

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

**INFLUÊNCIA LOCAL EM MODELOS DE
CALIBRAÇÃO COMPARATIVA**

Olga Lidia Solano Dávila

Orientador: Prof. Dr. Filidor E. Vilca Labra

- Campinas, Maio de 2000 -



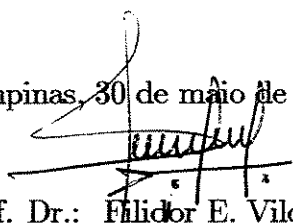
UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE

856670008

INFLUÊNCIA LOCAL EM MODELOS DE CALIBRAÇÃO COMPARATIVA

Este exemplar corresponde à redação final da dissertação devidamente corrigida e defendida por Olga Lidia Solano Dávila e aprovada pela comissão julgadora.

Campinas, 30 de maio de 2000.


Prof. Dr.: Filidor E. Vilca Labra

Banca Examinadora:

- 1 Filidor Edilfonso Vilca Labra
- 2 Heleno Bolfarine
- 3 Armando M. Infante
- 4 Hildete P. Pinheiro

Dissertação apresentado ao Instituto de Matemática, Estatística e Computação Científica, UNICAMP, como requisito parcial para obtenção do Título de MESTRE em ESTATÍSTICA.

UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE

**FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DO IMECC DA UNICAMP**

Solano Dávila, Olga Lidia

So41i Influência local em modelos de calibração comparativa / Olga
Lidia Solano Dávila – Campinas, [S.P. :s.n.], 2000.

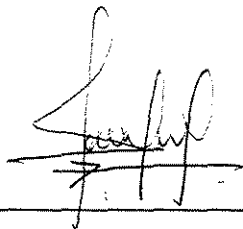
Orientador : Filidor E. Vilca Labra

Dissertação (mestrado) - Universidade Estadual de Campinas,
Instituto de Matemática, Estatística e Computação Científica.

I. Calibração. I. Vilca Labra, Filidor E. II. Universidade Estadual
de Campinas. Instituto de Matemática, Estatística e Computação
Científica. III. Título.

Dissertação de Mestrado defendida e aprovada em 30 de maio de 2000

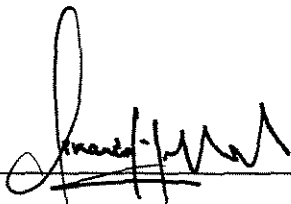
pela Banca Examinadora composta pelos Profs. Drs.



Prof (a). Dr (a). FILIDOR EDILFONSO VILCA LABRA



Prof (a). Dr (a). HELENO BONFARINE



Prof (a). Dr (a). ARMANDO MÁRIO INFANTE



Prof (a). Dr (a). HILDETE PRISCO PINHEIRO

UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE

Ao meu Deus, por tudo.
À toda a minha família
pelo apoio e compreensão.

UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE

Agradecimentos

- Gostaria de agradecer ao meu orientador Filidor Vilca pela colaboração, sugestões, apoio, amizade e orientação constante.
- Ao povo brasileiro, ao CAPES e à Unicamp (IMECC) pelo financiamento e condições de trabalho.
- Aos professores do Departamento de Estatística pelos ensinamentos e colaboração concedidos.
- Ao pessoal da secretaria e biblioteca pelo apoio e amizade demonstrados.
- A todos aqueles que de alguma forma, colaboraram à minha formação acadêmica e na execução da tese.

Muito Obrigada

Resumo

Neste trabalho estudamos alguns aspectos de diagnóstico em uma subclasse de modelos com erros nas variáveis; tais modelos são comumente usados para comparar instrumentos ou métodos de medição. Este problema é conhecido na literatura como Calibração Comparativa.

Consideramos inicialmente um estudo de análise de influência no modelo de regressão linear e do método de influência local proposta por Cook (1986).

Apresentamos um estudo de inferência no modelo de calibração comparativa sob a parametrização de análise de fatores. Estimativas de máxima verossimilhança são apresentadas e um teste assintótico é proposto utilizando a estatística de Wald.

Utilizando o método de influência local proposta por Cook (1986), apresentamos um estudo de diagnóstico no modelo de calibração comparativa considerando alguns esquemas de perturbação. Apresentamos uma ilustração da metodologia proposta com conjuntos de dados apresentados na literatura.

Abstract

In this work we consider some diagnostics aspects in a special sub-class of a general models with measurements of erros; such models are specially designed to compare measuring devices. This problem is known in the literature as Comparative Calibrations. First, we considered a study of analysis of influence in linear regression models and the method of local influence proposed by Cook (1986).

We introduce a study of inference to comparative calibration model on parametrization of factor analysis model. Estimates of maximum likelihood are considered and an asymptotic test is proposed based in the Wald statistic.

Using the method of local influence proposed by Cook (1986) and considering some perturbation schemes, we present an analysis of diagnostic in the comparative calibration model.

An illustration of the methodology is proposed with a sets of data taken from the literature.

Conteúdo

1 Introdução	1
1.1 Formulação do Problema.....	1
1.2 O Problema de Calibração Comparativa.....	5
1.3 Objetivo do Trabalho.....	8
1.4 Organização do Trabalho.....	8
2 Análise de Influência	10
2.1 Introdução.....	10
2.2 Diagnóstico em Modelos de Regressão.....	10
2.2.1 A Distância de Cook.....	17
2.2.2 Medida de DFFITS.....	18
2.2.3 Afastamento pela Verossimilhança para uma Observação....	18
2.3 Afastamento pela Verossimilhança no Caso Geral.....	30
2.3.1 Alguns Tipos de Perturbações.....	33
2.4 Influência Local.....	35
2.4.1 Aplicação.....	48
3 Modelo de Calibração Comparativa	55
3.1 Introdução.....	55
3.2 Modelo de Calibração Comparativa.....	56
3.3 O Modelo Estrutural de Regressão de Calibração.....	58
3.3.1 Definição de Confiabilidade e Precisão.....	61
3.3.2 Estimadores de Máxima Verossimilhança.....	63
3.4 O Modelo Estrutural de Calibração de Análise de Fatores.....	65
3.4.1 Estimadores de Máxima Verossimilhança.....	66
3.4.2 Matriz de Informação Observada.....	68
3.4.3 Matriz de Informação Esperada.....	75
3.4.4 Testes de Hipóteses.....	83
3.5 Aplicação.....	85

Lista de Figuras

2-1	Gráfico de dispersão e retas ajustadas.....	14
2-2	$LD_i(\beta/s^2)$ vs observações.....	26
2-3	$LD_i(\beta, s^2)$ vs observações	26
2-4	$LD_i(s^2/\beta)$ vs observações.....	27
2-5	Ilustração esquemática do afastamento pela verossimilhança.....	31
2-6	Gráfico de influência em R^3	41
2-7	Gráfico de dispersão dos dados de Ganso.....	54
2-8	Gráfico de $d_{\max}$ vs contagem do observador.....	54
3-1	Método gráfico para a medição do volume testicular.....	88
3-2	Gráfico da distância transformada para o modelo normal.....	91
4.1	Gráfico dos três métodos vs observações	110
4-2	Gráfico das medições do testículo direito.....	110
4-3	Perturbação de ponderação de casos.....	111
4-4	Perturbação aditiva no método 1.....	114
4-5	Perturbação aditiva no método 3	116
4-6	Perturbação multiplicativa no método 1.....	119
4-7	Perturbação multiplicativa no método 3.....	121
4-8	Perturbação de ponderação de casos quando o parâmetro de interesse é $\theta_1 = (\mu^T, \lambda^T)^T$	124
4-9	Perturbação de ponderação de casos quando o parâmetro de interesse é $\theta_2 = \phi$	126
4-10	Gráfico da capacidade vital para 72 pacientes, feitas por três instrumentos.....	131
4-11	Gráfico das medições da capacidade vital para 72 pacientes.....	131

Lista de Tabelas

2.1	Dados Hipotéticos.....	13
2.2	Dados de Clube de Saúde.....	24
2.3	Medidas baseadas sobre a função de Verossimilhança.....	25
2.4	Medidas de influência para os dados (a) de Anscombe (1973).....	28
2.5	Medidas de influência para os dados (b) de Anscombe (1973).....	28
2.6	Medidas de influência para os dados (c) de Anscombe (1973).....	29
2.7	Medidas de influência para os dados (d) de Anscombe (1973).....	29
2.8	Absorção de éter pelo fígado de ratos (Weisberg, 1985, p.122).....	39
2.9	Estimativa do número de gansos brancos pelo primeiro observador em Weisberg (1985).....	52
2.10	Autovetor correspondente ao maior autovalor de $\Delta^T \left(\tilde{\mathbf{L}} \right)^{-1} \Delta$ nos dados de ganso.....	53
3.1	Medidas do volume do testículo direito humano para 42 adolescentes entre 11 e 17 anos de idade, feita por três métodos de medição. Chipkevitch et al.(1996).....	89
3.2	Valores de z_j e $\tilde{d}_j$ do volume de testículo humano para 42 adolescentes entre 11 e 17 anos de idade, feitas por três métodos de medição	92
4.1	Elementos de $d_{\max}$ em valor absoluto correspondente às medidas do volume do testículo direito humano em 42 adolescentes, feitas por três métodos de medição em perturbação de ponderação de casos	112
4.2	Estimativa dos parâmetros.....	112
4.3	Elementos de $d_{\max}$ em valor absoluto correspondente às medidas do volume do testículo direito humano em 42 adolescentes, feitas por três métodos de medição em perturbação aditiva no método 1.....	115

4.4	Estimativa dos parâmetros.....	115
4.5	Elementos de $d_{\max}$ em valor absoluto correspondente às medidas do volume do testículo direito humano em 42 adolescentes, feitas por três métodos de medição em perturbação aditiva no método 3.....	117
4.6	Estimativa dos parâmetros.....	118
4.7	Elementos de $d_{\max}$ em valor absoluto correspondente às medidas do volume do testículo direito humano em 42 adolescentes, feitas por três métodos de medição em perturbação multiplicativa no método 1.....	120
4.8	Estimativa dos parâmetros.....	120
4.9	Elementos de $d_{\max}$ em valor absoluto correspondente às medidas do volume do testículo direito humano em 42 adolescentes, feitas por três métodos de medição em perturbação multiplicativa no método 3.. ..	122
4.10	Estimativa dos parâmetros.....	122
4.11	Elementos de $d_{\max}$ em valor absoluto correspondente às medidas do volume do testículo direito humano em 42 adolescentes, feitas por três métodos de medição, quando o parâmetro de interesse é $\theta_1 = (\mu^T, \lambda^T)^T$	125
4.12	Estimativa dos parâmetros.....	125
4.13	Elementos de $d_{\max}$ em valor absoluto correspondente às medidas do volume do testículo direito humano em 42 adolescentes, feitas por três métodos de medição, quando o parâmetro de interesse é $\theta_2 = \phi$	127
4.14	Medidas da capacidade vital para 72 pacientes, feitas por 3 instrumentos.....	129
4.15	Estimativas dos parâmetros.....	132
4.16	Elementos de $d_{\max}$ em valor absoluto correspondente aos dados Barnett (1969), em perturbação de ponderação de casos.....	133
4.17	Elementos de $d_{\max}$ em valor absoluto correspondente aos dados Barnett (1969), em perturbação aditiva	135
4.18	Estimativa dos parâmetros.....	136
4.19	Elementos de $d_{\max}$ em valor absoluto correspondente aos dados	

Capítulo 1

Introdução

1.1 Formulação do Problema

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. Por causa destas faltas de exatidão, 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 resultados que foram observados, a utilização do modelo pode conduzir a 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 dos modelos, e as estimativas dos seus parâmetros, ou seja, fazer uma *análise de diagnóstico*, que consiste de métodos para avaliar o grau de sensibilidade das inferências a pequenas perturbações nos dados ou no

modelo proposto.

Por exemplo, em um estudo de análise de regressão, freqüentemente o conjunto de dados contém algumas observações extremas ou discrepantes em relação ao restante dos dados. Durante a construção de um modelo de regressão é importante examinar os pontos que estão sendo usados, com o objetivo de determinar se existe alguns deles que esteja controlando propriedades importantes no modelo.

Se os pontos influentes detectados tivessem surgido como resultados 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 excluí-los do conjunto de dados.

Por outro lado, pode não ter havido nada de errado durante a obtenção destes pontos, mas mesmo assim é importante descobrir se estes controlam propriedades importantes do modelo de regressão, já que a ocorrência deste fato pode limitar o uso do modelo.

Na análise de regressão com uma ou duas variáveis explanatórias é fácil identificar observações extremas por meio de diagrama de dispersão ou gráficos de resíduos, e estuda assim sua influência no ajuste do modelo. Mas, quando mais de duas variáveis explanatórias estão envolvidas no estudo de análise de regressão, a identificação de pontos extremos e a avaliação de sua influência por meio de métodos gráficos simples não são fáceis, sendo necessário empregar procedimentos especiais de detecção.

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.

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

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.

Um dos métodos de diagnóstico empregado em modelos de regressão utiliza a estratégia da exclusão de casos, e consiste em identificar observações que influenciam, de maneira desmedida, um aspecto particular da análise.

Dentre as medidas geralmente usadas em regressão linear para detectar influência de observações sobre o modelo ajustado, duas se destacam:

1. A distância D_i sugerida por Cook (1977), conhecido como um dos precursores do estudo de influência em modelos de regressão. A distância de Cook consiste em comparar o valor ajustado para a variável resposta, quando a i -ésima observação é excluída da análise, com o valor ajustado obtido considerando o ajuste com todas as observações.
2. Os $DFITS_i$, são outras medidas de influência que usam a estratégia da exclusão de casos; foram proposta por Belsley, Kuh e Welsch (1980).

Cook, Peña e Weisberg (1988) comparam as duas medidas e mostram que as mesmas não medem as mesmas coisas e portanto podem ser usadas conjuntamente.

O problema de detecção de influência e avaliação da exclusão de algumas observações na estimação dos parâmetros no modelo de regressão têm sido investigado por muitos autores, dentre os quais citamos Cook (1979), Cook e Weisberg (1982), Johnson e Geisser (1982) e Polasek (1984).

Uma noção mais geral para avaliar a influência de pequenas perturbações nos modelos é proposta 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 permite avaliar o efeito da exclusão de alguma observação na estimação dos parâmetros.

O método proposto por Cook pode ser aplicado em uma grande variedade de modelos estatísticos, daí a relevância do seu estudo. 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 as predições nos modelos lineares generalizados, Paula (1996) faz uma extensão desse estudo para os modelos próprios de dispersão, propostos por Jorgensen (1997), Lee e Zhao (1996) consideram o método de influência local em modelos lineares generalizados quando as covariáveis são medidas com erros, Lesaffre e Verbeke (1998) consideram o método no efeito do modelo linear misto e 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.

Kim (1995) generaliza o método de influência local para modelos de regressão multivariada, considerando somente perturbação de ponderação de casos para todos os parâmetros de interesse, Fung e Tang (1997) consideram esquemas de perturbação nas variáveis explicativas e nas variáveis de respostas, Kim (1996) investiga a influência de casos sobre o estimador de máxima verossimilhança dos parâmetros de transformação em um modelo de regressão multivariado, Jiancheng e Xizhi (1998) usa o método sobre o modelo de Hazard proporcional de Cox e Galea, Paula e Bolfarine (1997) aplicam o método no modelo de regressão linear elíptico.

Na literatura existente o método de influência local sob o ponto de vista bayesiano, por exemplo, Pan, Fang e Liski (1996) utilizam a aproximação de influência local bayesiana

para diagnosticar a adequação do modelo da curva de crescimento com estrutura de covariância simples de Rao.

Lavine (1992) trabalhou em influência local de prognóstico em modelos lineares bayesianos com conjugadores a priori.

Recentemente Zhao e Lee (1998) e Kwan e Fung (1998) aplicaram o método de influência local em modelos de equações simultâneas e em análise de fatores, respectivamente, sob as suposições de normalidade. No entanto nenhuma das técnicas anteriores tem sido desenvolvida para modelos de calibração comparativa.

1.2 O Problema de Calibração Comparativa

Calibração estatística é bastante importante em aplicações práticas de diversas áreas.

Na física e 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 que são indicadores do estado de saúde individual (termômetros, medidores de pressão sanguínea, 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 cromatógrafos, etc.).

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

Em qualquer discussão sobre calibração é importante distinguir 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. Na calibração comparativa, pretende-se determinar a relação ente diversos testes ou instrumentos que fornecem medidas indiretas similares, não existindo, medidas padrões como referência. 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 calibração comparativa.

Na calibração absoluta, uma técnica para 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 Williams (1969), correspondia à medida do calibre (ou diâmetro interno) de um tubo.

Neste trabalho, será examinada essencialmente a calibração comparativa.

O problema da comparação de p instrumentos ou métodos de medição, quando cada um deles é usado para medir uma característica em um grupo comum de unidades experimentais, é conhecido na literatura estatística como *calibração comparativa*.

O termo método ou instrumento de medição refere-se às diferentes formas de medição de uma característica dada. O termo também pode ser aplicado quando a mesma forma de medição é feita por diferentes aparelhos ou diferentes operadores. Por exemplo, quando uma análise química é realizada em diferentes laboratórios, cada laboratório é identificado como um método ou instrumento de medição. Em alguns casos pode acontecer também, que os diferentes métodos sejam um mesmo instrumento avaliado em diferentes tempos.

A necessidade de comparar instrumentos de medição que diferem em custo, velocidade ou outros fatores tem aparecido com frequência em diversas áreas de conhecimento.

Grubbs (1948, 1973), por exemplo, compara três cronômetros. Barnett (1969) dá um exemplo onde quatro combinações instrumento-operador, concebidas para medir a capacidade vital em um grupo de pacientes, são avaliados. Leurgans (1980) compara dois métodos para medir a concentração da glicose no sangue. Vários outros exemplos, na área médica, são mencionados por Kelly (1984, 1985) e Kaaks et al. (1994). Já Jaech

(1985) dá vários exemplos na área industrial, Fuller (1987) apresenta exemplos na área da agricultura e Dunn (1992) dá aplicações em psicologia e educação.

Suponhamos que dispomos de p ($p \geq 2$) instrumentos ou métodos de medição 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 -ésima, $i = 1, \dots, p$, $j = 1, \dots, n$.

