$$

sabendo-se ainda que

$$I_0 = \sqrt{2\pi} \int_{-\infty}^{+\infty} n(x|0,1) dx = \sqrt{2\pi}.$$

e que, por simetria, $I_1 = 0$.

Igualmente por simetria vê-se que

$$\mu'_{2\ell+1}(|0,1,r) = \frac{1}{I_{2r}} \int_{-\infty}^{+\infty} x^{2\ell+1} x^{2r} e^{-\frac{x^2}{2}} dx = 0, \ell = 0,1,\dots,$$

e por recorrência que

$$\mu'_{2\ell}(|0, 1, r) = \frac{I_{2\ell+2r}}{I_{2r}}$$

vindo

$$\mu(|0, 1, r) = 0.$$

Pelo que os momentos relativos á origem das densidades estandardizadas coincidem com os momentos centrais.

Tem-se ainda

$$\mu_{2\ell+1}(|0, 1, r) = 0$$

igualmente por simetria, e

$$\begin{cases} \mu_2(|0, 1, r) = \frac{I_{2+2r}}{I_{2r}} = 2r + 1 \\ \mu_4(|0, 1, r) = \frac{I_{4+2r}}{I_{2r}} = 2r + 3, \end{cases}$$

pelo que, para a distribuição estandardizada temos os seguintes cumulantes:

- $\mathcal{X}_1(|0, 1, r) = 0$
- $\mathcal{X}_2(|0, 1, r) = \mu_2(|0, 1, r) = 2r + 1$
- $\mathcal{X}_3(|0, 1, r) = \mu_3(|0, 1, r) = 0$
- $\mathcal{X}_4(|0, 1, r) = \mu_4(|0, 1, r) - 3\mu_2^2(|0, 1, r) = (2r + 3) - 3(2r + 1)^2$

3.1.4.3 - Modelos Multi-Student Generalizados

Admitimos agora que as componentes de $\mathbf{Z}_1, \dots, \mathbf{Z}_w$ têm densidade t de Student generalizada

$$f(x|m, \lambda, \delta) = \frac{1}{\delta} \frac{\Gamma(m)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(m - \frac{1}{2}\right)} \frac{1}{\left[1 + \left(\frac{x - \lambda}{\delta}\right)^2\right]^m}; \quad -\infty < x < +\infty,$$

com $m \geq 3$ de forma a terem momentos de quarta ordem.

A densidade estandardizada é dada por, [ver Stuart, & Ord (2000, p.75)]

$$f(x|m, 0, 1) = \frac{\Gamma(m)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(m - \frac{1}{2}\right)} \frac{1}{(1 + x^2)^m}; \quad -\infty < x < +\infty,$$

tendo-se, por simetria

$$\mu(|m, 0, 1) = 0,$$

o que implica que

$$\mu_r(|m, 0, 1) = \mu'_r(|m, 0, 1).$$

Também por simetria vê-se que

$$\mu_{2r+1}(|m, 0, 1) = 0, r = 0, 1,$$

tendo-se ainda, [ver Stuart & Ord (2000, p.75)

$$\mu_{2r}(|m, 0, 1) = \frac{\Gamma\left(r + \frac{1}{2}\right) \Gamma\left(m - r - \frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right) \Gamma\left(m - \frac{1}{2}\right)}; 2r < 2m - 1.$$

Assim tem-se

$$\left\{ \begin{array}{l} \mu(|m, 0, 1) = 0 \\ \mu_2(|m, 0, 1) = \frac{\Gamma\left(\frac{3}{2}\right) \Gamma\left(m - \frac{3}{2}\right)}{\Gamma\left(\frac{1}{2}\right) \Gamma\left(m - \frac{1}{2}\right)} \\ \mu_3(|m, 0, 1) = 0 \\ \mu_4(|m, 0, 1) = \frac{\Gamma\left(\frac{5}{2}\right) \Gamma\left(m - \frac{5}{2}\right)}{\Gamma\left(\frac{1}{2}\right) \Gamma\left(m - \frac{1}{2}\right)}, \end{array} \right.$$

vindo

$$\left\{ \begin{array}{l} \mathcal{X}_r(|m, 0, 1) = 0; r = 1, 3 \\ \mathcal{X}_2(|m, 0, 1) = \mu_2(|m, 0, 1) \\ \mathcal{X}_4(|m, 0, 1) = \frac{\Gamma\left(\frac{5}{2}\right) \Gamma\left(m - \frac{5}{2}\right)}{\Gamma\left(\frac{1}{2}\right) \Gamma\left(m - \frac{1}{2}\right)} - 3 \left[\frac{\Gamma\left(\frac{3}{2}\right) \Gamma\left(m - \frac{3}{2}\right)}{\Gamma\left(\frac{1}{2}\right) \Gamma\left(m - \frac{1}{2}\right)} \right]^2, \end{array} \right.$$

bem como

$$\gamma_1(|m) = 0$$

e

$$\gamma_2(|m) = \frac{\Gamma\left(\frac{5}{2}\right) \Gamma\left(m - \frac{5}{2}\right)}{\Gamma\left(\frac{1}{2}\right) \Gamma\left(m - \frac{1}{2}\right)} / \left[\left(\frac{\Gamma\left(\frac{3}{2}\right) \Gamma\left(m - \frac{3}{2}\right)}{\Gamma\left(\frac{1}{2}\right) \Gamma\left(m - \frac{1}{2}\right)} \right)^2 - 3 \right],$$

observe-se que fazendo $\ell = 1$ se obtêm os resultados para modelos em que as componentes dos vetores $\mathbf{Z}_1, \dots, \mathbf{Z}_w$ têm densidades exponenciais

$$\ell(x|\lambda, \delta) = \frac{e^{-\frac{x-\lambda}{\delta}}}{\delta}, x \geq \lambda.$$

3.1.4.4 – Modelos Multi-Gama

Nestes modelos as componentes dos vetores $\mathbf{Z}_1, \dots, \mathbf{Z}_w$, terão densidades gama.

As densidades gama reduzidas são dadas por

$$f(y|0, 1, \ell) = \frac{1}{\Gamma(\ell)} y^{\ell-1} e^{-y}, y > 0.$$

Obtém-se assim a função geradora de momentos

$$\varphi(t|0, 1, \ell) = \frac{1}{\Gamma(\ell)} \int_{-\infty}^{+\infty} e^{ty} y^{\ell-1} e^{-y} dy.$$

Fazendo-se a transformação

$$u = (1 - t) y$$

vem

$$\begin{aligned} \varphi(t|0, 1, \ell) &= \frac{1}{\Gamma(\ell)} \int_0^{+\infty} \left(\frac{u}{1-t}\right)^{\ell-1} e^{-u} \frac{1}{1-t} du = \\ &= \frac{1}{(1-t)^\ell} \end{aligned}$$

bem como

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

sendo

$$\psi^{<r>}(t|0, 1, \ell) = (-1)^{r+1} (r-1)! \ell (1-t)^{-r},$$

o que dá

$$\chi_r(|0, 1, \ell) = (-1)^{r+1} (r-1)! \ell, r = 1, \dots, \ell = 1, \dots$$

[ver Nelder (1966)] e [ver Laplace (1836)].

3.1.4.5 – Modelos Multi-Gumbel

As componentes dos vetores $\mathbf{Z}_1, \dots, \mathbf{Z}_w$, terão agora distribuições de Gumbel. As densidades estandardizadas correspondentes a essas distribuições são dadas por

$$f(x|0, 1) = e^{-x} e^{-e^{-x}} dx; -\infty < x < +\infty,$$

com função geradora de momentos

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

fazendo-se a transformação

$$u = e^{-x}$$

vem

$$\varphi(t|0, 1) = \int_0^{+\infty} u^{-t} e^{-u} du = \Gamma(1 - t),$$

com Γ representando a função gama. Assim a função geradora de cumulantes será

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

vindo

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

bem como

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

logo, utilizando a fórmula de Euler da derivada do produto de funções obtêm-se

$$\begin{aligned} \Gamma^{<\ell+1>}(1 - t) &= (-1)^{\ell+1} \sum_{i=0}^{\ell} \binom{\ell}{i} (-1)^i \Gamma^{<i>}(1 - t) \psi^{<\ell-i+1>}(t|0, 1) \\ &= (-1)^{\ell+1} \Gamma(1 - t) \psi^{<\ell+1>}(t|0, 1) + (-1)^{\ell+1} \sum_{i=1}^{\ell} \binom{\ell}{i} (-1)^i \Gamma^{<i>}(1 - t) \psi^{<\ell-i+1>}(t|0, 1), \end{aligned}$$

vindo

$$\Gamma^{(\ell+1)}(1) = (-1)^{\ell+1} \psi^{<\ell+1>}(0|0, 1) + (-1)^{\ell+1} \sum_{i=1}^{\ell} \binom{\ell}{i} (-1)^i \Gamma^{<i>}(1) \psi^{<\ell-i+1>}(0|0, 1),$$

o que dá a fórmula de recorrência para os cumulantes das distribuições de Gumbel reduzida

$$\mathcal{X}_{\ell+1}(|0, 1) = (-1)^{\ell+1} \Gamma^{<\ell+1>}(1) - (-1)^{\ell+1}.$$

$$\sum_{i=1}^{\ell} \binom{\ell}{i} (-1)^i \Gamma^{<i>}(1) \mathcal{X}_{\ell+1-i}(|0,1),$$

o que permite obter os

$$\begin{cases} \mathcal{X}_1(|\lambda, \delta) = \lambda + \delta \mathcal{X}_1(|0,1) \\ \mathcal{X}_\ell(|\lambda, \delta) = \delta^\ell \mathcal{X}_\ell(|0,1) \quad ; \ell > 1 \end{cases}$$

Aqui consideramos índice ℓ ao invés de r porque λ apenas intervém na expressão do cumulante da 1ª ordem.

Para mais detalhes sobre modelos Multi-Gumbel [ver Fisher e Tippett (1928), Watanabe (2013), Barbosa et al. (2014), Santos et al. (2014) e Pinheiro (2014)].

3.1.4.6 - Modelos Multi-Beta

Começemos por observar que, [ver Taylor (1995, p.680)], se tem

$$B(p, q) = \int_0^1 x^{p-1} (1-x)^{q-1} dx = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)},$$

considerando agora modelos em que as componentes dos vetores $\mathbf{Z}_1, \dots, \mathbf{Z}_w$ têm densidade

$$f(x|p, q) = \frac{1}{B(p, q)} x^{p-1} (1-x)^{q-1}, 0 < x < r,$$

sem parâmetros de localização e dispersão e pares (p, q) de parâmetros de forma.

com

$$p^{[r]} = \prod_{\ell=0}^{r-1} (p + \ell),$$

tem-se

$$\Gamma(p+r) = \Gamma(p)p^{[r]},$$

vindo

$$\frac{B(p+r, q)}{B(p, q)} = \frac{\Gamma(p+r)\Gamma(q)}{\Gamma(p+r+q)} \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} = \frac{p^{[r]}}{(p+q)^{[r]}}.$$

Ora

$$\mu'_r(p, q) = \frac{B(p+r, q)}{B(p, q)} = \frac{p^{[r]}}{(p+q)^{[r]}}$$

logo

$$\begin{cases} \mu(|p, q) = \frac{p}{p+q} \\ \sigma^2(|p, q) = \frac{p(p+1)}{(p+q)(p+q+1)} - \frac{p^2}{(p+q)^2}, \end{cases}$$

bem como

$$\begin{aligned} \frac{\sigma^2(|p, q)}{\mu^2(|p, q)} &= \frac{p+1}{p} \frac{p+q}{p+q+1} - 1 = \frac{(p+1)(p+q) - p(p+q+1)}{p(p+q+1)} = \\ &= \frac{q}{p(p+q+1)}. \end{aligned}$$

Tomando

$$h = \frac{q}{p},$$

vem

$$\mu = \frac{1}{1+h},$$

bem como

$$h = \frac{\mu - 1}{\mu},$$

assim a partir de estimadores de μ obtém-se estimadores de h

Tem-se ainda

$$\frac{\sigma^2(|p, q)}{\mu^2(|p, q)} = \frac{h}{p(p+hq+r+1)},$$

o que, a partir dos estimadores de h , de $\mu(|p, q)$ e $\sigma^2(|p, q)$ permite estimar p e daí estimar q .

Assim a partir dos resultados do ajustamento podem estimar-se os parâmetros de forma destas distribuições.

3.1.4.7 – Modelo Multi-Típicos

Nos casos particulares que consideramos as distribuições das componentes dos vetores $\mathbf{Z}_1, \dots, \mathbf{Z}_w$ pertenciam todas ao mesmo tipo. Ora nada exige que isso se verifique.

Com efeito a técnica de ajustamento apresentada assenta no uso de cumulantes e as expressões abaixo

$$\begin{cases} \mathcal{X}_1(\lambda, \delta_1, \dots) = \lambda + \delta \mathcal{X}_1(0, 1, \dots) \\ \mathcal{X}_r(\lambda, \delta, \dots) = \delta^r \mathcal{X}_r(0, 1, \dots); r > 1 \end{cases}$$

são validas independentes do tipo a que pertencem as distribuições referidas acima.

Este facto permite considerar modelos com

$$\mathbf{X}_w = \mathbf{I}_n,$$

e tendo as componentes $Z_{w,1} \dots Z_{w,n}$ de \mathbf{Z}_w distribuição $\mathcal{N}(0, \sigma^2 \mathbf{I}_n)$.

Integra-se assim nos modelos com vetores de erros o pode ser útil se as componentes $Y_1 \dots Y_n$ do vetor das observações forem medições [ver Thiele (1903)]. Por exemplo os modelos com erro são modelos Multi-Típicos. Assim com $w = 2$ as componentes do vector Z_1 podem ter distribuição de Gumbel sendo as componentes de Z_2 normais com a mesma variância. Este será um ponto a explorar de futuro. Para já observamos que, com

$$\bar{\Omega} = R([X \ X_1 \dots X_{w-1}]) \subset R^n,$$

$\mathbf{Y}_{\bar{\Omega}^\perp}$ a projeção ortogonal de \mathbf{Y} sobre $\bar{\Omega}^\perp$, e

$$\bar{g} = \dim(\bar{\Omega}),$$

$\|\mathbf{Y}_{\bar{\Omega}^\perp}\|^2$ será o produto por σ^2 dum qui-quadrado central com $n - \bar{g}$ graus de liberdade.

Como exemplo de modelos Multi-Típicos podemos dar o dos modelos com vetor de erros ver secção (3,5) em que $\mathbf{X}_w = \mathbf{I}_n$ e \mathbf{Z}_w é normal, $\mathcal{N}(0, \sigma^2 \mathbf{I}_n)$.

3.1.4.8 – Simetria

Se as componentes dos $\mathbf{Z}_1 \dots \mathbf{Z}_w$ tiverem densidades simétricas, como as dos modelos: Multi-normais, Multi-normais generalizados ou Multi-Student, para que isso aconteça basta que os parâmetros de localização sejam nulos.

3.2 – Comutações

Nesta secção vamos introduzir comutação entre matrizes para assegurarmos a obtenção de estimadores ótimos.

3.2.1 – BQUE (Best Quadratic Unbiased Estimators)

Vamos agora obter resultados ótimos para a estimação começando por admitir que as matrizes

$$\mathbf{M}_i = \mathbf{X}_i \mathbf{X}_i^t, i = 1, \dots, w$$

comutam e em seguida que

$$\mathbf{M} = \mathbf{X} \mathbf{X}^t,$$

também comuta. Utilizaremos para isso Álgebras de Jordan comutativas, AJC, que serão espaços vectoriais constituídas por matrizes simétricas que comutam e contendo

os quadrados das respectivas matrizes. Ora, [ver Seely (1972)], toda a AJC \mathcal{A} , tem uma base única, a base principal $\mathcal{b}_p(\mathcal{A})$, constituída por matrizes da projeção ortogonal mutuamente ortogonais.

Por outro lado [ver Schott (2017)] se as matrizes $\mathbf{M}_1, \dots, \mathbf{M}_w$ [$\mathbf{M}, \mathbf{M}_1, \dots, \mathbf{M}_w$] comutam existe uma matriz ortogonal P [\bar{P}] que as diagonaliza, pertencendo as matrizes à AJC \mathcal{A} [$\bar{\mathcal{A}}$] das matrizes diagonalizadas por P [\bar{P}].

Sendo $\{\mathbf{Q}_1, \dots, \mathbf{Q}_m\} = \mathcal{b}_p(\mathcal{A})$ [$\{\bar{\mathbf{Q}}_1, \dots, \bar{\mathbf{Q}}_m\} = \mathcal{b}_p(\bar{\mathcal{A}})$], teremos

$$\mathbf{M}_i = \sum_{j=1}^m \mathcal{b}_{i,j} \mathbf{Q}_j, i = 1, \dots, w,$$

e se $\mathbf{M}, \mathbf{M}_1, \dots, \mathbf{M}_w$ comutam

$$\mathbf{M}_i = \sum_{j=1}^{\bar{m}} \bar{\mathcal{b}}_{i,j} \bar{\mathbf{Q}}_j, i = 1, \dots, w + 1,$$

com $\mathbf{M} = \mathbf{M}_{w+1}$

Como vimos temos

$$\Sigma(\mathbf{Y}) = \sum_{i=1}^w \sigma_i^2 \mathbf{M}_i,$$

assim também se tem

$$\Sigma(\mathbf{Y}) = \sum_{j=1}^w \sigma_j^2 \left(\sum_{i=1}^m \mathcal{b}_{i,j} \mathbf{Q}_j \right) = \sum_{j=1}^m \gamma_j \mathbf{Q}_j,$$

com

$$\gamma_j = \sum_{i=1}^w \mathcal{b}_{i,j} \sigma_i^2, j = 1, \dots, m,$$

o que é o mesmo que

$$\boldsymbol{\gamma} = \mathbf{B}^t \boldsymbol{\sigma}^2,$$

com $\mathbf{B} = [\mathcal{b}_{i,j}]$, $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_m)$ e $\boldsymbol{\sigma}^2 = (\sigma_1^2, \dots, \sigma_w^2)$. Tem-se então, [ver Schott (2017)]

$$\boldsymbol{\sigma}^2 = (\mathbf{B}^t)^+ \boldsymbol{\gamma},$$

o que reduz o problema de estimar $\boldsymbol{\sigma}^2$ a estimar $\boldsymbol{\gamma}$.

Ora como veremos teremos estimadores ótimos para os $\gamma_1, \dots, \gamma_m$ quando as matrizes $\mathbf{M}_1, \dots, \mathbf{M}_w$ comutam. Então, $\Sigma(\mathbf{Y})$ é combinação linear de matrizes de projeção ortogonal mutuamente ortogonais logo [ver McCullagh & Nelder (1989 e 1995)], os modelos têm “Orthogonal Block Structure”, OBS, para mostrarmos que temos estimadores BQUE “Best Quadratic Unbiased Estimators” para os $\gamma_1, \dots, \gamma_m$ iremos generalizar o teorema de HSU [ver HSU (1936)].

