UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE MATEMÁTICA,
ESTATÍSTICA E COMPUTAÇÃO
CIENTÍFICA

DEPARTAMENTO DE ESTATÍSTICA

INFERÊNCIA E DIAGNÓSTICO EM MODELOS DE GRUBBS

Víctor Hugo Lachos Dávila

Orientador : Prof. Dr. Filidor E. Vilca Labra

- Campinas, Fevereiro de 2002 -

.

À minha mãe Rosa, ao espírito imortal de meu pai Roque (in memoriam) e a meus inseparáveis irmãos.

Agradecimentos

- Em primeiro lugar, quero agradecer a minha mãe Rosa, por nunca medir esforços na tarefa de me-educar.
- Ao Prof. Filidor Vilca Labra pela Confiança, sugestões, apoio e orientação segura durante a elaboração deste trabalho.
- À CAPES pelo apoio financeiro e ao curso de pós-graduação do IME-UNICAMP, pelas excelentes condições de trabalho.
- Aos professores do Departamento de Estatística, pelos ensinamentos concedidos e aos funcionários do IME-UNICAMP pela amizade demonstrada.
- A todos os amigos da pós que dividiram estes dois anos de estudo e a todos aqueles que de alguma forma, colaboraram na minha formação acadêmica e na execução da tese.
- E acima de tudo, agradeço a Deus pela capacidade e dádiva da vida.

Muito obrigado.

Resumo

Neste trabalho estudamos uma subclasse de modelos de calibração comparativa, proposto por Grubbs (1948, 1973), que assume que os instrumentos medem a característica de interesse na mesma escala.

Apresentamos um estudo de inferência no modelo de calibração comparativa estrutural, baseado no método de máxima verossimilhança. Estudos de testes de hipóteses são considerados utilizando as estatísticas de Wald, Escore e da Razão de Verossimilhanças. Usamos o algoritmo EM para obter os estimadores de máxima verossimilhança.

Utilizando o método de influência local proposto por Cook (1986) e Poon & Poon (1999), apresentamos um estudo de diagnóstico no modelo de Grubbs. Ilustramos a metodologia proposta com dados encontrados na literatura.

Abstract

In this work we study a special subclass of comparative calibration models proposed by Grubbs (1948, 1973). This subclass of models assume that the instruments measure the interested characteristic in the same scale.

We presented a study o statistical inference in comparative calibration structural models based in the maximum likelihood methods. Hypotheses test studies are considered using the Wald, Score and Likelihood Ratios statistics. The EM algorithm is used to obtained the maximum likelihood estimators.

Using the method of local influence, proposed by Cook (1986) and Poon & Poon (1999), we presented o study of diagnostics in the Grubbs models. We illustrated the proposed methodology with data set taken from the literature.

Conteúdo

Lista de Figuras

Lista de Tabelas

Capítulo 1

Introdução

1.1 O Problema de Calibração Comparativa

Calibração estatística é bastante importante em aplicações práticas de diversas áreas. Na física como na engenharia tem despertado muito interesse na aferição de instrumentos de medição de grandezas físicas (como amperímetros, trenas, etc.). Na medicina, é muito importante para a calibração de instrumentos indicadores do estado de saúde individual (termômetros, medidores de pressão sangüínea, etc.). Na biologia e na química a técnica é aplicada para determinar concentrações de soluções e composições de materiais através de fotômetros de chama, de cromatrógrafos, etc.

A calibração estatística é o processo através do qual a escala de um instrumento de medição é determinada ou ajustada a partir de um instrumento de calibração, conhecido na inferência bayesiana como experimento informativo.

Em qualquer discussão sobre calibração é importante fazer distinção entre "calibração absoluta"e "calibração comparativa". Williams (1969) enfatiza que o termo calibração, na literatura, é aplicado sem distinção aos dois tipos de problemas que são conceitualmente diferentes, e conduzindo a modelos estatísticos diferentes. Na calibração absoluta, uma técnica para a obtenção de medidas de forma rápida ou não-padrão é calibrada por uma medida conhecida ou feita com erro experimental desprezível. O uso deste termo é o que está mais próximo do significado original da palavra calibração, que segundo Willians (1969) correspondia à medida do calibre (ou diâmetro interno) de um tubo. Neste contexto, a calibração é uma forma de predição inversa que alguns autores chamam de regressão inversa.

Na calibração comparativa, pretende-se determinar a relação entre diversos testes ou instrumentos que fornecem medidas indiretas similares. Um exemplo é a capacidade pulmonar humana, que pode ser medida, por exemplo, por dois instrumentos, um mais tradicional, e outro mais moderno, portátil e de operação mais simples. Submetendo um grupo de indivíduos aos dois instrumentos, as medidas (ambas sujeitas a erros) dos dois aparelhos podem ser calibradas. Barnett (1969); Theobald e Mallison (1978) e Galea-Rojas (1995) discutem o problema de calibração comparativa.

A necessidade de se comparar instrumentos de medição tem aparecido com freqüência em diversas áreas do conhecimento, por exemplo: Grubbs (1948,1973), compara três cronômetros; Barnett (1969) dá um exemplo onde quatro combinações instrumento-operador concebidos para medir a capacidade vital num grupo de pacientes são avaliados e Jaech (1985) dá vários exemplos na área industrial. Mas recentemente, Christensen e Blackwood (1993) compara cinco "thermocouples"; Galea-Rojas (1995) apresenta um estudo de inferência no modelo "t-Student"e Bedrick (2001) compara três métodos para medir sedimentos de solo.

Em muitas aplicações científicas, por exemplo, na área industrial e agricultura, é assumido que todos os instrumentos medem a característica de interesse na mesma escala. Uma classe particular do modelo utilizado para comparar instrumentos de medição foi introduzida por Grubbs (1948,1973). Outros trabalhos mais recentes são, por exemplo, Christensen e Blackwood (1993) e Bedrick (2001), este último apresentando um estudo de inferência baseado na estatística de "Escore", usando uma parametrização diferente da qual será considerada neste trabalho. De fato, a parametrização que iremos considerar, corresponde a uma subclasse dos modelos de calibração comparativa introduzida por Barnett (1969), que assume que os instrumentos medem a caraterística de interesse na mesma escala. Este modelo é definido como segue:

Suponha que dispomos de p ($p \geq 3$) instrumentos para medir uma característica ou resposta x num grupo de n unidades. Sejam

x_j : O verdadeiro valor da característica de interesse na unidade j;

y_{ij} : Medida fornecida pelo instrumento i para a característica da unidade j, $i = 1, \ldots, p$ e $j = 1, \ldots, n$.

O modelo conhecido na literatura, como modelo de Grubbs, é definido pela relação

$$y_{ij} = \alpha_i + x_j + \varepsilon_{ij} \quad , \quad i = 1, \ldots, p, \quad j = 1, \ldots, n, \tag{1.1}$$

onde os erros aleatórios de medição ε_{ij} são não-correlacionados e para cada i, $\varepsilon_{i1}, \ldots, \varepsilon_{in}$ são tais que

$$E(\varepsilon_{ij}) = 0 \quad , \quad Var(\varepsilon_{ij}) = \phi_i \, , \tag{1.2}$$

e α_i é o vício aditivo do instrumento i , $i = 1, \ldots, p$.

O modelo definido em (1.1-1.2) representa uma classe particular de modelos de regressão linear com erros nas variáveis, amplamente discutida em Fuller (1987). Deste ponto de vista, o modelo definido em (1.1-1.2) pode ser especificado segundo as quantidades x_j como:

1. As $x_j's$ são constantes desconhecidas, chamadas de parâmetros incidentais.
2. As $x_j's$ são variáveis aleatórias não-correlacionadas com média μ_x e variância ϕ_x e são não-correlacionadas com ε_{ij} .
3. As $x_j's$ são variáveis aleatórias não-correlacionadas com média μ_{xj} e variância ϕ_x e são não-correlacionadas com os ε_{ij} .

O modelo descrito por (1) e (1.1-1.2) é denominado modelo funcional de Grubbs, neste caso o número de parâmetros cresce com o número de observações. Para fazer um estudo de inferência neste modelo é necessário considerar suposições adicionais de alguns parâmetros, especialmente sobre as quantidades desconhecidas $x_1, \ldots, x_n$; veja por exemplo, Vilca Labra et al. (1998). Neste caso os parâmetros do modelo são

$$(x_1, \ldots, x_n \, , \, \boldsymbol{\theta}')', \text{ onde } \boldsymbol{\theta} = (\alpha_1, \ldots, \alpha_p, \phi_1, \ldots, \phi_p)'.$$

O modelo descrito por (2) e (1.1-1.2) é denominado modelo estrutural de Grubbs, ao contrário do modelo funcional, neste modelo não é possível encontrar uma solução de máxima verossimilhança para os parâmetros do modelo, pois o mesmo não é identificável, no sentido que dois diferentes vetores de parâmetros podem dar origem a uma mesma distribuição. Assim, faz-se necessário o conhecimento adicional de algum parâmetro do modelo para possibilitar a estimação dos parâmetros de maior interesse. Neste caso o vetor de parâmetros do modelo é

$$\boldsymbol{\theta} = (\mu_x, \alpha_1, \ldots, \alpha_p, \phi_x, \phi_1, \ldots, \phi_p)'.$$

Finalmente, o modelo descrito por (3) e (1.1-1.2) é denominado modelo ultraestrutural. Neste caso, os modelos funcional e estrutural podem ser vistos como casos especiais deste modelo. Sob o modelo ultraestrutural os parâmetos do modelo são

$$(\mu_{x_1}, \ldots, \mu_{x_n}, \boldsymbol{\theta}')', \text{ onde } \boldsymbol{\theta} = (\alpha_1, \ldots, \alpha_p, \phi_x, \phi_1, \ldots, \phi_p)'.$$

Do ponto de vista de aplicação, o modelo funcional é adequado quando temos interesse apenas nos indivíduos participantes do estudo, enquanto que o estrutural é adequado quando o interesse é generalizar os resultados do estudo para a população de onde vem os indivíduos envolvidos no estudo.

O modelo estrutural definido em (1.1-1.2) e (2) pode ser escrito em forma matricial como:

$$\mathbf{Y}_j = \mathbf{a} + \mathbf{1}_p\, x_j + \boldsymbol{\epsilon}_j, \quad j = 1, \ldots, n, \tag{1.3}$$

onde $\mathbf{a} = (\alpha_1, \ldots, \alpha_p)'$, $\mathbf{Y}_j = (y_{1j}, \ldots, y_{pj})'$, $\boldsymbol{\epsilon}_j = (\varepsilon_{1j}, \ldots, \varepsilon_{pj})'$ são vetores $p \times 1$ e $\mathbf{1}_p$ é um vetor $p \times 1$ com todos os elementos iguais à unidade. Note que

$$E(\mathbf{Y}_j) = \mathbf{a} + \mathbf{1}_p\, \mu_x = \boldsymbol{\mu} \;\; e \;\; Var(\mathbf{Y}_j) = \mathrm{D}(\boldsymbol{\phi}) + \phi_x\, \mathbf{1}_p \mathbf{1}_p' = \boldsymbol{\Sigma}, \tag{1.4}$$

onde $\mathrm{D}(\boldsymbol{\phi}) = \mathrm{Diag}(\phi_1, \ldots, \phi_p)$ é uma matriz diagonal $p \times p$, $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_p)'$ e A' significa a transposta de A, com

$$\boldsymbol{\theta} = (\mu_x, \mathbf{a}', \phi_x, \boldsymbol{\phi}')'. \tag{1.5}$$

Sob as condições do modelo estrutural definido em (1.3), $\boldsymbol{\theta}$ não é identificável, isto é, existem $\boldsymbol{\theta}_*$ e $\boldsymbol{\theta}_{**}$ tais que $F(\mathbf{y}, \boldsymbol{\theta}_*) = F(\mathbf{y}, \boldsymbol{\theta}_{**})$, para todo $\mathbf{y} \in \mathbf{R}^p$, onde F é a função de distribuição de $\mathbf{Y}_j$, $j = 1, \ldots, n$. Por exemplo, para o caso de $p = 3$, sejam

$$\boldsymbol{\theta}_* = (1, 1, 1, 1, 1, 1, 1, 1)'$$

e

$$\boldsymbol{\theta}_{**} = (4, -2, -2, -2, 1, 1, 1, 1)' ,$$

de (1.4) é facil verificar que para ambos valores de $\boldsymbol{\theta}$ definidos acima, geram a mesma média e a mesma matriz de covariância de $\mathbf{Y}_j$, isto é

$$E(\mathbf{Y}_j) = (2, 2, 2)' ,$$

e

$$Var(\mathbf{Y}_j) = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} .$$

A falta de identificabilidade implica ausência de estimadores consistentes, o que tem serias implicações na teoria assintótica de estimação e testes de hipóteses. Contudo, o fato de ser identificável não implica que exista um estimador consistente para $\boldsymbol{\theta}$. Uma das formas de contornar o problema de identificabilidade é colocar restrições sobre o parâmetro $\boldsymbol{\theta}$. Similarmente como em Barnett (1969) vamos assumir que exista um método de medição, o qual chamamos instrumento de referência ou de controle, que mede x sem vício. Especificamente, supondo que o instrumento de referência seja o número um, vamos assumir que

$$\alpha_1 = 0 . \tag{1.6}$$

Assim, o modelo proposto por Grubbs (1948,1973) considerando um instrumento de referência, é dado por

$$\mathbf{Y}_j = \mathbf{a} + \mathbf{1}_p \, x_j + \boldsymbol{\epsilon}_j \, , j = 1, \ldots, n \, , \tag{1.7}$$

onde $\mathbf{a} = (0, \boldsymbol{\alpha}')'$, com $\boldsymbol{\alpha} = (\alpha_2, \ldots, \alpha_p)'$, $\mathbf{Y}_j$ e $\boldsymbol{\epsilon}_j$ como em (1.3). Agora, sob a condição de identificabilidade, o vetor de parâmetros dado em (1.5) reduz-se a um vetor de dimensão $(2p + 1) \times 1$ dado por

$$\boldsymbol{\theta} = (\mu_x, \boldsymbol{\alpha}', \phi_x, \boldsymbol{\phi}')' . \tag{1.8}$$

Em muitos trabalhos encontrados na literatura utilizando o modelo de Grubbs, a análise é feita desde um ponto de vista multivariado, onde não é necessário preocupar-se com a falta de identificabilidade; veja por exemplo, Christensen e Blackwood (1993), Deutler (1991), Brindley e Bradley (1995), entre outros.

Consideremos o modelo estrutural de Grubbs definido em (1.1) ou (1.7), isto é,

$$y_{ij} = \alpha_i + x_j + \varepsilon_{ij} \, , \tag{1.9}$$

onde $i = 1, \ldots, p$ e $j = 1, \ldots, n$. Na literatura é assumido que os ε_{ij} têm distribuição normal, são independentes e identicamente distribuídos (iid) através dos p instrumentos com média zero e variância $\phi_i \, , i = 1, \ldots, p$, que os verdadeiros valores x_j são variáveis aleatórias, tendo distribuição normal com média μ_x e variância ϕ_x, e independente dos ε_{ij}. Assim, $y_{ij} \sim N(\alpha_i + \mu_x, \phi_x + \phi_i)$ e $cov(y_{ij}, y_{kj}) = \phi_x \, , i \neq k$, $i, k = 1, ..., p \, , \, j = 1, ..., n$, com $\alpha_1 = 0$.

O estudo apresentado por Grubbs (1948, 1973) é baseado considerando as diferenças pareadas $d_{ijk} = y_{ik} - y_{jk}$ e em um estimador tipo momentos, que tem sido contestado, devido às dificuldades encontradas no desenvolvimento de procedimentos de inferência estatística. Jaech (1985) faz um estudo baseado também nas diferenças pareadas e nos dados originais, mas aplica o princípio de máxima verossimilhança somente nas variâncias $\phi_i \, , \; i = 1, \ldots, p, \;$ e ϕ_x. Portanto, uma contribuição importante neste trabalho é a apresentação de um procedimento estatístico baseado no método de máxima verossimilhança para os dados originais, aplicável em todos os parâmetros do modelo.

No contexto do modelo de Grubbs a qualidade de um instrumento de medição é avaliada em termos do vício ou acurácia e da precisão (inversa da variância). Contudo é importante distinguir os conceitos de acurácia e precisão; precisão se refere à proximidade de cada observação à sua própria média, enquanto que acurácia mede a proximidade das medidas ao verdadeiro valor. Assim, a acurácia está relacionada com os vícios, enquanto que precisão está relacionada com a dimensão dos erros de medida, ou seja, um instrumento é preciso se ϕ_i é pequena.

Pelo exposto acima, o problema de comparar instrumentos ou métodos de medição pode ser estudado fazendo inferência sobre os vícios $\alpha_2, \ldots, \alpha_p$ dos instrumentos $2, \ldots, p \, ,$ respectivamente, e comparando-os com o instrumento de referência. A precisão de um instrumento, que é baseado na variancia é um índice ou medida para avaliar a qualidade de um instrumento, veja por exemplo, Shyr e Gleser (1986).

A precisão do i-ésimo instrumento, denotada por π_i, é definida como sendo:

$$\pi_i = \frac{1}{\phi_i} \quad , \quad i = 1, \ldots, p\,. \tag{1.10}$$

Outra medida ou índice utilizado na literatura, no contexto de calibração comparativa, é a confiabilidade do instrumento, veja por exemplo, Carter (1981). A confiabilidade do i-ésimo instrumento, denotado por ρ_i, define-se por

$$\rho_i = \frac{\phi_x}{\phi_x + \phi_i} \quad , \quad i = 1, \ldots, p\,, \tag{1.11}$$

o coeficiente ρ_i é uma medida tipo coeficiente de determinação usada em regressão linear e pode ser interpretada como a proporção da variação observada "explicada"pelo i-ésimo instrumento, ou seja, $\rho_i = \rho^2_{y_{ij}\,,x_j}$ e $0 \leq \rho_i \leq 1$, $i = 1, \ldots, p$.

Das relações em (1.10) e (1.11), pode-se mostrar facilmente que

$$\rho_k = \rho_l \Leftrightarrow \pi_k = \pi_l \Leftrightarrow \phi_k = \phi_l \quad , l \neq k = 1, 2, ..., p \quad , \tag{1.12}$$

isto é, comparar confiabilidade dos instrumentos no modelo de Grubbs é equivalente a comparar as precisões dos instrumentos ou as variâncias dos erros.

Assim, comparar p instrumentos de medição utilizando o modelo de Grubbs reduz-se a fazer inferências sobre os vícios α_i e as variâncias $\phi_1, \ldots, \phi_p$. Considerar os vícios e as variâncias simultaneamente é importante para verificar se o instrumento i fornece medidas exatas da característica x, pois um valor de ϕ_i próximo de zero implica só que y_i (medida fornecida pelo instrumento i) é uma medida confiável ou de alta precisão de $\alpha_i + x$. O instrumento i forneceria medidas exatas da característica x só se ϕ_i for próximo de zero e se ele é não-viciado ($\alpha_i = 0$).

Assim, uma hipótese inicial de interesse usando o modelo (1.7) é

$$H_{01} \;:\; \alpha_2 = \alpha_3 = ... = \alpha_p = 0 \;\text{ e }\; \pi_1 = \pi_2 = ... = \pi_p \quad , \tag{1.13}$$

ou equivalentemente, de (1.12)

$$H_{01} \;:\; \alpha_2 = \alpha_3 = ... = \alpha_p = 0 \;\text{ e }\; \phi_1 = \phi_2 = ... = \phi_p \quad , \tag{1.14}$$

ou seja, estamos interessados em testar se os p instrumentos medem a característica x sem vício (aditivo) ou acurácia e com a mesma precisão.

Outras hipóteses que poderiam ser consideradas são:

$$H_{02} \ : \ \alpha_2 = \alpha_3 = ... = \alpha_p = 0 \quad , \tag{1.15}$$

isto é, estamos avaliando a accurácia das medições feitas pelos diferentes instrumentos. Para avaliar se os instrumentos são igualmente precisos, a hipótese poderia ser

$$H_{03} \ : \ \phi_1 = \phi_2 = ... = \phi_p \quad . \tag{1.16}$$

Logicamente, dependendo do problema, poderia ser de interesse testar outro tipo de hipótese, embora as mencionadas acima sejam de interesse geral.

Para testar as hipóteses acima, vamos utilizar as estatísticas da Razão de Verossimilhanças, Escore e de Wald. Estimadores de máxima verossimilhança sob H_0 e sob o modelo irrestrito são apresentados no decorrer deste trabalho.

Para $p = 2$, Grubbs (1948,1973) apresenta um estudo de inferência estatística baseado em técnicas multivariadas. Maloney e Rastogi (1970) mostram que o teste de Pitman (1939) é equivalente a testar H_{03}. O teste de Pitman é baseado em avaliar a significância do coeficiente de correlação entre as somas e diferenças do conjunto de pares de medida. Blackwood e Bradley (1991) desenvolveram um teste para testar H_{01} usando regressão simples.

Para $p \geq 2$ Choi e Wette (1972) desenvolveram um teste para avaliar a igualdade da variância dos instrumentos (H_{03}) e mostra que é equivalente ao teste de Pitman (1939) para o caso $p = 2$. Para $p = 3$, Grubbs (1973) apresenta um estudo considerando conceitos de acurácia e precisão dos instrumentos. Jaech (1985, pag. 215) desenvolve um teste para avaliar a igualdade das variâncias H_{03} baseado no teste da Razão das Verossimilhanças, assumindo que a soma dos vícios α_i , $i = 1, \ldots, p$, é zero e aplica o princípio de máxima verossimilhança somente nas variâncias do modelo. Recentemente, Christensen e Blackwood (1993) propõe um modelo linear multivariado para testar H_{01}, H_{02} e H_{03}, e mostram que o teste para H_{01} é equivalente a testar $H_0 : \gamma_i = \delta_i = 0\,, \quad i = 1, \ldots, p-1$, no modelo de regressão multivariado

$$y_{ij} - \bar{y}_{.j} = \delta_i + \gamma_i \bar{y}_{.j} + \varepsilon_{ij}\,, \qquad j = 1, \ldots, n, \;\; i = 1, \ldots, p-1\,, \tag{1.17}$$

Bedrick (2001) desenvolve um teste baseado na estatística de "Escore", mas usando uma parametrização diferente à parametrização considerada neste trabalho.

1.2 Diagnóstico

Modelos estatísticos são ferramentas úteis para extrair e compreender as características essenciais de um conjunto de dados. Modelos são, quase sempre, descrições aproximadas de processos reais bem complexos e portanto são quase sempre inexatos. Devido a estas inexatitudes, o estudo de uma análise sobre pequenas modificações da formulação do problema torna-se importante. Se uma modificação pequena de uma descrição aproximada afeta seriamente os resultados de uma análise, esta deve ser o centro de nosso interesse. Por outro lado, se as alterações encontradas são desprezíveis, a amostra é robusta com respeito às perturbações induzidas e nossa ignorância do modelo preciso poderia não causar dano.

Se o modelo ajustado aos dados não apresentar uma boa descrição dos dados que foram observados, a utilização do modelo pode conduzir à inferências errôneas. Por esta razão é importante que se faça um estudo sobre robustez dos resultados obtidos, em termos dos vários aspectos que envolvem a formulação do modelo, e as estimativas dos seus parâmetros. Ou seja, uma *análise de diagnóstico* consiste de métodos para avaliar o grau de sensibilidade das inferências à pequenas perturbações nos dados ou no modelo proposto.

Durante a construção de um modelo é importante examinar os pontos amostrais que estão sendo usados, com o objetivo de determinar se existe alguns deles que estejam controlando propriedades importantes no modelo. Se os pontos influentes detectados tivessem surgido como resultado da ocorrência de eventos não usuais (tais como registros incorretos dos dados, leitura errônea no levantamento da informação, erros de medição ou de cálculo ou funcionamento inadequado de algum equipamento), então estes pontos devem ser corrigidos, se isto for possível, ou excluilos do conjunto de dados. Por outro lado, pode não ter havido nada errado durante a obtenção destes pontos, mas mesmo assim é importante descobrir se estes controlam propriedades importantes do modelo, já que a ocorrência deste fato pode afetar o uso do modelo. Numa análise estatística é importante levar em consideração a estabilidade dos resultados obtidos e das inferências de um modelo geral, com relação a possíveis perturbações nos dados ou no modelo estatístico.

As análises de resíduos são feitas para investigar eventuais problemas com o modelo ajustado, enquanto que uma análise de diagnóstico é feita assumindo o modelo como correto e então investigando a robustez das conclusões a pequenas perturbações.

Os elementos do conjunto de dados que efetivamente controlam aspectos da análise são ditos influentes. Em particular, uma observação, é dita influente se ela produzir alterações relevantes no resultado de análise quando for excluída, ou submetida a uma pequena perturbação.

Uma noção geral para avaliar a influência de pequenas perturbações no modelo ou nos dados é proposto por Cook (1986), que se baseia na análise das curvaturas das seções normais do gráfico de influência, numa vizinhança do ponto em que o modelo postulado e o modelo perturbado coincidem. Este método é conhecido como *Método de Influência Local*, que permite avaliar a influência que pequenas perturbações podem exercer sobre os componentes do modelo, tais como estimativa dos parâmetros, e outros resultados da análise. Em particular, o método proposto por Cook (1986), permite avaliar o efeito da exclusão de alguma observação na estimativa dos parâmetros.

O método proposto por Cook (1986), pode ser aplicado em uma grande variedade de modelos estatísticos. Esta técnica de diagnóstico tem sido estudada por muitos outros autores, dentre os quais citamos: Thomas e Cook (1990), que aplicam o método de influência local para predições em modelos lineares generalizados; Laurent e Cook (1993) consideram a influência local em um modelo de regressão não-linear, enfatizando a relação entre o jacobiano do "leverage"e a influência local sobre os valores do modelo ajustado; Paula (1996) faz uma extensão desse estudo para os modelos próprios de dispersão, proposto por Jorgensen (1997); Lee e Zhao (1996) consideram o método de influência local em modelos lineares generalizados quando as covariáveis são medidas sem erro e Lesaffre e Verbeke (1998) consideram o método no efeito do modelo linear misto.

Alguns trabalhos recentes na área de modelos com erros nas variáveis são por exemplo: Kim (2000) que aplica o método de influência local no modelo com erro nas variáveis simples, amplamente estudado em Fuller (1987); Galea-Rojas, Bolfarine e Castro (1999) aplicam o método no modelo de calibração comparativa, estrutural e funcional considerado por Barnett (1969) e Kimura (1992), respectivamente e Aoki (2001) que apresenta um estudo de influência no modelo com erros nas variáveis com intercepto nulo.

1.3 Objetivo do Trabalho

O objetivo deste trabalho é apresentar um estudo de inferência estatística e diagnóstico no modelo estrutural de Grubbs, para $p \geq 3$ instrumentos. O estudo de diagnóstico é baseado no método de influência proposto por Cook (1986) e Poon e Poon (1999), cuja implementação depende da estimativa de máxima verossimilhança sob o modelo proposto e do esquema de perturbação considerado. A obtenção dos estimadores de máxima verossimilhança serão via algoritmo EM (Dempster et al. 1997). Baseado no método de máxima verossimilhança apresentamos um estudo de testes de hipóteses para aquelas hipóteses de interesse. Assim, os objetivos específicos deste trabalho podem ser resumidos como segue:

1. Apresentar um estudo de inferência estatística no modelo de calibração estrutural de Grubbs, o que inclui obter a matriz de Fisher observada e esperada que será de utilidade para testar hipóteses de interesse e para desenvolver os métodos de diagnóstico.

2. Aplicar o método de influência local proposto por Cook (1986) no modelo estrutural de Grubbs, considerando alguns esquemas de perturbação, por exemplo, ponderação de casos e perturbação na variável resposta. O método de influência local proposto por Poon e Poon (1999) também será considerado.

1.4 Organização do Trabalho

Este trabalho consta de 4 capítulos e 3 apêndices. No capítulo 2 apresentamos um estudo de inferência no modelo de calibração comparativa proposto por Grubbs (1948, 1973). São obtidas as matrizes de informação observada e esperada de Fisher. Estimação de parâmetros via algoritmo EM e testes de hipóteses são discutidos, baseados nas estatísticas de Wald, Escore e Razão das Verossimilhanças. Um pequeno estudo de simulação é apresentado para estudar o comportamento dos três testes. Estimadores de momentos são também considerados, assim como suas propriedades assintóticas.

No capítulo 3 tratamos da aplicação dos métodos de influência local propostos por Cook (1986) e Poon e Poon (1999). Consideramos dois tipos de perturbação: ponderação de casos e perturbação na variável resposta. Para cada caso é apresentada uma aplicação dos resultados obtidos a um conjunto de dados reais encontrados na literatura.

O capítulo 4 é dedicado à conclusões e alguns direcionamentos para estudos subseqüentes.

Finalmente, nos apêndices, apresentamos alguns resultados da literatura para obter a matriz de informação esperada, além de um programa implementado no software Matlab 6.0 utilizado para obter os EMV para os parâmetros do modelo.

Quando mencionarmos manipulações algébricas no decorrer da tese, estamos referindo à manipulações algébricas desenvolvidas "manualmente" e conferidas através do software Matlab 6.0.

Capítulo 2

Modelo de Grubbs

2.1 Introdução

Em muitas aplicações científicas e de engenharia é necessário avaliar a qualidade relativa de vários instrumentos utilizados para fazer uma medida particular. Métodos de medida freqüentemente diferem em custo, velocidade ou outros fatores. Estes vários instrumentos podem ser um instrumento avaliado várias vezes e nosso interesse seria se o vício e a precisão dos instrumentos variam com o tempo. Freqüentemente a avaliação dos instrumentos é feita baseada em dados obtidos utilizando cada instrumento para medir uma característica comum em varias unidades experimentais. Em muitas situações de aplicação, é necessário a comparação de instrumentos que medem a característica de interesse na mesma escala, por exemplo, na área industrial. Um modelo que tem sido muito utilizado na literatura é o modelo proposto por Grubbs (1948, 1973).

O objetivo principal deste capítulo é apresentar um estudo de inferência estatística sob o modelo estrutural de Grubbs, onde é de interesse comparar vários instrumentos de medição da mesma quantidade desconhecida x em um grupo comum de n unidades experimentais, supondo que as observações obtidas pelos diferentes instrumentos seguem uma distribuição normal multivariada e a característica de interesse x é medida na mesma escala. No contexto do modelo de Grubbs, a qualidade de um instrumento é avaliada em termos da precisão (inversa da variância) e da acurácia (vício) dos diferentes instrumentos.

Na Seção 2.2 apresentamos a matriz de informação observada e esperada de Fisher, a implementação do algoritmo EM para a obtenção do estimador de máxima verossimilhança (EMV) e sua distribuição assintótica.

A Seção 2.3 é dedicada à obtenção do EMV restrito. Em alguns casos é apresentada sua distribuição assintótica e a implementação do algoritmo EM para a obtenção dos EMV restritos.

A Seção 2.4 é dedicada a discussão de testes de hipóteses usando as estatísticas de Wald, Escore e Razão de Verossimilhanças, onde um estudo de simulação é apresentado.

Finalmente, na Seção 2.5 apresentamos um estudo do comportamento assintótico do estimador de momentos proposto por Grubbs (1948).

2.2 Modelo de Grubbs Estrutural Normal

Nesta seção apresentamos um estudo estatístico do modelo estrutural de Grubbs definido em (1.7), supondo que as observações da característica de interesse x seguem uma distribuição normal.

Consideremos novamente o modelo definido em (1.7)

$$\mathbf{Y}_j = \mathbf{a} + \mathbf{1}_p\, x_j + \boldsymbol{\epsilon}_j \quad j = 1, \ldots, n\,, \tag{2.1}$$

onde o modelo normal é obtido ao considerar

$$\boldsymbol{\epsilon}_j \stackrel{\text{iid}}{\sim} N_p\left(\mathbf{0}, \mathrm{D}(\boldsymbol{\phi})\right) \;\text{ e }\; x_j \stackrel{\text{iid}}{\sim} N(\mu_x\,,\, \phi_x)\,, j = 1, \ldots, n\,, \tag{2.2}$$

com $\boldsymbol{\epsilon}_j$ e x_j independentes. De (2.1), (2.2) e das propriedades da distribuição normal temos que

$$\mathbf{Y}_j \stackrel{\text{iid}}{\sim} N_p(\boldsymbol{\mu}\,,\, \boldsymbol{\Sigma})\,,\; j = 1, \ldots, n\,, \tag{2.3}$$

onde $\boldsymbol{\mu} = \mathbf{a} + \mathbf{1}_p\, \mu_x$ e $\boldsymbol{\Sigma} = \phi_x \mathbf{1}_p \mathbf{1}_p' + \mathrm{D}(\boldsymbol{\phi})$, com $\mathrm{D}(\boldsymbol{\phi})$ como em (1.4) e $\mathbf{a}$ como em (1.7).

2.2.1 Matriz de Informação Observada

Nesta subseção, vamos apresentar a matriz de informação observada baseada no modelo estrutural de Grubbs definido em (2.1). Esta matriz será utilizada posteriormente para aplicar o método de influência local. Além disso, será utilizada para obter a matriz de informação esperada, amplamente usada em procedimentos de inferência estatística baseado no método de máxima verossimilhança.

Sob normalidade, a função de densidade de probabilidade do modelo de Grubbs definido em (2.1-2.2) é dada por

$$f(\mathbf{y}_j;\boldsymbol{\theta}) = (2\pi)^{-p/2}|\boldsymbol{\Sigma}|^{-1/2}\exp\{-\frac{1}{2}(\mathbf{y}_j-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{y}_j-\boldsymbol{\mu})\}\,, \tag{2.4}$$

onde

$$\boldsymbol{\Sigma} = \phi_x\,\mathbf{1}_p\mathbf{1}_p' + \mathrm{D}(\boldsymbol{\phi})\ ,\ \ \boldsymbol{\Sigma}^{-1} = \mathrm{D}^{-1}(\boldsymbol{\phi}) - c^{-1}\mathbf{M}\ e\ |\boldsymbol{\Sigma}| = c|\mathrm{D}(\boldsymbol{\phi})|,$$

com $c = 1+\phi_x\mathbf{1}_p'D^{-1}(\boldsymbol{\phi})\mathbf{1}_p$, $\mathbf{M} = \phi_x\mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{1}_p\mathbf{1}_p'\mathrm{D}^{-1}(\boldsymbol{\phi})$ e $\boldsymbol{\mu}$ definido em (2.3).

Logo, para uma amostra aleatória $\mathbf{Y}_1,...,\mathbf{Y}_n$, a função de verossimilhança é dada por

$$L(\boldsymbol{\theta};\mathbf{y}_1,...,\mathbf{y}_n) = \prod_{j=1}^{n}(2\pi)^{-p/2}|\boldsymbol{\Sigma}|^{-1/2}\exp\{-\frac{1}{2}(\mathbf{y}_j-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{y}_j-\boldsymbol{\mu})\}\,,$$

e a função de log-verossimilhança pode ser escrita como

$$\ell(\boldsymbol{\theta}) = \sum_{j=1}^{n}\ell_j(\boldsymbol{\theta}) = \sum_{j=1}^{n}\log f(\mathbf{y}_j;\boldsymbol{\theta}), \tag{2.5}$$

onde

$$\ell_j(\boldsymbol{\theta}) = -\frac{p}{2}\log 2\pi - \frac{1}{2}\log|\boldsymbol{\Sigma}| - \frac{1}{2}\|\mathbf{T}_j\|^2,$$

com $\|\mathbf{T}_j\|^2 = (\mathbf{y}_j-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{y}_j-\boldsymbol{\mu})$.

A fim de usar os resultados no capítulo 3, relacionados com o método de influência local, precisaremos das derivadas parciais de primeira e segunda ordem da função $\ell(\boldsymbol{\theta})$ com relação a $\boldsymbol{\theta}$.

a) Derivadas parciais de primeira ordem de $\ell(\boldsymbol{\theta})$

Como $\ell(\boldsymbol{\theta}) = \sum_{j=1}^{n} \ell_j(\boldsymbol{\theta})$, então $\dfrac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \sum_{j=1}^{n} \dfrac{\partial \ell_j(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$, onde

$$\begin{aligned}
\frac{\partial \ell_j(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} &= \frac{\partial \ell_j}{\partial \boldsymbol{\theta}} = \begin{pmatrix} \dfrac{\partial \ell_j}{\partial \mu_x} \\ \dfrac{\partial \ell_j}{\partial \boldsymbol{\alpha}} \\ \dfrac{\partial \ell_j}{\partial \phi_x} \\ \dfrac{\partial \ell_j}{\partial \boldsymbol{\phi}} \end{pmatrix} = -\frac{1}{2}\frac{\partial}{\partial \boldsymbol{\theta}} \log|\boldsymbol{\Sigma}| - \frac{1}{2}\frac{\partial}{\partial \boldsymbol{\theta}} \|\mathbf{T}_j\|^2 \\
&= -\frac{1}{2}\left[\frac{\partial}{\partial \boldsymbol{\theta}} \log|\boldsymbol{\Sigma}| + \frac{\partial}{\partial \boldsymbol{\theta}} \|\mathbf{T}_j\|^2\right]. \qquad (2.6)
\end{aligned}$$

Assim, temos que as derivadas de primeira ordem de $\ell(\boldsymbol{\theta})$ dependem das derivadas de $\log|\boldsymbol{\Sigma}|$ e $\|\mathbf{T}_j\|^2$ com respeito a μ_x, $\boldsymbol{\alpha}$, ϕ_x *e* $\boldsymbol{\phi}$, respectivamente.