Um modelo frequentemente adaptado na literatura (veja, Jaech, 1964, Cochran, 1968, Barnett, 1969 e Williams, 1969) para este tipo de situação:

$$Y_{ij} = \alpha_i + \beta_i x_j + e_{ij}, \quad i = 1, \dots, p, \quad j = 1, \dots, n,$$

onde os erros aleatórios de medição e_{ij} são não-correlacionados e para cada $i = 1, \dots, p$, $e_{i1}, \dots, e_{in}$ são identicamente distribuídos com

$$E(e_{ij}) = 0 \quad \text{e} \quad \text{Var}(e_{ij}) = \sigma_i^2, \quad i = 1, \dots, p.$$

Além disso os parâmetros α_i e β_i representam

α_i : vício aditivo do instrumento i .

β_i : vício multiplicativo do instrumento i , $i = 1, \dots, p$.

Estes modelos correspondem a uma subclasse de modelos com erros nas variáveis (veja Fuller, 1987), logo podemos ter as versões estrutural e funcional do modelo de calibração comparativa.

Em termos gerais, o problema de comparar instrumentos de medição, reduz-se a fazer inferências sobre os parâmetros do modelo.

1.3 Objetivo do Trabalho

O objetivo deste trabalho é considerar um estudo de influência e diagnóstico no modelo estrutural normal de calibração comparativa, baseado no método proposto por Cook(1986). Para o desenvolvimento do método de diagnóstico no modelo de calibração comparativa consideramos a parametrização de análise de fatores introduzida por Theobald e Mallison (1978). Assim, os objetivos específicos deste trabalho podem ser resumidos como segue:

- 1) Apresentar um estudo do método de influência local proposto por Cook(1986), baseado no modelo de regressão linear. Avaliar a influência de uma observação utilizando os resíduos ordinários e a função de verossimilhança.
- 2) Apresentar um estudo de inferência sob o modelo de calibração comparativa, baseado na parametrização de análise de fatores do modelo. Obter a matriz de informação observada de Fisher, que será utilizada no desenvolvimento do método de diagnóstico para o modelo de calibração comparativa.
- 3) Aplicar o método de influência local proposto por Cook (1986) ao modelo de calibração comparativa, sob a parametrização de análise de fatores; propor alguns esquemas de perturbações (ponderação de casos e perturbação na observação).

1.4 Organização do Trabalho

Este trabalho consta de 5 capítulos e 2 apêndices. No capítulo 2 tratamos do problema de diagnóstico em modelos de regressão, com ênfase no método de influência local, através do afastamento pela verossimilhança (tradução livre de “likelihood displacement”, Cook e Weisberg (1982)).

O capítulo 3 está dedicado ao estudo do modelo de calibração comparativa para três instrumentos baseado no modelo proposto por Barnett (1969) e Williams (1969) a qual é apresentado como o modelo estrutural de regressão de calibração e o modelo estrutural de calibração de análise de fatores, proposto por Theobald e Mallison (1978). Utilizamos

este último modelo para achar a matriz de informação observada e esperada de Fisher. A estatística de Wald é proposta para testar hipóteses de interesse.

No capítulo 4 discutimos diferentes esquemas de perturbação para o modelo de análise de fatores e apresentamos um estudo de influência local usando esse modelo. Também fazemos aplicações dos resultados obtidos a um conjunto de dados reais.

O capítulo 5 está dedicado a nossas conclusões e alguns direcionamentos para estudos subseqüentes.

Finalmente, nos apêndices resumimos alguns resultados da literatura utilizados na tese e também apresentamos teoremas, lemas e corolários usados para obter a matriz de informação e a estatística de Wald.

Capítulo 2

Análise de Influência

2.1 Introdução

Neste capítulo apresentamos um estudo do problema de diagnóstico no modelo de regressão linear com especial atenção ao método de influência local proposto por Cook(1986).

Tal método é baseado no uso da função de verossimilhança.

O estudo de diagnóstico será baseado em análise de resíduos e a função de afastamento pela verossimilhança. Este último conceito é utilizado para desenvolver o método de influência local proposto por Cook (1986).

2.2 Diagnóstico em Modelos de Regressão

Considere o modelo de regressão linear definido por

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \quad (2.1)$$

onde $\mathbf{Y} = (Y_1, \dots, Y_n)^T$ é um vetor aleatório, $\mathbf{X} = (x_{ij})$ é uma matriz $n \times p$ não aleatória de posto menor que n , $p \leq n$, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^T$ é um vetor de parâmetros (de regressão)

desconhecidos e $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)^T$ é um vetor de erros aleatórios (não observáveis) tal que

$$E(\varepsilon) = 0 \quad \text{e} \quad \text{Var}(\varepsilon) = \sigma^2 \mathbf{I}_n,$$

onde $\mathbf{I}_n$ denota a matriz de identidade de ordem n .

O estimador de mínimos quadrados do vetor de parâmetros β é dado pela solução do sistema de equações normais

$$\mathbf{X}^T(\mathbf{Y} - \mathbf{X}\hat{\beta}) = 0. \quad (2.2)$$

Se a matriz $\mathbf{X}$ for de posto completo, isto é, $\text{posto}(\mathbf{X}) = p$, então o estimador de mínimos quadrados do vetor β fica dado por

$$\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}. \quad (2.3)$$

Se o posto da matriz $\mathbf{X}$ for de posto menor que p , isto é, $\text{posto}(\mathbf{X}) < p$, a solução do sistema de equações normais não é única, e neste caso, substituímos $(\mathbf{X}^T \mathbf{X})^{-1}$ em (2.3) por uma inversa generalizada (g-inversa), com a qual são obtidos resultados análogos ao caso em que se tem posto completo (veja, por exemplo, Rao, 1973). No decorrer deste trabalho assumiremos que a matriz $\mathbf{X}$ é de posto completo e $\varepsilon \sim N_n(0, \sigma^2 \mathbf{I}_n)$. Assim utilizando as propriedades da distribuição normal

$$\mathbf{Y} \sim N_n(\mathbf{X}\beta, \sigma^2 \mathbf{I}_n) \quad \text{e} \quad \hat{\beta} \sim N_p(\beta, \sigma^2 (\mathbf{X}^T \mathbf{X})^{-1}).$$

Se $\hat{\mathbf{Y}}$ denota o vetor de valores ajustados correspondente ao vetor de valores observados $\mathbf{Y}$, então

$$\hat{\mathbf{Y}} = \mathbf{X}\hat{\beta} = \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y} = \mathbf{H}\mathbf{Y}, \quad (2.4)$$

onde $\mathbf{H} = \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T$ é a matriz de projeção ortogonal sobre o espaço vetorial gerado pelas colunas da matriz $\mathbf{X}$. Na literatura, $\mathbf{H}$ é conhecida como matriz *chapéu*, uma vez

que transforma o vetor de valores observados $\mathbf{Y}$ no vetor $\hat{\mathbf{Y}}$ de valores ajustados Weisberg (1985).

O vetor de resíduos ordinários é dado por

$$\mathbf{e} = \mathbf{Y} - \hat{\mathbf{Y}} = (\mathbf{I}_n - \mathbf{H})\mathbf{Y}. \quad (2.5)$$

Das propriedades da distribuição normal

$$\mathbf{e} \sim N_n(\mathbf{0}, \sigma^2(\mathbf{I}_n - \mathbf{H}))$$

e

$$\frac{\mathbf{e}^T \mathbf{e}}{\sigma^2} \sim \chi_{(n-p)}^2.$$

Uma estimativa não viesada para a variância σ^2 é dada por

$$\hat{\sigma}^2 = s^2 = \frac{\mathbf{e}^T \mathbf{e}}{n - p}$$

e resíduos “studentizados” são obtidos por

$$t_i = \frac{e_i}{\{s^2(1 - h_{ii})\}^{1/2}},$$

onde $h_{11}, \dots, h_{nn}$ são os elementos da diagonal principal da matriz $\mathbf{H}$.

Se denotamos as linhas da matriz $\mathbf{X}$ por $\mathbf{x}_1^T, \dots, \mathbf{x}_n^T$, de (2.1) temos que

$$Y_j = \mathbf{x}_j^T \boldsymbol{\beta} + \varepsilon_j, \quad j = 1, \dots, n.$$

Assim, o (i, j) -ésimo elemento de $\mathbf{H}$ é dado por

$$h_{ij} = \mathbf{x}_i^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{x}_j, \quad i, j = 1, \dots, n.$$

Como ilustração, mostraremos um exemplo fictício, apresentado por Anscombe (1973) e Weisberg(1985), onde quatro conjuntos de dados (a,b,c,d) (Tabela 2.1) fornecem as mesmas estimativas de mínimos quadrados para os parâmetros da reta ajustada.

Tabela 2.1: Dados Hipotéticos

Nro. de Observações	a		b		c		d	
	x	y	x	y	x	y	x	y
1	10	8,04	10	9,14	10	7,46	8	6,58
2	8	6,95	8	8,14	8	6,77	8	5,76
3	13	7,58	13	8,74	13	12,70	8	7,71
4	9	8,81	9	8,77	9	7,11	8	8,84
5	11	8,33	11	9,26	11	7,81	8	8,47
6	14	9,96	14	8,10	14	8,84	8	7,04
7	6	7,24	6	6,13	6	6,08	8	5,25
8	4	4,26	4	3,10	4	5,39	8	12,50
9	12	10,80	12	9,13	12	8,15	8	5,56
10	7	4,82	7	7,26	7	6,42	8	7,91
11	5	5,68	5	4,74	5	5,73	8	6,89

Fonte: Anscombe (1973).

Uma análise visual desses conjuntos de dados sugere comportamentos diferentes; no entanto, se assumirmos um relacionamento linear $Y = \beta_0 + \beta_1 X$ entre as variáveis X e Y , um ajuste por mínimos quadrados fornece as mesmas estimativas para os parâmetros β_0 e β_1 nos quatro casos, ou seja, $\hat{Y} = 3 + 0.5X$, com uma estimativa comum do desvio padrão $s = \sqrt{1.53}$ e o coeficiente de determinação $R^2 = 0,67$. Com estes resultados tudo leva a crer que esse modelo de regressão linear é apropriado para os quatro casos. Uma análise gráfica utilizando diagramas de dispersão para as quatro situações mostra que o ajuste de um modelo de regressão linear simples não é adequado em todos os casos apresentados, como se observa nos gráficos da Figura 2.1. Essa figura mostra o

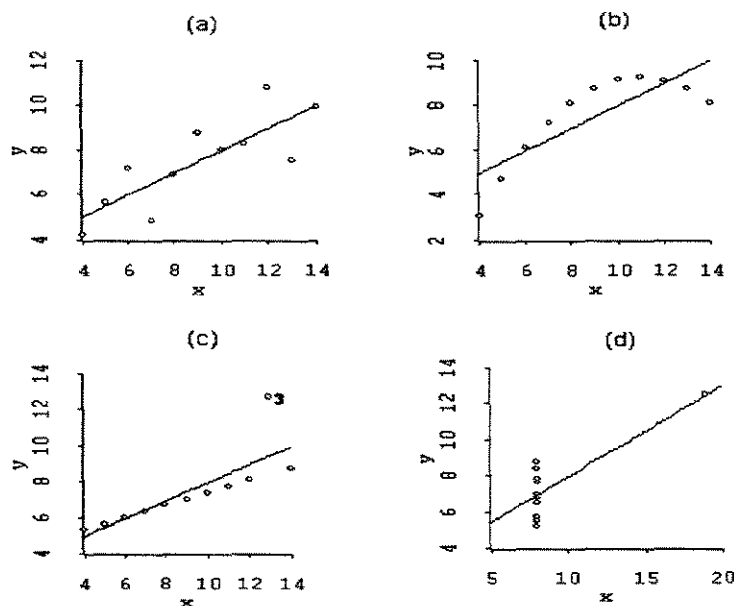


Figura 2-1: Gráfico de dispersão e retas ajustadas

diagrama de dispersão, para os quatro conjuntos de dados, assim como a correspondente reta ajustada. Analisando esses gráficos pode-se concluir:

1. Para o conjunto de dados (a) não há indícios de problemas de ajuste.
2. Para o conjunto de dado (b), falta uma componente quadrática em X no modelo, ou seja, o modelo linear simples não é adequado.
3. A terceira observação do conjunto (c) não está bem ajustada, havendo indícios de ser uma observação aberrante.
4. As estimativas dos parâmetros para o conjunto de dados (d) são altamente influenciadas pela oitava observação.

No conjunto de dados (d) toda a informação da inclinação da linha de regressão reside na oitava observação. Se essa observação fosse excluída a inclinação não poderia ser estimada.. No conjunto de dados (a), se alguma observação fosse desacreditada e excluída

do conjunto de dados, a análise com as observações restantes mostrariam os mesmos resultados. O que não aconteceria no conjunto de dados (d). Cada conjunto de dados b,c e d ilustra um efeito peculiar de uma forma extrema. Numa forma extrema menor tais efeitos são às vezes encontradas na análise estatística. Para o efeito do conjunto de dados (d), Anscombe (1973) mostra um exemplo.

Os resultados da análise de variância são similares nos quatro conjunto de dados.

Estas situações mostram a necessidade de se fazer uma análise minuciosa sobre vários aspectos do ajuste de um modelo de regressão.

Sabemos que o estimador de mínimos quadrados é bastante sensível a :

- Observações cujos resíduos ordinários são elevados, as quais são ditas pobremente ajustadas.
- Pontos extremos no espaço gerado pela matriz de especificação do modelo Pregibon(1981).

Para o i -ésimo resíduo ordinário definido na equação (2.5), tem-se que $Var(e_i) = \sigma^2(1 - h_{ii})$, e Cook e Weisberg (1982, p.12) mostram que $\frac{1}{n} \leq h_{ii} \leq \frac{1}{c}$, onde c representa o número de linhas na matriz X que são iguais a x_i^T . Assim, quando h_{ii} se aproxima de seu limite superior, tem-se que a $Var(e_i)$ torna-se pequena, e portanto o resíduo e_i fica próximo de seu valor esperado, que é zero. Com isto, segue-se que $\hat{Y}_i \approx Y_i$. Neste caso, sabe-se de antemão que a i -ésima observação deverá ter um resíduo pequeno, independente do valor observado para a variável resposta.

A exposição acima pode também ser feita algebricamente, usando as propriedades da matriz *chapéu*. Basta lembrar que H só depende de X e que é uma matriz de projeção, sendo portanto, simétrica e idempotente. Assim temos que

$$\begin{aligned} h_{ii} &= \sum h_{ij}h_{ji} \\ h_{ii} &= \sum h_{ij}^2 \end{aligned}$$

$$h_{ii} = h_{ii}^2 + \sum_{j \neq i} h_{ij}^2.$$

E daí segue-se que $h_{ii}(1 - h_{ii}) = \sum_{j \neq i} h_{ij}^2 \geq 0$. Portanto, se o elemento h_{ii} for próximo de 1, os demais elementos de sua linha serão próximos de zero. Fazendo uma decomposição na expressão do valor ajustado, equação (2.4) $\hat{Y}_i = h_{ii}Y_i + \sum h_{ij}Y_j$, vê-se que, se h_{ii} é próximo a 1, o valor observado Y_i irá influenciar fortemente a estimativa do próprio valor ajustado $\hat{Y}_i$.

O elemento h_{ii} é conhecido na literatura como *leverage* do i -ésimo caso, por representar o poder do impacto que uma observação Y_i exerce sobre o próprio valor ajustado $\hat{Y}_i$.

Quando uma matriz X é de posto completo tem-se que $\sum_{i=1}^n h_{ii} = p$. Baseado nisso, Hoaglin e Welsch (1978) sugerem que sejam considerados pontos com alto *leverage* aquelas observações cujos h_{ii} excedam o valor $2p/n$.

Com relação aos elementos que estão fora da diagonal principal da matriz H , eles trazem informação sobre a influência que dois casos exercem conjuntamente sobre o ajuste do modelo. Em particular, h_{ij} pode ser interpretado como a influência que Y_i exerce sobre $\hat{Y}_j$, e também a influência que Y_j exerce sobre $\hat{Y}_i$.

Pelo que vimos anteriormente, os elementos $h_{11}, \dots, h_{nn}$ são úteis para detectar os pontos extremos e a influência exercida por cada observação sobre a respectiva predição.

Atualmente tem-se várias medidas de diagnóstico para detectar observações que exercem influência sobre outros aspectos de ajuste, tais como estimativas dos parâmetros; em geral essas medidas combinam informações contidas tanto em h_{ii} quanto no resíduo ordinário e_i .

No que segue desta seção vamos considerar algumas medidas para avaliar a influência que uma observação exerce sobre as estimativas dos parâmetros. As medidas que vamos considerar são a distância de Cook, proposta por Cook (1977), a medida DFFITS que foi proposta por Belsley, Kuh e Welsh (1980) e o afastamento pela verossimilhança, esta última medida foi considerada por Cook e Peña (1988). Outras medidas de diagnóstico

podem ser encontradas por exemplo, em Chatterjee e Hadi (1988).

2.2.1 A Distância de Cook

Na análise de um modelo normal linear, quando a i -ésima observação é excluída do ajuste, a distância de Cook é definida como

$$D_i = \frac{(\hat{\beta} - \hat{\beta}_{(i)})^T (\mathbf{X}^T \mathbf{X}) (\hat{\beta} - \hat{\beta}_{(i)})}{ps^2} \quad (2.6)$$

onde (i) significa que a estimativa do vetor β está sendo feita sem a i -ésima observação. Essa medida está baseada em elipsóides de confiança, e sua análise é feita geralmente através de um “*index plot*”. Os valores D_i são comparados com o quantil da distribuição $F_{p,(n-p)}$, em uma analogia com regiões de confiança.

Usando álgebra matricial, podemos expressar a nova estimativa de β como $\hat{\beta}_{(i)} = \hat{\beta} - \frac{e_i}{1-h_{ii}} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{x}_i$, e com esta relação a distância de Cook pode ainda ser escrita na forma

$$D_i = \left\{ \frac{e_i}{s(1-h_{ii})^{1/2}} \right\}^2 \left(\frac{h_{ii}}{1-h_{ii}} \right) \frac{1}{p}, \quad (2.7)$$

onde $s^2 = \mathbf{e}^T \mathbf{e} / (n-p)$.

Da equação (2.7) podemos observar que D_i será grande quando o resíduo “Studentizado” $t_i = \frac{e_i}{s(1-h_{ii})^{1/2}}$ for grande e/ou h_{ii} for próximo de um. Nesse caso, dizemos que o ponto é ao mesmo tempo aberrante e influente. Para uma observação com e_i grande e h_{ii} pequeno (próximo de $1/n$), pode acontecer que D_i não detecte esse ponto. Por essa razão, a distância de Cook pode não ser adequada em situações como essa. Esse problema não existe com a medida $DFFITs_i$, que é definida a seguir.