Teorema 1: *Teorema de HSU para BQUE*

“Se o vetor aleatório \mathbf{Z} , tiver vetor médio $\boldsymbol{\mu}$, que pertence ao espaço Ω , com dimensão \hat{n} , e matriz de covariância $\sigma^2 \mathbf{I}_n$, o BQUE para σ^2 é

$$\tilde{\sigma}^2 = \frac{\mathbf{T}^t (\mathbf{I}_n - \mathbf{T}) \mathbf{T}}{n - \hat{n}},$$

sendo \mathbf{T} a matriz de projeção ortogonal sobre Ω ”.

Seja então \mathbf{Y} um vetor aleatório com vetor médio $X \boldsymbol{\beta}$ e matriz de covariância $\sum_{j=1}^m \gamma_j \mathbf{Q}_j$.

Sendo \mathbf{A}_j uma matriz cujos g_j vetores linha constituem uma base ortonormada para espaço imagem $\nabla_j = R(\mathbf{Q}_j)$, o vetor

$$\dot{\mathbf{Y}}_j = \mathbf{A}_j \mathbf{Y}, j = 1, \dots, m$$

terá vetor médio

$$\dot{\boldsymbol{\mu}}_j = \dot{\mathbf{X}}_j \boldsymbol{\beta}, j = 1, \dots, w,$$

com $\dot{\mathbf{X}}_j = \mathbf{A}_j \mathbf{X}$, e a matriz de covariância $\gamma_j \mathbf{I}_{g_j}, j = 1, \dots, w$.

Segundo o teorema de HSU e estimador BQUE para γ_j obtido a partir de $\dot{\mathbf{Y}}_j$, $j = 1, \dots, m$ será

$$\tilde{\gamma}_j = \frac{\dot{\mathbf{Y}}_j^t (\mathbf{I}_{g_j} - \mathbf{P}_j) \dot{\mathbf{Y}}_j}{g_j - p_j}, j = 1, \dots, m,$$

caso $p_j < g_j$, com \mathbf{P}_j a matriz de projeção ortogonal sobre $R(\dot{\mathbf{X}}_j)$ e $p_j = \text{car}(\mathbf{P}_j)$.

Tomemos agora

$$\mathbf{A} = [\mathbf{A}_1^t \dots \mathbf{A}_m^t]^t,$$

e

$$\dot{\mathbf{Y}} = [\dot{\mathbf{Y}}_1^t \dots \dot{\mathbf{Y}}_m^t]^t = (\dot{\mathbf{Y}}_1 \dots \dot{\mathbf{Y}}_m),$$

tendo-se

$$\mathbf{Y} = \mathbf{A}^t \dot{\mathbf{Y}}$$

e sendo os mesmos, os estimadores quadráticos para os $\gamma_1, \dots, \gamma_m$ obtidos a partir de \mathbf{Y} ou de $\dot{\mathbf{Y}}$. Então, com

$$\begin{cases} \bar{g}_0 = 0 \\ \bar{g}_j = \sum_{h=1}^j g_h, \quad j = 1, \dots, m, \end{cases}$$

e $\Sigma(\dot{\mathbf{Y}}) = [\dot{\gamma}_{\ell, h}]$, temos

$$\begin{cases} \dot{\gamma}_{\ell, h} = 0, & \ell \neq h \\ \dot{\gamma}_{\ell, \ell} = \gamma_j; & \bar{g}_{j-1} < \ell \leq \bar{g}_j, \quad j = 1, \dots, m. \end{cases}$$

Com $\mathbf{B}_j = [\mathfrak{b}_{j, \ell, h}]$, $j = 1, \dots, m$ a matriz do estimador quadrático

$$\tilde{\gamma}_j = \dot{\mathbf{Y}}^t \mathbf{B}_j \dot{\mathbf{Y}} = \sum_{\ell=1}^n \sum_{h=1}^n \mathfrak{b}_{j, \ell, h} \dot{Y}_\ell \dot{Y}_h, \quad j = 1, \dots, m$$

onde $\dot{Y}_1, \dots, \dot{Y}_n$ são as componentes de $\dot{\mathbf{Y}}$.

Como $\dot{\mathbf{Y}}$ tem a matriz de covariância diagonal por blocos

$$\Sigma(\dot{\mathbf{Y}}) = D(\gamma_1 \mathbf{I}_{g_1}, \dots, \gamma_m \mathbf{I}_{g_m}).$$

Com $\dot{\mu}_\ell$ o valor médio de \dot{Y}_ℓ , $\ell = 1, \dots, n$, temos os valores médios

$$\begin{cases} E(\dot{Y}_\ell \dot{Y}_h) = \dot{\mu}_\ell \dot{\mu}_h, & \ell \neq h \\ E(\dot{Y}_\ell^2) = \dot{\gamma}_j + \dot{\mu}_\ell^2; & \bar{g}_{j-1} < \ell \leq \bar{g}_j, \quad j = 1, \dots, m \end{cases}$$

logo

$$E(\tilde{\gamma}_j) = \sum_{\ell \neq h} \mathfrak{b}_{j, \ell, h} \dot{\mu}_\ell \dot{\mu}_h + \sum_{j=1}^m \sum_{\ell=\bar{g}_{j-1}+1}^{\bar{g}_j} \mathfrak{b}_{j, \ell, \ell} (\dot{\gamma}_j + \dot{\mu}_\ell^2)$$

pelo que, para $\tilde{\gamma}_j$ ser estimador centrado de γ_j , $j = 1, \dots, m$ temos de ter

$$\begin{cases} \mathfrak{b}_{j, \ell, h} = 0; & \ell \neq h, \quad j = 1, \dots, m \\ \mathfrak{b}_{j, \ell, \ell} = 0, & (\ell = 1, \dots, \bar{g}_{j-1}) \vee (\ell = \bar{g}_{j+1}), \quad j = 1, \dots, m, \end{cases}$$

bem como

$$\mathfrak{b}_{j, \ell, \ell} = 0, \quad \ell \in \varphi_j = \{h: \mu_h \neq 0\}, \quad j = 1, \dots, m.$$

Logo, com

$$D_j = \{\bar{g}_{j-1} + 1, \dots, \bar{g}_j\}, \quad j = 1, \dots, m,$$

para $\dot{\gamma}_j$ ser estimador centrado de γ_j , $j = 1, \dots, m$, tem de se ter

$$\tilde{\gamma}_j = \sum_{\ell \in D_j} \ell_{j,\ell,\ell} \dot{Y}_\ell^2; \quad j = 1, \dots, m,$$

com

$$\sum_{\ell \in D_j} \ell_{j,\ell,\ell} = 1; \quad j = 1, \dots, m.$$

Como $\dot{\gamma}_j$ pode ser obtido a partir de \dot{Y}_j os estimadores quadráticos centrados de γ_j obtidos a partir de Y são os mesmos que os obtidos a partir de $\dot{Y}_j, j = 1, \dots, m$. Ora os \dot{Y}_j satisfazem as condições do teorema original de HSU obtendo-se,

Teorema 2: Teorema de HSU para modelos com OBS

Se $\Sigma(Y) = \sum_{j=1}^m \gamma_j Q_j$, com $Q_j = A_j^t A_j$, se P_j for a matriz de projeção ortogonal sobre $R(\dot{X}_j)$ com $p_j = \text{car}(P_j)$ e se tiver $\dot{Y}_j = A_j Y$, o BQUE para $\gamma_j, j = 1, \dots, m$, será

$$\tilde{\gamma}_j = \frac{\dot{Y}_j^t (\mathbf{I}_{g_j} - P_j) \dot{Y}_j}{g_j - p_j}; \quad j = 1, \dots, m$$

se $p_j < g_j, j = 1, \dots, m$.

3.2.2 – BLUE (Best Linear Unbiased Estimators)

Admitimos agora que as matrizes M, M_1, \dots, M_w comutam pertencendo á CJA com base principal $\{\bar{Q}_1, \dots, \bar{Q}_m\}$, tendo-se

$$M_i = \sum_{j=1}^m \bar{\ell}_{i,j} \bar{Q}_j, \quad j = 1, \dots, w+1,$$

onde $M = X X^t = M_{w+1}$. Com

$$\varphi_i = \{j; \bar{\ell}_{i,j} \neq 0\}, \quad i = 1, \dots, w+1,$$

a matriz da projeção ortogonal sobre $R(M_i), i = 1, \dots, w+1$, é

$$P(M_i) = \sum_{j \in \varphi_i} \bar{Q}_j, \quad i = 1, \dots, w+1$$

tendo-se ainda

$$\Sigma(Y) = \sum_{i=1}^w \sigma_i^2 M_i = \sum_{j=1}^m \tilde{\gamma}_j \bar{Q}_j, \quad j = 1, \dots, m,$$

logo $\Sigma(Y)$ comuta com

$$T = P_{w+1}$$

que é a matriz de projeção ortogonal sobre $\Omega = R(\mathbf{X})$. Aliás \mathbf{T} comuta com as $\bar{Q}_1, \dots, \bar{Q}_m$ tendo o modelo *COBS* (Commutative Orthogonal Block Structure) o que, [ver Carvalho et al. (2015), Ferreira et al. (2013)], garante que

$$\tilde{\boldsymbol{\beta}} = (\mathbf{X}^t \mathbf{X})^+ \mathbf{X}^t \mathbf{Y}$$

é BLUE (Best Linear Unbiased Estimators) de $\boldsymbol{\beta}$, [ver Mexia (1990)].

3.3 – Elipsoides de Confiança

Sendo \mathbf{Y} o vetor das observações ou, no caso de pares de modelos, o respetivo valor médio, ponhamos, com $\|\mathbf{u}\|^2 = 1$

$$W(\mathbf{u}) = \mathbf{u}^t (\mathbf{Y} - \boldsymbol{\mu}) = \sum_{i=1}^w \alpha_i^t \mathbf{Z}_i,$$

onde, no caso de pares de modelos os $\mathbf{Z}_i(\ell)$, $\ell = 1, \dots, w$ são vetores médios, e

$$\alpha_i^t = \mathbf{u}^t \mathbf{X}_i = (a_{1,1}, \dots, a_{i,c_i}), i = 1, \dots, w.$$

Caso os $\mathbf{Z}_i(\ell)$, $\ell = 1, \dots, w$, têm vetores médios nulos temos

$$E(W(\mathbf{u})) = 0.$$

Sendo $\mathcal{V}_r(\mathbf{u})$ o r -ésimo cumulante de $W(\mathbf{u})$ temos

[ver Craig (1931)]

$$\mathcal{V}_r(\mathbf{u}) = \sum_{i=1}^w \mathfrak{h}_{r,i} \mathcal{X}_{r,i}, r = 2, 3, 4,$$

com

$$\mathfrak{h}_{r,i} = \sum_{j=1}^{c_i} a_{i,j}^r, r = 2, 3, 4, i = 1, \dots, w.$$

Teremos, pois, os estimadores

$$\tilde{\mathcal{V}}_r(\mathbf{u}) = \sum_{i=1}^w \mathfrak{h}_{r,i} \tilde{\mathcal{X}}_{r,i}, r = 2, 3, 4,$$

a partir dos quais para os coeficientes de assimetria e achatamento

$$\begin{cases} \rho_1(\mathbf{u}) = \mathcal{V}_3(\mathbf{u})/\mathcal{V}_2^{3/2}(\mathbf{u}) \\ \rho_2(\mathbf{u}) = \mathcal{V}_4(\mathbf{u})/\mathcal{V}_2^2(\mathbf{u}) - 3 \end{cases}$$

obtemos os estimadores

$$\begin{cases} \tilde{\rho}_1(\mathbf{u}) = \tilde{\mathcal{V}}_3(\mathbf{u})/\tilde{\mathcal{V}}_2^{3/2}(\mathbf{u}) \\ \tilde{\rho}_2(\mathbf{u}) = \tilde{\mathcal{V}}_4(\mathbf{u})/\tilde{\mathcal{V}}_2^2(\mathbf{u}) - 3, \end{cases}$$

pelo que para a distribuição $F_{\hat{w}(\mathbf{u})}(\cdot)$, de

$$\hat{w}(\mathbf{u}) = \frac{w(\mathbf{u})}{\sqrt{\hat{V}_2(\mathbf{u})}},$$

temos a expansão de Edgeworth, para detalhes [ver Martins (2008, pp.148-150) ou Simões (2017, pp.12-14)],

$$Ed_{\hat{w}(\mathbf{u})}(z) = N(z) - \frac{\tilde{\rho}_1(\mathbf{u})}{3!} n^{<2>}(z) + \frac{\tilde{\rho}_2(\mathbf{u})}{4!} n^{<3>}(z) + 10 \frac{\tilde{\rho}_1(\mathbf{u})^2}{6!} n^{<6>}(z),$$

o que nos permite obter o p-ésimo quantil estimado $\tilde{w}_p(\mathbf{u})$ de $Ed_{\hat{w}(\cdot)}(\cdot)$. Assim teremos os quantis estimados

$$\tilde{w}_p(\mathbf{u}) = \sqrt{\tilde{v}_2(\mathbf{u})} \tilde{w}_p(\mathbf{u}).$$

Vamos agora obter o elipsoide centrado na origem associado á forma quadrática $\mathbf{u}^t K \mathbf{u}$ que, dados os vetores $\mathbf{u}_1, \dots, \mathbf{u}_n$ minimiza

$$\Delta(K) = \sum_{\ell=1}^n (\tilde{w}_p(\mathbf{u}_\ell) - \mathbf{u}_\ell^t k \mathbf{u}_\ell)^2.$$

Ora, com

$$Y \in R^n$$

$$\Omega = R(X)$$

$$\{\alpha_1, \dots, \alpha_n\} = \mathcal{B} \perp (\Omega^\perp)$$

Dado um vetor $\mathbf{u} = (u_1, \dots, u_n)$ podemos considerar o vetor $\mathbf{u}^{[2]}$ com componentes $\mathbf{u}_1^2 \dots 2u_1 u_n, \mathbf{u}_2^2 \dots 2u_2 u_n, \dots, \mathbf{u}_n^2$ e dada a matriz $\mathbf{K} = [k_{\ell, \ell}]$ o vetor

$$\mathbf{s}(\mathbf{K}) = (k_{1,1} \dots k_{1,n} \quad k_{2,2} \dots k_{2,n} \dots k_{n,n}).$$

Temos

$$\mathbf{u}_\ell^t K \mathbf{u}_\ell = \mathbf{u}_\ell^{[2]t} \mathbf{s}(\mathbf{K}), \quad \ell = 1, \dots, m.$$

Observe-se que $\mathbf{u}^{[2]}$ e $\mathbf{s}(\mathbf{K})$ têm $\frac{n(n+1)}{2}$ componentes. Sendo $\{\mathbf{u}_1 \dots \mathbf{u}_n\}$ uma base ortonormada para R^n (constituída por n vetores com norma 1 ortogonais dois a dois), tomemos

$$\begin{cases} U^{[2]} = [\mathbf{u}_1^{[2]}, \dots, \mathbf{u}_n^{[2]}] \\ \tilde{\mathbf{w}}_p = [\tilde{w}_p(\mathbf{u}_1) \dots \tilde{w}_p(\mathbf{u}_n)], \end{cases}$$

pretendendo-se minimizar $\|\tilde{\mathbf{w}}_p - U^{[2]}\lambda\|^2$ com $\lambda = \mathbf{s}(K)$ e, portanto,

$$s(K(\lambda)) = \lambda.$$

Dado ter-se

$$E(\tilde{\mathbf{w}}_p) = U^{[2]}\lambda_p,$$

teremos os LSE

$$\begin{cases} \tilde{\lambda}_p = (U^{[2]t}U^{[2]})^{-1}U^{[2]t}\tilde{\mathbf{w}}_p \\ \tilde{K}_p = K(\tilde{\lambda}_p) \end{cases}$$

Observe-se que, dado um vetor \mathbf{u} com $\|\mathbf{u}\|^2 = 1$, o elipsoide passa pela direção dada por \mathbf{u} à distância $\mathbf{u}^t\tilde{K}_p\mathbf{u}$ do centro. Centrando-se o elipsoide no “ponto médio”

$$\tilde{\boldsymbol{\mu}} = X\tilde{\boldsymbol{\beta}},$$

o contorno passará pelo ponto de semi-recta com direção dada por \mathbf{u} e a origem em $\tilde{\boldsymbol{\mu}}$ à distância $\mathbf{u}^t\tilde{K}_p\mathbf{u}$ da sua origem.

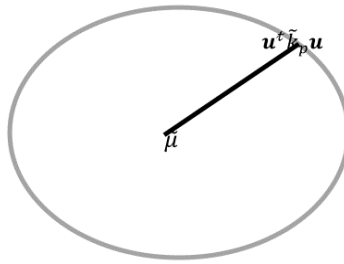


Fig. 3.1 – Elipsoide de nível p (aproximado) centrado no ponto médio

Atendendo á última expressão vemos que a precisão com que se “coloca” o ponto médio é tanto maior quanto mais precisa for a estimação de $\tilde{\boldsymbol{\beta}}$. Assim se o modelo tiver COBS e sendo $\tilde{\boldsymbol{\beta}}$ BLUE estaremos numa situação de maior precisão do que quando o modelo não tiver COBS. Com efeito o ponto central do elipsoide é o ponto a fim de $X\tilde{\boldsymbol{\beta}}$. Assim se o modelo tiver COBS (Cummutative Orthogonal Block Structure) tendo-se portando BLUE para $\tilde{\boldsymbol{\beta}}$ tem-se uma maior precisão na localização do elipsoide.

3.4 - Intervalos de Predição

A abordagem usual para predição fornece um BLUP (Best Linear Unbiased Predictor), centrado, com variância mínima para uma futura observação. Apresentamos agora

uma alternativa em que se obtém um intervalo de predição com uma probabilidade (aproximada) de conter a futura observação. Sejam $\hat{\mathbf{x}}, \hat{\mathbf{x}}_1, \dots, \hat{\mathbf{x}}_w$ com k, c_1, \dots, c_w componentes, que serão os vetores dos valores das variáveis controladas para os quais se pretende obter o intervalo de predição.