Consideremos as seguintes notações: Sejam $\mathbf{a} = (a_1, ..., a_p)'$ um vetor em $\mathbf{R}^p$ e $\mathbf{W}_j = \mathbf{Y}_j - \boldsymbol{\mu}$, $\mathbb{I}_{(p)} = [0, \mathbf{I}_q]$ matriz $q \times p$, com $\mathbf{I}_q$ matriz identidade de ordem $q = p - 1$, então definimos $\mathrm{D}(\mathbf{a}) = \mathrm{Diag}(a_1, ..., a_p)$ e $\mathrm{D}^{-k}(\mathbf{a}) = \mathrm{Diag}(a_1^{-k}, ..., a_p^{-k})$. A última notação também é valida se $\mathbf{a}$ for um vetor aleatório.

Considerando as notações acima, derivando em forma vetorial e após algumas manipulações algébricas, temos que

$$\frac{\partial}{\partial \mu_x} \log|\boldsymbol{\Sigma}| = 0 \ ,$$

$$\frac{\partial}{\partial \boldsymbol{\alpha}} \log|\boldsymbol{\Sigma}| = \mathbf{0}_q \ ,$$

$$\frac{\partial}{\partial \phi_x} \log|\boldsymbol{\Sigma}| = \frac{c-1}{c\phi_x} \ ,$$

$$\frac{\partial}{\partial \boldsymbol{\phi}} \log|\boldsymbol{\Sigma}| = -c^{-1}\phi_x \mathrm{D}^{-2}(\boldsymbol{\phi})\mathbf{1}_p + \mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{1}_p \ ,$$

$$\frac{\partial}{\partial \mu_x} \|\mathbf{T}_j\|^2 = -2\mathbf{1}_p'\boldsymbol{\Sigma}^{-1}\mathbf{W}_j \ ,$$

$$\frac{\partial}{\partial \boldsymbol{\alpha}} \|\mathbf{T}_j\|^2 = -2\mathbb{I}_{(p)}\boldsymbol{\Sigma}^{-1}\mathbf{W}_j \ ,$$

$$
\begin{aligned}
\frac{\partial}{\partial \phi_x}\|\mathbf{T}_j\|^2 &= -\frac{c^{-2}}{\phi_x}\mathbf{W}_j'\mathbf{M}\mathbf{W}_j \ , \\
\frac{\partial}{\partial \boldsymbol{\phi}}\|\mathbf{T}_j\|^2 &= -\mathrm{D}(\mathbf{W}_j)\mathrm{D}^{-2}(\boldsymbol{\phi})\mathbf{W}_j - c^{-2}\phi_x\mathbf{W}_j'\mathbf{M}\mathbf{W}_j D^{-2}(\boldsymbol{\phi})\mathbf{1}_p \\
&\quad + 2c^{-1}\phi_x\mathbf{W}_j'\mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{1}_p\mathrm{D}^{-2}(\boldsymbol{\phi})\mathbf{W}_j \ ,
\end{aligned}
$$

com c e $\mathbf{M}$ como em (2.4). Logo, o vetor de derivadas parciais de primeira ordem $\frac{\partial \ell_j}{\partial \boldsymbol{\theta}}$ é dado por

$$
\frac{\partial \ell_j}{\partial \boldsymbol{\theta}} = \begin{pmatrix}
\mathbf{1}_p'\,\boldsymbol{\Sigma}^{-1}\mathbf{W}_j \\
\mathbb{I}_{(p)}\boldsymbol{\Sigma}^{-1}\mathbf{W}_j \\
\frac{1}{2\phi_x}c^{-1}(1-c) + \frac{1}{2}\frac{c^{-2}}{\phi_x}\mathbf{W}_j'\mathbf{M}\mathbf{W}_j \\
\frac{1}{2}c^{-1}\phi_x\mathrm{D}^{-2}(\boldsymbol{\phi})\mathbf{1}_p - \frac{1}{2}\mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{1}_p + \frac{1}{2}\mathrm{D}(\mathbf{W}_j)\mathrm{D}^{-2}(\boldsymbol{\phi})\mathbf{W}_j \\
+\frac{c^{-2}}{2}\phi_x\mathbf{W}_j'\mathbf{M}\mathbf{W}_j\mathrm{D}^{-2}(\boldsymbol{\phi})\mathbf{1}_p - c^{-1}\phi_x\mathbf{W}_j'\mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{1}_p\,\mathrm{D}^{-2}(\boldsymbol{\phi})\mathbf{W}_j
\end{pmatrix}. \tag{2.7}
$$

Esta matriz é importante, na implementação do método de influência local proposto por Cook (1986), principalmente na perturbação de ponderação de casos, e na obtenção da estatística de Escore.

b) Derivadas parciais de segunda ordem de $\ell(\boldsymbol{\theta})$

Seja $\ell(\boldsymbol{\theta})$ a função de log-verossimilhança definida em (2.5). Então, a matriz de derivadas parciais de segunda ordem com respeito de $\boldsymbol{\theta}$, é dada por

$$
\ddot{L} = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'} = \sum_{j=1}^{n}\frac{\partial^2 \ell_j}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'}, \tag{2.8}
$$

onde

$$
\frac{\partial^2 \ell_j}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'} = -\frac{1}{2}\left[\frac{\partial^2 \log|\boldsymbol{\Sigma}|}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'} + \frac{\partial^2\,\|\mathbf{T}_j\|^2}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'}\right].
$$

Assim, temos que as derivadas de segunda ordem de $\ell(\boldsymbol{\theta})$ dependem das derivadas de $\log|\boldsymbol{\Sigma}|$ e de $\|\mathbf{T}_j\|^2$ com respeito a μ_x, $\boldsymbol{\alpha}$, ϕ_x e $\boldsymbol{\phi}$ respectivamente.

Considerando as propriedades de derivadas vetoriais e após algumas manipulações algébricas, temos que

$$\frac{\partial^2 \log|\boldsymbol{\Sigma}|}{\partial \mu_x \partial \lambda'} = 0 \ , \ \ \lambda = \mu_x, \boldsymbol{\alpha}, \phi_x, \boldsymbol{\phi} \, ,$$

$$\frac{\partial^2 \log|\boldsymbol{\Sigma}|}{\partial \boldsymbol{\alpha} \partial \lambda'} = 0 \ , \ \ \lambda = \boldsymbol{\alpha}, \phi_x, \boldsymbol{\phi} \, ,$$

$$\frac{\partial^2 \log|\boldsymbol{\Sigma}|}{\partial \phi_x \partial \phi_x} = -\frac{1}{\phi_x^2}\left(\frac{c-1}{c}\right)^2 , \qquad \frac{\partial^2 \log|\boldsymbol{\Sigma}|}{\partial \phi_x \partial \boldsymbol{\phi}'} = -c^{-2}\mathbf{1}_p \mathrm{D}^{-2}(\boldsymbol{\phi}) \ ,$$

$$\frac{\partial^2 \log|\boldsymbol{\Sigma}|}{\partial \boldsymbol{\phi} \partial \boldsymbol{\phi}'} = -\mathrm{D}^{-2}(\boldsymbol{\phi}) + 2\phi_x c^{-1}\mathrm{D}^{-3}(\boldsymbol{\phi}) - \phi_x c^{-2}\mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{M}\mathrm{D}^{-1}(\boldsymbol{\phi}),$$

$$\frac{\partial^2 \|\mathbf{T}_j\|^2}{\partial \mu_x \partial \mu_x} = 2\mathbf{1}_p'\boldsymbol{\Sigma}^{-1}\mathbf{1}_p \ ,$$

$$\frac{\partial^2 \|\mathbf{T}_j\|^2}{\partial \mu_x \partial \boldsymbol{\alpha}'} = 2\mathbf{1}_p'\boldsymbol{\Sigma}^{-1}\mathbb{I}_{(p)}' \ ,$$

$$\frac{\partial^2 \|\mathbf{T}_j\|^2}{\partial \mu_x \partial \phi_x} = 2\frac{c^{-2}}{\phi_x}(c-1)\mathbf{1}_p'\mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{W}_j \ ,$$

$$\frac{\partial^2 \|\mathbf{T}_j\|^2}{\partial \mu_x \partial \boldsymbol{\phi}'} = 2c^{-1}\mathbf{W}_j'\boldsymbol{\Sigma}^{-1}\mathrm{D}^{-1}(\boldsymbol{\phi}) \ ,$$

$$\frac{\partial^2 \|\mathbf{T}_j\|^2}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\alpha}'} = 2\mathbb{I}_{(p)}\boldsymbol{\Sigma}^{-1}\mathbb{I}_{(p)}' \ ,$$

$$\frac{\partial^2 \|\mathbf{T}_j\|^2}{\partial \boldsymbol{\alpha} \partial \phi_x} = 2\frac{c^{-2}}{\phi_x}\mathbb{I}_{(p)}\mathbf{M}\mathbf{W}_j \ ,$$

$$\begin{aligned}\frac{\partial^2 \|\mathbf{T}_j\|^2}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\phi}'} = {} & 2\mathbb{I}_{(p)}[\boldsymbol{\Sigma}^{-1}\mathrm{D}(\mathbf{W}_j)\mathrm{D}^{-1}(\boldsymbol{\phi}) + c^{-2}\phi_x\mathbf{M}\mathbf{W}_j\mathbf{1}_p'\mathrm{D}^{-2}(\boldsymbol{\phi}) \\ & - c^{-1}\phi_x\mathbf{1}_p'\,\mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{W}_j\mathrm{D}^{-2}(\boldsymbol{\phi})],\end{aligned}$$

$$\frac{\partial^2 \|\mathbf{T}_j\|^2}{\partial \phi_x \partial \phi_x} = 2\frac{c^{-3}}{\phi_x}\mathbf{W}_j'\mathbf{M}\mathbf{W}_j\mathbf{1}_p'\mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{1}_p \ ,$$

$$\frac{\partial^2\|\mathbf{T}_j\|^2}{\partial\phi_x\partial\boldsymbol{\phi}'} = -2c^{-3}\mathbf{W}_j'\mathbf{M}\mathbf{W}_j\mathbf{1}_p'\mathrm{D}^{-2}(\boldsymbol{\phi}) + 2c^{-2}\mathbf{W}_j'\mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{1}_p\mathbf{W}_j'\mathrm{D}^{-2}(\boldsymbol{\phi}) \ ,$$

$$\begin{aligned}\frac{\partial^2\|\mathbf{T}_j\|^2}{\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}'} &= 2\mathrm{D}^2(\mathbf{W}_j)\mathrm{D}^{-3}(\boldsymbol{\phi}) - 2c^{-3}\phi_x\mathbf{W}_j'\mathbf{M}\mathbf{W}_j\mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{M}\mathrm{D}^{-1}(\boldsymbol{\phi})\\ &+ 2\phi_x c^{-2}\mathrm{D}^{-2}(\boldsymbol{\phi})\mathbf{W}_j\mathbf{W}_j'\mathbf{M}\mathrm{D}^{-1}(\boldsymbol{\phi}) + 2c^{-2}\phi_x\mathbf{W}_j'\mathbf{M}\mathbf{W}_j\mathrm{D}^{-3}(\boldsymbol{\phi})\\ &+ 2c^{-2}\phi_x\mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{M}\mathbf{W}_j\mathbf{W}_j'\mathrm{D}^{-2}(\boldsymbol{\phi}) - 2c^{-1}\phi_x\mathrm{D}^{-2}(\boldsymbol{\phi})\mathbf{W}_j\mathbf{W}_j'\mathrm{D}^{-2}(\boldsymbol{\phi})\\ &- 4c^{-1}\phi_x\mathbf{1}_p'\mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{W}_j\mathrm{D}(\mathbf{W}_j)\mathrm{D}^{-3}(\boldsymbol{\phi})\,.\end{aligned}$$

Logo, de (2.8) e dos resultados acima, a matriz $\dfrac{\partial^2\ell_j}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}$ é dada por

$$\frac{\partial^2\ell_j}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'} = \begin{pmatrix} \dfrac{\partial^2\ell_j}{\partial\mu_x\partial\mu_x} & \dfrac{\partial^2\ell_j}{\partial\mu_x\partial\boldsymbol{\alpha}'} & \dfrac{\partial^2\ell_j}{\partial\mu_x\partial\phi_x} & \dfrac{\partial^2\ell_j}{\partial\mu_x\partial\boldsymbol{\phi}'}\\ & \dfrac{\partial^2\ell_j}{\partial\boldsymbol{\alpha}\partial\boldsymbol{\alpha}'} & \dfrac{\partial^2\ell_j}{\partial\boldsymbol{\alpha}\partial\phi_x} & \dfrac{\partial^2\ell_j}{\partial\boldsymbol{\alpha}\partial\boldsymbol{\phi}'}\\ & & \dfrac{\partial^2\ell_j}{\partial\phi_x\partial\phi_x} & \dfrac{\partial^2\ell_j}{\partial\phi_x\partial\boldsymbol{\phi}'}\\ Simetria & & & \dfrac{\partial^2\ell_j}{\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}'} \end{pmatrix}, \tag{2.9}$$

onde

$$\frac{\partial^2\ell_j}{\partial\mu_x\partial\mu_x} = -\mathbf{1}_P'\boldsymbol{\Sigma}^{-1}\mathbf{1}_p \ ,$$

$$\frac{\partial^2\ell_j}{\partial\mu_x\partial\boldsymbol{\alpha}'} = -\mathbf{1}_p'\boldsymbol{\Sigma}^{-1}\mathbb{I}_{(p)}' \ ,$$

$$\frac{\partial^2\ell_j}{\partial\mu_x\partial\phi_x} = -\frac{c^{-2}}{\phi_x}(c-1)\mathbf{1}_p'\mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{W}_j \ ,$$

$$\frac{\partial^2\ell_j}{\partial\mu_x\partial\boldsymbol{\phi}'} = -c^{-1}\mathbf{W}_j'\boldsymbol{\Sigma}^{-1}\mathrm{D}^{-1}(\boldsymbol{\phi}) \ ,$$

$$\frac{\partial^2\ell_j}{\partial\boldsymbol{\alpha}\partial\boldsymbol{\alpha}'} = -\mathbb{I}_{(p)}\boldsymbol{\Sigma}^{-1}\mathbb{I}_{(p)}' \ ,$$

$$\frac{\partial^2\ell_j}{\partial\boldsymbol{\alpha}\partial\phi_x} = -\frac{c^{-2}}{\phi_x}\mathbb{I}_{(p)}\mathbf{M}\mathbf{W}_j \ ,$$

$$\frac{\partial^2 \ell_j}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\phi}'} = -\mathbb{I}_{(p)}[\boldsymbol{\Sigma}^{-1}\mathrm{D}(\mathbf{W}_j)\mathrm{D}^{-1}(\boldsymbol{\phi}) + c^{-2}\phi_x \mathbf{M}\mathbf{W}_j \mathbf{1}_p' \mathrm{D}^{-2}(\boldsymbol{\phi}) - c^{-1}\phi_x \mathbf{1}_p' \mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{W}_j \mathrm{D}^{-2}(\boldsymbol{\phi})],$$

$$\frac{\partial^2 \ell_j}{\partial \phi_x \partial \phi_x} = \frac{1}{2\phi_x^2}\left(\frac{c-1}{c}\right)^2 - \frac{c^{-3}}{\phi_x}\mathbf{W}_j' \mathbf{M}\mathbf{W}_j \mathbf{1}_p' \mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{1}_p \ ,$$

$$\frac{\partial^2 \ell_j}{\partial \phi_x \partial \boldsymbol{\phi}'} = \frac{c^{-2}}{2}\mathbf{1}_p' \mathrm{D}^{-2}(\boldsymbol{\phi}) + c^{-3}\mathbf{W}_j' \mathbf{M}\mathbf{W}_j \mathbf{1}_p' \mathrm{D}^{-2}(\boldsymbol{\phi}) - c^{-2}\mathbf{W}_j' \mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{1}_p \mathbf{W}_j' \mathrm{D}^{-2}(\boldsymbol{\phi}) \ ,$$

$$\begin{aligned}
\frac{\partial^2 \ell_j}{\partial \boldsymbol{\phi} \partial \boldsymbol{\phi}'} \quad = \quad & \frac{1}{2}\mathrm{D}^{-2}(\boldsymbol{\phi}) - \phi_x c^{-1}\mathrm{D}^{-3}(\boldsymbol{\phi}) + \frac{\phi_x}{2}c^{-2}\mathrm{D}^{-1}(\boldsymbol{\phi})\mathrm{M}\mathrm{D}^{-1}(\boldsymbol{\phi}) - \mathrm{D}^2(\mathbf{W}_j)\mathrm{D}^{-3}(\boldsymbol{\phi}) \\
& + c^{-3}\phi_x \mathbf{W}_j' \mathbf{M}\mathbf{W}_j \mathrm{D}^{-1}(\boldsymbol{\phi})\mathrm{M}\mathrm{D}^{-1}(\boldsymbol{\phi}) - \phi_x c^{-2}\mathrm{D}^{-2}(\boldsymbol{\phi})\mathbf{W}_j \mathbf{W}_j' \mathrm{M}\mathrm{D}^{-1}(\boldsymbol{\phi}) \\
& - c^{-2}\phi_x \mathbf{W}_j' \mathbf{M}\mathbf{W}_j \mathrm{D}^{-3}(\boldsymbol{\phi}) - c^{-2}\phi_x \mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{M}\mathbf{W}_j \mathbf{W}_j' \mathrm{D}^{-2}(\boldsymbol{\phi}) \\
& + c^{-1}\phi_x \mathrm{D}^{-2}(\boldsymbol{\phi})\mathbf{W}_j \mathbf{W}_j' \mathrm{D}^{-2}(\boldsymbol{\phi}) + 2c^{-1}\phi_x \mathbf{1}_p' \mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{W}_j \mathrm{D}(\mathbf{W}_j)\mathrm{D}^{-3}(\boldsymbol{\phi}) \, .
\end{aligned}$$

Portanto, dos resultados acima, temos que a matriz de informação observada é dada por

$$\mathbf{I}_F(\boldsymbol{\theta}) = -\ddot{L} = -\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} = -\sum_{j=1}^{n} \frac{\partial^2 \ell_j}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'},$$

onde $\dfrac{\partial^2 \ell_j(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}$ é como em (2.9).

2.2.2 Matriz de Informação Esperada

Nesta seção vamos obter a matriz de informação esperada do modelo de Grubbs. A obtenção da matriz de informação esperada será importante para testar as hipóteses de interesse apresentadas na Seção 1.1 e para estudar o comportamento assintótico do estimador de máxima verossimilhança de $\boldsymbol{\theta}$ ou de uma função diferenciável de $\boldsymbol{\theta}$.

A matriz de informação esperada de Fisher, que denotamos por $\mathbf{J}_F(\boldsymbol{\theta})$, é definida como sendo

$$\mathbf{J}_F(\boldsymbol{\theta}) = E\left(-\frac{\partial^2 \ell}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}\right) = -\sum_{j=1}^{n} E\left(\frac{\partial^2 \ell_j}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}\right), \qquad (2.10)$$

como as variáveis aleatórias consideradas são identicamente distribuídas

$$\mathbf{J}_F(\boldsymbol{\theta}) = -nE\left(\frac{\partial^2 \ell_j}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}\right).$$

Para propósitos de inferência, a matriz de interesse será a matriz $\mathbf{J}_{F1}(\boldsymbol{\theta})$, chamada de matriz de informação esperada de Fisher por elemento e definida por

$$\mathbf{J}_{F1}(\boldsymbol{\theta}) = -E\left(\frac{\partial^2 \ell_j}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right). \tag{2.11}$$

Os elementos da matriz $-\mathbf{J}_{F1}(\boldsymbol{\theta})$ são obtidos calculando o valor esperado da matriz definida em (2.9). Logo, utilizando os resultados dos Lemas A.1 e A.2 (veja, Apêndice, A), e após as manipulações algébricas, temos que

$$E\left(\frac{\partial^2 \ell_j}{\partial\mu_x\partial\mu_x}\right) = -\mathbf{1}_p'\,\boldsymbol{\Sigma}^{-1}\mathbf{1}_p\ ,$$

$$E\left(\frac{\partial^2 \ell_j}{\partial\mu_x\partial\boldsymbol{\alpha}'}\right) = -\mathbf{1}_p'\boldsymbol{\Sigma}^{-1}\mathbb{I}_{(p)}'\ ,$$

$$E\left(\frac{\partial^2 \ell_j}{\partial\mu_x\partial\phi_x}\right) = 0\ ,$$

$$E\left(\frac{\partial^2 \ell_j}{\partial\mu_x\partial\boldsymbol{\phi}'}\right) = \mathbf{0}_{(1,p)}\ ,$$

$$E\left(\frac{\partial^2 \ell_j}{\partial\boldsymbol{\alpha}\partial\boldsymbol{\alpha}'}\right) = -\mathbb{I}_{(p)}\boldsymbol{\Sigma}^{-1}\mathbb{I}_{(p)}'\ ,$$

$$E\left(\frac{\partial^2 \ell_j}{\partial\boldsymbol{\alpha}\partial\phi_x}\right) = \mathbf{0}_{(q,1)}\ ,$$

$$E\left(\frac{\partial^2 \ell_j}{\partial\boldsymbol{\alpha}\partial\boldsymbol{\phi}'}\right) = \mathbf{0}_{(q,p)}\ ,$$

$$E\left(\frac{\partial^2 \ell_j}{\partial\phi_x\partial\phi_x}\right) = \frac{1}{2\phi_x^2}\left(\frac{c-1}{c}\right)^2 - \frac{c^{-2}}{\phi_x}(c-1)\mathbf{1}_p'\,\mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{1}_p\ ,$$

$$E\left(\frac{\partial^2 \ell_j}{\partial\phi_x\partial\boldsymbol{\phi}'}\right) = \frac{c^{-2}}{2}\mathbf{1}_p'\mathrm{D}^{-2}(\boldsymbol{\phi}) + c^{-2}(c-1)\mathbf{1}_p'\mathrm{D}^{-2}(\boldsymbol{\phi}) - c^{-2}\mathbf{1}_p'\mathrm{D}^{-1}(\boldsymbol{\phi})\boldsymbol{\Sigma}D^{-2}(\boldsymbol{\phi})\ ,$$

$$E\left(\frac{\partial^2 \ell_j}{\partial \boldsymbol{\phi} \partial \boldsymbol{\phi}'}\right) = -\frac{1}{2}\mathrm{D}^{-2}(\boldsymbol{\phi}) - \frac{1}{2}c^{-2}\phi_x \mathrm{D}^{-1}(\boldsymbol{\phi})\mathrm{M}\mathrm{D}^{-1}(\boldsymbol{\phi}) + c^{-1}\phi_x \mathrm{D}^{-1}(\boldsymbol{\phi})\mathrm{M}\mathrm{D}^{-1}(\boldsymbol{\phi})$$
$$-2c^{-2}\phi_x \mathrm{D}^{-2}(\boldsymbol{\phi})\boldsymbol{\Sigma}\mathrm{M}\mathrm{D}^{-1}(\boldsymbol{\phi}) + c^{-1}\boldsymbol{\phi}_x \mathrm{D}^{-2}(\boldsymbol{\phi})\boldsymbol{\Sigma}\mathrm{D}^{-2}(\boldsymbol{\phi}).$$

Dos resultados apresentados anteriormente e de (2.11), provamos o seguinte teorema.

Teorema 2.1. *Sob o modelo estrutural de Grubbs definido em (2.1). A matriz de informação esperada de Fisher por elemento é dada por:*

$$\mathbf{J}_{F1}(\theta) = \begin{pmatrix} J_{\mu_x\mu_x} & J_{\mu_x\alpha} & 0 & 0 \\ J_{\alpha\mu_x} & J_{\alpha\alpha} & 0 & 0 \\ 0 & 0 & J_{\phi_x\phi_x} & J_{\phi_x\phi} \\ 0 & 0 & J_{\phi\phi_x} & J_{\phi\phi} \end{pmatrix},$$

onde

$$J_{\mu_x\mu_x} = \boldsymbol{1}_p'\boldsymbol{\Sigma}^{-1}\boldsymbol{1}_p \ ,$$

$$J_{\mu_x\alpha} = \boldsymbol{1}_p'\boldsymbol{\Sigma}^{-1}\mathbb{I}_{(p)}' = J_{\alpha\mu_x}' \ ,$$

$$J_{\alpha\alpha} = \mathbb{I}_{(p)}\boldsymbol{\Sigma}^{-1}\mathbb{I}_{(p)}' \ ,$$

$$J_{\phi_x\phi_x} = -\frac{1}{2\phi_x^2}\left(\frac{c-1}{c}\right)^2 + \frac{c^{-2}}{\phi_x}(c-1)\boldsymbol{1}_p'D^{-1}(\boldsymbol{\phi})\boldsymbol{1}_p \ ,$$

$$J_{\phi_x\phi} = -\frac{c^{-2}}{2}\boldsymbol{1}_p'D^{-2}(\boldsymbol{\phi}) - c^{-2}(c-1)\boldsymbol{1}_p' D^{-2}(\boldsymbol{\phi}) + c^{-2}\boldsymbol{1}_p'D^{-1}(\boldsymbol{\phi})\boldsymbol{\Sigma}D^{-2}(\boldsymbol{\phi}) = J_{\phi\phi_x}'$$
$$J_{\phi\phi} = \frac{1}{2}D^{-2}(\boldsymbol{\phi}) + \frac{c^{-2}}{2}\phi_x D^{-1}(\boldsymbol{\phi})\boldsymbol{M}D^{-1}(\boldsymbol{\phi}) - c^{-1}\phi_x D^{-1}(\boldsymbol{\phi})\boldsymbol{M}D^{-1}(\boldsymbol{\phi})$$

$$+ 2c^{-2}\phi_x D^{-2}(\boldsymbol{\phi})\boldsymbol{\Sigma}\boldsymbol{M}D^{-1}(\boldsymbol{\phi}) - c^{-1}\phi_x D^{-2}(\boldsymbol{\phi})\boldsymbol{\Sigma}D^{-2}(\boldsymbol{\phi}).$$

Note que, esta matriz é diagonal em bloco, a qual significa que os EMV dos parâmetros (μ_x , $\boldsymbol{\alpha}'$), e os EMV dos parâmetros ($\phi_x, \boldsymbol{\phi}'$) são assintoticamente independentes, como veremos na seção seguinte.

2.2.3 Estimação de Máxima Verossimilhança (EMV)

Encontrar os EMV de $\boldsymbol{\theta}$ por métodos convencionais é complicado, dado que as equações resultantes não têm forma fechada. Embora existam muitas metodologias para encontrar os EMV, aqui mostramos como obter os EMV de $\boldsymbol{\theta}$ usando o algoritmo EM, o qual é fácil de implementar e computacionalmente conveniente, além disso, as estimativas das variâncias são todas não negativas.

Como é conhecido, o algoritmo EM (Dempster et al., 1977) aumenta os dados $\mathbf{Y} = (\mathbf{Y}_1, \ldots, \mathbf{Y}_n)'$ por adicionamento de dados hipotéticos, ou não-observáveis, $\mathbf{x} = (x_1, \ldots, x_n)'$ de tal forma que o EMV de $\boldsymbol{\theta}$ baseado em $\mathbf{Z} = (\mathbf{Y}, \mathbf{x})$ seja fácil de calcular. Dada a estimativa $\boldsymbol{\theta}^{(m)}$, na iteração m, a iteração $m+1$ do algoritmo EM consiste de duas etapas: uma etapa **E** de esperança e uma etapa **M** de maximização. Na etapa **E** calculamos o valor esperado do logaritmo da função de verossimilhança completa $\ell(\boldsymbol{\theta}/\mathbf{Z})$ com respeito à distribuição condicional de $\mathbf{x}$ dado $\mathbf{Y}$ e $\boldsymbol{\theta}^{(m)}$. Na etapa **M** maximizamos a função resultante com respeito a $\boldsymbol{\theta}$, e obtemos assim uma nova estimativa $\boldsymbol{\theta}^{(m+1)}$. Cada iteração do EM incrementa o logaritmo da função de verossimilhança observada $\ell(\boldsymbol{\theta}/\mathbf{Y})$; $\ell(\boldsymbol{\theta}^{(m)}/\mathbf{Y}) \leq \ell(\boldsymbol{\theta}^{(m+1)}/\mathbf{Y})$.

O algoritmo EM é utilizado em problemas de estimação onde a função de verossimilhança é dificultosa de maximizar, mas os dados observados $(\mathbf{y}_j)$ podem ser vistos como uma função de alguma variável não-observável (x_j) sob a qual o cômputo dos EMV podem ser obtidos diretamente. A função de verossimilhança com dados completos é uma família exponencial regular, a qual tipicamente simplifica a implementação do algoritmo EM. Desde que estamos tratando com família exponencial regular (distribuição normal), é suficiente calcular os valores esperados das estatísticas suficientes que seguem da função de log-verossimilhança dos dados completos. Portanto, desde que as estatísticas suficientes dependam somente de $\mathbf{x} = (x_1, \ldots, x_n)'$, então a etapa **E** é implementada estimando x_j e x_j^2. A etapa **M** obtém os próximos valores dos parâmetros desconhecidos maximizando a função de log-verossimilhança dos dados completos com as estatísticas suficientes substituídas pelos valores esperados obtidos na etapa **E**. A seguir, discutimos a implementação do algoritmo EM para encontrar o EMV de $\boldsymbol{\theta}$.

Seja $\mathbf{Z}_j = \begin{pmatrix} x_j \\ \mathbf{Y}_j \end{pmatrix}$, $j = 1, \ldots, n$. Então, de (2.2) temos que

$$\mathbf{Z}_j \overset{\text{iid}}{\sim} N_p(\boldsymbol{\mu}_z, \boldsymbol{\Sigma}_z)\,,\ j = 1, \ldots, n\,, \tag{2.12}$$

onde

$$\boldsymbol{\mu}_z = \begin{pmatrix} \mu_x \\ \boldsymbol{\mu} \end{pmatrix} \;\; e \;\; \boldsymbol{\Sigma}_z = \begin{pmatrix} \phi_x & \mathbf{1}_p'\phi_x \\ \mathbf{1}_p\phi_x & \phi_x\mathbf{1}_p\mathbf{1}_p' + \mathbf{D}(\boldsymbol{\phi}) \end{pmatrix},$$

com $\boldsymbol{\mu}$ como em (2.3).

Logo, segue que o logaritmo da função de verossimilhança completa, denotado por $\ell(\boldsymbol{\theta}/\mathbf{z})$ é dado por

$$\ell(\boldsymbol{\theta}/\mathbf{z}) = cte - \frac{n}{2}\log|\boldsymbol{\Sigma}_z| - \frac{1}{2}\sum_{j=1}^{n}(\mathbf{z}_j - \boldsymbol{\mu}_z)'\boldsymbol{\Sigma}_z^{-1}(\mathbf{z}_j - \boldsymbol{\mu}_z), \tag{2.13}$$

onde

$$|\boldsymbol{\Sigma}_z| = \phi_x \prod_{i=1}^{p}\phi_i \quad \text{e} \quad \boldsymbol{\Sigma}_z^{-1} = \begin{pmatrix} c/\phi_x & -\mathbf{1}_p'\mathrm{D}^{-1}(\boldsymbol{\phi}) \\ -\mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{1}_p & \mathrm{D}^{-1}(\boldsymbol{\phi}) \end{pmatrix}, \tag{2.14}$$

com c como em (2.4). Usando (2.13) e (2.14), temos após manipulações algébricas que

$$\ell(\boldsymbol{\theta}/\mathbf{z}) = cte - \frac{n}{2}\log\phi_x - \frac{n}{2}\sum_{i=1}^{p}\log\phi_i - \frac{1}{2}\sum_{j=1}^{n} f_j\,, \tag{2.15}$$

onde

$$\begin{aligned} f_j \;\; = \;\; & \frac{(x_j - \mu_x)c}{\phi_x} - 2\frac{(x_j - \mu_x)(y_{1j} - \mu_x)}{\phi_1} \\ & -2(x_j - \mu_x)[\mathbf{y}_{2j} - \boldsymbol{\alpha} - \mathbf{1}_q\mu_x]'\mathrm{D}_1^{-1}(\boldsymbol{\phi})\mathbf{1}_q \\ & +\frac{(y_{1j} - \mu_x)^2}{\phi_1} + [\mathbf{y}_{2j} - \boldsymbol{\alpha} - \mathbf{1}_q\mu_x]'\mathrm{D}_1^{-1}(\boldsymbol{\phi})[\mathbf{y}_{2j} - \boldsymbol{\alpha} - \mathbf{1}_q\mu_x]\,, \end{aligned}$$

com $\mathrm{D}_1(\boldsymbol{\phi}) = \mathrm{Diag}(\phi_2, \ldots, \phi_p)$ matriz diagonal $q \times q$, $\mathbf{y}_{2j} = (y_{2j}, \ldots, y_{pj})'$ vetor $q \times 1$ e $q = p - 1$.

Usando propriedades da distribuição normal multivariada (Muirhead, 1982) e a condição em (2.12), podemos implementar a etapa E do algoritmo EM, onde obtemos as estimativas de x_j e x_j^2 dadas por

$$E[x_j/\mathbf{y}_j, \boldsymbol{\theta}] = \widehat{x}_j = \mu_x + \phi_x\mathbf{1}_p'\,[\phi_x\mathbf{1}_p\mathbf{1}_p' + \mathrm{D}(\boldsymbol{\phi})]^{-1}(\mathbf{y}_j - \boldsymbol{\mu}) = \mu_x + \frac{\phi_x}{c}[\mathbf{1}_p'\mathrm{D}^{-1}(\boldsymbol{\phi})(\mathbf{y}_j - \boldsymbol{\mu})], \tag{2.16}$$

e

$$Var[x_j/\mathbf{y}_j, \boldsymbol{\theta}] = \phi_x - \phi_x\mathbf{1}_p'(\phi_x\mathbf{1}_p\mathbf{1}_p' + \mathrm{D}(\boldsymbol{\phi}))^{-1}\phi_x\mathbf{1}_p = \frac{\phi_x}{c}\,,$$

portanto

$$\widehat{x_j^2} = \widehat{x}_j^2 + \frac{\phi_x}{c} \,. \tag{2.17}$$

A etapa M do algoritmo, é implementada maximizando a função de log-verossimilhança completa dada em (2.15), com as estatísticas suficientes substituídas pelos valores esperados obtidas no passo E. Derivando $\ell(\boldsymbol{\theta}/\mathbf{z})$ com relação a $\boldsymbol{\theta}$ e após algumas manipulações algébricas, obtemos as seguintes estimativas

$$\widehat{\mu}_x = \frac{1}{n}\sum_{j=1}^{n} x_j = \bar{x} \,,$$

$$\widehat{\alpha}_i = \bar{y}_{i.} - \bar{x} \,,$$

$$\widehat{\phi}_x = \frac{1}{n}\sum_{j=1}^{n}(x_j - \bar{x})^2 \,,$$

$$\widehat{\phi}_1 = \frac{1}{n}\sum_{j=1}^{n}(y_{1j} - x_j)^2 \,,$$

e

$$\widehat{\phi}_i = \frac{1}{n}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{i.} - x_j + \bar{x})^2 \,, \tag{2.18}$$

onde $\bar{y}_{i.} = \dfrac{1}{n}\sum_{j=1}^{n} y_{ij}$ e $i = 2, \ldots, p$.

Portanto, se $\boldsymbol{\theta}^{(m)}$ maximiza (2.13) na m-ésima iteração, então na iteração $m+1$ o algoritmo EM procede como segue:

etapa E: Calcule $\widehat{x}_j(\boldsymbol{\theta}^{(m)})$ e $\widehat{x_j^2}(\boldsymbol{\theta}^{(m)})$, j=1,...,n , dados por (2.16) e (2.17) respectivamente, com $\boldsymbol{\theta} = \boldsymbol{\theta}^{(m)}$.

etapa M: Calcule $(\boldsymbol{\theta} = \boldsymbol{\theta}^{(m+1)})$ a partir de (2.18).

Note que as estimativas das variâncias são sempre não-negativas, o que não acontece necessariamente com os estimadores propostos por Grubbs (1973), mostrados na Seção 2.5, e a implementação do algoritmo EM é bastante simples e computacionalmente conveniente.

Como estimativas iniciais $\boldsymbol{\theta} = \boldsymbol{\theta}^{(0)}$ podemos utilizar os estimadores de momentos propostos por Grubbs (veja, Seção, 2.5) e como critério de convergência podemos usar

$$\|\boldsymbol{\theta}^{(m)} - \boldsymbol{\theta}^{(m-1)}\| < \varepsilon \,,$$

onde, $\|\mathbf{a}\|$ indica norma do vetor $\mathbf{a}$ e $\varepsilon > 0$.

2.3 Estimação de Máxima Verossimilhança Restrita

Nesta seção apresentamos expressões para os EMV quando alguma informação adicional é incluída no modelo de Grubbs definido em (2.1-2.2). Como veremos, em muitos casos os EMV são obtidos em forma exata e nenhum procedimento iterativo é requerido. Aqui consideramos cinco casos:

1. Os vícios são todos iguais a zero e as variâncias dos instrumentos são iguais, isto é, $\alpha_2 = \alpha_3 = \ldots = \alpha_p = 0$ e $\phi_1 = \phi_2 = \ldots = \phi_p$;

2. As variâncias dos instrumentos são iguais, isto é, $\phi_1 = \phi_2 = \ldots = \phi_p$;

3. Os vícios dos instrumentos são todos iguais a zero, isto é, $\alpha_2 = \alpha_3 = \ldots = \alpha_p = 0$;

4. O instrumento de referência mede sem erro a característica de interesse, isto é, $\phi_1 = 0$ e

5. Um dos instrumentos mede sem erro a característica de interesse, isto é, $\phi_i = 0\,,\ i = 2, \ldots, p$.

Para os casos 1, 2, 4, 5 os EMV são obtidos em forma exata e apresentamos expressões explícitas para os EMV. No caso 3 as estimativas dos parâmetros do modelo são calculadas via algoritmo EM. Para os casos 4 e 5, também, damos a matriz de covariância assintótica a qual pode ser usada para testar as hipóteses de interesse. Os casos 1,2 e 3 serão de grande utilidade na seção de testes de hipóteses, quando necessitaremos dos EMV sob as hipóteses nulas H_{01}, H_{02} e H_{03} definidas na Seção 1.1.