2.2.2 Medida de DFFITS

Esta medida permite avaliar a influência de uma observação individual sobre a estimativa da variância σ^2 . Foi proposta por Belsley, Kuh e Welsch (1980), e é definida como sendo

$$\begin{aligned} DFFITS_i &= \sqrt{\frac{(\hat{\beta} - \hat{\beta}_{(i)})^T (\mathbf{X}^T \mathbf{X}) (\hat{\beta} - \hat{\beta}_{(i)})^T}{s_{(i)}^2}} \\ &= \frac{|e_i|}{s_{(i)}(1 - h_{ii})^{1/2}} \left(\frac{h_{ii}}{1 - h_{ii}} \right)^{1/2}, \end{aligned} \quad (2.8)$$

onde $s_{(i)}^2$ é um estimador da variância σ^2 , quando a i -ésima observação é excluída.

A medida $DFFITS_i$ é analisada graficamente para cada observação e deve-se colocar especial atenção para aquelas observações que tem a propriedade $DFFITS_i \geq 2\sqrt{p/(n-p)}$. As medidas D_i e $DFFITS_i$ podem ser unificadas através do afastamento pela verossimilhança, como foi apontado por Cook, Peña e Weisberg (1988).

2.2.3 Afastamento pela Verossimilhança para uma Observação

Outra maneira de avaliar a influência que uma observação exerce sobre a estimativa dos parâmetros é utilizando a função de log-verossimilhança. Esta forma alternativa é proporcionada pelo afastamento da verossimilhança (distância), que foi denominada por Cook e Weisberg (1982) como “likelihood displacement”.

Seja θ um vetor de parâmetros $p \times 1$ particionado como

$$\theta = (\theta_1^T, \theta_2^T)^T, \quad (2.9)$$

onde θ_1 é $p_1 \times 1$ e θ_2 é $p_2 \times 1$, $p_1 + p_2 = p$, e seja $l(\theta) = l(\theta_1, \theta_2)$ a função de log-verossimilhança que fornece toda a informação que existe na amostra com respeito aos parâmetros, desde que a escolha do modelo seja verdadeira.

Em análise de influência freqüentemente desejamos comparar o EMV (Estimador de Máxima Verossimilhança) $\tilde{\theta}$ de θ utilizando todas as observações e o EMV $\tilde{\theta}_{(i)} = (\tilde{\theta}_{1(i)}^T, \tilde{\theta}_{2(i)}^T)^T$

obtido sem considerar a i -ésima observação.

Um método geral para comparar $\tilde{\theta}$ e $\tilde{\theta}_{(i)}$ é baseado no afastamento pela verossimilhança definido por

$$LD_i(\theta) = 2 \left[l(\tilde{\theta}) - l(\tilde{\theta}_{(i)}) \right]. \quad (2.10)$$

O uso da expressão em (2.10) fornece uma métrica fixa para comparar valores de LD para as diferentes observações. Outras medidas alternativas usando a função de log-verossimilhança podem ser encontradas em Cook, Peña e Weisberg (1988).

Se $LD_i(\theta)$ é grande, a i -ésima observação é chamada influente, pois a exclusão desta pode causar importantes mudanças nas conclusões.

A medida fornecida pela relação em (2.10) pode ser interpretada em termos da região de confiança assintótica.

$$\left\{ \theta \in \mathbb{R}^p : 2 \left[l(\tilde{\theta}) - l(\theta) \right] \leq \chi_{(1-\alpha, p)}^2 \right\},$$

onde $\chi_{(1-\alpha, p)}^2$ é o percentil superior de tamanho $1 - \alpha$ da distribuição qui-quadrado com p graus de liberdade (veja, Cox e Hinkley, 1974, cap. 9).

Agora, seja $\theta = (\theta_1^T, \theta_2^T)^T$ como em (2.9) e suponhamos que nosso interesse reside em θ_1 e portanto θ_2 representa um parâmetro secundário. Definimos a função implícita $g(\theta_1)$, tal que, para θ_1 fixado, a função $L(\theta_1, g(\theta_1))$ é maximizada. O afastamento pela verossimilhança para θ_1 , ignorando θ_2 , é definido como sendo

$$LD_i(\theta_1/\theta_2) = 2 \left[l(\tilde{\theta}) - l(\tilde{\theta}_{1(i)}, g(\tilde{\theta}_{1(i)})) \right]. \quad (2.11)$$

A interpretação é similar à medida dada em (2.10). Esta medida de influência foi utilizada por Cook e Wang (1982) para desenvolver medidas para transformações de parâmetros em modelos lineares.

Os resultados apresentados acima serão aplicados ao modelo de regressão linear definido em (2.1), onde $\theta = (\beta^T, \sigma^2)^T = (\theta_1^T, \theta_2^T)$ e a função de log-verossimilhança para θ é

dada por

$$l(\boldsymbol{\theta}) = -\frac{n}{2} \log 2\pi - \frac{n}{2} \log \sigma^2 - \frac{(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})}{2\sigma^2}.$$

Avaliando $l(\boldsymbol{\theta})$ no EMV $\tilde{\boldsymbol{\theta}} = (\tilde{\boldsymbol{\beta}}, \tilde{\sigma}^2)$, onde $\tilde{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}$ e $\tilde{\sigma}^2 = \mathbf{e}^T \mathbf{e} / n$, obtemos

$$l(\tilde{\boldsymbol{\theta}}) = -\frac{n}{2} \log 2\pi - \frac{n}{2} \log \tilde{\sigma}^2 \frac{(n-p)}{n} - \frac{n}{2},$$

e com a eliminação da i -ésima observação, $i = 1, \dots, n$, o EMV de $\boldsymbol{\theta}$ é dado por $\tilde{\boldsymbol{\theta}}_{(i)} = (\tilde{\boldsymbol{\beta}}_{(i)}, \tilde{\sigma}_{(i)}^2)$, onde

$$\tilde{\boldsymbol{\beta}}_{(i)} = (\mathbf{X}_{(i)}^T \mathbf{X}_{(i)})^{-1} \mathbf{X}_{(i)}^T \mathbf{Y}_{(i)}$$

e

$$\tilde{\sigma}_{(i)}^2 = \hat{\sigma}_{(i)}^2 \left(\frac{n-p-1}{n-1} \right) = \hat{\sigma}^2 \left(\frac{n-p-t_i^2}{n-1} \right), \quad (2.12)$$

onde $\hat{\sigma}_{(i)}^2$ é o estimador de mínimos quadrados de σ^2 com a eliminação da i -ésima observação. Assim, avaliando em $l(\boldsymbol{\theta})$, temos que

$$l(\tilde{\boldsymbol{\theta}}_{(i)}) = -\frac{n}{2} \log 2\pi - \frac{n}{2} \log \tilde{\sigma}_{(i)}^2 - \frac{n-1}{2} - \frac{(y_i - \mathbf{x}_i^T \tilde{\boldsymbol{\beta}}_{(i)})^2}{2\tilde{\sigma}_{(i)}^2},$$

e após manipulações algébricas

$$LD_i(\boldsymbol{\theta}) = n \log \left(\frac{n(n-p-t_i^2)}{(n-1)(n-p)} \right) + \frac{(n-1)(y_i - \mathbf{x}_i^T \tilde{\boldsymbol{\beta}}_{(i)})^2}{\tilde{\sigma}^2(n-p-t_i^2)} - 1.$$

A expressão $LD_i(\boldsymbol{\theta}) = LD_i(\boldsymbol{\beta}, \sigma^2)$ é útil quando estamos interessados na estimação conjunta de $\boldsymbol{\beta}$ e σ^2 , isto é $LD_i(\boldsymbol{\theta})$ é utilizado quando estamos interessados em analisar a influência da i -ésima observação na estimação conjunta dos parâmetros $\boldsymbol{\beta}$ e σ^2 . Se o interesse está centrado somente na estimação de $\boldsymbol{\beta}$, então utilizamos (2.11) que é equivalente a

$$LD_i(\boldsymbol{\beta} / \sigma^2) = 2 \left[l(\tilde{\boldsymbol{\theta}}) - \sup_{\sigma} l(\tilde{\boldsymbol{\beta}}_{(i)}, \sigma^2) \right].$$

Neste caso,

$$l\left(\tilde{\boldsymbol{\beta}}_{(i)}, \sigma^2\right) = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log \sigma^2 - \frac{1}{2\sigma^2} \sum_{j=1}^n \left(y_j - \mathbf{x}_j^T \tilde{\boldsymbol{\beta}}_{(i)}\right)^2$$

e o valor de σ^2 que maximiza $l\left(\tilde{\boldsymbol{\beta}}_{(i)}, \sigma^2\right)$ é dado por

$$\begin{aligned} \tilde{\sigma}^2\left(\tilde{\boldsymbol{\beta}}_{(i)}\right) &= \frac{1}{n} \sum_{j=1}^n \left(y_j - \mathbf{x}_j^T \tilde{\boldsymbol{\beta}}_{(i)}\right)^2 \\ &= \tilde{\sigma}_{(i)}^2 \left(\frac{n-1}{n}\right) + \frac{e_i^2}{n(1-h_{ii})^2} \end{aligned} \quad (2.13)$$

e

$$\sup_{\sigma} l\left(\tilde{\boldsymbol{\beta}}_{(i)}, \sigma^2\right) = -\frac{n}{2} \log 2\pi - \frac{n}{2} \log \tilde{\sigma}_{(i)}^2 - \frac{n}{2}.$$

Logo, de (2.12), (2.13) e (2.7), temos que

$$\begin{aligned} LD_i(\boldsymbol{\beta}/\sigma^2) &= n \log \left(\frac{\tilde{\sigma}_{(i)}^2}{\tilde{\sigma}^2} \right) \\ &= n \log \left(\frac{n-1}{n} \frac{\tilde{\sigma}_{(i)}^2}{\tilde{\sigma}^2} + \frac{e_i^2}{n\tilde{\sigma}^2(1-h_{ii})^2} \right) \\ &= n \log \left(\frac{n-p-1}{n-p} \frac{\hat{\sigma}_{(i)}^2}{\hat{\sigma}^2} + \frac{e_i^2}{(n-k)\hat{\sigma}^2(1-h_{ii})^2} \right) \\ &= n \log \left(\frac{n-p-t_i^2}{n-p} + \frac{t_i^2}{(n-p)(1-h_{ii})} \right) \\ &= \frac{n}{2} \log \left(1 + \frac{t_i^2}{(n-p)(1-h_{ii})} \right) \\ &= \frac{n}{2} \log \left(1 + \frac{p}{n-p} D_i \right), \end{aligned}$$

onde D_i é como em (2.7). Como $LD_i(\boldsymbol{\beta}/\sigma^2)$ é uma função monótona de D_i , então $LD_i(\boldsymbol{\beta}/\sigma^2)$ é equivalente à distância de Cook D_i .

Observação 2.1: Seja

$$b_i = \frac{e_i^2}{\mathbf{e}^T \mathbf{e} (1 - h_{ii})} \quad i = 1, \dots, n. \quad (2.14)$$

Então, sob o modelo de regressão definido em (2.1), temos que

$$b_i \sim \text{Beta} \left(1/2, \frac{n-p-1}{2} \right).$$

Além disso, utilizando a definição de b_i , (2.7) e fazendo operações algébricas obtemos que

$$D_i = \frac{(n-p)}{p} b_i \frac{h_{ii}}{(1-h_{ii})}, \quad (2.15)$$

$$\frac{\tilde{\sigma}_{(i)}^2}{\tilde{\sigma}^2} = \frac{s_{(i)}^2}{s^2} = \frac{n}{n-1} (1 - b_i), \quad (2.16)$$

$$LD_i(\boldsymbol{\beta}/\sigma^2) = n \log \left(\frac{b_i h_{ii}}{1 - h_{ii}} + 1 \right)$$

e

$$LD_i(\boldsymbol{\beta}, \sigma^2) = n \log \left(\frac{n}{n-1} \right) + n \log (1 - b_i) + \left(\frac{b_i}{1 - b_i} \right) \left(\frac{n-1}{1 - h_{ii}} \right) - 1.$$

Finalmente, se estamos interessados somente em σ^2 , então de (2.11), temos que

$$\begin{aligned} LD_i(\sigma^2/\boldsymbol{\beta}) &= 2 \left[l(\tilde{\boldsymbol{\theta}}) - \sup_{\boldsymbol{\beta}} l(\boldsymbol{\beta}, \tilde{\sigma}_{(i)}^2) \right] \\ &= n \log \left(\frac{n}{n-1} \frac{\tilde{\sigma}_{(i)}^2}{\tilde{\sigma}^2} \right) + n \left(\frac{\tilde{\sigma}^2}{\tilde{\sigma}_{(i)}^2} - 1 \right) \\ &= n \log \left(\frac{n}{n-1} (1 - b_i) \right) + \frac{nb_i - 1}{1 - b_i}. \end{aligned}$$

Além disso, podemos expressar os $DFFITs_i^2$ como

$$DFFITs_i^2 = (n-p-1) \left(\frac{b_i}{1 - b_i} \right) \left(\frac{h_{ii}}{1 - h_{ii}} \right),$$

logo o afastamento da verossimilhança pode ser relacionado com a medida de $DFFIT S_i$ ao considerar a diferença de $LD_i(\beta, \sigma^2)$ e $LD_i(\sigma^2/\beta)$, isto é

$$\begin{aligned} LD_i(\beta, \sigma^2) - LD_i(\sigma^2/\beta) &= \frac{b_i}{(1-b_i)} \frac{(n-1)}{(1-h_{ii})} - 1 - \frac{nb_i-1}{(1-b_i)} \\ &= (n-1) \frac{h_{ii}}{(1-h_{ii})} \frac{b_i}{(1-b_{ii})} \\ &= \frac{n-1}{n-p-1} DFFIT S_i^2. \end{aligned}$$

É importante observar que $LD_i(\beta, \sigma^2)$, $LD_i(\beta/\sigma^2)$ e $LD_i(\sigma^2/\beta)$ são considerados sob o modelo normal. Os resultados acima podem ser estendidos para modelos lineares não-normais, como por exemplo o modelo linear t -student (veja por exemplo, Galea et al., 1997). Para finalizar esta seção, a seguir apresentamos um exemplo onde utilizamos as medidas $LD_i(\beta, \sigma^2)$, $LD_i(\beta/\sigma^2)$ e $LD_i(\sigma^2/\beta)$ para identificar pontos influentes.

Exemplo 1

Para ilustração, consideramos os Dados do Clube de Saúde de Chatterjee e Hadi (1988, p. 128), que são apresentados na Tabela 2.2, consistindo de registros de saúde de 30 empregados, membros regulares do clube de saúde de uma companhia. As variáveis medidas são:

X_1 : Peso em libras.

X_2 : Taxa do pulso por minuto .

X_3 : Resistência de braços e pernas.

X_4 : Tempo (em segundos) de uma corrida de 1/4 de milha.

Y : Tempo (em segundos) de uma corrida de uma milha.

Na Tabela 2.3 apresentamos as medidas de influência baseadas nas funções de verossimilhança consideradas anteriormente e os correspondentes gráficos de influências dadas nas Figuras 2.2, 2.3 e 2.4. Observamos que na Figura 2.2 e 2.3 as observações de número 28 e 30 são mais influentes sob a função de verossimilhança e na Figura 2.4 as observações de número 8, 21 e 28 têm uma influência moderada sobre a estimação do

parâmetro embora a observação 30 seja a mais influente.

Tabela 2.2: Dados de Clube de Saúde

Nro.	Y	X ₁	X ₂	X ₃	X ₄	Nro.	Y	X ₁	X ₂	X ₃	X ₄
1	481	217	67	260	91	16	267	138	70	180	62
2	292	141	52	190	66	17	404	147	54	150	75
3	338	152	58	203	68	18	442	197	76	228	88
4	357	153	56	183	70	19	368	165	59	188	70
5	396	180	66	170	77	20	295	125	58	160	66
6	429	193	71	178	82	21	391	161	52	190	69
7	345	162	65	160	74	22	264	132	62	163	59
8	469	180	80	170	84	23	487	257	64	313	96
9	425	205	77	188	83	24	481	236	72	225	84
10	358	168	74	170	79	25	374	149	57	173	68
11	393	232	65	220	72	26	309	161	57	173	65
12	346	146	68	158	68	27	367	198	59	220	62
13	279	173	51	243	56	28	469	245	70	218	69
14	311	155	64	198	59	29	252	141	63	193	60
15	401	212	66	220	77	30	338	177	53	183	75

Fonte : Chatterjee e Hadi (1988, p. 128).

X₁ : Peso em libras.

X₂ : Taxa do pulso por minuto .

X₃ : Resistência de braços e pernas.

X₄ : Tempo (em segundos) de uma corrida de 1/4 de milha.

Y : Tempo (em segundos) de uma corrida de uma milha.

