

---

---

ANÁLISE DE INFLUÊNCIA BASEADA NA CURVATURA NORMAL  
CONFORME PARA UM MODELO DE REGRESSÃO DIRICHLET

LUZ MILENA ZEA FERNÁNDEZ

Orientador: Prof. Dr. Klaus Leite Pinto Vasconcellos  
Área de concentração: Estatística Matemática

Dissertação submetida como requerimento parcial para obtenção do grau  
de Mestre em Estatística pela Universidade Federal de Pernambuco

Recife/PE  
Dezembro de 2005

---

---

# **Livros Grátis**

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

---

ANÁLISE DE INFLUÊNCIA BASEADA NA CURVATURA NORMAL  
CONFORME PARA UM MODELO DE REGRESSÃO DIRICHLET

LUZ MILENA ZEA FERNÁNDEZ

Zea Fernández, Luz Milena

Análise de influência baseada na curvatura normal conforme para um modelo de regressão Dirichlet / Luz Milena Zea Fernández. – Recife : O Autor, 2005.

xviii, 95 folhas : il., fig., tab.

Dissertação (mestrado) – Universidade Federal de Pernambuco. CCEN. Estatística, 2005.

Inclui bibliografia e apêndices.

1. Estatística matemática – Análise de diagnóstico. 2. Modelos de regressão – Influência local – Modelo de Dirichlet. 3. Esquemas de perturbação – Variáveis. I. Título.

---

519.233.5  
519.535

CDU (2.ed.)  
CDD (22.ed.)

UFPE  
BC2005-652

---

Universidade Federal de Pernambuco  
Mestrado em Estatística

02 de dezembro de 2005

(data)

Nós recomendamos que a dissertação de mestrado de autoria de

Luz Milena Zea Fernández

intitulada

Análise de influência Baseada na Curvatura Normal Conforme para um

Modelo de Regressão Dirichlet

seja aceita como cumprimento parcial dos requerimentos para o grau de Mestre em Estatística.

Klaus Leite Pinto Vasconcelos

Coordenador da Pós-Graduação em Estatística

Banca Examinadora:

Klaus Leite P. Vasconcelos

Klaus Leite Pinto Vasconcelos

orientador

Valdério Anselmo Reisen

Valdério Anselmo Reisen (UFES)

Francisco José de Azevedo Cysneiros

Francisco José de Azevedo Cysneiros



Prof. Klaus L. P. Vasconcelos  
Coordenador da  
Pós-graduação em  
Estatística, UFPE

Este documento será anexado à versão final da dissertação.

*Aos meus pais, eles são a razão  
da minha vida e a força que me  
impulsiona a continuar cada dia.*

## Agradecimentos

A Deus por ter-me dado a vida, a capacidade para alcançar esta meta e sempre estar ao meu lado nesta caminhada que às vezes parecia muito difícil.

Aos meus pais e aos meus irmãos pelo amor, pela dedicação e por estar sempre torcendo em minha casa para que tudo desse certo.

Ao meu orientador Klaus Leite Pinto Vasconcellos, pela atenção, dedicação, boa orientação no decorrer deste trabalho e por ter acreditado em mim.

A Nicolás pelo amor, atenção, compressão, por toda ajuda, pelos nossos longos papos e por estar do meu lado sempre que precisei.

Ao professor Francisco Cribari Neto, pelos conhecimentos transmitidos, dedicação e eficiência que o caracterizam.

Aos professores do departamento de estatística da UFPE que colaboraram para meu crescimento profissional.

Aos meus grandes amigos que considero como meus irmãos Dany, Renata, Carlos e Tiago por todo o carinho, força, ajuda, amizade, companhia e atenção que me brindaram em minha estadia aqui no Brasil.

A todos os demais colegas do mestrado pela convivência e pelos momentos de descontração que vivemos juntos.

A todos meus colegas de Colômbia que moram aqui no Brasil, pelas nossas reuniões, as longas conversações e os momentos de diversão, especialmente a minha amiga Laura.

À Valéria Bittencourt, pela atenção que me deu desde minha chegada ao Brasil, pelo carinho e por ser muito competente em seu trabalho.

Aos participantes da banca examinadora, pelas sugestões.

Ao CNPq, pelo apoio financeiro.

*Se tens conhecimento, deixa que os  
demais acendam suas velas nele.*

*–Thomas Fuller*

## Resumo

Neste trabalho apresentamos um estudo de influência local no modelo de regressão Dirichlet proposto em Silva (2004) usando a curvatura normal conforme proposta por Poon & Poon (1999). Perturbamos o modelo segundo quatro esquemas de perturbação diferentes, a saber: a log-verossimilhança de forma multiplicativa, as variáveis explicativas de forma aditiva e multiplicativa e, finalmente, as variáveis resposta de forma aditiva. No desenvolvimento deste último esquema nos encontramos frente a um problema de maximização sujeito a restrições, resolvemos o problema e encontramos a solução. A partir daí conseguimos enunciar e provar um Teorema que nos dá uma forma de realizar análise de influência quando as perturbações satisfazem um conjunto de restrições lineares; além disso, estendemos o conceito de contribuição agregada para modelos multivariados. Apresentamos também um exemplo de análise de influência para dados reais usando o modelo de regressão Dirichlet. Neste exemplo, verificamos que uma observação se destaca como influente nos quatro esquemas de perturbação e que a influência das observações é maior para maiores valores das variáveis explicativas.



## Abstract

This work presents a local influence study for the Dirichlet regression model proposed in Silva (2004), by using the conformal normal curvature proposed in Poon & Poon (1999). Four different perturbation schemes are considered: the log-likelihood terms in the multiplicative form, the explanatory variables in additive and multiplicative forms and, finally, the response variables in additive form. In order to perform the last perturbation scheme, we solve a maximization problem with restrictions. We then obtain a theorem that gives us a form of performing influence analysis under linear constraints in the perturbation scheme. We also propose an influence analysis for multivariate data. An example with real data is presented for which we detect an observation as particularly influential and for which the influence of the observation is larger for larger values of the explanatory variables.

# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Introdução . . . . .	1
1.1.1	Organização da dissertação . . . . .	3
1.1.2	Suporte computacional . . . . .	3
1.2	Preliminares . . . . .	3
1.2.1	Estimadores de máxima verossimilhança . . . . .	4
1.3	Otimização não-linear em inferência estatística . . . . .	5
<b>2</b>	<b>O modelo de regressão Dirichlet</b>	<b>6</b>
2.1	Modelo beta . . . . .	7
2.1.1	Definição e estimação . . . . .	7
2.2	Modelo Dirichlet . . . . .	11
2.2.1	Definição e estimação . . . . .	12
<b>3</b>	<b>Curvatura normal conforme e influência local</b>	<b>18</b>
3.1	Influência local . . . . .	20
3.2	Curvatura normal conforme . . . . .	21

<i>SUMÁRIO</i>	vii
3.3 Avaliação da influência local . . . . .	22
3.3.1 Influência de autovetor individual . . . . .	23
3.3.2 Contribuição agregada de vetores de perturbação básicos . . . . .	23
3.4 Esquemas de perturbação . . . . .	25
<b>4 Aplicações</b>	<b>38</b>
4.1 Introdução . . . . .	38
4.2 Composição de sedimentos . . . . .	40
4.2.1 Esquema de perturbação I . . . . .	40
4.2.2 Esquema de perturbação II . . . . .	46
4.2.3 Esquema de perturbação III . . . . .	52
4.2.4 Esquema de perturbação IV . . . . .	58
<b>5 Conclusões</b>	<b>64</b>
<b>Apêndice</b>	<b>65</b>
<b>A Derivadas</b>	<b>66</b>
A.1 Esquema de perturbação I . . . . .	67
A.2 Esquema de perturbação II . . . . .	68
A.3 Esquema de perturbação III . . . . .	73
A.4 Esquema de perturbação IV . . . . .	80
<b>B Dados</b>	<b>83</b>
<b>C Programa</b>	<b>85</b>

*SUMÁRIO*

viii

**Refêrencias**

**91**

# Capítulo 1

## Introdução

### 1.1 Introdução

Os modelos de regressão dão aos pesquisadores ferramentas poderosas que permitem estudar e modelar relações entre variáveis, e fazer previsões de eventos do passado, do presente ou do futuro usando a informação do passado ou presente. Por isto, são largamente utilizados em muitas áreas do conhecimento. Durante muitos anos, os modelos normais lineares simples e multivariados foram utilizados para descrever a maioria dos fenômenos aleatórios, mesmo quando o fenômeno sob estudo apresentava uma resposta para a qual não fosse razoável a suposição de normalidade ou que estivesse restrita ao intervalo  $(0,1)$ . O modelo de regressão Dirichlet, recentemente proposto em Silva (2004), oferece grande utilidade quando as variáveis resposta são proporções, pois, diferentemente do modelo de regressão linear, permite assimetrias, o que é muito comum quando se trabalha com proporções. Existe uma grande quantidade de trabalhos teóricos que envolvem a distribuição Dirichlet, porém, são poucos os trabalhos que têm sido feitos em aplicações práticas. Esta falha pode ser devida ao fato de que esta distribuição não é familiar para muitos cientistas.

Um tópico importante na análise de diagnóstico em regressão é a detecção de observações influentes, ou seja, de pontos que podem exercer forte influência nas estimativas dos parâmetros do modelo. Uma das idéias mais inovadoras foi apresentada

por Cook (1986), que propõe verificar a existência de pontos que sob modificações modestas causam variações desproporcionais nos resultados, ao invés da avaliação pela retirada individual ou conjunta de pontos. Este método se denomina influência local e teve uma grande receptividade entre os usuários e pesquisadores da área de regressão, motivando muitos artigos onde se aplica a metodologia em modelos particulares ou se propõe extensões da técnica, como será visto no Capítulo 3. Apesar do método de Cook, baseado na curvatura normal, ser de grande utilidade, o método tem alguns inconvenientes; por exemplo, a curvatura normal pode tomar qualquer valor e não é invariante sob uma mudança uniforme de escala. Recentemente, Poon & Poon (1999) propuseram uma medida chamada “curvatura normal conforme”, e, sobre esta medida definiram a contribuição agregada para cada vetor de perturbação básico, fornecendo um ponto de partida objetivo para julgar a grandeza da curvatura.

Nesta dissertação apresentamos a análise de influência local utilizando a curvatura normal conforme para o modelo de regressão Dirichlet. Estudamos quatro esquemas de perturbação, onde perturbamos a log-verossimilhança multiplicativamente, as variáveis explicativas de forma aditiva e multiplicativa e por último perturbamos as variáveis resposta de forma aditiva. No desenvolvimento deste último esquema de perturbação nos encontramos com um problema de maximização com restrições, resolvemos o problema e encontramos sua solução analítica. Neste sentido enunciamos um Teorema e apresentamos a demonstração. A partir deste Teorema, temos como primeiro resultado desta dissertação uma proposta geral para análise de influência quando o espaço de perturbação é sujeito a um conjunto de restrições lineares homogêneas. O segundo resultado que apresentamos é uma proposta de análise de influência de observações para modelos multivariados. Por último, analisamos um conjunto de dados reais com o modelo de regressão Dirichlet, para o qual encontramos uma observação que se destaca das outras em todos os esquemas de perturbação, além de verificarmos que as observações correspondentes a valores maiores das variáveis explicativas são as mais influentes.

### 1.1.1 Organização da dissertação

A estrutura do restante desta dissertação está organizada da seguinte forma. No Capítulo 2 estão apresentadas as distribuições beta e Dirichlet, assim como algumas propriedades destas distribuições. No Capítulo 3 está apresentada a construção da curvatura normal conforme, a contribuição agregada individual dos vetores de perturbação básicos, o Teorema que enunciamos e demonstramos, e, por último a nossa proposta de análise de influência para modelos multivariados. No Capítulo 4 apresentamos a análise de influência aplicada ao modelo de regressão Dirichlet para um conjunto de dados reais tomado de Aitchison (2003). As conclusões desta dissertação são apresentadas no Capítulo 5. Finalmente, no Apêndice A estão apresentados os cálculos das derivadas necessárias para fazer a análise de influência segundo cada esquema de perturbação, no Apêndice B está apresentado o conjunto de dados reais estudado e o Apêndice C apresenta o programa utilizado para fazer a análise de influência no modelo de regressão Dirichlet.

### 1.1.2 Suporte computacional

As ferramentas computacionais utilizadas no desenvolvimento desta dissertação foram a linguagem de programação matricial Ox em sua versão 3.40 e o programa R em sua versão 2.0.1. Ox está disponível gratuitamente para uso acadêmico no site <http://www.nuff.ox.ac.uk/Users/Doornik> e possui uma abrangente biblioteca de funções estatísticas e matemáticas. Os gráficos apresentados nesta dissertação foram feitos utilizando a ferramenta computacional R.

## 1.2 Preliminares

Um dos principais problemas da Inferência Estatística é a estimação de parâmetros, pois, na Inferência Paramétrica, para inferir sobre o comportamento de uma população a partir do conhecimento de uma amostra, é necessário encontrar estimadores dos parâmetros. Mais especificamente, dado um vetor aleatório  $(Y_1, \dots, Y_N)$ , com dis-

tribuição conjunta conhecida e caracterizada por uma função de probabilidade ou de densidade conjunta  $f(\mathcal{Y}; \theta)$ , dependente de um vetor de parâmetros desconhecidos  $\theta$ ,  $\theta \in \Theta \subset \mathbb{R}^p$ , a tarefa de fazer inferência inclui estimar o valor deste vetor de parâmetros. Para este fim, na literatura estatística existem diversos métodos que podem ser usados, uns dos métodos mais aplicados sendo a estimação por máxima verossimilhança. Os estimadores de máxima verossimilhança podem ser calculados através da solução de um problema de otimização; em alguns casos este problema pode ser transformado em um problema de solução de um sistema de equações, porém, é muito comum que os estimadores de máxima verossimilhança não possuam forma fechada. Assim, as estimativas devem ser obtidas a partir da maximização numérica da função de verossimilhança ou da função de log-verossimilhança.

### 1.2.1 Estimadores de máxima verossimilhança

O método de máxima verossimilhança objetiva escolher os valores dos parâmetros que dão a chance mais provável de que novamente ocorram os fatos que ocorreram. A função de verossimilhança para  $\theta$ , baseada na observação  $\mathcal{Y} = (y_1, \dots, y_N)$ , é definida como

$$L(\theta) \equiv L(\theta; \mathcal{Y}) = f(\mathcal{Y}, \theta), \quad \theta \in \Theta,$$

sendo interpretada como função de  $\theta$  para  $\mathcal{Y}$  fixo. Dada a função  $L(\theta)$ , o estimador de máxima verossimilhança  $\hat{\theta}$  é o valor que maximiza a função de verossimilhança  $L(\theta)$ . Formalmente, o estimador de máxima verossimilhança é definido como a variável aleatória

$$\hat{\theta} = \underset{\theta \in \Theta}{\operatorname{argmax}} L(\theta).$$

Geralmente o processo de maximização é melhor conduzido trabalhando-se com o logaritmo natural da função de verossimilhança, denotado por  $l(\theta) = \log(L(\theta))$ . Como a função logaritmo é monótona crescente, maximizar  $L(\theta)$  e  $l(\theta)$  em  $\Theta$  são processos equivalentes. Quando as observações são independentes segue-se que  $l(\theta)$  é aditiva, ou seja,  $l(\theta) = \sum_{i=1}^N \log f_i(y_i; \theta)$ , onde  $f_i$  representa a densidade individual da observação  $i$ .



### 1.3 Otimização não-linear em inferência estatística

Para encontrar as estimativas de máxima verossimilhança em modelos estatísticos é preciso fazer maximização de uma função específica. Quando o estimador de máxima verossimilhança não possui forma fechada, as estimativas devem ser obtidas a partir da maximização numérica da função de verossimilhança ou da função de log-verossimilhança. Para este propósito, algoritmos iterativos podem ser empregados.

A literatura divide as técnicas de maximização em classes, sendo as mais utilizadas a *classe de métodos gradiente* e a *classe de algoritmos quasi-Newton*. Na primeira classe se faz uso da matriz hessiana e os métodos mais conhecidos são steepest ascent, Newton-Raphson e escore de Fischer. A segunda classe elimina a necessidade do cálculo de segundas derivadas, pois usa uma aproximação da matriz hessiana construída de forma iterativa. Entre os algoritmos que pertencem a esta classe tem-se o algoritmo DFP (Davidon, Fletcher e Powell) e o algoritmo BFGS (Broyden, Fletcher, Goldfarb e Shanno). No decorrer deste trabalho usa-se a técnica de maximização BFGS.

## Capítulo 2

# O modelo de regressão Dirichlet

Os modelos de regressão dão aos pesquisadores ferramentas poderosas que permitem estudar e modelar relações entre variáveis, e fazer previsões de eventos do passado, do presente ou do futuro usando a informação do passado ou presente. Por isto, são largamente utilizados em muitas áreas do conhecimento, tais como: saúde, biologia, agronomia, computação, administração, engenharias, sociologia, etc. Durante muitos anos, os modelos normais lineares foram utilizados para descrever a maioria dos fenômenos aleatórios, mesmo quando o fenômeno sob estudo apresentava uma resposta para a qual não fosse razoável a suposição de normalidade ou que estivesse restrita ao intervalo  $(0,1)$ . No primeiro caso tentava-se algum tipo de transformação no sentido de alcançar a normalidade procurada, e, no segundo, uma possibilidade era transformar a variável resposta para que assumisse valores na reta real e então modelar a média da resposta transformada com um preditor linear baseado em um conjunto de variáveis explicativas. Porém, sob esta modelagem os parâmetros do modelo nem sempre podem ser facilmente interpretados em termos da resposta original; além disso, as proporções tipicamente apresentam assimetria. Assim, inferências baseadas na suposição de normalidade podem estar erradas.

Uma proposta muito interessante e inovadora no assunto foi apresentada por Nelder e Wedderburn (1972), que propuseram os modelos lineares generalizados (MLGs). A idéia básica consiste em abrir o leque de opções para a distribuição da variável resposta, permitindo que a mesma pertença à família exponencial de dis-

tribuições, bem como dar maior flexibilidade para a relação funcional entre a média da variável resposta e o preditor linear. Uma outra possibilidade são os modelos aditivos generalizados, veja Hastie & Tibshirani (1990).

## 2.1 Modelo beta

A distribuição beta é uma das mais usadas para modelar experimentos aleatórios que produzem resultados no intervalo  $(0, 1)$  devido à grande flexibilidade de ajuste de seus parâmetros. Ferrari & Cribari-Neto (2004) usam uma parametrização definida pela média e por um parâmetro de dispersão para a modelagem de taxas e proporções. Achcar & Netto (2003) estudam a prevalência da tuberculose usando métodos bayesianos onde a distribuição beta é tomada como distribuição a priori. Bury (1999) lista um conjunto de aplicações da distribuição beta em engenharia. Janardan & Padmanabhan (1986) modelam variáveis hidrológicas usando a distribuição beta. Mc-Nally (1990) utiliza a distribuição beta no estudo de algumas variáveis que afetam a reprodutibilidade de vacas. Graham & Hollands (1990) e Milyutin & Yaromenko (1991) usam a distribuição beta para estudar índices relacionados à transmissão de radiação solar. Wiley, Herschokoru & Padiou (1989) desenvolvem um modelo beta para estimar a probabilidade de transmissão de HIV durante o contato sexual entre um indivíduo infectado e um indivíduo sadio. A seguir, consideram-se algumas propriedades da distribuição beta, como sua função de densidade e seus primeiros momentos.

### 2.1.1 Definição e estimação

Seja  $Y$  uma variável aleatória contínua. Diz-se que  $Y$  tem distribuição beta com parâmetros  $\alpha_1, \alpha_2 > 0$ , denotado por  $Y \sim \beta(\alpha_1, \alpha_2)$ , se sua função de densidade é da seguinte forma

$$f(y; \alpha_1, \alpha_2) = \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} y^{\alpha_1-1} (1-y)^{\alpha_2-1} \mathbf{I}_{(0,1)}(y), \quad (2.1)$$

onde  $\Gamma(\cdot)$  é a função gama, definida como  $\Gamma(t) = \int_0^\infty y^{t-1}e^{-y}dy$  e  $\mathbf{I}_A(\cdot)$  representa a função indicadora, que para um conjunto  $A \subset \mathbb{R}$  corresponde a

$$\mathbf{I}_A(x) = \begin{cases} 1 & \text{se } x \in A, \\ 0 & \text{se } x \notin A. \end{cases}$$

Note inicialmente que  $f(y) \geq 0$ , pois é o produto de funções não-negativas; além disso,  $\int_{-\infty}^\infty f(y)dy = 1$ . De fato,

$$\begin{aligned} \int_{-\infty}^\infty f(y)dy &= \int_0^1 \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} y^{\alpha_1-1}(1-y)^{\alpha_2-1} dy, \\ &= \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \int_0^1 y^{\alpha_1-1}(1-y)^{\alpha_2-1} dy. \end{aligned} \quad (2.2)$$

Usando a função Beta definida por

$$B(a, b) = \frac{1}{c^{a+b-1}} \int_0^c y^{a-1}(c-y)^{b-1} dy, \quad \forall c > 0 \quad (2.3)$$

e a seguinte propriedade desta função:

$$B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}, \quad (2.4)$$

tem-se que para  $c = 1, a = \alpha_1$  e  $b = \alpha_2$ ,

$$\int_0^1 y^{\alpha_1-1}(1-y)^{\alpha_2-1} dy = \frac{\Gamma(\alpha_1)\Gamma(\alpha_2)}{\Gamma(\alpha_1 + \alpha_2)},$$

portanto, temos na equação (2.2) que  $\int_{-\infty}^\infty f(y)dy = 1$ . Logo, a função  $f(\cdot)$  é efetivamente uma função de densidade.

O momento de ordem  $n$  em torno de zero da distribuição  $\beta(\alpha_1, \alpha_2)$  é obtido como

segue:

$$\begin{aligned}
E(Y^n) &= \int_0^1 y^n f(y) dy, \\
&= \int_0^1 y^n \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} y^{\alpha_1-1} (1-y)^{\alpha_2-1} dy, \\
&= \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \int_0^1 y^{n+\alpha_1-1} (1-y)^{\alpha_2-1} dy, \\
&= \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} B(n + \alpha_1, \alpha_2), \\
&= \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \frac{\Gamma(n + \alpha_1)\Gamma(\alpha_2)}{\Gamma(n + \alpha_1 + \alpha_2)}, \\
&= \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)} \frac{\Gamma(n + \alpha_1)}{\Gamma(n + \alpha_1 + \alpha_2)}.
\end{aligned}$$

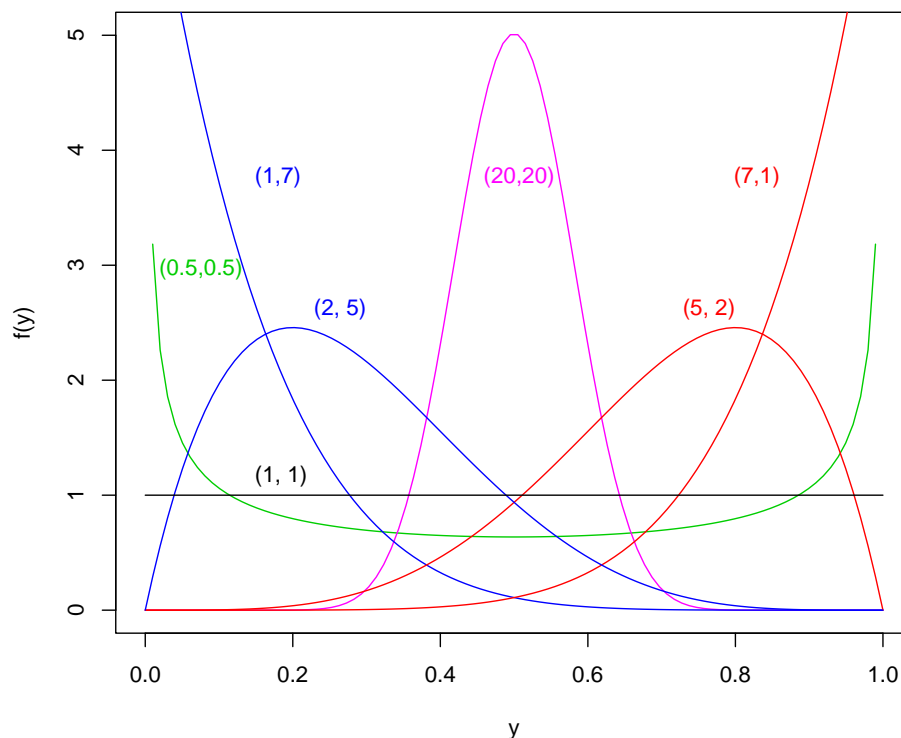
Esta expressão pode ser simplificada usando a propriedade da função gama que afirma que  $\Gamma(z + 1) = z\Gamma(z)$ ,  $\forall z > 0$ . Assim,

$$E(Y^n) = \frac{\alpha_1(\alpha_1 + 1)(\alpha_1 + 2) \cdots (\alpha_1 + n - 1)}{(\alpha_1 + \alpha_2)(\alpha_1 + \alpha_2 + 1)(\alpha_1 + \alpha_2 + 2) \cdots (\alpha_1 + \alpha_2 + n - 1)}.$$

Daí, pode-se encontrar que a média e a variância da variável  $Y$  com distribuição  $\beta(\alpha_1, \alpha_2)$  são respectivamente

$$\begin{aligned}
E(Y) &= \frac{\alpha_1}{\alpha_1 + \alpha_2}, \\
V(Y) &= \frac{\alpha_1\alpha_2}{(\alpha_1 + \alpha_2)^2(\alpha_1 + \alpha_2 + 1)}.
\end{aligned}$$

Os parâmetros  $\alpha_1$  e  $\alpha_2$  são parâmetros de ajuste, pois através dos diferentes valores dados para  $\alpha_1$  e  $\alpha_2$  podem ser obtidas diferentes distribuições no intervalo  $(0, 1)$ . Quando  $\alpha_1 = \alpha_2 = 1/2$  a distribuição beta é conhecida como “distribuição arco seno”, a qual é de grande utilidade no estudo de passeios aleatórios. No caso em que  $\alpha_1 + \alpha_2 = 1$  com  $\alpha_1 \neq 1/2$  tem-se as distribuições “arco seno generalizadas”. As densidades simétricas em torno de  $1/2$  correspondem a  $\alpha_1 = \alpha_2$  e quando  $\alpha_1 = \alpha_2 = 1$  tem-se a conhecida distribuição uniforme no intervalo  $(0, 1)$ . Assim, a distribuição beta é na realidade uma família de distribuições, isto pode ser observado na Figura 2.1.

**Figura 2.1.1:** Gráfico da densidade beta para diferentes valores de  $\alpha_1$  e  $\alpha_2$ .

Note-se que à medida que os parâmetros  $\alpha_1$  e  $\alpha_2$  aumentam, a variabilidade da densidade diminui; este fato é confirmado pela expressão da variância.

Para encontrar os estimadores de máxima verossimilhança dos parâmetros  $\alpha_1$  e  $\alpha_2$  é necessário encontrar a função de verossimilhança e a seguir encontrar os valores dos parâmetros que maximizam esta função. Para isto se procede da seguinte maneira.

Sejam  $Y_1, \dots, Y_N$ , variáveis aleatórias independentes, onde  $Y_i \sim \beta(\alpha_1, \alpha_2)$ , para  $i = 1, \dots, N$ , e sejam  $y_1, \dots, y_N$  os valores observados destas variáveis. Então a função de verossimilhança é dada pela seguinte expressão

$$L(\alpha_1, \alpha_2) = \prod_{i=1}^N f(y_i; \alpha_1, \alpha_2) = \prod_{i=1}^N \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} y_i^{\alpha_1-1} (1 - y_i)^{\alpha_2-1}.$$

Logo a função de log-verossimilhança é da forma

$$\begin{aligned}\ell(\alpha_1, \alpha_2) &= \log(L(\alpha_1, \alpha_2)) = \sum_{i=1}^N \log \left\{ \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} y_i^{\alpha_1-1} (1 - y_i)^{\alpha_2-1} \right\} \\ &= \sum_{i=1}^N \left\{ \log \left( \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \right) + (\alpha_1 - 1) \log y_i + (\alpha_2 - 1) \log(1 - y_i) \right\} \\ &= N \log \left( \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \right) + (\alpha_1 - 1) \sum_{i=1}^N \log y_i + (\alpha_2 - 1) \sum_{i=1}^N \log(1 - y_i).\end{aligned}$$

Dessa forma,

$$\ell(\alpha_1, \alpha_2) = N \log \left( \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \right) + (\alpha_1 - 1)g_1 + (\alpha_2 - 1)g_2$$

onde  $g_1 = \sum_{i=1}^N \log y_i$  e  $g_2 = \sum_{i=1}^N \log(1 - y_i)$ . Os estimadores de máxima verossimilhança  $\hat{\alpha}_1, \hat{\alpha}_2$  são obtidos pela solução do sistema não linear

$$N(\psi(\alpha_1) - \psi(\alpha_1 + \alpha_2)) = g_1,$$

$$N(\psi(\alpha_2) - \psi(\alpha_1 + \alpha_2)) = g_2,$$

onde  $\psi(\cdot)$  é a função digamma, ou seja,  $\psi(z) = \frac{d \log \Gamma(z)}{dz}$  para  $z > 0$ .