Pondo

$$\hat{\mathbf{u}} = [\hat{\mathbf{x}}^t \hat{\mathbf{x}}_1^t \dots \hat{\mathbf{x}}_w^t]^t$$

, caso os $\mathbf{Z}_1 \dots \mathbf{Z}_w$ tenham vetores médios nulos, e utilizando os resultados da secção anterior podemos obter $(\tilde{w}_{p/2}(\hat{\mathbf{u}}), \tilde{w}_{1-p/2}(\hat{\mathbf{u}}))$ e a partir daqui o intervalo de predição de nível $(1 - p)$,

$$[\hat{\mathbf{x}}^t \tilde{\boldsymbol{\beta}} + \tilde{w}_{p/2}(\hat{\mathbf{u}}); \hat{\mathbf{x}}^t \tilde{\boldsymbol{\beta}} + \tilde{w}_{1-p/2}(\hat{\mathbf{u}})].$$

Se o modelo tiver COBS teremos um BLUE para $\boldsymbol{\beta}$ tendo-se maior precisão na estimação dos extremos dos intervalos de predição.

3.5 - Modelos com vetores de erros

3.5.1 - Estimação

Vamos agora considerar modelos para os quais se conseguem obter testes F aproximando-nos da análise de variância. Consegue-se assim ter testes de nível conhecido em vez dos testes obtidos a partir de elipsoide de confiança em que o nível é aproximado.

Nestes modelos tem-se $w = 2, X_2 = I_n$ e $\mathbf{Z}_2 = \mathbf{e}$ com $\mathbf{e} \sim N(0, \sigma^2 I_n)$ (\mathbf{e} normal com vetor médio $\mathbf{0}_n$ e matriz de covariância $\sigma^2 I_n$). Quanto a \mathbf{Z}_1 apenas se exige que tenha componentes *i. i. d* e cumulantes $\mathcal{X}_r, r = 2, 3$, tendo-se em particular $\mathcal{X}_1 = 0$ e $\mathcal{X}_2 = \sigma_2^2$. Teremos então

$$\mathbf{Y} = \mathbf{X} \boldsymbol{\beta} + X_1 \mathbf{Z}_1 + \mathbf{e}$$

$\mathbf{M}_2 = \mathbf{X}_2 \mathbf{X}_2^t = I_n I_n^t = I_n$ observação: porque as matrizes $\mathbf{M}_1 \mathbf{M}_2$ comutam, como $\mathbf{M}_1 = \mathbf{X}_1 \mathbf{X}_1^t$ comuta com $\mathbf{M}_2 = I_n$ o modelo terá OBS. Sendo $\mathbf{M}_i, i = 1, 2$, simétrica temos

$$\mathbf{M}_1 = \sum_{j=1}^{m-1} \varrho_{1,j} \mathbf{Q}_j,$$

com $\mathbf{Q}_1, \dots, \mathbf{Q}_{m-1}$ matrizes de projeção ortogonal mutuamente ortogonais e, se $R(\mathbf{M}_i)$ for sub-espaço próprio de R^n , teremos ainda

$$\mathbf{M}_2 = I_n = \sum_{j=1}^m \mathbf{Q}_j,$$

com \mathbf{Q}_m matriz de projeção ortogonal sobre $R(\mathbf{M}_1)^\perp$, logo ortogonal às $\mathbf{Q}_1, \dots, \mathbf{Q}_{m-1}$. Assim ter-se-á

$$\mathbf{M}_i = \sum_{j=1}^m \ell_{i,j} \mathbf{Q}_j, i = 1, 2$$

com $\ell_{1,m} = 0$ e $\ell_{2,1} = \dots = \ell_{2,m} = 1$. Com

$$\mathbf{B} = [\ell_{i,j}],$$

teremos como vimos na subalínea **3.2.1**

$$\boldsymbol{\gamma} = \mathbf{B}^t \boldsymbol{\sigma}^2,$$

com $\boldsymbol{\gamma} = (\gamma_1 \dots \gamma_m)$ e $\boldsymbol{\sigma}^2 = (\sigma_1^2, \sigma_2^2)$, utilizando os resultados dessa sub-alínea obtemos BQUE para os $\gamma_j, j = 1, \dots, m$. Com $\tilde{\boldsymbol{\gamma}} = (\tilde{\gamma}_1, \dots, \tilde{\gamma}_m)$ o vetor desses estimadores, teremos

$$\tilde{\boldsymbol{\sigma}}^2 = (\mathbf{B}^t)^+ \tilde{\boldsymbol{\gamma}} = (\mathbf{B} \mathbf{B}^t)^+ \mathbf{B} \tilde{\boldsymbol{\gamma}}.$$

Observe-se que

$$\mathbf{T} = \mathbf{B}^t (\mathbf{B} \mathbf{B}^t)^+ \mathbf{B},$$

é a matriz de projeção ortogonal sobre $R(\mathbf{B}^t)$, logo

$$\mathbf{B}^t \tilde{\boldsymbol{\sigma}}^2 = \mathbf{T} \tilde{\boldsymbol{\gamma}},$$

pelo que $\tilde{\boldsymbol{\sigma}}^2$ será o LSE de $\boldsymbol{\sigma}^2$.

Como, com $\boldsymbol{\ell}_1 = (\ell_{1,1} \dots \ell_{1,m-1}, 0)$, temos

$$\mathbf{B} = \begin{bmatrix} \boldsymbol{\ell}_1^t & 0 \\ \mathbf{1}_{m-1}^t & 1 \end{bmatrix},$$

vindo

$$\mathbf{B} \mathbf{B}^t = \begin{bmatrix} \|\boldsymbol{\ell}_1\|^2 & \sum_{j=1}^{m-1} \ell_{1,j} \\ \sum_{j=1}^{m-1} \ell_{1,j} & m \end{bmatrix},$$

pelo que

$$|\mathbf{B} \mathbf{B}^t| = m \|\boldsymbol{\ell}_1\|^2 - \left(\sum_{j=1}^{m-1} \boldsymbol{\ell}_{1,j} \right)^2,$$

tendo-se $|\mathbf{B} \mathbf{B}^t| > 0$ a menos que

$$\boldsymbol{\ell}_{1,1} = \dots = \boldsymbol{\ell}_{1,m-1}$$

Quando esta condição não se verifica $\mathbf{B} \mathbf{B}^t$ é invertível com

$$(\mathbf{B} \mathbf{B}^t)^+ = (\mathbf{B} \mathbf{B}^t)^{-1} = \frac{1}{|\mathbf{B} \mathbf{B}^t|} \begin{bmatrix} m & - \sum_{j=1}^{m-1} \boldsymbol{\ell}_{1,j} \\ - \sum_{j=1}^{m-1} \boldsymbol{\ell}_{1,j} & \|\boldsymbol{\ell}_1\|^2 \end{bmatrix},$$

vindo

$$\begin{aligned} (\mathbf{B}^t)^+ &= (\mathbf{B} \mathbf{B}^t)^{-1} \mathbf{B} = \frac{1}{|\mathbf{B} \mathbf{B}^t|} \begin{bmatrix} m & \sum_{j=1}^{m-1} \boldsymbol{\ell}_{1,j} \\ \sum_{j=1}^{m-1} \boldsymbol{\ell}_{1,j} & \|\boldsymbol{\ell}_1\|^2 \end{bmatrix} \begin{bmatrix} \boldsymbol{\ell}_1^t & 0 \\ - \mathbf{1}_{m-1}^t & 1 \end{bmatrix} \\ &= \frac{1}{|\mathbf{B} \mathbf{B}^t|} \begin{bmatrix} m \boldsymbol{\ell}_1^t - \left(\sum_{j=1}^{m-1} \boldsymbol{\ell}_{1,j} \right) \mathbf{1}_{m-1}^t & \sum_{j=1}^{m-1} \boldsymbol{\ell}_{1,j} \\ \left(\sum_{j=1}^{m-1} \boldsymbol{\ell}_{1,j} \right) \boldsymbol{\ell}_1^t - \|\boldsymbol{\ell}_1\|^2 \mathbf{1}_{m-1}^t & \|\boldsymbol{\ell}_1\|^2 \end{bmatrix}, \end{aligned}$$

o que permite obter $\tilde{\boldsymbol{\sigma}}^2$ resolvendo se assim o problema de estimar as componentes de variâncias $\gamma_1, \dots, \gamma_{m-1}, \sigma_1^2$ e σ_2^2 .

Para vermos se o modelo tem COBS observemos que $\mathbf{M}_2 = \mathbf{I}_n$ comuta com $\mathbf{M} = \mathbf{X} \mathbf{X}^t$ e $\mathbf{M}_1 = \mathbf{X}_1 \mathbf{X}_1^t$, logo o modelo terá COBS se e só se \mathbf{M} comutar com $\mathbf{M}_1 = \mathbf{X}_1 \mathbf{X}_1^t$ e que se pode verificar caso a caso. Se isso se verificar,

$$\tilde{\boldsymbol{\beta}} = (\mathbf{X}^t \mathbf{X})^+ \mathbf{X}^t \mathbf{Y},$$

será BLUE.

3. 5. 2 – Testes F

Vamos agora utilizar o facto de se ter $e \sim N(\mathbf{0}, \sigma^2 I_n)$ para construir testes cujas estatísticas, quando se verificam as hipóteses testadas têm distribuições F centrais pode-se, portanto, fixar exactamente o nível desses testes, isto é, a probabilidade de rejeitar falsamente as hipóteses testadas. Mostraremos ainda que esses testes são não distorcidos, isto é, que a probabilidade de falsa rejeição da hipótese testada é sempre menor do que o nível do teste.

Ponhamos

$$\begin{cases} \Omega = R(\mathbf{X}) \\ \Omega_1 = R(\mathbf{X}_1) \\ \bar{\Omega} = R([\mathbf{X} \ \mathbf{X}_1]) \end{cases}$$

Temos agora testes F para os seguintes casos:

3.5.2.1 – Caso aninhado

Neste caso tem-se

$$\Omega \subset \Omega_1 = \Omega \boxplus \nabla = \bar{\Omega},$$

com

$$\nabla = \Omega^\perp \cap \Omega, \text{ ou } \nabla = \Omega^\perp \cap \bar{\Omega}$$

e

$$\begin{cases} d = \dim(\nabla) \\ g = \dim(\bar{\Omega}^\perp) \end{cases}$$

Estamos interessados em testar

$$H_{0,1}; \sigma_1^2 = 0,$$

observando que, quando $H_{0,1}$ se verifica, como

$$\begin{cases} E(\mathbf{Z}_1) = \mathbf{0}_{c_1} \\ \Sigma(\mathbf{Z}_1) = \mathbf{0}_{c_1 \times c_1}, \end{cases}$$

tem-se

$$1 \geq pr(\mathbf{Z}_1 = \mathbf{0}) = pr[\cap_{i=1}^{c_1} (Z_{1,i} = 0)] \geq$$

$$\sum_{i=1}^{c_1} pr(\mathbf{Z}_{1,i} = 0) - (c_1 - 1) = 1,$$

visto

$$pr(Z_{1,i} = 0) = 1, i = 1, \dots, c_1.$$

Logo, quando $H_{0,1}$ se verifica tem-se

$$Y = X\beta + e.$$

Sendo \mathbf{v}_∇ a projeção ortogonal de \mathbf{v} sobre ∇ , temos, quando $H_{0,1}$ se verifica

$$\begin{cases} \mathbf{Y}_\nabla = \mathbf{e}_\nabla \\ \mathbf{Y}_{\bar{\Omega}^\perp} = \mathbf{e}_{\bar{\Omega}^\perp} \end{cases}$$

com matriz de covariância cruzada

$$\Sigma(\mathbf{Y}_\nabla; \mathbf{Y}_{\bar{\Omega}^\perp}) = \Sigma(\mathbf{e}_\nabla, \mathbf{e}_{\bar{\Omega}^\perp}) = \mathbf{0}_{d \times g},$$

pelo que \mathbf{Y}_∇ e $\mathbf{Y}_{\bar{\Omega}^\perp}$, tendo distribuição conjunta normal e matriz de covariância cruzada nula, são independentes.

Com $\theta\mathcal{X}_\hbar^2$ representando o produto por θ dum qui-quadrado central com \hbar graus de liberdade tem-se, quando $H_{0,1}$ se verifica,

$$\begin{cases} \|\mathbf{Y}_\nabla\|^2 \sim \sigma_2^2 \mathcal{X}_d^2 \\ \|\mathbf{Y}_{\bar{\Omega}^\perp}\|^2 \sim \sigma_2^2 \mathcal{X}_g^2 \end{cases}$$

sendo estes qui-quadrados independentes. Logo, quando $H_{0,1}$ se verifica

$$F_1 = \frac{g}{d} \frac{\|\mathbf{Y}_\nabla\|^2}{\|\mathbf{Y}_{\bar{\Omega}^\perp}\|^2},$$

terá distribuição F central com d e g graus de liberdade, que representamos por $F_{d,g,\delta(\mathbf{u})}$. Assim para um teste de nível q o valor crítico será o quantil $f_{1-q,d,g}$, para a probabilidade $1 - q$, dessa distribuição.

Vejamos o que se passa quando a alternativa

$$H_{0,1}: \sigma_1^2 > 0$$

se verifica, tendo-se então

$$\mathbf{Y}_\nabla = \mathbf{e}_\nabla + (X_1\mathbf{Z}_1)_\nabla,$$

com $X_1\mathbf{Z}_1$ e, conseqüentemente, $(X_1\mathbf{Z}_1)_\nabla$ independentes de \mathbf{e}_∇ .

Seja

$$\mathbf{U} = (X_1\mathbf{Z}_1)_\nabla$$

quando $\mathbf{U} = \mathbf{u}$ temos

$$\|\mathbf{Y}_\nabla\|^2 \sim \sigma_2^2 \mathcal{X}_{d,\delta(\mathbf{u})}^2,$$

com

$$\delta(\mathbf{u}) = \frac{\|\mathbf{u}\|^2}{\sigma_2^2},$$

vindo

$$pr(F_1 \leq f_{1-q,d,g} | \mathbf{U} = \mathbf{u}) = 1 - F_{d,g,\delta(\mathbf{u})}(f_{1-q,d,g}).$$

Onde $F_{d,g,\delta(\mathbf{u})}$ é a distribuição F com d e g graus de liberdade e parâmetros de não centralidade $\delta(\mathbf{u})$.

Ora, [ver Mexia (1989, p.36)]

$$pr(F_1 \leq f_{1-q,d,g} | \mathbf{U} = \mathbf{u}) > q,$$

sempre que $u \neq 0$.

Temos agora a

Proposição 3.1

O teste F para $H_{0,1}$ é não distorcido.

Devido á última desigualdade para qualquer alternativa $H_{0,1}$, tem-se

$$pr(F_1 \leq f_{1-q,d,g}) > q,$$

o que estabelece a tese.

Admitamos agora que se tem

$$\Omega = \boxplus_{j=1}^m \Delta_j,$$

com

$$h_j = \dim(\Delta_j), j = 1, \dots, m,$$

e que se queremos testar as

$$H_{0,2,j}: (\sigma_1^2 = 0) \wedge (\boldsymbol{\mu} \in \Delta_j^\perp \cap \Omega), j = 1, \dots, m.$$

Com $\boldsymbol{\mu} = X \boldsymbol{\beta}$, o vetor médio de \mathbf{Y} , quando $H_{0,2,j}$ se verifica, $j = 1, \dots, m$, ter-se-á

$$pr(\mathbf{Z}_1 = \mathbf{0}) = 1,$$

visto $H_{0,1}$ se verificar. Ter-se-á, então

$$\mathbf{Y}_{\Delta_j} = \mathbf{e}_{\mathbf{v}_j}, j = 1, \dots, m,$$

vindo

$$\|\mathbf{Y}_{\Delta_j}\|^2 \sim \sigma^2 \chi_{h_j}^2, j = 1, \dots, m,$$

independente de

$$\|\mathbf{Y}_{\bar{\Omega}^\perp}\|^2 \sim \sigma^2 \chi_g, j = 1, \dots, m.$$

Assim, quando $H_{0,2,j}$ se verifica, $j = 1, \dots, m$, a estatística

$$F_{1,j} = \frac{g}{h_j} \frac{\|\mathbf{Y}_{\Delta_j}\|^2}{\|\mathbf{Y}_{\bar{\Omega}^\perp}\|^2}, j = 1, \dots, m,$$

terá distribuição F central com $h_j, j = 1, \dots, m$, e g graus de liberdade $F_{h_j, g}$. Logo, para um teste de nível q , para $H_{0,1,j}$ com estatística $F_{1,j}, j = 1, \dots, m$, ter-se-á o valor crítico $f_{1-q, h_j, g}, j = 1, \dots, m$.

3.5.2.2 – Cruzamento

Teremos agora, como atrás,

$$\begin{cases} \Omega = R(X) \\ \Omega_1 = R(X_1) \\ \bar{\Omega} = R([X \ X_1]) \end{cases}$$

tendo-se também

$$\begin{cases} \Delta = \Omega_1^\perp \cap \Omega \\ \Delta_1 = \Omega^\perp \cap \Omega_1, \end{cases}$$

com

$$\begin{cases} d = \dim(\Delta) \\ d_1 = \dim(\Delta_1), \end{cases}$$

e além disso

$$\begin{cases} \mathbf{Y}_\Delta = (X \boldsymbol{\beta})_\Delta + \mathbf{e}_\Delta \\ \mathbf{Y}_{\Delta_1} = (X_1 \mathbf{Z})_{\Delta_1} + \mathbf{e}_{\Delta_1}. \\ \mathbf{Y}_{\bar{\Omega}^\perp} = \mathbf{e}_{\bar{\Omega}^\perp} \end{cases}$$

Observe-se que \mathbf{e}_Δ e $\mathbf{e}_{\bar{\Omega}^\perp}$ [\mathbf{e}_{Δ_1} e $\mathbf{e}_{\bar{\Omega}^\perp}$] serão independentes pois têm distribuição conjunta normal e a matriz de covariância cruzada nula.

Admitimos que se têm as partições ortogonais

$$\begin{cases} \Delta = \boxplus_{j=1}^m \nabla_j \\ \Delta_1 = \boxplus_{j=1}^{m_1} \nabla_{1,j} \end{cases}$$

e que se pretende testar as hipóteses

$$\begin{cases} H_{0,j}: (X\boldsymbol{\beta})_{\nabla_j} = \mathbf{0}, j = 1, \dots, m \\ H_{0,1,j} = pr\left((X_1\mathbf{Z})_{\nabla_{1,j}} = \mathbf{0}\right) = 1, j = 1, \dots, m_1' \end{cases}$$

quando $H_{0,j}$ se verifica tem-se

$$\mathbf{Y}_{\nabla_j} = \mathbf{e}_{\nabla_j}, j = 1, \dots, m,$$

vindo

$$\|\mathbf{Y}_{\nabla_j}\|^2 \sim \sigma^2 \chi_{h_j}^2, j = 1, \dots, m,$$

com

$$h_j = \dim(\nabla_j), j = 1, \dots, m,$$

independente de

$$\|\mathbf{Y}_{\bar{\Omega}^\perp}\|^2 \sim \sigma^2 \chi_g^2, j = 1, \dots, m,$$

pelo que

$$F_j = \frac{g}{h_j} \frac{\|\mathbf{Y}_{\nabla_j}\|^2}{\|\mathbf{Y}_{\bar{\Omega}^\perp}\|^2}, j = 1, \dots, m,$$

terá distribuição F central com $h_j, j = 1, \dots, m$ e g graus de liberdade $F_{h_j, g}$ pelo que o valor crítico para um teste de nível q para $H_{0,j}, j = 1, \dots, m$, será $f_{1-q, h_j, g}, j = 1, \dots, m$.