Tabela 2.3: Medidas baseadas sobre a função de verossimilhança

Nº	$LD_i(\beta/\sigma^2)$	$LD_i(\beta, \sigma^2)$	$LD_i(\sigma^2/\beta)$	Nº	$LD_i(\beta/\sigma^2)$	$LD_i(\beta, \sigma^2)$	$LD_i(\sigma^2/\beta)$
1	0,280	0,280	0,000	16	0,198	0,198	0,002
2	0,073	0,076	0,004	17	0,676	0,732	0,022
3	0,043	0,048	0,005	18	0,031	0,044	0,014
4	0,032	0,038	0,007	19	0,017	0,024	0,007
5	0,012	0,025	0,013	20	0,008	0,023	0,015
6	0,005	0,020	0,016	21	0,441	0,586	0,103
7	0,137	0,144	0,004	22	0,053	0,058	0,007
8	1,030	1,304	0,148	23	1,026	1,044	0,000
9	0,115	0,116	0,002	24	0,023	0,035	0,013
10	0,301	0,334	0,021	25	0,201	0,263	0,049
11	0,552	0,599	0,021	26	0,114	0,116	0,001
12	0,038	0,045	0,007	27	0,136	0,137	0,002
13	0,029	0,043	0,015	28	2,276	2,763	0,152
14	0,154	0,154	0,001	29	0,243	0,247	0,001
15	0,042	0,045	0,003	30	1,291	2,018	0,455

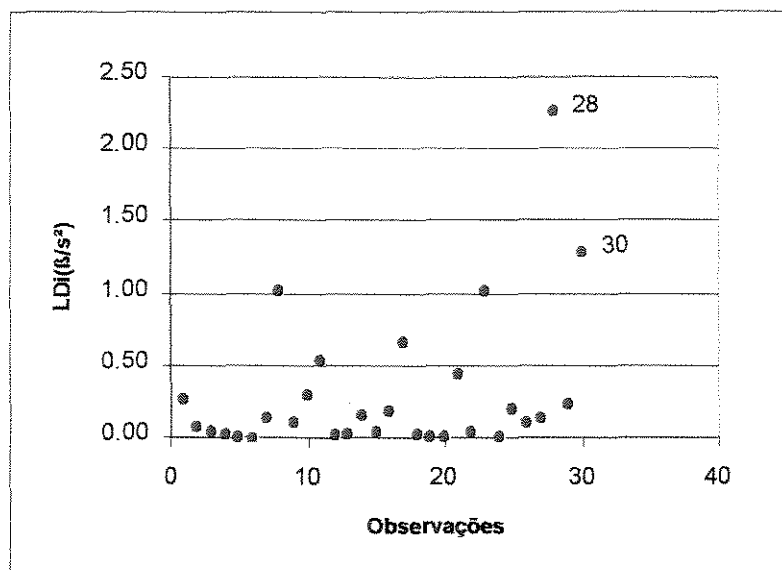


Figura 2-2: $LD_i (\beta/s^2)$ vs observações

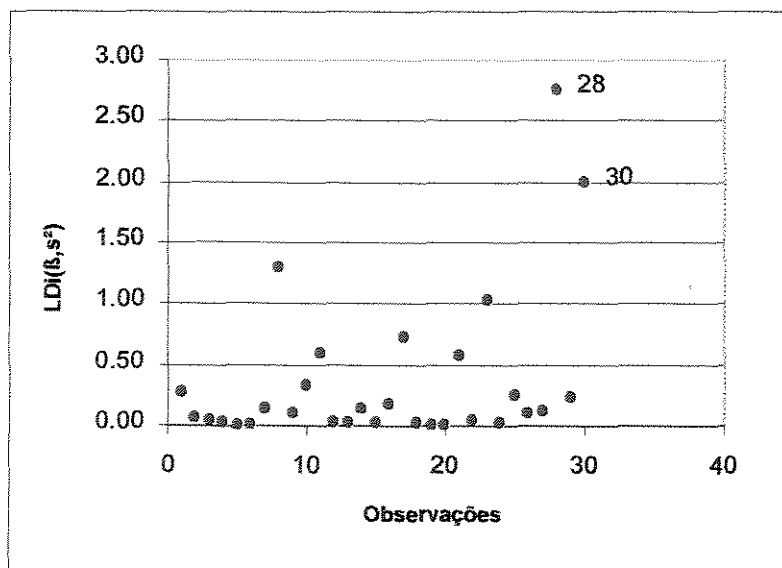


Figura 2-3: $LD_i (\beta, s^2)$ vs observações

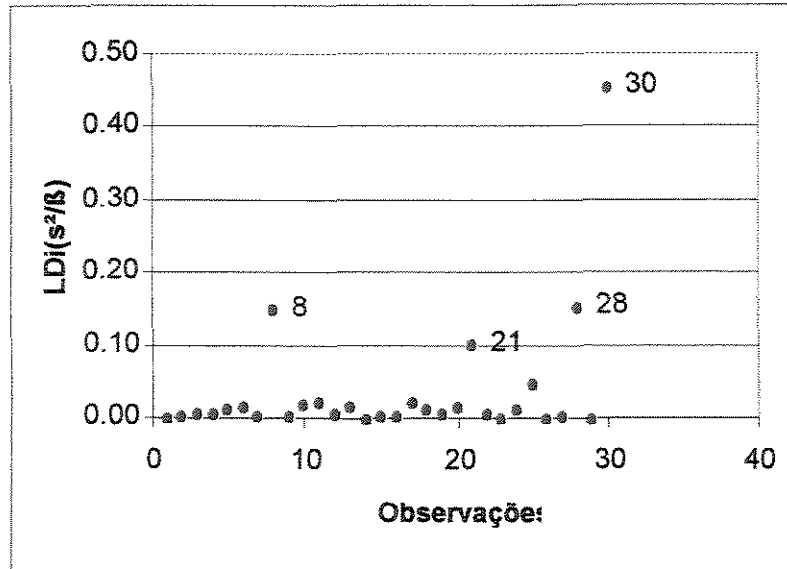


Figura 2-4: $LD_i(s^2/\beta)$ vs observações

Exemplo 2

Para ilustrar os resultados obtidos anteriormente utilizaremos os dados apresentados na seção 2.2.

- 1) Para o conjunto de dados (a) a observação três se destaca do resto (veja, Tabela 2.4).
A observação três e nove são mais influentes sob a função de verossimilhança quando nosso interesse é a estimação conjunta dos parâmetros β e σ^2 (veja, a coluna 4 da Tabela 2.4). A observação três é influente sob a função de verossimilhança quando nosso interesse reside na estimação de σ^2 (veja coluna 5).
- 2) Para o conjunto de dados (b) a observação 6 e 8 se destacam do resto (Tabela 2.5).
- 3) Para o conjunto de dados (c) a observação três é influente. (veja, Tabela 2.6)
- 4) Para o conjunto de dados (d) a observação oito é influente baseada nas cinco medidas de influência consideradas.

Tabela 2.4: Medidas de influência para os dados (a) de Anscombe (1973).

Nº	D_i	$DFFITs_i$	$LD_i(\beta/\sigma^2)$	$LD_i(\beta, \sigma^2)$	$LD_i(\sigma^2/\beta)$
1	0,000	0,012	0,000	-0,544	-0,544
2	0,000	-0,013	0,000	-0,544	-0,544
3	0,490	-1,160	0,247	5,145	3,264
4	0,063	0,360	0,033	0,394	0,250
5	0,002	-0,052	0,001	-0,530	-0,533
6	0,000	-0,020	0,000	-0,543	-0,544
7	0,128	0,514	0,067	0,651	0,325
8	0,125	-0,485	0,065	0,229	-0,130
9	0,272	0,825	0,140	2,703	1,863
10	0,156	-0,602	0,081	1,538	1,119
11	0,004	0,087	0,002	-0,517	-0,527

Tabela 2.5: Medidas de influência para os dados (b) de Anscombe (1973).

Nº	D_i	$DFFITs_i$	$LD_i(\beta/\sigma^2)$	$LD_i(\beta, \sigma^2)$	$LD_i(\sigma^2/\beta)$
1	0,052	0,322	0,028	0,090	-0,017
2	0,052	0,322	0,028	0,090	-0,017
3	0,077	-0,380	0,040	-0,055	-0,234
4	0,058	0,344	0,031	0,237	0,116
5	0,031	0,242	0,017	-0,249	-0,313
6	0,808	-1,528	0,394	7,786	4,285
7	0,001	0,049	0,001	-0,535	-0,538
8	0,808	-1,528	0,394	7,786	4,285
9	0,001	0,049	0,001	-0,535	-0,538
10	0,031	0,242	0,017	-0,249	-0,313
11	0,077	-0,380	0,040	-0,055	-0,234

Tabela 2.6: Medidas de influência para os dados (c) de Anscombe (1973).

Nº	D_i	$DFFITs_i$	$LD_i(\beta/\sigma^2)$	$LD_i(\beta, \sigma^2)$	$LD_i(\sigma^2/\beta)$
1	0,012	-0,100	0,006	-0,415	-0,437
2	0,002	-0,100	0,001	-0,522	-0,526
3	1,393	10000	0,644	*	*
4	0,005	-0,100	0,003	-0,480	-0,490
5	0,026	-0,200	0,014	-0,302	-0,354
6	0,301	-0,800	0,154	1,423	0,531
7	0,001	0,000	0,000	-0,541	-0,542
8	0,034	0,200	0,018	-0,362	-0,448
9	0,060	-0,300	0,031	-0,081	-0,211
10	0,000	0,000	0,000	-0,542	-0,542
11	0,007	0,100	0,004	-0,503	-0,518

Tabela 2.7: Medidas de influência para os dados (d) de Anscombe (1973).

Nº	D_i	$DFFITs_i$	$LD_i(\beta/\sigma^2)$	$LD_i(\beta, \sigma^2)$	$LD_i(\sigma^2/\beta)$
1	0,007	-0,114	0,004	-0,466	-0,480
2	0,062	-0,356	0,033	0,235	0,104
3	0,020	0,194	0,011	-0,315	-0,355
4	0,137	0,578	0,071	1,554	1,218
5	0,087	0,433	0,046	0,622	0,429
6	0,000	0,010	0,000	-0,544	-0,544
7	0,124	-0,541	0,065	1,289	0,993
8	*	*	*	*	*
9	0,084	-0,423	0,044	0,568	0,383
10	0,033	0,252	0,018	-0,155	-0,222
11	0,000	-0,030	0,000	-0,539	-0,540

2.3 Afastamento pela Verossimilhança no Caso Geral

Nesta seção vamos apresentar uma maneira mais geral para avaliar perturbação nas observações e/ou no modelo.

Para um conjunto de dados observados, seja $l(\boldsymbol{\theta})$ a função de log-verossimilhança do modelo postulado (o modelo assumido correto), onde $\boldsymbol{\theta}$ é um vetor de parâmetros $p \times 1$ e seja $\boldsymbol{\omega}$ um vetor de perturbações $q \times 1$ assumindo valores em um subconjunto aberto $\Omega \subseteq \mathbb{R}^q$. Denotaremos por $l(\boldsymbol{\theta}/\boldsymbol{\omega})$ a função de log-verossimilhança do modelo perturbado. Assumiremos que a função de log-verossimilhança do modelo perturbado $l(\boldsymbol{\theta}/\boldsymbol{\omega})$ é duas vezes diferenciável em $(\boldsymbol{\theta}^T, \boldsymbol{\omega}^T)^T$, e que o modelo postulado está encaixado no modelo perturbado, isto é, existe $\boldsymbol{\omega}_0 \in \Omega$ tal que $l(\boldsymbol{\theta}/\boldsymbol{\omega}_0) = l(\boldsymbol{\theta})$, $\forall \boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^p$. Se p e q (dimensão de $\boldsymbol{\omega}$) são pequenos, é suficiente comparar $l(\boldsymbol{\theta})$ e $l(\boldsymbol{\theta}/\boldsymbol{\omega})$ graficamente para vários valores de $\boldsymbol{\omega}$ em Ω . Em situações mais gerais o problema é mais complexo.

É importante observar que em geral, a dimensão q (do vetor de perturbação) está relacionada com a dimensão do vetor $\boldsymbol{\theta}$ ou com o tamanho da amostra. Como ilustração vamos considerar o problema da influência de uma determinada observação, que foi estudada nas seções anteriores. Considere o modelo de regressão linear definido em (2.1), com função de log-verossimilhança dada por

$$l(\boldsymbol{\theta}) = -\frac{n}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{j=1}^n (y_j - \mathbf{x}_j^T \boldsymbol{\beta})^2,$$

onde $\boldsymbol{\theta} = (\boldsymbol{\beta}^T, \sigma^2)^T$. O esquema de perturbação correspondente à exclusão da i -ésima observação é obtido ao considerar $\boldsymbol{\omega} = \omega_i$, $\omega_i > 0$ e a função de log-verossimilhança $l(\boldsymbol{\theta}/\boldsymbol{\omega})$ é dada por

$$l(\boldsymbol{\theta}/\boldsymbol{\omega}) = l_1(\boldsymbol{\theta}) + \dots + l_{i-1}(\boldsymbol{\theta}) + w_i l_i(\boldsymbol{\theta}) + l_{i+1}(\boldsymbol{\theta}) + \dots + l_n(\boldsymbol{\theta}),$$

onde $\omega_i \rightarrow 0$ e

$$l_j(\theta) = -\frac{1}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} (y_j - \mathbf{x}_j^T \beta)^2.$$

Neste caso, o vetor ω_0 é dado por $\omega_0 = 1$, $l(\theta/\omega_0) = l(\theta)$ e $\omega \in \Omega \subseteq \mathbf{R}$.

Para o problema de exclusão da i -ésima observação considerado acima, comparamos $\tilde{\theta}$ e $\tilde{\theta}_{(i)}$ utilizando a distância LD_i (veja seção 2.2.3). O estimador $\tilde{\theta}_{(i)}$ obtido ao considerar a exclusão da i -ésima observação sob a perturbação $\omega = \omega_i$ com $\omega_i \rightarrow 0$ pode ser denotado por $\tilde{\theta}_\omega = \tilde{\theta}_{(i)}$, ou seja comparamos $\tilde{\theta}$ e $\tilde{\theta}_\omega$ utilizando a função de log-verossimilhança.

Em geral, se $\tilde{\theta}$ é o EMV sob o modelo postulado e $\tilde{\theta}_\omega$ o EMV sob o modelo perturbado, nosso objetivo é comparar $\tilde{\theta}$ e $\tilde{\theta}_\omega$, quando ω varia em Ω . Se a distância entre eles permanece “pequena” quando ω varia em Ω , é uma indicação que existe uma estabilidade do modelo ajustado no que diz respeito ao esquema particular de perturbação utilizado, e portanto ao correspondente aspecto da análise que está sendo monitorado.

A diferença entre $\tilde{\theta}$ e $\tilde{\theta}_\omega$ pode depender da forma de $l(\theta)$. Se $l(\theta)$ é suficientemente horizontal, podemos dizer que $\tilde{\theta}$ e $\tilde{\theta}_\omega$ são bem “próximos”, embora se $l(\theta)$ é suficientemente concentrado em torno de $\tilde{\theta}$, então $\tilde{\theta}$ e $\tilde{\theta}_\omega$ podem aparecer muito distantes, veja Figura 2.5.

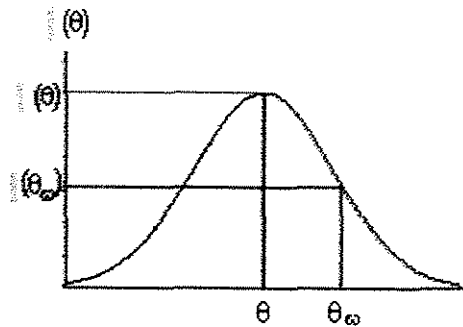


Figura 2-5: Ilustração esquemática do afastamento pela verossimilhança

Assim, a comparação direta de $\tilde{\theta}$ e $\tilde{\theta}_\omega$ pode não ser simples, devido a diversos fatores tais como diferença de escala, unidade de medida, erros de medição, etc. Uma proposta para comparar $\tilde{\theta}$ e $\tilde{\theta}_\omega$ é sugerida por Cook (1986) e consiste no uso da função

$$LD(\omega) = 2 \left[l(\tilde{\theta}) - l(\tilde{\theta}_\omega) \right], \quad \omega \in \Omega,$$

que é denominada afastamento pela verossimilhança, e é uma generalização do afastamento pela verossimilhança LD_i definido em 2.10. Desta forma, a distância entre $\tilde{\theta}$ e $\tilde{\theta}_\omega$ pode ser avaliado utilizando a função LD . Se $l(\theta)$ tiver um grande achatamento, então podemos julgar que $\tilde{\theta}$ e $\tilde{\theta}_\omega$ estão próximos entre si. Por outro lado, se $l(\theta)$ for suficientemente concentrada em torno de $\tilde{\theta}$, então $\tilde{\theta}_\omega$ deverá ficar distante de $\tilde{\theta}$.

Como $\tilde{\theta}$ é o EMV para θ no modelo não perturbado, segue-se que $LD(\omega) \geq 0$ para $\omega \in \Omega$. Além disso, como $LD(\omega_0) = 0$, podemos concluir que ω_0 é um ponto de mínimo local da função $LD(\omega)$. Uma análise sobre o comportamento geométrico da função $LD(\omega)$, quando ω varia em Ω , pode fornecer informações de características relevantes ao modelo sob investigação; por exemplo, se os valores de $LD(\omega)$, para diferentes valores de $\omega \in \Omega$, são “bem” próximos, isto indica que existe estabilidade no modelo ajustado sob a perturbação considerada. Por esta razão, antes de fazer um estudo da função $LD(\omega)$, $\omega \in \Omega$, é importante que o esquema de perturbação seja bem definido de modo que responda a questões previamente estabelecidas.

A definição de $LD(\omega)$ pode ser interpretada em termos da região de confiança assintótica para θ

$$\left\{ \theta \in \mathbf{R}^p : 2 \left[l(\tilde{\theta}) - l(\theta) \right] \leq \chi^2_{(1-\alpha, p)} \right\},$$

onde $\chi^2_{(1-\alpha, p)}$ é o percentil superior de tamanho $1 - \alpha$ da distribuição qui-quadrado com p graus de liberdade veja por exemplo, Cox e Hinkley (1974, capítulo 9).

2.3.1 Alguns Tipos de Perturbações

Na literatura estatística não há uma definição clara de perturbação. Segundo Billor e Loynes (1993), perturbação é qualquer arranjo da mudança da suposição do modelo ou dados perturbados para constatar alguma mudança substancial que ocorre nos resultados da análise. Também mencionam que pode-se perturbar o modelo ou os dados.

Alguns esquemas de perturbações mais comuns considerados na literatura são:

Ponderação de Casos: A função de log-verossimilhança perturbada tem a forma

$$l(\boldsymbol{\theta}/\boldsymbol{\omega}) = \sum_{j=1}^n \omega_j \log f(y_j; \boldsymbol{\theta}) = \sum_{j=1}^n \omega_j l_j(\boldsymbol{\theta}),$$

onde $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n)^T$ pertence a um subconjunto aberto em $\mathbf{R}^n$, com $\omega_i > 0$, $i = 1, \dots, n$.

Com este tipo de perturbação, deseja-se avaliar se a contribuição das observações com ponderações diferentes afeta na estimação do parâmetro $\boldsymbol{\theta}$. Segundo Cook (1987), a perturbação de ponderação de casos é talvez o método mais comum usado em investigações para avaliar influência em uma modificação pequena de um modelo.

Consideremos o modelo postulado como sendo o modelo de regressão linear definido em (2.1). Então, sob normalidade, a função de log-verossimilhança sob o modelo perturbado é dada por

$$l(\boldsymbol{\beta}, \sigma^2/\boldsymbol{\omega}) = \sum_{j=1}^n \omega_j \log 2\pi\sigma^2 - \frac{1}{2\sigma^2} \sum_{j=1}^n \omega_j (y_j - \mathbf{x}_j^T \boldsymbol{\beta})^2,$$

onde, $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n)^T$ e ω_i representa a ponderação fixa para o i -ésimo caso. Quando $\boldsymbol{\omega} = \boldsymbol{\omega}_0 = (1, \dots, 1)^T$, o modelo perturbado se reduz ao modelo postulado, isto é, $l(\boldsymbol{\theta}) = l(\boldsymbol{\theta}/\boldsymbol{\omega})$.