2.3.1 O Caso $\alpha_2 = \alpha_3 = \ldots = \alpha_p = 0$ e $\phi_1 = \phi_2 = \ldots = \phi_p$

Nesta seção vamos considerar a possibilidade de que os instrumentos meçam a característica de interesse sem vício e com a mesma variabilidade, as quais consideramos como sendo iguais a ϕ, isto é, $\alpha_2 = \alpha_3 = \ldots = \alpha_p = 0$ e $\phi_1 = \phi_2 = \ldots = \phi_p = \phi$. Logo, segue que

$$\mathbf{Y}_j \overset{\text{iid}}{\sim} N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})\,, \tag{2.19}$$

onde $\boldsymbol{\mu} = \mu_x \mathbf{1}_p$ *e* $\boldsymbol{\Sigma} = \phi_x \mathbf{1}_p \mathbf{1}_p' + \phi \mathbf{I}_p$, com $\mathbf{I}_p$ matriz identidade de ordem p e a função de log-verossimilhança fica dada por

$$\ell(\boldsymbol{\theta}) = \sum_{j=1}^{n} \ell_j(\boldsymbol{\theta})\,, \tag{2.20}$$

onde $\ell_j(\boldsymbol{\theta}) = cte - \frac{1}{2}\log|\boldsymbol{\Sigma}| - \frac{1}{2}\|\mathbf{T}_j\|^2$, com $\|\mathbf{T}_j\|^2 = \mathbf{W}_j'\boldsymbol{\Sigma}^{-1}\mathbf{W}_j$, $\mathbf{W}_j = \mathbf{y}_j - \boldsymbol{\mu}$, $|\boldsymbol{\Sigma}| = c(\phi)^p$, $\boldsymbol{\Sigma}^{-1} = \frac{1}{\phi}\mathbf{I}_p - c^{-1}\mathbf{M}$, $c = 1 + p\frac{\phi_x}{\phi}$ e $\mathbf{M} = \frac{\phi_x}{\phi^2}\mathbf{1}_p\mathbf{1}_p'$. Neste caso

$$\boldsymbol{\theta} = (\mu_x, \phi_x, \phi)'\,.$$

Maximizando a função de log-verossimilhança em (2.20), encontramos que os EMV de $\boldsymbol{\theta}$ para este modelo são dados por

$$\widehat{\mu}_x = \bar{y}_{..}\ ,$$
$$\widehat{\phi} = \frac{1}{n(p-1)}\sum_{j=1}^{n}\sum_{i=1}^{p}(y_{ij} - \bar{y}_{.j})^2\,, \tag{2.21}$$

e

$$\widehat{\phi}_x = \frac{1}{n}\sum_{j=1}^{n}(\bar{y}_{.j} - \bar{y}_{..})^2 - \frac{\widehat{\phi}}{p} = \frac{1}{pn}\sum_{j=1}^{n}\sum_{i=1}^{p}(y_{ij} - \bar{y}_{..}) - \widehat{\phi}\,,$$

onde $\bar{y}_{..} = \frac{1}{np}\sum_{i=1}^{p}\sum_{j=1}^{n} y_{ij}$.

Note que ϕ_x pode ser estimado negativamente, então procedemos como em Jaech (1985), fazendo $\widehat{\phi}_x = 0$ e conseqüentemente o EMV de ϕ será dado por

$$\widehat{\phi} = \frac{1}{pn}\sum_{j=1}^{n}\sum_{i=1}^{p}(y_{ij} - \bar{y}_{..})^2\,. \tag{2.22}$$

2.3.2 O Caso $\phi_1 = \phi_2 = \ldots = \phi_p$

Sem perda de generalidade, vamos considerar que as variâncias são iguais a ϕ, isto é

$$\phi_1 = \phi_2 = \ldots = \phi_p = \phi\,.$$

Então, a distribuição dos $\mathbf{Y}_j$ é dada por

$$\mathbf{Y}_j \overset{\text{iid}}{\sim} N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})\,, \tag{2.23}$$

onde $\boldsymbol{\Sigma} = \phi_x \mathbf{1}_p \mathbf{1}_p' + \phi \mathbf{I}_p$ $\;e\;$ $\boldsymbol{\mu} = \mathbf{a} + \mathbf{1}_p \mu_x$, com $\mathbf{I}_p$ matriz identidade de ordem p e $\mathbf{a}$ definido como em (1.7). Neste caso

$$\boldsymbol{\theta} = (\mu_x,\, \boldsymbol{\alpha}',\, \phi_x,\, \phi).$$

A função de densidade de probabilidade do modelo é dada por

$$f(\mathbf{y}_j; \boldsymbol{\theta}) = (2\pi)^{-p/2} |\boldsymbol{\Sigma}|^{-1/2} \exp\{-\frac{1}{2}(\mathbf{y}_j - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{y}_j - \boldsymbol{\mu})\}\,,$$

onde $\boldsymbol{\Sigma}^{-1} = \frac{1}{\phi}\mathbf{I}_p - c^{-1}\mathbf{M}$ $\;e\;$ $|\boldsymbol{\Sigma}| = c(\phi)^p$, com $c = 1 + p\frac{\phi_x}{\phi}$ $\;e\;$ $M = \frac{\phi_x}{\phi^2}\mathbf{1}_p\mathbf{1}_p'$. Logo, a função de log-verossimilhança é dada por

$$\ell(\boldsymbol{\theta}) = \sum_{j=1}^{n} \ell_j(\boldsymbol{\theta}) = \sum_{j=1}^{n} \log f(\mathbf{y}_j; \boldsymbol{\theta})\,, \tag{2.24}$$

onde

$$\ell_j(\boldsymbol{\theta}) = cte - \frac{1}{2}\log|\boldsymbol{\Sigma}| - \frac{1}{2}\|\mathbf{T}_j\|^2\,,$$

com $\|\mathbf{T}_j\|^2 = \mathbf{W}_j' \boldsymbol{\Sigma}^{-1} \mathbf{W}_j$ e $\mathbf{W}_j = \mathbf{y}_j - \boldsymbol{\mu}$.

De (2.24), segue após algumas manipulações algébricas que o EMV dos parâmetros do modelo tem uma forma fechada, dado por

$$\widehat{\mu}_x = \bar{y}_{1.}\;,$$

$$\widehat{\alpha}_i = \bar{y}_{i.} - \widehat{\mu}_x\,,\;\; i = 2, \ldots, p\;,$$

$$\widehat{\phi} = \frac{1}{n(p-1)} \sum_{j=1}^{n} \sum_{i=1}^{p} \left[(y_{ij} - \bar{y}_{i.})^2 - (\bar{y}_{.j} - \bar{y}_{..})^2\right]\;,$$

$$\widehat{\phi}_x = \frac{1}{n} \sum_{j=1}^{n} (\bar{y}_{.j} - \bar{y}_{..})^2 - \frac{\widehat{\phi}}{p}\;. \tag{2.25}$$

Note também que, ϕ_x pode ser estimado negativamente, assim consideramos $\widehat{\phi}_x = 0$ e conseqüentemente o EMV de ϕ será dado nesse caso por

$$\widehat{\phi} = \frac{1}{pn} \sum_{j=1}^{n} \sum_{i=1}^{p} (y_{ij} - \bar{y}_{i.})^2 \,. \tag{2.26}$$

2.3.3 O Caso $\alpha_2 = \alpha_3 = \ldots = \alpha_p = 0$

Considerando a possibilidade de que os diferentes instrumentos medem a característica de interesse sem vício, isto é, $\alpha_2 = \alpha_3 = \ldots = \alpha_p = 0$, segue que

$$\mathbf{Y}_j \overset{\text{iid}}{\sim} N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \,, \tag{2.27}$$

onde

$$\boldsymbol{\mu} = \mu_x \mathbf{1}_p \;\; e \;\; \boldsymbol{\Sigma} = \phi_x \mathbf{1}_p \mathbf{1}_p' + \mathrm{D}(\boldsymbol{\phi}) \,.$$

Neste caso

$$\boldsymbol{\theta} = (\mu_x, \phi_x, \boldsymbol{\phi}')' = (\mu_x, \phi_x, \phi_1, \ldots, \phi_p)' \,.$$

É fácil verificar que as equações de verossimilhança não têm uma solução fechada, isto é, não conduzem a nenhuma expressão isolada dos parâmetros do modelo em função dos valores amostrais. Para encontrar os EMV, vamos implementar o algoritmo EM sob este modelo , segundo os passos da estimação feita na Seção 2.2.3 .

Seja $\mathbf{Z}_j = (x_j \,, \mathbf{Y}_j')'$. Então o logaritmo da função de verossimilhança completa é dada por

$$\ell(\boldsymbol{\theta}/\mathbf{z}) = cte - \frac{n}{2}\log|\boldsymbol{\Sigma}_z| - \frac{1}{2}\sum_{j=1}^{n}(\mathbf{z}_j - \boldsymbol{\mu}_z)'\boldsymbol{\Sigma}_z^{-1}(\mathbf{z}_j - \boldsymbol{\mu}_z), \tag{2.28}$$

onde

$$|\boldsymbol{\Sigma}_z| = \phi_x \prod_{i=1}^{p} \phi_i \quad e \quad \boldsymbol{\Sigma}_z^{-1} = \begin{pmatrix} c/\phi_x & -\mathbf{1}_p'\mathrm{D}^{-1}(\boldsymbol{\phi}) \\ -\mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{1}_p & \mathrm{D}^{-1}(\boldsymbol{\phi}) \end{pmatrix} ,$$

com c como em (2.4).

Da relação acima, $\ell(\boldsymbol{\theta}/\mathbf{z})$ pode ser escrita como

$$\ell(\boldsymbol{\theta}/\mathbf{z}) = cte - \frac{n}{2}\log\phi_x - \frac{n}{2}\sum_{i=1}^{p}\log\phi_i - \frac{1}{2}\sum_{j=1}^{n} f_j \,,$$

onde

$$f_j = \frac{c}{\phi_x}(x_j - u_x)^2 - 2(x_j - \mu_x)\mathbf{1}_p' \mathrm{D}^{-1}(\boldsymbol{\phi})(\mathbf{y}_j - \mathbf{1}_p\mu_x) + (\mathbf{y}_j - \mathbf{1}_p\mu_x)' \mathrm{D}^{-1}(\boldsymbol{\phi})(\mathbf{y}_j - \mathbf{1}_p\mu_x)\,.$$

De uma maneira similar à discutida na Seção 2.2.3, temos que o algoritmo EM procede como segue

Etapa E: Esta etapa é implementada calculando

$$\widehat{x}_j = E(x_j/\mathbf{y}_j, \boldsymbol{\theta}) = \mu_x + \frac{\phi_x}{c}\mathbf{1}_p' \mathrm{D}^{-1}(\boldsymbol{\phi})(\mathbf{y}_j - \boldsymbol{\mu})\,,$$

e

$$\widehat{x_j^2} = Var(x_j/\mathbf{y}_j, \boldsymbol{\theta}) + \{E(x_j/\mathbf{y}_j, \boldsymbol{\theta})\}^2 = \frac{\phi_x}{c} + \widehat{x}_j^2\,.$$

Etapa M: Nesta etapa maximizamos a função $\ell(\boldsymbol{\theta}/\mathbf{z})$, e obtemos as seguintes soluções

$$\widehat{\mu}_x = \frac{1}{n}\sum_{j=1}^{n} x_j = \bar{x}\,,$$

$$\widehat{\phi}_x = \frac{1}{n}\sum_{j=1}^{n}(x_j - \bar{x})^2\,,$$

e

$$\widehat{\phi}_i = \frac{1}{n}\sum_{j=1}^{n}(y_{ij} - x_j)^2\ , \quad i = 1, \ldots, p\,.$$

Note que, a implementação deste algoritmo é simples e as estimativas das variâncias $(\phi_x, \boldsymbol{\phi})$ são sempre não-negativas. Como estimativas iniciais podemos usar os estimadores de momentos propostos na Seção 2.5, com exceção da estimativa de $\boldsymbol{\alpha}$, que como sabemos é considerada como sendo zero.

Segundo Jaech (1985), quando uma das variâncias é estimada como sendo zero (perto ou menor que zero), a conclusão pode ser que o correspondente instrumento meça a característica de interesse sem erro ou em forma precisa. No caso onde um dos instrumentos meça precisamente a característica de interesse, a variância dos erros de medida correspondente a tal instrumento, pode ser considerada zero. No que se segue desta seção consideraremos mais dois casos $\phi_1 = 0$ e $\phi_p = 0$, como veremos, os EMV são obtidos em forma exata e nenhum procedimento iterativo é requerido. Também é obtida a matriz de covariância assintótica dos EMV.

2.3.4 O Caso $\phi_1 = 0$

Existem situações, especialmente na área da medicina, onde é possível considerar, numa primeira aproximação, que o instrumento de referência mede a característica de interesse sem erro de medição, esta suposição se for razoável, é atrativa pois simplifica a obtenção dos EMV. No caso p=2, a restrição $\phi_1 = 0$ é usada para identificar o modelo estrutural, veja por exemplo, Kaaks et al.(1994).

Considerando $\phi_1 = 0$, segue que $\mathbf{Y}_j$ é normal multivariada com média $\boldsymbol{\mu}$ dada como em (2.3) e matriz de covariância dada por

$$\boldsymbol{\Sigma} = \begin{pmatrix} \phi_x & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix},$$

onde $\boldsymbol{\Sigma}_{12} = \phi_x \mathbf{1}_q' = \boldsymbol{\Sigma}_{21}'$ e $\boldsymbol{\Sigma}_{22} = \phi_x \mathbf{1}_q \mathbf{1}_q' + D(\boldsymbol{\psi})$, com $\mathrm{D}(\boldsymbol{\psi}) = \mathrm{Diag}(\phi_2, ..., \phi_p)$ matriz diagonal $q \times q$, $\boldsymbol{\psi} = (\phi_2, \ldots, \phi_p)'$ vetor de dimensão $q = p - 1$.

Utilizando propriedades de matrizes (veja, Graybill, 1983), segue que

$$|\boldsymbol{\Sigma}| = \phi_x |\mathrm{D}(\boldsymbol{\psi})|,$$

e

$$\boldsymbol{\Sigma}^{-1} = \begin{pmatrix} c_1/\phi_x & -\mathbf{1}_q'\mathrm{D}^{-1}(\boldsymbol{\psi}) \\ -\mathrm{D}^{-1}(\boldsymbol{\psi})\mathbf{1}_q & \mathrm{D}^{-1}(\boldsymbol{\psi}) \end{pmatrix},$$

onde $c_1 = 1 + \phi_x \mathbf{1}_q' D^{-1}(\boldsymbol{\psi}) \mathbf{1}_q$. Agora, o vetor de parâmetros do modelo que é de dimensão $2p$ é dado por $\boldsymbol{\theta} = (\mu_x, \alpha_2, ..., \alpha_p, \phi_x, \phi_2, ..., \phi_p)' = (\mu_x, \boldsymbol{\alpha}', \phi_x, \boldsymbol{\psi}')'$ e a função de log-verossimilhança pode ser escrita como

$$\ell(\boldsymbol{\theta}) = \sum_{j=1}^{n} \ell_j(\boldsymbol{\theta}) = cte - \frac{n}{2} \log|\boldsymbol{\Sigma}| - \frac{1}{2} \sum_{j=1}^{n} \|\mathbf{T}_j\|^2, \tag{2.29}$$

onde

$$\begin{aligned} \|\mathbf{T}_j\|^2 &= (\mathbf{y}_j - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{y}_j - \boldsymbol{\mu}) \\ &= \frac{W_{1j}^2 c_1}{\phi_x} - 2W_{1j} \mathbf{W}_{2j}' \mathrm{D}^{-1}(\boldsymbol{\psi}) \mathbf{1}_q + \mathbf{W}_{2j}' \mathrm{D}^{-1}(\boldsymbol{\psi}) \mathbf{W}_{2j}, \end{aligned} \tag{2.30}$$

com $W_{1j} = y_{1j} - \mu_x$, $\mathbf{W}_{2j} = (y_{2j} - \alpha_2 - \mu_x, ..., y_{pj} - \alpha_p - \mu_x)'$ vetor de dimensão $q \times 1$, $q = p - 1$ e $j = 1, ..., n$. Considerando a função de log-verossimilhança em

(2.29) e após manipulações algébricas, encontramos que os estimadores de máxima verossimilhança são dados em forma exata por

$$\widehat{\mu}_x = \bar{y}_{1.} \ ,$$

$$\widehat{\alpha}_i = \bar{y}_{i.} - \bar{y}_{1.} \ ,$$

$$\widehat{\phi}_i = S_{11} + S_{ii} - 2S_{1i}, \ \ i = 2, \ldots, p \ ,$$

e

$$\widehat{\phi}_x = S_{11} \ , \tag{2.31}$$

onde $\bar{y}_{i.} = \dfrac{1}{n}\displaystyle\sum_{j=1}^{n} y_{ij}$, $i = 1, 2, ..., p$ e $S_{kl} = \dfrac{1}{n}\displaystyle\sum_{j=1}^{n}(y_{kj} - \bar{y}_k)(y_{lj} - \bar{y}_l)$, $k, l = 1, ..., p$. Além disso, é fácil verificar que os EMV das variâncias são sempre não-negativas.

Como em (2.6), o vetor de derivadas parciais de primeira ordem, $\dfrac{\partial \ell_j}{\partial \boldsymbol{\theta}}$, é dado por

$$\frac{\partial \ell_j}{\partial \boldsymbol{\theta}} = \left(\begin{array}{c} \dfrac{W_{1j}}{\phi_x} \\ \mathrm{D}^{-1}(\boldsymbol{\psi})\mathbf{W}_{2j} - W_{1j}\mathrm{D}^{-1}(\boldsymbol{\psi})\mathbf{1}_q \\ \dfrac{W_{1j}^2}{2\phi_x^2} - \dfrac{1}{2\phi_x} \\ \dfrac{1}{2}W_{1j}^2\mathrm{D}^{-2}(\boldsymbol{\psi})\mathbf{1}_q - W_{1j}\mathrm{D}^{-2}(\boldsymbol{\psi})\mathbf{W}_{2j} + \\ +\dfrac{1}{2}\mathrm{D}(\mathbf{W}_{2j})\mathbf{D}^{-2}(\boldsymbol{\psi})\mathbf{W}_{2j} - \dfrac{1}{2}D^{-1}(\boldsymbol{\psi})\mathbf{1}_q \end{array} \right) , \tag{2.32}$$

e a matriz de derivadas parciais de segunda ordem, como foi definido em (2.9), é da forma

$$\frac{\partial^2 \ell_j}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'} = \left(\begin{array}{cccc} \dfrac{\partial^2 \ell_j}{\partial \mu_x \partial \mu_x'} & \dfrac{\partial^2 \ell_j}{\partial \mu_x \partial \boldsymbol{\alpha}'} & \dfrac{\partial^2 \ell_j}{\partial \mu_x \partial \phi_x} & \dfrac{\partial^2 \ell_j}{\partial \mu_x \partial \boldsymbol{\psi}'} \\ & \dfrac{\partial^2 \ell_j}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\alpha}'} & \dfrac{\partial^2 \ell_j}{\partial \boldsymbol{\alpha} \partial \phi_x} & \dfrac{\partial^2 \ell_j}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\psi}'} \\ & & \dfrac{\partial^2 \ell_j}{\partial \phi_x \partial \phi_x} & \dfrac{\partial^2 \ell_j}{\partial \phi_x \partial \boldsymbol{\psi}'} \\ \mathit{simetria} & & & \dfrac{\partial^2 \ell_j}{\partial \boldsymbol{\psi} \partial \boldsymbol{\psi}'} \end{array} \right) , \tag{2.33}$$

onde os elementos da matriz são dados por

$$\frac{\partial^2 \ell_j}{\partial \mu_x \partial \mu_x} = -\frac{1}{\phi_x} \ ,$$

$$\frac{\partial^2 \ell_j}{\partial \mu_x \partial \boldsymbol{\alpha}'} = \mathbf{0}_{1\times q} \ ,$$

$$\frac{\partial^2 \ell_j}{\partial \mu_x \partial \phi_x} = -\frac{W_{1j}}{\phi_x^2} \ ,$$

$$\frac{\partial^2 \ell_j}{\partial \mu_x \partial \boldsymbol{\psi}'} = \mathbf{0}_{1\times q} \ ,$$

$$\frac{\partial^2 \ell_j}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\alpha}'} = -\mathrm{D}^{-1}(\boldsymbol{\psi}) \ ,$$

$$\frac{\partial^2 \ell_j}{\partial \boldsymbol{\alpha} \partial \phi_x} = \mathbf{0}_{q\times 1} \ ,$$

$$\frac{\partial^2 \ell_j}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\psi}'} = W_{1j}\mathrm{D}^{-2}(\boldsymbol{\psi}) - \mathrm{D}^{-2}(\boldsymbol{\psi})\mathrm{D}(\mathbf{W}_{2j}) \ ,$$

$$\frac{\partial^2 \ell_j}{\partial \phi_x \partial \phi_x} = \frac{1}{2\phi_x^2} - \frac{W_{1j}^2}{\phi_x^3} \ ,$$

$$\frac{\partial^2 \ell_j}{\partial \phi_x \partial \boldsymbol{\psi}'} = \mathbf{0}_{1\times q} \ ,$$

$$\frac{\partial^2 \ell_j}{\partial \boldsymbol{\psi} \partial \boldsymbol{\psi}'} = \frac{1}{2}\mathrm{D}^{-2}(\boldsymbol{\psi}) - W_{1j}^2\mathrm{D}^{-3}(\boldsymbol{\psi}) + 2W_{1j}\mathrm{D}(\mathbf{W}_{2j})\mathrm{D}^{-3}(\boldsymbol{\psi}) - \mathrm{D}^2(\mathbf{W}_{2j})\mathrm{D}^{-3}(\boldsymbol{\psi}) \ .$$

Logo, a matriz de informação observada de $\boldsymbol{\theta} = (\mu_x,\, \alpha_2, ...\alpha_p,\, \phi_x,\, \phi_2, ..., \phi_p)'$, é obtida ao considerar

$$\mathbf{I}_F(\boldsymbol{\theta}) = -\ddot{\mathbf{L}} = -\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} = -\sum_{j=1}^{n} \frac{\partial^2 \ell_j}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}.$$

Dos resultados anteriores, tem-se o seguinte Teorema.

Teorema 2.2. *Sob o modelo de Grubbs definido em (2.1) e supondo que $\phi_1 = 0$, a matriz de informação esperada de $\boldsymbol{\theta}$, é dada por*

$$\mathbf{J}_{F1}(\boldsymbol{\theta}) = \begin{pmatrix} J_{\mu_x\mu_x} & 0 & 0 & 0 \\ 0 & J_{\alpha\alpha} & 0 & 0 \\ 0 & 0 & J_{\phi_x\phi_x} & 0 \\ 0 & 0 & 0 & J_{\psi\psi} \end{pmatrix},$$

onde

$$J_{\mu_x\mu_x} = \frac{1}{\phi_x} \quad ,$$

$$J_{\alpha\alpha} \quad = D^{-1}(\boldsymbol{\psi}) \quad ,$$

$$J_{\phi_x\phi_x} = \frac{1}{2\phi_x^2} \quad ,$$

$$J_{\psi\psi} \quad = \frac{1}{2} D^{-2}(\boldsymbol{\psi}) \ .$$

Demonstração. A prova segue da relação em (2.11) e dos resultados dos Lemas A.1 e A.2 dados no Apêndice A. □

O fato importante de ser salientado do Teorema anterior é que os EMV dos parâmetros do modelo considerando $\phi_1 = 0$, são assintoticamente independentes. Assim, usando a estatística de Wald, podemos testar hipóteses similares a H_{01}, H_{02} e H_{03} consideradas na Seção 1.1.

2.3.5 O Caso $\phi_p = 0$

Considerando agora a possibilidade que uma das variâncias $\phi_2, ..., \phi_p$ é estimada como zero, sem perda de generalidade vamos considerar $\phi_p = 0$. Assim $\mathbf{Y}_j = (y_{1j}, ..., y_{pj})'$ tem distribuição normal multivariada com vetor de médias $\boldsymbol{\mu}$ dada como em (2.3) e matriz de covariâncias dada por

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \phi_x \end{pmatrix} , \tag{2.34}$$

onde $\boldsymbol{\Sigma}_{11} = \phi_x \mathbf{1}_q \mathbf{1}_q' + \mathrm{D}(\boldsymbol{\tau})$ e $\boldsymbol{\Sigma}_{12} = \phi_x \mathbf{1}_q = \boldsymbol{\Sigma}_{21}'$, com $\mathrm{D}(\boldsymbol{\tau}) = \mathrm{Diag}(\phi_1, ..., \phi_q)$ matriz diagonal $q \times q$, $\boldsymbol{\tau} = (\phi_1, ..., \phi_q)'$ vetor $q \times 1$ e $q = p - 1$. Utilizando propriedades de matrizes (ver, Graybill, 1983), segue que

$$|\boldsymbol{\Sigma}| = \phi_x |\mathrm{D}(\boldsymbol{\tau})| = \phi_x \prod_{i=1}^{q} \phi_i \, , \tag{2.35}$$

e

$$\boldsymbol{\Sigma}^{-1} = \begin{pmatrix} \mathrm{D}^{-1}(\boldsymbol{\tau}) & -\mathrm{D}^{-1}(\boldsymbol{\tau})\mathbf{1}_q \\ -\mathbf{1}_q'\mathbf{D}^{-1}(\boldsymbol{\tau}) & c_2/\phi_x \end{pmatrix} , \tag{2.36}$$

onde $c_2 = 1 + \phi_x \mathbf{1}_q' \mathrm{D}^{-1}(\boldsymbol{\tau}) \mathbf{1}_q$.

Neste caso, o vetor de parâmetros do modelo que tem dimensão $2p$ é dado por $\boldsymbol{\theta} = (\mu_x,\ \alpha_2, \ldots, \alpha_p,\ \phi_x,\ \phi_1, \ldots, \phi_q)' = (\mu_x,\ \boldsymbol{\alpha}',\ \phi_x,\ \boldsymbol{\tau}')'$ e a função de log-verossimilhança pode ser escrita como

$$\ell(\boldsymbol{\theta}) = \sum_{j=1}^{n} \ell_j(\boldsymbol{\theta}) = \ cte - \frac{n}{2}\log|\boldsymbol{\Sigma}| - \frac{1}{2}\sum_{j=1}^{n} \|\mathbf{T}_j\|^2\,, \tag{2.37}$$

onde

$$\|\mathbf{T}_j\|^2 = \frac{W_{1j}^2 c_2}{\phi_x} - 2W_{1j}\mathbf{W}_{2j}'\mathrm{D}^{-1}(\boldsymbol{\tau})\mathbf{1}_q + \mathbf{W}_{2j}'\mathrm{D}^{-1}(\boldsymbol{\psi})\mathbf{W}_{2j}\,,$$

com $W_{1j} = y_{pj} - \alpha_p - \mu_x$ e $\mathbf{W}_{2j} = (y_{1j} - \mu_x, y_{2j} - \alpha_2 - \mu_x, ..., y_{qj} - \alpha_q - \mu_x)'$ vetor de dimensão $q \times 1$, $j = 1, ..., n$.

Para efeitos de estimação vamos considerar o vetor $\boldsymbol{\alpha}$ como sendo $\boldsymbol{\alpha} = (\alpha_2, \ldots, \alpha_p)' = (\boldsymbol{\alpha}_*', \alpha_p)'$, com $\boldsymbol{\alpha}^* = (\alpha_2, ..., \alpha_{p-1})'$ e $\boldsymbol{\theta}$ pode ser representado por $(\mu_x, \boldsymbol{\alpha}_*', \alpha_p, \phi_x, \boldsymbol{\tau}')'$. Neste caso o EMV de $\boldsymbol{\theta}$ tem forma fechada e é dado por

$$\widehat{\mu}_x = \bar{y}_{1.}\ ,$$

$$\widehat{\alpha}_i = \bar{y}_{i.} - \bar{y}_{1.}\ ,\ i = 2, \ldots, p\,,$$

$$\widehat{\phi}_i = S_{pp} + S_{ii} - 2S_{pi}\ ,\ i = 1, \ldots, q\,,$$

e

$$\widehat{\phi}_x = S_{pp}\ \ , \tag{2.38}$$

onde $\overline{y}_{i.} = \dfrac{1}{n}\displaystyle\sum_{j=1}^{n} y_{ij}\ ,\ \ i = 1, 2, ..., p$ e $S_{kl} = \dfrac{1}{n}\displaystyle\sum_{j=1}^{n}(y_{kj} - \bar{y}_k)(y_{lj} - \bar{y}_l)\ ,\ k, l = 1, ..., p.$ Observe que os EMV das variâncias são não-negativos.

No que se segue desta seção vamos estudar o comportamento assintótico do EMV $\widehat{\boldsymbol{\theta}}$ de $\boldsymbol{\theta}$. Derivando a função da log-verossimilhança $\ell(\boldsymbol{\theta})$ definido em (2.37) temos que o vetor de derivadas parciais de primeira ordem, é dado por

$$\frac{\partial \ell_j}{\partial \boldsymbol{\theta}} = \begin{pmatrix} \dfrac{\partial \ell_j}{\partial \mu_x} \\ \dfrac{\partial \ell_j}{\partial \boldsymbol{\alpha}_*} \\ \dfrac{\partial \ell_j}{\partial \alpha_p} \\ \dfrac{\partial \ell_j}{\partial \phi_x} \\ \dfrac{\partial \ell_j}{\partial \boldsymbol{\tau}} \end{pmatrix} = \begin{pmatrix} \dfrac{W_{1j}}{\phi_x} \\ -W_{1j}\mathbb{I}_{(q)}\mathrm{D}^{-1}(\boldsymbol{\tau})\mathbf{1}_q + \mathbb{I}_{(q)}\mathrm{D}^{-1}(\boldsymbol{\tau})\mathbf{W}_{2j} \\ \dfrac{c_2 W_{1j}}{\phi_x} - \mathbf{W}'_{2j}\mathbf{D}^{-1}(\boldsymbol{\tau})\mathbf{1}_q \\ -\dfrac{1}{2\phi_x} + \dfrac{W_{1j}^2}{2\phi_x^2} \\ -\dfrac{1}{2}\mathrm{D}^{-1}(\boldsymbol{\tau})\mathbf{1}_q + \dfrac{1}{2}W_{1j}^2 D^{-2}(\boldsymbol{\tau})\mathbf{1}_q - W_{1j}\mathrm{D}^{-2}(\boldsymbol{\tau})\mathbf{W}_{2j} + \\ +\dfrac{1}{2}\mathrm{D}(\mathbf{W}_{2j})\mathrm{D}^{-2}(\boldsymbol{\tau})\mathbf{W}_{2j} \end{pmatrix}.$$

A matriz de derivadas parciais de segunda ordem $\dfrac{\partial^2 \ell_j}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}}$ é da forma

$$\begin{pmatrix} \dfrac{\partial^2 \ell_j}{\partial \mu_x \partial \mu_x} & \dfrac{\partial^2 \ell_j}{\partial \mu_x \partial \boldsymbol{\alpha}'_*} & \dfrac{\partial^2 \ell_j}{\partial \mu_x \partial \alpha_p} & \dfrac{\partial^2 \ell_j}{\partial \mu_x \partial \phi_x} & \dfrac{\partial^2 \ell_j}{\partial \mu_x \partial \boldsymbol{\tau}'} \\ & \dfrac{\partial^2 \ell_j}{\partial \boldsymbol{\alpha}_* \partial \boldsymbol{\alpha}'_*} & \dfrac{\partial^2 \ell_j}{\partial \boldsymbol{\alpha}_* \partial \alpha_p} & \dfrac{\partial^2 \ell_j}{\partial \boldsymbol{\alpha}_* \partial \phi_x} & \dfrac{\partial^2 \ell_j}{\partial \boldsymbol{\alpha}_* \partial \boldsymbol{\tau}'} \\ & & \dfrac{\partial^2 \ell_j}{\partial \alpha_p \partial \alpha_p} & \dfrac{\partial^2 \ell_j}{\partial \alpha_p \partial \phi_x} & \dfrac{\partial^2 \ell_j}{\partial \alpha_p \partial \boldsymbol{\tau}'} \\ & & & \dfrac{\partial^2 \ell_j}{\partial \phi_x \partial \phi_x} & \dfrac{\partial^2 \ell_j}{\partial \phi_x \partial \boldsymbol{\tau}'} \\ \mathit{Simetria} & & & & \dfrac{\partial^2 \ell_j}{\partial \boldsymbol{\tau} \partial \boldsymbol{\tau}'}, \end{pmatrix}, \tag{2.39}$$

cujos elementos são dados por

$$\frac{\partial^2 \ell_j}{\partial \mu_x \partial \mu_x} = -\frac{1}{\phi_x},$$

$$\frac{\partial^2 \ell_j}{\partial \mu_x \partial \boldsymbol{\alpha}'_*} = \mathbf{0}_{1\times(p-2)},$$

$$\frac{\partial^2 \ell_j}{\partial \mu_x \partial \alpha_p} = -\frac{1}{\phi_x},$$

$$\frac{\partial^2 \ell_j}{\partial \mu_x \partial \phi_x} = -\frac{W_{1j}}{\phi_x^2},$$

$$\frac{\partial^2 \ell_j}{\partial \mu_x \partial \boldsymbol{\tau}'} = \mathbf{0}_{1\times q},$$

$$\frac{\partial^2 \ell_j}{\partial \boldsymbol{\alpha}_* \partial \boldsymbol{\alpha}'_*} = -\mathbb{I}_{(q)}\mathrm{D}^{-1}(\boldsymbol{\tau})\mathbb{I}'_{(q)},$$

$$\frac{\partial^2\ell_j}{\partial\boldsymbol{\alpha}_*\partial\alpha_p} = \mathbb{I}_{(q)}\mathrm{D}^{-1}(\boldsymbol{\tau})\mathbf{1}_q \ ,$$

$$\frac{\partial^2\ell_j}{\partial\boldsymbol{\alpha}_*\partial\phi_x} = \mathbf{0}_{(p-2)\times 1} \ ,$$

$$\frac{\partial^2\ell_j}{\partial\boldsymbol{\alpha}_*\partial\boldsymbol{\tau}'} = W_{1j}\mathbb{I}_{(q)}\mathrm{D}^{-2}(\boldsymbol{\tau}) - \mathbb{I}_{(q)}\mathrm{D}^{-2}(\boldsymbol{\tau})\mathrm{D}(\mathbf{W}_{2j}) \ ,$$

$$\frac{\partial^2\ell_j}{\partial\alpha_p\partial\alpha_p} = -\frac{c_2}{\phi_x} \ ,$$

$$\frac{\partial^2\ell_j}{\partial\alpha_p\partial\phi_x} = -\frac{W_{1j}}{\phi_x^2} \ ,$$

$$\frac{\partial^2\ell_j}{\partial\alpha_p\partial\boldsymbol{\tau}'} = -W_{1j}\mathbf{1}_q'\mathrm{D}^{-2}(\boldsymbol{\tau}) + \mathbf{W}_{2j}'\mathbf{D}^{-2}(\boldsymbol{\tau}) \ ,$$

$$\frac{\partial^2\ell_j}{\partial\phi_x\partial\phi_x} = \frac{1}{2\phi_x^2} - \frac{W_{1j}^2}{\phi_x^3} \ ,$$

$$\frac{\partial^2\ell_j}{\partial\phi_x\partial\boldsymbol{\tau}'} = \mathbf{0}_{1\times q} \ ,$$

$$\frac{\partial^2\ell_j}{\partial\boldsymbol{\tau}\partial\boldsymbol{\tau}'} = \frac{\mathrm{D}^{-2}(\boldsymbol{\tau})}{2} - W_{1j}^2\mathrm{D}^{-3}(\boldsymbol{\tau}) + 2W_{1j}\mathrm{D}(\mathbf{W}_{2j})\mathrm{D}^{-3}(\boldsymbol{\tau}) - \mathrm{D}^2(\mathbf{W}_{2j})\mathrm{D}^{-3}(\boldsymbol{\tau}) \ .$$

Logo, a matriz de informação observada de $\boldsymbol{\theta} = (\mu_x, \alpha_*', \alpha_p, \phi_x, \phi_1, ..., \phi_q)'$ é obtida ao considerar

$$\mathbf{I}_F(\boldsymbol{\theta}) = -\ddot{\mathbf{L}} = -\frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'} = -\sum_{j=1}^{n}\frac{\partial^2\ell_j}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}}.$$

Dos resultados acima, da relação em (2.11), e dos resultados nos Lemas A.1 e A.2, dados no apêndice A, temos o seguinte Teorema.