Quando $H_{0,j}, j = 1, \dots, m$, não se verifica tem-se

$$\mathbf{Y}_{\nabla_j} = (X\boldsymbol{\beta})_{\nabla_j} + \mathbf{e}_{\nabla_j}, j = 1, \dots, m,$$

vindo

$$\|\mathbf{Y}_{\nabla_j}\|^2 \sim \sigma^2 \chi_{h_j, \delta_j}^2, j = 1, \dots, m,$$

independente de $\|\mathbf{Y}_{\bar{\Omega}^\perp}\|^2 \sim \sigma^2 \chi_g^2$, pelo que F_j (terá distribuição F com h_j e g graus de liberdade e parâmetros de não centralidade

$$\delta_j = \frac{1}{\sigma^2} \|(X\boldsymbol{\beta})_{\nabla_j}\|^2, j = 1, \dots, m,$$

pelo que, [ver Mexia (1989, p.36)], estes testes F são não distorcidos.

Passando às hipóteses $H_{0,1,j}, j = 1, \dots, m$, temos para as mesmas as estatísticas

$$F_{1,j} = \frac{g}{h_{1,j}} \frac{\|\mathbf{Y}_{\nabla_{1,j}}\|^2}{\|\mathbf{Y}_{\bar{\Omega}^\perp}\|^2}, j = 1, \dots, m,$$

que, quando se verificam as hipóteses testadas, têm distribuição F central com $h_{1,j} = \dim(\Delta_{1,j}), j = 1, \dots, m$ e g graus de liberdade $F_{h_{1,j},g}$. Assim, para testar com nível q os valores críticos são $f_{1-q, h_{1,j}, g}, j = 1, \dots, m$.

Um último aspeto que podemos analisar é a decomposição das hipóteses

$H_{0,j} [H_{0,1,j}] j = 1, \dots, m [j = 1, \dots, m_1]$. Sejam $\{\boldsymbol{\alpha}_{j,1}, \dots, \boldsymbol{\alpha}_{j,h_j}\}, j = 1, \dots, m$
 $[\{\boldsymbol{\alpha}_{1,j}, \dots, \boldsymbol{\alpha}_{1,j,h_j}\}, j = 1, \dots, m_1]$, bases ortonormadas para os $\nabla_j, j = 1, \dots, m$
 $[\nabla_{1,j}, j = 1, \dots, m_1]$, então podemos testar as hipóteses

$$\begin{cases} H_{0,j,\ell}: \boldsymbol{\alpha}_{j,\ell}^t(X \boldsymbol{\beta}) = 0, \ell = 1, \dots, h_j, j = 1, \dots, m \\ H_{0,1,j,\ell}: pr[\boldsymbol{\alpha}_{1,j,\ell}^t(X_1 \mathbf{Z}) = 0] = 1, \ell = 1, \dots, h_{1,j}, j = 1, \dots, m_1, \end{cases}$$

tendo-se as estatísticas

$$\begin{cases} F_{j,\ell} = \frac{g (\boldsymbol{\alpha}_{j,\ell}^t \mathbf{Y})^2}{\|\mathbf{Y}_{\bar{\Omega}^\perp}\|^2}, \ell = 1, \dots, h_j, j = 1, \dots, m \\ F_{1,j,\ell} = \frac{g (\boldsymbol{\alpha}_{1,j,\ell}^t \mathbf{Y})^2}{\|\mathbf{Y}_{\bar{\Omega}^\perp}\|^2}, \ell = 1, \dots, h_{1,j}, j = 1, \dots, m_1, \end{cases}$$

que, quando se verificam as hipóteses testadas têm distribuições F centrais com 1 e g graus de liberdade $F_{1,g}$. Assim para estes testes com nível q o valor crítico será $f_{1-q,1,g}$.

Ao considerar-se estas sub-hipóteses consegue-se refinar a análise já que

$$H_{0,j}, j = 1, \dots, m [H_{0,1,j}, j = 1, \dots, m_1]$$

se verifica se e só se as respetivas sub-hipóteses se verificam.

Capítulo 4

Neste capítulo descrevem-se as famílias de modelos bi-aditivos, começando pelas famílias OBS e COBS cujos modelos têm Orthogonal Block Structure ou Commutative Orthogonal Block Structure (ver secções 3.2.1 e 3.2.2). Em seguida famílias regressionais e por fim famílias estruturadas cujos modelos correspondem aos tratamentos dum delineamento base com cruzamento de fatores e cruzamento e aninhamento de fatores. Seguem-se fatoriais de base prima.

4 - Famílias de Modelos Bi-aditivos

4.1 - Ajustamento

Os d modelos duma destas famílias terão as mesmas matrizes $\mathbf{X}, \mathbf{X}_1, \dots, \mathbf{X}_w$ podendo ser escritos como

$$\mathbf{Y}(\mathcal{h}) = \mathbf{X} \boldsymbol{\beta}(\mathcal{h}) + \sum_{i=1}^w X_i \mathbf{Z}_i(\mathcal{h}), \mathcal{h} = 1, \dots, d,$$

com os vetores $\mathbf{Z}_i(\mathcal{h}), i = 1, \dots, w, \mathcal{h} = 1, \dots, d$, independentes com vetores médios nulos. Tem-se ainda

$$\begin{cases} \boldsymbol{\beta}(\mathcal{h}) = (\beta_1(\mathcal{h}), \dots, \beta_k(\mathcal{h})), \mathcal{h} = 1, \dots, d \\ \mathbf{Z}_1(\mathcal{h}) = (Z_{i,1}(\mathcal{h}) \dots Z_{i,c_i}(\mathcal{h})), i = 1, \dots, w, \mathcal{h} = 1, \dots, d, \end{cases}$$

sendo *i.i.d* as componentes de $Z_{i,1}(\mathcal{h}) \dots Z_{i,c_i}(\mathcal{h})$ com cumulantes $\mathcal{X}_{r,i}$ de ordem r , $i = 1, \dots, w, \mathcal{h} = 1, \dots, d, r = 2, 3, 4, \dots$ os mesmos para todos os modelos.

Observe-se que as componentes dos $\mathbf{Z}_i(\mathcal{h}), i = 1, \dots, w, \mathcal{h} = 1, \dots, d$, tendo valores nulos têm $\mathcal{X}_{1,i} = 0, i = 1, \dots, w$.

Por outro lado, como as matrizes $\mathbf{M} = \mathbf{X} \mathbf{X}^t, \mathbf{M}_i = \mathbf{X}_i \mathbf{X}_i^t, i = 1, \dots, w$, são as mesmas para todos os modelos logo tem-se a

Proposição 4.1:

Todos ou nenhum dos modelos duma família têm OBS [COBS]. A proposição 4.1 resulta da observação feita antes de que sendo as matrizes $\mathbf{X}, \mathbf{X}_1, \dots, \mathbf{X}_w$ as mesmas para todos os modelos de família todos ou nenhum têm OBS e ou COBS.

As famílias OBS [COBS] serão constituídas por modelos com OBS [COBS]. Por outro lado, da identidade da matriz \mathbf{X} para todos os modelos resulta que os vetores $\alpha_1, \dots, \alpha_n$, da base ortonormada da $R(\mathbf{X})^\perp$ são os mesmos, então as

$$\dot{Y}_i(\mathcal{h}) = \alpha_i^t \mathbf{Y}(\mathcal{h}), i = 1, \dots, n, \mathcal{h} = 1, \dots, d,$$

terão os cumulantes

$$\Theta_{r,i} = \Theta_r(\dot{Y}_i(\mathcal{h})) = \sum_{j=1}^w \mathcal{b}_{i,j}(r) \mathcal{X}_{r,j}, i = 1, \dots, n, \mathcal{h} = 1, \dots, d,$$

mantendo-se as matrizes $\mathbf{B}(r) = [\mathcal{b}_{i,j}(r)], r = 2, 3$ e continuando a ter-se

$$\Theta_r = \mathbf{B}(r) \mathbf{X}_r, r = 2, 3$$

e sendo os vetores Θ_r e \mathbf{X}_r os mesmos para todos os modelos de família. Temos ainda para os vários modelos os estimadores.

$$\tilde{\Theta}_r(\mathcal{h}) = (\dot{Y}_1^r(\mathcal{h}), \dots, \dot{Y}_n^r(\mathcal{h})), r = 2, 3, \mathcal{h} = 1, \dots, d,$$

o que origina os *LSE*

$$\tilde{\mathbf{X}}_r(\mathcal{h}) = (\mathbf{B}(r)^t \mathbf{B}(r))^+ \mathbf{B}(r)^t \tilde{\Theta}_r(\mathcal{h}), r = 2, 3,$$

a partir dos quais se obtêm os

$$\tilde{\mathbf{X}}_r = \frac{1}{d} \sum_{\mathcal{h}=1}^d \tilde{\mathbf{X}}_r(\mathcal{h}), r = 2, 3,$$

em particular tem-se

$$\tilde{\mathbf{X}}_2 = (\tilde{\sigma}_1^2 \dots \tilde{\sigma}_w^2),$$

podendo estimar-se as matrizes de covariância dos vetores de observações dos modelos, utilizando-se

$$\tilde{\Sigma}(Y(h)) = \sum_{j=1}^w \tilde{\sigma}_j^2 \mathbf{M}_j = \tilde{\Sigma}(Y(\mathbf{h})), \mathbf{h} = 1, \dots, d,$$

isto é, os modelos da família têm a mesma matriz de covariância [estimada]. Resulta daqui ter-se os *LSE*.

$$\tilde{\boldsymbol{\beta}}(h) = (\mathbf{X}^t \mathbf{X})^+ \mathbf{X}^t \mathbf{Y}(h), \mathbf{h} = 1, \dots, d \quad (4.1)$$

que são *BLUE* nas famílias com *COBS*. No caso geral temos os *GLSE*

$$\tilde{\boldsymbol{\beta}}(h) = (\mathbf{X}^t \tilde{\Sigma}(\mathbf{Y})^+ \mathbf{X}) \mathbf{X}^t \tilde{\Sigma}(\mathbf{Y})^+ \mathbf{X} \mathbf{Y}(h), \mathbf{h} = 1, \dots, d,$$

com $\tilde{\Sigma}(\mathbf{Y}) = \frac{1}{d} \sum_{i=1}^d \tilde{\Sigma}(\mathbf{Y}(h))$; onde $\mathbf{Y} = \frac{1}{d} \sum_{\mathbf{h}=1}^d \mathbf{Y}(\mathbf{h})$.

4.2 - Famílias *COBS*

Os modelos destas famílias têm *COBS* com estimadores dos vetores dos coeficientes

$$\tilde{\boldsymbol{\beta}}(h) = (\mathbf{X}^t \mathbf{X})^+ \mathbf{X}^t \mathbf{Y}(h), \mathbf{h} = 1, \dots, d,$$

com matrizes de covariância estimadas idênticas

$$\tilde{\Sigma}(\boldsymbol{\beta}(h)) = (\mathbf{X}^t \mathbf{X})^+ \mathbf{X}^t \tilde{\Sigma}(\mathbf{Y}) \mathbf{X} (\mathbf{X}^t \mathbf{X}), \mathbf{h} = 1, \dots, d,$$

e tendo os mesmos pares de valores e vetores próprios $(\xi_j, \mathbf{v}_j), j = 1, \dots, k$.

Teremos então as funções estimáveis principais estimadas.

$$\tilde{\eta}_{j,\mathbf{h}} = \mathbf{v}_j^t \tilde{\boldsymbol{\beta}}(h), j = 1, \dots, k, \quad \mathbf{h} = 1, \dots, d,$$

com variâncias estimadas $\xi_j, j = 1, \dots, k$.

Observe-se que, com $\boldsymbol{v} \in R^k$, se tem

$$\boldsymbol{v} = \sum_{j=1}^k (\boldsymbol{v}^t \mathbf{v}_j) \mathbf{v}_j,$$

vindo

$$\boldsymbol{v}^t \tilde{\boldsymbol{\beta}}(h) = \sum_{j=1}^k (\boldsymbol{v}^t \mathbf{v}_j) \tilde{\eta}_{j,\mathbf{h}}, \mathbf{h} = 1, \dots, d.$$

Como $\tilde{\eta}_{j,1}, \dots, \tilde{\eta}_{j,d}$ têm a mesma variância ao comparar $\eta_{j,1}, \dots, \eta_{j,d}, j = 1, \dots, k$ utilizamos testes χ^2 para as

$$H_{0,j}: \eta_{j,1} = \dots = \eta_{j,d}, j = 1, \dots, k,$$

dado estarmos no caso equilibrado em que, [ver Scheffé (1959) cp. X], a ANOVA e técnicas relacionadas são robustas relativamente à não normalidade.

Esses testes terão estatísticas

$$T_j = \frac{1}{\xi_j} \sum_{h=1}^d (\tilde{\eta}_{j,h} - \tilde{\eta}_j)^2, j = 1, \dots, k,$$

com

$$\tilde{\eta}_j = \frac{1}{d} \sum_{h=1}^d \tilde{\eta}_{j,h},$$

e quando $H_{0,j}, j = 1, \dots, k$, se verifica aproximadamente distribuição de qui-quadrado com $d - 1$ graus de liberdade.

Da mesma maneira se podem testar as

$$H_0(\boldsymbol{v}): \boldsymbol{v}^t \boldsymbol{\beta}_1 = \dots = \boldsymbol{v}^t \boldsymbol{\beta}_d$$

No caso geral pode-se aplicar este tratamento desde que se continuem a utilizar os *LSE* dados por (4.1) já que os mesmos continuam a ter a mesma matriz de covariância.

A concluir esta alínea vamos considerar os testes χ^2 (Qui-Quadrado) que propusemos para testar as $H_{0,j}, j = 1, \dots, k$ e $H_0(\boldsymbol{v})$.

Admitimos para isso ter as variáveis aleatórias independentes $L_1 \dots L_k$ normais com valores médios μ_1, \dots, μ_k e variância σ^2 .

Tomando-se

$$L. = \frac{1}{k} \sum_{j=1}^k L_j,$$

a estatística

$$U = \sum_{j=1}^k (L_j - L.)^2$$

será o produto $(\sigma^2 \chi_{k-1, \delta}^2)$ por σ^2 dum qui-quadrado com $k - 1$ graus de liberdade e parâmetro de não centralidade

$$\delta = \frac{1}{\sigma^2} \sum_{\ell=1}^k (\mu_\ell - \mu.)^2.$$

Se σ^2 for conhecida pode-se utilizar U/σ^2 como estatística de teste e o $(1 - p)$ - ésimo quantil $\chi_{k-1, 1-p}^2$ dum qui-quadrado central com $k - 1$ graus de liberdade, como valor crítico para um teste de nível p .

No caso das famílias *COBS* tomamos

$$L_i = \tilde{\eta}_{j,i}, [L_i = \boldsymbol{v}^t \tilde{\boldsymbol{\beta}}_i; i = 1, \dots, d] \quad (4.2)$$

e

$$\sigma^2 = \xi_j, j = 1, \dots, k$$

para as $H_{0,j}, j = 1, \dots, k$. Para $H_0(\boldsymbol{v})$ toma-se

$$\sigma^2 = \boldsymbol{v}^t \boldsymbol{\xi},$$

com

$$\boldsymbol{\xi} = (\xi_1 \dots \xi_k).$$

4.3 – Famílias Regressionais

Para cada modelo duma destas famílias tem-se um vetor cujas componentes são os valores de \boldsymbol{u} variáveis controladas exógenas, isto é, variáveis controladas distintas das inicialmente consideradas. As iniciais serão endógenas visto entrarem na determinação dos modelos enquanto as últimas são exógenas.

Observemos que os

$$\tilde{\boldsymbol{\beta}}_{\hbar} = (\boldsymbol{X}^t \boldsymbol{X})^+ \boldsymbol{X}^t \boldsymbol{Y}_{\hbar}, \hbar = 1, \dots, d,$$

têm todos a mesma matriz de covariância e, como as matrizes $\boldsymbol{X}, \boldsymbol{X}_1, \dots, \boldsymbol{X}_w$ são idênticas para todos os modelos, podemos voltar a apelar robustez da ANOVA e técnicas afins para admitir a normalidade destes estimadores com vetores médios $\boldsymbol{\beta}_{\hbar}, \hbar = 1, \dots, d$, a matriz de covariância $\Sigma(\tilde{\boldsymbol{\beta}}_{\hbar})$.

Tendo-se $\tilde{\boldsymbol{\beta}}_{\hbar} = (\tilde{\beta}_{1,\hbar}, \dots, \tilde{\beta}_{k,\hbar})$ e $\boldsymbol{\beta}_{\hbar} = (\beta_{1,\hbar}, \dots, \beta_{k,\hbar}), \hbar = 1, \dots, d$, ponhamos

$$\begin{cases} \dot{\boldsymbol{\beta}}_{\ell} = (\beta_{\ell,1}, \dots, \beta_{\ell,d}), \ell = 1, \dots, k \\ \tilde{\boldsymbol{\beta}}_{\ell} = (\tilde{\beta}_{\ell,1}, \dots, \tilde{\beta}_{\ell,d}), \ell = 1, \dots, k \end{cases}$$

Como os $\tilde{\boldsymbol{\beta}}_{\hbar}, \hbar = 1, \dots, d$, são independentes com a mesma matriz de covariância, representando por $\dot{\sigma}^2(\ell)$ a variância das ℓ -ésima componentes dos

$$\tilde{\boldsymbol{\beta}}_{\hbar}, \hbar = 1, \dots, d,$$

tem-se

$$\Sigma(\tilde{\boldsymbol{\beta}}_{\ell}) = \dot{\sigma}^2(\ell) I_d; \ell = 1, \dots, k.$$

Seja $\dot{\boldsymbol{X}}$ a matriz cujos vetores linha são os vetores de valores das variáveis controladas.