## 2.2 Modelo Dirichlet

A distribuição Dirichlet é uma generalização multivariada da distribuição beta que oferece muita flexibilidade e facilidade de uso. Em contraste com outras distribuições, a distribuição Dirichlet permite múltiplas assimetrias e simetrias; de fato, pode ser viesada à direita, viesada à esquerda ou simétrica. Existe uma grande quantidade de trabalhos teóricos que envolvem a distribuição Dirichlet, porém, são poucos os trabalhos que têm sido feitos em aplicações práticas. Esta falha pode ser devida ao fato de que esta distribuição não é familiar para muitos cientistas. Entretanto, Bouguila, Ziou

& Vaillancourt (2004) propõem um modelo para estimar os parâmetros de misturas de Dirichlet e testam o método para diferentes aplicações tais como estimação de histogramas, descrição sucinta de dados de imagem via retenção eficiente de informação e detecção de pele humana em conjuntos de dados de multimídia. Bouguila & Ziou (2004) apresentam um modelo misto robusto de probabilidade baseado em uma generalização da distribuição Dirichlet, neste caso, a proposta para estudar os parâmetros de um modelo Dirichlet misto é baseada em métodos de máxima verossimilhança e Fisher score e os resultados experimentais envolvem modelamento da cor da pele humana e aplicações à detecção da pele em imagens. Esta distribuição é também de grande utilidade na estatística bayesiana. Merwe & Pretorius (2003) apresentam um estudo de estimação bayesiana na reprodução de animais usando um processo de Dirichlet a priori para os efeitos aleatórios correlacionados. Boys, Henderson & Wilkinson (2000) apresentam uma solução bayesiana ao problema da detecção de segmentos homogêneos de DNA usando como informação a priori a distribuição Dirichlet. Mazzuchi & Soyer (1993) apresentam uma proposta completa baseada em uma versão bayesiana do modelo de crescimento-confiabilidade de Barlow-Scheuer para analisar a confiabilidade de produtos durante o desenvolvimento da fase, a distribuição Dirichlet sendo usada como distribuição a priori para uma transformação das probabilidades de fracasso.

### 2.2.1 Definição e estimação

**Definição 1.** *Sejam  $Y_1, \dots, Y_p$  variáveis aleatórias. Diz-se que o vetor aleatório  $(Y_1, \dots, Y_{p-1})$ , com  $Y_1 + \dots + Y_p = 1$ , tem distribuição Dirichlet com parâmetros  $\alpha_1 > 0, \dots, \alpha_p > 0$ , se sua função de densidade é da forma*

$$f(y_1, \dots, y_{p-1}; \alpha_1, \dots, \alpha_p) = \begin{cases} \frac{\Gamma(\sum_{j=1}^p \alpha_j)}{\prod_{j=1}^p \Gamma(\alpha_j)} \prod_{j=1}^{p-1} y_j^{\alpha_j-1} (1 - \sum_{j=1}^{p-1} y_j)^{\alpha_p-1} & \text{se } \underline{y} \in S, \\ 0 & \text{se } \underline{y} \notin S, \end{cases} \quad (2.5)$$

onde  $\underline{y} = (y_1, \dots, y_{p-1})$  e  $S$  é a região dada por

$$S = \left\{ (y_1, \dots, y_{p-1}) \in \mathbb{R}^{p-1}; y_j > 0 \text{ e } \sum_{j=1}^{p-1} y_j < 1 \right\}.$$



Note inicialmente que  $f(y_1, \dots, y_{p-1}) \geq 0$ , pois é o produto de funções positivas em  $S$  e  $(1 - \sum_{j=1}^{p-1} y_j) > 0$ . Além disso,

$$\int \cdots \int_{\mathbb{R}^{p-1}} f(y_1, \dots, y_{p-1}; \alpha_1, \dots, \alpha_p) dy_1 \dots dy_{p-1} = 1.$$

Por simplicidade nas operações mostra-se para o caso em que  $p = 3$ ; a prova para o caso geral é uma simples generalização.

Então, se quer provar que

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(y_1, y_2; \alpha_1, \alpha_2, \alpha_3) dy_1 dy_2 = 1,$$

Tem-se que

$$\begin{aligned} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(y_1, y_2; \alpha_1, \alpha_2, \alpha_3) dy_1 dy_2 &= \frac{\Gamma(\sum_{j=1}^3 \alpha_j)}{\prod_{j=1}^3 \Gamma(\alpha_j)} \int_0^1 \int_0^{1-y_2} \prod_{j=1}^2 y_j^{\alpha_j-1} \\ &\quad (1 - \sum_{j=1}^2 y_j)^{\alpha_3-1} dy_1 dy_2 \\ &= \frac{\Gamma(\alpha_1 + \alpha_2 + \alpha_3)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)} \int_0^1 \int_0^{1-y_2} y_1^{\alpha_1-1} y_2^{\alpha_2-1} (1 - y_1 - y_2)^{\alpha_3-1} dy_1 dy_2 \\ &= \frac{\Gamma(\alpha_1 + \alpha_2 + \alpha_3)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)} \int_0^1 \left[ \int_0^{1-y_2} y_1^{\alpha_1-1} (1 - y_1 - y_2)^{\alpha_3-1} dy_1 \right] y_2^{\alpha_2-1} dy_2. \end{aligned} \quad (2.6)$$

Lembrando a definição da função beta  $B(a, b)$  em (2.3) e considerando  $a = \alpha_1, b = \alpha_3$  e  $c = 1 - y_2$ , tem-se que

$$\int_0^{1-y_2} y_1^{\alpha_1-1} (1 - y_1 - y_2)^{\alpha_3-1} dy_1 = B(\alpha_1, \alpha_3) (1 - y_2)^{(\alpha_1 + \alpha_3 - 1)}.$$

Substituindo este resultado, (2.6) pode ser escrita como

$$\frac{\Gamma(\alpha_1 + \alpha_2 + \alpha_3)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)} B(\alpha_1, \alpha_3) \int_0^1 y_2^{\alpha_2-1} (1 - y_2)^{(\alpha_1 + \alpha_3 - 1)} dy_2$$

e recorrendo novamente à função beta  $B(a, b)$  com  $a = \alpha_2, b = \alpha_1 + \alpha_3$  e  $c = 1$  tem-se que

$$\int_0^1 y_2^{\alpha_2-1} (1 - y_2)^{(\alpha_1 + \alpha_3 - 1)} dy_2 = B(\alpha_2, \alpha_1 + \alpha_3).$$

Portanto,

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(y_1, y_2; \alpha_1, \alpha_2, \alpha_3) dy_1 dy_2 = \frac{\Gamma(\alpha_1 + \alpha_2 + \alpha_3)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)} B(\alpha_1, \alpha_3) B(\alpha_2, \alpha_1 + \alpha_3).$$

Aplicando a propriedade da função beta em (2.4)

$$\begin{aligned} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(y_1, y_2; \alpha_1, \alpha_2, \alpha_3) dy_1 dy_2 &= \frac{\Gamma(\alpha_1 + \alpha_2 + \alpha_3)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)} \frac{\Gamma(\alpha_1)\Gamma(\alpha_3)}{\Gamma(\alpha_1 + \alpha_3)} \frac{\Gamma(\alpha_2)\Gamma(\alpha_1 + \alpha_3)}{\Gamma(\alpha_1 + \alpha_2 + \alpha_3)}, \\ &= \frac{\Gamma(\alpha_1 + \alpha_2 + \alpha_3)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)} \frac{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)}{\Gamma(\alpha_1 + \alpha_2 + \alpha_3)} = 1. \end{aligned}$$

□

Portanto, pode-se concluir que a função dada em (2.5) é efetivamente uma função de densidade.

A seguir serão apresentadas algumas propriedades da Distribuição Dirichlet, como as médias, as variâncias e covariâncias. Começando pelo momento de ordem  $n$  em torno de zero na distribuição Dirichlet, tem-se que

$$E(Y_j^n) = \frac{\Gamma(\alpha_1 + \dots + \alpha_p)\Gamma(\alpha_j + 1)}{\Gamma(\alpha_1 + \dots + \alpha_p + n)\Gamma(\alpha_j)}, \text{ com } j = 1, \dots, p. \quad (2.7)$$

Daí, o valor esperado de  $Y_j$ , com  $j = 1, \dots, p$  é

$$E(Y_j) = \frac{\alpha_j}{\sum_{t=1}^p \alpha_t},$$

a variância de  $Y_j$ , com  $j = 1, \dots, p$ , corresponde a

$$\begin{aligned} V(Y_j) &= E(Y_j^2) - E(Y_j)^2 \\ &= \frac{\alpha_j(\alpha_j + 1)}{(\sum_{t=1}^p \alpha_t)(\sum_{t=1}^p \alpha_t + 1)} - \frac{\alpha_j^2}{(\sum_{t=1}^p \alpha_t)^2} \\ &= \frac{\alpha_j(\sum_{t \neq j}^p \alpha_t)}{(\sum_{t=1}^p \alpha_t)^2(\sum_{t=1}^p \alpha_t + 1)} \end{aligned}$$

e a covariância entre  $Y_j$  e  $Y_{j^*}$ , com  $j \neq j^*$ ,  $j, j^* = 1, \dots, p$  é

$$Cov(Y_j, Y_{j^*}) = -\frac{\alpha_j \alpha_{j^*}}{(\sum_{t=1}^p \alpha_t)^2(\sum_{t=1}^p \alpha_t + 1)}.$$

Sejam  $(Y_{i1}, \dots, Y_{ip})$ ,  $i = 1, \dots, N$ ,  $N$  vetores aleatórios independentes com distribuição Dirichlet e parâmetros  $\alpha_{i1}, \dots, \alpha_{ip}$ , com função de densidade dada por (2.5). Então, a log-densidade para cada vetor aleatório é dada pela seguinte expressão

$$\ell_i(\alpha_{i1}, \dots, \alpha_{ip}) = \log \Gamma \left( \sum_{j=1}^p \alpha_{ij} \right) - \sum_{j=1}^p \log \Gamma(\alpha_{ij}) + \sum_{j=1}^p (\alpha_{ij} - 1) \log y_{ij}; \quad i = 1, \dots, N. \quad (2.8)$$

Assim, a função de log-verossimilhança é dada por

$$\ell(\underline{\alpha}_1, \dots, \underline{\alpha}_p) = \sum_{i=1}^N \ell_i(\alpha_{i1}, \dots, \alpha_{ip}),$$

onde  $\underline{\alpha}_1 = (\alpha_{11}, \dots, \alpha_{N1})^\top, \dots, \underline{\alpha}_p = (\alpha_{1p}, \dots, \alpha_{Np})^\top$ , ou seja,

$$\ell(\underline{\alpha}_1, \dots, \underline{\alpha}_p) = \sum_{i=1}^N \left\{ \log \Gamma \left( \sum_{j=1}^p \alpha_{ij} \right) - \sum_{j=1}^p \log \Gamma(\alpha_{ij}) + \sum_{j=1}^p (\alpha_{ij} - 1) \log y_{ij} \right\}. \quad (2.9)$$

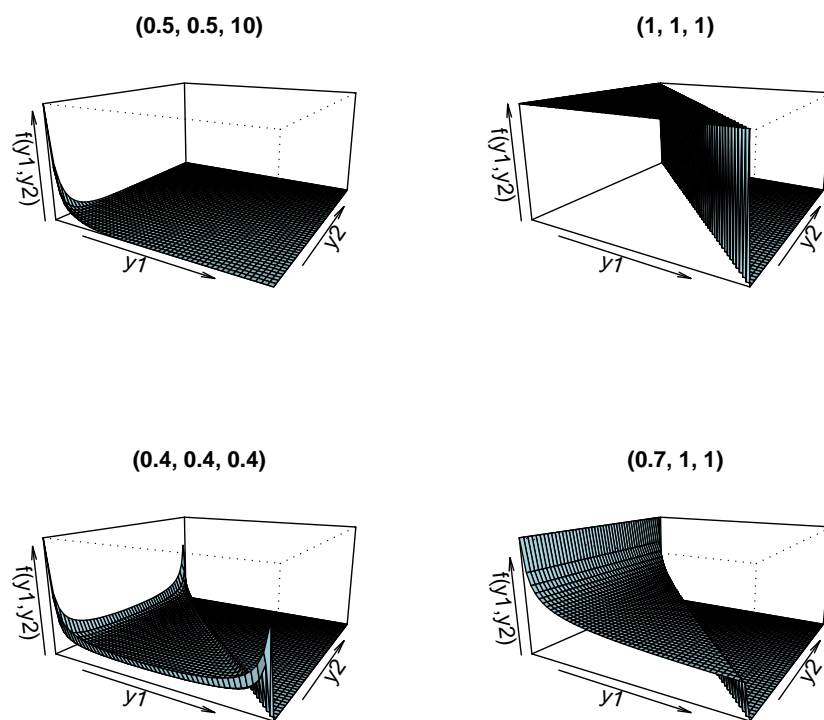
Para obter uma estrutura de regressão que permita modelar variáveis aleatórias regidas pela distribuição Dirichlet, considera-se  $\alpha_{ij} = g(\eta_{ij})$ , onde  $\eta_{ij} = x_{i1}\beta_{1j} + \dots + x_{ik}\beta_{kj}$ ;  $x_{i1}, \dots, x_{ik}$  são  $k$  covariáveis,  $\beta_{1j}, \dots, \beta_{kj}$  são os parâmetros da regressão  $j = 1, \dots, p$  e  $g(\cdot)$  é uma função de ligação conhecida que assume valores reais em  $(0, \infty)$ , estritamente monotônica e duas vezes diferenciável; neste trabalho considera-se  $g(\eta_{ij}) = e^{\eta_{ij}}$ . Ou seja, a média  $\mu_{ij}$  de cada  $Y_{ij}$  pode ser escrita como

$$\mu_{ij} = \frac{g(\eta_{ij})}{g(\eta_{i1}) + \dots + g(\eta_{ip})}.$$

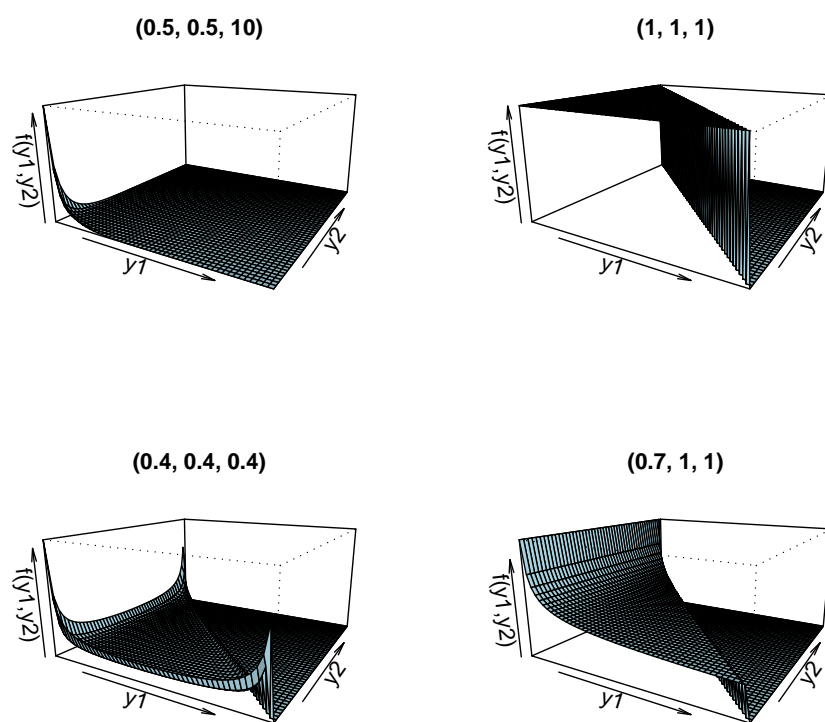
A parametrização do modelo de regressão Dirichlet apresentada aqui corresponde ao modelo 2 proposto em Silva (2004), onde também é estudada outra parametrização em que se modela a média das variáveis resposta; este modelo é chamado naquele trabalho modelo 1. Nesta dissertação não consideramos o modelo 1 dado que pela construção daquele modelo se usa a média da variável  $Y_1$  para modelar a média da variável  $Y_2$  e assim por diante, criando uma hierarquia entre as componentes.

Nas Figuras 2.2.1 e 2.2.2 estão apresentados os gráficos das densidades Dirichlet quando  $p$  é três, para diferentes valores dos parâmetros  $\alpha_1, \alpha_2$  e  $\alpha_3$ . Note-se que à medida que os parâmetros aumentam, a variabilidade da densidade diminui; este fato pode ser verificado pela expressão da variância.

**Figura 2.2.1:** Gráficos das densidades Dirichlet quando  $p = 3$  para diferentes valores de  $\alpha_1, \alpha_2$  e  $\alpha_3$ .



**Figura 2.2.2:** Gráficos das densidades Dirichlet quando  $p = 3$  para diferentes valores de  $\alpha_1, \alpha_2$  e  $\alpha_3$ .



## Capítulo 3

# Curvatura normal conforme e influência local

Um tópico importante na análise de diagnóstico em regressão é a detecção de observações influentes, ou seja, de pontos que podem exercer forte influência nas estimativas dos coeficientes da regressão. Durante a década de 70 surgiram várias propostas relacionadas com a influência das observações nas estimativas dos coeficientes do modelo normal linear. A eliminação de pontos talvez seja a técnica mais conhecida para avaliar o impacto de uma observação particular nas estimativas da regressão. Um problema que pode ocorrer com a retirada individual de pontos é o que se denomina em inglês *masking effect*, ou seja, deixar de detectar pontos conjuntamente discrepantes.

Uma das idéias mais inovadoras foi apresentada por Cook (1986), que propõe avaliar a influência conjunta das observações sob pequenas perturbações nos dados ou no modelo, isto é, verificar a existência de pontos que sob modificações modestas causam variações desproporcionais nos resultados, ao invés da avaliação pela retirada individual ou conjunta de pontos. Este método se denomina influência local e teve uma grande receptividade entre os usuários e pesquisadores da área de regressão, motivando muitos artigos onde se aplica a metodologia em modelos particulares ou se propõe extensões da técnica. Por exemplo, Galea, Paula & Bolfarine (1997), Liu (2000) e Galea, Paula & Uribe-Opazo (2003) apresentam estudos de influência local em modelos elípticos lineares e novamente Liu (2002) apresenta um estudo da influência lo-

cal destes modelos no caso multivariado, enquanto Kwan & Fung (1998) aplicam a metodologia em análise fatorial, Gu & Fung (1998) em análise de correlação canônica, Paula (1996) em modelos próprios de dispersão, Ortega, Bolfarine & Paula (2003) em modelos log-gama generalizados com dados censurados, Jansen, Molenberghs, Aerts, Thijs & Steen (2003) em um estudo psiquiátrico de dados binários e Shi & Wang (1999) em regressão ridge. Na classe de erros normais, Lawrence (1988) investiga a aplicação de influência local em modelos lineares com parâmetros na transformação da resposta, Beckman, Nachtsheim & Cook (1987) apresentam estudos de influência em modelos de análise de variância com efeito misto, Tsai & Wu (1992) investigam influência local em modelos auto-regressivos de primeira ordem e modelos heteroscedásticos e Paula (1993) aplica influência local em modelos lineares com restrições nos parâmetros na forma de desigualdades lineares. Para maiores detalhes veja Paula (2004), Cook & Weisberg (1982) e Wei (1998).

Mais especificamente, o conceito de influência local está baseado na curvatura da superfície de log-verossimilhança. Apesar do método de Cook, baseado na curvatura normal, ser de grande utilidade, apresenta alguns inconvenientes; por exemplo, a curvatura normal pode tomar qualquer valor e não é invariante sob uma mudança uniforme de escala. Assim, não foi ao princípio proposto um critério objetivo para avaliar a grandeza da curvatura normal. Com o decorrer do tempo, diferentes autores tentaram dar uma solução a este problema; por exemplo, Fung & Kwan (1997) apresentam uma discussão do uso da influência local baseada na curvatura normal calculada a partir da perturbação de outras quantidades além da verossimilhança deslocada, tais como estimadores de parâmetros ou estatísticas de teste. Eles encontraram que a influência local baseada na curvatura normal não é suficientemente bem fundamentada para gerar resultados úteis quando a primeira derivada com respeito aos parâmetros de perturbação não é nula. Schall & Dunne (1992) mostraram que existe uma relação íntima entre os conceitos de colinearidade de parâmetros e influência local em diagnóstico de regressão, e introduziram uma modificação da curvatura normal de Cook, a saber, a curvatura escalonada. Poon & Poon (1999) construíram uma medida que é uma função injetiva da curvatura normal assumindo valores no intervalo  $[0, 1]$ , chamada curvatura normal conforme, e, sobre esta medida construíram a contribuição agregada para cada vetor de perturbação básico, fornecendo um ponto de partida

objetivo para julgar a grandeza da curvatura.

### 3.1 Influência local

Seja  $\ell(\underline{\theta})$  a log-verossimilhança para um modelo postulado, onde  $\underline{\theta} \in \mathbb{R}^p$  é um vetor de parâmetros desconhecidos. Seja  $\ell(\underline{\theta} | \underline{w})$  a log-verossimilhança correspondente ao modelo perturbado para um  $\underline{w}$  dado, onde  $\underline{w} \in \Omega$ ,  $\Omega \subset \mathbb{R}^r$  representando o conjunto de perturbações relevantes ( $r$  não necessariamente representa o tamanho da amostra). É assumido que existe  $\underline{w}_0$  tal que  $\ell(\underline{\theta} | \underline{w}_0) = \ell(\underline{\theta})$  para todo  $\underline{\theta}$ . Sejam  $\widehat{\underline{\theta}}$  e  $\widehat{\underline{\theta}}_{\underline{w}}$  os estimadores de máxima verossimilhança de  $\underline{\theta}$  sob  $\ell(\underline{\theta} | \underline{w}_0)$  e  $\ell(\underline{\theta} | \underline{w})$  respectivamente. Cook (1986) sugeriu que a influência da perturbação  $\underline{w}$  poderia ser estudada usando a função de verossimilhança deslocada

$$LD(\underline{w}) = 2 \left( \ell(\widehat{\underline{\theta}}) - \ell(\widehat{\underline{\theta}}_{\underline{w}}) \right). \quad (3.1)$$

Os valores de  $LD$  como uma função de  $\underline{w}$  podem ser representados por uma superfície de pontos  $(\underline{w}, LD(\underline{w}))$ ;  $LD(\underline{w})$  tem um valor mínimo de 0 em  $\underline{w} = \underline{w}_0$ . O comportamento de (3.1) é investigado tomando seções planas, isto é, considerando seus valores ao longo de uma reta em  $\Omega$  na direção de um vetor unitário  $\underline{d}$  passando por  $\underline{w}_0$ , ou seja, ao longo de  $\underline{w}(a) = \underline{w}_0 + a\underline{d}$ ,  $a \in \mathbb{R}$ . O resultado é uma curva plana  $(a, LD(\underline{w}(a)))$  cujo comportamento como uma função de  $a$  e em torno de  $a = 0$  é estudado. A curvatura normal na direção  $\underline{d}$  no ponto  $\underline{w} = \underline{w}_0$  é estudada por Cook (1986) como uma medida de influência e é dada por

$$C_d = -2(\underline{d}^\top \ddot{F} \underline{d}) \quad (3.2)$$

onde

$$\ddot{F} = \left. \frac{\partial^2 \ell(\widehat{\underline{\theta}}_{\underline{w}})}{\partial \underline{w} \partial \underline{w}^\top} \right|_{\underline{w} = \underline{w}_0} \quad \text{e} \quad \|\underline{d}\| = 1,$$

$\|\cdot\|$  denotando a norma euclidiana em  $\mathbb{R}^r$ .

Sejam as matrizes  $\Delta$  e  $\ddot{L}$  definidas como

$$\Delta = \left. \frac{\partial^2 \ell(\underline{\theta} | \underline{w})}{\partial \underline{\theta} \partial \underline{w}^\top} \right|_{\underline{\theta} = \widehat{\underline{\theta}}, \underline{w} = \underline{w}_0}, \quad \ddot{L} = \left. \frac{\partial^2 \ell(\underline{\theta} | \underline{w})}{\partial \underline{\theta} \partial \underline{\theta}^\top} \right|_{\underline{\theta} = \widehat{\underline{\theta}}, \underline{w} = \underline{w}_0};$$



então, segundo Cook (1986), a equação (3.2) pode ser escrita como

$$C_d = -2(\underline{d}^\top \Delta^\top (\ddot{L})^{-1} \Delta \underline{d}) \Big|_{\underline{\theta}=\hat{\underline{\theta}}, \underline{w}=\underline{w}_0}.$$

Cook (1986) sugeriu que a direção  $d_{max}$  da curvatura máxima  $C_{max}$  contém valiosa informação sobre diagnóstico. A curvatura máxima é igual ao maior autovalor  $\lambda_{max}$  de  $-\Delta^\top (\ddot{L})^{-1} \Delta$  e  $d_{max}$  é o autovetor associado. Se o valor absoluto do  $i$ -ésimo elemento de  $d_{max}$  é o maior, então, deve prestar-se atenção ao elemento perturbado por  $w_i$ , a  $i$ -ésima componente do vetor  $\underline{w}$ . Embora esta proposta seja muito útil, apresenta alguns inconvenientes como os mencionados anteriormente. A seguir introduz-se a curvatura normal conforme proposta por Poon & Poon (1999), como uma possível solução para estes inconvenientes.

## 3.2 Curvatura normal conforme

A curvatura normal conforme  $B_{\underline{d}}$  na direção  $\underline{d}$  no ponto  $\underline{w}=\underline{w}_0$  é definida por Poon & Poon (1999) como

$$B_{\underline{d}} = -\frac{\underline{d}^\top \ddot{F} \underline{d}}{\sqrt{\text{tr}(\ddot{F}^2)}} \Big|_{\underline{w}=\underline{w}_0} = -\frac{(\underline{d}^\top \Delta^\top (\ddot{L})^{-1} \Delta \underline{d})}{\sqrt{\text{tr}\{(\Delta^\top (\ddot{L})^{-1} \Delta)^2\}}} \Big|_{\underline{\theta}=\hat{\underline{\theta}}, \underline{w}=\underline{w}_0} \quad (3.3)$$

onde as matrizes  $\ddot{F}$ ,  $\Delta$ ,  $\ddot{L}$  são definidas como na seção (3.1) e  $\|\underline{d}\|=1$ .

Note-se da equação (3.3) que o cálculo de  $B_{\underline{d}}$  não requer mais esforço que o de  $C_{\underline{d}}$ . Além disso, a curvatura normal conforme desfruta de muitas propriedades, que são resumidas em poucos teoremas e estão apresentadas em Poon & Poon (1999). Duas destas propriedades são apresentadas a seguir.

Seja  $\Omega$  o conjunto de todas as perturbações. Uma reparametrização é uma função suave  $\phi : \Omega \rightarrow \Lambda$ ,  $\Lambda$  com a mesma dimensão de  $\Omega$ , tal que a matriz jacobiana de  $\phi$  é não singular em todo  $\Omega$ . Considera-se reparametrizações como “modificações do esquema de perturbação”. Além disso, diz-se que uma matriz  $M$  de dimensão  $r \times r$  é uma matriz conforme se existe um número inteiro positivo  $a$  tal que  $MM^\top = aI_r$ ,

onde  $I_r$  representa a matriz identidade de dimensão  $r \times r$ . Uma reparametrização é conforme em  $\mathcal{W}_0$  se a matriz jacobiana em  $\mathcal{W}_0$  é uma matriz conforme; além disso, se uma reparametrização de  $\Omega$  é conforme em um ponto crítico  $\mathcal{W}_0$  no gráfico de  $LD$  sobre  $\Omega$ , então a curvatura normal conforme em qualquer direção em  $\mathcal{W}_0$  é invariante sob a reparametrização. Um exemplo de reparametrização conforme é  $\phi(\mathcal{W}) = M \mathcal{W} + c$ , tal que  $M$  é uma matriz conforme.

Para qualquer direção  $\underline{d}$ ,  $\mathbf{B}_{\underline{d}}$  satisfaz a condição  $0 \leq |\mathbf{B}_{\underline{d}}| \leq 1$ , onde  $|\cdot|$  representa a função valor absoluto. Ou seja,  $\mathbf{B}_{\underline{d}}$  é uma medida normalizada e, assim, torna-se mais fácil interpretar seu valor. Ainda, se  $\{\lambda_1, \dots, \lambda_r\}$  são os autovalores da matriz  $\ddot{F}$  com correspondentes autovetores normalizados  $\{\underline{e}_1, \dots, \underline{e}_r\}$ , tem-se que  $\mathbf{B}_{\underline{e}_j}$  é igual ao autovalor normalizado  $\hat{\lambda}_j$ , que corresponde a

$$\hat{\lambda}_j = \frac{\lambda_j}{\sqrt{\sum_{n=1}^r \lambda_n^2}}.$$

Daí,  $\sum_{j=1}^r \mathbf{B}_{\underline{e}_j}^2 = 1$ , e, assim, se a curvatura normal conforme é igual para todos os autovetores, então o valor comum é  $\frac{1}{\sqrt{r}}$ . Com estas propriedades da curvatura normal conforme, a avaliação da influência local pode ser feita de forma mais objetiva e sistemática.

### 3.3 Avaliação da influência local

Como foi visto na seção (3.1), a curvatura normal  $C_{\underline{d}}$  e a direção  $\underline{d}$  são usadas para avaliar a influência local. Mais especificamente, Cook (1986) sugeriu inspecionar o autovetor  $e_{max}$  com máxima curvatura normal  $C_{max}$  sem considerar seu tamanho, lembrando que  $C_{max}$  é igual ao maior autovalor  $\lambda_{max}$  da matriz  $-\ddot{F}$ . Como a curvatura normal e a curvatura normal conforme diferem só por um fator positivo, as duas curvaturas são medidas de diagnóstico equivalentes e o autovetor  $e_{max}$  dá também a máxima curvatura normal conforme. Entretanto, a curvatura normal conforme é preferida porque suas propriedades de invariância e natureza normalizada facilitam sua interpretação. Todavia, a discussão na seção anterior sugere que um valor de referência para julgar o efeito de  $\mathbf{B}_{\underline{e}_j}$  e  $\mathbf{B}_{\underline{d}}$  em diferentes níveis pode ser determinado

usando o conceito geométrico de curvatura. Assim, a seguinte definição é estabelecida: um autovetor  $\underline{\ell}$  é  $q$ -influyente se  $|\mathbf{B}_{\underline{\ell}}| \geq \frac{q}{\sqrt{r}}$ .

### 3.3.1 Influência de autovetor individual

Um autovetor influente pode ser examinado. Seja  $\underline{\zeta} = \{\underline{E}_1, \dots, \underline{E}_r\}$  a base canônica de  $\mathbb{R}^r$ . Chama-se  $\underline{E}_t$  de  $t$ -ésimo vetor de perturbação básico do espaço de perturbação. Para analisar a contribuição de vetores de perturbação básicos à influência de um autovetor influente  $\underline{\ell}$ , pode-se encontrar os vetores de perturbação básicos que estão próximos a  $\underline{\ell}$ .