Teorema 2.3. *Sob o modelo de Grubbs definido em (2.1), com $\phi_p = 0$, a matriz de informação esperada de $\boldsymbol{\theta}$ é dada por*

$$\mathbf{J}_{F1}(\boldsymbol{\theta}) = \begin{pmatrix} J_{\mu_x\mu_x} & 0 & J_{\mu_x\alpha_p} & 0 & 0 \\ 0 & J_{\alpha_*\alpha_*} & J_{\alpha_*\alpha_p} & 0 & 0 \\ J_{\alpha_p\mu_x} & J_{\alpha_p\alpha_*} & J_{\alpha_p\alpha_p} & 0 & 0 \\ 0 & 0 & 0 & J_{\phi_x\phi_x} & 0 \\ 0 & 0 & 0 & 0 & J_{\tau\tau} \end{pmatrix},$$

onde

$$J_{\mu_x\mu_x} = \frac{1}{\phi_x} \ ,$$

$$J_{\mu_x\alpha_p} \ = \frac{1}{\phi_x} \ ,$$

$$J_{\alpha_*\alpha_*} = \mathbb{I}_{(q)} D^{-1}(\boldsymbol{\tau})\mathbb{I}'_{(q)} \ ,$$

$$J_{\alpha_*\alpha_p} = -\mathbb{I}_{(q)} D^{-1}(\boldsymbol{\tau})\boldsymbol{1}_q \ ,$$

$$J_{\alpha_p\alpha_p} = \frac{c_2}{\phi_x} \ ,$$

$$J_{\phi_x\phi_x} = \frac{1}{2\phi_x^2} \ ,$$

$$J_{\tau\tau} \ = \frac{1}{2} D^{-2}(\boldsymbol{\tau}) \ .$$

Dos resultados nos Teoremas 2.2 e 2.3, e de resultados gerais de teoria assintótica (veja, Sen e Singer, 1993), segue que o EMV $\widehat{\boldsymbol{\theta}}$ de $\boldsymbol{\theta}$ em ambos casos, têm distribuição assintótica com média $\boldsymbol{\theta}$ e matriz de covariância $V = \mathbf{J}_{F1}^{-1}(\boldsymbol{\theta})/\mathrm{n}$, que pode ser utilizado, por exemplo, para testar alguma hipótese de interesse.

É importante observar que sob a condição $\phi_1 = 0$, os estimadores $\widehat{\boldsymbol{\alpha}}$, $\widehat{\mu}_x$, $\widehat{\phi}_x$ e $\widehat{\boldsymbol{\psi}}$ são assintoticamente independentes, pois a matriz de covariância assintótica é diagonal em bloco. No caso em que $\phi_p = 0$, os estimadores $\widehat{\mu}_x$ e $\widehat{\boldsymbol{\alpha}}$ são assintoticamente independentes de $\widehat{\phi}_x$ e $\widehat{\boldsymbol{\psi}}$.

2.4 Testes de Hipóteses

Nesta seção apresentamos um estudo de testes de hipóteses para avaliar as hipóteses definidas na Seção 1.1 utilizando as estatísticas da Razão de Verossimilhanças (Q), Escore (E) e de Wald (W). Estas estatísticas são derivadas para algumas suposições sobre o vetor de parâmetros $\boldsymbol{\theta}$ do modelo definido em (2.1-2.2).

Sejam $\widehat{\boldsymbol{\theta}}$ e $\widetilde{\boldsymbol{\theta}}$ os EMV de $\boldsymbol{\theta}$ sob o modelo irrestrito e sob a hipótese nula, respectivamente. Como as hipóteses definidas na Seção 1.1, podem ser escritas como $H_0 \ : \ \mathbf{A}\boldsymbol{\theta} \ = \ \boldsymbol{\varphi}_o$, onde a matriz $\mathbf{A}$ de dimensão $r \times (2p+1)$ tem posto $pos(\mathbf{A}) = r \leq 2p+1$ e $\boldsymbol{\varphi}_o \ : \ r \times 1$ é um vetor conhecido. Logo as estatísticas Q, E e W podem ser escritas, respectivamente como

$$Q = 2[\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widetilde{\boldsymbol{\theta}})], \ E = \frac{1}{n}[\frac{\partial}{\partial \boldsymbol{\theta}}\ell(\widetilde{\boldsymbol{\theta}})]'[\mathbf{J}_{F1}^{-1}(\widetilde{\boldsymbol{\theta}})][\frac{\partial}{\partial \boldsymbol{\theta}}\ell(\widetilde{\boldsymbol{\theta}})]$$

e

$$W = n[\mathbf{A}\widehat{\boldsymbol{\theta}} - \boldsymbol{\varphi}_o]'[\mathbf{A}'\mathbf{J}_{F1}^{-1}(\widehat{\boldsymbol{\theta}})\mathbf{A}]^{-1}[\mathbf{A}\widehat{\boldsymbol{\theta}} - \boldsymbol{\varphi}_o],$$

onde $\ell(\boldsymbol{\theta})$ é a função de log-verossimilhança definida em (2.5) e $\mathbf{J}_{F1}$ é a matriz de informação de Fisher esperada definida em (2.11).

As três estatísticas definidas acima são assintoticamente equivalentes sob H_0, de fato elas convergem em distribuição para uma variável aleatória tendo distribuição qui-quadrado com r graus de liberdade (χ_r^2). Isto é, sob H_0, Q, E ou W tem a mesma distribuição assintótica, veja por exemplo, Sen e Singer (1993).

2.4.1 Teste de Wald

Seja $\boldsymbol{\theta} = (\boldsymbol{\theta}_L', \boldsymbol{\theta}_E')'$, onde $\boldsymbol{\theta}_L = (\mu_x, \alpha_2, \ldots, \alpha_p)'$ e $\boldsymbol{\theta}_E = (\phi_x, \phi_1, \ldots, \phi_p)'$. Então a matriz de informação de Fisher esperada, dada no Teorema 2.1, pode ser representada por $\mathbf{J}_{F1}(\boldsymbol{\theta}) = \mathrm{Diag}(\mathbf{J}_L, \mathbf{J}_E)$, onde

$$\mathbf{J}_L = \begin{pmatrix} J_{\mu_x\mu_x} & J_{\mu_x\alpha} \\ J_{\alpha\mu_x} & J_{\alpha\alpha} \end{pmatrix} \text{ e } \mathbf{J}_E = \begin{pmatrix} J_{\phi_x\phi_x} & J_{\phi_x\phi} \\ J_{\phi\phi_x} & J_{\phi\phi} \end{pmatrix}$$

são as submatrizes de informação dos vetores na locação $\boldsymbol{\phi}_L = (\mu_x, \boldsymbol{\alpha}')'$ e na escala $\boldsymbol{\phi}_E = (\phi_x, \boldsymbol{\phi}')'$, respectivamente.

Agora, se $\widehat{\boldsymbol{\theta}}$ é o EMV de $\boldsymbol{\theta}$, então utilizando o Teorema 5.2.1 dado em Sen e Singer (1993), temos que

$$\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{\mathrm{d}} N_{2p+1}(0, \mathbf{J}_{F1}^{-1}(\boldsymbol{\theta})), \tag{2.40}$$

onde $\mathbf{J}_{F1}^{-1}(\boldsymbol{\theta}) = \mathrm{Diag}(\mathbf{J}_L^{-1}, \mathbf{J}_E^{-1})$.

Por outro lado, se $\boldsymbol{\theta}_* = g(\boldsymbol{\theta}) = (\alpha_2, \ldots, \alpha_p, \phi_1, \ldots, \phi_p)' = (g_1(\boldsymbol{\theta}), g_2(\boldsymbol{\theta}), \ldots, g_{2p-1}(\boldsymbol{\theta}))'$, então, usando o método delta temos que

$$\sqrt{n}(\widehat{\boldsymbol{\theta}}_* - \boldsymbol{\theta}_*) \xrightarrow{\mathrm{d}} N_{2p-1}(\mathbf{0}, \boldsymbol{\Omega}_*), \tag{2.41}$$

onde $\boldsymbol{\Omega}_* = \mathbf{D}\mathbf{J}_{F1}^{-1}\mathbf{D}'$, com $\mathbf{J}_{F1}$ a matriz de informação esperada, definida no Teorema 2.1 e $\mathbf{D}$ é uma matriz $(2p-1) \times (2p+1)$, tal que

$$\mathbf{D} = [\frac{\partial g_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_j}] = \begin{pmatrix} \mathbb{I}_{(p)} & \mathbf{0} \\ \mathbf{0} & \mathbb{I}_{(p+1)} \end{pmatrix},$$

assim de (2.40) e (2.41), temos que

$$\boldsymbol{\Omega}_* = \begin{pmatrix} \mathbb{I}_{(p)}\mathbf{J}_L^{-1}\mathbb{I}'_{(p)} & \mathbf{0} \\ \mathbf{0} & \mathbb{I}_{(p+1)}\mathbf{J}_E^{-1}\mathbb{I}'_{(p+1)} \end{pmatrix}, \tag{2.42}$$

onde $\mathbb{I}_{(r)}$, $r = p,\ p+1$, como definido na Seção 2.2.1 (a).

Para testar se os instrumentos medem a característica de interesse sem vício e com a mesma precisão, a hipótese é

$$H_{01} \ : \ \alpha_2 = \alpha_3 = ... = \alpha_p = 0 \ \text{ e } \ \phi_1 = \phi_2 = ... = \phi_p \ . \tag{2.43}$$

Note que, H_{01} pode ser escrita como

$$H_{01} \ : \ \mathbf{C}\boldsymbol{\theta}_* = \mathbf{q}_o\,, \tag{2.44}$$

onde $\mathbf{C} = \begin{pmatrix} \mathbf{I}_q & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_1 \end{pmatrix}$ é uma matriz $(2p-2) \times (2p-1)$ de posto $2p-2$, com $q = p-1$ e

$$\mathbf{A}_1 = \begin{pmatrix} 1 & -1 & 0 & \dots & 0 & 0 \\ 0 & 1 & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & -1 \end{pmatrix}, \tag{2.45}$$

matriz $(p-1) \times p$ e $\mathbf{q}_o$ vetor nulo de dimensão $2(p-1)$. Logo, a estatística de Wald para testar a hipótese H_{01} é dada por

$$W_{01} = n(\mathbf{C}\widehat{\boldsymbol{\theta}}_* - \mathbf{q}_o)'[\mathbf{C}\widehat{\boldsymbol{\Omega}}_*\,\mathbf{C}']^{-1}(\mathbf{C}\widehat{\boldsymbol{\theta}}_* - \mathbf{q}_o), \tag{2.46}$$

onde $\widehat{\boldsymbol{\Omega}}_*$ denota a matriz $\boldsymbol{\Omega}_*$ com os parâmetros desconhecidos $\boldsymbol{\theta}_*$, substituídos pelo EMV $\widehat{\theta}_*$ de $\boldsymbol{\theta}_*$. Portanto, H_{01} é rejeitada ao nível α se $W_{01} > \chi^2_{1-\alpha}$, onde $\chi^2_{1-\alpha}$ denota o percentil $100(1-\alpha)\%$ da distribuição qui-quadrado com $2(p-1)$ graus de liberdade.

Observe que de (2.42) e após as manipulações algébricas, $\mathbf{C}\mathbf{\Omega}_*\mathbf{C}'$ pode ser escrita por

$$\mathbf{C}\mathbf{\Omega}_*\mathbf{C}' = \begin{pmatrix} \mathbb{I}_{(p)}\mathbf{J}_L^{-1}\mathbb{I}'_{(p)} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_1\mathbb{I}_{(p+1)}\mathbf{J}_E^{-1}\mathbb{I}'_{(p+1)}\mathbf{A}'_1 \end{pmatrix},$$

e conseqüentemente a estatística W_{01} pode ser escrita como

$$W_{01} = n\widehat{\boldsymbol{\alpha}}'(\mathbb{I}_{(p)}\widehat{\mathbf{J}}_L^{-1}\mathbb{I}'_{(p)})^{-1}\widehat{\boldsymbol{\alpha}} + n\widehat{\boldsymbol{\phi}}'\mathbf{A}'_1(\mathbf{A}_1\mathbb{I}_{(p+1)}\widehat{\mathbf{J}}_E^{-1}\mathbb{I}'_{(p+1)}\mathbf{A}'_1)^{-1}\mathbf{A}_1\widehat{\boldsymbol{\phi}}. \tag{2.47}$$

Para testar a hipótese de interesse

$$H_{02} \ : \ \alpha_2 = \alpha_3 = ... = \alpha_p = 0 \ \ , \tag{2.48}$$

isto é, para testar se os instrumentos medem sem vício a característica de interesse x, vamos escrever H_{02} como

$$H_{02} \ : \ \mathbf{C}\boldsymbol{\theta}_* = \mathbf{0}_q \, , \tag{2.49}$$

onde $\mathbf{C} = [\mathbf{I}_q \ \mathbf{0}]$ é uma matriz $(p-1) \times (2p-1)$ de posto $q = p-1$. Assim, a estatística de Wald para testar H_{02}, denotada por W_{02}, é dada por

$$\begin{aligned} W_{02} &= n(\mathbf{C}\widehat{\boldsymbol{\theta}}_*)'[\mathbf{C}\widehat{\mathbf{\Omega}}_*\mathbf{C}']^{-1}(\mathbf{C}\widehat{\boldsymbol{\theta}}_*) \\ &= n\widehat{\boldsymbol{\alpha}}'(\mathbb{I}_{(p)}\widehat{\mathbf{J}}_L^{-1}\mathbb{I}'_{(p)})^{-1}\widehat{\boldsymbol{\alpha}} \, , \end{aligned} \tag{2.50}$$

que converge em distribuição para uma variável aleatória $\chi^2_{(p-1)}$ sob H_{02}. Portanto, H_{02} é rejeitado ao nível α se $W_{02} > \chi^2_{1-\alpha}$, onde $\chi^2_{1-\alpha}$ denota o percentil $100(1-\alpha)\%$ da distribuição qui-quadrado com $p-1$ graus de liberdade.

Finalmente, para testar se os instrumentos são igualmente precisos, a hipótese é

$$H_{03} \ : \ \phi_1 = \phi_2 = ... = \phi_p \quad \text{ou} \ \ \mathbf{C}\boldsymbol{\theta}_* = \mathbf{0}_q \, , \tag{2.51}$$

onde $\mathbf{C} = [\mathbf{0} \ \mathbf{A}_1]$ é uma matriz $(p-1) \times (2p-1)$ de posto $q = p-1$, com $\mathbf{A}_1$ como em (2.45). A estatística de Wald para testar H_{03}, denotada por W_{03} é dada por

$$\begin{aligned} W_{03} &= n(\mathbf{C}\widehat{\boldsymbol{\theta}}_*)'[\mathbf{C}\widehat{\mathbf{\Omega}}_*\mathbf{C}']^{-1}(\mathbf{C}\widehat{\boldsymbol{\theta}}_*) \\ &= n\widehat{\boldsymbol{\phi}}'\mathbf{A}'_1(\mathbf{A}_1\mathbb{I}_{(p+1)}\widehat{\mathbf{J}}_E^{-1}\mathbb{I}'_{(p+1)}\mathbf{A}'_1)^{-1}\mathbf{A}_1\widehat{\boldsymbol{\phi}} \, , \end{aligned} \tag{2.52}$$

que converge em distribuição para uma variável aleatória $\chi^2_(p-1)$ sob H_{03}. Portanto, H_{03} é rejeitada ao nível α se $W_{03} > \chi^2_{1-\alpha}$, onde $\chi^2_{1-\alpha}$ denota o percentil $100(1-\alpha)\%$

da distribuição qui-quadrado com $p-1$ graus de liberdade.

É importante observar que das relações em (2.47), (2.50) e (2.52), segue que

$$W_{01} = W_{02} + W_{03}, \tag{2.53}$$

onde W_{02} e W_{03} são as estatísticas de Wald utilizadas para testar H_{02} e H_{03}, respectivamente.

2.4.2 Teste de Escore

Rao (1948) introduz a estatística de Escore dada por

$$E = \frac{1}{n}\mathbf{S}(\widetilde{\boldsymbol{\theta}})'\widetilde{\mathbf{J}}_{F1}^{-1}\mathbf{S}(\widetilde{\boldsymbol{\theta}})\,, \tag{2.54}$$

onde $\mathbf{S}(\boldsymbol{\theta})$ é o vetor de derivadas parciais da log-verossimilhança $\ell(\boldsymbol{\theta})$ definido em (2.5), $\widetilde{\boldsymbol{\theta}}$ é o vetor de EMV restrito sob a hipótese nula, e $\widetilde{\mathbf{J}}_{F1}$ é a matriz de informação esperada avaliado em $\widetilde{\boldsymbol{\theta}}$. Esta estatística é atrativa porque somente requer calcular os EMV sob a hipótese nula e é assintoticamente equivalente ao teste de Wald e teste da Razão de Verossimilhanças sob a hipótese nula.

A estatística de Escore tem sido utilizada no trabalho desenvolvido por Bedrick (2001), que apresenta um estudo de inferência no modelo de Grubbs sob uma parametrização diferente. É importante observar que os EMV sob H_{01}, apresentados abaixo, coincidem com o EMV proposto por Bedrick (2001).

Sob H_{01}, os EMV foram encontrados em (2.21) e são dados por

$$\begin{aligned}
\widetilde{\mu}_x &= \bar{y}_{..}\ , \\
\widetilde{\phi} &= \frac{1}{n(p-1)}\sum_{j=1}^{n}\sum_{i=1}^{p}(y_{ij}-\bar{y}_{.j})^2\,, \\
\widetilde{\phi}_x &= \frac{1}{n}\sum_{j=1}^{n}(\bar{y}_{.j}-\bar{y}_{..})^2 - \frac{\widetilde{\phi}}{p} = \frac{1}{pn}\sum_{j=1}^{n}\sum_{i=1}^{p}(y_{ij}-\bar{y}_{..})^2 - \widetilde{\phi}\,,
\end{aligned} \tag{2.55}$$

onde ϕ é considerado como sendo a variância comum na hipótese

$$H_{01}\ :\ \alpha_2 = \alpha_3 = ... = \alpha_p = 0 \ \text{ e } \ \phi_1 = \phi_2 = ... = \phi_p = \phi\ \ . \tag{2.56}$$

Se ϕ_x é estimado negativamente, consideramos $\widetilde{\phi}_x = 0$ e o EMV restrito para ϕ é dado por

$$\widetilde{\phi} = \frac{1}{pn}\sum_{j=1}^{n}\sum_{i=1}^{p}(y_{ij} - \bar{y}_{..})^2 \,. \tag{2.57}$$

Mostramos no Apêndice C que a estatística de Escore para testar H_{01} é dada por

$$E_{01} = E_1 + E_2 \,, \tag{2.58}$$

onde

$$E_1 = \frac{n}{\widetilde{\phi}}\sum_{i=1}^{p}(\bar{y}_{i.} - \bar{y}_{..})^2 \;\text{ e }\; E_2 = \frac{1}{n}\mathbf{A}'\,\mathbf{B}\,\mathbf{A}\,,$$

com

$$\mathbf{A} = \frac{1}{2\widetilde{\phi}^2}\left[n(\widetilde{\phi}_x - \widetilde{\phi})\mathbf{1}_p + \sum_{j=1}^{n}\mathrm{D}(\mathbf{W}_j)\mathbf{W}_j - 2\widetilde{\tau}\sum_{j=1}^{n}\mathbf{W}_j'\mathbf{1}_p\mathbf{W}_j\right]$$

e

$$\mathbf{B} = \frac{2\widetilde{\phi}^2}{1-2\widetilde{\tau}}\left[\mathbf{I}_p - \frac{1-\dfrac{2\widetilde{\tau}\widetilde{\phi}}{\widetilde{\phi}_x}}{p(p-1)}\mathbf{1}_p\mathbf{1}_p'\right],$$

onde $\mathbf{W}_j = \mathbf{y}_j - \mathbf{1}_p\widetilde{\mu}_x$ e $\widetilde{\tau} = \dfrac{\widetilde{\phi}_x}{\widetilde{\phi} + p\,\widetilde{\phi}_x}$.

Como $\mathbf{J}_{F1}$ é uma matriz diagonal em bloco (veja, Teorema, 2.1), então as estatísticas E_1 e E_2 são assintoticamente independentes. Além disso, as estatísticas E_1 e E_2 são as estatísticas de Escore para testar H_{02} e H_{03} individualmente, assumindo que a outra hipótese é verdadeira. Resultado similar é obtido no trabalho de Bedrick (2001), para o modelo de Grubbs, usando uma parametrização diferente deste trabalho.

Para amostras grandes, temos que E_{01} converge em distribuição a uma variável aleatória $\chi^2_{2(p-1)}$ sob H_{01}. Portanto, H_{01} é rejeitada ao nível α se $E_{01} > \chi^2_{1-\alpha}$, onde $\chi^2_{1-\alpha}$ denota o percentil $100(1-\alpha)\%$ da distribuição qui-quadrado com $2(p-1)$ graus de liberdade.

Sob H_{03} : $\phi_1 = \phi_2 = ... = \phi_p = \phi$ os EMV tem uma forma fechada e são dados por

$$\widetilde{\mu}_x = \bar{y}_{1.} \ ,$$

$$\widetilde{\alpha}_i = \bar{y}_{i.} - \widetilde{\mu}_x \quad i = 2, \ldots, p \ ,$$

$$\widetilde{\phi} = \frac{1}{n(p-1)} \sum_{j=1}^{n} \sum_{i=1}^{p} \left[(y_{ij} - \bar{y}_{i.})^2 - (\bar{y}_{.j} - \bar{y}_{..})^2 \right] \ ,$$

$$\widetilde{\phi}_x = \frac{1}{n} \sum_{j=1}^{n} (\bar{y}_{.j} - \bar{y}_{..})^2 - \frac{\widetilde{\phi}}{p} \ . \qquad (2.59)$$

Se ϕ_x é estimado negativamente, consideramos $\widetilde{\phi}_x = 0$ e o EMV restrito para ϕ, dado por

$$\widetilde{\phi} = \frac{1}{pn} \sum_{j=1}^{n} \sum_{i=1}^{p} (y_{ij} - \bar{y}_{i.})^2.$$

Neste caso a estatística de Escore para testar H_{03} é dada por

$$E_{03} = \frac{1}{n} \mathbf{A}' \, \mathbf{B} \, \mathbf{A} \, ,$$

onde

$$\mathbf{A} = \frac{1}{2\widetilde{\phi}^2} \left[n(\widetilde{\phi}_x - \widetilde{\phi}) \mathbf{1}_p + \sum_{j=1}^{n} \mathrm{D}(\mathbf{W}_j) \mathbf{W}_j - 2\widetilde{\tau} \sum_{j=1}^{n} \ \mathbf{W}_j' \mathbf{1}_p \mathbf{W}_j \right]$$

e

$$\mathbf{B} = \frac{2\widetilde{\phi}^2}{1 - 2\widetilde{\tau}} \left[\mathbf{I}_p - \frac{1 - \dfrac{2\widetilde{\tau}\widetilde{\phi}}{\widetilde{\phi}_x}}{p(p-1)} \mathbf{1}_p \mathbf{1}_p' \right] ,$$

com $\mathbf{W}_j = \mathbf{y}_j - (0, \widetilde{\boldsymbol{\alpha}}')' - \mathbf{1}_p \, \widetilde{\mu}_x$, $\widetilde{\tau} = \dfrac{\widetilde{\phi}_x}{\widetilde{\phi} + p \, \widetilde{\phi}_x}$ e $\widetilde{\mu}_x, \widetilde{\alpha}_i, \widetilde{\phi}_x, \widetilde{\phi}$ como em (2.59).

Para amostras grandes temos que a E_{03} converge em distribuição a uma variável aleatória $\chi^2_{(p-1)}$ sob H_{03}. Portanto, H_{03} é rejeitada ao nível α se $E_{03} > \chi^2_{1-\alpha}$, onde $\chi^2_{1-\alpha}$ denota o percentil $100(1-\alpha)\%$ da distribuição qui-quadrado com $(p-1)$ graus de liberdade.

Finalmente, sob H_{02} o EMV de $\boldsymbol{\theta}$ não tem uma forma explícita, neste caso o EMV de $\boldsymbol{\theta}$ pode ser obtido via um procedimento iterativo, neste trabalho é obtido via algoritmo EM (veja, Seção 2.3.3) e a estatística de Escore é da forma

$$E_{02} = \frac{1}{n}\mathbf{S}(\widetilde{\boldsymbol{\theta}})'\widetilde{\mathbf{J}}_{F1}^{-1}\mathbf{S}(\widetilde{\boldsymbol{\theta}})\,, \tag{2.60}$$

A implementação pode ser facilmente manipulada com programas computacionais, como por exemplo, Mathematica ou Matlab. Note que sob H_{02} a estatística E_{02} converge em distribuição para uma variável aleatória χ^2_{p-1}

2.4.3 Teste da Razão de Verossimilhanças

A estatística da Razão de Verossimilhanças é definida como sendo

$$Q = 2\left[\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widetilde{\boldsymbol{\theta}})\right]\,, \tag{2.61}$$

onde $\ell(\boldsymbol{\theta})$ é a função de log-verossimilhança definida em (2.5), $\widehat{\boldsymbol{\theta}}$ é o EMV do modelo irrestrito , e $\widetilde{\boldsymbol{\theta}}$ é o EMV sob a hipótese nula que tem a forma $\mathbf{A}\boldsymbol{\theta}_*$, onde $\boldsymbol{\theta}_*$ é como em (2.41) e $\mathbf{A}$ uma matriz de ordem $r \times (2p-1)$, com posto r.

Neste caso requeremos os EMV de $\boldsymbol{\theta}$ sob a hipótese nula e sob o modelo irrestrito, cujo método de implementação para obter $\widehat{\boldsymbol{\theta}}$ e $\widetilde{\boldsymbol{\theta}}$ é apresentado na Seção 2.2.

Embora não seja possível obter uma expressão fechada da estatística Q, sua implementação é computacionalmente simples, dado que os EMV sob as hipóteses nulas H_{01}, H_{02} e H_{03} têm uma forma fechada ou é de fácil implementação.

2.4.4 Ilustração Numérica

Consideremos os seguintes dados, tomados de Jaech (1985), relativos a um experimento, no qual a densidade de 43 pastilhas de combustível de urânio sintetizado para uso em reatores nucleares foram medidas por seis "instrumentos".

Instrumento 1: Método geométrico, operador 1

Instrumento 2: Método geométrico, operador 2

Instrumento 3: Método geométrico, operador 3

Instrumento 4: Método de imersão, operador 4

Tabela 2.1: Medidas de densidade(%TD) de pastilhas de urânio sintetizado, fornecidas por seis instrumentos.

obs.	Instrum.1	Instrum. 2	Instrum. 3	Instrum. 4	Instrum. 5	Instrum. 6
1	4.30	4.06	4.14	4.68	5.22	4.75
2	4.51	4.35	4.53	5.15	4.45	4.79
3	4.42	4.52	4.57	4.78	4.79	4.72
4	4.90	4.67	4.75	5.55	5.10	4.75
5	4.60	4.70	4.84	5.08	4.82	5.05
6	4.29	4.30	4.21	4.80	4.47	4.93
7	4.30	4.47	4.01	4.41	4.27	4.66
8	4.69	4.59	4.47	5.32	4.95	4.75
9	4.02	4.05	4.09	4.12	4.48	4.40
10	4.61	4.78	4.60	4.41	4.75	5.18
11	4.35	4.25	4.40	4.37	4.06	4.39
12	4.46	4.55	4.70	4.73	5.14	4.75
13	4.08	4.12	3.85	4.60	4.22	4.33
14	3.95	3.95	4.13	4.62	4.49	3.94
15	4.21	4.27	3.94	4.34	4.29	4.62
16	4.62	4.47	4.50	5.09	5.44	4.70
17	4.81	4.81	4.98	5.26	4.86	5.22
18	4.47	4.34	4.33	5.05	4.62	4.38
19	4.38	4.24	4.47	4.60	4.27	4.2
20	4.43	4.37	4.57	5.23	5.04	4.51
21	4.74	4.52	4.58	5.24	5.36	4.77
22	4.35	4.37	4.45	4.42	4.66	4.54
23	4.11	4.23	4.07	4.46	4.48	4.26
24	3.97	3.98	4.19	4.07	4.69	4.29
25	4.40	4.36	4.33	4.55	4.73	4.62
26	4.24	4.30	4.36	4.84	4.56	4.97
27	4.51	4.52	4.57	4.70	4.20	4.69
28	4.49	4.33	4.61	4.85	4.26	4.67
29	4.42	4.14	4.22	4.35	4.39	4.46
30	4.49	4.59	4.69	5.06	5.50	4.75
31	4.14	4.32	4.45	4.02	4.30	4.58
32	4.38	4.25	4.44	4.86	4.81	4.63
33	4.45	4.60	4.28	4.40	4.17	4.81
34	4.39	4.38	4.64	4.86	4.98	4.54
35	4.56	4.35	4.83	4.53	5.36	4.46
36	4.38	4.40	4.60	4.96	4.92	4.56
37	4.44	4.34	4.66	4.60	5.59	4.66
38	4.21	4.24	4.07	4.21	4.39	4.40
39	4.51	4.52	4.54	4.79	4.85	4.77
40	4.23	4.12	4.35	4.62	4.45	4.15
41	4.33	4.47	4.68	5.05	4.29	4.92
42	4.27	4.21	4.45	4.74	5.24	4.41
43	4.67	4.52	4.60	4.42	5.06	4.73

Instrumento 5: Método de imersão, operador 5

Instrumento 6: Método de imersão, operador 6.

O método geométrico para medir a densidade, neste experimento, é baseado em pesar uma pastilha e encontrar seu volume, medindo o diâmetro e longitude da pastilha. O método de imersão é baseado em notar a mudança no peso quando as pastilhas são pesadas no ar e num determinado líquido.

Seja y_{ij} a densidade fornecida pelo instrumento i e a pastilha j, com $i = 1, \ldots, 6$ e $j = 1, \ldots, 43$, expressada como densidade teórica percentual (%TD) menos 90% por conveniência. O conjunto de observações y_{ij} é apresentado na Tabela 2.1.

A suposição de normalidade requerida para o estudo estatístico foi avaliada fazendo o gráfico de probabilidade normal para cada um dos seis instrumentos. Os gráficos indicam que os dados seguem uma distribuição normal de maneira razoável.

Seguindo Jaech (1985), vamos supor um modelo de Grubbs estrutural normal, ver (2.1), com p=6 e n=43. Adotaremos o instrumento 1, como sendo de referência, isto é, $\alpha_1 = 0$.

A Tabela 2.2 mostra algumas iterações do algoritmo EM, indicando que a convergência é atingida rapidamente e aproximadamente na iteração 50, dando estimativas para ϕ_i, $i = 1, \ldots, 6$.

Tabela 2.2: Convergência do algoritmo EM.

iter	$\widehat{\phi}_1$	$\widehat{\phi}_2$	$\widehat{\phi}_3$	$\widehat{\phi}_4$	$\widehat{\phi}_5$	$\widehat{\phi}_6$
11	0.0095	0.0042	0.0275	0.0813	0.1394	0.0274
21	0.0072	0.0073	0.0251	0.0759	0.1319	0.0304
31	0.0069	0.0076	0.0249	0.0754	0.1313	0.0308
41	0.0068	0.0076	0.0249	0.0753	0.1313	0.0309
51	0.0068	0.0076	0.0249	0.0753	0.1313	0.0309
desv. pad.	(0.0024)	(0.0026)	(0.0060)	(0.0168)	(0.0289)	(0.0073)

Note que as estimativas das variâncias são aproximadas às obtidas por Jaech (1985, pags. 162 e 193) considerando diferenças pareadas e modelo de vícios constantes para os dados originais. Além disso, os seguintes EMV de μ_x , ϕ_x e $\boldsymbol{\alpha}$ foram obtidos

$$\widehat{\mu}_x = 4.3972\,(0.0316);$$

$$\widehat{\phi}_x = 0.0361\,(0.0084);$$

$$\widehat{\alpha}_2 = -0.0270\,(0.0183),\ \widehat{\alpha}_3 = 0.0386\,(0.0272),\ \widehat{\alpha}_4 = 0.3188\,(0.0437),$$

$$\widehat{\alpha}_5 = 0.3230\,(0.0567),\ \widehat{\alpha}_6 = 0.2228\,(0.0296),$$

onde os números entre parênteses indicam os desvios padrões assintóticos estimados dos respectivos estimadores dos parâmetros.

Do Teorema 2.1, segue que a matriz de informação esperada de Fisher é dada por

$$\mathbf{J}_{F1}(\widehat{\boldsymbol{\theta}}) = 10^3 \times \begin{pmatrix} \widehat{\mathbf{J}}_1 & \mathbf{0} \\ \mathbf{0} & \widehat{\mathbf{J}}_2 \end{pmatrix},$$

onde

$$\widehat{J}_1 = \begin{pmatrix} 0.0258 & 0.0091 & 0.0028 & 0.0009 & 0.0005 & 0.0022 \\ 0.0091 & 0.0882 & -0.0133 & -0.0044 & -0.0025 & -0.0107 \\ 0.0028 & -0.0133 & 0.0361 & -0.0013 & -0.0008 & -0.0033 \\ 0.0009 & -0.0044 & -0.0013 & 0.0128 & -0.0003 & -0.0011 \\ 0.0005 & -0.0025 & -0.0008 & -0.0003 & 0.0075 & -0.0006 \\ 0.0022 & -0.0107 & -0.0033 & -0.0011 & -0.0006 & 0.0298 \end{pmatrix},$$

$$\widehat{J}_2 = \begin{pmatrix} 0.3324 & 0.0515 & 0.0418 & 0.0039 & 0.0004 & 0.0001 & 0.0025 \\ 0.0515 & 4.2839 & 1.1622 & 0.1080 & 0.0118 & 0.0039 & 0.0703 \\ 0.0418 & 1.1622 & 3.8894 & 0.0878 & 0.0096 & 0.0032 & 0.0571 \\ 0.0039 & 0.1080 & 0.0878 & 0.6515 & 0.0009 & 0.0003 & 0.0053 \\ 0.0004 & 0.0118 & 0.0096 & 0.0009 & 0.0823 & 0.0000 & 0.0006 \\ 0.0001 & 0.0039 & 0.0032 & 0.0003 & 0.0000 & 0.0279 & 0.0002 \\ 0.0025 & 0.0703 & 0.0571 & 0.0053 & 0.0006 & 0.0002 & 0.4426 \end{pmatrix}.$$

Para calcular as estatísticas de Escore e Razão de Verossimilhanças será necessário calcular os EMV sob as hipóteses nulas H_{01}, H_{02} e H_{03}. Assim, dos resultados da Seção 2.3, segue que: Sob H_{01}, as estimativas de máxima verossimilhança de $\boldsymbol{\theta}$ são

$$\widetilde{\mu}_x = 4.5433\,,\ \widetilde{\phi} = 0.0709\,,\ \widetilde{\phi}_x = 0.0350\,.$$

Sob H_{02}, os EMV de $\boldsymbol{\theta}$ são calculados mediante o algoritmo EM e são dados por

$$\widetilde{\mu}_x = 4.4152\,,\ \widetilde{\phi}_x = 0.0360\,,\ \widetilde{\phi}_1 = 0.0063\,,\ \widetilde{\phi}_2 = 0.0118\,,\ \widetilde{\phi}_3 = 0.0246\,,$$

$$\widetilde{\phi}_4 = 0.1665\,,\ \widetilde{\phi}_5 = 0.2231\,,\ \widetilde{\phi}_6 = 0.0776\,,$$

notando que a convergência é atingida aproximadamente na iteração 40, e finalmente, sob H_{03} os EMV são dados por

$$\widetilde{\mu}_x = 4.3972\,,\ \widetilde{\alpha}_2 = -0.0270\,,\ \widetilde{\alpha}_3 = 0.0386\,,\ \widetilde{\alpha}_4 = 0.3188\,,\ \widetilde{\alpha}_5 = 0.3230\,,$$

$$\widetilde{\alpha}_6 = 0.2228\,,\ \widetilde{\phi}_x = 0.0393\,,\ \widetilde{\phi} = 0.0449\,.$$

Usando estes resultados e as estatística de Wald, Escore e Razão de Verossimilhanças dadas nas subseções precedentes, podemos testar hipóteses de interesse H_{01}, H_{02} e H_{03} definidas no Capitulo 1. Com efeito, dos resultados da Seção 2.4.1, com p=6 e n=43, temos que $W_{01} = 200.8226$, com 10 graus de liberdade, para testar H_{01}, para H_{02} temos que $W_{02} = 151.5859$, com cinco graus de liberdade, para H_{03} temos que $W_{03} = 49.2366$, com cinco graus de liberdade, significa que mediante uso da estatística de Wald, todas as hipóteses são rejeitadas a um nível de 5%. Dos resultados da Seção 2.4.2 e 2.4.3, temos que $E_{01} = 124.7358$ e $Q_{01} = 192.7844$, com 10 graus de liberdade, para testar H_{01}, para testar H_{02} temos que $E_{02} = 74.8692$ e $Q_{02} = 103.9554$, com cinco graus de liberdade, e para testar H_{03} temos que $E_{03} = 77.9824$ e $Q_{03} = 94.7256$, com cinco graus de liberdade, indicando, que todas as hipóteses são rejeitadas a um nível de 5%. Note que $W_{01} = W_{02} + W_{03}$, o qual verifica o argumento prévio dado em (2.53). A conclusão geral é que os instrumentos são viciados, i.e., medem com vicio a densidade das pastilhas de urânio e não são igualmente precisos.

A fim de ilustrar os resultados da Seção 2.3.4, referente à EMV restrita e segundo Jaech (1985, pag. 242), vamos supor que $\phi_1 = 0$ para os dados da Tabela 2.1. Note que a estimativa correspondente a esta variância foi calculada como sendo perto de zero.

Da equação (2.31), as estimativas de máxima verossimilhança são

$$\widehat{\mu}_x = 4.3972\,(0.0320);$$

$$\widehat{\alpha}_2 = -0.0270\,(0.0187),\ \widehat{\alpha}_3 = 0.0386\,(0.0263),\ \widehat{\alpha}_4 = 0.3188\,(0.0408),$$

$$\widehat{\alpha}_5 = 0.3230\,(0.0549),\ \widehat{\alpha}_6 = 0.2228\,(0.0318);$$

$$\widehat{\phi}_x = 0.0439\,(0.0095);$$

$$\widehat{\phi}_2 = 0.0151\,(0.0032),\ \widehat{\phi}_3 = 0.0297\,(0.0064),\ \widehat{\phi}_4 = 0.0714\,(0.0154),$$

$$\widehat{\phi}_5 = 0.1294\,(0.0279),\ \widehat{\phi}_6 = 0.0435\,(0.0094),$$

onde os números entre parênteses são os desvios padrões assintóticos estimados dos respectivos estimadores dos parâmetros. Note que estas estimativas obtidas são ligeiramente aproximadas às obtidas usando o algoritmo EM e as variâncias dos erros coincidem com as obtidas por Jaech (1985, pag. 242).