Ponhamos ainda

$$\begin{cases} \mathbf{z}_\ell = \tilde{\boldsymbol{\beta}}_\ell, \ell = 1, \dots, k \\ \boldsymbol{\xi}_\ell = \tilde{\boldsymbol{\beta}}_\ell, \ell = 1, \dots, k \end{cases}$$

Se ajustarmos modelos lineares da forma

$$\boldsymbol{\xi}_\ell = \dot{\mathbf{X}} \boldsymbol{\Theta}_\ell, \ell = 1, \dots, k$$

temos os estimadores *LSE*

$$\tilde{\boldsymbol{\Theta}}_\ell = (\dot{\mathbf{X}}^t \dot{\mathbf{X}})^+ \dot{\mathbf{X}}^t \mathbf{z}_\ell, \ell = 1, \dots, k,$$

com

$$\Sigma(\tilde{\boldsymbol{\Theta}}_\ell) = \hat{\sigma}^2(\ell) (\dot{\mathbf{X}}^t \dot{\mathbf{X}})^+, \ell = 1, \dots, k.$$

Tem-se ainda (aproximadamente)

$$S_\ell = \|\mathbf{z}_\ell - \dot{\mathbf{X}} \tilde{\boldsymbol{\Theta}}_\ell\|^2 \sim \hat{\sigma}^2(\ell) \chi_g^2, \ell = 1, \dots, k,$$

com

$$g = d - \text{car}(\dot{\mathbf{X}}),$$

podendo admitir-se que

$$\tilde{\boldsymbol{\Theta}}_\ell \sim \mathcal{N}(\boldsymbol{\Theta}_\ell, \hat{\sigma}^2(\ell) (\dot{\mathbf{X}}^t \dot{\mathbf{X}})^+), \ell = 1, \dots, k.$$

Para as hipóteses

$$H_0(\boldsymbol{\Theta}_{\ell,0}): \boldsymbol{\Theta}_\ell = \boldsymbol{\Theta}_{\ell,0}, \ell = 1, \dots, k$$

temos os testes com estatísticas

$$F(\boldsymbol{\Theta}_{\ell,0}) = \frac{g}{g'} \frac{(\tilde{\boldsymbol{\Theta}}_\ell - \boldsymbol{\Theta}_{\ell,0})^t (\dot{\mathbf{X}}^t \dot{\mathbf{X}})^+ (\tilde{\boldsymbol{\Theta}}_\ell - \boldsymbol{\Theta}_{\ell,0})}{S_\ell}, \ell = 1, \dots, k,$$

com (aproximadamente) distribuição *F* com

$$g' = \text{car}(\dot{\mathbf{X}}),$$

e *g* graus de liberdade e parâmetros de não centralidade

$$\delta_\ell = \frac{1}{\hat{\sigma}^2(\ell)} (\boldsymbol{\Theta}_\ell - \boldsymbol{\Theta}_{\ell,0})^t (\dot{\mathbf{X}}^t \dot{\mathbf{X}})^+ (\boldsymbol{\Theta}_\ell - \boldsymbol{\Theta}_{\ell,0}), \ell = 1, \dots, k,$$

que se anulam quando as $H_{0,\ell}, \ell = 1, \dots, k$ se verificam.

Como a função da distribuição de estatística é, [ver Mexia (1989)], decrescente em δ o teste é não distorcido, já que quando $H_{0,\ell}$ se verifica, se tem

$$\delta_\ell = 0, \quad \ell = 1, \dots, k.$$

Os modelos das famílias regressionais podem ter ou não *COBS*, no entanto utilizaremos para as não *COBS* os *LSE* de forma a garantir que os $\tilde{\beta}_{\mathcal{h}}, \mathcal{h} = 1, \dots, d$, são independentes e conseqüentemente que (4.2) se verifica.

4.4 - Famílias estruturadas

Nas famílias estruturadas a inferência centra-se nos vetores de coeficientes. Estes têm a mesma matriz de covariância para os seus estimadores em todos os modelos. Assim os vetores próprios $\alpha_1 \dots \alpha_k$ dessas matrizes serão os mesmos para todos os modelos, tendo-se, pois, as mesmas funções estimáveis principais

$$\psi_{\ell, \mathcal{h}} = \alpha_{\ell} \beta_{\mathcal{h}}, \ell = 1, \dots, k, \mathcal{h} = 1, \dots, d.$$

Os modelos destas famílias correspondem aos tratamentos dum delineamento base com efeitos fixos.

Admitimos que o delineamento base tem estrutura ortogonal associada a uma partição ortogonal.

$$R^d = \boxplus_{\ell=1}^u \nabla_{\ell} \quad (4.3)$$

onde \boxplus indica soma direta ortogonal de sub-espacos.

As matrizes de projeção ortogonal Q_1, \dots, Q_u sobre $\nabla_1, \dots, \nabla_u$ são mutuamente ortogonais constituindo a base principal duma AJC associada ao delineamento. Como

$$\sum_{\ell=1}^u Q_{\ell} = I_d,$$

essa álgebra será completa contendo matrizes invertíveis.

Se os g_{ℓ} vetores linha de A_{ℓ} constituírem uma base ortonormada para $\nabla_{\ell}, \ell = 1, \dots, u$, teremos

$$Q_{\ell} = A_{\ell}^t A_{\ell} = \sum_{\mathcal{h}=1}^{g_{\ell}} \alpha_{\ell, \mathcal{h}} \alpha_{\ell, \mathcal{h}}^t,$$

com

$$A_{\ell} = [\alpha_{\ell, 1} \dots \alpha_{\ell, g_{\ell}}]^t$$

Sejam $\tilde{\eta}_{j, \mathcal{h}}, j = 1, \dots, k, \mathcal{h} = 1, \dots, d$, os valores das funções estimáveis principais para os diferentes modelos. Como as $\tilde{\eta}_{j, 1} \dots \tilde{\eta}_{j, d}$ são independentes com variância ξ_j ,

$j = 1, \dots, k$ e

$$\tilde{\psi}_j = (\tilde{\eta}_{j, 1} \dots \tilde{\eta}_{j, d}), j = 1, \dots, k,$$

tem vector médio

$$\psi_j = (\eta_{j,1} \dots \eta_{j,d}), j = 1, \dots, k,$$

utilizaremos a robustez da ANOVA no caso equilibrado para admitir que

$$\tilde{\psi}_j \sim \mathcal{N}(\psi_j, \xi_j I_d), j = 1, \dots, k$$

Admitamos ainda que existe

$$\mathcal{D} \subset \bar{d} = \{1, \dots, d\},$$

tal que podemos admitir a priori que

$$\psi_j \in \bigoplus_{\ell \in \mathcal{D}} \nabla_\ell, j = 1, \dots, k,$$

isto é, que as projeções ortogonais dos vetores médios $\psi_1 \dots \psi_k$ sobre os $\nabla_j, j \in \mathcal{D}$, são nulas. Este tipo de pressuposto é usual por exemplo, nos planos com vários fatores, admitir-se, a priori, a nulidade das interações de mais alta ordem. Recorde-se que a ordem duma interação é o número dos fatores que nela participam menos um. Por exemplo havendo 3 fatores que cruzam os sub-espacos da partição ortogonal correspondem aos subconjuntos de $\bar{3} = \{1, 2, 3\}$ havendo $2^3 = 8$ sub-espacos. O valor médio geral estará associado a ϕ , os efeitos dos três fatores aos conjuntos $\{j\}, j = 1, 2, 3$ as interações entre pares de fatores aos $\{j, j'\}, j \neq j'$, havendo 3 conjuntos associados a efeitos ou as interações associadas a interação entre pares de fatores. Finalmente, $\{1, 2, 3\}$ será o conjunto associado á interação entre os três fatores.

Com

$$g_\ell = \dim(\nabla_\ell), \ell = 1, \dots, u,$$

teremos os graus de liberdade para o erro

$$g = \sum_{\ell \in \mathcal{D}} g_\ell.$$

No caso dos 3 fatores sendo J_1, J_2, J_3 os números dos níveis dos mesmos, temos

$$g_\ell = \prod_{h \in \varphi_\ell} (J_h - 1), \ell = 1, \dots, 8,$$

tendo os sub-espacos de $\bar{3}$ os índices

$$\ell(\varphi) = 1 + \sum_{h \in \varphi} 2^{h-1}.$$

Em particular teremos

$$\ell[\{1, 2, 3\}] = 8$$

e

$$g_8 = (J_1 - 1)(J_2 - 1)(J_3 - 1).$$

A seguir vamos considerar vários casos para o delineamento base.

4.4.1 – Cruzamento dos fatores

Teremos f fatores com J_1, \dots, J_f níveis. Haverá

$$u = 2^f,$$

subconjuntos de $\bar{f} = \{1, \dots, f\}$ dos quais $\binom{f}{h}$ conterão h índices, $h = 0, \dots, f$. Os subconjuntos de \bar{f} recebem os índices $\ell(\varphi)$ dado por (4.3) e [ver Ferreira et al., (2013)], tendo-se ainda

$$\mathbf{A}_\ell = \bigotimes_{h=1}^u \mathbf{A}_{\ell, h},$$

com $\ell(\varphi) = \ell$ temos

$$\begin{cases} \mathbf{A}_{\ell, h} = \frac{1}{\sqrt{J_h}} \mathbf{1}_{J_h}^t & , h \notin \varphi \\ \mathbf{A}_{\ell, h} = \mathbf{T}_{J_h} & , h \in \varphi, \end{cases}$$

sendo \mathbf{T}_r a matriz que se obtém retirando a primeira linha igual a $\frac{1}{\sqrt{r}} \mathbf{1}_r^t$ duma matriz ortogonal $r \times r$.

Para obter as matrizes \mathbf{T}_r encontramos várias possibilidades. Uma destas possibilidades é.

$$\mathbf{T}_r = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & \dots & 0 & \dots & 0 \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & & & \\ \frac{1}{\sqrt{r(r+1)}} & -\frac{1}{\sqrt{r(r+1)}} & -\frac{1}{\sqrt{r(r+1)}} & \dots & \frac{r}{\sqrt{r(r+1)}} \end{bmatrix}$$

tem-se

$$g_\ell = \text{car}(\mathbf{A}_\ell) = \prod_{h=1}^f \text{car}(\mathbf{A}_{\ell, h}) = \prod_{h \in \varphi_\ell} (J_h - 1),$$

já que

$$\begin{cases} \text{car}\left(\frac{1}{\sqrt{J_{\mathfrak{h}}}} 1_{J_{\mathfrak{h}}}^t\right) = 1 \\ \text{car}(T_{J_{\mathfrak{h}}}) = J_{\mathfrak{h}} - 1. \end{cases}$$

Se se admitir a priori que

$$\mathcal{D} = \{\ell(\varphi); \#(\varphi) \geq \mathfrak{h}\},$$

isto é, se se utilizamos as interações entre $\mathfrak{h}'(\geq \mathfrak{h})$ fatores para estimar o erro tem-se

$$g = \sum_{\ell \in \mathcal{D}} g_{\ell}.$$

Em particular se for \bar{f} o único conjunto “utilizado” para estimar o erro tem-se

$$g = \prod_{\mathfrak{h}=1}^f (J_{\mathfrak{h}} - 1).$$

Sejam $\tilde{\psi}_j, j = 1, \dots, k$, os vetores cujas componentes são os estimadores das funções estimáveis principais, para as hipóteses

$$H_{0,j,\mathfrak{h}}: \boldsymbol{\psi}_j \in \nabla_{\mathfrak{h}}^{\perp}, j = 1, \dots, \mathfrak{h}, \mathfrak{h} \notin \mathcal{D},$$

teremos a estatística F

$$F_{j,\mathfrak{h}} = \frac{g}{g_{\mathfrak{h}}} \frac{S_{j,\mathfrak{h}}}{S_j}, j = 1, \dots, \mathfrak{h}, \mathfrak{h} \notin \mathcal{D},$$

com

$$S_{j,\mathfrak{h}} = \|A_{\mathfrak{h}} \tilde{\psi}_j\|^2, j = 1, \dots, k, \mathfrak{h} = 1, \dots, f,$$

e

$$S_j = \sum_{\mathfrak{h} \in \mathcal{D}} S_{j,\mathfrak{h}}, j = 1, \dots, k.$$

Damos agora alguns exemplos de cruzamento de fatores.

Começamos por considerar o caso de $u = 4$ fatores. Na tabela 1 indicamos os conjuntos φ e os respetivos graus de liberdade.

Tabela 1: conjuntos de fatores, caso de quatro fatores

| Conjunto | índice | $g(D) = \prod_{h \in D} (J_h - 1); \quad (J = 1 + \sum_{h \in D} 2^{h-1})$ |
|--------------|--------|--|
| \emptyset | 1 | 1 |
| {1} | 2 | $(J_1 - 1)$ |
| {2} | 3 | $(J_2 - 1)$ |
| {1, 2} | 4 | $(J_1 - 1)(J_2 - 1)$ |
| {3} | 5 | $(J_3 - 1)$ |
| {1, 3} | 6 | $(J_1 - 1)(J_3 - 1)$ |
| {2, 3} | 7 | $(J_2 - 1)(J_3 - 1)$ |
| {1, 2, 3} | 8 | $(J_1 - 1)(J_2 - 1)(J_3 - 1)$ |
| {4} | 9 | $(J_4 - 1)$ |
| {1, 4} | 10 | $(J_1 - 1)(J_4 - 1)$ |
| {2, 4} | 11 | $(J_2 - 1)(J_4 - 1)$ |
| {1, 2, 4} | 12 | $(J_1 - 1)(J_2 - 1)(J_4 - 1)$ |
| {3, 4} | 13 | $(J_3 - 1)(J_4 - 1)$ |
| {1, 3, 4} | 14 | $(J_1 - 1)(J_3 - 1)(J_4 - 1)$ |
| {2, 3, 4} | 15 | $(J_2 - 1)(J_3 - 1)(J_4 - 1)$ |
| {1, 2, 3, 4} | 16 | $(J_1 - 1)(J_2 - 1)(J_3 - 1)(J_4 - 1)$ |

tem-se as matrizes:

$$A(\emptyset) = \frac{1}{\sqrt{J_1}} 1_{J_1}^t \otimes \frac{1}{\sqrt{J_2}} 1_{J_2}^t \otimes \frac{1}{\sqrt{J_3}} 1_{J_3}^t \otimes \frac{1}{\sqrt{J_4}} 1_{J_4}^t$$

$$A(\{1\}) = T_{J_1} \otimes \frac{1}{\sqrt{J_2}} 1_{J_2}^t \otimes \frac{1}{\sqrt{J_3}} 1_{J_3}^t \otimes \frac{1}{\sqrt{J_4}} 1_{J_4}^t$$

$$A(\{2\}) = \frac{1}{\sqrt{J_1}} 1_{J_1}^t \otimes T_{J_2} \otimes \frac{1}{\sqrt{J_3}} 1_{J_3}^t \otimes \frac{1}{\sqrt{J_4}} 1_{J_4}^t$$

$$A(\{1, 2\}) = T_{J_1} \otimes T_{J_2} \otimes \frac{1}{\sqrt{J_3}} 1_{J_3}^t \otimes \frac{1}{\sqrt{J_4}} 1_{J_4}^t$$

$$A(\{3\}) = \frac{1}{\sqrt{J_1}} 1_{J_1}^t \otimes \frac{1}{\sqrt{J_2}} 1_{J_2}^t \otimes T_{J_3} \otimes \frac{1}{\sqrt{J_4}} 1_{J_4}^t$$

$$A(\{1, 3\}) = T_{J_1} \otimes \frac{1}{\sqrt{J_2}} 1_{J_2}^t \otimes T_{J_3} \otimes \frac{1}{\sqrt{J_4}} 1_{J_4}^t$$

$$A(\{2, 3\}) = \frac{1}{\sqrt{J_1}} 1_{J_1}^t \otimes T_{J_2} \otimes T_{J_3} \otimes \frac{1}{\sqrt{J_4}} 1_{J_4}^t$$

$$A(\{1, 2, 3\}) = T_{J_1} \otimes T_{J_2} \otimes T_{J_3} \otimes \frac{1}{\sqrt{J_4}} 1_{J_4}^t$$

$$A(\{4\}) = \frac{1}{\sqrt{J_1}} 1_{J_1}^t \otimes \frac{1}{\sqrt{J_2}} 1_{J_2}^t \otimes \frac{1}{\sqrt{J_3}} 1_{J_3}^t \otimes T_{J_4}$$

$$A(\{1, 4\}) = T_{J_1} \otimes \frac{1}{\sqrt{J_2}} 1_{J_2}^t \otimes \frac{1}{\sqrt{J_3}} 1_{J_3}^t \otimes T_{J_4}$$

$$A(\{2, 4\}) = \frac{1}{\sqrt{J_1}} 1_{J_1}^t \otimes T_{J_2} \otimes \frac{1}{\sqrt{J_3}} 1_{J_3}^t \otimes T_{J_4}$$

$$A(\{1, 2, 4\}) = T_{J_1} \otimes T_{J_2} \otimes \frac{1}{\sqrt{J_3}} 1_{J_3}^t \otimes T_{J_4}$$

$$A(\{3, 4\}) = \frac{1}{\sqrt{J_1}} 1_{J_1}^t \otimes \frac{1}{\sqrt{J_2}} 1_{J_2}^t \otimes T_{J_3} \otimes T_{J_4}$$

$$A(\{1, 3, 4\}) = T_{J_1} \otimes \frac{1}{\sqrt{J_2}} 1_{J_2}^t \otimes T_{J_3} \otimes T_{J_4}$$

$$A(\{2, 3, 4\}) = \frac{1}{\sqrt{J_1}} 1_{J_1}^t \otimes T_{J_2} \otimes T_{J_3} \otimes T_{J_4}$$

$$A(\{1, 2, 3, 4\}) = T_{J_1} \otimes T_{J_2} \otimes T_{J_3} \otimes T_{J_4}.$$

No caso de 5 fatores tem-se:

Tabela 2: conjuntos de fatores, caso de cinco fatores

| Conjunto | índice | $g(D) = \prod_{h \in D} (J_h - 1); \quad (J = 1 + \sum_{h \in D} 2^{h-1})$ |
|--------------|--------|--|
| \emptyset | 1 | 1 |
| {1} | 2 | $(J_1 - 1)$ |
| {2} | 3 | $(J_2 - 1)$ |
| {1, 2} | 4 | $(J_1 - 1)(J_2 - 1)$ |
| {3} | 5 | $J_3 - 1$ |
| {1, 3} | 6 | $(J_1 - 1)(J_3 - 1)$ |
| {2, 3} | 7 | $(J_2 - 1)(J_3 - 1)$ |
| {1, 2, 3} | 8 | $(J_1 - 1)(J_2 - 1)(J_3 - 1)$ |
| {4} | 9 | $(J_4 - 1)$ |
| {1, 4} | 10 | $(J_1 - 1)(J_4 - 1)$ |
| {2, 4} | 11 | $(J_2 - 1)(J_4 - 1)$ |
| {1, 2, 4} | 12 | $(J_1 - 1)(J_2 - 1)(J_4 - 1)$ |
| {3, 4} | 13 | $(J_3 - 1)(J_4 - 1)$ |
| {1, 3, 4} | 14 | $(J_1 - 1)(J_3 - 1)(J_4 - 1)$ |
| {2, 3, 4} | 15 | $(J_2 - 1)(J_3 - 1)(J_4 - 1)$ |
| {1, 2, 3, 4} | 16 | $(J_1 - 1)(J_2 - 1)(J_3 - 1)(J_4 - 1)$ |
| {5} | 17 | $(J_1 - 1)$ |
| {1, 5} | 18 | $(J_1 - 1)(J_5 - 1)$ |

| | | |
|-----------------|----|---|
| {2, 5} | 19 | $(J_2 - 1)(J_5 - 1)$ |
| {1, 2, 5} | 20 | $(J_1 - 1)(J_2 - 1)(J_5 - 1)$ |
| {3, 5} | 21 | $(J_3 - 1)(J_5 - 1)$ |
| {1, 3, 5} | 22 | $(J_1 - 1)(J_3 - 1)(J_5 - 1)$ |
| {2, 3, 5} | 23 | $(J_2 - 1)(J_3 - 1)(J_5 - 1)$ |
| {4, 5} | 24 | $(J_4 - 1)(J_5 - 1)$ |
| {2, 4, 5} | 25 | $(J_2 - 1)(J_4 - 1)(J_5 - 1)$ |
| {1, 4, 5} | 26 | $(J_1 - 1)(J_4 - 1)(J_5 - 1)$ |
| {3, 4, 5} | 27 | $(J_3 - 1)(J_4 - 1)(J_5 - 1)$ |
| {1, 3, 4, 5} | 28 | $(J_1 - 1)(J_3 - 1)(J_4 - 1)(J_5 - 1)$ |
| {2, 1, 4, 5} | 29 | $(J_2 - 1)(J_1 - 1)(J_4 - 1)(J_5 - 1)$ |
| {1, 2, 3, 5} | 30 | $(J_1 - 1)(J_2 - 1)(J_3 - 1)(J_5 - 1)$ |
| {2, 3, 4, 5} | 31 | $(J_2 - 1)(J_3 - 1)(J_4 - 1)(J_5 - 1)$ |
| {1, 2, 3, 4, 5} | 32 | $(J_1 - 1)(J_2 - 1)(J_3 - 1)(J_4 - 1)(J_5 - 1)$ |

bem como

$$A(\emptyset) = \frac{1}{\sqrt{J_1}} 1_{J_1}^t \otimes \frac{1}{\sqrt{J_2}} 1_{J_2}^t \otimes \frac{1}{\sqrt{J_3}} 1_{J_3}^t \otimes \frac{1}{\sqrt{J_4}} 1_{J_4}^t \otimes \frac{1}{\sqrt{J_5}} 1_{J_5}^t$$

$$A(\{1\}) = T_{J_1} \otimes \frac{1}{\sqrt{J_2}} 1_{J_2}^t \otimes \frac{1}{\sqrt{J_3}} 1_{J_3}^t \otimes \frac{1}{\sqrt{J_4}} 1_{J_4}^t \otimes \frac{1}{\sqrt{J_1}} 1_{J_5}^t$$

$$A(\{2\}) = \frac{1}{\sqrt{J_1}} 1_{J_1}^t \otimes T_{J_2} \otimes \frac{1}{\sqrt{J_3}} 1_{J_3}^t \otimes \frac{1}{\sqrt{J_4}} 1_{J_4}^t \otimes \frac{1}{\sqrt{J_1}} 1_{J_5}^t$$

$$A(\{1, 2\}) = T_{J_1} \otimes T_{J_2} \otimes \frac{1}{\sqrt{J_3}} 1_{J_3}^t \otimes \frac{1}{\sqrt{J_4}} 1_{J_4}^t \otimes \frac{1}{\sqrt{J_1}} 1_{J_5}^t$$

$$A(\{3\}) = \frac{1}{\sqrt{J_1}} 1_{J_1}^t \otimes \frac{1}{\sqrt{J_2}} 1_{J_2}^t \otimes T_{J_3} \otimes \frac{1}{\sqrt{J_4}} 1_{J_4}^t \otimes \frac{1}{\sqrt{J_1}} 1_{J_5}^t$$

$$A(\{1, 3\}) = T_{J_1} \otimes \frac{1}{\sqrt{J_2}} 1_{J_2}^t \otimes T_{J_3} \otimes \frac{1}{\sqrt{J_4}} 1_{J_4}^t \otimes \frac{1}{\sqrt{J_1}} 1_{J_5}^t$$

$$A(\{2, 3\}) = \frac{1}{\sqrt{J_1}} 1_{J_1}^t \otimes T_{J_2} \otimes T_{J_3} \otimes \frac{1}{\sqrt{J_4}} 1_{J_4}^t \otimes \frac{1}{\sqrt{J_1}} 1_{J_5}^t$$

$$A(\{1, 2, 3\}) = T_{J_1} \otimes T_{J_2} \otimes T_{J_3} \otimes \frac{1}{\sqrt{J_4}} 1_{J_4}^t \otimes \frac{1}{\sqrt{J_1}} 1_{J_5}^t$$

$$A(\{4\}) = \frac{1}{\sqrt{J_1}} 1_{J_1}^t \otimes \frac{1}{\sqrt{J_2}} 1_{J_2}^t \otimes \frac{1}{\sqrt{J_3}} 1_{J_3}^t \otimes T_{J_4} \otimes \frac{1}{\sqrt{J_1}} 1_{J_5}^t$$

$$A(\{1, 4\}) = T_{J_1} \otimes \frac{1}{\sqrt{J_2}} 1_{J_2}^t \otimes \frac{1}{\sqrt{J_3}} 1_{J_3}^t \otimes T_{J_4} \otimes \frac{1}{\sqrt{J_1}} 1_{J_5}^t$$

$$A(\{2, 4\}) = \frac{1}{\sqrt{J_1}} 1_{J_1}^t \otimes T_{J_2} \otimes \frac{1}{\sqrt{J_3}} 1_{J_3}^t \otimes T_{J_4} \otimes \frac{1}{\sqrt{J_1}} 1_{J_5}^t$$

$$A(\{1, 2, 4\}) = T_{J_1} \otimes T_{J_2} \otimes \frac{1}{\sqrt{J_3}} 1_{J_3}^t \otimes T_{J_4} \otimes \frac{1}{\sqrt{J_1}} 1_{J_5}^t$$

$$\begin{aligned}
A(\{3, 4\}) &= \frac{1}{\sqrt{J_1}} 1_{J_1}^t \otimes \frac{1}{\sqrt{J_2}} 1_{J_2}^t \otimes T_{J_3} \otimes T_{J_4} \otimes \frac{1}{\sqrt{J_1}} 1_{J_5}^t \\
A(\{1, 3, 4\}) &= T_{J_1} \otimes \frac{1}{\sqrt{J_2}} 1_{J_2}^t \otimes T_{J_3} \otimes T_{J_4} \otimes \frac{1}{\sqrt{J_1}} 1_{J_5}^t \\
A(\{2, 3, 4\}) &= \frac{1}{\sqrt{J_1}} 1_{J_1}^t \otimes T_{J_2} \otimes T_{J_3} \otimes T_{J_4} \otimes \frac{1}{\sqrt{J_1}} 1_{J_5}^t \\
A(\{5\}) &= \frac{1}{\sqrt{J_1}} 1_{J_1}^t \otimes \frac{1}{\sqrt{J_2}} 1_{J_2}^t \otimes \frac{1}{\sqrt{J_3}} 1_{J_3}^t \otimes \frac{1}{\sqrt{J_4}} 1_{J_4}^t \otimes T_{J_5} \\
A(\{1, 5\}) &= T_{J_1} \otimes \frac{1}{\sqrt{J_2}} 1_{J_2}^t \otimes \frac{1}{\sqrt{J_3}} 1_{J_3}^t \otimes \frac{1}{\sqrt{J_4}} 1_{J_4}^t \otimes T_{J_5} \\
A(\{2, 5\}) &= T_{J_2} \otimes \frac{1}{\sqrt{J_2}} 1_{J_2}^t \otimes \frac{1}{\sqrt{J_3}} 1_{J_3}^t \otimes \frac{1}{\sqrt{J_4}} 1_{J_4}^t \otimes T_{J_5} \\
A(\{1, 2, 5\}) &= T_{J_1} \otimes T_{J_2} \otimes \frac{1}{\sqrt{J_3}} 1_{J_3}^t \otimes \frac{1}{\sqrt{J_4}} 1_{J_4}^t \otimes T_{J_5} \\
A(\{3, 5\}) &= \frac{1}{\sqrt{J_1}} 1_{J_1}^t \otimes \frac{1}{\sqrt{J_2}} 1_{J_2}^t \otimes T_{J_3} \otimes \frac{1}{\sqrt{J_4}} 1_{J_4}^t \otimes T_{J_5} \\
A(\{1, 3, 5\}) &= T_{J_1} \otimes \frac{1}{\sqrt{J_2}} 1_{J_2}^t \otimes T_{J_3} \otimes \frac{1}{\sqrt{J_4}} 1_{J_4}^t \otimes T_{J_5} \\
A(\{2, 3, 5\}) &= \frac{1}{\sqrt{J_1}} 1_{J_1}^t \otimes T_{J_2} \otimes T_{J_3} \otimes \frac{1}{\sqrt{J_4}} 1_{J_4}^t \otimes T_{J_5} \\
A(\{4, 5\}) &= \frac{1}{\sqrt{J_1}} 1_{J_1}^t \otimes \frac{1}{\sqrt{J_2}} 1_{J_2}^t \otimes \frac{1}{\sqrt{J_3}} 1_{J_3}^t \otimes T_{J_4} \otimes T_{J_5} \\
A(\{2, 4, 5\}) &= \frac{1}{\sqrt{J_1}} 1_{J_1}^t \otimes T_{J_2} \otimes \frac{1}{\sqrt{J_3}} 1_{J_3}^t \otimes T_{J_4} \otimes T_{J_5} \\
A(\{1, 4, 5\}) &= T_{J_1} \otimes \frac{1}{\sqrt{J_2}} 1_{J_2}^t \otimes \frac{1}{\sqrt{J_3}} 1_{J_3}^t \otimes T_{J_4} \otimes T_{J_5} \\
A(\{3, 4, 5\}) &= \frac{1}{\sqrt{J_1}} 1_{J_1}^t \otimes \frac{1}{\sqrt{J_2}} 1_{J_2}^t \otimes T_{J_3} \otimes T_{J_4} \otimes T_{J_5} \\
A(\{1, 3, 4, 5\}) &= T_{J_1} \otimes \frac{1}{\sqrt{J_2}} 1_{J_2}^t \otimes T_{J_3} \otimes T_{J_4} \otimes T_{J_5} \\
A(\{2, 1, 4, 5\}) &= T_{J_1} \otimes T_{J_2} \otimes \frac{1}{\sqrt{J_3}} 1_{J_3}^t \otimes T_{J_4} \otimes T_{J_5} \\
A(\{1, 2, 3, 5\}) &= T_{J_1} \otimes T_{J_2} \otimes T_{J_3} \otimes \frac{1}{\sqrt{J_4}} 1_{J_4}^t \otimes T_{J_5} \\
A(\{1, 2, 3, 4, 5\}) &= T_{J_1} \otimes T_{J_2} \otimes T_{J_3} \otimes T_{J_4} \otimes T_{J_5}.
\end{aligned}$$

4.4.2- Cruzamento e aninhamento de fatores

Suponhamos haver u grupos de c_1, \dots, c_u fatores aninhados que cruzam, e que o ℓ –ésimo fator do h -ésimo grupo terá $J_{\ell,h}$ níveis. Dado aninharem os fatores de cada grupo não interagem. Teremos: de considerar

$$\begin{cases} \text{.efeitos dos fatores} \\ \text{.interação entre fatores de grupos distintos} \end{cases}$$

Estes efeitos e interações juntamente com o valor médio geral corresponderão aos vetores de

$$\Gamma = \{\boldsymbol{\ell}; \ell_i = 0, \dots, c_i, i = 1, \dots, u\},$$

tendo -se [ver Fonseca et al. (2007)], para obter as somas de quadrados $S_{\ell}, \ell \in \Gamma$, as matrizes

$$A_{\ell} = \bigotimes_{h=1}^u A_{\ell,h}, \ell \in \Gamma,$$

com

$$A_{\ell,h} = \bigotimes_{h=1}^{\ell_h-1} I_{J_{\ell,h}} \otimes T_{J_{\ell,h}} \otimes \left(\bigotimes_{h=\ell_j+1}^{c_j} \frac{1}{\sqrt{J_{\ell,h}}} 1_{J_{\ell,h}}^t \right), \ell \in \Gamma, j = 1, \dots, f$$

onde se terá

$$\begin{cases} \bigotimes_{h=1}^0 I_{J_{\ell,h}} = [1] \\ \bigotimes_{h=c_j+1}^{c_j} \frac{1}{\sqrt{J_{\ell,h}}} 1_{J_{\ell,h}}^t = [1]. \end{cases}$$

Observe-se que qualquer que seja a matriz W se tem

$$\begin{cases} W \otimes [1] = W \\ [1] \otimes W = W \end{cases}$$

Temos então

$$g_{\ell} = \text{car}(A_{\ell}) = \prod_{h=1}^u \text{car}(A_{\ell,h}); \ell \in \Gamma,$$

com

$$g_{\ell,h} = \text{car}(A_{\ell,h}) = \prod_{h'=1}^{\ell_h-1} J_{\ell,h'} \cdot (J_{\ell,h'} - 1), \ell \in \Gamma, h = 1, \dots, f,$$

as somas de quadrados serão dados por

$$S_{\ell,j} = \|A_{\ell} \tilde{\psi}_j\|^2, \ell \in \Gamma, j = 1, \dots, k.$$

com g_{ℓ} graus de liberdade.

Para estimar o erro com

$$Q_{\ell} = A_{\ell}^t A_{\ell}, \ell \in \Gamma,$$

e

$$\nabla_{\ell} = R(Q_{\ell}),$$

admite-se a priori

$$\psi_j \in (\otimes_{\ell \in \Gamma} (\nabla_{\ell}))^{\perp}, j = 1, \dots, h,$$

com

$$\dot{\Gamma} \subset \Gamma.$$

Definimos ordem da interação associado a $\ell \in \Gamma$ como

$$o(\ell) = \left(\sum_{j=1}^u \ell_j \right) - 1$$

o que generaliza a noção que se tem para os modelos com cruzamento de fatores.

Poderá em geral coloca-se em $\dot{\Gamma}$ os ℓ associado as interações de mais alta ordem

Teremos então as estatísticas F

$$F_{\ell,j} = \frac{g}{g_{\ell}} \frac{S_{\ell,j}}{S_j}, \ell \in \Gamma / \dot{\Gamma}, j = 1, \dots, k,$$

com

$$\begin{cases} S_j = \sum_{\ell \in \dot{\Gamma}} S_{\ell,j}, j = 1, \dots, k \\ g = \sum_{\ell \in \dot{\Gamma}} g_{\ell} \end{cases},$$

sendo g_{ℓ} e g os graus de liberdade.

As hipóteses testadas serão

$$H_{0,\ell,j}: \psi_{\ell,j} \in \nabla_{\ell}^{\perp}, \ell \notin \mathcal{D}, j = 1, \dots, k,$$

com

$$\nabla_{\ell} = R(Q_{\ell}), \ell \notin \mathcal{D},$$

e

$$Q_{\ell} = A_{\ell}^t A_{\ell}, \ell \notin \mathcal{D}.$$

Nesta e na subalínea **4.4.1** privilegiamos as hipóteses sobre funções estimáveis principais, no entanto os nossos resultados adaptam-se ao caso de funções estimáveis quaisquer.

4.4.2.1 – Um exemplo

Admitamos ter $u = 2$ grupos de fatores que cruzam. O primeiro grupo tem dois fatores, $c_1 = 2$, com $J_{1,1} = 3$ e $J_{2,1} = 2$, tendo-se igualmente $c_2 = 2$ mas agora com $J_{1,2} = 2$ e $J_{2,2} = 2$.

Teríamos então as matrizes

$$\begin{cases} T_2 = [-1/\sqrt{2} & 1/\sqrt{2}] \\ T_3 = \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -1/\sqrt{6} & -1/\sqrt{6} & 2/\sqrt{6} \end{bmatrix} \end{cases}$$

que nos permitem obter as

$$\left[\begin{array}{l} A_{0,1} = \frac{1}{\sqrt{3}} 1_3^t \otimes \frac{1}{\sqrt{2}} 1_2^t \\ A_{1,1} = T_3 \otimes \frac{1}{\sqrt{2}} 1_2^t \\ A_{2,1} = I_3 \otimes T_2 \\ A_{0,2} = \frac{1}{\sqrt{2}} 1_2^t \otimes \frac{1}{\sqrt{2}} 1_2^t \\ A_{1,2} = T_2 \otimes \frac{1}{\sqrt{2}} 1_2^t \\ A_{2,2} = I_2 \otimes T_2 \end{array} \right.$$

Por outro lado, temos

$$\Gamma = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}; \begin{bmatrix} 0 \\ 1 \end{bmatrix}; \begin{bmatrix} 0 \\ 2 \end{bmatrix}; \begin{bmatrix} 1 \\ 0 \end{bmatrix}; \begin{bmatrix} 1 \\ 1 \end{bmatrix}; \begin{bmatrix} 1 \\ 2 \end{bmatrix}; \begin{bmatrix} 2 \\ 0 \end{bmatrix}; \begin{bmatrix} 2 \\ 1 \end{bmatrix}; \begin{bmatrix} 2 \\ 2 \end{bmatrix} \right\}$$

, vindo

$$\left[\begin{array}{l} A_{\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}}^{[0]} = A_{0,1} \otimes A_{0,2} \\ A_{\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}}^{[0]} = A_{0,1} \otimes A_{1,2} \\ A_{\begin{smallmatrix} 0 \\ 2 \end{smallmatrix}}^{[0]} = A_{0,1} \otimes A_{2,2} \\ A_{\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}}^{[1]} = A_{1,1} \otimes A_{0,2} \\ A_{\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}}^{[1]} = A_{1,1} \otimes A_{1,2} \\ A_{\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}}^{[1]} = A_{1,1} \otimes A_{2,2} \\ A_{\begin{smallmatrix} 2 \\ 0 \end{smallmatrix}}^{[2]} = A_{2,1} \otimes A_{0,2} \\ A_{\begin{smallmatrix} 2 \\ 1 \end{smallmatrix}}^{[2]} = A_{2,1} \otimes A_{1,2} \\ A_{\begin{smallmatrix} 2 \\ 2 \end{smallmatrix}}^{[2]} = A_{2,1} \otimes A_{2,2} \end{array} \right.$$

as características destas matrizes sendo dadas por

$$\text{car} \left(A_{\begin{smallmatrix} \ell_1 \\ \ell_2 \end{smallmatrix}} \right) = \text{car}(A_{\ell_1,1}) \text{car}(A_{\ell_2,2}),$$

já que $\text{car}(W_1 \otimes W_2) = \text{car}(W_1) \text{car}(W_2)$, tendo-se portanto

$$\left[\begin{array}{l} \text{car}(A_{0,1}) = 1 \cdot 1 = 1 \\ \text{car}(A_{1,1}) = 2 \cdot 1 = 2 \\ \text{car}(A_{2,1}) = 3 \cdot 1 = 3 \\ \text{car}(A_{0,2}) = 1 \cdot 1 = 1 \\ \text{car}(A_{1,2}) = 1 \cdot 1 = 1 \\ \text{car}(A_{2,2}) = 2 \cdot 1 = 2 \end{array} \right.$$

Logo

$$\left[\begin{array}{l} \text{car} \left(A_{\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}} \right) = 1.1 = 1 \\ \text{car} \left(A_{\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}} \right) = 1.1 = 1 \\ \text{car} \left(A_{\begin{smallmatrix} 0 \\ 2 \end{smallmatrix}} \right) = 1.2 = 2 \\ \text{car} \left(A_{\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}} \right) = 1.1 = 1 \\ \text{car} \left(A_{\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}} \right) = 1.1 = 1 \\ \text{car} \left(A_{\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}} \right) = 1.2 = 2 \\ \text{car} \left(A_{\begin{smallmatrix} 2 \\ 0 \end{smallmatrix}} \right) = 2.1 = 2 \\ \text{car} \left(A_{\begin{smallmatrix} 2 \\ 1 \end{smallmatrix}} \right) = 2.1 = 2 \\ \text{car} \left(A_{\begin{smallmatrix} 2 \\ 2 \end{smallmatrix}} \right) = 2.2 = 4 \end{array} \right.$$

4.4.3 – Fatoriais de base prima

Estes delineamentos encontram-se estudados na dissertação de Doutoramento Oliveira (2018).