Sejam  $\hat{\lambda}_1, \dots, \hat{\lambda}_r$  os autovalores normalizados de  $\ddot{F}$  com correspondentes autovetores normalizados  $\underline{\ell}_1, \dots, \underline{\ell}_r$ . Como  $\underline{\zeta}$  é uma base de  $\mathbb{R}^r$  tem-se para cada  $i = 1, \dots, r$

$$\underline{\ell}_i = \sum_{t=1}^r a_{it} \underline{E}_t;$$

além disso, como  $\|\underline{\ell}_i\| = 1$ , então,  $\sum_{t=1}^r a_{it}^2 = 1$ .

Daí, para cada  $i$  fixo, se a contribuição de todos os  $a_{it}$  é uniforme, então,  $|a_{it}| = \frac{1}{\sqrt{r}}$ . Isto pode ser usado na construção de um ponto de referência para julgar a grandeza. Além do mais, a reta gerada por  $\underline{\ell}_i$  é próxima à reta gerada por  $\underline{E}_t$  se  $|a_{it}|$  é próximo de 1. Este método pode ser usado para estudar  $e_{max}$  ou qualquer autovetor influente individual.

### 3.3.2 Contribuição agregada de vetores de perturbação básicos

De modo geral, pode-se analisar a contribuição de vetores de perturbação básicos para todos os autovetores influentes. Para isto, Poon & Poon (1999) definem a contribuição agregada dos vetores de perturbação básicos para todos os autovetores  $q$ -influyentes e propõem um ponto de referência para avaliar a influência. A seguir, apresenta-se a metodologia usada por Poon & Poon (1999).

Define-se  $\mu_i = |\hat{\lambda}_i|$  e posteriormente ordena-se estes valores da seguinte forma

$$\mu_{max} = \mu_1 \geq \dots \geq \mu_v \geq \frac{q}{\sqrt{r}} > \mu_{v+1} \dots \geq \mu_r \geq 0.$$

Da seção anterior tem-se que

$$\underline{\ell}_j = \sum_{t=1}^r a_{it} \underline{E}_t = a_{i1} \underline{E}_1 + \dots + a_{ir} \underline{E}_r = (a_{i1}, \dots, a_{ir}),$$

logo,  $a_{it}$  denota a  $t$ -ésima coordenada do autovetor normalizado correspondente a  $\mu_i$ . Além disso, como  $\underline{\ell}_1, \dots, \underline{\ell}_r$  são autovetores normalizados, então, para cada  $i = 1, \dots, r$ ,  $\|\underline{\ell}_i\| = 1$ , daí  $\sum_{t=1}^r a_{it}^2 = 1$ .

A contribuição agregada do  $t$ -ésimo vetor de perturbação básico para todos os autovetores  $q$ -influentes é definida como

$$m[q]_t = \sqrt{\sum_{i=1}^v \mu_i a_{it}^2}.$$

Como

$$\sum_{t=1}^r (m[q]_t)^2 = \sum_{t=1}^r \left( \sum_{i=1}^v \mu_i a_{it}^2 \right) = \sum_{i=1}^v \mu_i \left( \sum_{t=1}^r a_{it}^2 \right) = \sum_{i=1}^v \mu_i \cdot (1) = \sum_{i=1}^v \mu_i,$$

então, se a contribuição de todos os vetores de perturbação básicos é a mesma, cada uma é igual a

$$\bar{m}[q] = \sqrt{\frac{1}{r} \left( \sum_{i=1}^v \mu_i \right)}.$$

Conseqüentemente, para determinar a significância da contribuição individual dos vetores de perturbação básicos usa-se  $\bar{m}[q]$ .

Há dois casos extremos a considerar ao usar esta metodologia. O primeiro é permitir  $q$  suficientemente grande tal que se considere a contribuição individual dos vetores de perturbação básicos somente para  $e_{max}$ . Neste caso,

$$m[q]_t = \sqrt{\mu_1 a_{1t}^2} = \sqrt{\mu_1} |a_{1t}|,$$

e

$$\bar{m}[q] = \sqrt{\frac{1}{r} \mu_1} = \frac{1}{\sqrt{r}} \sqrt{\mu_1},$$

logo, este método é equivalente a comparar  $|a_{1t}|$  com  $\frac{1}{\sqrt{r}}$  como foi sugerido na seção (3.3.1). O outro caso extremo é permitir  $q = 0$ , assim, todos os autovalores são incluídos na análise. Neste caso  $m[0]_t$  é chamada a contribuição total, que se denota por  $m_t$  e corresponde a

$$m_t = \sqrt{\sum_{i=1}^r \mu_i a_{it}^2}.$$

Se a contribuição de todas as perturbações básicas é a mesma, então cada uma é igual a

$$\bar{m} = \bar{m}[0] = \sqrt{\frac{1}{r} \sum_{i=1}^r \mu_i} = \sqrt{\frac{\sum_{i=1}^r |\lambda_i|}{r \sqrt{\sum_{i=1}^r \lambda_i^2}}}.$$

A contribuição total  $m_t$  e a curvatura normal conforme  $B_{\underline{E}_t}$  do vetor de perturbação básico  $\underline{E}_t$  estão altamente relacionadas. Como todos os autovalores são não-negativos,  $B_{\underline{E}_t}$  é igual ao quadrado da contribuição total do  $t$ -ésimo vetor de perturbação básico. Ou seja, como a matriz  $-\ddot{F}$  é semi-definida positiva e todos os autovalores são não negativos, então  $m_t^2 = B_{\underline{E}_t}$ , para todo  $t$ .

Se a contribuição total de todos os  $B_{\underline{E}_t}$  é a mesma, então cada uma é igual a  $b = \frac{\text{tr}(\ddot{F})}{r \sqrt{\text{tr}(\ddot{F}^2)}}$ . Portanto, tem-se que  $\bar{m} = B_{\underline{E}_t}$ . Assim  $b$  pode ser utilizado como valor crítico para as curvaturas dos vetores de perturbação básicos.

Segundo a prática comum, se a contribuição agregada é maior que duas vezes a média, então a observação correspondente é considerada influente. Utiliza-se  $2b$  e  $\sqrt{2\bar{m}[q]}$  como valores críticos para  $B_{\underline{E}_t}$  e  $m[q]_t$ , respectivamente. Usa-se  $\sqrt{2}$  ao invés de 2 nos valores críticos para  $m[q]_t$  devido à relação quadrática entre  $m_t$  e  $B_{\underline{E}_t}$ .

### 3.4 Esquemas de perturbação

Existem diferentes formas de se fazer a avaliação da influência de pequenas perturbações num modelo. Pode-se, por exemplo, perturbar a função de log-verossimilhança, ou as variáveis explicativas de maneira coletiva ou individualmente, ou pode-se também perturbar as variáveis resposta. Além disso, essas perturbações podem ser feitas de maneira aditiva ou de maneira multiplicativa.

Neste trabalho os esquemas de perturbação que se consideram são os seguintes:

- *Esquema de perturbação I*  
Perturbação da função de log-verossimilhança de forma multiplicativa.
- *Esquema de perturbação II*  
perturbação nas variáveis explicativas de forma multiplicativa.
- *Esquema de perturbação III*  
perturbação nas variáveis explicativas de forma aditiva.
- *Esquema de perturbação IV*  
perturbação nas variáveis resposta de forma aditiva.

Considere-se o modelo de regressão Dirichlet definido na seção (2.2.1). Seja  $\ell(\alpha_1, \dots, \alpha_p) = \sum_{i=1}^N \ell_i(\alpha_{i1}, \dots, \alpha_{ip})$  a função de log-verossimilhança correspondente a uma amostra de  $N$  observações,  $i = 1, \dots, N$ , seja  $X$  a matriz de dimensão  $N \times k$  cuja  $i$ -ésima linha representa os valores das  $k$  variáveis explicativas para a  $i$ -ésima observação, seja  $Y$  a matriz de dimensão  $N \times p$  cuja  $i$ -ésima linha representa os valores das  $p$  variáveis resposta para a  $i$ -ésima observação e seja  $B$  a matriz de dimensão  $k \times p$  correspondente aos parâmetros desconhecidos da regressão. A seguir apresenta-se a forma como cada esquema foi trabalhado.

#### *Esquema de perturbação I*

Neste caso a dimensão do espaço de perturbação  $\Omega$  é igual ao tamanho da amostra, ou seja,  $r = N$ . Dado  $\underline{w}$ ,  $\underline{w} \in \Omega$ , seja  $\ell(B | \underline{w})$  a log-verossimilhança correspondente ao modelo perturbado, então,

$$\ell(B | \underline{w}) = \sum_{i=1}^N w_i \ell_i(\alpha_{i1}, \dots, \alpha_{ip}),$$

onde  $w_i$  é o  $i$ -ésimo elemento do vetor  $\underline{w}$ .

#### *Esquema de perturbação II*

Como no esquema anterior, a dimensão de  $\Omega$ , o espaço de perturbação, coincide com o tamanho da amostra, ou seja,  $r = N$ . No modelo perturbado, para um  $\underline{w}$  dado, onde  $\underline{w} \in \Omega$ , a matriz  $X$  é substituída pela matriz  $X_{\underline{w}}$ , que corresponde a

$$\mathbf{X}_{\underline{w}} = \begin{pmatrix} w_1 x_{11} & w_1 x_{12} & \cdots & w_1 x_{1k} \\ w_2 x_{21} & w_2 x_{22} & \cdots & w_2 x_{2k} \\ \vdots & \vdots & & \vdots \\ w_N x_{N1} & w_N x_{N2} & \cdots & w_N x_{Nk} \end{pmatrix}.$$

### Esquema de perturbação III

Diferentemente dos dois esquemas de perturbação anteriores, neste esquema a dimensão do espaço de perturbação varia de acordo com o número de variáveis perturbadas. Mais especificamente, no caso em que se perturbam as  $k$  covariáveis, a dimensão de  $\Omega$  será  $Nk$ , ou seja,  $r = Nk$ . Assim, a dimensão do espaço de perturbação vai coincidir com o tamanho da amostra quando se perturba uma única variável explicativa; neste caso, ter-se-á  $r = N$ .

No modelo perturbado, para um  $\underline{w}$  dado,  $\underline{w} \in \Omega$ , a matriz  $\mathbf{X}$  é substituída pela matriz  $\mathbf{X}_{\underline{w}} = \mathbf{X} + \mathbf{W}\mathbf{S}$ , onde  $\underline{w} = \text{vec}(\mathbf{W})$ ,  $\mathbf{W}$  é a matriz de perturbações de dimensão  $N \times k$ , e  $\mathbf{S}$  é uma matriz diagonal  $\mathbf{S} = \text{diag}\{s_1, \dots, s_k\}$  tal que cada  $s_m$  converte a perturbação  $w_{im}$  no tamanho e unidade apropriados para que  $w_{im}s_m$  seja compatível com o  $(i, m)$ -ésimo elemento de  $\mathbf{X}$ , ou seja, com  $x_{im}$ . Tem-se que

$$\mathbf{X}_{\underline{w}} = \begin{pmatrix} x_{11} + w_{11}s_1 & x_{12} + w_{12}s_2 & \cdots & x_{1k} + w_{1k}s_k \\ x_{21} + w_{21}s_1 & x_{22} + w_{22}s_2 & \cdots & x_{2k} + w_{2k}s_k \\ \vdots & \vdots & & \vdots \\ x_{N1} + w_{N1}s_1 & x_{N2} + w_{N2}s_2 & \cdots & x_{Nk} + w_{Nk}s_k \end{pmatrix}.$$

Fazendo alguns dos  $s_m$  iguais a 0, somente algumas das variáveis explicativas são perturbadas.

### Esquema de perturbação IV

Este método de perturbação é diferente dos anteriores, pois ao perturbar as variáveis resposta de forma aditiva, os vetores perturbados devem continuar a ser proporções, assim, é necessário considerar algumas restrições, e em consequência estaremos frente a um problema de maximização com restrições. Neste caso, a máxima curvatura já não será mais o maior autovalor da matriz  $-\Delta^\top (\ddot{L})^{-1} \Delta$ . Ao contrário, como veremos mais

adiante, a máxima curvatura sujeita às restrições impostas estará associada ao maior autovalor de uma nova matriz e à direção normalizada associada a este autovalor. Cabe ressaltar que tanto o desenvolvimento da solução da maximização com restrições como a definição de observação influente e contribuição agregada neste novo esquema de perturbação foram propostos por nós.

Para este esquema de perturbação, a dimensão de  $\Omega$ , o espaço de perturbação, é  $r = Np$ . A seguir explicamos em detalhe qual é o procedimento desenvolvido neste caso.

No modelo perturbado, para um  $\underline{w}$  dado,  $\underline{w} \in \Omega$ , a matriz  $\mathbf{Y}$  é substituída pela matriz  $\mathbf{Y}_{\underline{w}}$ , que corresponde a

$$\mathbf{Y}_{\underline{w}} = \begin{pmatrix} y_{11} + w_{11} & y_{12} + w_{12} & \cdots & y_{1p} + w_{1p} \\ y_{21} + w_{21} & y_{22} + w_{22} & \cdots & y_{2p} + w_{2p} \\ \vdots & \vdots & & \vdots \\ y_{N1} + w_{N1} & y_{N2} + w_{N2} & \cdots & y_{Np} + w_{Np} \end{pmatrix},$$

sujeita às restrições

$$\sum_{j=1}^p w_{1j} = 0, \sum_{j=1}^p w_{2j} = 0, \dots, \sum_{j=1}^p w_{Nj} = 0$$

e a direção que dá a máxima curvatura deve ser buscada entre os vetores de norma um. Seja  $\mathbf{W} \equiv (w_{ij})$ , ou seja,  $\mathbf{W}$  é uma matriz de dimensão  $N \times p$  onde o  $(i, j)$ -ésimo elemento corresponde a  $w_{ij}$ . Fazendo  $\underline{z} = \text{vec}(\mathbf{W})$  podemos escrever nosso problema da seguinte maneira:

maximizar a forma quadrática

$$- \underline{z}^\top (\Delta^\top \ddot{L}^{-1} \Delta) \underline{z}$$

sujeita às restrições

$$\sum_{j=1}^p z_{1j} = 0, \sum_{j=1}^p z_{2j} = 0, \dots, \sum_{j=1}^p z_{Nj} = 0, \|\underline{z}\| = 1. \quad (3.4)$$



Logo, podemos escrever nosso problema da seguinte forma:

$$\text{maximizar } f(\underline{z}) = \underline{z}^\top A \underline{z} \text{ sujeito às restrições } g_n(\underline{z}) = 0, n = 1, \dots, N; \quad h(\underline{z}) = 0$$

onde  $A = -(\Delta^\top \ddot{L}^{-1} \Delta)$ ,  $g_n(\underline{z}) = \sum_{j=1}^p z_{nj}$ ,  $n = 1, \dots, N$  e  $h(\underline{z}) = \underline{z}^\top \underline{z} - 1$ .

Para encontrar a solução do problema anterior vamos utilizar o método dos multiplicadores de Lagrange. Assim, precisamos encontrar a solução do sistema

$$\begin{aligned} g_n(\underline{z}) &= 0, \quad n = 1, \dots, N \\ h(\underline{z}) &= 0 \\ \nabla f(\underline{z}) &= \sum_{n=1}^N \delta_n \nabla g_n(\underline{z}) + \phi \nabla h(\underline{z}) \end{aligned}$$

onde  $\nabla f$  representa o gradiente da função  $f$  e  $\delta_1, \dots, \delta_N, \phi$  são conhecidos como os *multiplicadores de Lagrange*.

Se definimos a matriz  $C$  de dimensão  $N \times Np$  como

$$C = \begin{pmatrix} 1 0 \dots 0 & 1 0 \dots 0 & \dots & 1 0 \dots 0 \\ 0 1 \dots 0 & 0 1 \dots 0 & \dots & 0 1 \dots 0 \\ \vdots & \vdots & & \vdots \\ 0 0 \dots 1 & 0 0 \dots 1 & \dots & 0 0 \dots 1 \end{pmatrix}_{N \times Np}, \quad (3.5)$$

ou seja,

$$C = \left( I_N \quad I_N \quad \dots \quad I_N \right)_{N \times Np},$$

tem-se satisfeito que

$$C \underline{z} = \begin{pmatrix} g_1(\underline{z}) \\ g_2(\underline{z}) \\ \vdots \\ g_N(\underline{z}) \end{pmatrix}.$$

Portanto, para qualquer  $\underline{z}$  é válido que  $\nabla g_n(\underline{z})$  é igual à  $n$ -ésima linha da matriz  $C$ ,  $n = 1, \dots, N$ . Seja  $\underline{E}_n$  o  $n$ -ésimo vetor da base canônica de  $\mathbb{R}^N$ . Então a  $n$ -ésima linha da matriz  $C$  é precisamente o vetor escrito de forma particionada como

$$\left( \underline{E}_n \quad \underline{E}_n \quad \dots \quad \underline{E}_n \right),$$

logo,

$$\nabla g_n(\underline{z}) = \begin{pmatrix} \underline{E}_n & \underline{E}_n & \cdots & \underline{E}_n \end{pmatrix}.$$

Por outro lado, como  $A$  é simétrica, então  $\nabla f(\underline{z}) = 2 \underline{z}^\top A$  e  $\nabla h(\underline{z}) = 2 \underline{z}^\top$ . Logo, o sistema acima pode ser escrito da seguinte forma

$$C\underline{z} = 0 \quad (3.6)$$

$$\underline{z}^\top \underline{z} = 1 \quad (3.7)$$

$$2 \underline{z}^\top A = \sum_{n=1}^N \delta_n \begin{pmatrix} \underline{E}_n & \underline{E}_n & \cdots & \underline{E}_n \end{pmatrix} + 2\phi \underline{z}^\top \quad (3.8)$$

Escrevendo  $\delta_n = 2\phi_n$ , temos que (3.8) fica como segue

$$\underline{z}^\top A = \sum_{n=1}^N \phi_n \begin{pmatrix} \underline{E}_n & \underline{E}_n & \cdots & \underline{E}_n \end{pmatrix} + \phi \underline{z}^\top. \quad (3.9)$$

Mas

$$\begin{aligned} \sum_{n=1}^N \phi_n \begin{pmatrix} \underline{E}_n & \underline{E}_n & \cdots & \underline{E}_n \end{pmatrix} &= \begin{pmatrix} \sum_{n=1}^N \phi_n \underline{E}_n & \sum_{n=1}^N \phi_n \underline{E}_n & \cdots & \sum_{n=1}^N \phi_n \underline{E}_n \end{pmatrix} \\ &= \left( (\phi_1, \dots, \phi_N), (\phi_1, \dots, \phi_N), \dots, (\phi_1, \dots, \phi_N) \right) = (\Phi, \Phi, \dots, \Phi), \end{aligned}$$

onde  $\Phi = (\phi_1, \dots, \phi_N)$ . Então a equação (3.9) fica da seguinte forma

$$\underline{z}^\top A = (\Phi, \Phi, \dots, \Phi) + \phi \underline{z}^\top. \quad (3.10)$$

Multiplicando à direita por  $C^\top$  a equação (3.10) e usando a equação (3.6) obtemos

$$\underline{z}^\top AC^\top = (\Phi, \Phi, \dots, \Phi)C^\top. \quad (3.11)$$

Por outro lado é fácil ver que  $CC^\top = pI_N$ , enquanto  $C^\top C$  é uma matriz composta por  $p \times p$  blocos iguais a  $I_N$ .

$$C^\top C = \begin{pmatrix} I_N & I_N & \cdots & I_N \\ \vdots & \vdots & & \vdots \\ I_N & I_N & \cdots & I_N \end{pmatrix}_{Np \times Np}. \quad (3.12)$$

Multiplicando à direita por  $C$  na equação (3.11) obtemos que

$$\underline{z}^\top AC^\top C = (\Phi, \Phi, \dots, \Phi) \cdot \begin{pmatrix} I_N & I_N & \cdots & I_N \\ \vdots & \vdots & & \vdots \\ I_N & I_N & \cdots & I_N \end{pmatrix} = p(\Phi, \Phi, \dots, \Phi). \quad (3.13)$$

Multiplicando (3.10) à direita por  $\underline{z}$  e usando (3.6), (3.7) e (3.13)

$$f(\underline{z}) = \underline{z}^\top A \underline{z} = (\Phi, \Phi, \dots, \Phi) \underline{z} + \phi = \frac{1}{p} \underline{z}^\top AC^\top C \underline{z} + \phi = \phi. \quad (3.14)$$

Substituindo (3.13) em (3.10) obtemos

$$\underline{z}^\top A = \frac{1}{p} \underline{z}^\top AC^\top C + \phi \underline{z}^\top.$$

Tomando transposta nesta última equação

$$A \underline{z} = \frac{1}{p} C^\top C A \underline{z} + \phi \underline{z}.$$

Podemos escrever esta última equação como

$$QA \underline{z} = \phi \underline{z} = f(\underline{z}) \underline{z}, \quad \text{onde} \quad Q = I_{Np} - \frac{1}{p} C^\top C \quad (3.15)$$

Assim, podemos concluir que  $f(\underline{z})$  é um autovalor  $\phi$  de  $QA$ , e que  $\underline{z}$  é um autovetor correspondente.

Dado que  $S = \{\underline{z} \in \mathbb{R}^{Np} \mid \underline{z}^\top \underline{z} = 1, C \underline{z} = 0\}$  é fechado e limitado em  $\mathbb{R}^{Np}$ , logo, compacto, então a função contínua  $f(\underline{z}) = \underline{z}^\top A \underline{z}$  alcança um máximo  $\underline{z}_0$  neste conjunto, veja Apostol (1977), Teorema 4-20. Tal ponto maximizador satisfaz as equações (3.6)-(3.8) para alguns valores  $\phi_1, \dots, \phi_N, \phi$ , veja Lima (1981).

Dado que o posto de  $QA$  é no máximo  $kp$ , temos não mais do que  $kp$  autovalores de  $QA$  diferentes de zero. Seja  $\tau$  o maior autovalor de  $QA$ , suponha  $\tau \neq 0$ . Queremos provar que  $f(\underline{z}_0) = \tau$ . Pelo argumento anterior, vimos que  $f(\underline{z}_0)$  será um autovalor de  $QA$ , logo  $f(\underline{z}_0) \leq \tau$ . Falta provar agora que existe  $\underline{v} \in S$  tal que  $f(\underline{v}) = \tau$ , assim, por ser  $\underline{z}_0$  maximizador global de  $f$  em  $S$  teremos que  $f(\underline{z}_0) \geq f(\underline{v}) = \tau$  e ficará provado o resultado.

Seja então  $\underline{v}$  que satisfaz  $h(\underline{v}) = \underline{v}^\top \underline{v} - 1 = 0$  e tal que  $QA \underline{v} = \tau \underline{v}$ . Queremos provar que  $C \underline{v} = 0$  e além disso que  $f(\underline{v}) = \tau$ .  $QA \underline{v} = \tau \underline{v}$  pode escrever-se como

$$A \underline{v} = \frac{1}{p} C^\top C A \underline{v} + \tau \underline{v}. \quad (3.16)$$

Multiplicando (3.16) à esquerda por  $C$  e usando que  $CC^\top = pI_N$  obtemos que

$$CA \underline{v} = \frac{1}{p} pI_N CA \underline{v} + \tau C \underline{v} = CA \underline{v} + \tau C \underline{v},$$

logo  $\tau C \underline{v} = 0$ . Como  $\tau \neq 0$ , então  $C \underline{v} = 0$  e assim  $\underline{v} \in S$ ; multiplicando (3.16) por  $\underline{v}^\top$  obtemos

$$\underline{v}^\top A \underline{v} = f(\underline{v}) = \frac{1}{p} \underline{v}^\top C^\top C A \underline{v} + \tau \underline{v}^\top \underline{v} = \frac{1}{p} \underline{v}^\top C^\top C A \underline{v} + \tau.$$

Dado que  $C \underline{v} = 0$ , então  $f(\underline{v}) = \tau$ .

□

Vamos denotar por  $\underline{z}_1$  a direção que dá a máxima curvatura sujeita às restrições descritas anteriormente. Queremos encontrar agora a segunda direção, maximizando  $\underline{z}^\top A \underline{z}$  sujeita às restrições  $h(\underline{z}) = 0, g_n(\underline{z}) = 0, n = 1, \dots, N$  e  $\underline{z}_1^\top \underline{z} = 0$ ; isto pode ser escrito da seguinte maneira

$$\text{maximizar } f(\underline{z}) = \underline{z}^\top A \underline{z} \text{ sujeito a } \begin{pmatrix} C \\ \underline{z}_1^\top \end{pmatrix} \underline{z} = 0 \text{ e } \underline{z}^\top \underline{z} = 1.$$

Fazendo  $C_1 = \begin{pmatrix} C \\ \underline{z}_1^\top \end{pmatrix}$ , podemos escrever o problema da seguinte forma

$$\text{maximizar } f(\underline{z}) = \underline{z}^\top A \underline{z} \text{ sujeito a } C_1 \underline{z} = 0 \text{ e } \underline{z}^\top \underline{z} = 1.$$

Sob estas novas condições, usando novamente o método de multiplicadores de Lagrange, obtemos que a solução é o vetor  $\underline{z}$  que satisfaz

$$Q^* A \underline{z} = \phi^* \underline{z} = f(\underline{z}) \underline{z}$$

onde  $Q^* = (I_{Np} - C_1^\top (C_1 C_1^\top)^{-1} C_1)$  e  $\phi^* = (\underline{z}^\top A \underline{z})$  ou seja,  $\phi^*$  é autovalor de  $Q^* A$  e  $\underline{z}$  é o autovetor associado. Assim, usando a mesma argumentação acima, a solução  $\underline{z}_2$  do nosso segundo problema é o autovetor unitário associado ao maior autovalor  $\phi^*$  de  $Q^* A$ .

Analogamente podemos encontrar  $\underline{z}_3$ , que nos dá o maior valor da forma quadrática  $\underline{z}^\top A \underline{z}$  sujeita às restrições  $\|\underline{z}\| = 1, C \underline{z} = 0$  e que é ortogonal às direções  $\underline{z}_1$  e  $\underline{z}_2$ . Usando este procedimento podemos encontrar as direções  $\underline{z}_1, \underline{z}_2, \dots$ , que dão a máxima curvatura, a segunda máxima curvatura e assim sucessivamente, sujeitas às restrições descritas anteriormente. Veremos em seguida que, se o posto da matriz  $QA$  é  $\rho$ , o posto da matriz  $Q^*A$  é  $\rho - 1$  e, assim, sucessivamente, vamos ter  $\rho$  matrizes  $Q^{*\dots*}A$

de posto  $\rho, \rho - 1, \dots, 1$  das quais vamos encontrar o maior autovalor e um autovetor de norma um associado a esse autovalor, e, assim, poderemos encontrar as  $\rho$  direções  $\tilde{z}_1, \tilde{z}_2, \dots, \tilde{z}_\rho$  descritas acima.

Vale a pena destacar que o problema sem restrições resulta ser um caso particular do problema com restrições, considerando o problema com a única restrição de que a maximização seja feita na superfície da esfera unitária e que a matriz  $C$  seja a matriz identicamente nula.

Nosso problema geral pode ser escrito assim:

$$\text{encontrar as direções que maximizam } f(\tilde{z}) = \tilde{z}^\top A \tilde{z} \text{ sujeitas a } C_i \tilde{z} = 0 \text{ e } \tilde{z}^\top \tilde{z} = 1$$

onde  $C_i$  é uma matriz de dimensão  $(N + i) \times Np$ ,  $0 \leq i \leq \rho - 1$  e  $N + i + 1 = \text{posto}(C_i) + i + 1$  representa o número total de restrições. Note-se que para  $i = 0$ ,  $C_i$  corresponde à matriz dada pela equação (3.5) e temos  $\text{posto}(C_i) + 1 = N + 1$  restrições. Quando a matriz  $C_i$  é a matriz identicamente nula, temos  $\text{posto}(C_i) + 1 = 1$  restrição, a matriz  $QA$  corresponde a  $QA = I_{Np}A = A$  e, lembrando que a matriz  $A$  é a matriz  $-\ddot{F}$  vista na seção 3.1, podemos afirmar que obtivemos o problema original, que tem como única restrição a direção que dá o máximo ser de norma um. Quando  $i = 1$ , temos  $\text{posto}(C_i) + 1 = N + 1 + 1 = N + 2$  restrições, a matriz  $C_i$  corresponde à matriz  $C_1$  descrita anteriormente e a direção que dá a segunda maior curvatura é o autovetor associado ao maior autovalor da matriz  $Q^*A$ .

O Teorema a seguir mostra que todas as direções influentes correspondentes ao problema restrito podem ser encontradas utilizando apenas a matriz  $QA$ . Vale a pena destacar que este Teorema foi proposto e provado por nós e podemos dizer que é o resultado mais importante do nosso trabalho.

**Teorema.** *Sob as condições descritas acima, sejam  $QA$  e  $Q^*A$  as matrizes mencionadas anteriormente, sejam  $\lambda_1 > \dots > \lambda_\rho$  os autovalores não nulos de  $QA$  com correspondentes autovetores unitários  $\tilde{z}_1, \tilde{z}_2, \dots, \tilde{z}_\rho$ . Então, os autovetores  $\tilde{z}_1, \tilde{z}_2, \dots, \tilde{z}_\rho$  são ortogonais,  $\lambda_\rho > 0$ ,  $Q^*A(\tilde{z}_1) = 0$ ,  $\lambda_2$  é o maior autovalor de  $Q^*A$  e o posto de  $Q^*A$  é  $\rho - 1$ .*

*Demonstração.* Seja  $P = C^\top (CC^\top)^{-1}C$ , então  $P$  é idempotente, ou seja,  $P^2 = P$ . Como a matriz  $Q = I_{Np} - P$ , então  $Q$  também é idempotente, ou seja,  $Q^2 = Q$ . Além disso,

todo autovetor de  $QA$  associado a um autovalor não nulo  $\lambda_s, 1 \leq s \leq \rho$ , é também um autovetor de  $Q$ . De fato, suponha que

$$QA\tilde{z}_s = \lambda_s \tilde{z}_s;$$

multiplicando à esquerda por  $Q$  e usando que  $Q$  é idempotente, temos que

$$QA\tilde{z}_s = \lambda_s Q\tilde{z}_s,$$

então  $\lambda_s \tilde{z}_s = \lambda_s Q\tilde{z}_s$ , e assim  $Q\tilde{z}_s = \tilde{z}_s$ , pois  $\lambda_s \neq 0$ .