Do Teorema 2.2, segue que a matriz de informação esperada de Fisher do vetor de parâmetros $\boldsymbol{\theta}_* = \mathrm{g}(\boldsymbol{\theta}) = (\alpha_2, \ldots, \alpha_p, \phi_2, \ldots, \phi_p)'$ é dada por

$$\mathbf{J}_{F1}(\widehat{\boldsymbol{\theta}}) = 10^3 \times \mathrm{Diag}(\widehat{\mathbf{V}}),$$

onde

$$\widehat{\mathbf{V}} = \begin{pmatrix} 0.0228 \\ 0.0664 \\ 0.0337 \\ 0.0140 \\ 0.0077 \\ 0.0230 \\ 0.2593 \\ 2.2064 \\ 0.5681 \\ 0.0980 \\ 0.0298 \\ 0.2646 \end{pmatrix}.$$

Para testar as hipóteses de interesse, temos que neste caso

$$\mathbf{D} = [\frac{\partial g_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_j}] = \begin{pmatrix} \mathbb{I}_{(p)} & \mathbf{0} \\ \mathbf{0} & \mathbb{I}_{(p)} \end{pmatrix},$$

é uma matriz $(2p-2)\times(2p)$ e $\widehat{\boldsymbol{\theta}}_* = g(\widehat{\boldsymbol{\theta}})$ é $AN(g(\boldsymbol{\theta}); \mathbf{D}J_{F1}^{-1}(\boldsymbol{\theta})\mathbf{D}'/n)$.

Assim, para testar $H_{01} : \alpha_2 = \ldots = \alpha_p = 0$, $\phi_2 = \ldots = \phi_p$, temos que $W_{01} = 185.4772$, com 9 graus de liberdade, para testar $H_{02} : \alpha_2 = \ldots = \alpha_p = 0$, temos que $W_{02} = 149.2166$, com 5 graus de liberdade e para testar $H_{03} : \phi_2 = \ldots = \phi_p$, temos que $W_{03} = 36.2605$, com 4 graus de liberdade, significa que mediante uso da estatística de Wald e considerando $\phi_1 = 0$, todas as hipótese também são rejeitadas a um nível de 5%.

2.4.5 Um Estudo de Simulação

Nesta seção apresentamos um pequeno estudo de simulação, no qual os níveis de significância e o poder dos três testes (Wald, Escore e Razão de Verossimilhanças) para a hipótese H_{01} de igualdade de variâncias e vícios são considerados.

Para cada estimação Monte Carlo do nível de significância, 1000 amostras independentes foram geradas de acordo com o modelo (2.1-2.2), para $\alpha_2 = \ldots = \alpha_p = 0$; $\phi_1 = \phi_2 = \ldots = \phi_p = 1$, com $p = 3$, $p = 5$ e tamanhos de amostra n=25, 50, 100. As estatísticas de Wald (W), Escore (E) e Razão de Verossimilhanças (Q) foram calculadas de acordo com (2.46), (2.58) e (2.61), respectivamente. O poder dos três testes foram estimados sob o modelo (2.1-2.2), para $p = 3$ e tamanhos de amostra $n = 25, 50, 100, 200$, considerando $H_1 : \alpha_2 = 0, \alpha_3 = 0.5, \phi_1 = \phi_2 = 1, \phi_3 = 1.5$ como hipótese alternativa. Em ambos casos a característica de interesse x_j, $j = 1, \ldots, n$, foi gerada de acordo a uma distribuição normal $N(0,\, \phi_x)$, com $\phi_x = 0.25,\ 1$, acrescentando $\phi_x = 0.01$ para o estudo do poder dos três testes.

Na Tabela 2.3 apresentamos os níveis de significância estimados, em porcentagens, dos três testes estatísticos, correspondentes ao nível nominal de 5%. Podemos observar que o teste de Wald converge lentamente para o nível nominal a medida que n cresce com relação aos outros dois testes. O teste de Escore é ligeiramente consistente e converge rapidamente, ainda para valores pequenos de n, enquanto que os outros dois testes superestimam o nível nominal do 5%(principalmente o teste de Wald), esta tabela parece indicar nenhuma diferença notável nos resultados quando variamos valores de p e ϕ_x. Isto também foi notado em outros estudos de simulação e pode ser observado na Figura 2.1 que representa para $n = 25,\ 50,$ **e** 100, com $p = 5$

Tabela 2.3: Nível de significância para os testes de Wald (W), Escore (E) e Razão de Verossimilhanças (Q) com nível nominal de 5%.

	Tamanho de amostra	$p=3$			$p=5$		
x_j	n	W	E	Q	W	E	Q
	25	7.6	4.9	7.9	10.5	4.5	7.2
$N(0, 0.25)$	50	5.2	4.4	5.0	7.1	4.9	6.0
	100	6.6	5.4	6.3	7.1	4.9	6.3
	25	7.6	4.5	6.9	9.8	4.0	6.0
$N(0, 1.0)$	50	5.8	4.9	5.6	7.4	4.0	5.8
	100	5.8	4.4	5.3	6.3	4.4	5.3

comparação das freqüências acumuladas estimadas dos três testes estatísticos, para H_{01}, com as freqüências acumuladas teóricas das correspondentes distribuições limites das estatísticas de Wald, Escore e Razão de Verossimilhanças, isto é χ^2_8.

Tabela 2.4: Poder estimado dos três testes para a hipótese alternativa $H_1 : \alpha_2 = 0$, $\alpha_3 = 0.5$, $\phi_1 = \phi_2 = 1$, $\phi_3 = 1.5$, com $p = 3$.

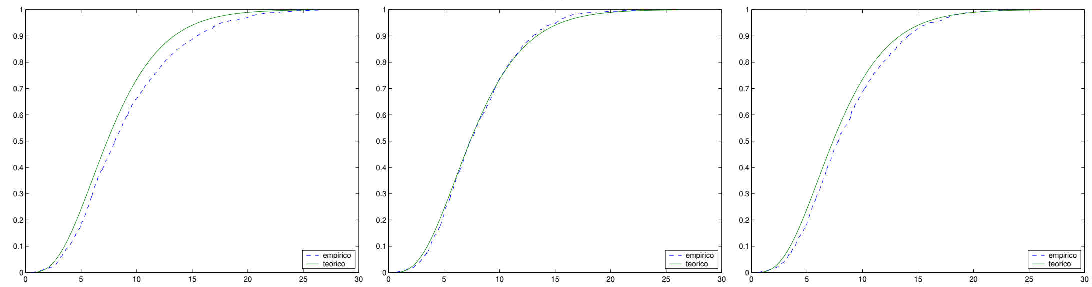
x_j	$N(0,\ 0.01)$			$N(0,\ 0.25)$			$N(0,\ 1.0)$		
n	W	E	Q	W	E	Q	W	E	Q
25	37.2	36.2	37.5	34.6	35.6	37.4	34.5	32.0	35.6
50	67.1	70.4	66.4	60.8	64.1	64.4	57.6	60.3	60.2
100	94.9	96.5	94.9	91.5	93.0	92.7	90.8	93.1	92.0
200	100	100	100	100	100	100	99.9	99.9	99.9

A Tabela 2.4 apresenta o poder estimado dos três testes para os valores dos parâmetros mostrados acima. O teste de Escore parece apresentar similar comportamento em termos de poder que o teste da Razão de Verossimilhanças, enquanto que o comportamento do teste de Wald é ligeiramente menor. Note que para n grande o comportamento dos três testes é otimizado consideravelmente. Nenhuma diferença relevante é notada quando ϕ_x cresce.

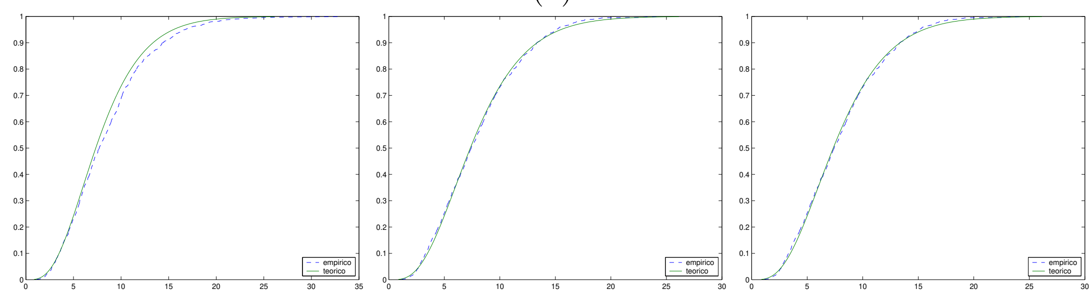
Assim, para o modelo de Grubbs estrutural normal recomendamos usar o teste de Escore, por apresentar um melhor comportamento em termos de nível de significância e é claro que sua distribuição assintótica é aproximada a uma distribuição qui-quadrado, ainda para pequenos valores de n, como a Figura 2.1 parece indicar.

Figura 2.1: Distribuição empírica (...) e teórica (–) das estatísticas de Wald, Escore e Razão de Verossimilhanças, respectivamente para testar H_{01}, onde a) $n = 25$, b) $n = 50$ e c) $n = 100$, com $p = 5$.

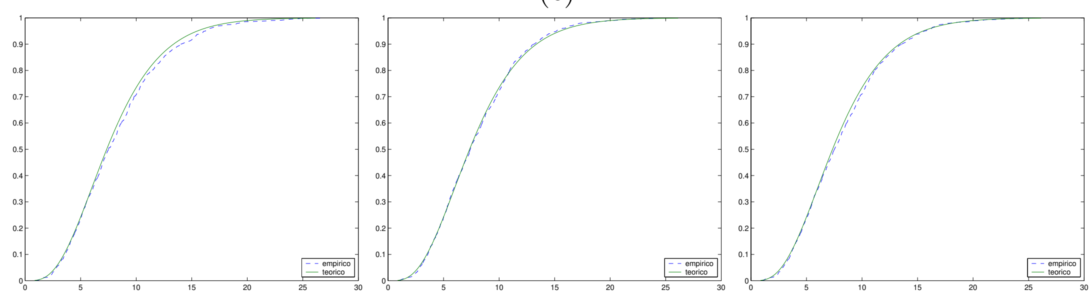
(a)



(b)



(c)



2.5 Estudo Assintótico dos Estimadores de Momentos

Considere a média amostral $\bar{\mathbf{Y}}$ e a matriz de covariância amostral $\mathbf{S}_n$ da seqüência de observações $\mathbf{Y}_1, \ldots, \mathbf{Y}_n$ definidos em (2.1), isto é

$$\bar{\mathbf{Y}} = \frac{1}{n}\sum_{j=1}^{n}\mathbf{Y}_j \quad \text{e} \quad \mathbf{S}_n = \frac{1}{n}\sum_{j=1}^{n}(\mathbf{Y}_j - \bar{\mathbf{Y}})(\mathbf{Y}_j - \bar{\mathbf{Y}})' . \tag{2.62}$$

Sob o modelo estrutural, Grubbs (1948), considera os seguintes estimadores tipo momentos para o vetor de parâmetros $\boldsymbol{\theta}$ dado por $\widetilde{\boldsymbol{\theta}} = (\widetilde{\mu}_x, \widetilde{\boldsymbol{\alpha}}', \widetilde{\phi}_x, \widetilde{\boldsymbol{\phi}}')' = (\widetilde{\mu}_x, \widetilde{\alpha}_2, ...\widetilde{\alpha}_p, \widetilde{\phi}_x, \widetilde{\phi}_1, ..., \widetilde{\phi}_p)'$, onde

$$\widetilde{\mu}_x = \bar{y}_{1.} = \frac{1}{n}\sum_{j=1}^{n} y_{1j} ,$$

$$\widetilde{\alpha}_i = \frac{1}{n}\sum_{j=1}^{n} y_{ij} - \widetilde{\mu}_x \ , \ \ 2, \ldots, p ,$$

$$\widetilde{\phi}_x = \frac{2}{p(p-1)}\sum_{i=1}^{p}\sum_{j>i}^{p} S_{ij} ,$$

e

$$\widetilde{\phi}_i = S_{ii} - \widetilde{\phi}_x , \tag{2.63}$$

para $i = 1, \ldots, p$, onde $S_{ij},\ i, j = 1, \ldots, p$, são os elementos da matriz de covariância amostral $\mathbf{S}_n$. Grubbs (1948) não faz um estudo do comportamento assintótico do estimador $\widetilde{\boldsymbol{\theta}}$, nem outros autores, por esta razão nesta parte do trabalho, vamos apresentar um estudo assintótico do estimador $\widetilde{\boldsymbol{\theta}}$. Note que as estimativas $\widetilde{\phi}_i$ e $\widetilde{\phi}_x$ têm o inconveniente de poder assumir valores negativos e por esta razão têm sido discutidas por alguns autores. Em particular, para $p = 2$ o estimador $\widetilde{\boldsymbol{\theta}}$ corresponde ao estimador de máxima verossimilhança de $\boldsymbol{\theta}$, se $\widetilde{\phi}_i > 0$ e $\widetilde{\phi}_x > 0$.

Para estudar o comportamento assintótico dos estimadores em (2.63), primeiro apresentaremos alguns resultados do comportamento assintótico de $\bar{\mathbf{Y}}_n$ e $\mathbf{S}_n$. As seguintes notações serão necessárias para uma explicação mais detalhada desta parte do trabalho. Sejam $\mathbf{A} = (a_{ij})$ e $\mathbf{B} = (b_{ij})$ matrizes de ordem $r \times s$ e $p \times q$, respectivamente. Então

i) $\mathbf{e}_i(j)$ denota o vetor de dimensão $i \times 1$ com um na j-ésima posição e zero nos outros casos;

ii) $\text{Vec}(\mathbf{A})$ denota o vetor de dimensão $rs \times 1$ dado por $\text{Vec}(\mathbf{A}) = (\mathbf{A}'_{1\bullet}, \ldots, \mathbf{A}'_{s\bullet})'$, onde $\mathbf{A}_{i\bullet}$ denota a i-ésima coluna de $\mathbf{A}$;

iii) $\mathbf{A} \otimes \mathbf{B}$ é uma matriz de ordem $rp \times sq$, que denota o produto Kroneker de $\mathbf{A}$ e $\mathbf{B}$, isto é, $\mathbf{A} \otimes \mathbf{B} = (a_{ij}\mathbf{B})$;

iv) Existe uma matriz $\mathbf{K}_{pq}$, chamada de matriz de permutação tal que (veja, Henderson e Searle, 1979)

$$\text{Vec}(\mathbf{A}') = \mathbf{K}_{p\,q}\text{Vec}(\mathbf{A}). \tag{2.64}$$

Quando $p = q$, escrevemos $\mathbf{K}_p$ em lugar de $\mathbf{K}_{pp}$. Além disso, se $\mathbf{A} = \mathbf{A}'$, então $\text{Vech}(\mathbf{A})$ contém os $p(p+1)/2$ elementos distintos de $\mathbf{A}$ e, $\text{Vec}(\mathbf{A})$ e $\text{Vech}(\mathbf{A}')$ são uma transformação linear um do outro. Assim existe uma matriz $\mathbf{H}$ de ordem $p(p+1)/2 \times p^2$ tal que

$$\text{Vech}(\mathbf{A}) = \mathbf{H}\,\text{Vec}(\mathbf{A}). \tag{2.65}$$

Agora, consideremos a seqüência $\mathbf{Y}_1, \ldots, \mathbf{Y}_n$ definida em (2.1) e $\boldsymbol{\epsilon}_j = \mathbf{Y}_j - \boldsymbol{\mu}$, $j = 1, \ldots, n$. No seguinte resultado apresentamos o comportamento assintótico de $\bar{\mathbf{Y}}_n$ e $\mathbf{S}_n$.

Teorema 2.4. *Sejam $\bar{\boldsymbol{Y}}_n$ e $\boldsymbol{S}_n$ a média e matriz de covariância amostral dos vetores $\mathbf{Y}_1, \ldots, \mathbf{Y}_n$ definidos em (2.1). Então*

i) $\bar{\mathbf{Y}} \xrightarrow{\text{q.c}} \boldsymbol{a} + \boldsymbol{1}_p\,\mu_x = \boldsymbol{\mu}$ *e* $\mathbf{S}_n \xrightarrow{\text{q.c}} \boldsymbol{\Sigma}$.

ii) $\sqrt{n}(\bar{\mathbf{Y}} - \mathbf{a} - \boldsymbol{1}_p\,\mu_x) \xrightarrow{\text{d}} N_p(\boldsymbol{0}, \boldsymbol{\Sigma})$ *e* $\sqrt{n}\,Vec(\mathbf{S}_n - \boldsymbol{\Sigma}) \xrightarrow{\text{d}} N_{p^2}(\mathbf{0}, \boldsymbol{\Delta})$, *onde* $\boldsymbol{\Delta} = Var(vec(\boldsymbol{\epsilon}\boldsymbol{\epsilon}')) = (\boldsymbol{I}_{p^2} + \mathbf{K}_p)(\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma})$.

iii) $\bar{\mathbf{Y}}$ *e* $\boldsymbol{S}_n$ *são assintoticamente independentes.*

A prova deste Teorema pode ser encontrada, por exemplo em Vilca-Labra et al.(1998).

Note que, como $\mathbf{S}_n$ é uma matriz simétrica, então o $\text{Vec}(\mathbf{S}_n)$ contém elementos que são iguais, de modo que a distribuição assintótica de $\sqrt{n}\text{Vec}(\mathbf{S}_n)$ é singular. Contudo, se consideramos somente os $p(p+1)/2$ elementos diferentes de $\mathbf{S}_n$, isto é, os elementos do vetor Vech$\mathbf{S}_n$, vamos ter que a distribuição assintótica de $\sqrt{n}\text{Vech}(\mathbf{S}_n)$ é não-singular.

O seguinte resultado segue como conseqüência imediata do Teorema 2.4 e das relações em (2.64) e (2.65).

Corolário 2.1. *Sob a condição do Teorema 2.4,*
$\sqrt{n}(\bar{\mathbf{Y}} - \mathbf{a} - \boldsymbol{1}_p\,\mu_x) \xrightarrow{\text{d}} N_p(\boldsymbol{0}, \boldsymbol{\Sigma})$ *e* $\sqrt{n}\,Vech(\mathbf{S}_n - \boldsymbol{\Sigma}) \xrightarrow{\text{d}} N_{p(p+1)/2}(\mathbf{0}, \boldsymbol{\nabla})$,
onde $\boldsymbol{\nabla} = \mathbf{H}(\boldsymbol{I}_{p^2} + \mathbf{K}_p)(\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma})\mathbf{H}'$.

Note que o estimador $\widetilde{\boldsymbol{\theta}}$ de $\boldsymbol{\theta}$, apresentado em (2.63) depende da média amostral $\bar{\mathbf{Y}}$ e da matriz de covariância amostral $\mathbf{S}_n$, isto é, $\widetilde{\boldsymbol{\theta}} = F(\bar{\mathbf{Y}}, \text{Vech}(\mathbf{S}_n))$, onde $\text{Vech}(\mathbf{S}_n) = (S_{11}, S_{21}, \ldots, S_{p\,1}, S_{22}, S_{32}, \ldots, S_{p\,2}, \ldots, S_{p-1\,p-1}, S_{p-1\,p}, S_{pp})'$ e $\bar{\mathbf{Y}} = (\bar{Y}_1, \ldots, \bar{Y}_p)'$.

A seguir apresentamos o comportamento assintótico do estimador $\widetilde{\boldsymbol{\theta}}$ de $\boldsymbol{\theta}$ para $p \geq 2$.

Teorema 2.5. *Seja o modelo de Grubbs estrutural normal definido em (2.1). Então* $\widetilde{\boldsymbol{\theta}}$ *é um estimador consistente de* $\boldsymbol{\theta}$ *e*

$$\sqrt{n}(\widetilde{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{\text{d}} N_{2p+1}(\mathbf{0}, \boldsymbol{\Psi}),$$

onde $\boldsymbol{\Psi} = \begin{pmatrix} \boldsymbol{A}'\boldsymbol{\Sigma}\boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{C}'\boldsymbol{\nabla}\boldsymbol{C} \end{pmatrix}$, $\boldsymbol{A} = \boldsymbol{I}_p - \boldsymbol{B}$ *e* $\boldsymbol{C} = \begin{pmatrix} \boldsymbol{c}_1 \\ \vdots \\ \boldsymbol{c}_p \end{pmatrix}_{\frac{p(p+1)}{2} \times (p+1)}$, *com*

$\boldsymbol{B} = \begin{pmatrix} \boldsymbol{0} & -\boldsymbol{1}_q' \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}$ *matriz* $p \times p$, $\boldsymbol{\nabla}$ *como no Corolário 2.1,*

$$\boldsymbol{c}_i = \begin{pmatrix} 0 & \boldsymbol{e}_p'(i) \\ \dfrac{2}{p(p-1)}\boldsymbol{1}_{p-i} & -\dfrac{2}{p(p-1)}\boldsymbol{1}_{p-i}\boldsymbol{1}_p' \end{pmatrix}$$

matriz $(p - i + 1) \times (p + 1)$ *para* $i = 1, \ldots, q$, $q = p - 1$ *e* $\boldsymbol{c}_p = \boldsymbol{e}_{p+1}'(p + 1)$; $e_i(j)$ *foi definido no início da seção.*

Demonstração. Seja $(\bar{\mathbf{Y}}', \text{Vech}'(\mathbf{S}_n))' = \Upsilon$ um vetor de dimensão $(p+p(p+1)/2) \times 1$. Como $\widetilde{\boldsymbol{\theta}} = F(\bar{\mathbf{Y}}, \text{Vech}(\mathbf{S}_n))$ é uma função diferenciável de $\bar{\mathbf{Y}}$ e $\text{Vech}(\mathbf{S}_n)$, então

$$\widetilde{\boldsymbol{\theta}} = F(\bar{\mathbf{Y}}, \text{Vech}(\mathbf{S}_n)) \xrightarrow{\text{q.c}} F(\boldsymbol{\mu}\,, \text{Vech}(\boldsymbol{\Sigma})) = \boldsymbol{\theta}.$$

Por outro lado, do Corolário 2.1 e Teorema 2.4, temos que

$$\sqrt{n}\begin{pmatrix} \bar{\mathbf{Y}} - \boldsymbol{\mu} \\ \text{Vech}(\mathbf{S}_n) - \text{Vech}(\boldsymbol{\Sigma}) \end{pmatrix} \xrightarrow{\text{d}} N_{p+p(p+1)/2}(\mathbf{0}, \boldsymbol{\nabla}_*), \tag{2.66}$$

onde $\boldsymbol{\nabla}_* = \text{Diag}(\boldsymbol{\Sigma}, \mathbf{H}(I_{p^2} + \mathbf{K}_p)(\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma})\mathbf{H}')$. Aplicando o método delta, temos que

$$\sqrt{n}(\widetilde{\boldsymbol{\theta}} - \boldsymbol{\theta}) = \sqrt{n}(F(\bar{\mathbf{Y}}, \text{Vech}(\mathbf{S}_n)) - F(\boldsymbol{\mu}\,, \text{Vech}(\boldsymbol{\Sigma}))) \xrightarrow{\text{d}} N_{2p+1}(0, \mathbf{G}'\boldsymbol{\nabla}_*\mathbf{G})\,,$$

onde $\mathbf{G} = \dfrac{\partial F_i(\boldsymbol{\Upsilon})}{\partial \Upsilon_j}$ $, i = 1, \ldots, 2p+1\ ,\ j = 1, \ldots, p + p(p+1)/2$ tem a forma

$$\mathbf{G} = \begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{C} \end{pmatrix},$$

onde $\mathbf{A} = \dfrac{\partial F_i(\boldsymbol{\Upsilon})}{\partial \Upsilon_j}$ $i, j = 1, \ldots, p$, $\mathbf{C} = \dfrac{\partial F_i(\boldsymbol{\Upsilon})}{\partial \Upsilon_j}$ $i = p+1, \ldots, 2p+1$, $j = p+1, \ldots, p + p(p+1)/2$. Após as manipulações algébricas obtém-se $\mathbf{A}$ e $\mathbf{C}$ como no enunciado do teorema, para mais detalhes sobre a obtenção das matrizes $\mathbf{K}_p$ e $\mathbf{H}$, veja por exemplo, Henderson e Searle (1979). $\square$

Observação : Note que para $p = 3$, as matrizes $\mathbf{A}$, $\mathbf{C}$, $\mathbf{H}$ e $\mathbf{K}_p$, no Teorema acima, são respectivamente dadas por

$$\mathbf{A} = \begin{pmatrix} 1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$\mathbf{C} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1/3 & -1/3 & -1/3 & -1/3 \\ 1/3 & -1/3 & -1/3 & -1/3 \\ 0 & 0 & 1 & 0 \\ 1/3 & -1/3 & -1/3 & -1/3 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

$$\mathbf{H} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 0 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$\mathbf{K}_3 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

identicamente, para $p = 4$, as matrizes $\mathbf{A}$, $\mathbf{C}$, $\mathbf{H}$ e $\mathbf{K}_p$ são respectivamente

$$\mathbf{A} = \begin{pmatrix} 1 & -1 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$
\mathbf{C} = \begin{pmatrix}
0 & 1 & 0 & 0 & 0 \\
1/6 & -1/6 & -1/6 & -1/6 & -1/6 \\
1/6 & -1/6 & -1/6 & -1/6 & -1/6 \\
1/6 & -1/6 & -1/6 & -1/6 & -1/6 \\
0 & 0 & 1 & 0 & 0 \\
1/6 & -1/6 & -1/6 & -1/6 & -1/6 \\
1/6 & -1/6 & -1/6 & -1/6 & -1/6 \\
0 & 0 & 0 & 1 & 0 \\
1/6 & -1/6 & -1/6 & -1/6 & -1/6 \\
0 & 0 & 0 & 0 & 1
\end{pmatrix},
$$

$$
\mathbf{H} = \begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0.5 & 0 & 0 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0 & 0.5 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix},
$$

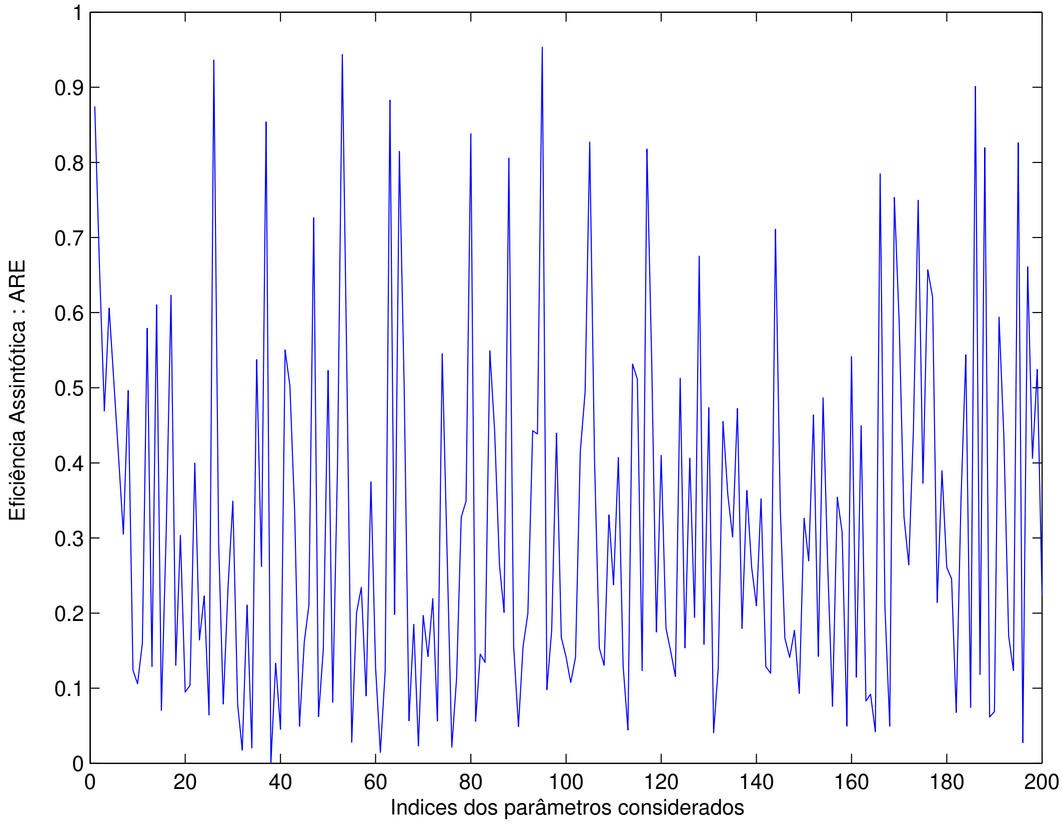
$$
\mathbf{K}_4 = \begin{pmatrix}
1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0\\
0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0\\
0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0\\
0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1
\end{pmatrix}.
$$

Na Figura 2.2, mostramos um estudo de comparação da eficiência assintótica relativa "ARE"dos EMV ($\widehat{\boldsymbol{\theta}}$) e os estimadores de momentos ($\widetilde{\boldsymbol{\theta}}$), definido como o quociente dos determinantes das respectivas variâncias assintóticas, isto é

$$
ARE(\widetilde{\boldsymbol{\theta}}, \widehat{\boldsymbol{\theta}}) = \frac{|J_{F1}^{-1}(\boldsymbol{\theta})|}{|\boldsymbol{\Psi}(\boldsymbol{\theta})|}\,. \tag{2.67}
$$

Foram gerados 200 valores arbitrários para o vetor de parâmetros $\boldsymbol{\theta}$, com $p = 3$ e variâncias positivas. Para cada valor do vetor de parâmetros gerado ($\boldsymbol{\theta}_{7\times 1}$), avaliamos na relação em (2.67). Note nesta figura, que os valores da ARE são menores que um (< 1), indicando que o EMV é assintoticamente mais eficiente que o estimador de momentos propostos por Grubbs. Resultados similares são obtidos quando tomamos os traços das respectivas variâncias assintóticas.

Figura 2.2: Valores de ARE para distintos valores do parâmetro $\boldsymbol{\theta}$.



Capítulo 3

Diagnóstico no Modelo de Grubbs

3.1 Introdução

Modelos estatísticos que são aplicados para analisar as características essenciais de um conjunto de dados são quase sempre descrições aproximadas dos mais complicados processos. Numa análise estatística é importante levar em consideração a estabilidade dos resultados obtidos e das inferências de um modo geral, com relação a possíveis perturbações nos dados ou no modelo estatístico.

Enquanto as análises de resíduos são feitas para investigar problemas com o modelo ajustado, uma análise de diagnóstico é feita assumindo o modelo como correto, e investigando a robustez das conclusões a pequenas perturbações dos dados.

Os elementos de um conjunto de dados que efetivamente controlam aspectos da análise são ditos influentes. Em particular, uma observação é dita influente se ela produzir alterações relevantes no resultado da análise, quando a mesma for excluída ou submetida a uma pequena perturbação. De fato, uma investigação completa de observações influentes é somente possível uma vez que elas tenham sido identificadas.

Na literatura não existem trabalhos desenvolvidos sobre influência e diagnóstico em modelos de Grubbs. Alguns trabalhos precursores em influência e diagnóstico são dados para modelos com erros nas variáveis e foram desenvolvidos por Kelly (1984) que propõe um procedimento de diagnóstico baseado em uma função de influência; Tanaka, Watadoni e Moon (1991) usam a função de influência de Hampel para desenvolver métodos de análise de influência de uma observação individual em estruturas de covariância. Abdullah (1995) apresenta vários métodos para detectar observações influentes no modelo de regressão funcional com erro nas variáveis. Mas essas medidas de diagnóstico avaliam o efeito dos parâmetros de regressão estimados depois de excluir uma única observação do conjunto de dados.

Mais do que eliminar casos, Cook (1986) propõe um método geral para avaliar o efeito, sob o estimador de máxima verossimilhança, de pequenas perturbações no modelo estatístico e/ou nos dados, baseado no afastamento pela verossimilhança. O método de influência local proposto por Cook (1986) tem sido aplicado em diversos modelos estatísticos. Recentemente, Kim (2000) e Galea et al. (2001) aplicam o método de influência local no modelo com erro nas variáveis simples (p=2). Além disso, Galea et al. (1999) aplicam o método de diagnóstico no modelo de calibração comparativa proposto por Barnett (1969).

Embora o método influência local proposto por Cook (1986) tenha sido aplicado com sucesso, alguns autores tem criticado o método. Recentemente, Poon e Poon (1999) propõem um método de influência local baseado na curvatura normal conformalizada, que é uma medida padronizada da curvatura normal utilizada por Cook (1986) e Loynes (2001) propõe uma nova medida de influência local que chama de "propagação de curvatura"(tradução livre de "*spread of curvature*") baseado num estudo probabilístico da curvatura normal.

Neste capítulo vamos desenvolver um estudo de diagnóstico no modelo de Grubbs, baseado no método de influência local proposto por Cook (1986) e a metodologia de Poon e Poon (1999). Vamos considerar dois esquemas de perturbação: ponderação de casos e perturbação na resposta. Os resultados serão aplicados a dados reais.

3.2 Influência Local

Nesta seção vamos apresentar em forma resumida o método de influência local proposto por Cook (1986); e alguns métodos alternativos que têm sido utilizados para avaliar a influência de determinada observação.

Para um conjunto de dados, seja $\ell(\boldsymbol{\theta}) = \sum_{i=1}^{n} \ell_i(\boldsymbol{\theta})$ a função de log-verossimilhança do modelo postulado, onde $\boldsymbol{\theta}$ é um vetor $p \times 1$ de parâmetros desconhecidos e $\ell_i(\boldsymbol{\theta})$ é a contribuição da i-ésima observação à log-verossimilhança. A perturbação no modelo é introduzida através de um vetor $\boldsymbol{\omega}$ de dimensão $q \times 1$, onde $\boldsymbol{\omega} \in \Omega \subseteq \mathbb{R}^q$, Ω um aberto. Seja $\ell(\boldsymbol{\theta}/\boldsymbol{\omega})$ a função de log-verossimilhança do modelo perturbado. Assumiremos que existe um $\boldsymbol{\omega}_o \in \Omega$ tal que $\ell(\boldsymbol{\theta}) = \ell(\boldsymbol{\theta}/\boldsymbol{\omega}_o)$, para todo $\boldsymbol{\theta}$ e que $\ell(\boldsymbol{\theta}/\boldsymbol{\omega})$ é duas vezes diferenciável em $(\boldsymbol{\theta}', \boldsymbol{\omega}')'$.

Sejam $\widehat{\boldsymbol{\theta}}$ e $\widehat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}$ os EMV de $\boldsymbol{\theta}$ sob o modelo postulado e perturbado, respectivamente. O objetivo é comparar $\widehat{\boldsymbol{\theta}}$ e $\widehat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}$ quando $\boldsymbol{\omega}$ varia em $\boldsymbol{\Omega}$. Cook (1986) sugere que a comparação entre $\widehat{\boldsymbol{\theta}}$ e $\widehat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}$ seja feita através do afastamento pela verossimilhança $LD(\boldsymbol{\omega})$, definida por

$$LD(\boldsymbol{\omega}) = 2[\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widehat{\boldsymbol{\theta}}_{\boldsymbol{\omega}})], \; \boldsymbol{\omega} \in \boldsymbol{\Omega}\,. \tag{3.1}$$

A função definida acima, pode ser vista como uma generalização do afastamento pela verossimilhança LD_i, considerado para avaliar a exclusão da i-ésima observação como proposto por Cook e Weisberg (1982), e definido por

$$LD_i(\boldsymbol{\theta}) = 2[\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widehat{\boldsymbol{\theta}}_{(i)})]\,, \tag{3.2}$$

onde $\widehat{\boldsymbol{\theta}}_{(i)}$ é o EMV de $\boldsymbol{\theta}$ obtido após se eliminar a i-ésima observação, $i = 1, \ldots, n$.

O sentido da distância entre $\widehat{\boldsymbol{\theta}}$ e $\widehat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}$, baseado na função $LD(\boldsymbol{\omega})$, pode depender da concavidade da função de log-verossimilhança $\ell(\boldsymbol{\theta})$. Se $\ell(\boldsymbol{\theta})$ é suficientemente achatada, pode-se dizer que $\widehat{\boldsymbol{\theta}}$ e $\widehat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}$ estão próximos entre si, enquanto que se $\ell(\boldsymbol{\theta})$ for suficientemente concentrada em torno de $\widehat{\boldsymbol{\theta}}$ estas estimativas podem estar distantes entre si (veja, Cook, 1987).