Diagnóstico popular baseado sob exclusão de casos, dá uma aplicação do esquema de ponderação de casos. Por exemplo, para comparar resultados baseados nos dados completos com esses baseados sob dados sem o i -ésimo caso. Kim (1995) e Fung e Tang (1997) tam-

bém estudaram influência local em regressão multivariada. Kim (1995) investiga somente o esquema de perturbação de ponderação de casos para todos os parâmetros de interesse, embora Fung e Tang (1997), além disso investiguem subconjuntos dos parâmetros de interesse e casos individuais.

Perturbação dos Dados: Billor e Loynes (1993) consideram que os tipos de perturbação possíveis podem depender do modelo. Em regressão linear normal pode-se perturbar a matriz de variáveis explanatórias e o vetor resposta. Há duas razões para considerar perturbação dos dados: a existência de medidas com erros de medição e a existência de “outliers” (em uma proporção relativamente pequena dos dados).

Cook (1987) considera perturbação dos valores da p -ésima variável explanatória, e desta forma o modelo perturbado passa a ser

$$\mathbf{Y} = [\mathbf{X} + (0, \boldsymbol{\omega})] \boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

onde $\boldsymbol{\varepsilon} \sim N_n(0, \sigma^2 \mathbf{I}_n)$ e $\boldsymbol{\omega}$ é um vetor $n \times 1$ de perturbações. Esse tipo de esquema de perturbação é muitas vezes proveitoso para detectar “outliers” nas colunas de $\mathbf{X}$, ou para avaliar influência devido à colinearidade. (veja também, Fung e Tang, 1997). Neste caso $\boldsymbol{\omega}_0 = (0, \dots, 0)^T$.

Perturbação na Resposta: Em geral este tipo de perturbação é obtido ao substituir a variável resposta Y_j por uma função de Y_j e ω , $\omega \in \Omega$. Em regressão linear, o esquema de perturbação considerado é dado por

$$Y_j(\omega) = Y_j + \tilde{\sigma} \omega_j, \quad j = 1, \dots, n,$$

onde $\tilde{\sigma}$ é o EMV de σ e σ o desvio padrão de Y_j .

Em análise de sobrevivência, Escobar e Meeker (1992) consideram perturbações no vetor

de respostas para o j -ésimo caso, que é dado por

$$Y_j(\omega) = \begin{cases} Y_j + \delta_j = Y_j + \hat{\sigma}\omega_j, & j = 1, \dots, n, \text{ se } \omega_j \in [-1, 0) \cup (0, 1] \\ Y_j & \text{se } \omega_j = 0. \end{cases}$$

e neste caso $\omega_0 = (0, \dots, 0)^T$. Emerson, Hoaglin e Kempthorne (1984) também usam perturbações na resposta para estudar “leverage” em um modelo não-linear para uma tabela de duas classificações.

Pelo exposto acima, os esquemas de perturbações, sem considerar a ponderação de casos, podem ser classificados em dois grandes grupos: perturbação no modelo (estudar as modificações das suposições) e perturbação nos dados. Essencialmente o esquema de perturbação a ser considerado deve ser formulado de modo que responda a questões previamente estabelecidas pelo pesquisador. Por exemplo, alguns esquemas de perturbação podem envolver a correlação dos elementos do vetor de erros ϵ , o vetor de respostas, parâmetros a priori em análise bayesiana baseada no modelo de regressão linear (veja, por exemplo Lavine, 1992).

2.4 Influência Local

A idéia de influência local foi introduzida por Cook (1986) e consiste em investigar o comportamento da função de afastamento pela verossimilhança $LD(\omega)$ em uma vizinhança de ω_0 , que é o ponto onde as duas funções de log-verossimilhanças são iguais, isto é, $l(\theta) = l(\theta/\omega)$.

Note que o gráfico da função $LD(\omega)$, $\omega \in \Omega$, pode ser representado pela superfície

geométrica $(q + 1)$ –dimensional formado pelos valores do vetor

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

quando ω varia em Ω . No contexto estatístico essa superfície é chamada de gráfico de influência, pois o gráfico da função $LD(\omega)$, $\omega \in \Omega$ mostra a influência do esquema de perturbação veja Cook (1986) e Cook (1987).

O estudo de influência local consiste em analisar como a superfície $\alpha(\omega) = (\omega^T, LD(\omega))^T$ desvia-se de seu plano tangente em ω_0 . Essa análise pode ser feita estudando-se as curvaturas das seções normais da superfície $\alpha(\omega)$ em ω_0 , que são intersecções de $\alpha(\omega)$ com o plano contendo o vetor normal a seu plano tangente em ω_0 .

Se a classe de todas as seções normais da superfície $\alpha(\omega)$ em ω_0 tiverem curvatura normal “próxima” de zero, isto é, as seções normais são “próximas” a uma linha reta em torno de ω_0 , então $LD(\omega)$ é “próximo” de zero em torno de ω_0 , o que indica que existe uma estabilidade do modelo ajustado sob o particular esquema de perturbação que está sendo considerado. Por outro lado, se existir uma seção normal com curvatura em módulo não “próximo” de zero, então nessa direção vamos ter maior oscilação da função $LD(\omega)$ e portanto, sob essa seção normal, vamos ter maior influência do particular esquema de perturbação. Se projetarmos perpendicularmente esta seção normal sobre o plano $\mathbf{R}^q$, vamos obter um vetor $\mathbf{d} \in \mathbf{R}^q$ tal que a reta $\omega_0 + t\mathbf{d}$, $t \in \mathbf{R}$ e a seção normal pertencem ao mesmo plano e portanto será equivalente considerar uma seção normal a $\alpha(\omega)$ em ω_0 com um vetor $\mathbf{d} \in \mathbf{R}^q$ para indicar a direção onde existe maior ou menor influência em torno de ω_0 .

Consideremos uma direção arbitrária $\mathbf{d}$ em $\mathbf{R}^q$ (sem perda de generalidade, vamos considerar o vetor $\mathbf{d}$ como sendo um vetor unitário, $\|\mathbf{d}\| = 1$) e uma linha reta em Ω ,

passando por ω_0 na mesma direção de $\mathbf{d}$. Essa reta é dada por

$$\omega(t) = \omega_0 + t\mathbf{d}, \quad t \in \mathbf{R}.$$

Calculando-se $LD(\omega(t))$ com $t \in \mathbf{R}$, podemos obter uma linha levantada (tradução livre de lifted line) sobre a superfície $\alpha(\omega)$ (que é equivalente a considerar uma seção normal da superfície $\alpha(\omega)$ em ω_0). Essa linha levantada é essencialmente a curva plana dada pela parametrização

$$\rho(t) = (\rho_1, \rho_2)^T = [t, LD\{\omega(t)\}]^T$$

e pode ser vista como o conjunto dos pontos de $\alpha(\omega)$ que são comuns ao plano gerado pelos vetores $\mathbf{b}_{q+1}$ e $(\mathbf{d}^T, 0)^T$, onde $\mathbf{b}_{q+1}$ é um vetor que tem o valor 1 na coordenada $q+1$ e zero nas demais coordenadas.

Na figura (2.6) apresentamos um exemplo de um gráfico de influência; os dados considerados são os dados de Weisberg (1985), que apresenta dados de absorção de éter em fígados de ratos.

Descrição dos dados

Consideremos os dados de Weisberg (1985). Um experimento foi conduzido para investigar a quantidade de uma droga particular apresentada no fígado de um rato. 19 ratos foram aleatoriamente selecionados e pesados sob o efeito anestésico do éter administrado por via oral. Ainda que fígados grandes absorvam mais de uma dada dose do que fígados pequenos, a dose atual que recebe um animal foi aproximadamente determinada como 40mg da droga por kilograma do peso de corporal (X_1). Depois de um determinado tempo cada rato foi sacrificado. O peso do fígado (X_2) e a porcentagem da dose no fígado (Y) foram determinados. Em resumo, o conjunto de dados consiste de 19 casos e 4 variáveis explanatórias, X_0 : constante, X_1 : peso do corpo, X_2 : peso do fígado, X_3 : dose relativa e a variável resposta Y : porcentagem da dose no fígado. A Tabela 2.4 mostra o conjunto de dados. O modelo postulado considerado é o modelo (2.1) e o modelo perturbado é definido por

$$\mathbf{Y} = [\mathbf{X} + (\mathbf{0}, \boldsymbol{\omega})] \boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

onde $\boldsymbol{\varepsilon} \sim N_n(\mathbf{0}, \sigma^2 \mathbf{I}_n)$, $\mathbf{0}$ é uma matriz 19×3 e $\boldsymbol{\omega} = (\omega_1, 0, \omega_3, 0, \dots, 0)^T$ é um vetor 19×1 de perturbações, construído adicionando as perturbações ω_i às doses para os casos 1 e 3. Para esses casos, as perturbações são construídas para ficar no intervalo $(-0.04, 0.04)$. As doses reportadas para caso 1 e 3 são 0.88 e 1.0, respectivamente (ver, Tabela 2.8) e desta forma as perturbações induzidas são relativamente pequenas. A Figura 2.6 mostra o exemplo de gráfico de influência de $LD(\boldsymbol{\omega})$ vs $\boldsymbol{\omega}$, quando $\boldsymbol{\omega}$ varia em Ω , para esse caso, onde o espaço de perturbações é $\Omega = (-0.04, 0.04) \times (-0.04, 0.04) \subseteq \mathbf{R}^2$. Nessa figura podemos ver um plano gerado pela direção $\mathbf{d} = (-0.44, 0.99)^T$. Este vetor fornece a curva plana em destaque, que é a linha projetada. Esse exemplo é discutido em Cook(1987).

Tabela 2.8: Absorção de éter pelo fígado de ratos (Weisberg 1985, p. 122)

Nº	X_1 : Peso do corpo	X_2 :Peso do fígado	X_3 : Dose	Y : % dose no fígado
1	176	6,50	0,88	0,42
2	176	9,50	0,88	0,25
3	190	9,00	1,00	0,56
4	176	8,90	0,88	0,23
5	200	7,20	1,00	0,23
6	167	8,90	0,83	0,32
7	188	8,00	0,94	0,37
8	195	10,00	0,98	0,41
9	176	8,00	0,88	0,33
10	165	7,90	0,84	0,38
11	158	6,90	0,80	0,27
12	148	7,30	0,74	0,36
13	149	5,20	0,75	0,21
14	163	8,40	0,81	0,28
15	170	7,20	0,85	0,34
16	186	6,80	0,94	0,28
17	146	7,30	0,73	0,30
18	181	9,00	0,90	0,37
19	149	6,40	0,75	0,46

Para fazer o gráfico de influência $LD(\omega)$, $\omega \in \Omega$ precisamos obter os estimadores de máxima verossimilhança $\tilde{\beta}$ e $\tilde{\beta}_\omega$ sob os modelos postulado e perturbado, respectivamente. No modelo postulado definido em (2.1), a função de log-verossimilhança para σ^2 fixo é dada por

$$l(\beta) = -\frac{n}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} (\mathbf{y} - \mathbf{X}\beta)^T (\mathbf{y} - \mathbf{X}\beta),$$

e

$$l(\tilde{\beta}) = -\frac{n}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} (\mathbf{y} - \mathbf{X}\tilde{\beta})^T (\mathbf{y} - \mathbf{X}\tilde{\beta}),$$

onde $\tilde{\beta} = (X^T X)^{-1} X^T Y$.

Seja $\mathbf{Y} = \mathbf{X}_\omega \beta + \varepsilon$ o modelo perturbado, onde $\mathbf{X}_\omega = \mathbf{X} + (0, \omega)$. Então a função de log-verossimilhança sob o modelo perturbado é definida por

$$l(\beta/\omega) = -\frac{n}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} (\mathbf{y} - \mathbf{X}_\omega \beta)^T (\mathbf{y} - \mathbf{X}_\omega \beta),$$

e o EMV de β sob o modelo perturbado é dado por

$$\tilde{\beta}_\omega = (\mathbf{X}_\omega^T \mathbf{X}_\omega)^{-1} \mathbf{X}_\omega^T \mathbf{Y}.$$

Assim, fazendo $\hat{\mathbf{Y}}_\omega = \mathbf{X} \tilde{\beta}_\omega$, temos que

$$\begin{aligned} LD(\omega) &= 2 \left[l(\tilde{\beta}) - l(\tilde{\beta}_\omega) \right] \\ &= \frac{1}{\sigma^2} \left[\|\mathbf{Y} - \hat{\mathbf{Y}}_\omega\|^2 - \|\mathbf{Y} - \hat{\mathbf{Y}}\|^2 \right]. \end{aligned}$$

Para o exemplo considerado aqui, desenharemos o gráfico $LD(\omega)$ vs ω (veja Figura 2.6). No que segue desta seção, vamos apresentar os principais resultados do método de influência proposto por Cook (1986). Será importante supor que para um conjunto de dados observados a matriz de informação de Fisher avaliada em $\theta = \tilde{\theta}$ é definida positiva.

Lembremos que a linha levantada consiste da curva plana

$$\rho(t) = (\rho_1(t), \rho_2(t))^T = (t, LD(\omega(t))),$$

onde $\omega(t) = \omega_0 + t\mathbf{d}$, $t \in \mathbf{R}$ e $\mathbf{d} \in \mathbf{R}^q$. Será importante supor que para um conjunto de dados observados a matriz de informação de Fisher avaliado em $\theta = \tilde{\theta}$ é definida positiva.

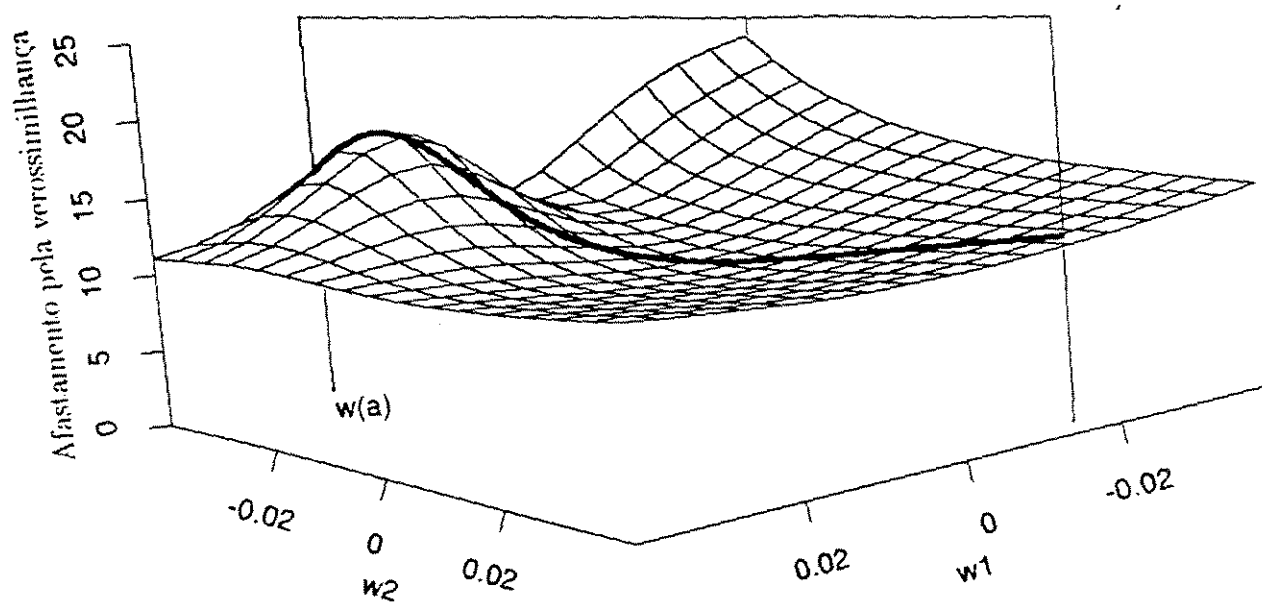


Figura 2-6: Gráfico de influência em R^3

Assim, a curvatura normal da curva $\rho(t)$ no ponto $\rho(0)$ na direção $\mathbf{d} \in \mathbf{R}^q$ é dada por

$$C_d = \frac{|\ddot{\rho}_1 \ddot{\rho}_2 - \ddot{\rho}_2 \ddot{\rho}_1|}{(\dot{\rho}_1^2 + \dot{\rho}_2^2)^{3/2}}, \quad (2.17)$$

avaliada em ω_0 , (veja por exemplo, Stoker, 1969, p. 26). Sob a suposição de que a função de log-verossimilhança perturbada $l(\theta/\omega)$ é duas vezes diferenciável em $(\theta^T, \omega^T)^T$, e portanto $\rho(t) = \left[t, LD\{\omega(t)\} \right]^T$ é também duas vezes diferenciável, então vamos ter os seguintes resultados:

i) $\dot{\rho}_1 = \frac{dt}{dt} = 1$ e $\ddot{\rho}_1 = \frac{d^2t}{dt^2} = 0$;

ii) $\frac{d}{dt}\omega(t) = \frac{d}{dt}(\omega_0 + t\mathbf{d}) = \mathbf{d}$ e $\frac{d^2}{dt^2}\omega(t) = 0$;

iii) Pela regra da cadeia (veja apêndice A) e (ii)

$$\begin{aligned} \frac{d}{dt}LD\{\omega(t)\} &= \frac{d\omega^T}{dt} \times \frac{\partial LD(\omega)}{\partial \omega} \\ &= \mathbf{d}^T \underbrace{\left[\frac{\partial}{\partial \omega} LD(\omega) \right]}_{\text{Número}} \\ &= \left[\frac{\partial}{\partial \omega^T} LD(\omega) \right] \mathbf{d} \end{aligned}$$

iv) $\dot{\rho}_2(0) = \frac{d}{dt}LD\{\omega(t)\}|_{t=0} = \left(\left[\frac{\partial}{\partial \omega} LD(\omega) \right]^T \mathbf{d} \right)|_{\omega=\omega_0} = 0$, uma vez que ω_0 é o ponto mínimo local da função LD .