Sejam  $\tilde{z}_s, \tilde{z}_t$  autovetores associados aos autovalores não nulos  $\lambda_s, \lambda_t$  de  $QA$ . Então  $Q\tilde{z}_s = \tilde{z}_s, Q\tilde{z}_t = \tilde{z}_t$ , ou seja,  $\tilde{z}_s$  e  $\tilde{z}_t$  são autovetores de  $Q$ . Como  $Q$  é simétrica,  $\tilde{z}_s$  e  $\tilde{z}_t$  são ortogonais, isto é,  $\tilde{z}_s^\top \tilde{z}_t = 0$ .

Temos ainda que se  $\tilde{z}_s$  é um autovetor normalizado associado ao autovalor não nulo  $\lambda_s, s = 1, \dots, \rho$ , então,

$$\begin{aligned} QA\tilde{z}_s &= \lambda_s \tilde{z}_s, \\ (I_{Np} - P)A\tilde{z}_s &= \lambda_s \tilde{z}_s, \text{ daí} \\ \tilde{z}_s^\top A\tilde{z}_s - \tilde{z}_s^\top PA\tilde{z}_s &= \lambda_s. \end{aligned}$$

Como  $(I_{Np} - P)\tilde{z}_s = \tilde{z}_s$ , então  $P\tilde{z}_s = 0$  e usando que  $P$  é simétrica, temos que  $\tilde{z}_s^\top P = 0$ , logo podemos afirmar que

$$\lambda_s = \tilde{z}_s^\top A\tilde{z}_s$$

para todo  $s = 1, \dots, \rho$ , ou seja, os autovalores não nulos  $\lambda_s$  de  $QA$  são sempre iguais a  $f(\tilde{z}_s)$ , onde  $\tilde{z}_s$  é um autovetor de norma um associado a  $\lambda_s$ . Dado que  $A$  é semidefinida positiva, então,  $\lambda_s > 0$ , pois assumimos  $\lambda_s \neq 0$ .

Lembrando agora que  $Q^* = (I_{Np} - C_1^\top (C_1 C_1^\top)^{-1} C_1)$ , onde  $C_1 = \begin{pmatrix} C \\ \tilde{z}_1^\top \end{pmatrix}$ , então, temos que

$$C_1 C_1^\top = \begin{pmatrix} C \\ \tilde{z}_1^\top \end{pmatrix} \begin{pmatrix} C^\top & \tilde{z}_1 \end{pmatrix} = \begin{pmatrix} CC^\top & 0 \\ 0 & 1 \end{pmatrix},$$

assim,

$$\begin{pmatrix} C^\top & \tilde{z}_1 \end{pmatrix} \begin{pmatrix} CC^\top & 0 \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} C \\ \tilde{z}_1^\top \end{pmatrix} = C^\top (CC^\top)^{-1} C + \tilde{z}_1 \tilde{z}_1^\top.$$

Então

$$Q^* = (I_{Np} - C^\top (CC^\top)^{-1}C - \tilde{z}_1 \tilde{z}_1^\top) = (I_{Np} - P) - \tilde{z}_1 \tilde{z}_1^\top = Q - \tilde{z}_1 \tilde{z}_1^\top.$$

Como  $\tilde{z}_1$  é autovetor de norma um associado ao autovalor  $\lambda_1$  de  $QA$ , então

$$\begin{aligned} Q^* A \tilde{z}_1 &= QA \tilde{z}_1 - \tilde{z}_1 (\tilde{z}_1^\top A \tilde{z}_1) \\ &= \lambda_1 \tilde{z}_1 - \tilde{z}_1 (\tilde{z}_1^\top A \tilde{z}_1) = (\tilde{z}_1^\top A \tilde{z}_1) \tilde{z}_1 - \tilde{z}_1 (\tilde{z}_1^\top A \tilde{z}_1) = 0. \end{aligned}$$

Logo  $Q^* A \tilde{z}_1 = 0$ , assim, o posto da matriz  $Q^* A$  é pelo menos um a menos que o posto de  $QA$ .

Por outro lado, para  $2 \leq s \leq \rho$ ,

$$Q^* A \tilde{z}_s = QA \tilde{z}_s - \tilde{z}_1 \tilde{z}_1^\top A \tilde{z}_s,$$

porém,

$$\begin{aligned} \tilde{z}_1^\top A \tilde{z}_s &= (Q \tilde{z}_1)^\top A \tilde{z}_s \\ &= \tilde{z}_1^\top QA \tilde{z}_s = \tilde{z}_1^\top \lambda_s \tilde{z}_s = \lambda_s \tilde{z}_1^\top \tilde{z}_s = 0; \end{aligned}$$

esta última igualdade é válida porque  $\tilde{z}_1$  e  $\tilde{z}_s$  são autovetores de  $QA$  e já vimos que os autovetores de  $QA$  são ortogonais. Então, temos que

$$Q^* A \tilde{z}_s = QA \tilde{z}_s = \lambda_s \tilde{z}_s,$$

ou seja,  $\lambda_s$  é autovalor de  $Q^* A$  com correspondente autovetor de norma um  $\tilde{z}_s$ , para todo  $s = 2, \dots, \rho$ . Como o posto de  $Q^* A$  é no máximo  $\rho - 1$ , então  $\lambda_2, \dots, \lambda_\rho$  correspondem a todos os autovalores não nulos de  $Q^* A$  e o posto de  $Q^* A$  é exatamente  $\rho - 1$ . Ainda,  $\lambda_2$  é o maior autovalor de  $Q^* A$ .  $\square$

Usando sucessivamente este argumento, concluímos que  $\tilde{z}_s$  é a direção que maximiza  $\tilde{z}^\top A \tilde{z}$  sujeita às restrições  $C \tilde{z} = 0$ ,  $\tilde{z}^\top \tilde{z} = 1$  e ortogonal a  $\tilde{z}_1, \dots, \tilde{z}_{s-1}$ .

Na prática o Teorema apresentado implica que para encontrar as direções maximizadoras sujeitas às restrições descritas acima, basta trabalhar com a matriz  $QA$  ao invés da matriz  $A = -\ddot{F}$ ; o trabalho computacional será o mesmo, pois o posto de  $QA$

é no máximo o posto de  $A$ . Além disso, vale a pena ressaltar que este resultado que encontramos é válido para qualquer matriz  $C$ , embora a matriz  $C$  no caso do modelo de regressão Dirichlet tenha a forma da matriz dada em (3.5). As direções influentes  $\mathbf{z}_1, \dots, \mathbf{z}_\rho$  que serão usadas na análise de influência serão os autovetores associados aos autovalores não nulos de  $QA$ .

A seguir, apresentamos nossa proposta para estender o conceito de contribuição agregada no caso de perturbação de vetores.

Sejam  $\tau_1, \tau_2, \dots, \tau_r$ , lembrando que  $r = Np$ , os autovalores da matriz  $QA$  com correspondentes autovetores unitários  $\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_r$ , onde para cada  $i = 1, \dots, r$

$$\mathbf{z}_j^\top = (z_{11}^{(i)}, z_{12}^{(i)}, \dots, z_{1p}^{(i)}, z_{21}^{(i)}, z_{22}^{(i)}, \dots, z_{2p}^{(i)}, \dots, z_{N1}^{(i)}, z_{N2}^{(i)}, \dots, z_{Np}^{(i)}).$$

Além disso, definimos

$$\bar{\mathbf{z}}_{i1}^\top = (z_{11}^{(i)}, z_{12}^{(i)}, \dots, z_{1p}^{(i)}), \bar{\mathbf{z}}_{i2}^\top = (z_{21}^{(i)}, z_{22}^{(i)}, \dots, z_{2p}^{(i)}), \dots, \bar{\mathbf{z}}_{iN}^\top = (z_{N1}^{(i)}, z_{N2}^{(i)}, \dots, z_{Np}^{(i)}).$$

Sejam  $\hat{\tau}_1, \dots, \hat{\tau}_r$  os autovalores normalizados de  $QA$ , isto é,

$$\hat{\tau}_i = \frac{\tau_i}{\sqrt{\sum_{n=1}^r \tau_n^2}},$$

para  $i = 1, \dots, r$ , com correspondentes autovetores normalizados  $\mathbf{z}_1, \dots, \mathbf{z}_r$ . Ordenamos estes valores da seguinte forma

$$\hat{\tau}_{(max)} = \hat{\tau}_{(1)} \geq \dots \geq \hat{\tau}_{(v)} \geq \frac{q}{\sqrt{r}} > \hat{\tau}_{(v+1)} \dots \hat{\tau}_{(r)} \geq 0.$$

Da mesma forma que no caso sem restrições, propomos definir um autovetor  $\mathbf{z}$  como  $q$ -influyente se  $|\mathbf{B}_{\mathbf{z}}| \geq \frac{q}{\sqrt{r}}$ , lembrando que neste esquema a dimensão do espaço de perturbação é  $r = Np$ .

Propomos a seguinte definição para a  $t$ -ésima contribuição agregada dos autovetores  $q$ -influyentes, conjuntamente para os  $p$  vetores de perturbação básicos do  $\mathbb{R}^{Np}$   $\mathbf{E}_{(t-1)p+1}, \dots, \mathbf{E}_{tp}$ , que corresponde à  $t$ -ésima observação de dimensão  $p$ . Esta contribuição agregada conjunta será por nós definida como

$$m[q]_t = \sqrt{\sum_{i=1}^v \hat{\tau}_{(i)} \|\bar{\mathbf{z}}_{it}\|^2}.$$



Como

$$\sum_{t=1}^N (m[q]_t)^2 = \sum_{t=1}^N \left( \sum_{i=1}^v \hat{\tau}_{(i)} \|\bar{z}_{it}\|^2 \right) = \sum_{i=1}^v \hat{\tau}_{(i)} \left( \sum_{t=1}^N \|\bar{z}_{it}\|^2 \right) = \sum_{i=1}^v \hat{\tau}_{(i)} \cdot (1) = \sum_{i=1}^v \hat{\tau}_{(i)},$$

então, se a contribuição de todos os vetores de perturbação básicos é a mesma, cada uma é igual a

$$\bar{m}[q] = \sqrt{\frac{1}{N} \left( \sum_{i=1}^v \hat{\tau}_{(i)} \right)}.$$

Então, para determinar a significância da contribuição conjunta dos vetores de perturbação básicos para cada observação de dimensão  $p$ , usamos  $\bar{m}[q]$ .

Neste caso, analogamente ao caso sem restrições, se a contribuição agregada é maior que duas vezes a média, então a observação correspondente é considerada *influyente*. Utilizamos  $\sqrt{2\bar{m}[q]}$  como valor crítico para  $m[q]_t$  devido à relação  $\bar{m}[q] = \sqrt{\frac{1}{N} \left( \sum_{i=1}^v \hat{\tau}_{(i)} \right)}$ .

No caso em que encontramos somente um autovetor  $q$ -influyente vamos ter  $m[q]_t = \sqrt{\hat{\tau}_{(1)} \|\bar{z}_{1t}\|^2}$  e  $\sqrt{2\bar{m}[q]} = \sqrt{\frac{1}{N} \hat{\tau}_{(1)}}$ , logo, na presença de um único autovetor  $q$ -influyente, basta comparar  $\|\bar{z}_{1t}\|$  com  $\sqrt{\frac{2}{N}}$ , para  $t = 1, \dots, N$ , para encontrar as observações influyentes.

# Capítulo 4

## Aplicações

### 4.1 Introdução

Como ilustração, considera-se o conjunto de dados apresentados por Aitchison (2003), p. 359, ao qual pode ser ajustado modelo de regressão Dirichlet. Os valores destes dados encontram-se na Tabela 1 do Apêndice B.

Para fazer o estudo de influência local aplicando a metodologia mencionada nos capítulos anteriores, procede-se da seguinte forma: primeiramente os parâmetros do modelo de regressão Dirichlet são estimados através da maximização da função de log-verossimilhança. A seguir, as contribuições agregadas  $m[q]_t$  para todos os autovetores  $q$ -influentes, com  $q$  variando de 0 a 3, assim como a curvatura normal conforme  $B_{\underline{E}_t}$  do vetor de perturbação básico  $\underline{E}_t$  são calculadas para  $t = 1, \dots, N$ , onde  $N$  é o tamanho da amostra. Os cálculos das derivadas necessárias para encontrar a curvatura normal conforme segundo os diferentes esquemas de perturbação aqui analisados estão apresentados nos Apêndices A.1, A.2, A.3 e A.4. Serão apresentadas as contribuições agregadas  $m[q]_t$  e a curvatura normal conforme  $B_{\underline{E}_t}$  do vetor de perturbação básico  $\underline{E}_t$  para os esquemas de perturbação *I*, *II* e *III*. A contribuição total  $m[0]_t$  não é apresentada, pois é equivalente a  $B_{\underline{E}_t}$  nos esquemas *I*, *II* e *III*. Dado que o *esquema de perturbação IV* tem um tratamento diferente dos outros esquemas, serão apresentadas as medidas  $m[q]_t$  para  $q = 0$  e para  $q = 1$ , pois para  $q = 2, 3$  não se encontraram autovetores

$q$ -influentes. No caso em que um único autovetor,  $\mu_1$ , é encontrado como 3-influente ter-se-á que  $m[1]_t = m[2]_t = m[3]_t = \sqrt{|\mu_1|} |a_{1t}|$ .

A partir das observações detectadas como influentes, faz-se uma retirada de tais observações, uma a uma e também conjuntamente com o intuito de comparar as novas estimativas com as estimativas obtidas ao usar a amostra completa e verificar o impacto que cada observação influente causa nessas estimativas. Para isto, é razoável determinar medidas onde se possa avaliar quão precisas são as estimativas e o impacto causado nelas a partir da retirada individual ou coletiva das observações influentes. A seguir mostra-se algumas medidas que podem ser utilizadas para este fim.

⊗ Desvio quadrático médio (*DQM*):

$$DQM_j = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_{ij} - \hat{y}_{ij})^2};$$

⊗ Erro absoluto médio (*EAM*):

$$EAM_j = \frac{1}{N} \sum_{i=1}^N |y_{ij} - \hat{y}_{ij}|;$$

⊗ Erro absoluto percentual médio (*EPAM*):

$$EPAM_j = \left( \frac{1}{N} \sum_{i=1}^N \frac{|y_{ij} - \hat{y}_{ij}|}{y_{ij}} \right) \times 100\%,$$

⊗ Desvio relativo percentual (*DRP*):

$$DRP_{hj} = \left( \frac{\hat{\beta}_{hj}^{(i)} - \hat{\beta}_{hj}}{\hat{\beta}_{hj}} \right) \times 100\%,$$

onde  $N$  é o tamanho da amostra,  $\hat{y}_{ij}$  é o valor ajustado para cada  $y_{ij}$ ,  $j = 1, \dots, p$  e  $\hat{\beta}_{hj}^{(i)}$  é o estimador do parâmetro  $\beta_{hj}$  encontrado ao tirar a  $i$ -ésima observação,  $h = 1, \dots, k$ .

## 4.2 Composição de sedimentos

O conjunto de dados aqui analisado consiste da composição de areia, lodo e argila de 39 amostras de sedimentos a diferentes profundidades de água, medidas em hectômetros, em um lago ártico. O vetor de variáveis resposta  $(Y_{i1}, Y_{i2}, Y_{i3})$  é constituído pelas proporções de areia, lodo e argila dos sedimentos, e as covariáveis  $(x_{i1}, x_{i2})$  correspondem ao intercepto e às profundidades de água para  $i = 1, \dots, 39$ . Logo, neste exemplo, tem-se o tamanho de amostra  $N = 39$ , o número de componentes  $p = 3$  e  $k$ , o número de covariáveis, é 2.

### 4.2.1 Esquema de perturbação I

Os autovalores normalizados não nulos para o *esquema de perturbação I* são 0.6704, 0.5318, 0.4283, 0.2501, 0.1373 e 0.0542. Neste conjunto de dados foram encontrados quatro autovetores 1-influentes, três autovetores 2-influentes e dois autovetores 3-influentes. Na Tabela 4.2.1.1 estão apresentados para  $t = 12, 18, 25, 35$  e 37 os valores da curvatura normal conforme do vetor de perturbação básico  $B_{\underline{E}_t}$  e a contribuição agregada  $m[q]_t$  para todos os autovetores  $q$ -influentes com  $q$  variando de 1 a 3. Para todas as medidas, as observações 12, 18 e 37 são consideradas como influentes.

**Tabela 4.2.1.1:** Medidas de observações influentes para o *esquema de perturbação I*

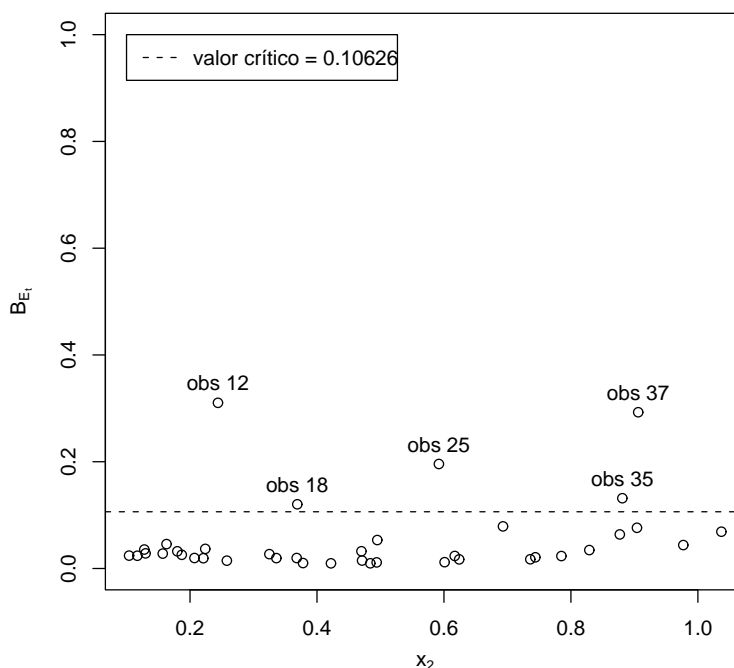
Medidas	$N(\underline{\mathcal{L}})$	$\bar{m}[q]$	Valores críticos	Valores das seguintes obs.				
				obs 12	obs 18	obs 25	obs 35	obs 37
$B_{\underline{E}_t}$	39	0.0531	0.1063	0.3103	0.1203	0.1958	0.1317	0.2927
$m[1]_t$	4	0.2196	0.3106	0.5567	0.3457	0.4421	0.2930	0.4968
$m[2]_t$	3	0.2045	0.2892	0.5417	0.3309	0.3922	0.2930	0.4752
$m[3]_t$	2	0.1756	0.2483	0.3901	0.2616	0.2367	0.2912	0.4713

$N(\underline{\mathcal{L}})$  representa a quantidade de autovetores influentes

Nas Figuras 4.2.1.1-4.2.1.4 estão apresentados os gráficos das curvaturas normais conformes  $B_{\underline{E}_t}$  e as medidas  $m[q]_t$  contra a variável explicativa  $x_2$  que corresponde às profundidades de água, para o *esquema de perturbação I*. Note-se que para as profundidades de água das observações  $t = 12, 18$  e 37 as medidas  $m[q]_t$  e  $B_{\underline{E}_t}$  estão acima

dos valores críticos respectivos, ou seja, são influentes, enquanto que para  $q = 0, 1, 2$  a observação  $t = 25$  é incluída como influente e a observação  $t = 35$  é considerada como marginalmente influente, enquanto que para  $q = 3$  a observação 25 é considerada marginalmente influente e a observação  $t = 35$  é incluída como influente. Diz-se que uma observação é marginalmente influente se a diferença absoluta entre o valor crítico e o valor de  $m[q]_t$  para a observação está muito próxima de zero.

**Figura 4.2.1.1:** Gráfico de  $B_{E_t}$  contra  $x_2$  para o esquema de perturbação I



Ao tirar as observações mais influentes, ou seja, as observações 12 e 37 da amostra completa, ficando assim com uma amostra de 37 observações, obtém-se 4 autovetores 1-influente, três autovetores 2-influente e dois autovetores 3-influente. Neste caso, as observações 18, 25, 35 e 36 são influente, enquanto que a observação 34 é considerada marginalmente influente para  $q = 0, 1$  e influente para  $q = 2, 3$ , e a observação  $t = 38$  é encontrada como marginalmente influente para  $q = 0, 1, 3$  na nova amostra com 37 observações. Na Tabela 4.2.1.2 estão apresentados estes resultados.

Figura 4.2.1.2: Gráfico de  $m[1]_t$  contra  $x_2$  para o esquema de perturbação I

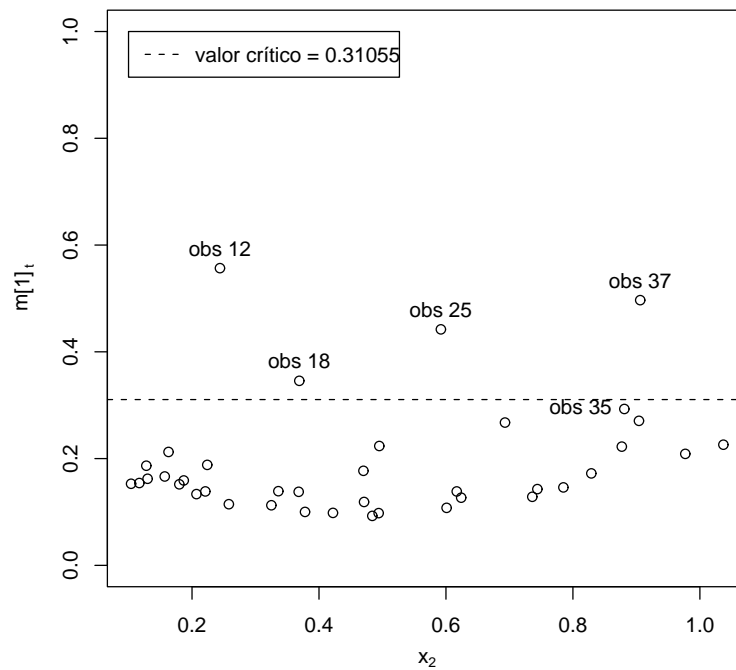


Figura 4.2.1.3: Gráfico de  $m[2]_t$  contra  $x_2$  para o esquema de perturbação I

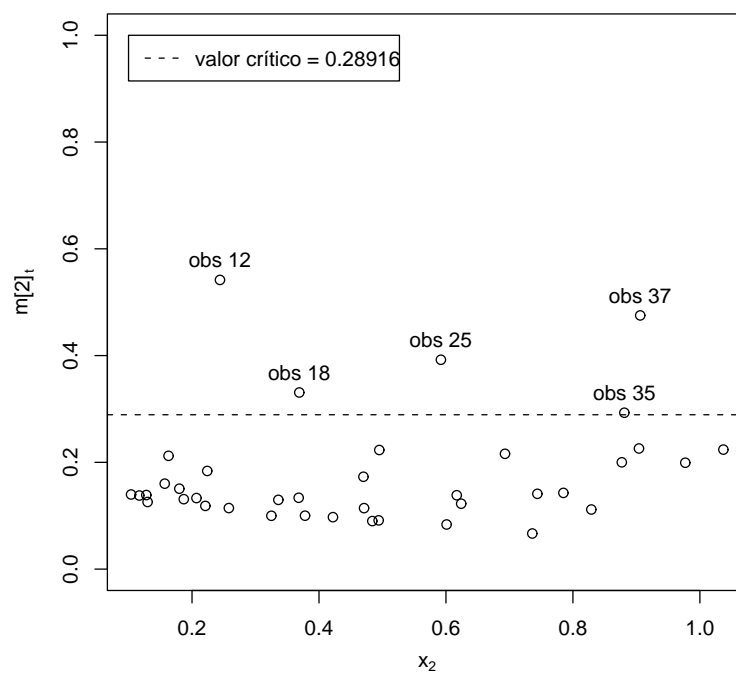


Figura 4.2.1.4: Gráfico de  $m[3]_t$  contra  $x_2$  para o esquema de perturbação I

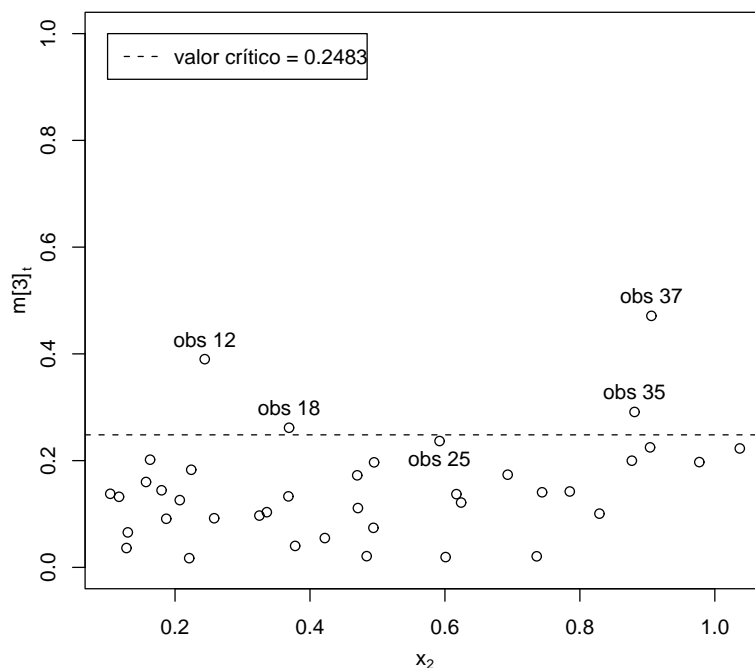


Tabela 4.2.1.2: Medidas de observações influentes para o esquema de perturbação I tirando as observações 12 e 37.

Medidas	$N(\underline{\rho})$	$\bar{m}[q]$	Valores críticos	Valores das seguintes obs.		
				obs 18	obs 25	obs 34
$B_{\underline{\rho}_t}$	37	0.0533	0.1066	0.1416	0.2444	0.1066
$m[1]_t$	4	0.2233	0.3157	0.3759	0.4895	0.3213
$m[2]_t$	3	0.2100	0.2970	0.3408	0.4877	0.3200
$m[3]_t$	2	0.1855	0.2623	0.3405	0.3600	0.3177

Medidas	$N(\underline{\rho})$	$\bar{m}[q]$	Valores críticos	Valores das seguintes obs.		
				obs 35	obs 36	obs 38
$B_{\underline{\rho}_t}$	37	0.0533	0.1066	0.2356	0.1276	0.1047
$m[1]_t$	4	0.2233	0.3157	0.4827	0.3465	0.2828
$m[2]_t$	3	0.2100	0.2970	0.4776	0.3334	0.2671
$m[3]_t$	2	0.1855	0.2623	0.4707	0.3164	0.2633

$N(\underline{\rho})$  representa a quantidade de autovetores influentes

Note-se que na amostra com 37 observações obteve-se como as mais influentes as observações 25 e 35. Tirando então as observações 12, 25, 35 e 37 da amostra original e ficando assim com uma amostra de tamanho 35, se encontraram 4 autovetores 1-influente e 2 autovetores 2-influente e 3-influente, logo  $m[2]_t = m[3]_t$ . Além disso, as observações 18, 34 e 36 são consideradas como influente para todas as medidas, enquanto que a observação 29 é considerada como influente unicamente para  $q = 0, 1$ . A Tabela 4.2.1.3 apresenta estes resultados.

**Tabela 4.2.1.3:** Medidas de observações influentes para o *esquema de perturbação I* tirando as observações 12, 25, 35 e 37.

Medidas	$N(\underline{\rho})$	$\bar{m}[q]$	Valores críticos	Valores das seguintes obs.			
				obs 18	obs 29	obs 34	obs 36
$B_{E_t}$	35	0.0546	0.1093	0.1767	0.1673	0.2330	0.2429
$m[1]_t$	4	0.2270	0.3204	0.4195	0.4073	0.4771	0.4873
$m[2]_t$	2	0.1934	0.2735	0.3663	0.1982	0.4718	0.4841
$m[3]_t$	2	0.1934	0.2735	0.3663	0.1982	0.4718	0.4841

$N(\underline{\rho})$  representa a quantidade de autovetores influentes

A Tabela 4.2.1.4 apresenta as estimativas, o desvio padrão e o desvio relativo percentual  $DRP_{hj}$  dos  $\beta_{hj}$  tirando coletivamente as observações encontradas como influente para o *esquema de perturbação I*. Note-se que as estimativas dos parâmetros  $\beta_{11}$ ,  $\beta_{12}$  e  $\beta_{13}$  foram as que tiveram maior impacto, chegando a ter uma variação de 43% da estimativa correspondente ao parâmetro  $\beta_{11}$ . Fizemos um teste para testar a significância dos parâmetros correspondentes ao intercepto e também a significância dos parâmetros correspondentes à variável explicativa a um nível de 5% de confiança e pudemos verificar que em todos os casos a hipótese nula é rejeitada, ou seja, os parâmetros são significativos.



**Tabela 4.2.1.4:** Estimativas, desvio padrão e  $DRP_{hj}$  dos  $\beta_{hj}$  com a amostra completa e tirando as observações influentes.

	$\hat{\beta}_{hj}$ , $dp(\hat{\beta}_{hj})$ e $DRP_{hj}$ tirando as obs		
	nenhuma	obs 12, 37	obs 12, 25, 35, 37
$\hat{\beta}_{11}$	0.1457	0.2087	0.2112
$dp(\hat{\beta}_{11})$	0.3245	0.3323	0.3335
$DRP_{11}$		43.25%	45.01%
$\hat{\beta}_{12}$	-0.2876	-0.1610	-0.2105
$dp(\hat{\beta}_{12})$	0.3147	0.3231	0.3237
$DRP_{12}$		-44.01%	-26.80%
$\hat{\beta}_{13}$	-1.1349	-0.9515	-1.0210
$dp(\hat{\beta}_{13})$	0.3109	0.3202	0.3211
$DRP_{13}$		-16.16%	-10.03%
$\hat{\beta}_{21}$	2.2590	2.3771	2.7354
$dp(\hat{\beta}_{21})$	0.5778	0.6003	0.6188
$DRP_{21}$		5.23%	21.09%
$\hat{\beta}_{22}$	5.4886	5.5815	6.0463
$dp(\hat{\beta}_{22})$	0.5664	0.5896	0.6083
$DRP_{22}$		1.69%	10.16%
$\hat{\beta}_{23}$	6.3690	6.4286	6.9006
$dp(\hat{\beta}_{23})$	0.5631	0.5871	0.6060
$DRP_{23}$		0.94%	8.35%

$dp(\hat{\beta}_{hj})$  representa o desvio padrão de  $\hat{\beta}_{hj}$

Nas medidas  $DQM_j$ ,  $EAM_j$  e  $EPAM_j$  descritas na seção 4.1,  $y_{ij}$  corresponde aos valores estimados pelo modelo quando se usa a amostra com as 39 observações e  $\hat{y}_{ij}$  são os valores estimados a partir da amostra que resulta depois de tirar uma ou várias das observações influentes obtidas em cada esquema de perturbação.