4.4.3.1 – Fatoriais completos

Vamos considerar o caso em que o delineamento base é um fatorial de base prima tendo n fatores com p níveis que cruzam.

A base principal, da AJC relevante, $\mathcal{A}(p^n)$, é, como vimos,

$$\wp(\mathcal{A}(p^n)) = \{\mathbf{Q}_0, Q(\mathbf{L}), L \in \mathcal{L}_r([p]^n)\},$$

com

$$\mathbf{Q}_0 = \frac{1}{p^n} 1_{p^n} 1_{p^n}^t$$

$$Q(\mathbf{L}) = A(\mathbf{L})^t A(\mathbf{L}); L \in \mathcal{L}_r([p]^n),$$

onde

$$A(\mathbf{L}) = \mathbf{T}_p B(\mathbf{L}), L \in \mathcal{L}_r([p]^n),$$

com $B(\mathbf{L})$ dada anteriormente [ver Oliveira (2018)].

Sendo as componentes dos $\boldsymbol{\psi}_j$ e $\tilde{\boldsymbol{\psi}}_j, j = 1, \dots, k$ as funções estimáveis principais e respectivos estimadores, ponhamos

$$S_j(\mathbf{L}) = \|A(\mathbf{L})\tilde{\boldsymbol{\psi}}_j\|^2, j = 1, \dots, k,$$

bem como

$$\delta_j(\mathbf{L}) = \frac{1}{\sigma^2} \|A(\mathbf{L})\boldsymbol{\psi}_j\|^2, j = 1, \dots, k.$$

Observe-se agora que as ordens das

$$L(\mathbf{x}|\mathbf{a}) = \left(\sum_{i=1}^n a_i x_i \right)_{(p)},$$

são os números dos seus coeficientes não nulos menos 1.

Seja $\sqrt{d}, d = 0, \dots, n - 1$ a família das aplicações lineares reduzidas de ordem não inferior a $d, d = 0, \dots, n - 1$.

Considere-se o pressuposto P_d que consiste em admitir-se que

$$\delta_j(\mathbf{L}) = 0, j = 1, \dots, k, L \in \sqrt{d},$$

então, se $L \notin \sqrt{d}$, podemos testar as hipóteses

$$H_{0,j}(\mathbf{L}): \delta_j(\mathbf{L}) = 0, j = 1, \dots, k, L \notin \sqrt{d},$$

sendo $c = \#(\sqrt{d})$ teremos as estatísticas de teste

$$F_j = c \frac{S_j(\mathbf{L})}{S_j}; L \notin \sqrt{d}, j = 1, \dots, k$$

com, aproximadamente, $p - 1$ e $c(p - 1)$ graus de liberdade, e

$$S_j = \sum_{L \in \sqrt{d}} S_j(L), j = 1, \dots, k.$$

Para mais detalhes ver [5]

Vejamos um exemplo com $p = 3$ e $n = 3$. As aplicações reduzidas podem agrupar-se de acordo com o índice do primeiro coeficiente não nulo. As aplicações reduzidas neste caso e as respetivas ordens são

$$\left[\begin{array}{l} 1: \quad x_1(0) \quad x_1 + x_2 (1) \quad x_1 + x_3 (1) \quad x_1 + x_2 + x_3 (2) \\ \quad \quad \quad x_1 + 2x_2 (1) \quad x_1 + 2x_3 (1) \quad x_1 + x_2 + 2x_3 (2) \\ \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad x_1 + 2x_2 + x_3 (2) \\ \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad x_1 + 2x_2 + 2x_3 (2) \\ \\ 2: \quad x_2(0) \quad x_2 + x_3(1) \\ \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad x_2 + 2x_3(1) \end{array} \right.$$

$$3: \quad x_3(0),$$

sendo fácil de a partir daqui obter \sqrt{d} . por exemplo

$$\sqrt{(2)} = \{x_1 + x_2 + x_3; x_1 + x_2 + 2x_3; x_1 + 2x_2 + x_3; x_1 + 2x_2 + 2x_3\},$$

com $c = 4$.

4.4.3.2 – Exemplo de construção de Blocos

Interessa-nos agora considerar um exemplo de construção de blocos.

Tendo-se um fatorial completo p^n teremos $p^m (m < n)$ blocos com p^{n-m} tratamentos.

Sendo $L(|\mathbf{a}_1) \dots L(|\mathbf{a}_m)$, m aplicações lineares reduzidas linearmente independentes utilizadas na construção de blocos temos

$$A(L)^t = [\mathbf{a}_1 \dots \mathbf{a}_m],$$

com

$$\text{car}(A(L)) = \text{car}(A(L)^t) = m.$$

Por exemplo com $m = 2$ e $n = 4$ e $p = 3$ poderá ser

$$A(L) = \begin{bmatrix} 1 & 2 & 1 & 2 \\ 1 & 1 & 2 & 1 \end{bmatrix},$$

isto é o bloco genérico será construído pelas soluções de

$$A(L)\mathbf{x} = \mathbf{b},$$

com $\mathbf{b} = (b_1, b_2)$, ter-se-á o sistema

$$\begin{cases} x_1 + 2x_2 + x_3 + 2x_4 = b_1 \\ x_1 + x_2 + 2x_3 + x_4 = b_2 \end{cases}$$

Observe-se que na aritmética modulo 3 têm-se os simétricos e os inversos dados por

$$\begin{array}{r|l} x & -x \\ \hline 0 & 0 \\ 1 & 2 \\ 2 & 1 \end{array}$$

$$\begin{array}{r|l} x & x^{-1} \\ \hline 1 & 1 \\ 2 & 2 \end{array}$$

, o que permite obter as soluções

$$\begin{cases} x_1 = 2b_2 + 2b_1 + 2x_4 \\ x_2 = 2b_2 + b_1 + x_3 + x_4. \end{cases}$$

Ora para cada par (b_1, b_2) tem-se um bloco. Por exemplo para o par $(2, 0)$ tem-se

$$\begin{cases} x_1 = 1 + 2x_4 \\ x_2 = 1 + x_3 + x_4' \end{cases}$$

e os tratamentos que constituem o bloco são

| x_1 | x_2 | x_3 | x_4 |
|-------|-------|-------|-------|
| 1 | 1 | 0 | 0 |
| 0 | 2 | 0 | 1 |
| 2 | 0 | 0 | 2 |
| 1 | 2 | 1 | 0 |
| 0 | 0 | 1 | 1 |
| 2 | 1 | 1 | 2 |
| 1 | 0 | 2 | 0 |
| 0 | 1 | 0 | 1 |
| 2 | 2 | 2 | 2 |

4.4.3.3 – Replicações Fracionárias

Utiliza-se apenas um dos blocos. Sendo $L_i(\mathbf{x}) = L(\mathbf{x}|\mathbf{a}_i), i = 1, \dots, s$ as aplicações reduzidas utilizadas para construir os blocos, teremos apenas $\frac{1}{p^s} \cdot p^n$ tratamentos no bloco escolhido.

Com

$$\begin{cases} \mathbf{L}(\mathbf{x}) = (L_1(\mathbf{x}), \dots, L_s(\mathbf{x})) \\ \mathbf{b} = (b_1 \dots b_s), \end{cases}$$

teremos os blocos

$$[L|\boldsymbol{\theta}] = \{\mathbf{x}: L_i(\mathbf{x}) = \theta_i, i = 1, \dots, s\}$$

Dado $\mathcal{L}([p]^n)$ ter dimensão n e as $L_i(\mathbf{x}), i = 1, \dots, s$ serem linearmente independentes podemos completar $\{L_1, \dots, L_s\}$ para obter uma base $\{L_1 \dots L_s, L_{s+1}, \dots, L_n\}$. Então, dada $L \in \mathcal{L}_r([p]^n)$, temos

$$L(\mathbf{x}) = \left(\sum_{i=1}^n a_i L_i(\mathbf{x}) \right)_{(p)},$$

com $a_1, \dots, a_n \in G[p]$ e $\mathbf{x} \in G[p]^n$.

Caso $\mathbf{x} \in [L|\boldsymbol{\theta}]$ ter-se-á

$$L(\mathbf{x}) = \left(\sum_{i=1}^s a_i \theta_i + \sum_{i=s+1}^n a_i L_i(\mathbf{x}) \right)_{(p)},$$

pelo que, se $\mathbf{x}_1, \mathbf{x}_2 \in [L|\boldsymbol{\theta}]$, tem-se

$$L(\mathbf{x}_2) - L(\mathbf{x}_1) = \left[\left(\sum_{i=s+1}^n a_i L_i(\mathbf{x}_2) \right) - \left(\sum_{i=s+1}^n a_i L_i(\mathbf{x}_1) \right) \right].$$

Então com

$$\dot{L}_{\hbar}(\mathbf{x}) = \sum_{i=1}^n a_{i,\hbar} L_i(\mathbf{x}), \hbar = 1, 2,$$

tem-se

$$\dot{L}_2(\mathbf{x}_2) - \dot{L}_2(\mathbf{x}_1) = \dot{L}_1(\mathbf{x}_2) - \dot{L}_1(\mathbf{x}_1) \quad (4.4)$$

se e só se

$$a_{i,2} = a_{i,1} = a_i, i = 1, \dots, s,$$

teremos então

$$L'_{\hbar}(\mathbf{x}) = L''(\mathbf{x}) + \sum_{i=s+1}^n a_{i,\hbar} L_i(\mathbf{x}), \hbar = 1, 2,$$

com

$$L''(\mathbf{x}) = \sum_{i=1}^s a_i L_i(\mathbf{x}).$$

Pondo-se

$$\dot{L}_1 \theta \dot{L}_2,$$

quando (4.4) se verifica. Define-se assim uma relação de equivalência θ em $\mathcal{L}_r([p]^n)$.

Caso $\dot{L}_1 \theta \dot{L}_2$, tem-se $\dot{L}_1(\mathbf{x}_1) = \dot{L}_1(\mathbf{x}_2)$ se e só se $\dot{L}_2(\mathbf{x}_1) = \dot{L}_2(\mathbf{x}_2)$ pelo que as duas aplicações \dot{L}_1 e \dot{L}_2 agrupam os p^n tratamentos nas mesmas classes. Assim ao desagregar-se a soma de quadrados para tratamentos pode-se escolher uma aplicação por classe de equivalência θ , procurando-se que os efeitos (ordem 0) e as interações fatoriais de baixa ordem estejam entre as escolhidas.

Vejamos um exemplo com $p = 3$ e $n = 4$. Considere-se uma replicação fracionária $\frac{1}{3} x 3^4$. Tem-se então $s = 1$ e suponhamos que se utilizem

$$L_1(\mathbf{x}) = x_1 + 2x_2 + 2x_3 + 2x_4,$$

para construir os blocos. Suponhamos ainda que foi escolhido o bloco $[L_1|2]$, tendo-se pois

$$x_1 + 2x_2 + 2x_3 + 2x_4 = 2$$

vindo dado (na aritmética modelo 3 ter-se $-0 = 0$, $-1 = 2$ e $-2 = 1$)

$$2x_4 = 2 + 2x_1 + x_2 + x_3,$$

e como, nessa aritmética, $2^{-1} = 2$ e $2 x 2 = 1$, temos

$$x_4 = 1 + x_1 + 2x_2 + 2x_3,$$

obtêm-se assim os tratamentos do bloco escolhido

| x_1 | x_2 | x_3 | x_4 | x_1 | x_2 | x_3 | x_4 | x_1 | x_2 | x_3 | x_4 |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 0 | 0 | 0 | 1 | 1 | 0 | 0 | 2 | 2 | 0 | 0 | 0 |
| 0 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 2 | 0 | 1 | 2 |
| 0 | 0 | 2 | 2 | 1 | 0 | 2 | 0 | 2 | 0 | 2 | 1 |
| 0 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 2 | 1 | 0 | 2 |
| 0 | 1 | 1 | 2 | 1 | 1 | 1 | 0 | 2 | 1 | 1 | 1 |
| 0 | 1 | 2 | 1 | 1 | 1 | 2 | 2 | 2 | 1 | 2 | 0 |
| 0 | 2 | 0 | 2 | 1 | 2 | 0 | 0 | 2 | 2 | 0 | 1 |
| 0 | 2 | 1 | 1 | 1 | 2 | 1 | 2 | 2 | 2 | 1 | 0 |
| 0 | 2 | 2 | 0 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 |

Observe-se que dado os blocos serem construídos a partir duma única aplicação reduzida.

$$L_1 = x_1 + 2x_2 + 2x_3 + 2x_4,$$

a classe da equivalência θ , a que pertence uma aplicação L é

$$\{L; L + L_1; L + 2L_1\},$$

tendo-se

$$2L_1 = 2x_1 + x_2 + x_3 + x_4.$$

Vamos obter essas classes para as aplicações reduzidas de ordem 0 e 1. Restringimo-nos às aplicações reduzidas pertencente á classe. Assim se o primeiro coeficiente não nulo for 2 multiplicam-se os coeficientes por $2 = 2^{-1}$.

| Ordem | Aplicação | Classe |
|-------|--------------|---|
| 0 | x_1 | $x_1; x_1 + x_2 + x_3 + x_4; x_2 + x_3 + x_4$ |
| | x_2 | $x_2; x_1 + 2x_3 + 2x_4; x_1 + x_2 + 2x_3 + 2x_4$ |
| | x_3 | $x_3; x_1 + 2x_2 + 2x_4; x_1 + 2x_2 + x_3 + 2x_4$ |
| | x_4 | $x_4; x_1 + 2x_2 + 2x_3; x_1 + 2x_2 + 2x_3 + x_4$ |
| 1 | $x_1 + x_2$ | $x_1 + x_2; x_1 + x_3 + x_4; x_2 + 2x_3 + 2x_4$ |
| | $x_1 + 2x_2$ | $x_1 + 2x_2; x_1 + 2x_2 + x_3 + x_4; x_3 + x_4$ |
| | $x_1 + x_3$ | $x_1 + x_3; x_1 + x_2 + x_4; x_3 + x_4$ |
| | $x_1 + 2x_3$ | $x_1 + 2x_3; x_1 + x_2 + 2x_3 + x_4; x_2 + x_4$ |
| | $x_1 + x_4$ | $x_1 + x_4; x_1 + x_2 + x_3; x_2 + x_3 + x_4$ |
| | $x_1 + 2x_4$ | $x_1 + 2x_4; x_1 + 2x_2 + 2x_3 + x_4; x_2 + x_3$ |
| | $x_2 + x_3$ | $x_2 + x_3; x_1 + x_4; x_1 + x_2 + x_4$ |
| | $x_2 + 2x_3$ | $x_2 + 2x_3; x_1 + 2x_3 + x_4; x_1 + x_2 + x_4$ |
| | $x_2 + x_4$ | $x_2 + x_4; x_1 + x_3; x_1 + x_2 + x_3 + x_4$ |
| | $x_2 + 2x_4$ | $x_2 + 2x_4; x_1 + x_3 + x_4; x_1 + x_2 + x_3$ |
| | $x_3 + x_4$ | $x_3 + x_4; x_1 + x_2; x_1 + x_2 + x_3 + x_4$ |
| | $x_3 + 2x_4$ | $x_3 + 2x_4; x_1 + x_2 + x_4; x_1 + x_2 + x_3$ |

vendo-se que as aplicações reduzidas com ordem 1 são as únicas com essas ordens nas respectivas classes de equivalência θ podendo ser escolhida.

4.4.4. 1 – Um algoritmo útil

Vamos agora introduzir um algoritmo útil para obter as matrizes $B(L)$. Se observamos as expressões (2.1) vemos que a chave dessa construção está nos vetores

$$\boldsymbol{v}(L) = (v_1(L), \dots, v_{p^n}(L)),$$

onde

$$v_j(L) = L(x_j), j = 1, \dots, p^n,$$

sendo os índices $j = j(\mathbf{x})$ são dados por (2.2).