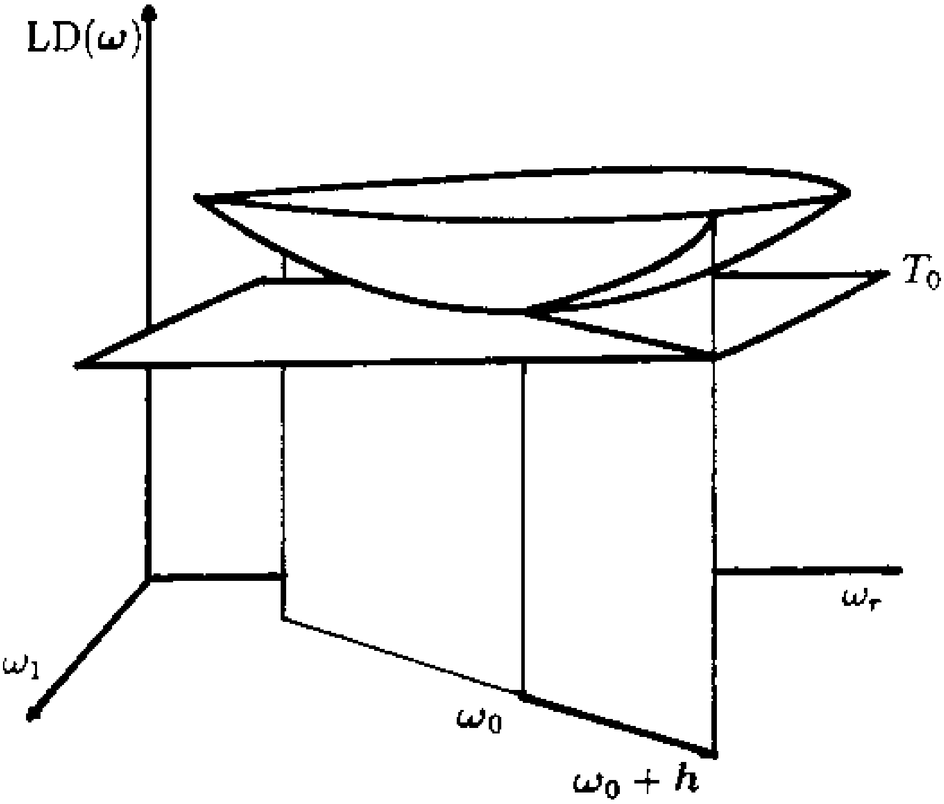
A idéia de Cook (1986) é investigar o comportamento da função $LD(\boldsymbol{\omega})$ numa vizinhança de $\boldsymbol{\omega}_o$. Para isso considerou o gráfico de $LD(\boldsymbol{\omega})$ versus $\boldsymbol{\omega}$ que pode ser visto como a superfície geométrica formada pelos valores do vetor $(q+1)$ dimensional

$$\boldsymbol{\alpha}(\boldsymbol{\omega}) = \begin{pmatrix} \boldsymbol{\omega} \\ LD(\boldsymbol{\omega}) \end{pmatrix},$$

onde $\boldsymbol{\omega}$ varia em $\boldsymbol{\Omega}$. A função $\boldsymbol{\alpha}(\boldsymbol{\omega})$ é chamada de gráfico de influência local, esta é uma superfície em $\mathbb{R}^{q+1}$ e pode ser usada para avaliar a influência ao variar $\boldsymbol{\omega}$

através de $\mathbf{\Omega}$. Uma ilustração gráfica é dada na Figura 3.1.

Figura 3.1: Representação gráfica do enfoque de influência local.



Aqui gostaríamos ter um gráfico da influência completa (isto é, gráfico de $\boldsymbol{\alpha}(\boldsymbol{\omega})$ variando $\boldsymbol{\omega}$) para avaliar a influência de um modelo particular e um conjunto particular de dados, no entanto, isto somente é possível nos casos onde $q \leq 2$. O método de influência local utiliza a curvatura normal de $\boldsymbol{\alpha}(\boldsymbol{\omega})$ em $\boldsymbol{\omega}_o$ que é um ponto de mínimo local da função $LD(\boldsymbol{\omega})$.

A idéia principal do método de Cook (1986) baseado na superfície $\boldsymbol{\alpha}(\boldsymbol{\omega})$ em torno de $\boldsymbol{\omega}_o$, consiste em considerar o plano tangente ($\mathbf{T}_o$) à superfície $\boldsymbol{\alpha}(\boldsymbol{\omega})$ em $\boldsymbol{\omega}_o$. Como

$LD(\boldsymbol{\omega})$ atinge o mínimo em $\boldsymbol{\omega}_o$, temos que $\mathbf{T}_o$ é paralelo a $\boldsymbol{\Omega} \subseteq \mathbb{R}^q$. Cada vetor $\mathbf{h}$ em $\boldsymbol{\Omega}$, de norma um, determina um plano que contém $\mathbf{h}$ e que é ortogonal a $\mathbf{T}_o$. A intersecção, chamada de seção normal, deste plano com a superfície é chamada de linha levantada (tradução livre de "lifted line") e pode ser obtida ao considerar o gráfico de $LD(\boldsymbol{\omega}_o + a\mathbf{h})$ em função de a, onde $a \in \mathbb{R}$. A curvatura normal da linha levantada, denotada por $C_{\mathbf{h}}$, é agora definida como a curvatura da curva plana $(a, LD(\boldsymbol{\omega}_o + a\mathbf{h}))$ em $a = 0$ e pode ser visualizado como a inversa do raio do círculo que melhor se ajusta em $\boldsymbol{\omega}_o$. Valores "grandes"de $C_{\mathbf{h}}$ indicam sensibilidade em relação à perturbação considerada na direção $\mathbf{h}$.

Cook (1986) mostra que a curvatura normal na direção $\mathbf{h}$ pode ser expressa na seguinte forma:

$$C_{\mathbf{h}} = 2|\mathbf{h}'\,\ddot{\mathbf{F}}\,\mathbf{h}|\,, \tag{3.3}$$

onde $\ddot{\mathbf{F}} = \boldsymbol{\Delta}'(\ddot{\mathbf{L}})^{-1}\boldsymbol{\Delta}$, com $\boldsymbol{\Delta}$ sendo uma matriz $p \times q$ com elementos $\boldsymbol{\Delta}_{ij} = \dfrac{\partial^2 \ell(\boldsymbol{\theta}/\boldsymbol{\omega})}{\partial\theta_i \partial\omega_j}$, $i = 1, \ldots, p$, $j = 1, \ldots, q$ e $-\ddot{\mathbf{L}}$ é a matriz de informação observada para o modelo postulado avaliada no ponto $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ e $\boldsymbol{\omega} = \boldsymbol{\omega}_o$.

Existem muitas maneiras na qual (3.3) pode ser usada para estudar $\boldsymbol{\alpha}(\boldsymbol{\omega})$, cada uma correspondendo a uma específica escolha do vetor unitário $\mathbf{h}$. Seja $C_{max} = \max_{\mathbf{h}:\, \|\mathbf{h}\|=1} C_h$ que corresponde ao maior autovalor de $-2\mathbf{F}$ (veja, Seber, 1984; pag. 526) e seja $\mathbf{h}_{max}$ o autovetor associado ao maior autovalor C_{max} . Neste caso $\mathbf{h}_{max}$ é a direção que produz a maior mudança local nas estimativas dos parâmetros, sendo que os elementos mais influentes podem ser identificados através do "tamanho"relativo dos elementos em $\mathbf{h}_{max}$, enquanto que a sensibilidade em relação a perturbação causada pode ser indicada pelo valor de C_{max}.

O afastamento pela verossimilhança descrito acima é uma medida de influência da perturbação $\boldsymbol{\omega}$ na estimação de máxima verossimilhança do vetor paramétrico completo $\boldsymbol{\theta}$. Quando o interesse for um subconjunto $\boldsymbol{\theta}_1$ de $\boldsymbol{\theta}' = (\boldsymbol{\theta}_1', \boldsymbol{\theta}_2')'$, o equivalente a (3.1) é dado por

$$LD_s(\boldsymbol{\omega}) = 2[\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widehat{\boldsymbol{\theta}}_{1\boldsymbol{\omega}},\, g(\widehat{\boldsymbol{\theta}}_{1\boldsymbol{\omega}}))]\,, \tag{3.4}$$

onde g é a função que maximiza $\ell(\boldsymbol{\theta}_1,\, \boldsymbol{\theta}_2)$ para cada $\boldsymbol{\theta}_1$ fixo de $\widehat{\boldsymbol{\theta}}_{1\boldsymbol{\omega}}$ e é determinada pela partição de $\widehat{\boldsymbol{\theta}}'_{\boldsymbol{\omega}} = (\widehat{\boldsymbol{\theta}}'_{1\boldsymbol{\omega}}, \widehat{\boldsymbol{\theta}}'_{2\boldsymbol{\omega}})'$. Neste caso, a curvatura normal na direção $\mathbf{h}$ é dada por (Cook, 1986)

$$C_{\mathbf{h}}(\boldsymbol{\theta}_1) = 2[\mathbf{h}'\boldsymbol{\Delta}'(\ddot{\mathbf{L}}^{-1} - \mathbf{B}_{22})\boldsymbol{\Delta}\,\mathbf{h}], \tag{3.5}$$

onde $\|\mathbf{h}\| = 1$, $\mathbf{B}_{22} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddot{\mathbf{L}}_{22}^{-1} \end{pmatrix}$ e $\ddot{\mathbf{L}}_{22}$ é determinada pela partição de $\ddot{\mathbf{L}} = \begin{pmatrix} \ddot{\mathbf{L}}_{11} & \ddot{\mathbf{L}}_{12} \\ \ddot{\mathbf{L}}_{21} & \ddot{\mathbf{L}}_{22} \end{pmatrix}$.

Segundo Seber (1984, pag. 526), para qualquer vetor $\mathbf{v}$

$$0 \geq \mathbf{v}' \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddot{\mathbf{L}}_{22}^{-1} \end{pmatrix} \mathbf{v} \geq \mathbf{v}'\ddot{\mathbf{L}}^{-1}\mathbf{v}$$

e, portanto,

$$\begin{aligned} C_{\mathbf{h}}(\boldsymbol{\theta}_1) &= -2\mathbf{h}'\boldsymbol{\Delta}'\ddot{\mathbf{L}}^{-1}\boldsymbol{\Delta}\mathbf{h} + 2\mathbf{h}'\boldsymbol{\Delta}' \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddot{\mathbf{L}}_{22}^{-1} \end{pmatrix} \boldsymbol{\Delta}\mathbf{h} \\ &= C_{\mathbf{h}} + 2\mathbf{h}'\boldsymbol{\Delta}' \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddot{\mathbf{L}}_{22}^{-1} \end{pmatrix} \boldsymbol{\Delta}\mathbf{h} \\ &\leq C_{\mathbf{h}}, \end{aligned} \tag{3.6}$$

isto significa que a curvatura normal para $\boldsymbol{\theta}_1$ na direção $\mathbf{h}$, nunca pode ser maior que a curvatura normal para $\boldsymbol{\theta}$ na mesma direção.

Note também que se $\ddot{\mathbf{L}}_{12} = 0$, então segue de (3.5) que

$$\begin{aligned} C_{\mathbf{h}} &= -2\mathbf{h}'\boldsymbol{\Delta}' \begin{pmatrix} \ddot{\mathbf{L}}_{11}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \boldsymbol{\Delta}\mathbf{h} - 2\mathbf{h}'\boldsymbol{\Delta}' \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddot{\mathbf{L}}_{22}^{-1} \end{pmatrix} \boldsymbol{\Delta}\mathbf{h} \\ &= C_{\mathbf{h}}(\boldsymbol{\theta}_1) + C_{\mathbf{h}}(\boldsymbol{\theta}_2). \end{aligned} \tag{3.7}$$

Portanto, temos que para qualquer direção $\mathbf{h}$, a curvatura normal para $\boldsymbol{\theta}$ nesta direção é então a soma das curvaturas normais para $\boldsymbol{\theta}_1$ e $\boldsymbol{\theta}_2$ na mesma direção. Expressando o fato de que a influência para $\boldsymbol{\theta}_1$ é independente da influência para $\boldsymbol{\theta}_2$.

Uma outra escolha evidente corresponde à perturbação na direção do i-ésimo caso que é obtida tomando $\mathbf{h}$ igual ao vetor $\mathbf{h}_i$ com um na i-ésima posição e zero nas outras posições, dando como resultado a seguinte medida de influência local

$$C_i = C_{h_i} = 2|\mathbf{\Delta}_i \, \ddot{\mathbf{L}}^{-1} \, \mathbf{\Delta}_i|, \; i = 1, \ldots, q\,, \tag{3.8}$$

Verbeke e Molenberghs (1997), propõe considerar a i-observação como influente se C_i é maior que o valor $2\sum_{i=1}^{q} C_i/q$. Note que C_i é de interesse particular desde que representa a influência local de cada observação separadamente na estimação de $\boldsymbol{\theta}$.

Finalmente, existem muitas formas possíveis de escolher o vetor $\mathbf{h}$. Por exemplo, a influência local na estimação de $\boldsymbol{\theta}_1$ na direção do i-ésimo caso, denotado por $\mathbf{h}_i$, o qual é um vetor com um na posição i e zero nas outras posições . A correspondente curvatura normal será denotada por $C_i(\boldsymbol{\theta}_1)$.

Embora, o método de influência local proposto por Cook (1986) tem sido aplicado com sucesso em muitos problemas estatísticos, alguns aspectos têm sido questionados por alguns autores (veja, discussão em Cook, 1986), por exemplo, a falta de um critério objetivo para avaliar a grandeza da curvatura normal e tamanho relativo das componentes da direção correspondente à maior curvatura normal.

Recentemente, Poon e Poon (1999) apresentaram um método de influência local similar ao método proposto por Cook (1986), que no lugar de usar a curvatura normal, recomenda o uso da curvatura normal conformalizada, sendo esta invariante sob reparametrização e assume valores entre 0 e 1.

Segundo Poon e Poon (1999), a curvatura normal conformalizada em $\boldsymbol{\omega}_o$ do gráfico $\boldsymbol{\alpha}(\boldsymbol{\omega})$ na direção $\mathbf{h}$ é definida por

$$B_{\mathbf{h}} = -\frac{\mathbf{h}'\mathbf{\Delta}'(\ddot{\mathbf{L}})^{-1}\mathbf{\Delta}\mathbf{h}}{\sqrt{tr(\mathbf{\Delta}'(\ddot{\mathbf{L}})^{-1}\mathbf{\Delta})^2}} \tag{3.9}$$

avaliadas em $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ e $\boldsymbol{\omega} = \boldsymbol{\omega}_o$. De (3.3) e (3.9) podemos observar que o trabalho para obter $C_{\mathbf{h}}$ e $B_{\mathbf{h}}$ são os mesmos, mas $B_{\mathbf{h}}$ tem várias propriedades importantes, que serão apresentadas mais adiante.

Seja $\mathbf{\Omega}$ o conjunto de vetores de perturbação definido em (3.1) e seja $\mathbf{\Omega}^*$ outro conjunto de vetores de perturbação com a mesma dimensão de $\mathbf{\Omega}$. Uma reparametrização é uma função diferenciável $\phi : \mathbf{\Omega} \rightarrow \mathbf{\Omega}^*$ tal que o jacobiano da transformação é não-singular, isto é, uma reparametrização é uma modificação do esquema de perturbação. Diremos que uma reparametrização é conformalizada em $\boldsymbol{\omega}_o$ se seu jacobiano que denotamos por $\mathbf{J}$ é tal que $\mathbf{J}\mathbf{J}' = \tau\mathbf{I}_q$, para algum número positivo τ.

As propriedades mais importantes de $B_{\mathbf{h}}$ são:

1. Se uma reparametrização de $\mathbf{\Omega}$ é normal conformalizada em $\boldsymbol{\omega}_o$, então a curvatura normal conformalizada em $\boldsymbol{\omega}_o$ é invariante sob reparametrização, isto é

   $$B_{\phi(\mathbf{h})} = B_{\mathbf{h}}\,,$$

   para qualquer direção h.

2. Para qualquer direção $\mathbf{h}$, $B_{\mathbf{h}}$ satisfaz a relação

   $$0 \leq |B_{\mathbf{h}}| \leq 1\,,$$

   ou seja, $B_{\mathbf{h}}$ é uma medida normalizada e é fácil de interpretar sua magnitude.

As provas dos resultados 1) e 2) podem ser encontradas em Poon e Poon (1999).

Por Poon e Poon (1999), a curvatura normal conformalizada total em $\boldsymbol{\omega}_0$ da superfície $\boldsymbol{\alpha}(\boldsymbol{\omega})$ na direção $\mathbf{h}_i$ pode ser expressa como

$$B_{\mathbf{h}_i} = -\frac{\mathbf{\Delta}_i'(\ddot{\mathbf{L}})^{-1}\mathbf{\Delta}_i}{\sqrt{tr(\mathbf{\Delta}'(\ddot{\mathbf{L}})^{-1}\mathbf{\Delta})^2}}, \tag{3.10}$$

onde $\mathbf{h}_i$ e $\mathbf{\Delta}_i$ são como em (3.8), com $i = 1, \ldots, q$. Se $\lambda_1 \geq \ldots \geq \lambda_k > 0$ denota os k autovalores diferentes de zero de

$$-\frac{\mathbf{\Delta}'(\ddot{\mathbf{L}})^{-1}\mathbf{\Delta}}{\sqrt{tr(\mathbf{\Delta}'(\ddot{\mathbf{L}})^{-1}\mathbf{\Delta})^2}} \tag{3.11}$$

e se $\mathbf{a}_1, \ldots, \mathbf{a}_k$ são os correspondentes autovetores, é fácil verificar que

$$B_{\mathbf{h}_i} = \sum_{j=1}^{k} \lambda_j a_{ji}^2, \tag{3.12}$$

onde a_{ji} é o i-ésimo componente do vetor $\mathbf{a}_j$, $i = 1, \ldots, q$. Isto mostra que casos individuais podem ter uma grande curvatura local total $B_{\mathbf{h}_i}$ sem ter uma grande componente $|a_{1i}|$ na direção $\mathbf{h}_{max}$. Por isso, a metodologia de Poon e Poon (1999) é usada para avaliar a influência local. Por Poon e Poon (1999), a contribuição conjunta da j-ésima base de vetores de perturbações para todo autovetor r-influente (o autovetor $\mathbf{b}$ é r-influente, se $|B_{\mathbf{b}}| \geq r/\sqrt{q}$) é definido como

$$m(r)_j = \sum_{i=1}^{s} \lambda_i a_{ij}^2, \tag{3.13}$$

onde s é o número de autovalores maiores a $r/\sqrt{q}$, $j = 1, \ldots, q$ e uma adequada marca de nível para julgar a grandeza da influência local é $2\bar{m}(r)$ para $r > 0$ e $2b$ para $r = 0$, onde $\bar{m}(r) = \sum_{i=1}^{s} \lambda_i/q$ e $b = tr(\mathbf{A})/(q\sqrt{tr(\mathbf{A}^2)})$, com $\mathbf{A} = -\boldsymbol{\Delta}'(\ddot{\mathbf{L}})^{-1}\boldsymbol{\Delta}$.

Dos argumentos precedentes, encontramos que se $s = 1$(ou r suficientemente grande) as observações influentes identificadas são as mesmas que as encontradas por $\mathbf{h}_{max}$. Por outro lado $r = 0$ identifica as mesmas observações que a medida de influência local C_i, isto é, os gráficos resultantes são proporcionais (veja, ilustração numérica) com a diferença que esta nova metodologia é mas objetiva, assim Poon e Poon (1999) propõe um método de diagnóstico unificado para avaliar a influência local de pequenas perturbações no conjunto de dados.

Um dos métodos de diagnóstico empregados em modelos de regressão utiliza a exclusão de casos que consiste em comparar os EMV $\widehat{\boldsymbol{\theta}}$ e $\widehat{\boldsymbol{\theta}}_{(i)}$. Estes métodos são utilizados com freqüência para diagnosticar globalmente possíveis observações influentes e podem ser facilmente adaptados ao modelo de Grubbs, dado que podemos calcular os EMV com e sem a i-ésima observação.

Neste trabalho, utilizaremos duas medidas de diagnóstico de influência global, uma chamada de LD_i definida em (3.2), e outra chamada de Distância de Cook Generalizada, definida como

$$D_i = (\widehat{\boldsymbol{\theta}} - \widehat{\boldsymbol{\theta}}_{(i)})'(-\ddot{\mathbf{L}}(\widehat{\boldsymbol{\theta}}))(\widehat{\boldsymbol{\theta}} - \widehat{\boldsymbol{\theta}}_{(i)}), \tag{3.14}$$

onde $-\ddot{\mathbf{L}}$ é a matriz de informação de Fisher observada e $\widehat{\boldsymbol{\theta}}_{(i)}$ como em (3.2). Comparações entre (3.8) e (3.14) mostram que $C_i/2$ e D_i são aproximadamente as mesmas, para n suficientemente grande. Assim, a medida de influência local C_i pode ser interpretada como uma aproximação de medida da influência global. Para mais detalhes sob esta medida, veja por exemplo, Verbeke e Molenberghs (1997) e Bo-Cheng Wei (1998, pag. 103).

Outro método de diagnóstico que é utilizado na literatura de análise multivariada, e que pode ser adaptado para detectar *outliers* no modelo de Grubbs, é a distância quadrada generalizada d_j^2, definida por

$$d_j^2 = (\mathbf{Y}_j - \bar{\mathbf{Y}})'\mathbf{S}_n^{-1}(\mathbf{Y}_j - \bar{\mathbf{Y}}),\ j = 1, \ldots, n, \tag{3.15}$$

onde $\mathbf{S}_n$ é como em (2.58). Um valor grande de d_j^2, poderia indicar que a observação j é um *outlier*.

Diversas aplicações e resultados adicionais sobre o método de influência local podem ser encontrados em Lawrence (1988); Tomas e Cook (1989, 1990); Paula (1993, 1995, 1996); Galea, Paula e Bolfarine (1997); Kim (2000), entre outros.

Thomas e Cook (1990) empregam esse método para fazer diagnóstico de influência sobre predições nos modelos lineares generalizados, utilizando uma função objetiva baseada em diferença de predições. Paula (1996) faz uma extensão desse estudo para os modelos próprios de dispersão de Jorgensen (1997).

Fung e Kwan (1997) apresentam uma discussão interessante sobre a aplicação de influência local para outras medidas de influência, digamos $\widehat{T}_\omega$, é invariante de escala se $\ddot{\lambda} = \left.\dfrac{\partial \widehat{T}_{\boldsymbol{\omega}}}{\partial \boldsymbol{\omega}}\right|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0} = 0$, quando esta derivada é diferente de zero a ordem entre os componentes de $\mathbf{h}_{max}$ não é necessariamente preservada sob mudanças de escala. Em particular, para o afastamento pela verossimilhança, tem-se que $\ddot{\lambda} = \left.\dfrac{\partial LD(\omega)}{\partial \boldsymbol{\omega}}\right|_{\boldsymbol{\omega}=\boldsymbol{\omega}_o} = 0$. Esta propriedade segue também, por exemplo, para as medidas de influência propostos por Thomas e Cook (1990) e Paula (1993). Porém ela não vale para outras medidas de influência discutida por Fung e Kwan (1997).

Cadigan e Farrel (1999) sugerem o uso do valor esperado da influência local como uma ferramenta útil na interpretação da análise de diagnóstico e Loynes (2001) propõe uma nova medida de influência local, chamada de propagação de curvatura, que tem a propriedade de ser invariante sob reparametrização das perturbações.

3.3 Aplicação no Modelo de Grubbs

Nesta seção, vamos desenvolver o método de influência local para o modelo de Grubbs estrutural normal, considerando dois tipos de perturbações: ponderação de casos e perturbação na variável resposta.

3.3.1 Perturbação de Ponderação de Casos

Considere o modelo definido em (2.1-2.2) e o vetor de pesos $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_n)'$. A função de verossimilhança do modelo perturbado, considerando ponderação de casos, é dada por

$$\ell(\boldsymbol{\theta}/\boldsymbol{\omega}) = \sum_{j=1}^{n} \omega_j \ell_j(\boldsymbol{\theta}), \tag{3.16}$$

onde $\ell_j(\boldsymbol{\theta})$, $j = 1, \ldots, n$, foi definido em (2.5) e $\boldsymbol{\theta} = (\mu_x, \alpha_2, \ldots, \alpha_p, \phi_x, \phi_1, \ldots, \phi_p)'$. Note que, neste caso o modelo não perturbado é obtido ao considerar $\boldsymbol{\omega}_o = \mathbf{1}_n$.

Derivando (3.16) com respeito a $\boldsymbol{\omega}'$ temos que

$$\frac{\partial \ell(\boldsymbol{\theta}/\boldsymbol{\omega})}{\partial \boldsymbol{\omega}'} = (\ell_1(\boldsymbol{\theta}), \ldots, \ell_n(\boldsymbol{\theta})),$$

logo, derivando com respeito a $\boldsymbol{\theta}$ a expressão anterior, temos que

$$\begin{aligned} \boldsymbol{\Delta} &= \frac{\partial^2 \ell(\boldsymbol{\theta}/\boldsymbol{\omega})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\omega}'} = \frac{\partial}{\partial \boldsymbol{\theta}} \left(\frac{\partial \ell(\boldsymbol{\theta}/\boldsymbol{\omega})}{\partial \boldsymbol{\omega}'} \right) = \frac{\partial}{\partial \boldsymbol{\theta}} (\ell_1(\boldsymbol{\theta}), \ldots, \ell_n(\boldsymbol{\theta})) \\ &= \left(\frac{\partial \ell_1(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}, \ldots, \frac{\partial \ell_n(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) = (\boldsymbol{\Delta}_1, \ldots, \boldsymbol{\Delta}_n), \end{aligned} \tag{3.17}$$

onde $\boldsymbol{\Delta}_j = \dfrac{\partial \ell_j(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$, $j = 1, \ldots, n$, com $\dfrac{\partial \ell_j(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ obtida na Seção 2.2.1 dada por

$$\frac{\partial \ell_j}{\partial \boldsymbol{\theta}} = \begin{pmatrix} \mathbf{1}_p' \, \boldsymbol{\Sigma}^{-1} \mathbf{W}_j \\ \mathbb{I}_{(p)} \boldsymbol{\Sigma}^{-1} \mathbf{W}_j \\ \dfrac{1}{2\phi_x} c^{-1}(1-c) + \dfrac{1}{2}\dfrac{c^{-2}}{\phi_x} \mathbf{W}_j' \mathbf{M} \mathbf{W}_j \\ \begin{array}{l} \dfrac{1}{2} c^{-1} \phi_x \mathrm{D}^{-2}(\boldsymbol{\phi}) \mathbf{1}_p - \dfrac{1}{2}\mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{1}_p + \dfrac{1}{2}\mathrm{D}(\mathbf{W}_j)\mathrm{D}^{-2}(\boldsymbol{\phi})\mathbf{W}_j \\ + \dfrac{c^{-2}}{2} \phi_x \mathbf{W}_j' \mathbf{M} \mathbf{W}_j \mathrm{D}^{-2}(\boldsymbol{\phi}) \mathbf{1}_p - c^{-1}\phi_x \mathbf{W}_j' \mathrm{D}^{-1}(\boldsymbol{\phi}) \mathbf{1}_p \, \mathrm{D}^{-2}(\boldsymbol{\phi}) \mathbf{W}_j \end{array} \end{pmatrix}, \tag{3.18}$$

avaliadas em $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}} = (\widehat{\mu}_x, \widehat{\alpha}_2, \ldots, \widehat{\alpha}_p, \widehat{\phi}_x, \widehat{\phi}_1, \ldots, \widehat{\phi}_p)'$.

Observe que, para o esquema de perturbação de casos, a matriz $\boldsymbol{\Delta}$ não depende do vetor $\boldsymbol{\omega}$. O vetor $\mathbf{h}_{max}$ é o autovetor normalizado correspondente ao maior autovalor da matriz $-2\boldsymbol{\Delta}'\ddot{\mathbf{L}}^{-1}\boldsymbol{\Delta}$ que contém informação sobre quais são as observações que exercem maior influência sobre as estimativas dos parâmetros.

3.3.2 Perturbação na Variável Resposta

Vamos considerar aqui, a seguinte perturbação na variável resposta

$$\mathbf{Y}_{\omega_j} = \mathbf{Y}_j + \mathbf{S}_y \, \omega_j \,, \;\; j = 1, \ldots, n, \tag{3.19}$$

onde $\mathbf{S}_y = (s_1, \ldots, s_p)'$ é um vetor $p \times 1$, com s_i, fator de escala correspondente ao i-ésimo instrumento, $i = 1, \ldots, p$ e ω_j um número real de pequenas perturbações. Neste caso $\boldsymbol{\omega}_o = (0, \ldots, 0)'_{n \times 1}$ e podemos utilizar, por exemplo, como fator de escala, o desvio padrão amostral correspondente ao i-ésimo instrumento.

A função de log-verossimilhança para o modelo perturbado $\ell(\boldsymbol{\theta}/\boldsymbol{\omega})$ é dado por

$$\ell(\boldsymbol{\theta}/\boldsymbol{\omega}) = \sum_{j=1}^{n} \ell_j(\boldsymbol{\theta}/\omega_j), \tag{3.20}$$

onde

$$\ell_j(\boldsymbol{\theta}/\omega_j) = \log(2\pi)^{-p/2} - \frac{1}{2}\log|\boldsymbol{\Sigma}| - \frac{1}{2}\|\mathbf{T}_{\omega_j}\|^2 \,,$$

e

$$\|\mathbf{T}_{\omega_j}\|^2 = (\mathbf{Y}_{\omega_j} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{Y}_{\omega_j} - \boldsymbol{\mu})\,,$$

com $\boldsymbol{\Sigma}$, $\boldsymbol{\mu}$ como em (2.3) e $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_n)'$. Derivando (3.20) com respeito a $\boldsymbol{\omega}'$, obtemos

$$\frac{\partial \ell(\boldsymbol{\theta}/\boldsymbol{\omega})}{\partial \boldsymbol{\omega}'} = (\mathbf{A}_1(\boldsymbol{\theta}, \omega_1), \ldots, \mathbf{A}_n(\boldsymbol{\theta}, \omega_n)), \tag{3.21}$$

onde

$$\mathbf{A}_j(\boldsymbol{\theta}, \omega_j) = -\frac{1}{2}\frac{\partial}{\partial \omega_j}\|\mathbf{T}_{\omega_j}\|^2 = -(\mathbf{Y}_{\boldsymbol{\omega}} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}\mathbf{S}_y,\ j = 1, \ldots, n.$$

Então derivando (3.21) com respeito a $\boldsymbol{\theta}$, temos que a matriz $\boldsymbol{\Delta}$ é dada por

$$\begin{aligned}
\boldsymbol{\Delta} &= \frac{\partial^2 \ell(\boldsymbol{\theta}/\boldsymbol{\omega})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\omega}'} = \frac{\partial}{\partial \boldsymbol{\theta}}(A_1(\boldsymbol{\theta}, \omega_1), \ldots, A_n(\boldsymbol{\theta}, \omega_n)) \\
&= \left(\frac{\partial}{\partial \boldsymbol{\theta}} A_1(\boldsymbol{\theta}, \omega_1), \ldots, \frac{\partial}{\partial \boldsymbol{\theta}} A_n(\boldsymbol{\theta}, \omega_n)\right) \\
&= \begin{pmatrix}
\frac{\partial}{\partial \mu_x} A_1(\boldsymbol{\theta}, \omega_1) & \ldots & \frac{\partial}{\partial \mu_x} A_n(\boldsymbol{\theta}, \omega_n) \\
\frac{\partial}{\partial \boldsymbol{\alpha}} A_1(\boldsymbol{\theta}, \omega_1) & \ldots & \frac{\partial}{\partial \boldsymbol{\alpha}} A_n(\boldsymbol{\theta}, \omega_n) \\
\frac{\partial}{\partial \phi_x} A_1(\boldsymbol{\theta}, \omega_1) & \ldots & \frac{\partial}{\partial \phi_x} A_n(\boldsymbol{\theta}, \omega_n) \\
\frac{\partial}{\partial \boldsymbol{\phi}} A_1(\boldsymbol{\theta}, \omega_1) & \ldots & \frac{\partial}{\partial \boldsymbol{\phi}} A_n(\boldsymbol{\theta}, \omega_n)
\end{pmatrix}.
\end{aligned} \tag{3.22}$$

Segue após as manipulações algébricas, que os elementos da matriz $\boldsymbol{\Delta}$ são dados por

$$\begin{aligned}
&\frac{\partial}{\partial \mu_x} A_j(\boldsymbol{\theta}, \omega_j) = \mathbf{1}_p'\boldsymbol{\Sigma}^{-1}\mathbf{S}_y, \\
&\frac{\partial}{\partial \boldsymbol{\alpha}} A_j(\boldsymbol{\theta}, \omega_j) = \mathbb{I}_{(p)}\boldsymbol{\Sigma}^{-1}\mathbf{S}_y, \\
&\frac{\partial}{\partial \phi_x} A_j(\boldsymbol{\theta}, \omega_j) = \frac{c^{-2}}{\phi_x}(\mathbf{Y}_{\omega_j} - \boldsymbol{\mu})'\mathbf{M}\mathbf{S}_y, \\
&\frac{\partial}{\partial \boldsymbol{\phi}} A_j(\boldsymbol{\theta}, \omega_j) = \mathrm{D}(\mathbf{Y}_{\omega_j} - \boldsymbol{\mu})\mathrm{D}^{-2}(\boldsymbol{\phi})\mathbf{S}_y + c^{-2}\phi_x\mathrm{D}^{-2}(\boldsymbol{\phi})\mathbf{1}_p(\mathbf{Y}_{\omega_j} - \boldsymbol{\mu})'\mathbf{M}\mathbf{S}_y \\
&\qquad - c^{-1}\phi_x\mathrm{D}^{-2}(\boldsymbol{\phi})(\mathbf{Y}_{\omega_j} - \boldsymbol{\mu})\mathbf{1}_p'\mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{S}_y \\
&\qquad - c^{-1}\phi_x(\mathbf{Y}_{\omega_j} - \boldsymbol{\mu})'\mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{1}_p\mathrm{D}^{-2}(\boldsymbol{\phi})\mathbf{S}_y,
\end{aligned}$$

para $j = 1, \ldots, n$, onde c, $\mathbf{M}$ são como em (2.4) e a matriz $\boldsymbol{\Delta}$ é avaliada em

$\boldsymbol{\omega} = \boldsymbol{\omega}_o = (0, \ldots, 0)'$ e $\widehat{\boldsymbol{\theta}} = (\widehat{\mu}_x, \widehat{\alpha}_2, \ldots, \widehat{\alpha}_p, \phi_x, \widehat{\phi}_1, \ldots, \widehat{\phi}_p)'$.

Segundo o método de Cook (1986), o vetor normalizado correspondente ao maior autovalor da matriz $-2\boldsymbol{\Delta}'\ddot{\mathbf{L}}^{-1}\boldsymbol{\Delta}$ fornece a direção de maior influência local. Este valor denotado por $\mathbf{h}_{max}$ contém informação sobre quais são as observações que exercem maior influência sobre as estimativas dos parâmetros.

Uma outra forma de perturbar a variável resposta é considerar

$$\mathbf{Y}_{\omega_j} = \mathbf{Y}_j + \mathrm{D}(\mathbf{S}_y)\,\boldsymbol{\omega}_j\,, \quad j = 1, \ldots, n, \tag{3.23}$$

onde $D(\mathbf{S}_y) = \mathrm{Diag}(s_1, \ldots, s_p)$ é uma matriz de ordem p, com s_i, fator de escala correspondente ao i-ésimo instrumento, $i = 1, \ldots, p$ e $\boldsymbol{\omega}_j$ um vetor $p \times 1$ de pequenas perturbações. Neste caso $\boldsymbol{\omega}_o = (0, \ldots, 0)'_{np\times 1}$ e podemos utilizar como fator de escala, o desvio padrão amostral correspondente ao i-ésimo instrumento. Neste caso a matriz $\boldsymbol{\Delta}$ é similar a encontrada em (3.22), com os mesmos elementos só que agora o vetor $\mathbf{S}_y$ é substituído em todos os casos pela matriz diagonal $\mathrm{D}(\mathbf{S}_y)$.

3.4 Ilustração Numérica

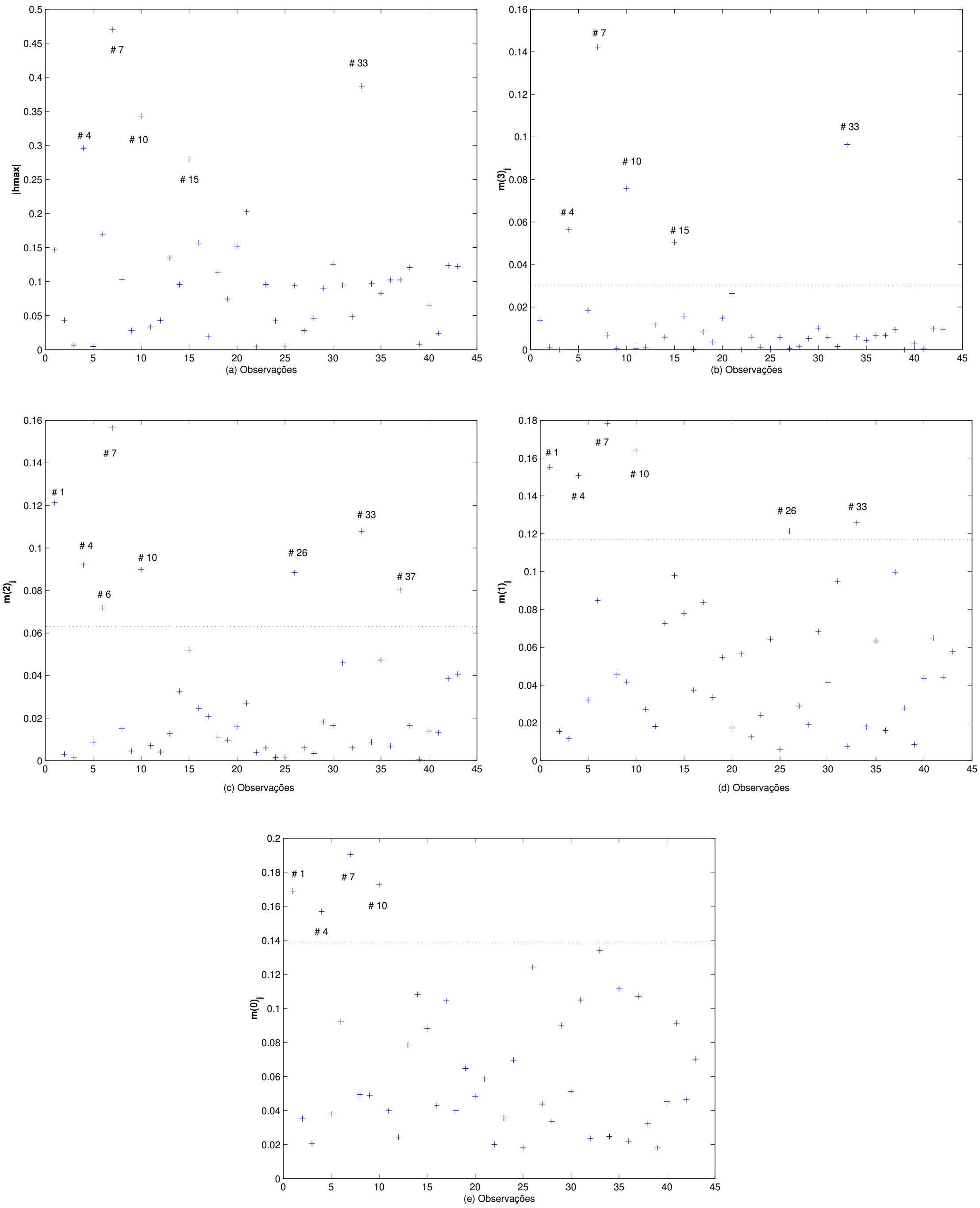
Nesta seção, dois exemplos serão usados para ilustrar as técnicas introduzidas nas seções anteriores. Será desenvolvida a análise de diagnóstico com ênfase especial no método de influência local proposto por Cook (1986) e Poon e Poon (1999), considerando os dois tipos de perturbação desenvolvida na seção prévia.