A Tabela 4.2.1.5 mostra o  $DQM_j$ , o  $EAM_j$  e o  $EPAM_j$  considerando a retirada conjunta das observações mais influentes obtidas no *esquema de perturbação I*. Observe-se que a retirada das observações influentes 12 e 37 causa os maiores  $DQM_j$ ,  $EAM_j$  e  $EPAM_j$  para a primeira e segunda coordenada, enquanto que para a primeira coordenada o impacto obtido com a retirada simultânea das observações influentes 12, 25, 35 e 37 é um pouco maior que o obtido com a retirada somente das observações 12 e 37.

**Tabela 4.2.1.5:**  $DQM_j$ ,  $EAM_j$  e  $EPAM_j$  para os  $\hat{y}_j$ , sob *esquema de perturbação I*.

	Medidas de erro para os $\hat{y}_j$ sem observações					
	obs 12, 37			obs 12, 25, 35 e 37		
	1 Coord.	2 Coord.	3 Coord.	1 Coord.	2 Coord.	3 Coord.
$DQM_j$	0.0120	0.0042	0.0115	0.0075	0.0009	0.0072
$EAM_j$	0.0103	0.0038	0.0113	0.0072	0.0009	0.0071
$EPAM_j$	5.16%	0.82%	4.07%	4.73%	0.19%	2.54%

## 4.2.2 Esquema de perturbação II

Na Tabela 4.2.2.1 estão apresentados os valores da curvatura normal conforme  $B_{E_t}$  do vetor de perturbação básico  $E_t$  e a contribuição agregada  $m[q]_t$  para todos os autovetores  $q$ -influentes com  $q$  variando de 1 a 3. Observe-se que os valores de  $m[2]_t$  e  $m[3]_t$  coincidem; isto é devido ao fato de que somente o autovetor  $e_{max}$  foi encontrado como 2-influente e 3-influente. Observe-se que para todas as medidas as observações  $t = 34, 35, 36$  e  $37$  são encontradas como influentes, sendo a observação 37 a mais influente, enquanto somente para  $q = 0$  e  $q = 1$  a observação 39 é incluída como influente e a observação 38 é marginalmente influente.

Para o *esquema de perturbação II*, os autovalores normalizados não nulos são 0.9637, 0.2495, 0.0819, 0.0393, 0.0230 e 0.0143; foram encontrados dois autovetores 1-influentes e somente o autovetor  $e_{max}$  foi encontrado como 2-influente e 3-influente.

**Tabela 4.2.2.1:** Medidas de observações influentes para o *esquema de perturbação II*

Medidas	$N(\underline{\rho})$	$\bar{m}[q]$	Valores críticos	Valores das seguintes obs.		
				obs 34	obs 35	obs 36
$\mathbf{B}_{\underline{E}_t}$	39	0.0352	0.0703	0.1416	0.2831	0.1476
$m[1]_t$	2	0.1764	0.2494	0.3658	0.5126	0.3831
$m[2]_t$	1	0.1572	0.2223	0.3628	0.5122	0.3717
$m[3]_t$	1	0.1572	0.2223	0.3628	0.5122	0.3717
Medidas	$N(\underline{\rho})$	$\bar{m}[q]$	Valores críticos	Valores das seguintes obs.		
				obs 37	obs 38	obs 39
$\mathbf{B}_{\underline{E}_t}$	39	0.0352	0.0703	0.4218	0.0733	0.1602
$m[1]_t$	2	0.1764	0.2494	0.6239	0.2654	0.3956
$m[2]_t$	1	0.1572	0.2223	0.6238	0.0535	0.0763
$m[3]_t$	1	0.1572	0.2223	0.6238	0.0535	0.0763

$N(\underline{\rho})$  representa a quantidade de autovetores influentes

Nas Figuras 4.2.2.1-4.2.2.3 estão apresentados os gráficos das curvaturas normais conformes  $\mathbf{B}_{\underline{E}_t}$  e as medidas  $m[q]_t$  contra a variável explicativa  $x_2$  que corresponde às profundidades de água, para o *esquema de perturbação II*, ou seja, para o esquema onde se perturba a variável explicativa de forma multiplicativa. Não é apresentado o gráfico correspondente às medidas  $m[3]_t$  porque somente o autovetor normalizado  $e_{max}$ , correspondente ao maior autovalor de  $-\ddot{F}$ , foi encontrado como 2-influente e 3-influente. Note-se que para as profundidades de água das observações  $t = 34, 35, 36, 37$  e  $39$  as medidas  $m[q]_t$  e  $\mathbf{B}_{\underline{E}_t}$  estão acima dos valores críticos respectivos, ou seja, são influentes, e a observação  $t = 38$  é considerada marginalmente influente. Assim, para este esquema de perturbação, as observações com grandes profundidades são as encontradas como influentes.

Figura 4.2.2.1: Gráfico de  $B_{E_t}$  contra  $x_2$  para o esquema de perturbação II

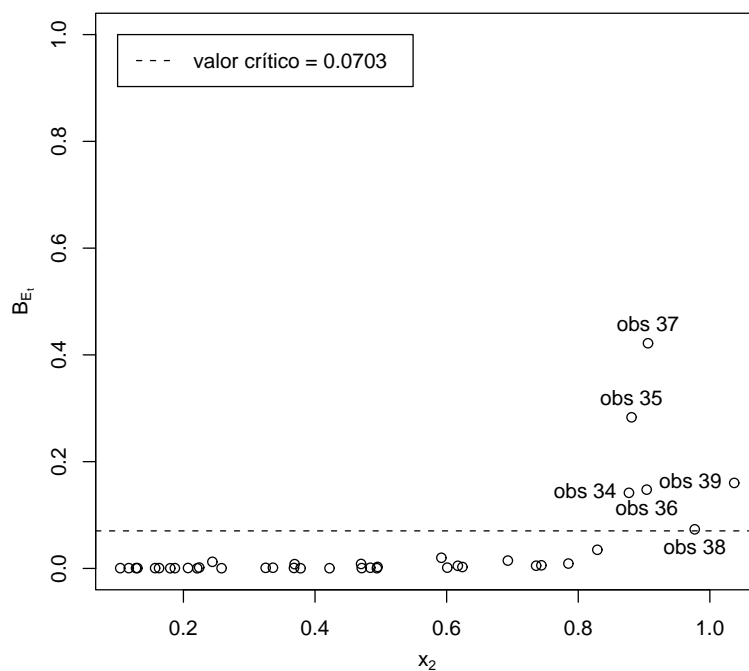
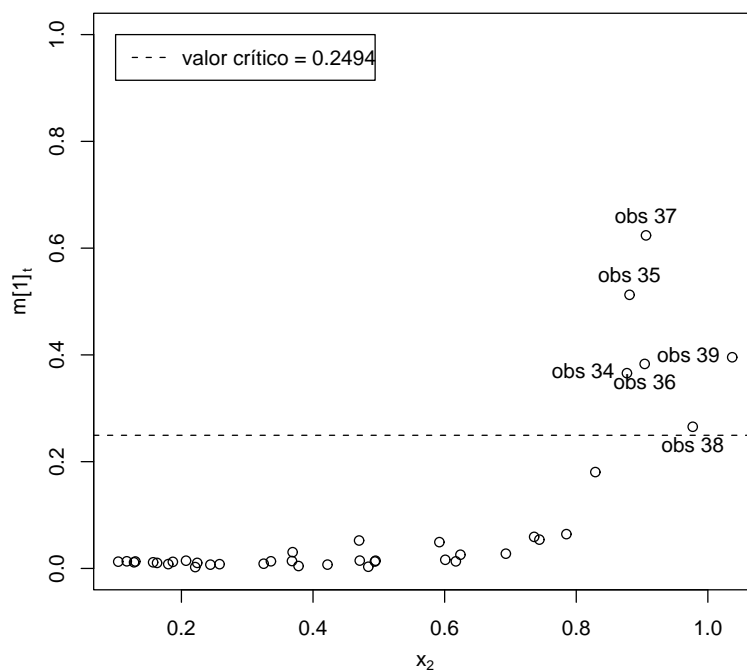
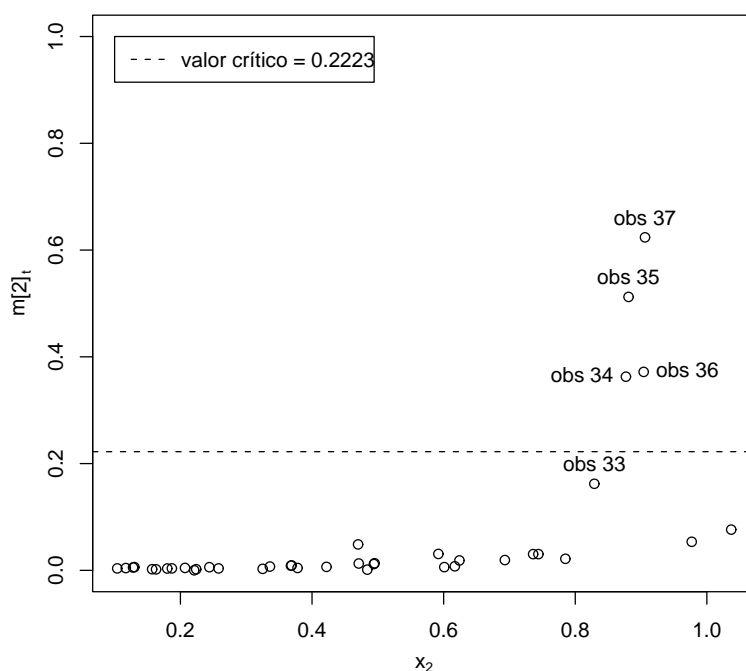


Figura 4.2.2.2: Gráfico de  $m[1]_t$  contra  $x_2$  para o esquema de perturbação II



**Figura 4.2.2.3:** Gráfico de  $m[2]_t$  contra  $x_2$  para o esquema de perturbação II



Tirando a observação mais influente, isto é, a observação 37 e ficando assim com uma amostra de 38 observações, obtém-se como influentes as observações 34, 35 e 36, ao passo que para  $q = 0$  e  $q = 1$  a observação 39 é incluída como influente e a observação 38 é considerada marginalmente influente. Note-se também que as medidas  $m[1]_t$  e  $m[2]_t$  coincidem, isto é devido a que são encontrados dois autovetores 1-influente e 2-influente. Estes resultados estão apresentados na Tabela 4.2.2.2.

**Tabela 4.2.2.2:** Medidas de observações influentes para o esquema de perturbação II tirando a observação 37.

Medidas	$N(\underline{\xi})$	$\bar{m}[q]$	Valores críticos	Valores das seguintes obs.				
				obs 34	obs 35	obs 36	obs 38	obs 39
$\mathbf{B}_{\underline{\xi}_t}$	38	0.0362	0.0724	0.2177	0.4265	0.2476	0.0928	0.2327
$m[1]_t$	2	0.1832	0.2591	0.4660	0.6514	0.4953	0.2940	0.4793
$m[2]_t$	2	0.1832	0.2591	0.4660	0.6514	0.4953	0.2940	0.4793
$m[3]_t$	1	0.1573	0.2225	0.4644	0.6506	0.4859	0.0569	0.0838

$N(\underline{\xi})$  representa a quantidade de autovetores influentes

Continuando com esta análise, ao tirar as observações mais influentes, isto é, as observações 35 e 37 da amostra completa, encontram-se dois autovetores 1-influente e 2-influente, e somente o autovetor  $e_{max}$  como 3-influente. Além disso, as observações: 34, 36 para todo  $q$ , a observação 39 para  $q = 0$  e  $q = 1$  e a observação 38 para  $q = 1$ , são consideradas como influentes, enquanto que a observação 33 para todo  $q$  e a observação 38 para  $q = 0$  são encontradas como marginalmente influentes. Isto pode ser visto na Tabela 4.2.2.3.

**Tabela 4.2.2.3:** Medidas de observações influentes para o esquema de perturbação II tirando as observações 35 e 37.

Medidas	$N(\underline{\epsilon})$	$\bar{m}[q]$	Valores críticos	Valores das seguintes obs.				
				obs 33	obs 34	obs 36	obs 38	obs 39
$B_{\underline{\epsilon}_t}$	37	0.0381	0.0761	0.0920	0.3868	0.4396	0.1124	0.2709
$m[1]_t$	2	0.1882	0.2661	0.2951	0.6198	0.6622	0.3258	0.5171
$m[2]_t$	2	0.1882	0.2661	0.2951	0.6198	0.6622	0.3258	0.5171
$m[3]_t$	1	0.1576	0.2228	0.2903	0.6194	0.6616	0.0125	0.0191

$N(\underline{\epsilon})$  representa a quantidade de autovetores influentes

A Tabela 4.2.2.4 apresenta as estimativas, o desvio padrão e o desvio relativo percentual  $DRP_{hj}$  dos  $\beta_{hj}$  tirando as observações encontradas como influentes para o esquema de perturbação II. Observe-se que com a retirada simultânea das observações 35 e 37 a estimativa do parâmetro  $\beta_{11}$  foi a que apresentou maior impacto, chegando a mudar de sinal com uma variação de -101.40%, as estimativas correspondentes ao intercepto apresentaram menor desvio padrão que as estimativas dos parâmetros correspondentes à variável explicativa e houve maior variação nas estimativas ao tirar conjuntamente as observações 35 e 37. Da mesma forma que no esquema de perturbação anterior, se testou a significância dos parâmetros correspondentes ao intercepto e também à variável explicativa a um nível de 5% de confiança e pudemos verificar que em todos os casos a hipótese nula é rejeitada, ou seja, os parâmetros são significativos.

**Tabela 4.2.2.4:** Estimativas, desvio padrão e  $DRP_{hj}$  dos  $\beta_{hj}$  com a amostra completa e tirando as observações influentes.

	$\hat{\beta}_{hj}$ , $dp(\hat{\beta}_{hj})$ e $DRP_{hj}$ tirando as obs		
	nenhuma	obs 37	obs 35, 37
$\hat{\beta}_{11}$	0.1457	0.0637	-0.0020
$dp(\hat{\beta}_{11})$	0.3245	0.3261	0.3277
$DRP_{11}$		-56.24%	-101.40%
$\hat{\beta}_{12}$	-0.2876	-0.3818	-0.4704
$dp(\hat{\beta}_{12})$	0.3147	0.3154	0.3163
$DRP_{12}$		32.76%	63.55%
$\hat{\beta}_{13}$	-1.1349	-1.2081	-1.2733
$dp(\hat{\beta}_{13})$	0.3109	0.3119	0.3130
$DRP_{13}$		6.45%	12.19%
$\hat{\beta}_{21}$	2.2590	2.5219	2.7357
$dp(\hat{\beta}_{21})$	0.5778	0.5942	0.6112
$DRP_{21}$		11.64%	21.10%
$\hat{\beta}_{22}$	5.4886	5.8109	6.1284
$dp(\hat{\beta}_{22})$	0.5664	0.5821	0.5986
$DRP_{22}$		5.87%	11.66%
$\hat{\beta}_{23}$	6.3690	6.6957	6.9778
$dp(\hat{\beta}_{23})$	0.5631	0.5791	0.5958
$DRP_{23}$		5.13%	9.56%

$dp(\hat{\beta}_{hj})$  representa o desvio padrão de  $\hat{\beta}_{hj}$

A Tabela 4.2.2.5 mostra o  $DQM_j$ , o  $EAM_j$  e o  $EPAM_j$  considerando a retirada individual e coletiva das observações influentes obtidas no *esquema de perturbação II*. Note-se que as medidas dos erros foram um pouco maiores quando se tiraram as duas observações mais influentes conjuntamente que quando se retirou somente a observação influente 37.

**Tabela 4.2.2.5:**  $DQM_j$ ,  $EAM_j$  e  $EPAM_j$  para os  $\hat{y}_j$ , sob esquema de perturbação II.

	Medidas de erro para os $\hat{y}_j$ sem observações					
	obs 37			obs 35, 37		
	1 Coord.	2 Coord.	3 Coord.	1 Coord.	2 Coord.	3 Coord.
$DQM_j$	0.0025	0.0032	0.0055	0.0053	0.0030	0.0074
$EAM_j$	0.0023	0.0030	0.0053	0.0047	0.0026	0.0073
$EPAM_j$	2.15%	0.63%	1.69%	4.41%	0.59%	2.46%

### 4.2.3 Esquema de perturbação III

No esquema de perturbação III, a dimensão do espaço de perturbação coincide com o tamanho da amostra, pois só tem sentido perturbar a variável explicativa que corresponde à profundidade da água. Neste esquema os elementos  $s_1, s_2$  da matriz diagonal  $S$  são os desvios padrão das variáveis explicativas para as 39 observações. Assim  $s_1 = 0$  e  $s_2 = 0.277$ ; isto garante que o intercepto não vai ser perturbado.

Os autovalores normalizados não nulos são 0.9062, 0.3370, 0.2425, 0.0694, 0.0388, 0.0047. São encontrados três autovetores 1-influentes, dois autovetores 2-influentes e um autovetor 3-influente, isto é, o único vetor considerado para encontrar as medidas  $m[3]_t$  é o autovetor correspondente ao maior autovalor da matriz  $-\ddot{F}$ .

A Tabela 4.2.3.1 apresenta para  $t = 12, 34, 35, 36, 37, 38$  e  $39$  os valores da curvatura normal conforme  $B_{\underline{E}_t}$  do vetor de perturbação básico  $\underline{E}_t$  e a contribuição agregada  $m[q]_t$  para todos os autovetores  $q$ -influentes com  $q$  variando de 1 a 3. Note-se que para as profundidades de água das observações  $t = 35, 36$  e  $37$  as medidas  $m[q]_t$  e  $B_{\underline{E}_t}$  estão acima dos valores críticos respectivos, ou seja, são influentes e para  $q = 0, 1$  e  $2$ , a observação 39 também é considerada como influente, sendo a mais influente a observação 37. Por outro lado as observações:  $t = 34$  para todo  $q$ , a observação 12 para  $q = 0$  e  $q = 1$ , e a observação 38 para  $q = 0, 1$  e  $2$ , são incluídas como marginalmente influentes. Assim, para este esquema de perturbação, as observações com grandes profundidades são as encontradas como influentes, ou marginalmente influentes.



**Tabela 4.2.3.1:** Medidas de observações influentes para o *esquema de perturbação III*

Medidas	$N(\underline{\rho})$	$\bar{m}[q]$	Valores críticos	Valores das seguintes obs.			
				obs 12	obs 34	obs 35	obs 36
$B_{E_t}$	39	0.0410	0.0820	0.1010	0.0972	0.1655	0.1264
$m[1]_t$	3	0.1959	0.2760	0.3148	0.2957	0.3856	0.3512
$m[2]_t$	2	0.1786	0.2525	0.1033	0.2954	0.3846	0.3508
$m[3]_t$	1	0.1524	0.2156	0.0797	0.2953	0.3785	0.3465

Medidas	$N(\underline{\rho})$	$\bar{m}[q]$	Valores críticos	Valores das seguintes obs.			
				obs 37	obs 38	obs 39	
$B_{E_t}$	39	0.0410	0.0820	0.5002	0.0848	0.2067	
$m[1]_t$	3	0.1959	0.2760	0.6944	0.2866	0.4504	
$m[2]_t$	2	0.1786	0.2525	0.6933	0.2799	0.4302	
$m[3]_t$	1	0.1524	0.2156	0.6923	0.0092	0.0526	

$N(\underline{\rho})$  representa a quantidade de autovetores influentes

**Figura 4.2.3.1:** Gráfico de  $B_{E_t}$  contra  $x_2$  para o *esquema de perturbação III*

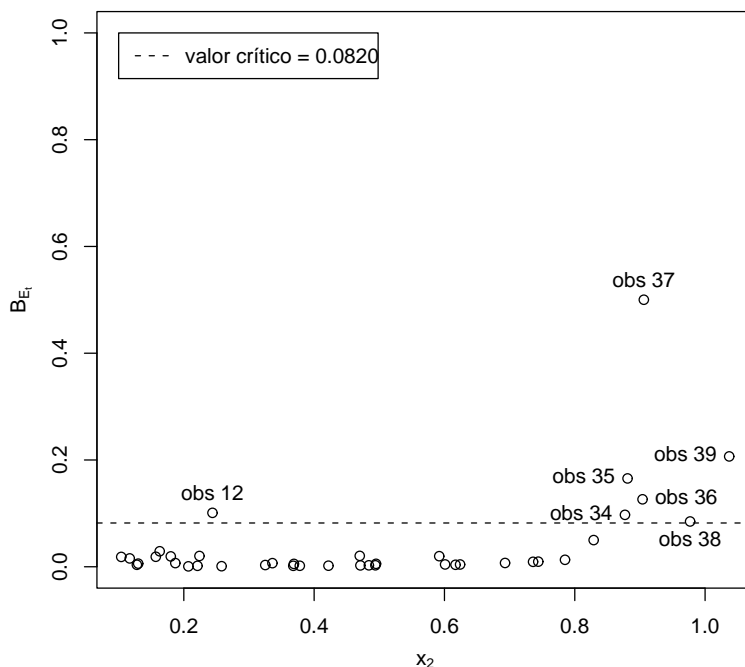


Figura 4.2.3.2: Gráfico de  $m[1]_t$  contra  $x_2$  para o esquema de perturbação III

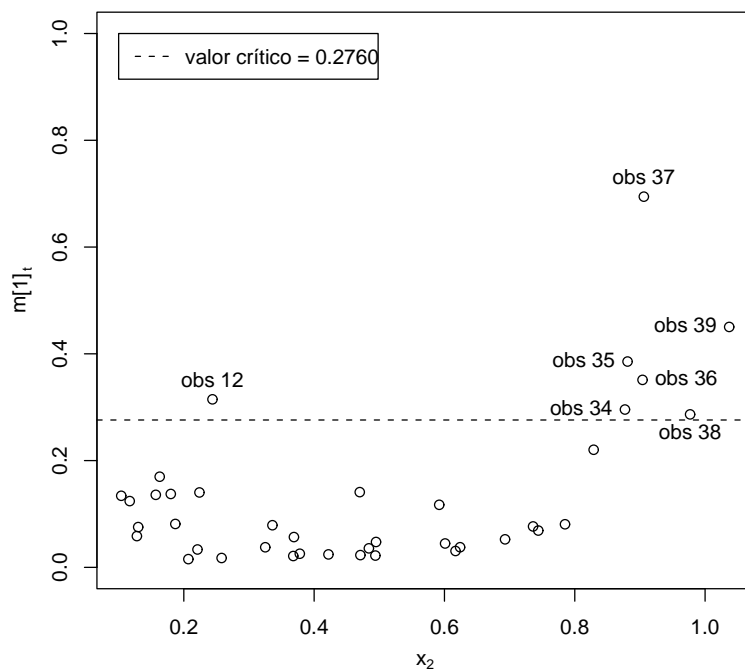
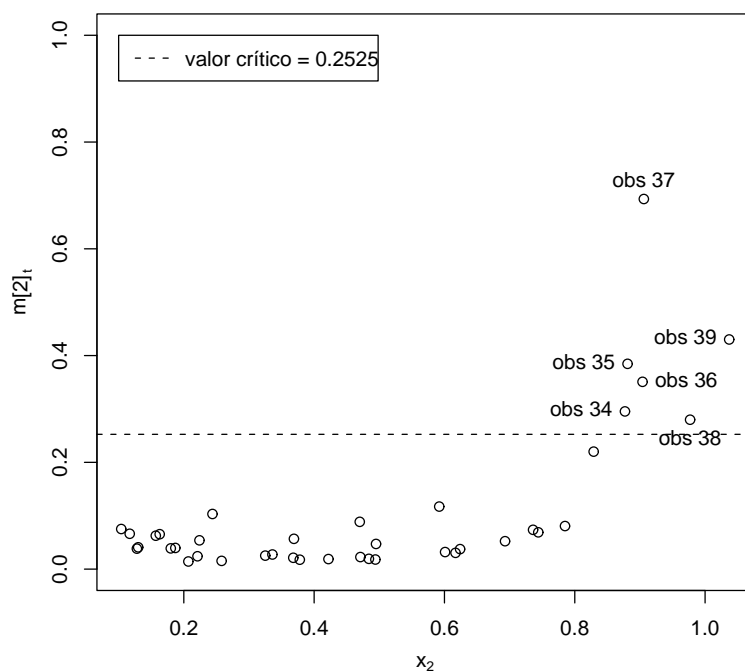
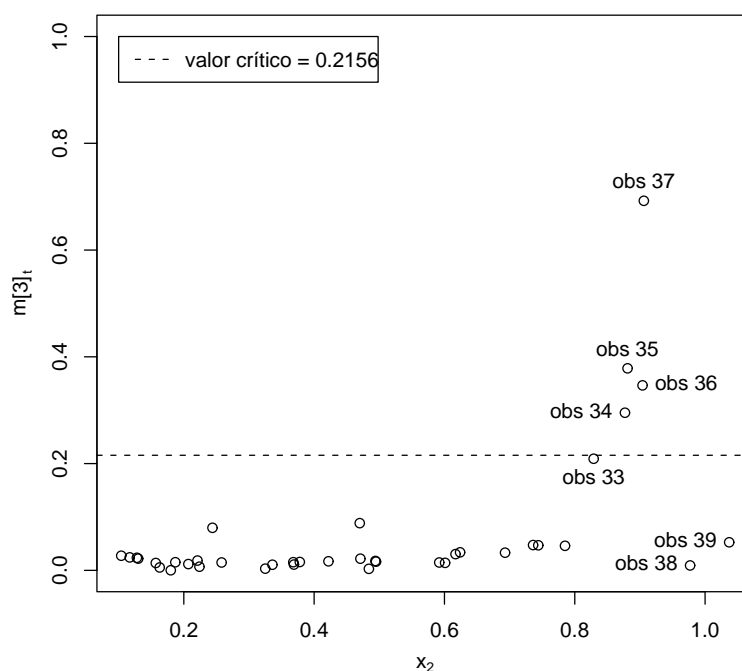


Figura 4.2.3.3: Gráfico de  $m[2]_t$  contra  $x_2$  para o esquema de perturbação III



**Figura 4.2.3.4:** Gráfico de  $m[3]_t$  contra  $x_2$  para o esquema de perturbação III

Nas Figuras 4.2.3.1-4.2.3.4 estão apresentados os gráficos das curvaturas normais conformes  $B_{E_t}$  e as medidas  $m[q]_t$  contra a variável explicativa  $x_2$ , que corresponde às profundidades de água, para o esquema de perturbação III.

Continuando com a mesma metodologia, ao tirar a observação mais influente, isto é, a observação 37 da amostra com 39 observações, se encontraram três autovetores 1-influente, dois autovetores 2-influente e somente um autovetor 3-influente, a saber, o autovetor  $e_{max}$  associado ao maior autovalor da matriz  $-\tilde{F}$ . Além disso, para todas as medidas, as observações 35 e 36 são consideradas como influente, ao passo que as observações 34 para  $q = 1, 2, 3$  e a observação 39 para  $q = 0, 1, 2$ , são incluídas como influente. Por outro lado, as observações: 33 para todo  $q$ , a observação 12 para  $q = 0, 1$ , e a observação 38 para  $q = 0$ , são encontradas como marginalmente influente. A Tabela 4.2.3.2 apresenta estes resultados.

**Tabela 4.2.3.2:** Medidas de observações influentes para o *esquema de perturbação III* tirando a observação 37.

Medidas	$N(\underline{\rho})$	$\bar{m}[q]$	Valores críticos	Valores das seguintes obs.			
				obs 12	obs 33	obs 34	obs 35
$B_{\underline{\rho}t}$	38	0.0434	0.0867	0.1035	0.0906	0.1789	0.3015
$m[1]_t$	3	0.2031	0.2872	0.3182	0.2977	0.4217	0.5477
$m[2]_t$	2	0.1863	0.2635	0.0869	0.2972	0.4216	0.5475
$m[3]_t$	1	0.1487	0.2103	0.0869	0.2951	0.4190	0.5269

Medidas	$N(\underline{\rho})$	$\bar{m}[q]$	Valores críticos	Valores das seguintes obs.			
				obs 36	obs 38	obs 39	
$B_{\underline{\rho}t}$	38	0.0434	0.0867	0.2440	0.1205	0.3320	
$m[1]_t$	3	0.2031	0.2872	0.4933	0.3391	0.5719	
$m[2]_t$	2	0.1863	0.2635	0.4924	0.3390	0.5689	
$m[3]_t$	1	0.1487	0.2103	0.4923	0.0447	0.1376	

$N(\underline{\rho})$  representa a quantidade de autovetores influentes

Ao tirar as observações 35, 36 e 37 da amostra completa, e ficando, assim, com uma amostra de tamanho 36, encontram-se como influentes as observações 33, 34, 38 e 39. Além disso, somente para  $q = 0$  e  $q = 1$ , a observação 12 é incluída como marginalmente influente. Estes resultados estão apresentados na Tabela 4.2.3.3.

**Tabela 4.2.3.3:** Medidas de observações influentes para o *esquema de perturbação III* tirando as observações 35, 36 e 37.