Substituindo essas derivadas na equação (2.17) tem-se que

$$\begin{aligned} C_d &= \left| \ddot{\rho}_2 \right| \\ &= \left| \ddot{LD}\{\omega(t)\} \right|_{t=0}. \end{aligned}$$

Usando a regra da cadeia (veja Apêndice A, propriedade A2), segue-se que

$$\begin{aligned}
\ddot{LD} \{ \omega(t) \} &= \frac{d^2}{dt^2} LD(\omega(t)) = \frac{d}{dt} \left(\frac{d}{dt} LD(\omega(t)) \right), \quad \text{por definição} \\
&= \frac{d}{dt} \left[\mathbf{d}^T \frac{\partial}{\partial \omega} LD(\omega(t)) \right], \quad \text{por iii)} \\
&= \frac{d}{dt} \left[\left(\frac{\partial}{\partial \omega} LD(\omega(t)) \right)^T \mathbf{d} \right] \\
&= \frac{d}{dt} \left[\frac{\partial}{\partial \omega^T} LD(\omega(t)) \mathbf{d} \right] \\
&= \frac{d}{dt} \left[\frac{\partial}{\partial \omega^T} LD(\omega(t)) \right] \mathbf{d} \\
&= \frac{d\omega^T}{dt} \cdot \frac{\partial}{\partial \omega} \left(\frac{\partial}{\partial \omega^T} LD(\omega) \right) \mathbf{d}, \quad \text{regra da cadeia} \\
&= \mathbf{d}^T \frac{\partial^2 LD \{ \omega(t) \}}{\partial \omega \partial \omega^T} \mathbf{d}, \quad \text{por ii)}.
\end{aligned}$$

Quando essa derivada de segunda ordem é avaliada em $t = 0$, obtém-se a curvatura normal

$$C_d = \left| \mathbf{d}^T \ddot{LD}(\omega_0) \mathbf{d} \right|. \quad (2.18)$$

Mas, pela definição de afastamento pela verossimilhança,

$$\ddot{LD}(\omega_0) = \frac{\partial^2}{\partial \omega \partial \omega^T} \left[2 \left\{ l(\tilde{\theta}) - l(\tilde{\theta}_\omega) \right\} \right]_{\omega=\omega_0} = -2 \frac{\partial^2}{\partial \omega \partial \omega^T} l(\tilde{\theta}_\omega) \Big|_{\omega=\omega_0}$$

e portanto, a curvatura normal fica dada por

$$C_d = 2 \left| \mathbf{d}^T \ddot{\mathbf{F}} \mathbf{d} \right|,$$

onde $\ddot{\mathbf{F}}$ é uma matriz quadrada de dimensão $q \times q$, cujos elementos são dados pelas derivadas $\frac{\partial^2}{\partial \omega_r \partial \omega_s} l(\tilde{\theta}_\omega)$, avaliadas em $\omega = \omega_0$.

Os elementos (r, s) da matriz $\ddot{\mathbf{F}}$ podem ser obtidos através da regra da cadeia, a saber

$$\begin{aligned}
\ddot{F}_{rs} &= \frac{\partial^2}{\partial \omega_r \partial \omega_s} l(\tilde{\boldsymbol{\theta}}_\omega) \Big|_{\omega=\omega_0} \\
&= \frac{\partial}{\partial \omega_r} \left\{ \frac{\partial}{\partial \omega_s} l(\tilde{\boldsymbol{\theta}}_\omega) \right\} \Big|_{\omega=\omega_0} \\
&= \frac{\partial}{\partial \omega_r} \left[\left\{ \frac{\partial}{\partial \tilde{\boldsymbol{\theta}}_\omega} l(\tilde{\boldsymbol{\theta}}_\omega) \right\}^T \frac{\partial \tilde{\boldsymbol{\theta}}_\omega}{\partial \omega_s} \Big|_{\omega=\omega_0} \right], \quad \text{regra da cadeia} \\
&= \left[\frac{\partial}{\partial \omega_r} \left\{ \frac{\partial}{\partial \tilde{\boldsymbol{\theta}}_\omega} l(\tilde{\boldsymbol{\theta}}_\omega) \right\} \frac{\partial \tilde{\boldsymbol{\theta}}_\omega}{\partial \omega_s} \right] \Big|_{\omega=\omega_0} + \\
&\quad \left[\left\{ \frac{\partial}{\partial \tilde{\boldsymbol{\theta}}_\omega} l(\tilde{\boldsymbol{\theta}}_\omega) \right\} \frac{\partial}{\partial \omega_r} \left(\frac{\partial \tilde{\boldsymbol{\theta}}_\omega}{\partial \omega_s} \right) \right] \Big|_{\omega=\omega_0}, \quad \text{derivada do produto} \\
&= \left[\left\{ \frac{\partial^2}{\partial \tilde{\boldsymbol{\theta}}_\omega \partial \tilde{\boldsymbol{\theta}}_\omega} l(\tilde{\boldsymbol{\theta}}_\omega) \right\}^T \frac{\partial \tilde{\boldsymbol{\theta}}_\omega}{\partial \omega_r} \right]^T \frac{\partial \tilde{\boldsymbol{\theta}}_\omega}{\partial \omega_s} \Big|_{\omega=\omega_0} + \\
&\quad + \left\{ \frac{\partial}{\partial \tilde{\boldsymbol{\theta}}_\omega} l(\tilde{\boldsymbol{\theta}}_\omega) \right\} \frac{\partial^2 \tilde{\boldsymbol{\theta}}_\omega}{\partial \omega_r \partial \omega_s} \Big|_{\omega=\omega_0} \quad \text{regra da cadeia} \\
&= \left[\left(\frac{\partial \tilde{\boldsymbol{\theta}}_\omega}{\partial \omega_r} \right)^T \left\{ \frac{\partial^2}{\partial \tilde{\boldsymbol{\theta}}_\omega \partial \tilde{\boldsymbol{\theta}}_\omega} l(\tilde{\boldsymbol{\theta}}_\omega) \right\} \frac{\partial \tilde{\boldsymbol{\theta}}_\omega}{\partial \omega_s} \right] \Big|_{\omega=\omega_0} \\
&\quad + \left\{ \frac{\partial l(\tilde{\boldsymbol{\theta}}_\omega)}{\partial \tilde{\boldsymbol{\theta}}_\omega} \frac{\partial^2 \tilde{\boldsymbol{\theta}}_\omega}{\partial \omega_r \partial \omega_s} \right\} \Big|_{\omega=\omega_0}, \tag{2.19}
\end{aligned}$$

Como, por hipótese, $l(\boldsymbol{\theta}/\omega_0) = l(\boldsymbol{\theta})$, para todo $\boldsymbol{\theta} \in \mathbf{R}^p$, tem-se $\hat{\boldsymbol{\theta}}_{\omega_0} = \hat{\boldsymbol{\theta}}$, e portanto

$$\frac{\partial}{\partial \tilde{\boldsymbol{\theta}}_\omega} l(\tilde{\boldsymbol{\theta}}_\omega) \Big|_{\omega=\omega_0} = \frac{\partial}{\partial \boldsymbol{\theta}} l(\boldsymbol{\theta}/\omega_0) \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} = 0,$$

já que é o ponto de máximo da função de log-verossimilhança, tanto no modelo postulado $l(\boldsymbol{\theta})$, quanto no modelo perturbado $l(\boldsymbol{\theta}, \omega_0)$. Com isso, a segunda parcela da derivada

$\ddot{F}_{rs}$, da equação (2.19), é nula. Portanto

$$\ddot{F}_{rs} = \left[\left(\frac{\partial \tilde{\boldsymbol{\theta}}_{\omega}}{\partial \omega_r} \right)^T \left(\frac{\partial^2 l(\tilde{\boldsymbol{\theta}}_{\omega})}{\partial \tilde{\boldsymbol{\theta}}_{\omega} \partial \tilde{\boldsymbol{\theta}}_{\omega}^T} \right) \left(\frac{\partial \tilde{\boldsymbol{\theta}}_{\omega}}{\partial \omega_s} \right) \right] \Big|_{\omega=\omega_0}, \quad r, s = 1, \dots, q$$

Em notação matricial, podemos escrever

$$\ddot{\mathbf{F}} = \mathbf{J}^T \ddot{\mathbf{L}} \mathbf{J}, \quad (2.20)$$

onde $\ddot{\mathbf{F}} = (\ddot{F}_{rs})$, $\mathbf{J}^T = \frac{\partial \boldsymbol{\theta}^T}{\partial \boldsymbol{\omega}}$ e $\ddot{\mathbf{L}} = \frac{\partial^2 l(\tilde{\boldsymbol{\theta}}_{\omega})}{\partial \tilde{\boldsymbol{\theta}}_{\omega} \partial \tilde{\boldsymbol{\theta}}_{\omega}^T}$ avaliadas em $\omega = \omega_0$. A fim de obtermos uma expressão para a matriz $\mathbf{J}$, vamos usar o fato de que $\tilde{\boldsymbol{\theta}}_{\omega}$ é um ponto de máximo local da log-verossimilhança perturbada $l(\boldsymbol{\theta}/\omega)$. Assim, para cada $\omega \in \Omega$,

$$\frac{\partial l(\tilde{\boldsymbol{\theta}}/\omega)}{\partial \boldsymbol{\theta}} \Big|_{\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}_{\omega}} = 0.$$

O primeiro membro dessa expressão é uma função das q variáveis. Já o segundo membro dessa expressão é uma função constante, a função identicamente nula. Vamos agora derivar os dois membros da equação (2.19), com relação a ω , e em seguida avaliar a derivada no ponto ω_0 . Derivando o primeiro membro obtemos :

$$\begin{aligned} \frac{\partial}{\partial \omega} \left(\frac{\partial l(\tilde{\boldsymbol{\theta}}_{\omega}/\omega)}{\partial \boldsymbol{\theta}^T} \right) \Big|_{\omega=\omega_0} &= \left(\frac{\partial \tilde{\boldsymbol{\theta}}_{\omega}^T}{\partial \omega} \right) \frac{\partial^2 l(\tilde{\boldsymbol{\theta}}_{\omega}/\omega)}{\partial \tilde{\boldsymbol{\theta}}_{\omega} \partial \boldsymbol{\theta}^T} \Big|_{\omega=\omega_0} + \left(\frac{\partial \omega^T}{\partial \omega} \right) \frac{\partial^2 l(\tilde{\boldsymbol{\theta}}_{\omega}/\omega)}{\partial \omega \partial \boldsymbol{\theta}^T} \Big|_{\omega=\omega_0} \\ &= \left(\frac{\partial \tilde{\boldsymbol{\theta}}_{\omega}^T}{\partial \omega} \right) \frac{\partial^2 l(\tilde{\boldsymbol{\theta}}_{\omega}/\omega)}{\partial \tilde{\boldsymbol{\theta}}_{\omega} \partial \boldsymbol{\theta}^T} \Big|_{\omega=\omega_0} + \frac{\partial^2 l(\tilde{\boldsymbol{\theta}}_{\omega}/\omega)}{\partial \omega \partial \boldsymbol{\theta}^T} \Big|_{\omega=\omega_0}. \end{aligned}$$

Observe que

$$\frac{\partial^2 l(\tilde{\boldsymbol{\theta}}_{\omega}/\omega)}{\partial \tilde{\boldsymbol{\theta}}_{\omega} \partial \boldsymbol{\theta}^T} \Big|_{\omega=\omega_0} = \frac{\partial^2 l(\tilde{\boldsymbol{\theta}}_{\omega_0}/\omega_0)}{\partial \tilde{\boldsymbol{\theta}}_{\omega_0} \partial \boldsymbol{\theta}^T} = \frac{\partial^2 l(\tilde{\boldsymbol{\theta}})}{\partial \tilde{\boldsymbol{\theta}} \partial \boldsymbol{\theta}^T}$$

$$= \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}.$$

Assim, de (2.21) e dos resultados acima

$$\begin{aligned} \frac{\partial}{\partial \boldsymbol{\omega}} \left(\frac{\partial l(\tilde{\boldsymbol{\theta}}_{\boldsymbol{\omega}}/\boldsymbol{\omega})}{\partial \boldsymbol{\theta}^T} \right) \Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0} &= \frac{\partial \tilde{\boldsymbol{\theta}}_{\boldsymbol{\omega}}^T}{\partial \boldsymbol{\omega}} \Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0} \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \Big|_{\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}} + \frac{\partial^2 l(\boldsymbol{\theta}/\boldsymbol{\omega})}{\partial \boldsymbol{\omega} \partial \boldsymbol{\theta}^T} \Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0, \boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}}, \\ &= \mathbf{J}^T \ddot{\mathbf{L}} + \Delta^T = 0, \end{aligned}$$

onde $\mathbf{J}$ e $\ddot{\mathbf{L}}$ são como em (2.20) e $\Delta^T = \frac{\partial^2 l(\boldsymbol{\theta}/\boldsymbol{\omega})}{\partial \boldsymbol{\omega} \partial \boldsymbol{\theta}^T} \Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0, \boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}}$, ou seja $\mathbf{J}^T = -\Delta^T \ddot{\mathbf{L}}^{-1}$.

Agora, substituindo as expressões em $\ddot{\mathbf{F}}$ obtemos:

$$\begin{aligned} \ddot{\mathbf{F}} &= \left(\frac{\partial \tilde{\boldsymbol{\theta}}_{\boldsymbol{\omega}}^T}{\partial \boldsymbol{\omega}} \right) \Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0} \left(\frac{\partial^2 l(\tilde{\boldsymbol{\theta}}_{\boldsymbol{\omega}})}{\partial \tilde{\boldsymbol{\theta}}_{\boldsymbol{\omega}} \partial \tilde{\boldsymbol{\theta}}_{\boldsymbol{\omega}}^T} \right) \Big|_{\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}} \left(\frac{\partial \tilde{\boldsymbol{\theta}}_{\boldsymbol{\omega}}^T}{\partial \boldsymbol{\omega}} \right)^T \Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0} \\ &= \left(-\Delta^T \ddot{\mathbf{L}}^{-1} \right) \ddot{\mathbf{L}} \left(-\Delta^T \ddot{\mathbf{L}}^{-1} \right)^T \\ &= \Delta^T \ddot{\mathbf{L}}^{-1} \Delta \end{aligned}$$

com isto, a curvatura normal da superfície $\alpha(\boldsymbol{\omega})$, na direção de um vetor unitário $\mathbf{d}$, fica dada por

$$C_d = 2 \left| \mathbf{d}^T \Delta^T \ddot{\mathbf{L}}^{-1} \Delta \mathbf{d} \right|. \quad (2.21)$$

A curvatura normal C_d da superfície de afastamento pela verossimilhança, $\alpha(\boldsymbol{\omega})$, na direção $\mathbf{d}$ é sugerida para medir a sensibilidade da análise. Valores grandes de C_d indicam sensibilidade para as perturbações introduzidas na direção $\mathbf{d}$. A curvatura máxima, que denotamos por $C_{\max}$, quando $\mathbf{d}$ varia, na direção correspondente $\mathbf{d}_{\max}$ de uma linha projetada na direção $\mathbf{d}_{\max}$, isto é, $LD(\boldsymbol{\omega}_0 + t\mathbf{d}_{\max})$ versus t , é uma quantidade importante no método de diagnóstico do método de influência local proposto por Cook (1986). Valores grandes de $C_{\max}$ indicam influência local notável. Um valor grande de $C_{\max}$ é melhor determinado por experiência, por exemplo, o valor $C_{\max} = 2$ parece ser

um proveitoso guia quando usamos perturbação de ponderação de casos em regressão linear. (veja, por exemplo, Cook, 1987).

Com a equação (2.21) podemos avaliar a influência que pequenas perturbações podem exercer sobre os componentes do modelo, tais como estimativas dos parâmetros, e outros resultados da análise. Interesse particular está na direção que produz a maior influência local. Essa direção $\mathbf{d}_{\max}$ é o autovetor normalizado correspondente ao maior autovalor $C_{\max}$ da matriz $\ddot{\mathbf{F}} = \Delta^T \ddot{\mathbf{L}}^{-1} \Delta$. Com o vetor $\mathbf{d}_{\max}$ identificamos os fatores mais influentes para o esquema de perturbação em análise.

Suponhamos agora uma partição $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^T, \boldsymbol{\theta}_2^T)^T$, onde $\boldsymbol{\theta}_1$ é de dimensão $p_1 \times 1$ e $\boldsymbol{\theta}_2$ é de dimensão $p_2 \times 1$, com $p_1 + p_2 = p$, e suponhamos que nosso interesse reside em $\boldsymbol{\theta}_1^T$. Nesse caso, o gráfico de influência é

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

onde

$$LD_s(\boldsymbol{\omega}) = 2 \left[l(\tilde{\boldsymbol{\theta}}_1, \tilde{\boldsymbol{\theta}}_2) - l\{\tilde{\boldsymbol{\theta}}_{1\omega}, g(\tilde{\boldsymbol{\theta}}_{1\omega})\} \right] \quad (2.22)$$

com $g : \mathbf{R}^{p_1} \rightarrow \mathbf{R}^{p_2}$ uma função implícita tal que, para $\boldsymbol{\theta}_1$ fixado, a função $l\{\boldsymbol{\theta}_1, g(\boldsymbol{\theta}_1)\}$ é maximizada e representa a log-verossimilhança perfilada para $\boldsymbol{\theta}_1$. O estimador $\hat{\boldsymbol{\theta}}_{1\omega}$ é determinado a partir de $\tilde{\boldsymbol{\theta}}_{\omega}^T = \begin{pmatrix} \tilde{\boldsymbol{\theta}}_{1\omega}^T & \tilde{\boldsymbol{\theta}}_{2\omega}^T \end{pmatrix}$.

Consideremos a partição

$$\ddot{\mathbf{L}} = \begin{pmatrix} \ddot{\mathbf{L}}_{11} & \ddot{\mathbf{L}}_{12} \\ \ddot{\mathbf{L}}_{21} & \ddot{\mathbf{L}}_{22} \end{pmatrix},$$

onde $(-\ddot{\mathbf{L}})$ é a matriz observada de informação de Fisher, no modelo postulado. A curvatura normal da superfície $\alpha_s(\boldsymbol{\omega})$ na direção de um vetor unitário $\mathbf{d}$ é a expressão

$$C_d(\boldsymbol{\theta}_1) = 2 [\mathbf{d}^T \Delta (L^{-1} - \mathbf{B}_{22}) \Delta \mathbf{d}],$$

onde

$$\mathbf{B}_{22} = \begin{pmatrix} 0 & 0 \\ 0 & \mathbf{L}_{22}^{-1} \end{pmatrix}.$$

Deve-se dar atenção à direção $\mathbf{d}_{\max}$, que produz a maior influência local, a qual pode ser calculada a partir da equação (2.22), e que é dada pelo autovetor normalizado correspondente ao maior autovalor da matriz $\Delta (\mathbf{L}^{-1} - \mathbf{B}_{22})$.