3.4.1 Exemplo 1

A primeira aplicação que faremos consiste em analisar o conjunto de dados apresentados na Tabela 2.1, tomados de Jaech (1985), relativos a um experimento, na qual a densidade de 43 pastilhas de combustível de urânio sintetizado para uso em reatores nucleares foram medidas por seis "instrumentos".

Para estes dados, examinemos primeiro a influência local de ponderação de casos baseados na análise descrito por Cook (1986) e Poon & Poon (1999), discutidos na Seção 3.2, os resultados são mostrados na Figura 3.2. A Figura 3.2(a) mostra

Figura 3.2: Dados de urânio sintetizado. Gráfico de influência local para perturbação de ponderação de casos.



a direção $|hmax|$ correspondente à maior curvatura normal $C_{\mathbf{h}}$ definido em (3.3). Note que neste gráfico não é fácil distinguir que observações são verdadeiramente influentes na estimação do vetor de parâmetros $\boldsymbol{\theta}$. As Figuras 3.2(b), (c), (d) e (e) são os *index plot* das contribuições conjuntas $m(r)_j$, com $r = 3, 2, 1, 0$, respectivamente. A Figura 3.2(b) indica que as observações 4, 7, 10, 15 e 33 são influentes na estimação do vetor de parâmetro completo $\boldsymbol{\theta}$, note também que as figuras 3.2(a) e (b) são proporcionais, o qual verifica o argumento dado na Seção 3.2. As marcas de nível (*benchmarks*) para $r = 3, 2, 1, 0$ são 0.0299, 0.063, 0.117 e 0.1390, respectivamente, representadas nas figuras por linhas pontilhadas. Examinando as Figuras 3.2(c) e (d) encontramos que as observações 1, 4, 6, 7, 10, 26, 33 e 37 são influentes, enquanto que os casos 1, 4, 7 e 10 são encontradas como influentes na Figura 3.2(e). As observações 7 e 33 são encontradas como mais influentes na Figura 3.2(b), embora as observações 1 e 7 sejam as mais influentes para as Figuras 3.2(c), (d) e (e).

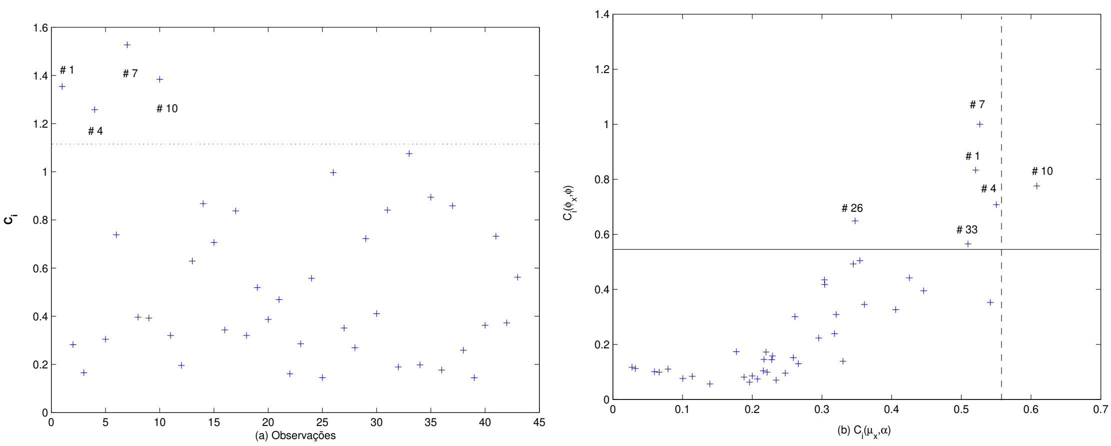


Figura 3.3: Ponderação de casos. (a) Gráfico de influência local total C_i vs. Observações. (b) Gráfico das medidas de influência local $C_i(\mu_x, \boldsymbol{\alpha})$ vs $C_i(\phi_x, \boldsymbol{\phi})$.

Na Figura 3.3(a) mostramos um *index plot* da influência local total C_i definida em (3.8), a marca de nível usada para avaliar as observações influentes dada por $2\sum_{i=1}^{n} C_i/n$ é igual a 1.1147 e é indicada no gráfico por uma línea pontilhada. O gráfico indica que as observações 1, 4, 7, e 10 são influentes na estimação completa do vetor de parâmetro $\boldsymbol{\theta}$. Na Figura 3.3(b), mostramos um *scatter plot* de $C_i(\mu_x,\, \boldsymbol{\alpha}')$ vs. $C_i(\phi_x,\, \boldsymbol{\phi}')$, as respectivas marcas de nível são dadas por 0.5581 e 0.5566, indicando que somente a observação 10 é localmente influente para ambas partes do parâmetro $\boldsymbol{\theta}$, isto é, $(\mu_x, \boldsymbol{\alpha}')'$ e $(\phi_x, \boldsymbol{\phi}')'$. Note que as observações localmente influentes encontradas em 3.3(a) coincidem com as encontradas na Figura

3.2(e) correspondente ao método de Poon & Poon (1999), para $r = 0$, indicando que ambas figuras são proporcionais e a marca de nível proposta aqui, é adequada.

Figura 3.4: Perturbação na variável resposta (a) $|hmax|$ vs. Observações. (b),(c), (d) e (e) Contribuições conjuntas. (f) Gráfico de influência total local C_i vs. observações.

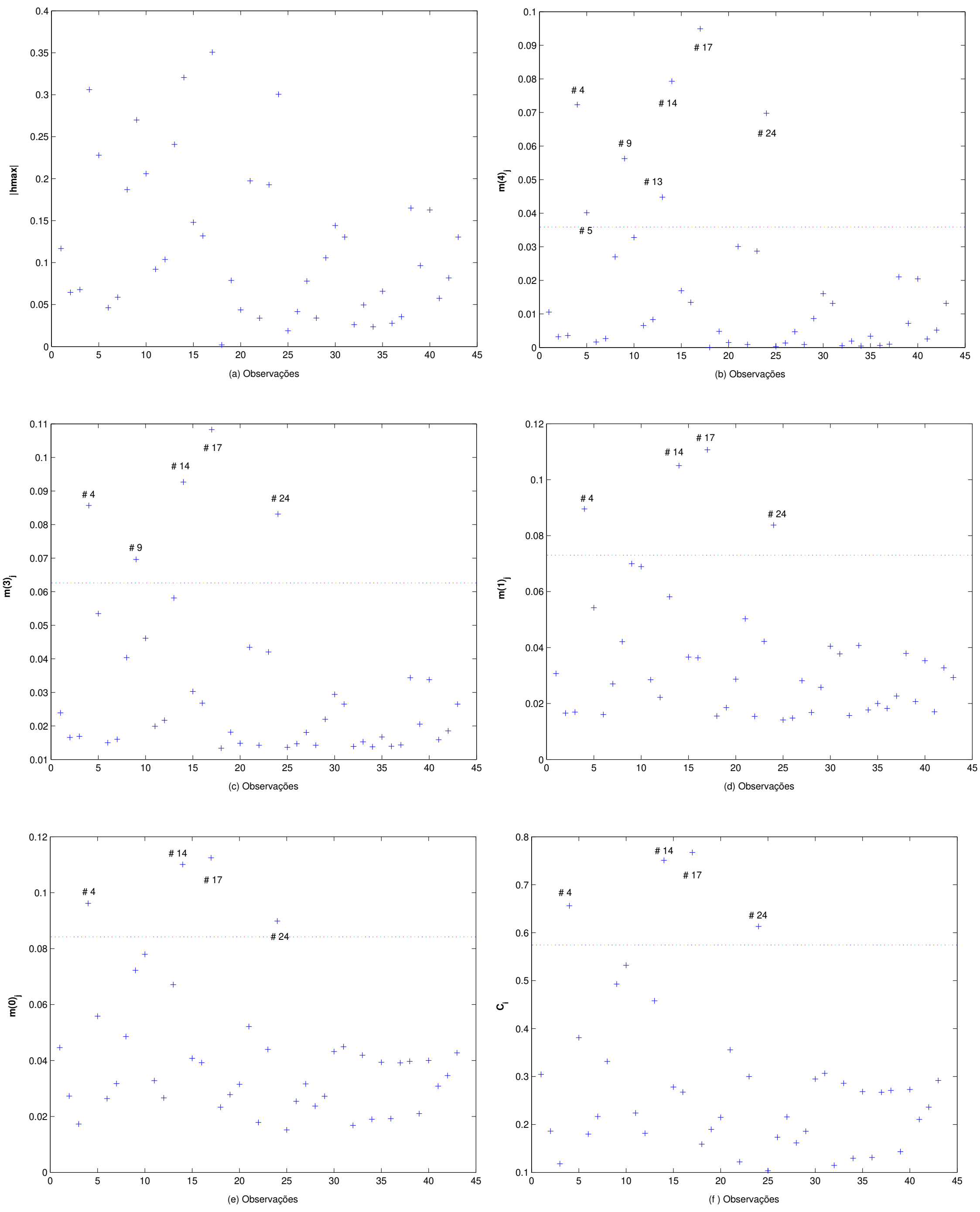


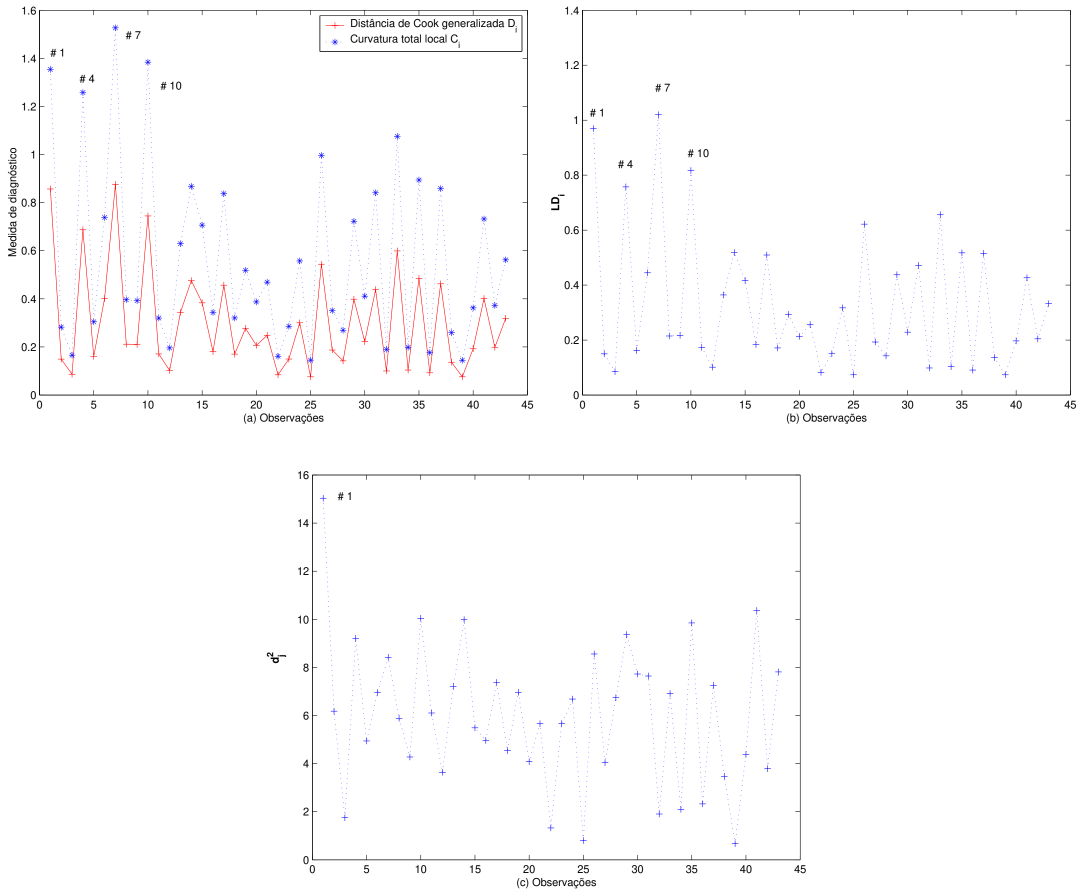
Tabela 3.1: Estimativas de máxima verossimilhança com os respectivos desvios padrões assintóticos, com a utilização dos dados completos, dados sem as observações A $(1, 4, 7, 10)$, sem as observações B $(4, 7, 10, 15, 33)$, e sem as observações C $(1, 4, 6, 7, 10, 15, 26, 33, 37)$.

$\widehat{\boldsymbol{\theta}}$	Todas as obs.	Sem as obs. A	Sem as obs. B	Sem as obs. C
$\widehat{\mu}_x$	4.3972 (0.0316)	4.3838 (0.0324)	4.3845 (0.0334)	4.3924 (0.0375)
$\widehat{\alpha}_2$	-0.0270 (0.0183)	-0.0264 (0.0174)	-0.0389 (0.0178)	-0.0356 (0.0188)
$\widehat{\alpha}_3$	0.0386 (0.0272)	0.0582 (0.0285)	0.0671 (0.0264)	0.0721 (0.0290)
$\widehat{\alpha}_4$	0.3188 (0.0437)	0.3274 (0.0444)	0.3439 (0.0438)	0.3359 (0.0491)
$\widehat{\alpha}_5$	0.3230 (0.0567)	0.3246 (0.0600)	0.3626 (0.0610)	0.3297 (0.0612)
$\widehat{\alpha}_6$	0.2228 (0.0296)	0.2141 (0.0291)	0.2113 (0.0310)	0.1762 (0.0272)
$\widehat{\phi}_x$	0.0361 (0.0084)	0.0330 (0.0080)	0.0340 (0.0083)	0.0374 (0.0094)
$\widehat{\phi}_1$	0.0068 (0.0024)	0.0080 (0.0025)	0.0085 (0.0026)	0.0103 (0.0029)
$\widehat{\phi}_2$	0.0076 (0.0026)	0.0038 (0.0018)	0.0036 (0.0018)	0.0017 (0.0014)
$\widehat{\phi}_3$	0.0249 (0.0060)	0.0237 (0.0059)	0.0180 (0.0046)	0.0183 (0.0048)
$\widehat{\phi}_4$	0.0753 (0.0168)	0.0688 (0.0160)	0.0644 (0.0152)	0.0717 (0.0177)
$\widehat{\phi}_5$	0.1313 (0.0289)	0.1326 (0.0305)	0.1330 (0.0310)	0.1168 (0.0286)
$\widehat{\phi}_6$	0.0309 (0.0073)	0.0249 (0.0061)	0.0280 (0.0069)	0.0148 (0.0040)

Na Tabela 3.1 encontram-se os valores as estimativas de máxima verossimilhança do vetor de parâmetros $\boldsymbol{\theta}$ e os respectivos desvios padrões assintóticos entre parênteses com a utilização dos dados completos, dos dados sem as observações 1, 4, 7 e 10 encontrados como influentes por C_i, dos dados sem as observações 4, 7, 10, 15, 33 encontrados como influentes por $|hmax|$ e dos dados sem as observações 1, 4, 6, 7, 10, 15, 26, 33 e 37 encontrados como influentes pelo método de Poon e Poon (1999), todas correspondentes à perturbação de ponderação de casos, para efeito de comparação dos resultados.

Agora vamos a examinar a influência local da perturbação na variável resposta detalhada na Seção 3.3.2. Na Figura 3.4, mostramos os *index plots* das contribuições conjuntas $m(r)_j$, $r = 4,\ 3,\ 1$ e 0, as marcas de nível respectivas são $0.0359,\ 0.0626,\ 0.0730$

Figura 3.5: Medidas de diagnóstico clássicas usadas em análises de regressão. (a) Comparação da distância de Cook generalizada e a curvatura local total C_i. (b) Afastamento pela verossimilhança LD_i vs. observações (c) Distância quadrada generalizada d_j^2 vs. observações.



e 0.0842, $r = 2$ não é mostrado por ser equivalente a $r = 3$. Note que $|hmax|$ é proporcional a $m(4)_j$ e todos os gráficos indicam que as observações 4, 14, 17 e 24 são influentes na estimação do vetor de parâmetro $\boldsymbol{\theta}$, sendo a observação 17 a mais influente. Na Figura 3.4(b) pode-se observar que as observações 5, 9 e 13 são também influente. Resultados similares a $m(0)_j$ são encontrados na Figura 3.4(f) (marca de nível igual a 0.5743), que mostra o *index plot* da influência local total C_i, indicando que ambas figuras são proporcionais como já tínhamos argumentado na seção prévia.

Como uma ilustração, na Figura 3.5 realizamos um análise de influência global com as medidas de diagnóstico clássicas, usadas geralmente em análises de regressão para detectar *outliers* e observações influentes, e discutidas na Seção 3.2; estas medidas são : afastamento pela verossimilhança LD_i; distância de Cook generalizada D_i e a distância quadrada generalizada d_j^2. Note que as medidas de influência global LD_i e D_i podem ser interpretadas como uma aproximação da medida de influência local C_i, além disso, a medida C_i é aproximadamente o dobro da medida D_i, como tínhamos argumentado na Seção 3.2.

3.4.2 Exemplo 2

Como uma segunda aplicação analisaremos o conjunto de dados apresentado em Grubbs (1948), que representa os tempos de queima de fusíveis (fuse-buring time), em segundos, que são reportados por três observadores A, B e C. O conjunto de observações y_{ij}, com $i = 1, 2, 3$ e $j = 1, \ldots, 30$ são apresentados na Tabela 3.2. Suporemos que o observador dois é o mais experiente, sendo assim e sem perda de generalidade consideraremos $\alpha_1 = 0$. A normalidade dos dados requerida para a análise é razoável (veja, Manuel Galea, 1995).

No conjunto de dados da Tabela 3.2, temos que a observação 19 correspondente ao segundo observador foi perdida, esta unidade perdida foi calculada usando o algoritmo EM e é dado por 10.0436. Para mais detalhes sobre este algoritmo veja, por exemplo, Johnson e Wichern (1998).

Como no exemplo anterior, analisaremos primeiro a ponderação de casos baseados na análise descrita por Cook (1986) e Poon e Poon (1999), este último com a finalidade de ter um critério objetivo para julgar as observações que verdadeiramente são influentes. Os resultados são mostrados na Figura 3.6. A Figura 3.6(a) mostra o *index plot* de $|hmax|$ correspondente à maior curvatura normal $C_{\mathbf{h}}$, este gráfico indica que as observações 2 e 4 são claramente influentes na estimação do vetor de parâmetros $\boldsymbol{\theta}$. As Figuras 3.6(b), (c), (d) e (e) são os *index plot* das contribuições conjuntas $m(r)_j$, com $r = 3, 2, 1$ e 0, respectivamente. As marcas de nível respectivas encontradas são 0.046, 0.1043, 0.1257 e 0.1475. A Figura 3.6(b) indica que as observações 2, 4 e 17 são influentes, sendo a observação 4 a mais influente. Nas Figuras 3.6(c), (d) e (e) notamos que as observações 9, 12, 15 são também influentes, sendo as observações 4 e 17 as mais influentes. C_i não é mostrado por ser

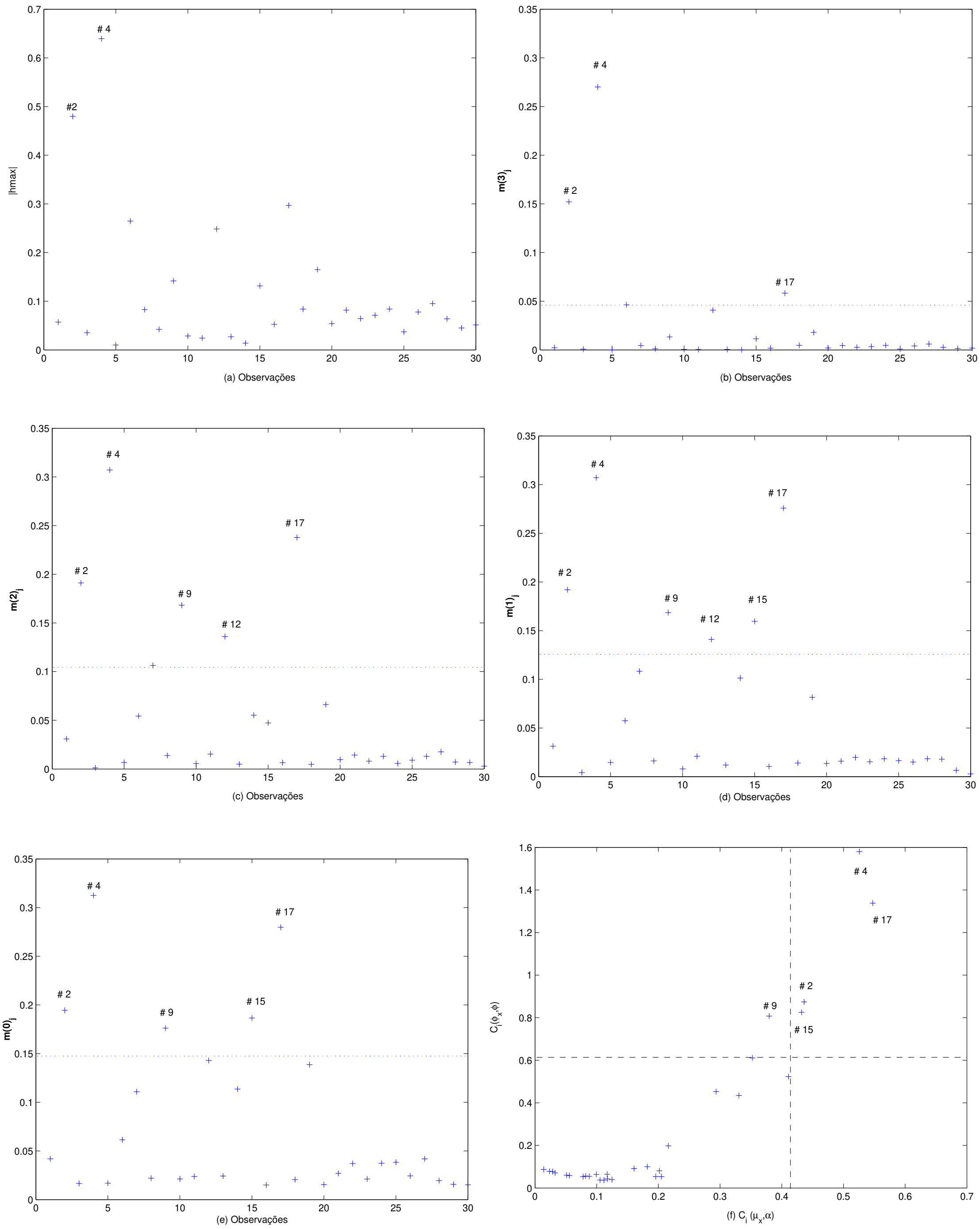
Tabela 3.2: Medidas simultâneas do tempo de queima de fusíveis , em segundos, tomadas por três observadores.

obs.	observador A	observador B	observador C
1	10.10	10.07	10.07
2	9.98	9.90	9.90
3	9.89	9.85	9.86
4	9.79	9.71	9.70
5	9.67	9.65	9.65
6	9.89	9.83	9.83
7	9.82	9.75	9.79
8	9.59	9.56	9.59
9	9.76	9.68	9.72
10	9.93	9.89	9.92
11	9.62	9.61	9.64
12	10.24	10.23	10.24
13	9.84	9.83	9.86
14	9.62	9.58	9.63
15	9.60	9.600	9.65
16	9.74	9.7300	9.74
17	10.32	10.32	10.34
18	9.86	9.86	9.86
19	10.01	–	10.03
20	9.65	9.64	9.65
21	9.50	9.49	9.50
22	9.56	9.56	9.55
23	9.54	9.53	9.54
24	9.89	9.89	9.88
25	9.53	9.52	9.51
26	9.52	9.52	9.53
27	9.44	9.43	9.45
28	9.67	9.67	9.67
29	9.77	9.76	9.78
30	9.86	9.84	9.86

equivalente a $m(0)_j$. Na figura 3.6(f) mostramos um *scatter plot* de $C_i(\mu_x, \boldsymbol{\alpha}')$ vs. $C_i(\phi_x, \boldsymbol{\phi}')$, as marcas de nível encontradas são 0.421 e 0.614, respectivamente, note que as observações 2, 4, 15 e 17 são encontradas como influentes em ambas partes do parâmetro e a observação 9 é somente influente na estimação de $(\phi_x, \boldsymbol{\phi}')$. Portanto, deve-se prestar especial atenção em todas as observações influentes encontradas e não somente nas observações 2 e 4, encontradas pelo método de influência local proposto por Cook (1986).

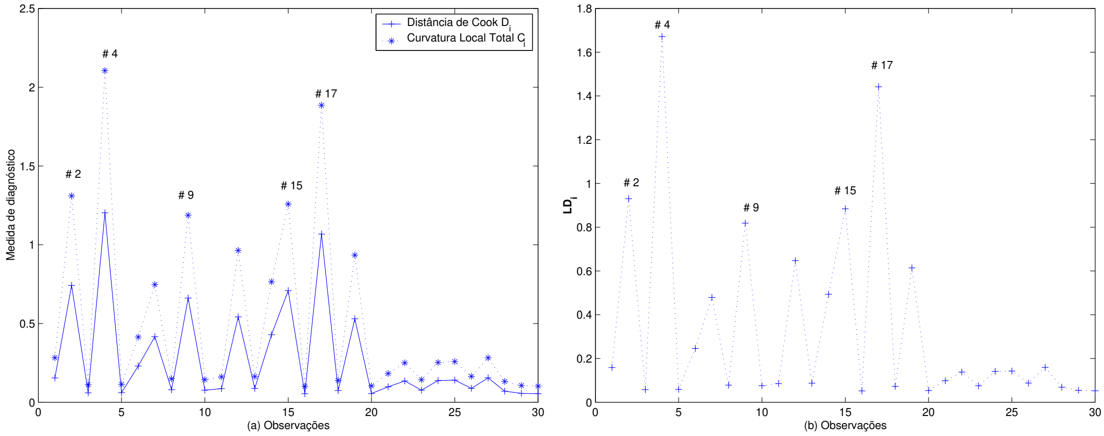
Vale a pena notar também que as observações influentes encontradas por C_i

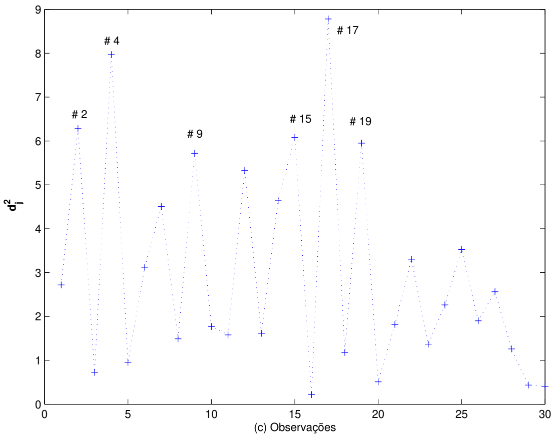
Figura 3.6: Dados de tempo de queima de fusíveis. Gráfico de influência local para perturbação de ponderação de casos.



(equivalente a $m(0)_j$) coincidem com as encontradas na Figura 3.7(a) e (b) correspondente a análise de influência global, mas estas observações foram encontradas

Figura 3.7: Medidas de diagnóstico clássicas usadas em análises de regressão. (a) Comparação da distância de Cook generalizada e a curvatura local total C_i. (b) Afastamento pela verossimilhança LD_i vs. observações (c) Distância quadrada generalizada d_j^2 vs. observações.





como influentes por critérios intuitivos e não por critérios objetivos como parece indicar a análise desenvolvida neste trabalho.

A Figura 3.8, mostra a análise de influência local para a perturbação na variável resposta correspondente ao método de Cook (1986) e à curvatura normal padronizada proposto por Poon e Poon (1999). As Figuras 3.8(b), (c), (d) são os *index plot* das contribuições conjuntas $m(r)_j$, $r = 3, 2, 0$, as marcas de nível respectivas são 0.0549, 0.0897 e 0.0916. $r = 1$ não é mostrado por ser equivalente a para $r = 2$. Estes gráficos indicam que as observações 1, 12, 17 e 27 são influentes, considerando a perturbação na variável resposta da relação em (3.19), na estimação do vetor paramétrico $\boldsymbol{\theta}$.

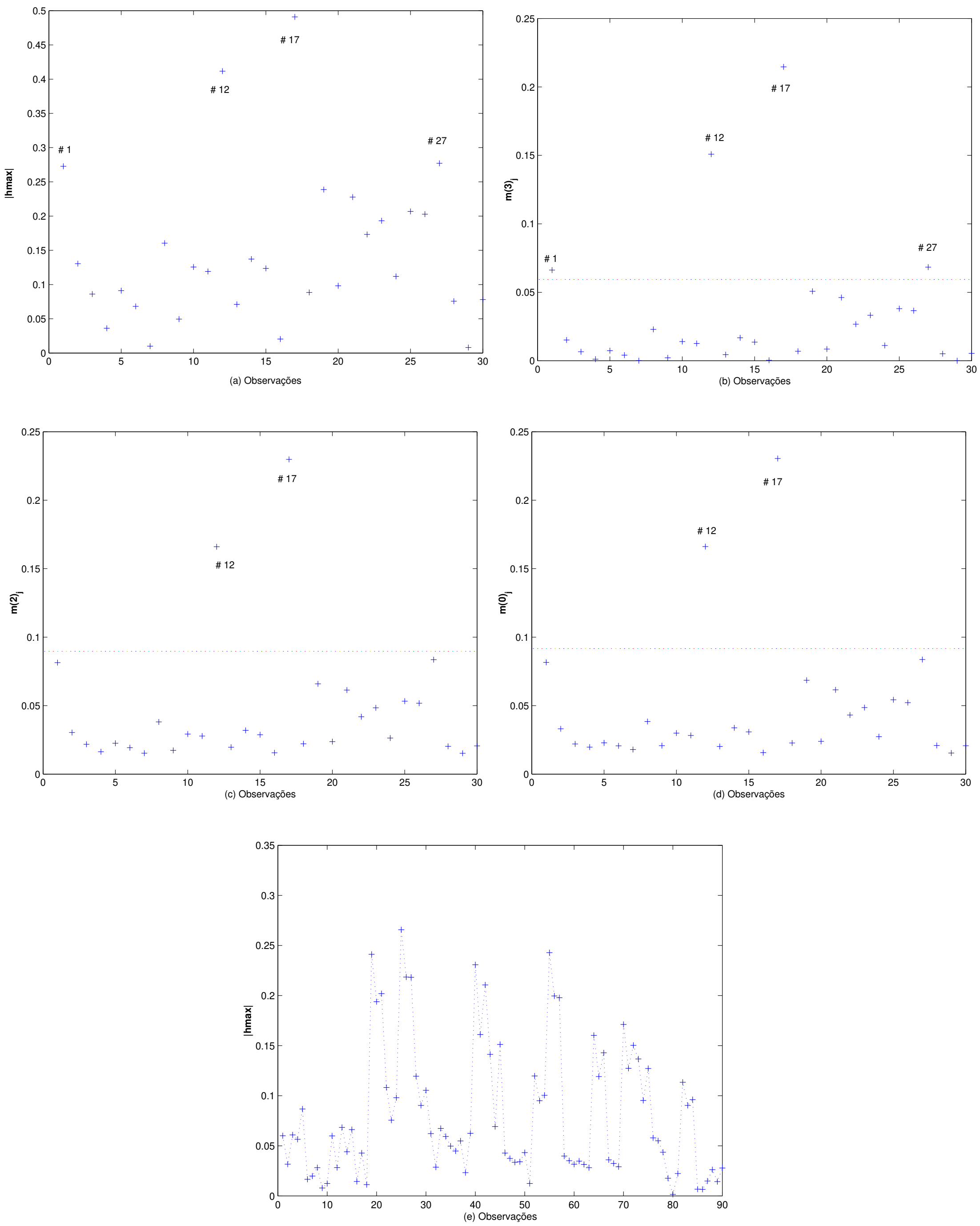
A Figura 3.8(e) mostra o *index plot* para a perturbação na variável resposta considerado em (3.23), note que este gráfico pode ajudar a encontrar dentro de cada observação influente, as unidades experimentais, y_{ij}, que verdadeiramente são influentes.

Na Tabela 3.3 encontram-se os valores as estimativas de máxima verossimilhança do vetor de parâmetros $\boldsymbol{\theta}$ e os respectivos desvios padrões assintóticos entre parênteses com a utilização dos dados completos, dos dados sem as observações 2, 4 encontrados como influentes por $|hmax|$, dos dados sem as observações 2, 4, 17, dos dados sem as observações 2, 4, 9, 15, 17, encontrados como influentes por C_i e dos dados sem as observações 2, 4, 9, 12, 15, 17, encontrados como influentes pelo método de Poon e Poon (1999), todas correspondentes à perturbação de ponderação de casos. Note que, em todos os casos as variâncias dos erros são as mais afetadas.

Tabela 3.3: Estimativas de máxima verossimilhança com os respectivos desvios padrões assintóticos, com a utilização dos dados completos, dados sem as observações A $(2, 4)$, sem as observações B (2, 4,17), sem as observações C $(2, 4, 9, 15, 17)$ e sem as observações D $(2, 4, 9, 12, 15, 17)$.

$\widehat{\boldsymbol{\theta}}$	Todas as obs.	Sem as obs. A	Sem as obs. B	Sem as obs. C	Sem as obs. D
$\widehat{\mu}_x$	9.75133 (0.03859)	9.74750 (0.04091)	9.72630 (0.03651)	9.73320 (0.03914)	9.73320 (0.03995)
$\widehat{\alpha}_2$	0.02200 (0.00502)	0.01786 (0.00443)	0.01852 (0.00457)	0.01680 (0.00418)	0.01680 (0.004268)
$\widehat{\alpha}_3$	0.01333 (0.00321)	0.01464 (0.00329)	0.01444 (0.00340)	0.01200 (0.00319)	0.01200 (0.00326)
$\widehat{\phi}_x$	0.04457 (0.01152)	0.04666 (0.01249)	0.03578 (0.00976)	0.03812 (0.01080)	0.03812 (0.01102)
$\widehat{\phi}_1$	0.00011 (0.00009)	0.00019 (0.00009)	0.00021 (0.00009)	0.00018 (0.00008)	0.00018 (0.00008)
$\widehat{\phi}_2$	0.00065 (0.00019)	0.00036 (0.00012)	0.00036 (0.00012)	0.00026 (0.00009)	0.00026 (0.00009)
$\widehat{\phi}_3$	0.00020 (0.00010)	0.00011 (0.00007)	0.00011 (0.00008)	0.00008 (0.00006)	0.00008 (0.00006)

Figura 3.8: Perturbação na variável resposta (a) $|hmax|$ vs. observações. (b), (c), (d) Contribuições conjuntas. (e) Gráfico $|hmax|$ vs Observações para a perturbação na variável resposta por unidade experimental considerado em (3.23)



Capítulo 4

Considerações Finais

4.1 Conclusões

Neste trabalho, o objetivo principal foi considerar um estudo de inferência e diagnóstico no modelo de Grubbs estrutural normal. No Capítulo 2 foi feito um estudo de inferência baseado em testes assintóticos para testar a igualdade de variâncias e vícios de p (≥ 3) instrumentos de medida, que medem a característica de interesse na mesma escala. Neste caso, os procedimentos de inferência foram baseados considerando três hipóteses H_{01}, H_{02} e H_{03}, os testes são baseados nas estatísticas de Wald, Escore e Razão de Verossimilhanças. Mostramos através de simulação que o teste de Escore tem melhor comportamento em relação aos outros dois testes, em termos de poder e nível de significância. A estatística de Escore utilizada para testar a hipóteses H_{01}, tem uma forma explicita e coincide com a expressão obtida por Bedrick (2001), quem utiliza uma parametrização diferente deste trabalho. Similarmente, a estatística de Escore para testar H_{03} tem forma explicita, este resultado deve-se a que em ambos casos os EMV sob as hipóteses nulas, tem forma explicita. Embora, a estatística de Escore para testar H_{02} não tenha forma fechada, sua implementação computacional é simples. Nos casos em que o EMV não tem uma forma fechada, a obtenção deles foi por meio do algoritmo EM.