Medidas	$N(\underline{\rho})$	$\bar{m}[q]$	Valores críticos	Valores das seguintes obs.				
				obs 12	obs 33	obs 34	obs 38	obs 39
$B_{\underline{\rho}t}$	36	0.0456	0.0912	0.0903	0.2131	0.5049	0.1496	0.4138
$m[1]_t$	3	0.2086	0.2950	0.2966	0.4597	0.7092	0.3779	0.6385
$m[2]_t$	2	0.1939	0.2741	0.0841	0.4596	0.7083	0.3778	0.6370
$m[3]_t$	2	0.1939	0.2741	0.0841	0.4596	0.7083	0.3778	0.6370

$N(\underline{\rho})$  representa a quantidade de autovetores influentes

Na Tabela 4.2.3.4 estão apresentadas as estimativas, o desvio padrão e o desvio relativo percentual  $DRP_{hj}$  dos  $\beta_{hj}$  tirando as observações influentes para o *esquema de perturbação III*. Similarmente aos casos anteriores, as estimativas dos parâmetros mu-

ram mais quando se retiraram conjuntamente as observações encontradas como mais influentes que quando se retirou somente a observação 37. A estimativa do parâmetro  $\beta_{11}$  foi a que apresentou maior impacto, chegando a mudar de sinal com uma variação de -110.31%, como no esquema anterior, as estimativas correspondentes ao intercepto apresentaram menor desvio padrão que as estimativas dos parâmetros correspondentes à variável explicativa. Finalmente testamos a significância dos parâmetros correspondentes ao intercepto e também à variável explicativa a um nível de 5% de confiança e pudemos verificar que em todos os casos a hipótese nula é rejeitada, ou seja, os parâmetros do modelo são significativos.

**Tabela 4.2.3.4:** Estimativas, desvio padrão e  $DRP_{hj}$  dos  $\beta_{hj}$  com a amostra completa e tirando as observações influentes.

	$\hat{\beta}_{hj}$ , $dp(\hat{\beta}_{hj})$ e $DRP_{hj}$ tirando as obs		
	nenhuma	obs 37	obs 35, 36, 37,
$\hat{\beta}_{11}$	0.1457	0.0637	-0.0150
$dp(\hat{\beta}_{11})$	0.3245	0.3261	0.3312
$DRP_{11}$		-56.24%	-110.31%
$\hat{\beta}_{12}$	-0.2876	-0.3818	-0.5295
$dp(\hat{\beta}_{12})$	0.3147	0.3154	0.3187
$DRP_{12}$		32.76%	84.09%
$\hat{\beta}_{13}$	-1.1349	-1.2081	-1.3150
$dp(\hat{\beta}_{13})$	0.3109	0.3119	0.3156
$DRP_{13}$		6.45%	15.87%
$\hat{\beta}_{21}$	2.2590	2.5219	2.8193
$dp(\hat{\beta}_{21})$	0.5778	0.5942	0.6350
$DRP_{21}$		11.64%	24.80%
$\hat{\beta}_{22}$	5.4886	5.8109	6.3459
$dp(\hat{\beta}_{22})$	0.5664	0.5821	0.6213
$DRP_{22}$		5.87%	15.62%
$\hat{\beta}_{23}$	6.3690	6.6957	7.1817
$dp(\hat{\beta}_{23})$	0.5631	0.5791	0.6187
$DRP_{23}$		5.13%	12.76%

$dp(\hat{\beta}_{hj})$  representa o desvio padrão de  $\hat{\beta}_{hj}$

A Tabela 4.2.3.5 mostra o  $DQM_j$ , o  $EAM_j$  e o  $EPAM_j$  considerando a retirada individual e coletiva das observações influentes obtidas no *esquema de perturbação III*. Note-se que as medidas dos erros foram maiores quando se tiraram simultaneamente as observações encontradas como as mais influentes do que quando se retirou somente a observação influente 37, a medida de erro maior se apresentou para a primeira coordenada.

**Tabela 4.2.3.5:**  $DQM_j$ ,  $EAM_j$  e  $EPAM_j$  para os  $\hat{y}_j$ , sob *esquema de perturbação III*.

	Medidas de erro para os $\hat{y}_j$ sem observações					
	obs 37			obs 35, 36, 37		
	1 Coord.	2 Coord.	3 Coord.	1 Coord.	2 Coord.	3 Coord.
$DQM_j$	0.0025	0.0032	0.0055	0.0077	0.0060	0.0098
$EAM_j$	0.0023	0.0030	0.0053	0.0070	0.0047	0.0095
$EPAM_j$	2.15%	0.63%	1.69%	6.31%	1.07%	3.16%

#### 4.2.4 Esquema de perturbação IV

No *esquema de perturbação IV* a dimensão do espaço de perturbação  $\Omega$  é  $39 \times 3 = 117$ . Neste esquema, perturba-se as variáveis resposta que correspondem às proporções de areia, lodo e argila dos sedimentos. Os autovalores normalizados não nulos da matriz  $QA$  descrita na seção 3.4 são 0.9983, 0.0524, 0.0272, 0.0044, 0.0021 e 0.000052. Então os autovetores unitários associados a estes autovalores são as direções que maximizam a curvatura sujeita às restrições dadas no *esquema de perturbação IV*, como foi visto na seção 3.4.

Para este conjunto de dados foi encontrado um único autovetor 1-influente, e, para  $q = 2$  e  $q = 3$ , não foram encontrados autovetores  $q$ -influentes quando usada a amostra com as 39 observações.

Na Figura 4.2.4.1 está apresentado o gráfico das normas dos vetores  $\bar{z}_{1t}$  definidos na seção 3.4, contra a variável explicativa  $x_2$  que corresponde às profundidades de água, para o *esquema de perturbação IV*. Note-se que para as profundidades de água das observações  $t = 37$  e  $t = 39$  as normas  $\|\bar{z}_{1t}\|$  estão acima do valor crítico  $\sqrt{\frac{2}{39}}$ , ou seja,

são influentes, isto também pode ser observado na Tabela 4.2.4.1.

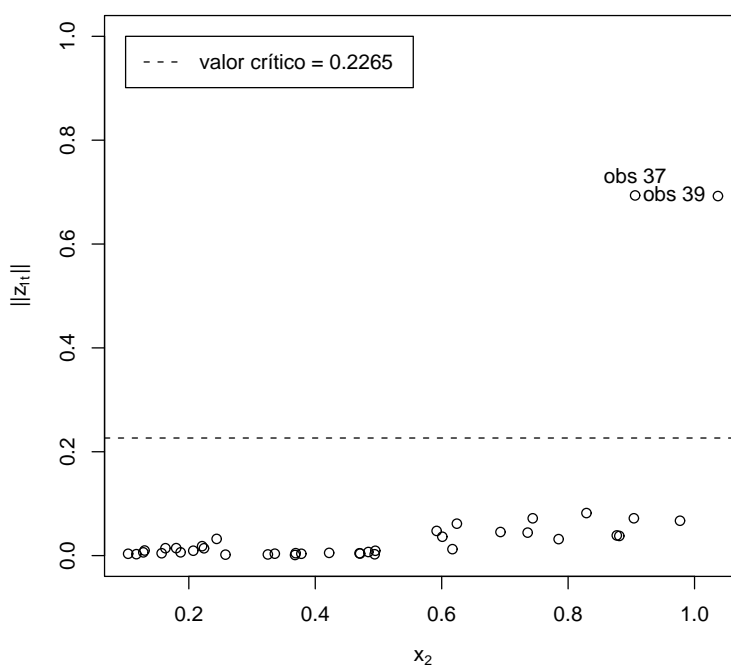
**Tabela 4.2.4.1:** Medidas de observações influentes para o esquema de perturbação IV

Medidas	$N(\underline{\rho})$	$\bar{m}[q]$	Valores críticos	Valores das obs.	
				obs 37	obs 39
$m[0]_t$	117	0.1667	0.2358	0.7026	0.7029
$m[1]_t$	1	0.1600	0.2263	0.6930	0.6918

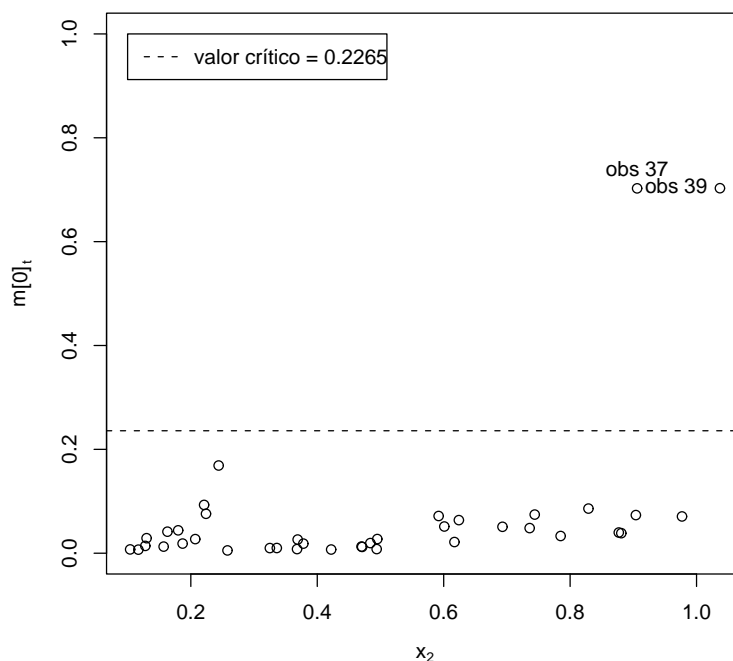
$N(\underline{\rho})$  representa a quantidade de autovetores influentes

A Figura 4.2.4.2 apresenta o gráfico de  $m[0]_t$  contra a variável explicativa  $x_2$  para o esquema de perturbação IV. Note-se que para esta medida as observações  $t = 37$  e  $t = 39$  ficaram acima do valor crítico, logo são consideradas como observações influentes. Além disso, dados os valores das medidas  $m[0]_t, m[1]_t$  pode-se afirmar que têm aproximadamente o mesmo nível de influência.

**Figura 4.2.4.1:** Gráfico de  $\|z_{1t}\|$  contra  $x_2$  para o esquema de perturbação IV



**Figura 4.2.4.2:** Gráfico de  $m[0]_t$  contra  $x_2$  para o esquema de perturbação IV



Fazendo a mesma análise que nos esquemas de perturbação anteriores, ao tirar as observações 37 e 39 conjuntamente, obtemos como influentes as observações 1, 2 e 38. Neste caso, se encontrou unicamente o autovetor unitário associado ao maior autovalor  $\tau_{max}$  da matriz  $QA$  como autovetor 1-influente, 2-influente e 3-influente. A Tabela 4.2.4.2 apresenta os valores das medidas para estas observações influentes.

**Tabela 4.2.4.2:** Medidas de observações influentes para o esquema de perturbação IV tirando as observações 37 e 39.

Medidas	$N(\underline{\epsilon})$	$\bar{m}[q]$	Valores críticos	Valores das seguintes obs.		
				obs 1	obs 2	obs 38
$m[0]_t$	111	0.1690	0.2389	0.4932	0.3330	0.7707
$m[1]_t$	1	0.1643	0.2324	0.4857	0.3311	0.7628
$m[2]_t$	1	0.1643	0.2324	0.4857	0.3311	0.7628
$m[3]_t$	1	0.1643	0.2324	0.4857	0.3311	0.7628

$N(\underline{\epsilon})$  representa a quantidade de autovetores influentes



Finalmente ao tirar conjuntamente as observações encontradas como as mais influentes, ou seja, ao tirar as observações 37, 38 e 39 simultaneamente da amostra completa, não foram encontrados autovetores 1-influente, 2-influente nem 3-influente. Neste caso, as observações  $t = 10, 12, 25, 28, 31, 33$  e  $36$  foram incluídas como influentes. A Tabela 4.2.4.3 apresenta estes resultados.

**Tabela 4.2.4.3:** Medidas de observações influentes para o *esquema de perturbação IV* tirando as observações 37, 38 e 39.

Medidas	$N(\underline{g})$	$\bar{m}[q]$	Valores críticos	Valores das seguintes obs.			
				obs 10	obs 12	obs 25	obs 28
$m[0]_t$	108	0.1985	0.2807	0.2941	0.5333	0.2907	0.3195
				obs 31	obs 33	obs 36	
				0.3595	0.4047	0.3746	

$N(\underline{g})$  representa a quantidade de autovetores influentes

Na Tabela 4.2.4.4 estão apresentadas as estimativas, o desvio padrão e o desvio relativo percentual  $DRP_{hj}$  dos  $\beta_{hj}$  tirando as observações influentes para o *esquema de perturbação IV*. Diferentemente dos esquemas de perturbação anteriores, em geral as estimativas dos parâmetros tiveram maior impacto com a retirada inicial das observações influentes 37 e 39 que com a retirada simultânea das observações 37, 38 e 39. Além disso, a variação das estimativas, segundo o  $DRP_{ij}$ , não foi tão alta (-38.91%) para o parâmetro  $\beta_{11}$  como foi verificado nos esquemas de perturbação anteriores. Analogamente ao procedimento aplicado nos esquemas anteriores, fizemos testes de hipóteses para provar a significância dos parâmetros, e verificamos que os parâmetros do modelo são significativos.

**Tabela 4.2.4.4:** Estimativas, desvio padrão e  $DRP_{hj}$  dos  $\beta_{hj}$  com a amostra completa e tirando as observações influentes.

	$\hat{\beta}_{hj}$ , $dp(\hat{\beta}_{hj})$ e $DRP_{hj}$ tirando as obs		
	nenhuma	obs 37, 39	obs 37, 38, 39
$\hat{\beta}_{11}$	0.1457	0.0913	0.0890
$dp(\hat{\beta}_{11})$	0.3245	0.3361	0.3456
$DRP_{11}$		-37.33%	-38.91%
$\hat{\beta}_{12}$	-0.2876	-0.2923	-0.2360
$dp(\hat{\beta}_{12})$	0.3147	0.3270	0.3379
$DRP_{12}$		1.61%	-17.94%
$\hat{\beta}_{13}$	-1.1349	-1.1875	-1.1757
$dp(\hat{\beta}_{13})$	0.3109	0.3224	0.3324
$DRP_{13}$		4.63%	3.59%
$\hat{\beta}_{21}$	2.2590	2.4385	2.4139
$dp(\hat{\beta}_{21})$	0.5778	0.6349	0.6761
$DRP_{21}$		7.95%	6.86%
$\hat{\beta}_{22}$	5.4886	5.6056	5.4539
$dp(\hat{\beta}_{22})$	0.5664	0.6235	0.6654
$DRP_{22}$		2.13%	-0.63%
$\hat{\beta}_{23}$	6.3690	6.5919	6.5114
$dp(\hat{\beta}_{23})$	0.5631	0.6193	0.6601
$DRP_{23}$		3.50%	2.24%

$dp(\hat{\beta}_{hj})$  representa o desvio padrão de  $\hat{\beta}_{hj}$

A Tabela 4.2.4.5 mostra o  $DQM_j$ , o  $EAM_j$  e o  $EPAM_j$  considerando a retirada coletiva das observações influentes obtidas no *esquema de perturbação IV*. Como nos esquemas de perturbação anteriores, as medidas dos erros foram maiores quando se tiraram simultaneamente as observações encontradas como as mais influentes que quando se retirou somente as observações influentes 37 e 39.

**Tabela 4.2.4.5:**  $DQM_j$ ,  $EAM_j$  e  $EPAM_j$  para os  $\hat{y}_j$ , sob esquema de perturbação IV.

	Medidas de erro para os $\hat{y}_j$ sem observações					
	obs 37, 39			obs 37, 38, 39		
	1 Coord.	2 Coord.	3 Coord.	1 Coord.	2 Coord.	3 Coord.
$DQM_j$	0.0045	0.0071	0.0053	0.0080	0.0125	0.0064
$EAM_j$	0.0038	0.0063	0.0038	0.0059	0.0110	0.0051
$EPAM_j$	1.74%	1.37%	1.06%	2.21%	2.44%	1.62%

Vale a pena ressaltar que ao ajustar o modelo de regressão Dirichlet para este conjunto de dados, a observação 37 foi encontrada como influente para os quatro esquemas de perturbação, sendo a mais influente em dois dos esquemas de perturbação.

# Capítulo 5

## Conclusões

Ao longo desta dissertação consideramos o modelo de regressão Dirichlet, recentemente proposto em Silva (2004), e a curvatura normal conforme proposta por Poon & Poon (1999). Encontramos dois resultados teóricos e um resultado prático que podemos descrever como segue. No lado teórico, no Capítulo 3, conseguimos estender a análise de influência apresentada por Poon & Poon (1999) em dois sentidos. O primeiro resultado nos permite encontrar a curvatura máxima no caso em que existem restrições lineares nas perturbações; neste sentido, enunciamos um Teorema e apresentamos a demonstração que prova sua validade. O segundo resultado teórico que apresentamos é a proposta da definição de contribuição agregada conjunta para modelos multivariados. No lado prático, no Capítulo 4, quando analisamos o conjunto de dados reais com o modelo de regressão Dirichlet, obtivemos uma observação que se destaca das outras em todos os esquemas de perturbação que consideramos, esta observação corresponde à observação 37. Cabe ainda destacar que para o conjunto de dados analisados as observações mais influentes de acordo com nossos resultados são aquelas que correspondem às maiores profundidades de água. Isto representa uma indicação de que o modelo é menos adequado para maiores profundidades.

Como trabalho futuro que poderia estender esta pesquisa, sugerimos encontrar a curvatura máxima sujeita a restrições lineares não homogêneas da forma  $Cz = \underline{b}$ , onde  $\underline{b} \neq 0$ . Por exemplo, este seria o caso em que perturbaríamos as observações multiplicativamente, neste caso, a matriz  $C$  incluiria as observações e teríamos  $\underline{b}$  como um

vetor de uns. Outra proposta é estudar a maximização quando as perturbações estão sujeitas a restrições de desigualdade.

# Apêndice A

## Derivadas

Neste apêndice encontram-se os cálculos das derivadas da função de log-verossimilhança perturbada com relação aos parâmetros desconhecidos e aos parâmetros de perturbação, com o intuito de encontrar as matrizes  $\check{L}$  e  $\Delta$  para os diferentes esquemas de perturbação vistos no Capítulo 3. As derivadas de primeira e segunda ordem com relação aos parâmetros  $\beta_{vu}$  necessárias para encontrar a matriz  $\check{L}$  estão apresentadas detalhadamente no Apêndice 2 de Silva (2004). Assim, neste apêndice somente serão apresentados os cálculos para encontrar a matriz  $\Delta$ .

A função de log-verossimilhança para uma amostra de  $N$  observações, com  $k$  covariáveis e  $p$  variáveis resposta através do modelo de regressão Dirichlet, é

$$\ell(\mathbf{B}) = \sum_{i=1}^N \left\{ \log \Gamma(\phi_i) - \sum_{j=1}^p \log(\Gamma(\phi_i \mu_{ij})) + \sum_{j=1}^p (\phi_i \mu_{ij} - 1) \log y_{ij} \right\},$$

onde  $\phi_i = \sum_{j=1}^p g(\eta_{ij})$ ,  $\mu_{ij} = \frac{g(\eta_{ij})}{\phi_i}$ ,  $\eta_{ij} = x_{i1}\beta_{1j} + \dots + x_{ik}\beta_{kj}$ , e,  $g(\eta_{ij}) = \exp(\eta_{ij})$ .

## A.1 Esquema de perturbação I

Neste esquema a função de log-verossimilhança perturbada é

$$\ell(\mathbf{B}|\mathbf{W}) = \sum_{i=1}^N w_i \left\{ \log \Gamma(\phi_i) - \sum_{j=1}^p \log(\Gamma(\phi_i \mu_{ij})) + \sum_{j=1}^p (\phi_i \mu_{ij} - 1) \log y_{ij} \right\}$$

e

$$\Delta_{vu,n} = \frac{\partial}{\partial \beta_{vu}} \left( \frac{\partial \ell(\mathbf{B}|\mathbf{W})}{\partial w_n} \right)$$

onde  $v = 1, \dots, k, u = 1, \dots, p, n = 1, \dots, N$  e  $\mathbf{W} = (w_1, \dots, w_n)$ .

A derivada com relação aos parâmetros de perturbação é

$$\frac{\partial \ell(\mathbf{B}|\mathbf{W})}{\partial w_n} = \log \Gamma(\phi_n) - \sum_{j=1}^p \log(\Gamma(\phi_n \mu_{nj})) + \sum_{j=1}^p (\phi_n \mu_{nj} - 1) \log y_{nj}$$

e usando as derivadas encontradas no Apêndice 2 de Silva (2004), tem-se que

$$\begin{aligned} \frac{\partial}{\partial \beta_{vu}} \left( \frac{\partial \ell(\mathbf{B}|\mathbf{W})}{\partial w_n} \right) &= \frac{\partial}{\partial \beta_{vu}} \left[ \log \Gamma(\phi_n) - \sum_{j=1}^p \log(\Gamma(\phi_n \mu_{nj})) + \sum_{j=1}^p (\phi_n \mu_{nj} - 1) \log y_{nj} \right] \\ &= x_{nv} g'(\eta_{nu}) [\psi(\phi_n) + A_{nu}] \end{aligned}$$

em que  $A_{ij} = \log y_{ij} - \psi(\phi_i \mu_{ij})$ .

Então,

$$\Delta_{vu,n} \Big|_{\mathbf{W}=\mathbf{1}, \mathbf{B}=\hat{\mathbf{B}}} = x_{nv} g'(\hat{\eta}_{nu}) [\psi(\hat{\phi}_n) + \hat{A}_{nu}]. \quad (\text{A.1})$$

## A.2 Esquema de perturbação II

Neste caso perturba-se as variáveis de forma multiplicativa, ou seja a matriz  $\mathbf{X}$  é substituída pela matriz  $\mathbf{X}_w = \mathbf{W}\mathbf{X}$ , onde  $\mathbf{W} = \text{diag}\{w_1, \dots, w_N\}$ , ou seja,  $\mathbf{X}_w$  é

$$\mathbf{X}_w = \begin{pmatrix} w_1 x_{11} & w_1 x_{12} & \cdots & w_1 x_{1k} \\ w_2 x_{21} & w_2 x_{22} & \cdots & w_2 x_{2k} \\ \vdots & \vdots & & \vdots \\ w_N x_{N1} & w_N x_{N2} & \cdots & w_N x_{Nk} \end{pmatrix}.$$

$$\ell(\mathbf{B}|\mathbf{W}) = \sum_{i=1}^N \left\{ \log \Gamma(\phi_i) - \sum_{j=1}^p \log(\Gamma(\phi_i \mu_{ij})) + \sum_{j=1}^p (\phi_i \mu_{ij} - 1) \log y_{ij} \right\}$$

onde  $\phi_i$  e  $\mu_{ij}$  são definidos como no esquema de perturbação I, mas  $\eta_{ij} = w_i x_{i1} \beta_{1j} + \cdots + w_i x_{ik} \beta_{kj}$ .

Como no esquema de perturbação anterior a matriz  $\Delta$  é

$$\Delta_{vu,n} = \frac{\partial}{\partial w_n} \left( \frac{\partial \ell(\mathbf{B}|\mathbf{W})}{\partial \beta_{vu}} \right)$$

A derivada de primeira ordem da função de log-verossimilhança perturbada com relação aos parâmetros desconhecidos para  $v = 1, \dots, k, u = 1, \dots, p$  e  $n = 1, \dots, N$  é

$$\begin{aligned} \frac{\partial \ell(\mathbf{B}|\mathbf{W})}{\partial \beta_{vu}} &= \sum_{i=1}^N \left\{ \frac{\partial}{\partial \beta_{vu}} \log \Gamma(\phi_i) + \sum_{j=1}^p \left[ \frac{\partial}{\partial \beta_{vu}} ((\phi_i \mu_{ij} - 1) \log y_{ij}) - \frac{\partial}{\partial \beta_{vu}} \log(\Gamma(\phi_i \mu_{ij})) \right] \right\} \\ &= \sum_{i=1}^N \left\{ \psi(\phi_i) \frac{\partial \phi_i}{\partial \beta_{vu}} + \sum_{j=1}^p \left[ \log y_{ij} \frac{\partial}{\partial \beta_{vu}} (\phi_i \mu_{ij}) - \psi(\phi_i \mu_{ij}) \frac{\partial}{\partial \beta_{vu}} (\phi_i \mu_{ij}) \right] \right\} \\ &= \sum_{i=1}^N \left\{ \psi(\phi_i) \frac{\partial \phi_i}{\partial \beta_{vu}} + \sum_{j=1}^p [\log y_{ij} - \psi(\phi_i \mu_{ij})] \frac{\partial}{\partial \beta_{vu}} (\phi_i \mu_{ij}) \right\} \\ &= \sum_{i=1}^N \left\{ \psi(\phi_i) \frac{\partial \phi_i}{\partial \beta_{vu}} + \sum_{j=1}^p A_{ij} \frac{\partial}{\partial \beta_{vu}} (\phi_i \mu_{ij}) \right\}, \end{aligned}$$

onde  $A_{ij} = \log y_{ij} - \psi(\phi_i \mu_{ij})$  Tem-se que

$$\frac{\partial}{\partial \beta_{vu}} (\phi_i \mu_{ij}) = \frac{\partial \phi_i}{\partial \beta_{vu}} \mu_{ij} + \phi_i \frac{\partial \mu_{ij}}{\partial \beta_{vu}} = \frac{\partial \phi_i}{\partial \beta_{vu}} \mu_{ij} + \phi_i \frac{\partial \mu_{ij}}{\partial \eta_{iu}} \frac{\partial \eta_{iu}}{\partial \beta_{vu}} = \frac{\partial \phi_i}{\partial \beta_{vu}} \mu_{ij} + \phi_i \frac{\partial \mu_{ij}}{\partial \eta_{iu}} w_i x_{iv}$$



e

$$\frac{\partial \phi_i}{\partial \beta_{vu}} = \sum_{j=1}^p \frac{\partial}{\partial \beta_{vu}} g(\eta_{ij}) = \sum_{\substack{j=1 \\ j \neq u}}^p g'(\eta_{ij}) \frac{\partial \eta_{ij}}{\partial \beta_{vu}} + g'(\eta_{iu}) \frac{\partial \eta_{iu}}{\partial \beta_{vu}} = g'(\eta_{iu}) w_i x_{iv}.$$

Assim

$$\frac{\partial}{\partial \beta_{vu}} (\phi_i \mu_{ij}) = g'(\eta_{iu}) w_i x_{iv} \mu_{ij} + \phi_i w_i x_{iv} \frac{\partial \mu_{ij}}{\partial \eta_{iu}} = w_i x_{iv} \left( g'(\eta_{iu}) \mu_{ij} + \phi_i \frac{\partial \mu_{ij}}{\partial \eta_{iu}} \right).$$

Por outro lado

$$\frac{\partial \mu_{ij}}{\partial \eta_{iu}} = \begin{cases} \frac{g'(\eta_{iu})(\phi_i - g(\eta_{iu}))}{\phi_i^2}, & \text{se } j = u, \\ -\frac{g(\eta_{ij})g'(\eta_{iu})}{\phi_i^2}, & \text{se } j \neq u. \end{cases} \quad (\text{A.2})$$

Portanto substituindo os resultados anteriores,

$$\begin{aligned} \frac{\partial \ell(\mathbf{B}|\mathbf{W})}{\partial \beta_{vu}} &= \sum_{i=1}^N \left\{ \psi(\phi_i) g'(\eta_{iu}) w_i x_{iv} + \sum_{j=1}^p A_{ij} w_i x_{iv} \left( g'(\eta_{iu}) \mu_{ij} + \phi_i \frac{\partial \mu_{ij}}{\partial \eta_{iu}} \right) \right\} \\ &= \sum_{i=1}^N w_i x_{iv} \left\{ \psi(\phi_i) g'(\eta_{iu}) + \sum_{j=1}^p A_{ij} g'(\eta_{iu}) \frac{g(\eta_{ij})}{\phi_i} \right. \\ &\quad \left. + \sum_{\substack{j=1 \\ j \neq u}}^p A_{ij} \phi_i \left( \frac{-g(\eta_{ij})g'(\eta_{iu})}{\phi_i^2} \right) + A_{iu} \phi_i \frac{g'(\eta_{iu})(\phi_i - g(\eta_{iu}))}{\phi_i^2} \right\} \\ &= \sum_{i=1}^N w_i x_{iv} \frac{g'(\eta_{iu})}{\phi_i} \left\{ \phi_i \psi(\phi_i) + \sum_{j=1}^p A_{ij} g(\eta_{ij}) - \sum_{\substack{j=1 \\ j \neq u}}^p A_{ij} g(\eta_{ij}) + A_{iu} (\phi_i - g(\eta_{iu})) \right\} \\ &= \sum_{i=1}^N w_i x_{iv} \frac{g'(\eta_{iu})}{\phi_i} \left\{ \phi_i \psi(\phi_i) + \sum_{j=1}^p A_{ij} g(\eta_{ij}) - \sum_{j=1}^p A_{ij} g(\eta_{ij}) + A_{iu} \phi_i \right\} \\ &= \sum_{i=1}^N w_i x_{iv} \frac{g'(\eta_{iu})}{\phi_i} \{ \phi_i \psi(\phi_i) + A_{iu} \phi_i \} \\ &= \sum_{i=1}^N w_i x_{iv} g'(\eta_{iu}) [\psi(\phi_i) + A_{iu}]. \end{aligned}$$

Logo

$$\frac{\partial \ell(\mathbf{B}|\mathbf{W})}{\partial \beta_{vu}} = \sum_{i=1}^N w_i x_{iv} g'(\eta_{iu}) [\psi(\phi_i) + A_{iu}].$$