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

2.4.1 Aplicação

Ponderação de casos:

Seja $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n)^T$ o vetor $n \times 1$ de perturbação de ponderação de casos para o modelo de regressão linear definido em (2.1), com σ^2 conhecido. Na seção 2.2.3, a função de log-verossimilhança para o modelo de regressão linear é dada por

$$\begin{aligned} l(\boldsymbol{\beta}) &= \sum_{j=1}^n l_j(\boldsymbol{\beta}), \\ &= -\frac{n}{2} \log 2\pi\sigma^2 - \frac{n}{2\sigma^2} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}), \end{aligned} \quad (2.23)$$

onde

$$l_j(\boldsymbol{\beta}) = -\frac{1}{2} \log 2\pi\sigma^2 - \frac{1}{2\sigma^2} (y_j - \mathbf{x}_j^T \boldsymbol{\beta})^2. \quad (2.24)$$

A função de log-verossimilhança do modelo perturbado é dada por

$$l(\boldsymbol{\beta}/\boldsymbol{\omega}) = \sum_{j=1}^n \omega_j l_j(\boldsymbol{\beta}) = -\frac{1}{2} \sum_{j=1}^n \omega_j \log (2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{j=1}^n \omega_j (y_j - \mathbf{x}_j^T \boldsymbol{\beta})^2, \quad (2.25)$$

onde ω_j e y_j são o j -ésimo componente de ω e $\mathbf{Y}$, respectivamente, e $\mathbf{x}_j^T$ é a j -ésima linha de $\mathbf{X}$. Derivando a expressão em (2.25) com respeito a β temos que,

$$\frac{\partial l(\beta/\omega)}{\partial \beta} = \frac{1}{\sigma^2} \sum_{j=1}^n \omega_j (y_j - \mathbf{x}_j^T \beta) \mathbf{x}_j.$$

Derivando a expressão anterior com respeito a ω obtemos que

$$\begin{aligned} \frac{\partial^2 l(\beta/\omega)}{\partial \omega \partial \beta^T} &= \frac{1}{\sigma^2} \frac{\partial}{\partial \omega} \left(\sum_{j=1}^n \omega_j (y_j - \mathbf{x}_j^T \beta) \mathbf{x}_j^T \right) \\ &= \frac{1}{\sigma^2} \begin{pmatrix} (y_1 - \mathbf{x}_1^T \beta) \mathbf{x}_1^T \\ \vdots \\ (y_n - \mathbf{x}_n^T \beta) \mathbf{x}_n^T \end{pmatrix} \end{aligned}$$

avaliando a expressão anterior em $\beta = \hat{\beta}$, temos que

$$\begin{aligned} \Delta^T &= \left. \frac{\partial^2 l(\beta/\omega)}{\partial \omega \partial \beta^T} \right|_{\beta=\hat{\beta}} = \frac{1}{\sigma^2} \begin{pmatrix} e_1 \mathbf{x}_1^T \\ \vdots \\ e_n \mathbf{x}_n^T \end{pmatrix} \\ &= \frac{1}{\sigma^2} D(\mathbf{e}) \mathbf{X}, \end{aligned}$$

onde $\mathbf{e} = (e_1, \dots, e_n)^T$ é o vetor de resíduos ordinários e $D(\mathbf{e}) = \text{diag}(e_1, \dots, e_n)$. Assim $\Delta = \frac{1}{\sigma^2} \mathbf{X}^T D(\mathbf{e})$.

Matriz Observada de Informação de Fisher

Seja $l(\beta)$ a função de log-verossimilhança definida em (2.23). Então a matriz de derivadas parciais de segunda ordem com respeito a β é dada por,

$$\bar{L} = \frac{\partial l(\beta)}{\partial \beta \partial \beta^T} = \sum_{j=1}^n \frac{\partial l_j(\beta)}{\partial \beta \partial \beta^T},$$

e derivando $l(\beta)$ com respeito a β , obtemos

$$\begin{aligned}
\frac{\partial l(\beta)}{\partial \beta} &= \frac{\partial}{\partial \beta} \left(-\frac{1}{2\sigma^2} (\mathbf{y} - \mathbf{X}\beta)^T (\mathbf{y} - \mathbf{X}\beta) \right) \\
&= -\frac{1}{2\sigma^2} \frac{\partial}{\partial \beta} (\mathbf{y} - \mathbf{X}\beta)^T \times \frac{\partial (\mathbf{y} - \mathbf{X}\beta)^T (\mathbf{y} - \mathbf{X}\beta)}{\partial (\mathbf{y} - \mathbf{X}\beta)} \\
&= \frac{1}{2\sigma^2} \mathbf{X}^T \times 2 (\mathbf{y} - \mathbf{X}\beta) \\
&= \frac{1}{\sigma^2} \mathbf{X}^T (\mathbf{y} - \mathbf{X}\beta).
\end{aligned}$$

Agora,

$$\begin{aligned}
\frac{\partial}{\partial \beta} \left(\frac{\partial l(\beta)}{\partial \beta^T} \right) &= \frac{\partial}{\partial \beta} \left(\frac{1}{\sigma^2} (\mathbf{y} - \mathbf{X}\beta)^T \mathbf{X} \right) = \frac{\partial}{\sigma^2 \partial \beta} (\mathbf{y}^T - \beta^T \mathbf{X}^T) \mathbf{X} \\
&= -\frac{1}{\sigma^2} (\mathbf{X}^T \mathbf{X}).
\end{aligned}$$

Logo,

$$\tilde{L}(\hat{\beta}) = -\frac{1}{\sigma^2} \mathbf{X}^T \mathbf{X} \quad (2.27)$$

Substituindo 2.26 e 2.27 em $C_d = 2 \left| \mathbf{d}^T \Delta^T \tilde{L}^{-1} \Delta \mathbf{d} \right|$, temos que

$$\begin{aligned}
C_d &= 2 \left| \mathbf{d}^T \left(\frac{\mathbf{X}^T D(\mathbf{e})}{\sigma^2} \right)^T \left(-\frac{\mathbf{X}^T \mathbf{X}}{\sigma^2} \right)^{-1} \left(\frac{\mathbf{X}^T D(\mathbf{e})}{\sigma^2} \right) \mathbf{d} \right| \\
&= 2 \left| \mathbf{d}^T D(\mathbf{e}) \underbrace{\mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T}_{\mathbf{H}} D(\mathbf{e}) \mathbf{d} \right| \\
&= \frac{2}{\sigma^2} \left| \mathbf{d}^T D(\mathbf{e}) \mathbf{H} D(\mathbf{e}) \mathbf{d} \right|,
\end{aligned} \quad (2.28)$$

onde $\mathbf{H} = \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T$ e $\|\mathbf{d}\| = 1$. Observe-se que a curvatura depende dos resíduos ordinários, da matriz chapéu e da direção.

Ilustração Numérica

A seguir, apresentamos uma ilustração numérica da metodologia proposta, considerando um conjunto de dados tomados de Weisberg (1985), relativos a um estudo de um método de pesquisa aérea para estimar o número de gansos brancos. Para obter a estimativa, um avião pequeno voa sobre a área do habitat dos gansos e quando um bando de gansos é visualizado, uma pessoa com experiência estima o número de gansos.

Para investigar a confiabilidade deste método de contagem, um experimento foi conduzido, considerando um avião, com dois observadores, voando sobre $n = 45$ bandos, e cada observador de forma independente estima o número de aves em cada bando. Uma fotografia do bando é tomada para obter-se uma contagem exata do número de aves.

Seja Y : verdadeiro tamanho do bando (coleção de gansos) que foi obtido por contagem de fotografia aérea. (foto por satélite) e X : tamanho do bando estimado visualmente para uma amostra de $n = 45$ bandos de gansos brancos, estimado pelo primeiro observador. Considerando o modelo de regressão linear mostramos na Figura 2.7 um gráfico de X vs Y , onde pode-se observar uma forte evidência de heterocedasticidade. Portanto, pode-se esperar que os coeficientes das linhas ajustadas poderiam ser sensíveis à menor modificação em perturbação de ponderação de casos.

Utilizando a metodologia de Cook (1986), sob o esquema de perturbação de ponderação de casos, encontramos que $C_{\max} = 14.37$ calculada de (2.28) com $\sigma^2 = \tilde{\sigma}^2$ (EMV), indica sensibilidade local extrema. Os elementos do autovetor $d_{\max}$ correspondente ao maior autovalor de $\Delta^T \left(\ddot{\mathbf{L}} \right)^{-1} \Delta$ são grandes para as observações 28, 29 e 41. A Figura 2.8 mostra um gráfico de $d_{\max}$ vs X . Observa-se claramente que $d_{\max}$ está respondendo à essencial heterocedasticidade nos dados. Geralmente, tais inspeções de $d_{\max}$ podem proporcionar informação de diagnóstico proveitoso que pode ser usado para guia de subseqüentes análises e experimentação futura. Para esses dados apresentados anteriormente há uma forte relação entre os resíduos ordinários e, a regressão linear simple de Y sobre X e os elementos de $d_{\max}$. Por exemplo, um gráfico (não mostrado) dos elementos em valor absoluto de $d_{\max}$ vs e mostra um padrão muito forte em forma de U . Os casos

com residuais em valor absoluto grandes tiendem a ser enfatizados por $d_{\max}$, enquanto casos com residuais em valor absoluto pequenos são desenfaturizados. O vetor $d_{\max}$ mostra atenção direta para os casos onde a especificação de ponderação é mais importante.

Tabela 2.9.: Estimativa do número de gansos brancos pelo primeiro observador em Weisberg (1985).

Obs.	X	Y	Obs.	X	Y	Obs.	X	Y
1	50	56	16	30	26	31	50	73
2	25	38	17	75	88	32	75	123
3	30	25	18	35	56	33	150	150
4	35	48	19	9	11	34	50	70
5	25	38	20	55	66	35	60	90
6	20	22	21	30	42	36	75	110
7	12	22	22	25	30	37	150	95
8	34	42	23	40	90	38	40	57
9	20	34	24	75	119	39	25	43
10	10	14	25	100	165	40	100	55
11	25	30	26	150	152	41	200	325
12	10	9	27	120	205	42	60	114
13	15	18	28	250	409	43	40	83
14	20	25	29	500	342	44	35	91
15	40	62	30	200	200	45	20	56

X : Tamanho do bando estimado visualmente pelo primeiro observador.

Y : Verdadeiro tamanho do bando (coleção de gansos) que foi obtido por contagem de fotografia aérea .

Tabela 2.10: Autovetor correspondente ao maior autovalor de $\Delta^T \left(\overset{\bullet\bullet}{L} \right)^{-1} \Delta$ nos dados de Ganso

Obs.	dmax	Obs.	dmax	Obs.	dmax
1	0,0011	16	-0,0058	31	-0,0002
2	-0,0030	17	0,0021	32	-0,0131
3	-0,0060	18	-0,0002	33	0,0136
4	-0,0013	19	-0,0121	34	0,0001
5	-0,0030	20	0,0013	35	-0,0023
6	-0,0079	21	-0,0024	36	-0,0074
7	-0,0072	22	-0,0053	37	0,0968
8	-0,0023	23	0,0019	38	-0,0003
9	-0,0037	24	-0,0114	39	-0,0016
10	-0,0107	25	-0,0397	40	0,0475
11	-0,0053	26	0,0106	41	-0,2716
12	-0,0132	27	-0,0783	42	-0,0075
13	-0,0094	28	-0,4767	43	0,0014
14	-0,0069	29	0,8234	44	0,0047
15	0,0000	30	0,0070	45	0,0042

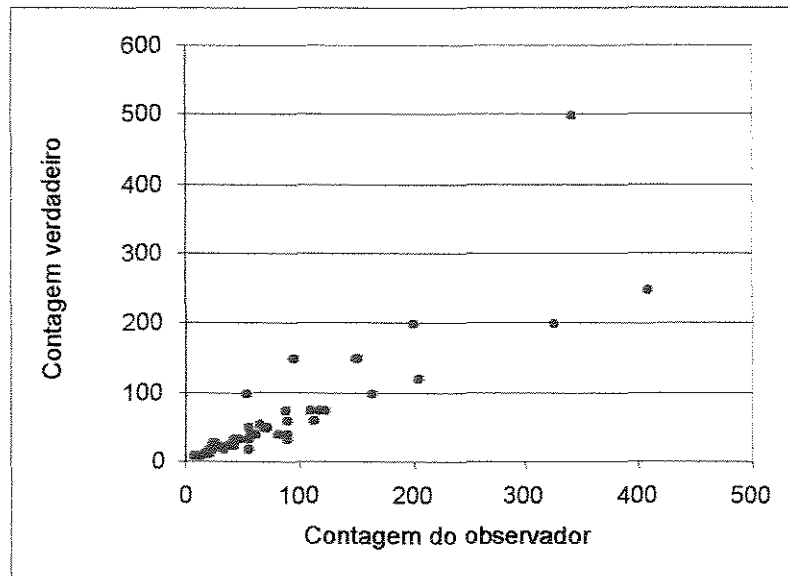


Figura 2-7: Gráfico de dispersão dos dados de Ganso

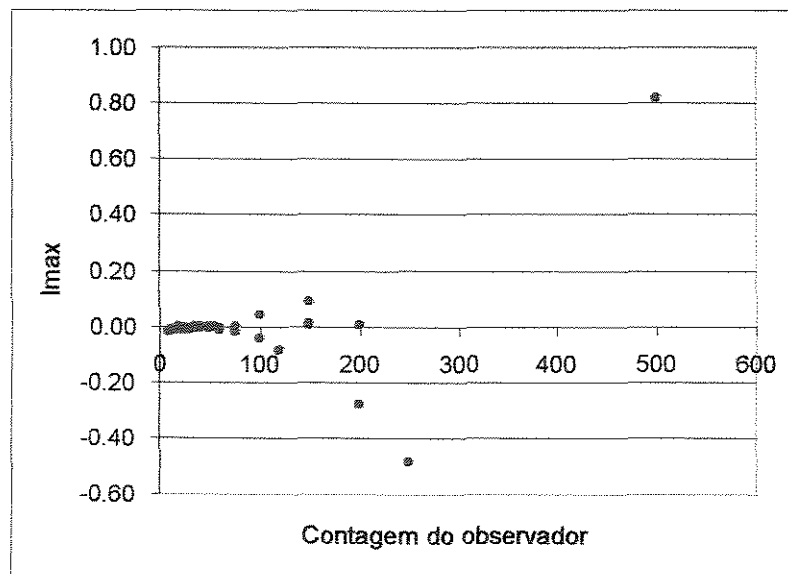


Figura 2-8: Gráfico de d_{max} vs contagem do observador

Capítulo 3

Modelo de Calibração Comparativa

3.1 Introdução

O objetivo principal deste capítulo é apresentar um estudo de inferência sob modelo de calibração comparativa, onde é de interesse comparar vários instrumentos de medição da mesma quantidade desconhecida x em um grupo comum de n unidades experimentais não destrutiva supondo que as medições dos instrumentos seguem uma distribuição normal multivariada. Este problema é conhecido na literatura como calibração comparativa, cujo objetivo é avaliar a qualidade dos diferentes instrumentos ou métodos de medição.

A necessidade de comparar instrumentos de medição que diferem em custo, velocidade ou outros fatores, tem aparecido com frequência em diversas áreas do conhecimento. Grubbs (1948, 1973), por exemplo, compara três cronômetros, Barnett (1969) dá um exemplo onde quatro combinações instrumento-operador concebidos para medir a capacidade vital em um grupo de pacientes são avaliadas, Leurgans (1980) compara dois métodos para medir a concentração de glicose no sangue; vários outros exemplos na área médica são mencionados por Kelly (1984, 1985) e Kaaks et al. (1994). Já Jaech (1985) dá vários exemplos na área industrial, Fuller (1987) apresenta exemplos na área da agricultura. Dunn (1992) dá aplicações em psicologia e educação.

Quando o processo de medição não é destrutivo todos os instrumentos podem ser usados em cada unidade experimental. Se o processo é destrutivo, assume-se que cada unidade (experimental) fornece partes homogêneas, que são atribuídas ao acaso a cada um dos métodos de medição.

O termo “método” ou “instrumento” de medição refere-se a diferentes formas de medição de uma característica dada. O termo também pode ser aplicado quando a mesma forma de medição é feita por diferentes aparelhos ou diferentes operadores. Por exemplo, quando uma análise química é realizada em diferentes laboratórios, cada laboratório é identificado como um método ou instrumento de medição. Em alguns casos pode acontecer, também, que os diferentes métodos sejam um mesmo instrumento avaliado em diferentes tempos. Na prática, a definição do método de medição terá influência sobre a construção do modelo estatístico.

3.2 Modelo de Calibração Comparativa

Nesta seção apresentamos um estudo do modelo de calibração comparativa para três instrumentos baseado no modelo proposto por Barnet (1969) e Williams (1969), que é definido por

$$\begin{cases} Y_{1j} = \alpha_1 + \beta_1 x_j + e_{1j} \\ Y_{2j} = \alpha_2 + \beta_2 x_j + e_{2j} \\ Y_{3j} = \alpha_3 + \beta_3 x_j + e_{3j}, \end{cases} \quad (3.1)$$

onde

Y_{ij} : medida fornecida pelo instrumento (método de medição) i associado ao indivíduo j ,

x_j : verdadeira medida desconhecida do indivíduo j , não observável,

α_i : vício aditivo do instrumento i ,

β_i : vício multiplicativo do instrumento i ,

e_{ij} : erros de medição do instrumento i associados ao indivíduo j , $i = 1, 2, 3$, $j = 1, \dots, n$.

Na literatura é assumido que os erros de medição $e_{i1}, \dots, e_{in}, i = 1, 2, 3$ são independentes com distribuição normal de média zero e variância σ_i^2 .

Como podemos observar da definição em (3.1), o modelo de calibração comparativa é uma classe particular do modelo com erros nas variáveis (veja por exemplo, Kendall e Stuart, 1979 e Fuller, 1987). Assim, o modelo de calibração comparativa pode ser estudado sob duas especificações diferentes:

•**Modelo de Calibração Funcional:** É obtido ao supor que as quantidades desconhecidas $x_1, \dots, x_n$ são constantes, denominados parâmetros incidentais. Neste caso, vamos ter $n + 9$ parâmetros, e como esse número cresce com o número de observações. Para fazer um estudo de inferência neste modelo é necessário considerar suposições adicionais de alguns parâmetros, especialmente nas quantidades desconhecidas $x_1, \dots, x_n$. Alguns resultados de inferência neste modelo podem ser encontrados, por exemplo, em Mak(1982), Kimura(1992), Vilca et al.(1998), Bolfarine e Rojas (1995) e Williams(1969). Sob o modelo de calibração funcional os parâmetros são $x_1, \dots, x_n, \alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3, \sigma_1^2, \sigma_2^2, \sigma_3^2$.