Notamos também que no modelo de Grubbs os parâmetros de locação e os parâmetros de escala tem propriedades ortogonais no sentido que a matriz de informação esperada de Fisher é diagonal em bloco. Assim a estatística de Wald pode ser escrita como: $W_{01} = W_{02} + W_{03}$. Além disso, neste capítulo verificamos a eficiência assintótica dos EMV, encontrados via algoritmo EM ($p \geq 3$), com relação aos estimadores de momentos propostos por Grubbs.

No Capítulo 3, discutimos um procedimento de diagnóstico, baseado no método de influência local proposto por Cook (1986) e a curvatura normal conformalizada de Poon e Poon (1999). Dois esquemas de perturbação são considerados na análise. Como observamos, ambas metodologias prevêem um estudo unificado para avaliar a influência local de pequenas perturbações e pode ser aplicado a modelos mais gerais. Os resultados obtidos foram comparados com outras medidas de diagnóstico usadas na literatura, com por exemplo, o afastamento pela verossimilhança LD_i e a distância de Cook generalizada D_i, notando que estas medidas prevêem quase as mesmas observações influentes que as obtidas pelas metodologias consideradas $(C_i, r = 0)$ para a perturbação de ponderação de casos, só que estas possuem um critério mais objetivo para avaliar influências de observações, além de outras propriedades importantes apresentadas neste capítulo.

4.2 Pesquisa Futura

Neste trabalho, foi considerado o modelo de Grubbs estrutural normal. Pode-se desenvolver um estudo similar no modelo de Grubbs funcional e sob o modelo estrutural de Grubbs, podemos considerar o caso em que os dados seguem uma distribuição "t-Student".

Pretendemos também fazer uma extensão do resultado do capitulo 2, considerando a robustez dos testes a dados não normais, digamos "t-Student", como é desenvolvido em Bedrick (2001). Também considerar outras hipóteses de interesse para fazer o estudo de inferência, como por exemplo, $H_0 \; : \; \alpha_2 = \ldots = \alpha_p = \alpha$ ou $H_0 \; : \; \alpha_2 = \ldots = \alpha_p = \alpha$ e $\phi_1 = \ldots = \phi_p = \phi$.

No contexto de diagnóstico, podemos considerar outros esquemas de perturbação para realizar a análise de influência local, por exemplo, $x_{w_j} = x_j + w_j$.

Apêndice A

Obtenção da Matriz de Informação de Fisher Esperada

Neste Apêndice apresentamos os Lemas principais para obter a matriz de informação de Fisher esperada $\mathbf{J}_{F1}$, utilizados na prova dos Teoremas 2.1 e 2.3. Alguns cálculos numéricos são apresentados para ilustrar seu uso.

Lema A.1. *Sejam $\boldsymbol{W}$ e $\boldsymbol{T}$ matrizes aleatórias de ordem $k_1 \times k_2$, $\boldsymbol{A}_1$ uma matriz constante $m_1 \times k_1$ e $\boldsymbol{A}_2$ uma matriz constante de Ordem $k_2 \times m_2$. Então*

1. $E(\boldsymbol{A}_1 \boldsymbol{W}) = \boldsymbol{A}_1 E(\boldsymbol{W})$, $E(\boldsymbol{W}\boldsymbol{A}_2) = E(\boldsymbol{W})\boldsymbol{A}_2$,
2. $E(\boldsymbol{A}_1 \boldsymbol{W}\boldsymbol{A}_2) = \boldsymbol{A}_1 E(\boldsymbol{W})\boldsymbol{A}_2$,
3. $E(\boldsymbol{T} + \boldsymbol{W}) = E(\boldsymbol{T}) + E(\boldsymbol{W})$

A prova deste Lema pode ser encontrada em Graybill (1983).

Lema A.2. *Se $c = 1+\phi_x \boldsymbol{1}_p' D^{-1}(\boldsymbol{\phi})\boldsymbol{1}_p$, $\boldsymbol{M} = \phi_x D^{-1}(\boldsymbol{\phi})\boldsymbol{1}_p'\boldsymbol{1}_p D^{-1}(\boldsymbol{\phi})$ e $\boldsymbol{W} \sim N_p(\boldsymbol{0}, \boldsymbol{\Sigma})$, onde $\boldsymbol{\Sigma} = \phi_x \boldsymbol{1}_p \boldsymbol{1}_p' + D(\boldsymbol{\phi})$ então:*

1. $E(\boldsymbol{W}\boldsymbol{M}\boldsymbol{W}) = c(c-1)$,
2. $E(D(\boldsymbol{W})D(\boldsymbol{W})) = \phi_x \boldsymbol{I}_p + D(\boldsymbol{\phi})$,
3. $E(D(\boldsymbol{W})) = \boldsymbol{0}_p$,
4. $E(\boldsymbol{W}' D^{-1}(\boldsymbol{\phi})\boldsymbol{1}_p D(\boldsymbol{W})) = c\boldsymbol{I}_p$.

Demonstração. 1. Sabemos que (veja, Graybill, 1983),

$$\begin{aligned} E(\mathbf{WMW}) &= tr(\mathbf{M\Sigma}) = tr(\phi_x(c-1)\mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{1}_p\mathbf{1}_p' + \phi_x\mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{1}_p\mathbf{1}_p') \\ &= c\phi_x tr(\mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{1}_p\mathbf{1}_p') = c(c-1). \end{aligned}$$

2. $E(\mathrm{D}(\mathbf{W})\mathrm{D}(\mathbf{W})) = \mathrm{D}(E(\mathbf{W} * \mathbf{W})) = \mathrm{D}(\mathbf{\Sigma}) = \phi_x\mathbf{I}_p + \mathrm{D}(\boldsymbol{\phi})$, onde $\mathbf{W} * \mathbf{W} = [(w_i^2)]$

3. $\mathrm{E}(\mathrm{D}(\mathbf{W}))=\mathrm{D}(\mathrm{E}(\mathbf{W}))=\mathrm{D}(\mathbf{0})=\mathbf{0}$.

4. $E(\mathbf{W}'\mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{1}_p\,\mathrm{D}(\mathbf{W})) = E(\mathrm{D}(\mathbf{W}(\mathbf{W}'\mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{1}_p))) = \mathrm{D}(\mathbf{\Sigma}\mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{1}_p) = c\mathbf{I}_p$

□

Para ilustrar os resultados acima, vamos obter alguns resultados dos elementos da matriz de informação de Fisher dada no Teorema 2.1, baseados na relação em (2.9), com $\mathbf{W} = \mathbf{W}_j \sim N_p(\mathbf{0}, \mathbf{\Sigma})$

(1) $-E(\dfrac{\partial^2 \ell_j}{\partial\mu_x\partial\boldsymbol{\alpha}'}) = J_{\mu_x\alpha} = -E(-\mathbf{1}_p\mathbf{\Sigma}^{-1}\mathbb{I}_{(p)}) = \mathbf{1}_p\mathbf{\Sigma}^{-1}\mathbb{I}_{(p)}\,,$

(2)

$$\begin{aligned} -E(\frac{\partial^2 \ell_j}{\partial\mu_x\partial\phi_x}) &= -E(-\frac{c^{-2}}{\phi_x}(c-1)\mathbf{1}_p'\mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{W}_j) = \frac{c^{-2}}{\phi_x}(c-1)\mathbf{1}_p'\mathrm{D}^{-1}(\boldsymbol{\phi})E(\mathbf{W}_j) \\ &= 0\,, \end{aligned}$$

(3)

$$-E(\frac{\partial^2 \ell_j}{\partial\boldsymbol{\alpha}\partial\boldsymbol{\phi}_x}) = -E(-\frac{c^{-2}}{\phi_x}\mathbb{I}_{(p)}\mathbf{M}\mathbf{W}_j) = \frac{c^{-2}}{\phi_x}\mathbb{I}_{(p)}\mathbf{M}E(\mathbf{W}_j) = \mathbf{0}_{p-1}\,,$$

(4)

$$\begin{aligned} -E(\frac{\partial^2 \ell_j}{\partial\phi_x\partial\phi_x}) &= J_{\phi_x\phi_x} = -E(\frac{1}{2\phi_x^2}\left(\frac{c-1}{c}\right)^2 - \frac{c^{-3}}{\phi_x}\mathbf{W}_j'\mathbf{M}\mathbf{W}_j\mathbf{1}_p'\mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{1}_p) \\ &= -\frac{1}{2\phi_x^2}\left(\frac{c-1}{c}\right)^2 + \frac{c^{-3}}{\phi_x}E(\mathbf{W}_j'\mathbf{M}\mathbf{W}_j)\mathbf{1}_p'\mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{1}_p \\ &= -\frac{1}{2\phi_x^2}\left(\frac{c-1}{c}\right)^2 + \frac{c^{-3}}{\phi_x}c(c-1)\mathbf{1}_p'\mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{1}_p \\ &= -\frac{1}{2\phi_x^2}\left(\frac{c-1}{c}\right)^2 + \frac{c^{-2}}{\phi_x}(c-1)\mathbf{1}_p'\mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{1}_p\,, \end{aligned}$$

(5)

$$\begin{aligned}
-E(\frac{\partial^2 \ell_j}{\partial \phi_x \partial \boldsymbol{\phi}'}) &= J_{\phi_x \phi} \\
&= -E(\frac{c^{-2}}{2}\mathbf{1}_p'\mathrm{D}^{-2}(\boldsymbol{\phi}) + c^{-3}\mathbf{W}_j'\mathbf{M}\mathbf{W}_j\mathbf{1}_p'\mathrm{D}^{-2}(\boldsymbol{\phi}) - c^{-2}\mathbf{W}_j'\mathrm{D}^{-1}(\boldsymbol{\phi})\mathbf{1}_p\mathbf{W}_j'\mathrm{D}^{-2}(\boldsymbol{\phi})) \\
&= -\frac{c^{-2}}{2}\mathbf{1}_p'\mathrm{D}^{-2}(\boldsymbol{\phi}) - c^{-3}E(\mathbf{W}_j'\mathbf{M}\mathbf{W}_j)\mathbf{1}_p'\mathrm{D}^{-2}(\boldsymbol{\phi}) + c^{-2}\mathbf{1}_p'\mathrm{D}^{-1}E(\mathbf{W}\mathbf{W}_j')\mathrm{D}^{-2}(\boldsymbol{\phi}) \\
&= -\frac{c^{-2}}{2}\mathbf{1}_p'\mathrm{D}^{-2}(\boldsymbol{\phi}) - c^{-3}c(c-1)\mathbf{1}_p'\mathrm{D}^{-2}(\boldsymbol{\phi}) + c^{-2}\mathbf{1}_p'\mathrm{D}^{-1}\boldsymbol{\Sigma}\mathrm{D}^{-2}(\boldsymbol{\phi}) \\
&= -\frac{c^{-2}}{2}\mathbf{1}_p'\mathrm{D}^{-2}(\boldsymbol{\phi}) \ - \ c^{-2}(c-1)\mathbf{1}_p'\,\mathrm{D}^{-2}(\boldsymbol{\phi}) + \ c^{-2}\mathbf{1}_p'\mathrm{D}^{-1}(\boldsymbol{\phi})\boldsymbol{\Sigma}D^{-2}(\boldsymbol{\phi}).
\end{aligned}$$

Os outros elementos de $\mathbf{J}_{F1}$, seguem facilmente de (2.9).

Apêndice B

Implementação do Algoritmo EM

Neste apêndice mostramos o programa escrito em Matlab 6.0, para calcular os estimadores de máxima verossimilhança do vetor de parâmetros $\boldsymbol{\theta}$ sob o modelo irrestrito usando o algoritmo EM, conforme descrito na Seção 2.2.3, para qualquer conjunto de dados.

```
function y=emv(b)
% b e o conjunto de dados
%encontra os EMV no modelo de Grubbs usando o algoritmo EM
% n e p qualquer
[n,p]=size(b); x=b; o=ones(p,1); f=zeros(n,1); f2=zeros(n,1);
alfa=zeros(p-1,1); phi=zeros(p,1); suma1 = sum(x)/n; x1=x(:,1);
ux=suma1(1); s=0;
for i=1:n
 s=s+((x(i,:))'-suma1')*(x(i,:)-suma1);
end c=s/n;
% estimadores iniciais usando os estimadores
% tipo momento proposto por Grubbs
for i=1:p-1
 alfa(i)=suma1(i+1)-ux;
end
 sx=sum(sum((c-diag(diag(c)))))/(p*(p-1));
for i=1:p
 phi(i)=abs(c(i,i)-sx);
end
```

```
cont=1;
%algoritmo EM

for j=1:120
 a=o'*diag(phi)^(-1)*o; c=1+sx*a; u=[0;alfa]+o*ux;
   for i=1:n
     f(i)=ux+sx/c*[o'*inv(diag(phi))]*((x(i,:))'-u);
     f2(i)=(f(i))^2+sx/c;
   end
 ux=sum(f)/n;
   for i=1:p-1
     alfa(i)=suma1(i+1)-ux;
   end
 sx=sum(f2)/n-ux^2;
 phi(1)= x(:,1)'*x(:,1)/n-2*x1'*f/n+sum(f2)/n;
   for i=2:p
     phi(i)=(x(:,i)-ones(n,1)*suma1(i))'*(x(:,i)-ones(n,1)*suma1(i))/n
             -2*(x(:,i)-ones(n,1)*suma1(i))'*(f-ones(n,1)*ux)/n+sx;
   end
   if j==cont
     itera=j
     theta=[ux,alfa',sx,phi']'
     cont=cont+10;
   end
end
y=[theta];
```

Apêndice C

Obtenção da Estatística de Escore

A estatística de Escore é definida como

$$E = \frac{1}{n}\mathbf{S}(\widetilde{\boldsymbol{\theta}})'\widetilde{\mathbf{J}}_{F1}^{-1}\mathbf{S}(\widetilde{\boldsymbol{\theta}})\,, \tag{C.1}$$

onde $\mathbf{S}(\boldsymbol{\theta})$ é o vetor de derivadas parciais da log-verossimilhança $\ell(\boldsymbol{\theta})$ definida em (2.5), $\widetilde{\boldsymbol{\theta}}$ é o EMV de $\boldsymbol{\theta}$ sob a hipótese nula H_o, e $\widetilde{\mathbf{J}}_{F1}$ é a matriz de informação esperada de Fisher definida no Teorema 2.1, avaliado em $\widetilde{\boldsymbol{\theta}}$. No Teorema 2.1 notamos que $\mathbf{J}_{F1}$ é uma matriz diagonal em blocos, após as manipulações algébricas obtemos que J_{F1}^{-1} é também diagonal em bloco (veja, Rao, 1973, pag. 33), isto é

$$\mathbf{J}_{F1}^{-1} = \begin{pmatrix} J_{11} & J_{12} & 0 & 0 \\ J_{21} & J_{22} & 0 & 0 \\ 0 & 0 & J_{11}^{*} & J_{12}^{*} \\ 0 & 0 & J_{21}^{*} & J_{22}^{*} \end{pmatrix}, \tag{C.2}$$

onde as submatrizes têm a forma
$J_{11} = [J_{\mu_x\mu_x} - J_{\mu_x\alpha}J_{\alpha\alpha}^{-1}J'_{\mu_x\alpha}]^{-1}$, $J_{12} = -J_{\mu_x\mu_x}J_{\mu_x\alpha}[J_{\alpha\alpha} - J'_{\mu_x\alpha}J_{\mu_x\mu_x}^{-1}J_{\mu_x\alpha}]^{-1} = J'_{21}$,
$J_{22} = [J_{\alpha\alpha} - J'_{\mu_x\alpha}J_{\mu_x\mu_x}^{-1}J_{\mu_x\alpha}]^{-1}$, $J_{11}^{*} = [J_{\phi_x\phi_x} - J_{\phi_x\phi}J_{\phi\phi}^{-1}J'_{\phi_x\phi}]^{-1}$, $J_{22}^{*} = [J_{\phi\phi} - J'_{\phi_x\phi}J_{\phi_x\phi_x}^{-1}J_{\phi_x\phi}]^{-1}$,
$J_{12}^{*} = -J_{\phi_x\phi_x}J_{\phi_x\phi}[J_{\phi\phi} - J'_{\phi_x\phi}J_{\phi_x\phi_x}^{-1}J_{\phi_x\phi}]^{-1} = J_{21}^{*\prime}$.

Se queremos testar a hipótese H_{01}, isto é, $H_{01}\ :\ \boldsymbol{\alpha} = \mathbf{0}_{(p-1)}, \boldsymbol{\phi} = \phi\mathbf{1}_p$, sabemos de 2.55 que $\widetilde{\boldsymbol{\theta}} = (\widetilde{\mu}_x, \widetilde{\phi}_x, \widetilde{\phi})$ é o EMV sob H_{01}. Logo

$$\mathbf{S}(\widetilde{\boldsymbol{\theta}}) = [0, \widetilde{\mathbf{S}}_1, 0, \widetilde{\mathbf{S}}_2] = \frac{\partial \ell(\boldsymbol{\theta})}{\partial(\boldsymbol{\theta})}\Big|_{\boldsymbol{\theta}=\widetilde{\boldsymbol{\theta}}}\,, \tag{C.3}$$

onde $\widetilde{\mathbf{S}}_1 = \frac{\partial \ell(\boldsymbol{\theta})}{\partial(\boldsymbol{\alpha})}\Big|_{\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}}$ e $\widetilde{\mathbf{S}}_2 = \frac{\partial \ell(\boldsymbol{\theta})}{\partial(\boldsymbol{\phi})}\Big|_{\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}}$.

Substituindo as relações $(C.2)$ e $(C.3)$ em $(C.1)$ e multiplicando convenientemente, temos que E_{01} é dada por:

$$E_{01} = \frac{1}{n}[\widetilde{\mathbf{S}}_1' \widetilde{J}_{22} \widetilde{\mathbf{S}}_1 + \widetilde{\mathbf{S}}_2' \widetilde{J^*}_{22} \widetilde{\mathbf{S}}_2] = E_1 + E_2, \tag{C.4}$$

onde $E_1 = \frac{1}{n}[\widetilde{\mathbf{S}}_1' \widetilde{J}_{22} \widetilde{\mathbf{S}}_1]$ e $E_2 = \frac{1}{n}[\widetilde{\mathbf{S}}_2' \widetilde{J^*}_{22} \widetilde{\mathbf{S}}_2]$.

a) Obtenção de $\mathbf{E_1}$.

Segue de (2.7) e do Teorema 2.1 que

$$\widetilde{S}_1 = \frac{\partial \ell(\boldsymbol{\theta})}{\partial(\boldsymbol{\alpha})}|_{\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}} = \sum_{j=1}^{p} \mathbb{I}_{(p)} \boldsymbol{\Sigma}^{-1} \mathbf{W}_j|_{\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}} = \sum_{j=1}^{p} \frac{\mathbb{I}_{(p)} \mathbf{W}_j}{\widetilde{\phi}},$$

$$\begin{aligned}
\widetilde{J}_{22} = J_{22}(\widetilde{\boldsymbol{\theta}}) &= [\mathbb{I}_{(p)} \widetilde{\boldsymbol{\Sigma}}^{-1} \mathbb{I}_{(p)}' - (\mathbf{1}_p' \widetilde{\boldsymbol{\Sigma}}^{-1} \mathbb{I}_{(p)}')' (\mathbf{1}_p \widetilde{\boldsymbol{\Sigma}}^{-1} \mathbf{1}_p)^{-1} (\mathbf{1}_p' \widetilde{\boldsymbol{\Sigma}}^{-1} \mathbb{I}_{(p)}')]^{-1} \\
&= [\frac{1}{\widetilde{\phi}} \mathbf{I}_q - \frac{c^{-1} \widetilde{\phi}_x}{\widetilde{\phi}^2} \mathbf{1}_q \mathbf{1}_q' - (\frac{\mathbf{1}_q}{\widetilde{\phi} c})(\frac{p}{c \widetilde{\phi}})^{-1} (\frac{\mathbf{1}_q}{\widetilde{\phi} c})']^{-1} \\
&= \widetilde{\phi}[\mathbf{I}_q + \mathbf{1}_q\, \mathbf{1}_q'],
\end{aligned}$$

onde $\mathbf{I}_q$ matriz identidade de ordem $q = p-1$, $\mathbf{W}_j = \mathbf{y}_j - \mathbf{1}_p \widetilde{\mu}_x$; c, $\widetilde{\boldsymbol{\Sigma}}$ como em (2.4) e avaliados em $\widetilde{\boldsymbol{\theta}}$. Logo, após de manipulações algébricas temos que

$$\begin{aligned}
E_1 &= \frac{1}{n}[\sum_{j=1}^{p} \frac{\mathbb{I}_{(p)} \mathbf{W}_j}{\widetilde{\phi}}]' \widetilde{\phi}[\mathbf{I}_q + \mathbf{1}_q\, \mathbf{1}_q'][\sum_{j=1}^{p} \frac{\mathbb{I}_{(p)} \mathbf{W}_j}{\widetilde{\phi}}] \\
&= \frac{n}{\widetilde{\phi}} \sum_{i=1}^{p} (\bar{y}_{i.} - \bar{y}_{..})^2.
\end{aligned}$$

b) Obtenção de $\mathbf{E_2}$.

Do Teorema 2.1, temos que

$$\begin{aligned}
\widetilde{J}_{\phi\phi} = J_{\phi\phi}(\widetilde{\boldsymbol{\theta}}) &= [-\frac{\mathbf{I}_p}{2\widetilde{\phi}^2} - \frac{\widetilde{\phi}_x^2}{2\widetilde{\phi}^2} c^{-2} \widetilde{\mathbf{M}} + \frac{\widetilde{\phi}_x}{\widetilde{\phi}^2} c^{-1} \widetilde{\mathbf{M}} - 2\frac{\widetilde{\phi}_x}{\widetilde{\phi}^3} c^{-2} \widetilde{\boldsymbol{\Sigma}} \widetilde{M} + \frac{\widetilde{\phi}_x}{\widetilde{\phi}^4} c^{-1} \widetilde{\boldsymbol{\Sigma}}] \\
&= \frac{1}{2\widetilde{\phi}^2}[(1 - 2\widetilde{\tau}) \mathbf{I}_p + \widetilde{\tau}^2 \mathbf{1}_p \mathbf{1}_p'],
\end{aligned}$$

$$\begin{aligned}
\widetilde{J}_{\phi_x\phi} = J_{\phi_x\phi}(\widetilde{\boldsymbol{\theta}}) &= -\frac{c^{-2}}{2\widetilde{\phi}^2} \mathbf{1}_p' - \frac{1}{\widetilde{\phi}^2} c^{-2}(c-1) \mathbf{1}_p' + \frac{c^{-2} \mathbf{1}_p'}{\widetilde{\phi}^3} (\mathbf{1}_p \mathbf{1}_p' \widetilde{\phi}_x + \mathbf{I}_p \widetilde{\phi}) \\
&= -\frac{c^{-2} \mathbf{1}_p'}{2\widetilde{\phi}^2},
\end{aligned}$$

$$\widetilde{J}_{\phi_x\phi_x} = J_{\phi_x\phi_x}(\widetilde{\boldsymbol{\theta}}) \;=\; -\frac{1}{2\widetilde{\phi}_x^2}(\frac{c-1}{c})^2 + \frac{c^{-2}}{\widetilde{\phi}_x\widetilde{\phi}}(c-1)p = \frac{c^{-2}p^2}{2\widetilde{\phi}^2},$$

onde $\widehat{\tau} = \dfrac{\widetilde{\phi}_x}{\widetilde{\phi} + p\,\widetilde{\phi}_x}$ e $\widetilde{\mathbf{M}}$ como em (2.4), avaliado em $\widetilde{\boldsymbol{\theta}}$. Segue, após as manipulações algébricas que

$$\widetilde{J}^*_{22} = [\widetilde{J}_{\phi\phi} - \widetilde{J}'_{\phi_x\phi}\widetilde{J}^{-1}_{\phi_x\phi_x}\widetilde{J}_{\phi_x\phi}]^{-1} = \frac{2\widetilde{\phi}^2}{1-2\widetilde{\tau}}\left[\mathbf{I}_p - \frac{1-\dfrac{2\widetilde{\tau}\widetilde{\phi}}{\widetilde{\phi}_x}}{p(p-1)}\mathbf{1}_p\mathbf{1}'_p\right] = \mathbf{B}\,.$$

Da relação em (2.7), temos que

$$\begin{aligned}\widetilde{\mathbf{S}}_2 = \frac{\partial \ell(\boldsymbol{\theta})}{\partial(\boldsymbol{\phi})}|_{\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}} \;&=\; -\frac{n}{2\widetilde{\phi}}\mathbf{1}_p + \frac{1}{2\widetilde{\phi}^2}\sum_{j=1}^{n}\mathrm{D}(\mathbf{W}_j)\mathbf{W}_j + \frac{n\widetilde{\phi}_x}{2\widetilde{\phi}^2}\mathbf{1}_p - \frac{\widetilde{\tau}}{\widetilde{\phi}^2}\sum_{j=1}^{n}\mathbf{W}'_j\mathbf{1}_p\mathbf{W}_j\\ &=\; \frac{1}{2\widetilde{\phi}^2}\left[n(\widetilde{\phi}_x - \widetilde{\phi})\mathbf{1}_p + \sum_{j=1}^{n}\mathrm{D}(\mathbf{W}_j)\mathbf{W}_j - 2\widetilde{\tau}\sum_{j=1}^{n}\mathbf{W}'_j\mathbf{1}_p\mathbf{W}_j\right] = \mathbf{A}.\end{aligned}$$

Portanto, de (C.4), segue que $E_2 = \dfrac{1}{n}\mathbf{A}'\mathbf{B}\mathbf{A}$.

Para testar a hipótese H_{03}, isto é, $H_{03} \;:\; \boldsymbol{\phi} = \phi\mathbf{1}_p$, sabemos de (2.59) que $\widetilde{\boldsymbol{\theta}} = (\widetilde{\mu}_x, \widetilde{\boldsymbol{\alpha}}, \widetilde{\phi}_x, \widetilde{\phi})'$ são os EMV sob H_{03}. Logo

$$\mathbf{S}(\widetilde{\boldsymbol{\theta}}) = [0,\, \mathbf{0}_{p-1}, 0, \widetilde{\mathbf{S}}_2] = \frac{\partial \ell(\boldsymbol{\theta})}{\partial(\boldsymbol{\theta})}|_{\boldsymbol{\theta}=\widetilde{\boldsymbol{\theta}}}\,, \tag{C.5}$$

onde $\widetilde{\mathbf{S}}_2 = \dfrac{\partial \ell(\boldsymbol{\theta})}{\partial(\boldsymbol{\phi})}|_{\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}}$. Substituindo as relações $(C.2)$ e $(C.5)$ em $(C.1)$ e multiplicando convenientemente, temos que E_{03} é dada por:

$$E_{03} = \frac{1}{n}[\widetilde{\mathbf{S}}'_2\widetilde{J^*}_{22}\widetilde{\mathbf{S}}_2]. \tag{C.6}$$

Assim, após as manipulações algébricas, temos que

$$E_{03} = \frac{1}{n}\mathbf{A}'\,\mathbf{B}\,\mathbf{A}\,,$$

onde

$$\mathbf{A} = \frac{1}{2\widetilde{\phi}^2}\left[n(\widetilde{\phi}_x - \widetilde{\phi})\mathbf{1}_p + \sum_{j=1}^{n}\mathrm{D}(\mathbf{W}_j)\mathbf{W}_j - 2\widetilde{\tau}\sum_{j=1}^{n}\;\mathbf{W}'_j\mathbf{1}_p\mathbf{W}_j\right]$$

e

$$\mathbf{B} = \frac{2\widetilde{\phi}^2}{1-2\widetilde{\tau}}\left[\mathbf{I}_p - \frac{1-\dfrac{2\widetilde{\tau}\widetilde{\phi}}{\widetilde{\phi}_x}}{p(p-1)}\mathbf{1}_p\mathbf{1}'_p\right],$$

com $\mathbf{W}_j = \mathbf{y}_j - (0, \widetilde{\boldsymbol{\alpha}}')' - \mathbf{1}_p\, \widetilde{\mu}_x$, $\widetilde{\tau} = \dfrac{\widetilde{\phi}_x}{\widetilde{\phi} + p\, \widetilde{\phi}_x}$ e $\widetilde{\mu}_x, \widetilde{\alpha}_i, \widetilde{\phi}_x, \widetilde{\phi}$ como em (2.59).

Bibliografia

[1] Abdullah, M. B. (1995). Detection of influental observations in functional errors in variables model. *Communication in Statistics - Theory and Methods* 24, 6, 1585-1595.

[2] Aoki, R. (2001). Modelos de Regressão com Erros nas Variáveis com Intercepto Nulo. Tese de doutorado. IME - USP.

[3] Barnett, V. D. (1969). Simultaneous pairwise linear structural relationships. *Biometrics* 25, 129-142.

[4] Bedrick, E. J. (2001). An efficient scores test for comparing several measuring devices. *Journal of Quality Technology* 33, 1, 96 - 103.

[5] Billor, N. e Loynes, R. M. (1993). Local influence: a new approach. *Communication in Statistics - Theory and Methods* 22, 6, 1595 - 1611.

[6] Blackwood, L. G. e Bradley, E. L. (1991). An omnibus test for comparing two measuring devices. *Journal of Quality Technology* 23, 12 - 16.

[7] Bo-Cheng, W. (1998). *Exponential Family Nonlinear Models.* Springer. Singapore.

[8] Bolfarine, H. e Galea-Rojas, M. (1995). Functional comparative calibration using on EM algorithm - comment. *Biometrics* 51, 4, 1579 - 1580.

[9] Bolfarine, H. e Galea-Rojas, M. (1995). Maximum likelihood estimation of simultaneous pairwise linear structural relationships. *Biometrical Journal* 37, 673 - 689.

[10] Brindley, D. A. e Bradley, R. A. (1985). Some new results on Grubb's estimators. *Journal of the American Statistical Association* 80, 711 - 714.

[11] Cardigan, N. G. e Farrell, P. J. (1999). Expected local in the normallinear regression model. *Statistics e Probability Letters* 41, 25 - 30.

[12] Carter, R. (1981). Restricted maximum likelihood estimation of several bias and reliability in the comparison measuring methodos. *Biometrics* 37, 733 - 741.

[13] Chatterjje, S. e Hadi, A. S. (1988). *Sensitivity Analysis in Linear Regression.* John Wiley. New York.

[14] Choi, S. C. and Wette, R. (1972). A test for the homogeneity of variances among correlated variables. *Biometrics* 28, 589 - 592.

[15] Christensen, R. e Blackwood, L. (1993). Tests for precision and accuracy of multiple measuring devices. *Technometrics* 35, 411 - 420.

[16] Cook, R. D. (1986). Assessment of local influence. *Journal of the Royal Statistical Society* B 48, 133 - 169.

[17] Cook, R. D. (1987). Influence assessment. *Journal of Applied Statistics* 14, 117 - 131.

[18] Cook, R. D. e Weisberg, S. (1982). *Residuals and Influence in Regression.* Chapman and Hall. london.

[19] Dempster, A., Lair, N., and Rubin, D. B. (1977). Maximum likelihood estimation from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society* B, 39, 1 - 38.

[20] Deutler, T. (1991). Grubbs - type estimators for reproducibility variances in an inter - laboratory test study. *Journal of Quality Technology* 23, 324 - 335.

[21] Fuller, W. A. (1987). *Measurement error models.* John Wiley. New York.

[22] Fung, W. K. e Kwan, C. W. (1997). A note on local influence based on normal curvature. *Journal of the Royal Statistical Society* B, 59, 839 - 843.

[23] Galea, M., Bolfarine, H. e Castro, M. (1999). Local influence in comparative calibration models. *RT - MAE 9925*, IME - USP.

[24] Galea, M., Bolfarine, H. e Vilca-Labra, F. (2001). Influence diagnostics for structural erros-in-variables model under the student-t distribuition. *RP 14/01* , IMECC - UNICAMP.

[25] Galea, M., Paula, G. A. e Bolfarine, H. (1997). Local influence in elliptical regression models. *The Statistician* 46, 1, 71 - 79

[26] Galea-Rojas, M. (1995). *Calibração Estrutural e Funcional.* Tese de doutorado. IME - USP.

[27] Graybill, F. A. (1983). *Matrices with applications in Statistics.* Second edition. Wadsworth. California.

[28] Grubbs, F. E. (1948). On estimating precision of mesuring instruments and product variability. *Journal of the American Statistical Association* 43, 243-264.

[29] Grubbs, F. E. (1973). Errors of measurement precision, accuraccy and the statistical comparison of measuring instruments. *Technometrics* 15, 53 - 66.

[30] Grubbs, F. E. (1983). Grubb's estimator. *Encyclopedia of Statistical Sciences* 3, 542 - 549.

[31] Henderson, H. V. and Searle, S. R. (1979). Vec and Vech operators for matrices, with some uses in jacobians and multivariate statistics. *The Canadian Journal of Statistics* 7, 1, 65 - 81.

[32] Jaech, J. L. (1976). Large sample tests for Grubb's estimators of instrument precision with more than two instruments. *Technometrics* 18, 2, 127 - 133.

[33] Jaech, J. L. (1985). Statistical analysis of measurement errors. *Exxon Monographs.* John Wiley. New York.

[34] Johnson, R. A. e Wichern, D. W. (1998). *Applied Multivariate Statistical Analysis.* Prentice - Hall. New Jersey.

[35] Jorgensen, B. (1997). *The Theory of Dispersion Models.* Chapman and Hall. London.

[36] Kaaks, R., Riboli, E., Estive, J., van Kappel and van Staveren (1994). Estimating the accuracy of questionnaire assessments: validation in terms of structural equation models. *Statistics in Medicine* 13, 127 - 142.

[37] Kelly, G. (1984). The influence function in the errors in variables problem. *Annals of Statistics* 12, 87 - 100.

[38] Kim, M. G. (2000). Outliers and influential observations in the structural errors in variables model. *Journal of Applied Statistics* 27, 4, 451 - 460.

[39] Kimura, D. K. (1992). Functional comparative calibration using an EM algorithm. *Biometrics* 48, 1263 - 1271.

[40] Laurent, R. T. e Cook, R. D. (1993). Leverage, local influence and curvature in nonlinear regression. *Biometrika* 80, 1, 99 - 106.

[41] Lawrence, A. J. (1998). Regression transformation diagnostic using local influence. *Journal of the American Statistical Association* 83, 1067 - 1072.

[42] Lee, A. H. e Zhao, Y. (1996). Assessing local influence in measurement error models. *Biometrical Journal* 38, 7, 829 - 841.

[43] Lesaffre, E. e Verbeke G. (1998). Local influence in linear mixed models. *Biometrics* 54, 570 - 582.

[44] Loynes, R. M. (2001). A new measure in local influence. *Journal of Statistical Planning and Inference* 92, 47 - 53.

[45] Maloney, C. J. e Rastogi, S. C. (1970). Significance test for grubbs' estimators. *Biometrics* 26, 671 - 676.

[46] Muirhead, R. (1982). *Aspects of Multivariate Statistical Theory.* John Wiley. New York.

[47] Paula, G. A. (1993). Assessing local influence in restricted regressions models. *Computational Statistics and Data Analysis* 16, 63 - 79.

[48] Paula, G. A. (1995). Influence and residuals in restricted generalized linear models. *Journal of Statistical Computation and Simulation* 51, 315 - 331.

[49] Paula, G. A. (1996). Influence diagnostics in proper dispersion models. *Australian Journal of Statistics* 38, 307 - 316.

[50] Pitman, E. G. (1939). A note on normal correlation. *Biometrika* 31, 9 - 12.

[51] Poon, W. Y. and Poon, Y. S. (1999). Conformal normal curvature and assessment of local influence. *Journal of the Royal Statistical Society* B, 61, 51 - 61.

[52] Rao, C. R. (1948). Large sample tests of statistical hypotheses concerning several parameters with aplications to problems of estimation. *Proceeding of the Cambridge Philosophical Society* 44, 50 - 57.

[53] Seber, G. A. F. (1984). *Multivariate Observations.*John Wiley e Sons. New York.

[54] Sen, P. K. e Singer, J. M. (1993). *Large Sample Methods in Statistics.* An Introduction with Applications. Chapman e Hall. New York.

[55] Shyr, J. Y. e Gleser, L. J. (1986). Inference about comparative precision in linear structural relationships. *Journal of the Statistical Planning and Inference* 14, 339 - 358.

[56] Solano - Dávila, O. L. (2000). *Influência Local em Modelos de Calibração Comparativa.* Tese de mestrado. IMECC-UNICAMP.

[57] Tanaka, Y., Watadoni, S. e Moon, S. H. (1991). Influence in covariance structure analysis with an application to confirmatory fator analysis. *Communication in Statistic - Theory and Methods* 20, 12, 3805 - 3821.

[58] Theobald, C. M. e Mallison, J. R. (1978). Comparative calibration, linear structural relationship and congeneric measurement. *Biometrics* 34, 39 - 45.

[59] Thomas, W. e Cook, R. D. (1989). Assessing influence on regression coefficients in generalized linear models. *Biometrika* 76, 741 - 749.

[60] Thomas, W. e Cook, R. D. (1990). Assessing influence on prediction from generalized linear models. *Technometrics* 32, 1, 59 - 65.

[61] Verveke, G. and Molenberghs, G. (1997). *Linear Mixed Models in Practice.* Springer-Verlag. New York.

[62] Vilca-Labra, F., Arellano-Valle, R. B., Bolfarine, H. (1998). Elliptical functional models. *Journal of Multivariate Analysis* 65, 1, 36 - 57.

[63] Willians, E. J. (1969). Regression methods in calibration problems. *Bulletin of the International Statistical Institute* 43, 17 - 28.

[64] Wu, X. e Luo, Z. (1993). Second - order approach to local influence. *Journal of the Royal Statistics Society* B, 55, 929 - 936.