Dessa forma,

$$\begin{aligned}
 \frac{\partial}{\partial w_n} \left( \frac{\partial \ell(\mathbf{B}|\mathbf{W})}{\partial \beta_{vu}} \right) &= \frac{\partial}{\partial w_n} \left( \sum_{i=1}^N w_i x_{iv} g'(\eta_{iu}) [\psi(\phi_i) + A_{iu}] \right) \\
 &= \sum_{\substack{i=1 \\ i \neq n}}^N w_i \frac{\partial}{\partial w_n} x_{iv} g'(\eta_{iu}) [\psi(\phi_i) + A_{iu}] + \frac{\partial}{\partial w_n} w_n x_{nv} g'(\eta_{nu}) [\psi(\phi_n) + A_{nu}] \\
 &= \sum_{\substack{i=1 \\ i \neq n}}^N w_i \frac{\partial}{\partial w_n} x_{iv} g'(\eta_{iu}) [\psi(\phi_i) + A_{iu}] + x_{nv} \left( g'(\eta_{nu}) [\psi(\phi_n) + A_{nu}] \right. \\
 &\quad \left. + w_n \frac{\partial}{\partial w_n} g'(\eta_{nu}) [\psi(\phi_n) + A_{nu}] \right) \\
 &= \sum_{i=1}^N w_i \frac{\partial}{\partial w_n} x_{iv} g'(\eta_{iu}) [\psi(\phi_i) + A_{iu}] + x_{nv} g'(\eta_{nu}) [\psi(\phi_n) + A_{nu}].
 \end{aligned}$$

onde

$$\begin{aligned}
 \sum_{i=1}^N w_i \frac{\partial}{\partial w_n} x_{iv} g'(\eta_{iu}) [\psi(\phi_i) + A_{iu}] &= \\
 \sum_{i=1}^N w_i x_{iv} \left\{ g''(\eta_{iu}) \frac{\partial \eta_{iu}}{\partial w_n} [\psi(\phi_i) + A_{iu}] + g'(\eta_{iu}) \left[ \psi'(\phi_i) \frac{\partial \phi_i}{\partial w_n} + \frac{\partial A_{iu}}{\partial w_n} \right] \right\}
 \end{aligned}$$

e

$$\begin{aligned}
 \frac{\partial \eta_{ij}}{\partial w_n} &= \begin{cases} \sum_{h=1}^k x_{ih} \beta_{hj}, & \text{se } i = n, \\ 0, & \text{se } i \neq n. \end{cases} \\
 \frac{\partial \phi_i}{\partial w_n} &= \begin{cases} \sum_{h=1}^k \sum_{j=1}^p g'(\eta_{ij}) x_{nh} \beta_{hj}, & \text{se } i = n, \\ 0, & \text{se } i \neq n. \end{cases}
 \end{aligned}$$

além disso,

$$\begin{aligned}
 \frac{\partial A_{iu}}{\partial w_n} &= \frac{\partial}{\partial w_n} \left( \log y_{iu} - \psi(\phi_i \mu_{iu}) \right) = -\psi'(\phi_i \mu_{iu}) \frac{\partial}{\partial w_n} (\phi_i \mu_{iu}) \\
 &= -\psi'(\phi_i \mu_{iu}) \left( \frac{\partial \phi_i}{\partial w_n} \mu_{iu} + \phi_i \frac{\partial \eta_{iu}}{\partial w_n} \right).
 \end{aligned}$$

Pela regra da cadeia tem-se que

$$\frac{\partial \mu_{iu}}{\partial w_i} = \frac{\partial \mu_{iu}}{\partial \eta_{i1}} \frac{\partial \eta_{i1}}{\partial w_i} + \frac{\partial \mu_{iu}}{\partial \eta_{i2}} \frac{\partial \eta_{i2}}{\partial w_i} + \dots + \frac{\partial \mu_{iu}}{\partial \eta_{ip}} \frac{\partial \eta_{ip}}{\partial w_i}$$

onde

$$\frac{\partial \mu_{iu}}{\partial \eta_{ij}} = \begin{cases} \frac{g'(\eta_{ij})(\phi_i - g(\eta_{ij}))}{\phi_i^2}, & \text{se } u = j, \\ -\frac{g(\eta_{iu})g'_j(\eta_{ij})}{\phi_i^2}, & \text{se } u \neq j. \end{cases}$$

Logo,

$$\frac{\partial \mu_{iu}}{\partial w_i} = -\frac{g(\eta_{iu})}{\phi_i^2} \sum_{h=1}^k \sum_{j=1}^p g'(\eta_{ij}) x_{ih} \beta_{hj} + \frac{g'(\eta_{iu})}{\phi_i} \sum_{h=1}^k x_{ih} \beta_{hu}$$

assim, se  $i = n$

$$\begin{aligned} \frac{\partial A_{nu}}{\partial w_n} &= -\psi'(\phi_n \mu_{nu}) \left[ \mu_{nu} \sum_{h=1}^k \sum_{j=1}^p g'(\eta_{nj}) x_{nh} \beta_{hj} + \phi_n \left( -\frac{g(\eta_{nu})}{\phi_n^2} \sum_{h=1}^k \sum_{j=1}^p g'(\eta_{nj}) x_{nh} \beta_{hj} \right. \right. \\ &\quad \left. \left. + \frac{g'(\eta_{nu})}{\phi_n} \sum_{h=1}^k x_{nh} \beta_{hu} \right) \right] \\ &= -\psi'(\phi_n \mu_{nu}) \left[ \mu_{nu} \sum_{h=1}^k \sum_{j=1}^p g'(\eta_{nj}) x_{nh} \beta_{hj} - \mu_{nu} \sum_{h=1}^k \sum_{j=1}^p g'(\eta_{nj}) x_{nh} \beta_{hj} \right. \\ &\quad \left. + g'(\eta_{nu}) \sum_{h=1}^k x_{nh} \beta_{hu} \right] \\ &= -\psi'(\phi_n \mu_{nu}) g'(\eta_{nu}) \sum_{h=1}^k x_{nh} \beta_{hu} \end{aligned}$$

Então, juntando todos estes resultados tem-se que

$$\begin{aligned} \Delta_{vu,n} &= w_n x_{nv} \left\{ g''(\eta_{nu}) [\psi(\phi_n) + A_{nu}] \sum_{h=1}^k x_{nh} \beta_{hu} + g'(\eta_{nu}) \left( \psi'(\phi_n) \sum_{h=1}^k \sum_{j=1}^p g'(\eta_{nj}) x_{nh} \beta_{hj} \right. \right. \\ &\quad \left. \left. - \psi'(\phi_n \mu_{nu}) g'(\eta_{nu}) \sum_{h=1}^k x_{nh} \beta_{hu} \right) \right\} + x_{nv} g'(\eta_{nu}) [\psi(\phi_n) + A_{nu}]. \end{aligned}$$

Daí,

$$\begin{aligned} \Delta_{vu,n} \Big|_{\mathbf{W}=\mathbf{1}, \mathbf{B}=\widehat{\mathbf{B}}} &= \\ &= x_{nv} \left\{ g''(\widehat{\eta}_{nu}) [\psi(\widehat{\phi}_n) + \widehat{A}_{nu}] \sum_{h=1}^k x_{nh} \widehat{\beta}_{hu} + g'(\widehat{\eta}_{nu}) \left( \psi'(\widehat{\phi}_n) \sum_{h=1}^k \sum_{j=1}^p g'(\widehat{\eta}_{nj}) x_{nh} \widehat{\beta}_{hj} \right. \right. \\ &\quad \left. \left. - \psi'(\widehat{\phi}_n \widehat{\mu}_{nu}) g'(\widehat{\eta}_{nu}) \sum_{h=1}^k x_{nh} \widehat{\beta}_{hu} \right) \right\} + x_{nv} g'(\widehat{\eta}_{nu}) [\psi(\widehat{\phi}_n) + \widehat{A}_{nu}]. \end{aligned}$$

### A.3 Esquema de perturbação III

Neste caso se perturba as variáveis de forma aditiva, mais especificamente a matriz  $\mathbf{X}$  é substituída pela matriz  $\mathbf{X}_w = \mathbf{X} + \mathbf{W}S$  onde  $\mathbf{W}$  corresponde à matriz de perturbações e  $S = \text{diag}\{s_1, \dots, s_k\}$ , de tal forma que cada  $s_m$  converte a perturbação  $w_{im}$  no tamanho e unidade apropriados para que  $w_{im}s_m$  seja compatível com  $x_{im}$  como foi visto no Capítulo 3.

$$\Delta_{vu, nm} = \frac{\partial}{\partial w_{nm}} \left( \frac{\partial \ell(\mathbf{B}|\mathbf{W})}{\partial \beta_{vu}} \right)$$

para  $n = 1, \dots, N, u = 1, \dots, p$  e  $v, m = 1, \dots, k$ .

A função de log-verossimilhança perturbada é

$$\ell(\mathbf{B}|\mathbf{W}) = \sum_{i=1}^N \left\{ \log \Gamma(\phi_i) - \sum_{j=1}^p \log(\Gamma(\phi_i \mu_{ij})) + \sum_{j=1}^p (\phi_i \mu_{ij} - 1) \log y_{ij} \right\}$$

onde  $\phi_i$  e  $\mu_{ij}$  são definidos como no esquema de perturbação I, mas  $\eta_{ij} = (x_{i1} + s_1 w_{i1})\beta_{1j} + \dots + (x_{ik} + s_k w_{ik})\beta_{kj}$ .

Analogamente ao esquema de perturbação II, a derivada de primeira ordem da função de log-verossimilhança perturbada com relação aos parâmetros  $\beta_{vu}$  é

$$\frac{\partial \ell(\mathbf{B}|\mathbf{W})}{\partial \beta_{vu}} = \sum_{i=1}^N \left\{ \psi(\phi_i) \frac{\partial \phi_i}{\partial \beta_{vu}} + \sum_{j=1}^p A_{ij} \frac{\partial}{\partial \beta_{vu}} (\phi_i \mu_{ij}) \right\}$$

onde

$$\frac{\partial \phi_i}{\partial \beta_{vu}} = \sum_{j=1}^p \frac{\partial}{\partial \beta_{vu}} g(\eta_{ij}) = \sum_{\substack{j=1 \\ j \neq u}}^p g'(\eta_{ij}) \frac{\partial \eta_{ij}}{\partial \beta_{vu}} + g'(\eta_{iu}) \frac{\partial \eta_{iu}}{\partial \beta_{vu}} = g'(\eta_{iu})(x_{iv} + s_v w_{iv})$$

e

$$\begin{aligned} \frac{\partial}{\partial \beta_{vu}} (\phi_i \mu_{ij}) &= \frac{\partial \phi_i}{\partial \beta_{vu}} \mu_{ij} + \phi_i \frac{\partial \mu_{ij}}{\partial \beta_{vu}} = \frac{\partial \phi_i}{\partial \beta_{vu}} \mu_{ij} + \phi_i \frac{\partial \mu_{ij}}{\partial \eta_{iu}} \frac{\partial \eta_{iu}}{\partial \beta_{vu}} \\ &= g'(\eta_{iu})(x_{iv} + s_v w_{iv}) \mu_{ij} + \phi_i \frac{\partial \mu_{ij}}{\partial \eta_{iu}} (x_{iv} + s_v w_{iv}). \end{aligned}$$

Então substituindo estas últimas derivadas e (A.2) encontra-se

$$\begin{aligned}
 \frac{\partial \ell(\mathbf{B}|\mathbf{W})}{\partial \beta_{vu}} &= \sum_{i=1}^N \left\{ \psi(\phi_i) g'(\eta_{iu}) (x_{iv} + s_v w_{iv}) \right. \\
 &\quad \left. + \sum_{j=1}^p A_{ij} \left[ g'(\eta_{iu}) (x_{iv} + s_v w_{iv}) \mu_{ij} + \phi_i \frac{\partial \mu_{ij}}{\partial \eta_{iu}} (x_{iv} + s_v w_{iv}) \right] \right\} \\
 &= \sum_{i=1}^N (x_{iv} + s_v w_{iv}) \left\{ \psi(\phi_i) g'(\eta_{iu}) + \sum_{j=1}^p A_{ij} g'(\eta_{iu}) \frac{g(\eta_{ij})}{\phi_i} + \sum_{j=1}^p A_{ij} \phi_i \frac{\partial \mu_{ij}}{\partial \eta_{iu}} \right\} \\
 &= \sum_{i=1}^N (x_{iv} + s_v w_{iv}) \left\{ \psi(\phi_i) g'(\eta_{iu}) + \frac{g'(\eta_{iu})}{\phi_i} \sum_{j=1}^p A_{ij} g(\eta_{ij}) + \phi_i \sum_{j=1}^p A_{ij} \frac{\partial \mu_{ij}}{\partial \eta_{iu}} \right\} \\
 &= \sum_{i=1}^N (x_{iv} + s_v w_{iv}) \left\{ \psi(\phi_i) g'(\eta_{iu}) + \frac{g'(\eta_{iu})}{\phi_i} \sum_{j=1}^p A_{ij} g(\eta_{ij}) \right. \\
 &\quad \left. + \phi_i \left( \sum_{\substack{j=1 \\ j \neq u}}^p -A_{ij} \frac{g(\eta_{ij}) g'(\eta_{iu})}{\phi_i^2} + A_{iu} \frac{g'(\eta_{iu}) (\phi_i - g(\eta_{iu}))}{\phi_i^2} \right) \right\} \\
 &= \sum_{i=1}^N (x_{iv} + s_v w_{iv}) \frac{g'(\eta_{iu})}{\phi_i} \left\{ \phi_i \psi(\phi_i) + \sum_{j=1}^p A_{ij} g(\eta_{ij}) - \sum_{\substack{j=1 \\ j \neq u}}^p A_{ij} g(\eta_{ij}) \right. \\
 &\quad \left. + A_{iu} (\phi_i - g(\eta_{iu})) \right\} \\
 &= \sum_{i=1}^N (x_{iv} + s_v w_{iv}) \frac{g'(\eta_{iu})}{\phi_i} \left\{ \phi_i \psi(\phi_i) + \sum_{j=1}^p A_{ij} g(\eta_{ij}) - \sum_{j=1}^p A_{ij} g(\eta_{ij}) + A_{iu} \phi_i \right\} \\
 &= \sum_{i=1}^N (x_{iv} + s_v w_{iv}) g'(\eta_{iu}) [\psi(\phi_i) + A_{iu}].
 \end{aligned}$$

Portanto,

$$\frac{\partial \ell(\mathbf{B}|\mathbf{W})}{\partial \beta_{vu}} = \sum_{i=1}^N (x_{iv} + s_v w_{iv}) g'(\eta_{iu}) [\psi(\phi_i) + A_{iu}].$$

A seguir estão apresentadas as derivadas necessárias para encontrar a matriz delta

$$\begin{aligned}\Delta_{vu, nm} &= \frac{\partial}{\partial w_{nm}} \left( \frac{\partial \ell(\mathbf{B}|\mathbf{W})}{\partial \beta_{vu}} \right) \\ \Delta_{vu, nm} &= \frac{\partial}{\partial w_{nm}} \left( \sum_{i=1}^N (x_{iv} + s_v w_{iv}) g'(\eta_{iu}) [\psi(\phi_i) + A_{iu}] \right) \\ &= \sum_{i=1}^N x_{iv} \frac{\partial}{\partial w_{nm}} (g'(\eta_{iu}) [\psi(\phi_i) + A_{iu}]) + s_v \sum_{i=1}^N \frac{\partial}{\partial w_{nm}} w_{iv} g'(\eta_{iu}) [\psi(\phi_i) + A_{iu}] \\ &= \sum_{i=1}^N x_{iv} \text{Der1} + s_v \sum_{i=1}^N \text{Der2}\end{aligned}$$

onde

$$\begin{aligned}\text{Der1} &= \frac{\partial}{\partial w_{nm}} [g'(\eta_{iu}) (\psi(\phi_i) + A_{iu})], \\ \text{Der2} &= \frac{\partial}{\partial w_{nm}} w_{iv} g'(\eta_{iu}) [\psi(\phi_i) + A_{iu}],\end{aligned}$$

$$\begin{aligned}\text{Der1} &= \left( \frac{\partial}{\partial w_{nm}} g'(\eta_{iu}) \right) [\psi(\phi_i) + A_{iu}] + g'(\eta_{iu}) \frac{\partial}{\partial w_{nm}} [\psi(\phi_i) + A_{iu}] \\ &= g''(\eta_{iu}) \frac{\partial \eta_{iu}}{\partial w_{nm}} [\psi(\phi_i) + A_{iu}] + g'(\eta_{iu}) \left( \psi'(\phi_i) \frac{\partial \phi_i}{\partial w_{nm}} + \frac{\partial A_{iu}}{\partial w_{nm}} \right)\end{aligned}$$

em que

$$\frac{\partial \phi_i}{\partial w_{nm}} = \sum_{j=1}^p \frac{\partial}{\partial w_{nm}} g(\eta_{ij}) = \sum_{j=1}^p g'(\eta_{ij}) \frac{\partial \eta_{ij}}{\partial w_{nm}}.$$

Mas

$$\frac{\partial \eta_{ij}}{\partial w_{nm}} = \begin{cases} s_m \beta_{mj}, & \text{se } i = n, \\ 0, & \text{se } i \neq n; \end{cases} \quad (\text{A.3})$$

logo,

$$\frac{\partial \phi_i}{\partial w_{nm}} = \begin{cases} s_m \sum_{j=1}^p g'(\eta_{ij}) \beta_{mj}, & \text{se } i = n, \\ 0, & \text{se } i \neq n. \end{cases} \quad (\text{A.4})$$

De outra parte, pela regra da cadeia tem-se que

$$\frac{\partial \mu_{iu}}{\partial w_{im}} = \sum_{j=1}^p \frac{\partial \mu_{iu}}{\partial \eta_{ij}} \frac{\partial \eta_{ij}}{\partial w_{im}},$$

onde

$$\frac{\partial \mu_{iu}}{\partial \eta_{ij}} = \begin{cases} \frac{g'(\eta_{ij})(\phi_i - g(\eta_{ij}))}{\phi_i^2}, & \text{se } u = j, \\ -\frac{g(\eta_{iu})g'(\eta_{ij})}{\phi_i^2}, & \text{se } u \neq j. \end{cases}$$

Assim,

$$\begin{aligned} \frac{\partial \mu_{iu}}{\partial w_{im}} &= -\frac{g(\eta_{iu})g'(\eta_{i1})}{\phi_i^2} s_m \beta_{m1} - \dots + \frac{g'(\eta_{iu})(\phi_i - g(\eta_{iu}))}{\phi_i^2} s_m \beta_{mu} \\ &\quad - \dots - \frac{g(\eta_{iu})g'(\eta_{ip})}{\phi_i^2} s_m \beta_{mp} \\ &= -s_m \frac{g(\eta_{iu})}{\phi_i^2} \sum_{j=1}^p g'(\eta_{ij}) \beta_{mj} + s_m \frac{g'(\eta_{iu})}{\phi_i} \beta_{mu}. \end{aligned}$$

Substituindo o resultado obtido em (A.4)

$$\frac{\partial \mu_{iu}}{\partial w_{nm}} = \begin{cases} -s_m \frac{g(\eta_{iu})}{\phi_i^2} \sum_{j=1}^p g'(\eta_{ij}) \beta_{mj} + s_m \frac{g'(\eta_{iu})}{\phi_i} \beta_{mu}, & \text{se } i = n, \\ 0, & \text{se } i \neq n. \end{cases} \quad (\text{A.5})$$

Portanto,

$$\begin{aligned} \sum_{i=1}^N x_{iv} \text{Der1} &= \sum_{i=1}^N x_{iv} \left\{ g''(\eta_{iu}) \frac{\partial \eta_{iu}}{\partial w_{nm}} [\psi(\phi_i) + A_{iu}] + g'(\eta_{iu}) \left( \psi'(\phi_i) \frac{\partial \phi_i}{\partial w_{nm}} + \frac{\partial A_{iu}}{\partial w_{nm}} \right) \right\} \\ &= \sum_{i=1}^N x_{iv} g''(\eta_{iu}) \frac{\partial \eta_{iu}}{\partial w_{nm}} [\psi(\phi_i) + A_{iu}] + \sum_{i=1}^N x_{iv} g'(\eta_{iu}) \left( \psi'(\phi_i) \frac{\partial \phi_i}{\partial w_{nm}} + \frac{\partial A_{iu}}{\partial w_{nm}} \right), \end{aligned}$$

mas, por (A.4), tem-se que

$$\begin{aligned} &\sum_{i=1}^N x_{iv} g''(\eta_{iu}) \frac{\partial \eta_{iu}}{\partial w_{nm}} [\psi(\phi_i) + A_{iu}] \\ &= \sum_{\substack{i=1 \\ i \neq n}}^N x_{iv} g''(\eta_{iu}) \frac{\partial \eta_{iu}}{\partial w_{nm}} [\psi(\phi_i) + A_{iu}] + x_{nv} g''(\eta_{nu}) \frac{\partial \eta_{nu}}{\partial w_{nm}} [\psi(\phi_n) + A_{nu}] \\ &= x_{nv} g''(\eta_{nu}) s_m \beta_{mu} [\psi(\phi_n) + A_{nu}] \end{aligned} \quad (\text{A.6})$$



e

$$\begin{aligned}
 & \sum_{i=1}^N x_{iv} g'(\eta_{iu}) \left( \psi'(\phi_i) \frac{\partial \phi_i}{\partial w_{nm}} + \frac{\partial A_{iu}}{\partial w_{nm}} \right) \\
 &= \sum_{i=1}^N x_{iv} g'(\eta_{iu}) \psi'(\phi_i) \frac{\partial \phi_i}{\partial w_{nm}} + \sum_{i=1}^N x_{iv} g'(\eta_{iu}) \frac{\partial A_{iu}}{\partial w_{nm}} \\
 &= \sum_{\substack{i=1 \\ i \neq n}}^N x_{iv} g'(\eta_{iu}) \psi'(\phi_i) \frac{\partial \phi_i}{\partial w_{nm}} + x_{nv} g'(\eta_{nu}) \psi'(\phi_n) \frac{\partial \phi_n}{\partial w_{nm}} + \sum_{i=1}^N x_{iv} g'(\eta_{iu}) \frac{\partial A_{iu}}{\partial w_{nm}} \\
 &= x_{nv} g'(\eta_{nu}) \psi'(\phi_n) \left[ s_m \sum_{j=1}^p g'(\eta_{nj}) \beta_{mj} \right] + \sum_{i=1}^N x_{iv} g'(\eta_{iu}) \frac{\partial A_{iu}}{\partial w_{nm}}; \tag{A.7}
 \end{aligned}$$

ainda

$$\begin{aligned}
 \frac{\partial A_{iu}}{\partial w_{nm}} &= \frac{\partial}{\partial w_{nm}} [\log y_{iu} - \psi(\phi_i \mu_{iu})] = -\psi'(\phi_i \mu_{iu}) \frac{\partial}{\partial w_{nm}} (\phi_i \mu_{iu}) \\
 &= -\psi'(\phi_i \mu_{iu}) \left( \frac{\partial \phi_i}{\partial w_{nm}} \mu_{iu} + \phi_i \frac{\partial \mu_{iu}}{\partial w_{nm}} \right).
 \end{aligned}$$

Substituindo (A.5) e (A.6) obtém-se,

$$\begin{aligned}
 \frac{\partial A_{iu}}{\partial w_{nm}} &= \begin{cases} -\psi'(\phi_i \mu_{iu}) \left[ s_m \sum_{j=1}^p g'(\eta_{ij}) \beta_{mj} \mu_{iu} \right. \\ \qquad \qquad \qquad \left. + \phi_i \left( s_m \frac{g'(\eta_{iu})}{\phi_i} \beta_{mu} - s_m \frac{g(\eta_{iu})}{\phi_i^2} \sum_{j=1}^p g'(\eta_{ij}) \beta_{mj} \right) \right] & \text{se } i = n, \\ -\psi'(\phi_i \mu_{iu}) [\mu_{ij} \cdot 0 + \phi_i \cdot 0], & \text{se } i \neq n, \end{cases} \\
 &= \begin{cases} -\psi'(\phi_i \mu_{iu}) \left[ s_m \mu_{iu} \sum_{j=1}^p g'(\eta_{ij}) \beta_{mj} \right. \\ \qquad \qquad \qquad \left. + s_m \beta_{mu} g'(\eta_{iu}) - s_m \frac{g(\eta_{iu})}{\phi_i} \sum_{j=1}^p g'(\eta_{ij}) \beta_{mj} \right] & \text{se } i = n, \\ 0, & \text{se } i \neq n. \end{cases}
 \end{aligned}$$

Logo,

$$\frac{\partial A_{iu}}{\partial w_{nm}} = \begin{cases} -\psi'(\phi_i \mu_{iu}) s_m \beta_{mu} g'(\eta_{iu}) & \text{se } i = n, \\ 0, & \text{se } i \neq n. \end{cases} \tag{A.8}$$

Assim,

$$\begin{aligned}
 \sum_{i=1}^N x_{iv} g'(\eta_{iu}) \frac{\partial A_{iu}}{\partial w_{nm}} &= \sum_{\substack{i=1 \\ i \neq n}}^N x_{iv} g'(\eta_{iu}) \frac{\partial A_{iu}}{\partial w_{nm}} + x_{nv} g'(\eta_{nu}) \frac{\partial A_{nu}}{\partial w_{nm}} \\
 &= \sum_{\substack{i=1 \\ i \neq n}}^N x_{iv} g'(\eta_{iu}) \cdot 0 + x_{nv} g'(\eta_{nu}) (-\psi'(\phi_n \mu_{nu}) s_m \beta_{mu} g'(\eta_{nu})) \\
 &= -x_{nv} g'(\eta_{nu})^2 \psi'(\phi_n \mu_{nu}) s_m \beta_{mu}.
 \end{aligned} \tag{A.9}$$

Então substituindo (A.10) em (A.8) e usando também o resultado encontrado em (A.7) tem-se que

$$\begin{aligned}
 \sum_{i=1}^N x_{iv} Der1 &= x_{nv} g''(\eta_{nu}) s_m \beta_{mu} [\psi(\phi_n) + A_{nu}] + x_{nv} g'(\eta_{nu}) \psi'(\phi_n) [s_m \sum_{j=1}^p g'(\eta_{nj}) \beta_{mj}] \\
 &\quad - x_{nv} g'(\eta_{nu})^2 \psi'(\phi_n \mu_{nu}) s_m \beta_{mu} \\
 &= x_{nv} s_m \left\{ g''(\eta_{nu}) \beta_{mu} [\psi(\phi_n) + A_{nu}] + g'(\eta_{nu}) \psi'(\phi_n) \left[ \sum_{j=1}^p g'(\eta_{nj}) \beta_{mj} \right] \right. \\
 &\quad \left. - g'(\eta_{nu})^2 \psi'(\phi_n \mu_{nu}) \beta_{mu} \right\} = \mathbf{K}.
 \end{aligned} \tag{A.10}$$

Para maior simplicidade na hora de escrever a matriz  $\Delta_{vu, nm}$  se vai denotar a expressão (A.11) por  $\mathbf{K}$ . Por outro lado

$$Der2 = \begin{cases} g'(\eta_{iu}) [\psi(\phi_i) + A_{iu}] + w_{iv} \frac{\partial}{\partial w_{nm}} g'(\eta_{iu}) [\psi(\phi_i) + A_{iu}], & \text{se } i = n, \text{ e, } v = m, \\ w_{iv} \frac{\partial}{\partial w_{nm}} g'(\eta_{iu}) [\psi(\phi_i) + A_{iu}], & \text{se } i \neq n, \text{ ou, } v \neq m. \end{cases}$$

Tem-se que

$$i \neq n \text{ ou } v \neq m \text{ quando } \begin{cases} i \neq n \text{ e } v \neq m, \text{ ou} \\ i \neq n \text{ e } v = m, \text{ ou} \\ i = n \text{ e } v \neq m. \end{cases}$$

Mas por (A.4), (A.5) e (A.9), se  $i \neq n$  então

$$\frac{\partial}{\partial w_{nm}} g'(\eta_{iu}) [\psi(\phi_i) + A_{iu}] = 0.$$

Assim,

$$s_v \sum_{i=1}^N Der2 = \begin{cases} s_m g'(\eta_{nu})[\psi(\phi_n) + A_{nu}] + s_m w_{nm} \frac{\partial}{\partial w_{nm}} g'(\eta_{nu})[\psi(\phi_n) + A_{nu}], & \text{se } v = m, \\ s_v w_{nv} \frac{\partial}{\partial w_{nm}} g'(\eta_{nu})[\psi(\phi_n) + A_{nu}], & \text{se } v \neq m. \end{cases} \quad (\text{A.11})$$

Note-se que ao avaliar (A.12) em  $\mathbf{W} = \mathbf{0}$  encontra-se que

$$s_v \sum_{i=1}^N Der2 = \begin{cases} s_m g'(\eta_{nu})[\psi(\phi_n) + A_{nu}], & \text{se } v = m, \\ 0, & \text{se } v \neq m. \end{cases} \quad (\text{A.12})$$

Portanto, pode-se concluir que

$$\Delta_{vu, nm} \Big|_{\mathbf{B}=\hat{\mathbf{B}}, \mathbf{W}=\mathbf{0}} = \begin{cases} \hat{\mathbf{K}} + s_m g'(\hat{\eta}_{nu})[\psi(\hat{\phi}_n) + \hat{A}_{nu}], & \text{se } v = m, \\ \hat{\mathbf{K}}, & \text{se } v \neq m. \end{cases} \quad (\text{A.13})$$

## A.4 Esquema de perturbação IV

Neste esquema a função de log-verossimilhança perturbada é

$$\ell(\mathbf{B}|\mathbf{W}) = \sum_{i=1}^N \left\{ \log \Gamma(\phi_i) - \sum_{j=1}^p \log(\Gamma(\phi_i \mu_{ij})) + \sum_{j=1}^p (\phi_i \mu_{ij} - 1) \log(y_{ij} + w_{ij}) \right\}$$

onde  $\phi_i = \sum_{j=1}^p g(\eta_{ij})$ ,  $\eta_{ij} = x_{i1}\beta_{1j} + \dots + x_{ik}\beta_{kj}$  e  $\mu_{ij} = \frac{g(\eta_{ij})}{\phi_i}$ . Como nos esquemas de perturbação anteriores, o objetivo é encontrar a matriz

$$\Delta_{vu,ns} = \frac{\partial}{\partial w_{ns}} \left( \frac{\partial \ell(\mathbf{B}|\mathbf{W})}{\partial \beta_{vu}} \right)$$

com  $v = 1, \dots, k$ ,  $n = 1, \dots, N$  e  $u, s = 1, \dots, p$ .