Convém-nos agora decompor a família $\mathcal{L}_r([p]^n)$ nas sub-famílias $\mathcal{L}_{r,i}([p]^n), i = 1, \dots, n$ das aplicações reduzidas cujo primeiro coeficiente não nulo, necessariamente igual a 1, é o i -ésimo. Por exemplo com $p = 3$ e $n = 3$ teremos

$$\begin{aligned} \mathcal{L}_{r,1}([3]^3) = \{ & x_1, x_1 + x_3, x_1 + 2x_3, x_1 + x_2, x_1 + x_2 + x_3, x_1 + x_2 + 2x_3, x_1 \\ & + 2x_2, x_1 + 2x_2 + x_3, x_1 + 2x_2 + 2x_3 \} \end{aligned}$$

$$\mathcal{L}_{r,2}([3]^3) = \{x_2, x_2 + x_3, x_2 + 2x_3\}$$

$$\mathcal{L}_{r,3}([3]^3) = \{x_3\}.$$

Em geral ter-se-á

$$\mathcal{L}_r([p]^n) = U_{i=1}^n \mathcal{L}_{r,i}([p]^n),$$

com

$$\begin{cases} \# (\mathcal{L}_r[p]^n) = \frac{p^n - 1}{p - 1} \\ \# (\mathcal{L}_{r,i}([p]^n)) = p^{n-i}, i = 1, \dots, n \end{cases}$$

Assim a matriz $V(n)$ cujos vetores linha são os $\boldsymbol{v}(L)$ das $L \in \mathcal{L}_r([p]^n)$ é dada por

$$V(n) = \begin{bmatrix} V(n, 1) \\ \vdots \\ V(n, n) \end{bmatrix},$$

onde $V(n, i)$ contém os $\boldsymbol{v}(L)$ das $L \in \mathcal{L}_{r,i}([p]^n), i = 1, \dots, n$.

Vamos agora obter uma fórmula de recorrência para as $V(n, 1)$ que são matrizes do tipo

$$p^{n-1} \cdot p^n.$$

Em particular ter-se-á

$$V(n+1, 1) = \begin{bmatrix} V_{1,1}(n+1, 1) & \dots & V_{1,p}(n+1, 1) \\ \vdots & \vdots & \vdots \\ V_{p,1}(n+1, 1) & \dots & V_{p,p}(n+1, 1) \end{bmatrix},$$

com

$$V_{\ell, h}(n+1, 1) = V_1(n, 1) + (\ell - 1)(h - 1) \cdot 1_{p^{n-1}} 1_{p^n}^t,$$

já que podemos atribuir á sub-matriz $V_{\ell, h}(n+1, 1)$ as $L \in \mathcal{L}_{r, n+1}([p]^{n+1})$ cujo $(n+1)$ -ésimo coeficiente é $\ell - 1$, e nessa sub-matriz colocar as componentes dos respetivos vetores $\mathbf{v}(L)$ com índices $(h-1)p^n + 1, \dots, hp^n$ em que a “contribuição” do último termo da aplicação reduzida é $(\ell - 1)(h - 1), \ell = 1, \dots, p, h = 1, \dots, p$.

Por exemplo com $n = 3$ e $p = 3$ teremos

$$V(1) = [0 \ 1 \ 2] = V(1, 1)$$

e

$$V(2) = \begin{bmatrix} V(2, 1) \\ V(2, 2) \end{bmatrix},$$

com

$$V(2, 1) = \begin{bmatrix} V_{1,1}(2, 1) & V_{1,2}(2, 1) & V_{1,3}(2, 1) \\ V_{2,1}(2, 1) & V_{2,2}(2, 1) & V_{2,3}(2, 1) \\ V_{3,1}(2, 1) & V_{3,2}(2, 1) & V_{3,3}(2, 1) \end{bmatrix},$$

onde

$$\begin{cases} V_{1,1}(2, 1) = V_{1,2}(2, 1) = V_{1,3}(2, 1) = V(1) \\ V_{2,1}(2, 1) = V_{3,1}(2, 1) = V(1) + 2 \cdot 1_1^t, \\ V_{2,2}(2, 1) = V_{3,3}(1)^{(3)} + 1 \cdot 1_3^t, \end{cases}$$

vindo

$$V(2, 1) = \begin{bmatrix} 0 \ 1 \ 2 & 0 \ 1 \ 2 & 0 \ 1 \ 2 \\ 0 \ 1 \ 2 & 1 \ 2 \ 0 & 2 \ 0 \ 1 \\ 0 \ 1 \ 2 & 2 \ 0 \ 1 & 1 \ 2 \ 0 \end{bmatrix}$$

Por outro lado, as aplicações reduzidas pertencentes a $\mathcal{L}_{r, i}([p]^n)$, têm os $\ell - 1$ primeiros coeficientes nulos pelo que as componentes dos respetivos vetores $\mathbf{v}(L)$ agrupam-se em grupos contíguais de $p^{\ell-1}$ componentes idênticas as de vetor $\mathbf{v}(\dot{L})$ com \dot{L} a aplicação de $\mathcal{L}_r([p]^{i-i+1})$ cujos coeficientes são os últimos coeficientes de L .

Assim

$$V(n, i) = V(n - i + 1, 1) \otimes 1_{p^{i-1}}^t,$$

em particular, com $p = 3$

$$V(3, 2) = V(2, 1) \otimes 1_3^t,$$

$$V(3, 3) = V(1, 1) \otimes 1_9^t.$$

Vê-se assim como aplicar este algoritmo numa forma de recorrência que pode ser programada.

Capítulo 5

No presente capítulo fez-se a síntese dos resultados da investigação, mencionando cada capítulo o que foi feito e como foi possível e por fim faz-se a apresentação de trabalho futuro.

5 – Considerações Finais

Os aspetos mais relevantes deste trabalho parecem-me ser, ter-se:

* considerado a comutação das matrizes M_1, \dots, M_w [M, M_1, \dots, M_w] para obter BQUE [BLUE] para as componentes de variância e covariância [o vetor dos coeficientes]. Este estudo completou-se com a introdução das famílias OBS e COBS de modelos;

** considerado as famílias regressionais de modelos de forma a estudar a ação de variáveis controladas distintas das que intervêm na construção dos modelos sobre os vetores de coeficientes;

*** o ter-se detalhado a estrutura dos delineamentos de efeitos fixos cujos tratamentos correspondem aos modelos das famílias estruturadas. Assim consideram-se os modelos com:

cruzamento de fatores

cruzamento e aninhamento de fatores, no cruzamento de fatores considerou-se em particular o caso de fatores com um número primo 1 de fatores:

* Explicitado a possibilidade das componentes dos vetores aleatórios Z_1, \dots, Z_w que intervêm nos modelos bi-aditivos poderemos ter distribuições de diferentes tipos: esta

possibilidade começou a poder ser explorada como a introdução dos modelos com vetores de erros da forma

$$\mathbf{Y} = X\boldsymbol{\beta} + X_1\mathbf{Z}_1 + \mathbf{E}$$

onde as componentes de \mathbf{Z}_1 podem ter qualquer distribuição e \mathbf{E} é normal com vetor médio nulo e matriz de covariância $\sigma^2\mathbf{I}_n$. Ao explorar as possibilidades dos modelos bi-aditivos obtivemos para os mesmos:

elipsoide de confiança com probabilidade (aproximadamente) dada de conterem realizações dos modelos;

intervalos de com probabilidade (aproximadamente) dada de conterem observações futuras associadas a conjuntos de valores das variáveis controladas intervenientes na construção dos modelos.

A obtenção de intervalos de predição parece-nos muito mais realista que a abordagem baseada em preditores. Assim passamos a ter um intervalo com probabilidade fixada por nós de conter uma observação futura em vez dum valor presente para a mesma.

Julgamos que a nossa abordagem aos modelos bi-aditivos evidencia as potencialidades dos mesmos.

De facto, pode pensar-se que em certos aspetos os mesmos serão difíceis de generalizar. Assim pretendemos estudar modelos de forma

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

Em que os vetores \mathbf{Z}_i são independentes com matrizes de covariância $\sigma_1^2\mathbf{C}_1, \dots, \sigma_w^2\mathbf{C}_w$ conhecidas a menos de $\sigma_1^2, \dots, \sigma_w^2$.

Ter-se-á então

$$\Sigma(\mathbf{Y}) = \sum_{i=1}^w \sigma_i^2 \mathbf{M}_i$$

com $\mathbf{M}_i = \mathbf{X}_i\mathbf{C}_i\mathbf{X}_i^t, i = 1, \dots, w$

o que mostra terem modelos terem matrizes de covariância com estrutura aditiva sendo, portanto, uma “generalização” dos modelos bi-aditivos.

Admitindo-se que as matrizes $\mathbf{C}_1, \dots, \mathbf{C}_w$ são invertíveis tendo-se

$$\mathbf{C}_i = \mathbf{P}_i\mathbf{D}_i\mathbf{P}_i^t, i = 1, \dots, w$$

onde as matrizes $\mathbf{P}_1, \dots, \mathbf{P}_w$ são ortogonais e as $\mathbf{D}_1, \dots, \mathbf{D}_w$ diagonais

com elementos principais positivos escalares as matrizes $\mathbf{D}_1^{-1/2}, \dots, \mathbf{D}_w^{-1/2}$ diagonais e as

$$C_i^{-1/2} = \mathbf{P}_i \mathbf{D}_i^{-1/2} \mathbf{P}_i^t, i = 1, \dots, w$$

Tomando-se

$$\begin{cases} \dot{\mathbf{X}}_i = \mathbf{X}_i C_i^{-1/2}, i = 1, \dots, w \\ \dot{\mathbf{Z}}_i = C_i^{-1/2} \mathbf{Z}_i, i = 1, \dots, w \end{cases}$$

tem-se

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

Com os vetores $\dot{\mathbf{Z}}_1, \dots, \dot{\mathbf{Z}}_w$ independentes com matrizes de covariância $\sigma_1^2 \mathbf{I}_{C_1}, \dots, \sigma_w^2 \mathbf{I}_{C_w}$.

De certa forma generalizamos os modelos anteriores, mas não podemos garantir a independência das componentes dos $\dot{\mathbf{Z}}_1, \dots, \dot{\mathbf{Z}}_w$ o que é importante, por exemplo no ajustamento dos cumulantes dessas componentes.

Pode ultrapassar-se esta dificuldade se admitirmos a normalidade dos $\dot{\mathbf{Z}}_1, \dots, \dot{\mathbf{Z}}_w$ e, conseqüentemente, a independência das respectivas componentes. No futuro esperamos explorar o que se passa no caso geral e quando admitimos a normalidade.

Referências Bibliográficas

- [01] Antunes, P., Ferreira, S., Nunes, C. & Mexia, J. (2020a). *Estimation in additive Models and ANOVA-like applications*, article, Journal of Applied Statistics.
- [02] Antunes, P., Ferreira, S., Ferreira, Dário. & Mexia, J. (2020b). Multiple Additive Models. *Communications in Statistics - Theory and methods*.
- [03] Balakrishnan, N., Johnson, N. L. & Kotz, S. (1998). *A note on relationships between moments, central moments and cumulants from multivariate distributions*. Stat Probabil Lett. pp.49-54.
- [04] Barbosa, E., Silva, C., Manuli, R., Tavares, R. & Nazaré, T. (2014). *Distribuição generalizada de Valores Extremos (GVE): Um estudo aplicado a valores de temperatura mínima da Cidade de Viçosa-MG*. Revista da Estatística da Universidade Federal de Ouro Preto, 3(3), pp.387-391.
- [05] Carvalho, F., Mexia, J. T., Sandro, C. & Nunes C. (2015). *Inference for types and structured families with commutative orthogonal block structure*. Metriks pp.337-372.
- [06] Cayley, A. (1858). *A memoir on the theory of matrices*. Philosophical Transactions of the Royal Society of London, v.148, pp.17-37.
- [07] Craig, C. C. (1931). *On a Property of Semi-Invariants of Thiele*. Ann Math Statist, pp.154-164.
- [08] Douglas, J. B. (1980). *Analysis with standard contagious distributions*. International Co-operative Publishing House Fairland, Maryland.

- [09] Ferreira, S. S., Ferreira, D., Nunes, C. & Mexia, J. T. (2013). *Estimation of variance components in linear mixed Models with commutating orthogonal Block*. Structure – Revista Colombiana de estatística: 3(2), pp.261-271.
- [10] Fisher, R. A. & Tippett, L. H. C. (1928). “*Limiting forms of the frequency distribution of the largest or smallest member of a sample*”. Proceedings of the Cambridge Philosophical Society, vol 24, pp. 180–290.
- [11] Fonseca, M., Mathew, T., Mexia, J. T. & Zmyslony, R. (2007). *Tolerance intervals in the two-way nested model with mixed or random effects*. Statistics 41, 5876-5885.
- [12] Fonseca, M.; Mexia, J. T. & Zmyslony, R. (2006). Binary operations on Jordan Algebras and orthogonal normal models. *Linear Algebra and Its Applications*, 411, 75-86.
- [13] Fonseca, M., Mexia, J. T. & Zmyslony, R. (2003) – *Estimators and Tests for variance components in cross nested orthogonal models design*. *Discussiones Mathematicae Probability and Statistics* 23(2), 175-201.
- [14] GAUSS, Carl Friedrich (1777-1855). *Theoria motus corporum coelestium in sectionibus conicis solem ambientium*. Hamburg: Friedrich Perthes and I.H. Besser, 1821.
- [15] GAUSS, Carl Friedrich (1777-1855). *Theoria motus corporum coelestium in sectionibus conicis solem ambientium*. Hamburg: Friedrich Perthes and I.H. Besser, 1809.
- [16] Graham, A. (1981). *Kronecker products and Matrix Calculus with Applications*. Halsted Press, John Willey and Sons, New York.
- [17] HSU, P. L. (1936). *On the least unbiased estimate of variance -statist*. Res. Mem 2, 91-104.
- [18] Kariya, T. & Kurata, H. (2004). *Generalized Least Squares*, Willey, França.
- [19] Laplace, P.S. (1836). *Théore analytique des probabilités*. Supplement to Third Edition, Couvier, Paris.
- [20] Leon, S. J. *Álgebra Linear, com Aplicações*, 8ª edição, Brasil, 2010, pg.198-205;

- [21] Martins, J. P. O. (2008). *Feira dos Momentos, Planeamento Experimental e Investigação de Localização e escala em Populações não Gaussianas*. [Tese do Doutoramento, Universidade de Lisboa, Portugal].
- [22] McCullagh, P. & Nelder, J.A. (1995). *Generalized Linear Models*. Chapman and Hall, London.
- [23] McCullagh, P. & Nelder, J.A. (1989). *Generalized Linear Models*. 2nd edition, Chapman and Hall, London.
- [24] Mexia, J. T. (1989). *Controlled Heteroscedasticity*. Quotient Vector Spaccs and F Tests for Hypothesis on Mean Vectors – Trabalhos de Investigação, nº 1. FCT/UNL.
- [25] Mexia, J. T. (1989). *Simultaneous Confidence Intervals: Generalization of the Scheffé Theorem* – Trabalhos de Investigação, nº 2. Dep. Matemática – FCT/UNL.
- [26] Mexia, J. T. (1990). *Best Linear Unbiased Estimates*. Duality of F tests and the Scheffé Multiple Comparison Method in presence of Controlled Heterocedasticity – Comp. Stat & Data Analysis – Vol. 10, nº3.
- [27] Mexia, J. T. (1995). *Introdução à Inferência Estatística Linear*, Centro de Estudo de Matemática Aplicada, edições Universidades Lusófonas, pp.1-55.
- [28] Mukherjea, A., Rão, M., & Suen, S. (2006). *A note on moment generating functions*, *Stat Probabil Lett*, pp.1185-1189.
- [29] Nelder, J. A. (1965). The analysis of randomized experiments with orthogonal block structure. II. *Treatment structure and the general analysis of variance*. Poceedings of Royal Society (Edimburgh) Serie A. doi.or/10.1098/rspa.1965.0013.
- [30] Nelder, J. A. (1965). *The analysis of randamized experiments with orthogonal block structure and null analysis of variance* – proceedings of Royal Society (Edimburgh) Serice A; 273, pp.147-168.
- [31] Nelder, J.A. (1966). *Inverse polynomials, a useful group of multi-factor response functions*. *Biometrics*, 22, pp.128-141.

- [32] Nelder, J.A. & Wedderburn, R.W.M. (1972). *Generalized linear models*. Journal of the Royal Statistical Society Series A 135, 370 - 384.
- [33] Oliveira, S. (2018). *Delineamento regressional múltiplo para um fatorial de base prima estritamente associado a uma álgebra de Jordan comutativa*. PhD Thesis. FCT – Universidade Nova de Lisboa.
- [34] Paliga, A. (2012). *Vetores. Álgebra Linear e Geometria Analítica*, Universidade Federal de Pelotas, Brasil.
- [35] Pinheiro, E. C. (2014). *Contribuições em Inferência e Modelagem de Valores Extremos*. (Tese de doutoramento – Universidade de São Paulo, Brasil).
- [36] Pistone, G. & Wynn, H.P. (2006). *Symb Comput*, pp.210-221.
- [37] Pollock, D. S. G. (1979). *The Algebra of Econometrics*. John Wiley & Sons;
- [38] Santos, W. D. O. et al. (2014). *Precipitações máximas para o município de Mossoró de 1964 a 2011 pela distribuição Gumbel*. IRRIGA – Brazilian Journal of Irrigation and Drainage.
- [39] Scheffé, H. (1959). *The Analysis of Variance*. John Wiley & Sons, New York.
- [40] Schott, J. R. (2017). *Matrix Analysis of Statistics*; 3rd ed. Wiley – Interscience.
- [41] Seber, G. A. F. (1980). *The Linear Hypothesis: a general Theory*. 2nd ed. Charles Griffin & Co-Londres.
- [42] Seely, J. (1972). *Completeness for a family of multivariate normal distribution*. Ann. Math, Stat. 43:1644-1647.
- [43] Seely, J. (1971). *Quadratic Subspaces and Completeness*. The Annals of Mathematical Statistics 42, 2, 710-721.
- [44] Simões, C. R. (2017). *Estimação de funções densidades de probabilidade neutras face ao risco: uma aplicação a opções S&P500*. [Dissertação do mestrado, Universidade de Coimbra, Portugal].
- [45] Smith, P. J. (2012). *A Recursive Formulation of the Old Problem of Obtaining Moments from Cumulants and Vice-Versa*, Am Stat, pp.217-218.
- [46] Steeb, W. H. (1991). *Kroneck Products of Matrices and Applications*, Germany-Bibliographisches Institut.

- [47] Steeb, W. H.; Hardy, Y. (2011). *Matrix calculus and Kronecker Product: a practical approach to linear and multilinear álgebra*. Second Edition. World Scientific. Publishing Company.
- [48] Stuart, A. & Ord, K. (2000). *Kendall's Advanced Theory of Statistics vol. 1 - Distribution Theory*. 6th ed. Willey.
- [49] Taylor, A. E. (1995). *Advamoed Calulus*. Xerox.
- [50] Thiele, T. N. (1903). *Theory of observations, Layton* (republicado em *Ann. Math. Stat.*, 2, pp.165-308, 1931).
- [51] Turkman, M.A.A. e Silva, G.L. (2000). *Modelos Lineares Generalizados, da Teoria à Prática*. Universidade de Lisboa.
- [52] Watanabe, F. M. (2013). *Análise do Método de Gumbel para cálculo de vazões de dimensionamento de vertedouros*. Universidade de São Paulo.