A derivada de primeira ordem da função de log-verossimilhança perturbada com relação aos parâmetros desconhecidos é

$$\begin{aligned} \frac{\partial \ell(\mathbf{B}|\mathbf{W})}{\partial \beta_{vu}} &= \sum_{i=1}^N \left\{ \frac{\partial}{\partial \beta_{vu}} \log \Gamma(\phi_i) + \sum_{j=1}^p \left[ \frac{\partial}{\partial \beta_{vu}} (-\log \Gamma(\phi_i \mu_{ij})) \right. \right. \\ &\quad \left. \left. + \frac{\partial}{\partial \beta_{vu}} (\phi_i \mu_{ij} - 1) \log(y_{ij} + w_{ij}) \right] \right\} \\ &= \sum_{i=1}^N \left\{ \psi(\phi_i) \frac{\partial \phi_i}{\partial \beta_{vu}} + \sum_{j=1}^p \left[ \left( \log(y_{ij} + w_{ij}) - \psi(\phi_i \mu_{ij}) \right) \frac{\partial}{\partial \beta_{vu}} (\phi_i \mu_{ij}) \right] \right\} \\ &= \sum_{i=1}^N \left\{ \psi(\phi_i) \frac{\partial \phi_i}{\partial \beta_{vu}} + \sum_{j=1}^p A_{ij}^* \frac{\partial}{\partial \beta_{vu}} (\phi_i \mu_{ij}) \right\} \end{aligned}$$

em que  $A_{ij}^* = \log(y_{ij} + w_{ij}) - \psi(\phi_i \mu_{ij})$ . A derivada de  $\phi_i$  com respeito aos parâmetros desconhecidos  $\beta_{vu}$  é

$$\frac{\partial \phi_i}{\partial \beta_{vu}} = \sum_{j=1}^p g'(\eta_{ij}) \frac{\partial \eta_{ij}}{\partial \beta_{vu}} = \sum_{\substack{j=1 \\ j \neq u}}^p g'(\eta_{ij}) \frac{\partial \eta_{ij}}{\partial \beta_{vu}} + g'(\eta_{iu}) \frac{\partial \eta_{iu}}{\partial \beta_{vu}} = g'(\eta_{iu}) x_{iv}$$

logo

$$\frac{\partial}{\partial \beta_{vu}} (\phi_i \mu_{ij}) = \frac{\partial \phi_i}{\partial \beta_{vu}} \mu_{ij} + \phi_i \frac{\partial \mu_{ij}}{\partial \beta_{vu}} = g'(\eta_{iu}) x_{iv} \mu_{ij} + \phi_i \frac{\partial \mu_{ij}}{\partial \eta_{iu}} x_{iv} = x_{iv} \left( g'(\eta_{iu}) \mu_{ij} + \phi_i \frac{\partial \mu_{ij}}{\partial \eta_{iu}} \right)$$

Assim, usando (A.2) e os resultados anteriores encontra-se que

$$\begin{aligned}
 \frac{\partial \ell(\mathbf{B}|\mathbf{W})}{\partial \beta_{vu}} &= \sum_{i=1}^N \left\{ \psi(\phi_i) g'(\eta_{iu}) x_{iv} + \sum_{j=1}^p A_{ij}^* x_{iv} \left( g'(\eta_{iu}) \mu_{ij} + \phi_i \frac{\partial \mu_{ij}}{\partial \eta_{iu}} \right) \right\} \\
 &= \sum_{i=1}^N x_{iv} \left\{ \psi(\phi_i) g'(\eta_{iu}) + \sum_{j=1}^p A_{ij}^* g'(\eta_{iu}) \frac{g(\eta_{ij})}{\phi_i} \right. \\
 &\quad \left. + \sum_{\substack{j=1 \\ j \neq u}}^p A_{ij}^* \phi_i \left( \frac{-g(\eta_{ij}) g'(\eta_{iu})}{\phi_i^2} \right) + A_{iu}^* \phi_i \frac{g'(\eta_{iu}) (\phi_i - g(\eta_{iu}))}{\phi_i^2} \right\} \\
 &= \sum_{i=1}^N x_{iv} \frac{g'(\eta_{iu})}{\phi_i} \left\{ \phi_i \psi(\phi_i) + \sum_{j=1}^p A_{ij}^* g(\eta_{ij}) - \sum_{\substack{j=1 \\ j \neq u}}^p A_{ij}^* g(\eta_{ij}) + A_{iu}^* (\phi_i - g(\eta_{iu})) \right\} \\
 &= \sum_{i=1}^N x_{iv} \frac{g'(\eta_{iu})}{\phi_i} \left\{ \phi_i \psi(\phi_i) + \sum_{j=1}^p A_{ij}^* g(\eta_{ij}) - \sum_{j=1}^p A_{ij}^* g(\eta_{ij}) + A_{iu}^* \phi_i \right\} \\
 &= \sum_{i=1}^N x_{iv} \frac{g'(\eta_{iu})}{\phi_i} \{ \phi_i \psi(\phi_i) + A_{iu}^* \phi_i \} \\
 &= \sum_{i=1}^N x_{iv} g'(\eta_{iu}) [\psi(\phi_i) + A_{iu}^*].
 \end{aligned}$$

Portanto

$$\frac{\partial \ell(\mathbf{B}|\mathbf{W})}{\partial \beta_{vu}} = \sum_{i=1}^N x_{iv} g'(\eta_{iu}) [\psi(\phi_i) + A_{iu}^*].$$

A seguir estão apresentadas as derivadas necessárias para encontrar a matriz  $\Delta_{vu,ns}$

$$\begin{aligned}
 \Delta_{vu,ns} &= \frac{\partial}{\partial w_{ns}} \left( \sum_{i=1}^N x_{iv} g'(\eta_{iu}) [\psi(\phi_i) + \log(y_{iu} + w_{iu}) - \psi(\phi_i \mu_{iu})] \right) \\
 \Delta_{vu,ns} &= \begin{cases} \sum_{i=1}^N \frac{\partial}{\partial w_{ns}} \left( x_{iv} g'(\eta_{iu}) [\psi(\phi_i) + \log(y_{iu} + w_{iu}) - \psi(\phi_i \mu_{iu})] \right), & \text{se } s = u, \\ 0, & \text{se } s \neq u. \end{cases} \\
 &= \begin{cases} \frac{\partial}{\partial w_{ns}} x_{nv} g'(\eta_{nu}) \log(y_{nu} + w_{nu}), & \text{se } s = u, \\ 0, & \text{se } s \neq u. \end{cases} \\
 &= \begin{cases} x_{nv} g'(\eta_{nu}) \frac{1}{y_{nu} + w_{nu}}, & \text{se } s = u, \\ 0, & \text{se } s \neq u. \end{cases}
 \end{aligned}$$

Dessa forma, pode-se concluir que

$$\Delta_{vu,ns} \Big|_{\mathbf{B}=\hat{\mathbf{B}}, \mathbf{W}=\mathbf{0}} = \begin{cases} x_{nv} g'(\hat{\eta}_{nu}) \frac{1}{y_{nu}}, & \text{se } s = u, \\ 0, & \text{se } s \neq u. \end{cases} \quad (\text{A.14})$$

# Apêndice B

## Dados

Neste apêndice apresenta-se o conjunto de dados reais que se utilizou no capítulo 4. O conjunto de dados consiste da composição de areia, lodo e argila de 39 amostras de sedimentos a diferentes profundidades de água em um lago ártico.

**Tabela 1:** (observações 1-12)

Observação	Profundidade ( <i>Hectômetros</i> )	Proporção de areia	Proporção de lodo	Proporção de argila
1	0.104	0.775	0.195	0.03
2	0.117	0.719	0.249	0.032
3	0.128	0.507	0.361	0.132
4	0.13	0.522	0.409	0.066
5	0.157	0.7	0.265	0.035
6	0.163	0.665	0.322	0.013
7	0.18	0.431	0.553	0.016
8	0.187	0.534	0.368	0.098
9	0.207	0.155	0.544	0.301
10	0.221	0.317	0.415	0.268
11	0.224	0.657	0.278	0.065
12	0.244	0.704	0.29	0.006

**Tabela 1:** (observações 13-39)

Observação	Profundidade (Hectômetros)	Proporção de areia	Proporção de lodo	Proporção de argila
13	0.258	0.174	0.536	0.29
14	0.325	0.106	0.698	0.196
15	0.336	0.382	0.431	0.187
16	0.368	0.108	0.527	0.365
17	0.378	0.184	0.507	0.309
18	0.369	0.046	0.474	0.48
19	0.422	0.156	0.504	0.34
20	0.47	0.319	0.451	0.23
21	0.471	0.095	0.535	0.37
22	0.484	0.171	0.48	0.349
23	0.494	0.105	0.554	0.341
24	0.495	0.048	0.547	0.41
25	0.592	0.026	0.452	0.522
26	0.601	0.114	0.527	0.359
27	0.617	0.067	0.469	0.464
28	0.624	0.069	0.497	0.434
29	0.693	0.04	0.449	0.511
30	0.736	0.074	0.516	0.409
31	0.744	0.048	0.495	0.457
32	0.785	0.045	0.485	0.47
33	0.829	0.066	0.521	0.413
34	0.877	0.067	0.473	0.459
35	0.881	0.074	0.456	0.469
36	0.904	0.06	0.489	0.451
37	0.906	0.063	0.538	0.399
38	0.977	0.025	0.48	0.495
39	1.037	0.02	0.478	0.502



# Apêndice C

## Programa

Neste apêndice apresenta-se o programa utilizado para encontrar os estimadores dos parâmetros  $\beta_{uv}$  e a matriz  $\Delta$  em cada esquema de perturbação para fazer a avaliação de influência local através da curvatura normal conforme.

```
/******  
*PROGRAMA: influencia.ox *  
* *  
*USO: Estimacao dos parametros do modelo de regressao Dirichlet *  
* usando o metodo de maxima verossimilhanca. Avaliacao da *  
* influencia local deste modelo atraves da curvatura normal *  
* conforme. *  
* *  
*AUTORA: Luz Milena Zea Fernandez *  
* *  
*DATA: Maio de 2005 *  
* *  
*ULTIMA MODIFICACAO: Setembro de 2005 *  
*****/  
  
#include <oxstd.h>  
#include <oxfloat.h>  
#import <maximize>  
  
//Variáveis globais  
static decl mX;//Matriz das variaveis explicativas  
static decl mY;//Matriz das variaveis explicadas  
static decl k;//Numero de variaveis explicativas  
static decl p;//Numero de variaveis explicadas  
//Funcao log-verossimilhanca do modelo Dirichlet nao perturbado  
flogDirichlet(const vP, const adFunc, const avScore, const amHess)  
{  
    decl i,j,B,eta,phi,mi,g_eta,derivadas,u,v,b;  
    B=shape(vP,k,p);  
    eta=mX*B;  
    g_eta=exp(eta);  
    phi=sumr(g_eta);
```

```

    adFunc[0] = sumc(loggamma(phi)-sumr(loggamma(g_eta))
        +sumr((g_eta-1).*log(mY)));
    if(isnan(adFunc[0]) || isdotinf(adFunc[0]))
        return 0;
    else
        return 1;
} //Fim funcao de log-verossimilhanca

main()
{
    decl vp,adfunc,dfunc,ir,delta,aux,B,eta,phi,g_eta,u,v,n,a,s,b;
    decl ele,m,F,A,NOR,amval,amvec,delta1,delta2,delta3,norma_obs_n;
    decl mq_barra,mj_q,limiar_critico,dimF,dim_q,M_salida,a1,a2,g2;
    decl i,j,amvec2,CN,mi_influ,esquema,C,S,delta4,QA,Q,mi,q,av;
    k=columns(mX);
    p=columns(mY);
    N=rows(mX);
    vp=zeros(k*p,1);
    ir=MaxBFGS(flogDirichlet,&vp,&dfunc,0,TRUE);
    B=shape(vp,k,p);
    eta=mX*B;
    g_eta=exp(eta);
    phi=sumr(g_eta);
    A=log(mY)-polygamma(g_eta,0);
    g2=g_eta.*g_eta;
    //Matriz L dois pontos
    ele=zeros(k*p,k*p);
    a=0;
    for(s=0;s<=p-1;s++){
        for(m=0;m<=k-1;m++){
            b=0;
            for(u=0;u<=p-1;u++){
                for(v=0;v<=k-1;v++){
                    if(u==s)
                        ele[a][b] = sumc(mX[][v].*mX[][m].*( g_eta[][u]
                            .*( polygamma(phi,0)+ A[][u])+ g2[][u]
                            .*(polygamma(phi,1)
                                -polygamma(g_eta[][u],1))));
                    else
                        ele[a][b] = sumc(mX[][v].*mX[][m].*(g_eta[][u]
                            .*g_eta[][s].*polygamma(phi,1)));
                    b=b+1;
                } //fim for v
            } //fim for u
            a=a+1;
        } //fim for m
    } //fim for s
    //Matriz Delta para esquema de perturbacao I
    delta1=zeros(k*p,N);
    aux=0;
    for(u=0;u<=p-1;u++){
        for(v=0;v<=k-1;v++){
            for(n=0;n<=N-1;n++){
                delta1[aux][n] = mX[n][v]*g_eta[n][u].*(polygamma(phi[n],0)
                    +A[n][u]);
            } //fim for n
        }
    }

```

```

        aux=aux+1;
    }//fim for v
}//fim for u
//Matriz Delta para esquema de perturbacao II
delta2=zeros(k*p,N);
aux=0;
for(u=0;u<=p-1;u++){
    for(v=0;v<=k-1;v++){
        for(n=0;n<=N-1;n++){
            delta2[aux][n] = mX[n][v]*g_eta[n][u]*((polygamma(phi[n],0)
                +A[n][u])*(sumr(mX[n][].*B[][u]')+1)
                +polygamma(phi[n],1)*(sumc(sumr(g_eta[n][]'
                .*(mX[n][].*B'))))-polygamma(g_eta[n][u],1)
                *g_eta[n][u]*(sumr(mX[n][].*B[][u]'))));
        }//fim for n
        aux=aux+1;
    }//fim for v
}//fim for u
//Matriz Delta para esquema de perturbacao III
delta3=zeros(k*p,N*k);
S=sqrt(diagonal(variance(mX)));
C=zeros(k*p,N*k);
a=0;
for(u=0;u<=p-1;u++){
    for(v=0;v<=k-1;v++){
        b=0;
        for(m=0;m<=k-1;m++){
            for(n=0;n<=N-1;n++){
                C[a][b] = mX[n][v]*g_eta[n][u]*S[m]*(B[m][u]
                    *(polygamma(phi[n],0)+A[n][u]-g_eta[n][u]
                    *polygamma(g_eta[n][u],1))+polygamma(phi[n],1)
                    *(sumr(g_eta[n][].*B[m][])));
                if(v==m)
                    delta3[a][b] = C[a][b]+S[v]*g_eta[n][u]
                        *(polygamma(phi[n],0)+A[n][u]);
                else
                    delta3[a][b] = C[a][b];
                b=b+1;
            }//fim for v
        }//fim for u
        a=a+1;
    }//fim for m
}//fim for s
//Matriz Delta para esquema de perturbacao IV
delta4=zeros(k*p,N*p);
a=0;
for(u=0;u<=p-1;u++){
    for(v=0;v<=k-1;v++){
        b=0;
        for(s=0;s<=p-1;s++){
            for(n=0;n<=N-1;n++){
                if(s==u)
                    delta4[a][b]=(mX[n][v]*g_eta[n][s])/mY[n][s];
                else
                    delta4[a][b]=0;
            }
        }
    }
}

```

```

        b=b+1;
    }//fim for n
} //fim for s
a=a+1;
} //fim for v
} //fim for u

for(esquema=1;esquema<=3;esquema++){
if(esquema==1)
    delta=delta1;
else if(esquema==2)
    delta=delta2;
else if(esquema==3)
    delta=delta3;
delta=selectifc(delta,delta.!=0);
F=delta'*invert(ele)*delta;
dimF=rows(F);//Recebe a dimensão de F
//Calculo dos autovalores e autovetores de -F
eigensym(-F, &amval, &amvec);
amval=(amval[] []/(norm(amval[] [],2)));//Normaliza os autovalores
mi=fabs(amval);
q=<1:3>/sqrt(dimF);
dim_q=columns(q);
for(j=0;j<=dimF-1;j++){
    amvec[] [j]=(amvec[] [j]/(norm(amvec[] [j],2)));
} //Normaliza autovetores
mi_influ=zeros(dim_q,dimF);
for(j=0;j<=dim_q-1;j++){
    mi_influ[j] []=mi.> q[j] .? mi .: 0;
}
NOR=sqrt(trace(F*F));
mq_barra=zeros(dim_q+1,1);
mq_barra[0]=trace(-F)/(dimF*NOR);
for(j=1;j<=dim_q;j++){
mq_barra[j]=sqrt(sumr(mi_influ[j-1] [])/dimF);
}
limiar_critico=zeros(dim_q+1,1);
limiar_critico[0]=2*mq_barra[0];
for(j=1;j<=dim_q;j++){
    limiar_critico[j]=sqrt(2)*mq_barra[j];
}
amvec2=amvec.*amvec;
mj_q=zeros(dim_q+1,dimF);//Recebe os valores das contribuicoes agregadas
for(j=0;j<=dimF-1;j++){
    for(i=0;i<=dim_q-1;i++){
        mj_q[i+1] [j]=sqrt(sumr(mi_influ[i] [].*amvec2[j] []));
    }
    mj_q[0] [j]=sumr(mi.*amvec2[j] []);
}
CN=zeros(dimF,1);//Curvatura normal conforme para todos os autovetores

```

```

for(i=0;i<=dimF-1;i++)
    CN[i]=-(amvec[] [i] '*F*amvec[] [i])/NOR;
b=zeros(4,1);
for(i=0;i<=dimF-1;i++){
    if(fabs(CN[i]).>0)
        b[0]=b[0]+1;
    if(fabs(CN[i]).>1/sqrt(dimF))
        b[1]=b[1]+1;//Autovetores 1-influentes
    if(fabs(CN[i]).>2/sqrt(dimF))
        b[2]=b[2]+1;//Autovetores 2-influentes
    if(fabs(CN[i]).>3/sqrt(dimF))
        b[3]=b[3]+1;//Autovetores 3-influentes
}
M_salida=selectifc(mj_q,mj_q.>limiar_critico);//Observacoes influentes
a=columns(M_salida);
aux=zeros(1,a);
i=0;
for(j=0;j<=dimF-1;j++){
    if(mj_q[0][j]>limiar_critico[0]||mj_q[1][j]>limiar_critico[1]
        ||mj_q[2][j]>limiar_critico[2]||mj_q[3][j]>limiar_critico[3]){
        aux[i]=j+1;
        i=i+1;
    }
}
a1=new array[dim_q+1];
a2=new array[a+3];
for(i=1;i<=dim_q;i++){
    a1[0]=sprintf("B_e=m_t[0]");
    a1[i]=sprintf("m_t[",i,"]");
}
for(i=0;i<=a-1;i++) {
    a2[0]=sprintf("N(e)");
    a2[1]=sprintf("mq_barra ");
    a2[2]=sprintf(" Valor crítico");
    a2[i+3]=sprintf("N = ", "%1.0f",aux[i]);
}
print("\n*****");
print("\n          ESQUEMA DE PERTURBAÇÃO ",esquema);
print("\n          TAMANHO DA AMOSTRA N = ",N);
print("\n          NÚMERO DE VARIÁVEIS EXPLICATIVAS k = ",k);
print("\n          NÚMERO DE VARIÁVEIS EXPLICADAS p = ",p);
print("\n*****\n");
print("  MEDIDAS DAS CONTRIBUIÇÕES AGREGADAS PARA AS OBSERVAÇÕES");
print("\n*****\n");
print("%r",a1,"%c",a2,b~mq_barra~limiar_critico~M_salida);
}// fin for esquema
if(esquema==4){
    delta=delta4;
    F=-delta*invert(ele)*delta;
    C=unit(N)**ones(1,p);
    Q=unit(N*p)-C'*invert(C*C')*C;
    QA=Q*F;
}

```

```

dimF=rows(QA);
eigen(QA, &amval, &amvec);
amval=(amval [] []/(norm(amval [] [],2)));
mi=fabs(amval);
q=<0:3>/sqrt(dimF);
dim_q=columns(q);
for(j=0;j<=dimF-1;j++){
    amvec [] [j]=(amvec [] [j]/(norm(amvec [] [j],2)));
}
mi_influ=zeros(dim_q,dimF);
for(j=0;j<=dim_q-1;j++){
    mi_influ[j] []=mi.> q[j] .? mi .: 0;
}
mq_barra=zeros(dim_q,1);
for(j=0;j<=dim_q-1;j++){
mq_barra[j]=sqrt(sumr(mi_influ[j] [])/N);
}
limiar_critico=zeros(dim_q,1);
for(j=0;j<=dim_q-1;j++){
    limiar_critico[j]=sqrt(2)*mq_barra[j];
}
amvec2=amvec.*amvec;
av=amvec2[0:N*p-1] [];
norma_obs_n=zeros(N,N*p);
aux=zeros(N,p);
for(s=0;s<=N*p-1;s++){
    for(i=0;i<=N-1;i++){
        for(j=0;j<=p-1;j++){
            aux[i] [j]=av[j*N+i] [s];
        }
        norma_obs_n[i] [s]=sumr(aux) [i];
    }
}
mj_q=zeros(dim_q,N);
for(j=0;j<=N-1;j++){
    for(i=0;i<=dim_q-1;i++){
        mj_q[i] [j]=sqrt(sumr(mi_influ[i] [] .*norma_obs_n[j] []));
    }
}
CN=zeros(dimF,1);
for(i=0;i<=dimF-1;i++) {
    CN[i]=-(amvec[0:dimF-1] [i] '*F*amvec[0:dimF-1] [i])/(sqrt(trace(F*F)));
}
b=zeros(4,1);
for(i=0;i<=dimF-1;i++){
    if(fabs(CN[i]).>0)
        b[0]=b[0]+1;
    if(fabs(CN[i]).>1/sqrt(dimF))
        b[1]=b[1]+1;
    if(fabs(CN[i]).>2/sqrt(dimF))
        b[2]=b[2]+1;
    if(fabs(CN[i]).>3/sqrt(dimF))
        b[3]=b[3]+1;
}

```

```

M_salida=selectifc(mj_q[][],mj_q[][].>limiar_critico);
a=columns(M_salida);
aux=zeros(1,a);
i=0;
for(j=0;j<=N-1;j++){
    if(mj_q[0][j]>limiar_critico[0]||mj_q[1][j]>limiar_critico[1]
        ||mj_q[2][j]>limiar_critico[2]||mj_q[3][j]>limiar_critico[3]){
        aux[i]=j+1;
        i=i+1;
    }//fim if
} //fim for j
a1=new array[dim_q+1];
a2=new array[a+3];
for(i=1;i<=dim_q;i++){
    a1[0]=sprintf("B_e=m_j[0]");
    a1[i]=sprintf("m_j[" ,i,"]");
}
for(i=0;i<=a-1;i++){
    a2[0]=sprintf("#auto-ve influ");
    a2[1]=sprintf("mq_barra");
    a2[2]=sprintf(" li_criti");
    a2[i+3]=sprintf("N = ", "%1.0f",aux[i]);
}
print("\n*****");
print("\n          ESQUEMA DE PERTURBAÇÃO ",esquema);
print("\n          TAMANHO DA AMOSTRA N = ",N);
print("\n          NÚMERO DE VARIÁVEIS EXPLICATIVAS k = ",k);
print("\n          NÚMERO DE VARIÁVEIS EXPLICADAS p = ",p);
print("\n*****\n");
print("  MEDIDAS DAS CONTRIBUIÇÕES AGREGADAS PARA AS OBSERVAÇÕES");
print("\n*****\n");
print("%r",a1,"%c",a2,b~mq_barra~limiar_critico~M_salida);
} //Fim if esquema IV
} //fim main

```

# Referências Bibliográficas

- Achcar, J. A. & Netto, A. R. (2003). Estudo da prevalência da tuberculose: uso de métodos bayesianos. *Revista Brasileira de Epidemiologia*, **6**(4), 380–387.
- Aitchison, J. (2003). *The Statistical Analysis of Compositional Data*. New Jersey: The Blackburn Press.
- Apostol, T. M. (1977). *Mathematical Analysis*. Addison Wesley Publishing Co., second edition.
- Beckman, R. J., Nachtsheim, C. J. & Cook, R. D. (1987). Diagnostics for mixed-model analysis of variance. *Technometrics*, **29**, 413–426.
- Bouguila, N. & Ziou, D. (2004). Dirichlet-based probability model applied to human skin detection. *IEEE Acoustics, Speech, and Signal Processing*, **5**, V521–V524. Proceedings. (ICASSP 2004).
- Bouguila, N., Ziou, D. & Vaillancourt, J. (2004). Unsupervised learning of a finite mixture model based on the Dirichlet distribution and its application. *IEEE Transactions on Image Processing*, **13**(11), 1533–1543.
- Boy, R. J., Henderson, D. A. & Wilkinson, D. J. (2000). Detecting homogeneous segments in DNA sequences by using hidden markov models. *Applied Statistics*, **49**(Part 2), 269–285.
- Bury, K. (1999). *Statistical Distributions in Engineering*. Cambridge University Press, New York.
- Cook, R. D. (1986). Assessment of local influence (with discussion). *Journal of the Royal Statistical Society - Series B*, **48**, 133–169.
- Cook, R. D. & Weisberg, S. (1982). *Residuals and Influence in Regression*. Chapman and Hall, London.



- Ferrari, S. L. P. & Cribari-Neto, F. (2004). Beta regression for modeling rates and proportions. *Journal of Applied Statistics*, **31**(7), 799–815.
- Fung, W. K. & Kwan, C. W. (1997). A note on local influence based on normal curvature. *Journal of the Royal Statistical Society - Series B*, **59**(4), 839–843.
- Galea, M., Paula, G. A. & Bolfarine, H. (1997). Local influence in elliptical linear regression models. *The Statistician*, **46**, 71–79.
- Galea, M., Paula, G. A. & Uribe-Opazo, M. (2003). On influence diagnostic in univariate elliptical linear regression models. *Statistical Papers*, **44**, 23–45.
- Graham, V. A. & Hollands, K. G. T. (1990). Method to generate synthetic hourly solar radiation globally. *Solar Energy*, **44**, 333–341.
- Gu, H. & Fung, W. K. (1998). Assessing local influence in canonical correlation analysis. *Annals of the Institute of Statistical Mathematics*, **50**, 755–772.
- Hastie, T. & Tibshirani, R. (1990). *Generalized Additive Models*. Chapman and Hall, London.
- Janardan, K. G. & Padmanabhan, G. (2004). Double bounded beta distribution for hydrologic variables. *Proc. 17th Annual Pittsburgh Conference*, **17**, 1107–1111.
- Jansen, I., Molenberghs, G., Aerts, G., Thijs, H. & Steen, K. V. (2003). A local influence approach applied to binary data from a psychiatric study. *Biometrics*, **59**, 410–419.
- Kwan, C. W. & Fung, W. K. (1998). Assessing local influence for specific restricted likelihood: Applications to factor analysis. *Psychometrika*, **63**, 35–46.
- Lawrence, A. J. (1998). Regression transformation diagnostics using local influence. *Journal of the American Statistical Association*, **84**, 125–141.
- Lima, E. L. (1981). *Curso de análise*, volume 2, quarta edição. Instituto de Matemática Pura e Aplicada, projeto Euclides, Rio de Janeiro.
- Liu, S. Z. (2000). On local influence for elliptical linear models. *Statistical Papers*, **41**, 211–224.
- Liu, S. Z. (2002). Local influence in multivariate elliptical linear regression models. *Linear Algebra and its Applications*, **354**, 159–174.

- Mazzuchi, T. A. & Soyer, R. (1993). A bayes method for assessing product-reliability during development testing. *IEEE Transactions on reliability*, **42**(3), 503–510.
- McNally, R. J. (1990). Maximum likelihood estimation of the parameters of the prior distributions of three variables that strongly influence reproductive performance in cows. *Biometrics*, **44**, 501–514.
- Merwe, A. J. V. D. & Pretorius, A. L. (2003). Bayesian estimation in animal breeding using the Dirichlet process prior for correlated random effects. *Genetics Selection Evolution*, **35**, 137–158.
- Milyutin, E. R. & Yaromenko, Y. I. (1991). Statistical characteristics of atmospheric transparency index over tilted routes. *Meteorologiya i Gidrologiya*, **12**, 72–76.
- Nelder, J. A. & Wedderburn, R. W. M. (1972). Generalized linear models. *Journal of the Royal Statistical Society - Series A*, **135**, 370–384.
- Ortega, E. M. M., Bolfarine, H. & Paula, G. A. (2003). Influence diagnostic in generalized log-gamma regression models. *Computational Statistics and Data Analysis*, **42**, 165–186.
- Paula, G. A. (1993). Assessing local influence in restricted regression models. *Computational Statistics and Data Analysis*, **16**, 63–79.
- Paula, G. A. (1996). Influence diagnostic in proper dispersion models. *Australian Journal of Statistics*, **38**, 307–316.
- Paula, G. A. (2004). *Modelos de regressão com apoio computacional*. Instituto de Matemática e Estatística, Universidade de São Paulo.
- Poon, W. Y. & Poon, Y. S. (1999). Conformal normal curvature and assessment of local influence. *Journal of the Royal Statistical Society - Serie B*, **61**, 51–61.
- Schall, R. & Dunne, T. T. (1992). A note on the relationship between parameter collinearity and local influence. *Biometrika*, **79**, 399–404.
- Shi, L. & Wang, X. (1999). Local influence in ridge regression. *Computational Statistics & Data Analysis*, **31**, 341–353.
- Silva, T. F. N. M. (2004). Estimação do Posto da Matriz dos Parâmetros do Modelo de Regressão Dirichlet. Dissertação de Mestrado em Estatística, Universidade Federal de Pernambuco.

Tsai, C. H. & Wu, X. (1992). Assessing local influence in linear regression models with first-order autoregressive or heteroscedastic error structure. *Statistics and Probability Letters*, **14**, 247–252.

Wei, B. C. (1998). *Exponential Family Nonlinear Models*. Springer-Verlag, Singapore.

Wiley, J. A., Herschokoru, S. J. & Padiau, N. S. (1989). Heterogeneity in the probability of HIV transmission per sexual contact: the case of male-to-female transmission in penilevaginal intercourse. *Statistics in Medicine*, **8**, 93–102.

# Livros Grátis

( <http://www.livrosgratis.com.br> )

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)  
[Baixar livros de Literatura de Cordel](#)  
[Baixar livros de Literatura Infantil](#)  
[Baixar livros de Matemática](#)  
[Baixar livros de Medicina](#)  
[Baixar livros de Medicina Veterinária](#)  
[Baixar livros de Meio Ambiente](#)  
[Baixar livros de Meteorologia](#)  
[Baixar Monografias e TCC](#)  
[Baixar livros Multidisciplinar](#)  
[Baixar livros de Música](#)  
[Baixar livros de Psicologia](#)  
[Baixar livros de Química](#)  
[Baixar livros de Saúde Coletiva](#)  
[Baixar livros de Serviço Social](#)  
[Baixar livros de Sociologia](#)  
[Baixar livros de Teologia](#)  
[Baixar livros de Trabalho](#)  
[Baixar livros de Turismo](#)