•**Modelo de Calibração Estrutural:** É obtido ao considerar que as variáveis não observadas $x_1, \dots, x_n$ são variáveis aleatórias não correlacionadas com média comum μ_x e variância σ_x^2 ; ao contrário do modelo funcional, neste modelo não é possível encontrar uma solução de máxima verossimilhança para os parâmetros, pois o modelo não é identificável, no sentido de que dois parâmetros diferentes podem dar origem à mesma distribuição. Logo, é necessário o conhecimento adicional de algum parâmetro para possibilitar a estimativa dos parâmetros de maior interesse. Alguns resultados de inferência baseados no modelo estrutural podem ser encontrados, em Barnett (1969), Theobald e Mallison (1978), Carter (1981), e Bolfarine e Galea(1995b), entre outros. Sob o modelo

de calibração estrutural, o vetor de parâmetros do modelo é dado por

$$\boldsymbol{\theta} = (\mu_x, \alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3, \sigma_x^2, \sigma_1^2, \sigma_2^2, \sigma_3^2)^T. \quad (3.2)$$

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 vêm os indivíduos envolvidos no estudo.

No que segue deste capítulo, vamos apresentar um estudo de inferência no modelo estrutural de calibração comparativa baseado no modelo considerado por Barnett (1969), especial atenção no modelo de calibração comparativa sob a parametrização de análise de fatores considerada por Theobald e Mallison (1978).

Sob o modelo de calibração estrutural são apresentados os estimadores de máxima verossimilhança dos parâmetros do modelo. Considerando a parametrização de análise de fatores, apresentamos a matriz de informação observada e esperada que serão de utilidade para obter regiões de confiança e testar hipóteses de interesse. Além disso, será de grande utilidade na aplicação do método de influência local proposto por Cook (1986), no modelo de calibração comparativa.

3.3 O Modelo Estrutural de Regressão de Calibração

O modelo proposto por Barnett (1969) definido em (3.1) pode ser escrito de forma matricial como

$$\mathbf{Y}_j = \mathbf{a} + \mathbf{b}x_j + \mathbf{e}_j = \mathbf{a} + \mathbf{B}\mathbf{r}_j, \quad (3.3)$$

onde $\mathbf{Y}_j = (Y_{1j}, Y_{2j}, Y_{3j})^T$, $\mathbf{e}_j = (e_{1j}, e_{2j}, e_{3j})^T$, $\mathbf{r}_j = (x_j, \mathbf{e}_j^T)^T$, $\mathbf{a} = (\alpha_1, \alpha_2, \alpha_3)^T$, $\mathbf{b} = (\beta_1, \beta_2, \beta_3)^T$, $\mathbf{B} = [\mathbf{b}, \mathbf{I}_3]$, $j = 1, \dots, n$.

O modelo estrutural normal de calibração comparativa é obtido ao supor que

$$\mathbf{r}_j \sim N_4(\boldsymbol{\eta}_x, \boldsymbol{\Omega}), \quad (3.4)$$

onde $\boldsymbol{\eta}_x = (\mu_x, 0, 0, 0)^T$ e $\boldsymbol{\Omega} = \text{Diag}(\sigma_x^2, \sigma_1^2, \sigma_2^2, \sigma_3^2)$. Neste caso,

$$\mathbf{E}(\mathbf{r}_j) = \boldsymbol{\eta}_x \quad \text{e} \quad \text{Var}(\mathbf{r}_j) = \boldsymbol{\Omega} = \begin{pmatrix} \sigma_x^2 & \mathbf{O}_3^T \\ \mathbf{O}_3 & D(\boldsymbol{\phi}) \end{pmatrix},$$

onde $D(\boldsymbol{\phi}) = \text{Diag}(\sigma_1^2, \sigma_2^2, \sigma_3^2)$, $\boldsymbol{\phi} = (\phi_1, \phi_2, \phi_3)^T = (\sigma_1^2, \sigma_2^2, \sigma_3^2)^T$ e $\mathbf{O}_3$ é um vetor nulo 3×1 , $j = 1, \dots, n$.

De (3.3) e (3.4) e das propriedades da distribuição normal temos que

$$\mathbf{Y}_j \sim N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma}),$$

onde

$$E(\mathbf{Y}_j) = \mathbf{a} + \mathbf{b}\mu_x = \boldsymbol{\mu} = \boldsymbol{\mu}(\boldsymbol{\theta}) \quad \text{e} \quad \text{Var}(\mathbf{Y}_j) = D(\boldsymbol{\phi}) + \sigma_x^2 \mathbf{b}\mathbf{b}^T = \boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\boldsymbol{\theta}),$$

com $\boldsymbol{\theta}$ como em (3.2).

Sob as condições do modelo de calibração definido em (3.3), $\boldsymbol{\theta}$ não é identificável (ver Bolfarine e Galea, 1995b), isto é, existem $\boldsymbol{\theta}_*$ e $\boldsymbol{\theta}_{**}$ tais que $F(\mathbf{y}; \boldsymbol{\theta}_*) = F(\mathbf{y}; \boldsymbol{\theta}_{**})$, para todo $\mathbf{y}$, onde F é a distribuição de $\mathbf{Y}_j$, $j = 1, \dots, n$. Por exemplo, se

$$\boldsymbol{\theta}_* = (1, 1, 1, 1, 1, 1, 1, 1, 1, 1)^T = \mathbf{1}_{11}^T,$$

e

$$\boldsymbol{\theta}_{**} = (2, -2, -2, -2, 2, 2, 2, 0.25, 1, 1)^T,$$

então

$$\mathbf{E}_{\theta.}(\mathbf{Y}_j) = \mathbf{E}_{\theta..}(\mathbf{Y}_j) = (2, 2, 2)^T$$

e

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

A falta de identificabilidade implica ausência de estimadores consistentes, o que tem sérias implicações na teoria assintótica de estimação e testes de hipóteses. Contudo, o fato de θ ser identificável não implica que existe um estimador consistente para θ . Para um exemplo interessante ver Gabielsen (1978).

Uma das formas de contornar o problema de identificabilidade é colocar restrições sobre o parâmetro θ . Neste caso, Barnett (1969) assume que existe um método de medição, que ele chama instrumento de referência, que mede x sem vício. Especificamente, supondo que o instrumento de referência é o número 1, Barnett (1969) assume que

$$\alpha_1 = 0 \text{ e } \beta_1 = 1. \quad (3.5)$$

Assim, o modelo definido em (2.3) pode ser escrito por

$$\mathbf{Y}_j = \mathbf{a} + \mathbf{b}x_j + \mathbf{e}_j = \mathbf{a} + \mathbf{B}r_j, \quad j = 1, \dots, n, \quad (3.6)$$

onde $\mathbf{a} = (0, \boldsymbol{\alpha}^T)^T$, $\mathbf{b} = (1, \boldsymbol{\beta}^T)^T$, com $\boldsymbol{\alpha} = (\alpha_2, \alpha_3)^T$ e $\boldsymbol{\beta} = (\beta_2, \beta_3)^T$.

Sob a condição de identificabilidade dada em (3.5), o vetor de parâmetros dado em (3.2) reduz-se a

$$\boldsymbol{\theta} = (\mu_x, \boldsymbol{\alpha}^T, \sigma_x^2, \boldsymbol{\beta}^T, \boldsymbol{\phi}^T)^T.$$

Sob as condições acima, Barnett (1969) consegue estimativas explícitas para os elementos de $\boldsymbol{\theta}$, usando o método de momentos. No entanto, pode ocorrer que algu-

mas estimativas das variâncias $(\sigma_x^2, \phi^T)^T$ sejam negativas com probabilidade positiva. Barnett (1969) também assume normalidade dos $\mathbf{r}'_j$ s e logo normalidade dos $\mathbf{Y}'_j$ s; neste caso, se as estimativas das variâncias são não-negativas, os estimadores propostos por Barnett correspondem aos de máxima verossimilhança.

A possibilidade dos estimadores propostos por Barnett(1969) serem negativos é contornada por Carter (1981), usando estimação restrita de máxima verossimilhança.

3.3.1 Definição de Confiabilidade e Precisão

O problema de comparar instrumentos ou métodos de medição pode ser estudado fazendo inferências sobre os vícios $\alpha_2, \alpha_3, \beta_2$ e β_3 dos instrumentos 2 e 3 e comparando-os com o instrumento de referência.

Dois conceitos importantes na comparação de instrumentos de medição, que permite obter mais informação para avaliar a qualidade do instrumento, é a confiabilidade e a precisão.

Definição 3.1: A confiabilidade do i -ésimo instrumento, denotada por ρ_i , define-se por:

$$\rho_i = \frac{\beta_i^2 \sigma_x^2}{\beta_i^2 \sigma_x^2 + \sigma_i^2}, \quad i = 1, 2, 3.$$

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” pela i -ésima linha de calibração e $0 \leq \rho_i \leq 1$, $i = 1, 2, 3$. Esta interpretação decorre do fato que

$$\rho_{Y_{ij}, x_j}^2 = \frac{\text{cov}^2(Y_{ij}, x_j)}{\text{Var}(Y_{ij}) \text{Var}(x_j)} = \rho_i,$$

onde $\text{Cov}(Y_{ij}, x_j) = \beta_i \sigma_x^2$, $\text{Var}(x_j) = \sigma_x^2$ e $\text{Var}(Y_{ij}) = \beta_i^2 \sigma_x^2 + \sigma_i^2$.

Uma outra medida ou índice usado para avaliar a qualidade de um instrumento ou método de medição é a precisão. Se conhecemos os parâmetros α_i e β_i , então as ob-

servações Y_{ij} , $j = 1, \dots, n$, no instrumento i podem ser calibradas como

$$Y_{ij}^* = \frac{1}{\beta_i} (Y_{ij} - \alpha_i) = x_j + \frac{1}{\beta_i} e_{ij} = x_j + e_{ij}^*,$$

as quais são “não-viciadas” para x_j , $j = 1, \dots, n$. Os erros de medição e_{ij}^* de Y_{ij}^* têm variância σ_i^2 / β_i^2 . Já que depois da calibração todos os instrumentos medem x_j sem vício, uma definição natural da precisão seria (ver Shyr e Gleser, 1986 e Cochran, 1968):

Definição 3.2 : A precisão do instrumento i , denotada por π_i , é definida por

$$\pi_i = \frac{1}{\text{var}(e_{ij}^*)} = \frac{\beta_i^2}{\sigma_i^2}, \quad i = 1, 2, 3; \beta_1 = 1.$$

Theobald e Mallison (1978) usam a “precisão relativa” (denotada por τ_i), definida por

$$\tau_i = \pi_i \sigma_x^2 = \beta_i^2 \sigma_x^2 / \sigma_i^2, \quad i = 1, 2, 3; \beta_1 = 1.$$

No entanto, Draper e Guttman(1975, 1976) definem a precisão relativa (denotada por δ_i) como

$$\delta_i = \frac{1}{\tau_i} = \sigma_i^2 / \beta_i^2 \sigma_x^2, \quad i = 1, 2, 3.$$

Pode-se mostrar facilmente que

$$\rho_i = \rho_k \iff \pi_i = \pi_k \iff \tau_i = \tau_k \iff \delta_i = \delta_k, \text{ para todo } i \neq k = 1, 2, 3. \quad (3.7)$$

Das medidas anteriores, prefere-se a confiabilidade, já que esta não depende das unidades de medição dos x_j e é mais fácil de ser interpretada. Contudo, de (3.7) vemos que comparar confiabilidades é equivalente a comparar precisões ou precisões relativas.

Em geral o problema de comparar 3 instrumentos reduz-se a fazer inferências sobre os vícios (α_i, β_i) e as confiabilidades (ρ_i) , desde que o modelo é assumido correto.

É necessário considerar o vício e a confiabilidade simultaneamente, pois um valor alto de ρ_i (próximo de 1) implica só que Y_i (medida feita pelo instrumento i) é uma medida confiável de $\alpha_i + \beta_i x$. O instrumento i forneceria medidas “exatas” da característica x só se ρ_i fosse próximo de um e se ele não fosse viciado ($\alpha_i = 0, \beta_i = 1$). Uma hipótese inicial de interesse, usando o modelo (3.6), poderia ser:

$$H_{01} : \alpha_2 = \alpha_3 = 0, \beta_2 = \beta_3 = 1, \rho_1 = \rho_2 = \rho_3.$$

Ou seja, estamos interessados em testar se os 3 instrumentos medem a característica x sem vício e com a mesma confiabilidade.

Outra hipótese que pode ser testada é:

$$H_{02} : \alpha_2 = \alpha_3 = 0, \beta_2 = \beta_3 = 1.$$

Aqui queremos testar se os instrumentos medem x sem vício.

Para testar se os instrumentos são igualmente confiáveis, a hipótese seria

$$H_{03} : \rho_1 = \rho_2 = \rho_3.$$

3.3.2 Estimadores de Máxima Verossimilhança

Sob o modelo normal de calibração comparativa, Barnett (1969) (veja também, Nice, 1996) mostra que o estimador de máxima verossimilhança de θ é dado por $\tilde{\theta} = (\tilde{\mu}_x, \tilde{\alpha}^T, \tilde{\sigma}^2, \tilde{\beta}^T, \tilde{\phi}^T)^T$, onde

$$\begin{aligned} \tilde{\mu}_x &= \bar{Y}_1, \\ \tilde{\alpha}_2 &= \bar{Y}_2 - \tilde{\beta}_2 \bar{Y}_1, \tilde{\alpha}_3 = \bar{Y}_3 - \tilde{\beta}_3 \bar{Y}_1, \\ \tilde{\sigma}_x^2 &= \frac{S_{12}S_{13}}{S_{23}}, \\ \tilde{\beta}_2 &= \frac{S_{23}}{S_{13}}, \tilde{\beta}_3 = \frac{S_{23}}{S_{12}}, \end{aligned}$$

$$\begin{aligned}\tilde{\sigma}_1^2 &= S_{11} - \tilde{\sigma}_x^2, \quad \tilde{\sigma}_2^2 = S_{22} - \tilde{\beta}_2^2 \tilde{\sigma}_x^2, \\ \tilde{\sigma}_3^2 &= S_{33} - \tilde{\beta}_3^2 \sigma_x^2,\end{aligned}$$

com

$$\begin{aligned}\bar{Y}_1 &= \frac{1}{n} \sum_{j=1}^n Y_{1j}, \quad \bar{Y}_2 = \frac{1}{n} \sum_{j=1}^n Y_{2j}, \quad \bar{Y}_3 = \frac{1}{n} \sum_{j=1}^n Y_{3j}, \\ S_{11} &= \frac{1}{n} \sum_{j=1}^n (Y_{1j} - \bar{Y}_1)^2, \quad S_{22} = \frac{1}{n} \sum_{j=1}^n (Y_{2j} - \bar{Y}_2)^2, \\ S_{33} &= \frac{1}{n} \sum_{j=1}^n (Y_{3j} - \bar{Y}_3)^2, \\ S_{12} &= \frac{1}{n} \sum_{j=1}^n (Y_{1j} - \bar{Y}_1) (Y_{2j} - \bar{Y}_2), \\ S_{13} &= \frac{1}{n} \sum_{j=1}^n (Y_{1j} - \bar{Y}_1) (Y_{3j} - \bar{Y}_3), \\ S_{23} &= \frac{1}{n} \sum_{j=1}^n (Y_{2j} - \bar{Y}_2) (Y_{3j} - \bar{Y}_3).\end{aligned}$$

É importante notar que os estimadores acima são EMV dos parâmetros, desde que seus valores pertencem ao espaço paramétrico, pois em algumas situações pode ocorrer que as estimativas das variâncias sejam negativas. Este problema pode ser contornado utilizando a estimação restrita de máxima verossimilhança como é apontado em Carter (1981).

Nos trabalhos de Nice (1996) e Bolfarine e Galea (1995b) mostra-se que os estimadores derivados anteriormente são consistentes e assintoticamente normal.

Pela propriedade de invariância do estimador de máxima verossimilhança, os estimadores

de máxima verossimilhança de ρ_1 , ρ_2 , e ρ_3 são

$$\tilde{\rho}_1 = \frac{\tilde{\sigma}_x^2 \tilde{\beta}_1^2}{\tilde{\sigma}_x^2 \tilde{\beta}_1^2 + \tilde{\sigma}_1^2}, \quad \tilde{\rho}_2 = \frac{\tilde{\sigma}_x^2 \tilde{\beta}_2^2}{\tilde{\sigma}_x^2 \tilde{\beta}_2^2 + \tilde{\sigma}_2^2} \quad \text{e} \quad \tilde{\rho}_3 = \frac{\tilde{\sigma}_x^2 \tilde{\beta}_3^2}{\tilde{\sigma}_x^2 \tilde{\beta}_3^2 + \tilde{\sigma}_3^2},$$

respectivamente. Similarmente, podem ser obtidos os estimadores de máxima verossimilhança para as precisões dos instrumentos.

A seguinte parametrização do modelo de calibração comparativa será de grande utilidade no decorrer deste trabalho e será utilizada na aplicação do método de influência local.

3.4 O Modelo Estrutural de Calibração de Análise de Fatores

Theobald e Mallison (1978) usam uma reparametrização do modelo (3.6), em termos de análise de fatores. Especificamente, seja

$$x_j = \mu_x + \sigma_x F_j \quad (3.8)$$

onde $F_j \sim N(0, 1)$, $j = 1, \dots, n$. Substituindo (3.8) em (3.6) temos que

$$\begin{aligned} \mathbf{Y}_j &= \mathbf{a} + \mathbf{b}(\mu_x + \sigma_x F_j) + \mathbf{e}_j, \\ &= \underbrace{\mathbf{a} + \mathbf{b}\mu_x}_{\boldsymbol{\mu}} + \underbrace{\sigma_x \mathbf{b}}_{\boldsymbol{\lambda}} F_j + \mathbf{e}_j, \\ &= \boldsymbol{\mu} + \boldsymbol{\lambda} F_j + \mathbf{e}_j = \boldsymbol{\mu} + \mathbf{L} \mathbf{v}_j, \end{aligned} \quad (3.9)$$

onde $\boldsymbol{\mu} = \mathbf{a} + \mathbf{b}\mu_x$, $\boldsymbol{\lambda} = \sigma_x \mathbf{b}$, $\mathbf{L} = [\boldsymbol{\lambda}, \mathbf{I}_3]$, $\mathbf{v}_j = (F_j, \mathbf{e}_j^T)^T$, com $\mathbf{a}$ e $\mathbf{b}$ definidos em (3.6), $j = 1, \dots, n$.

O modelo normal é obtido ao considerar

$$\mathbf{v}_j \sim N_4(\mathbf{0}, \mathbf{\Omega}),$$

onde $\mathbf{\Omega} = \text{Diag}(1, D(\phi))$, com $D(\phi)$ como em (3.4). Assim, pelas propriedades da distribuição normal temos de (3.9) que $\mathbf{Y}_j \sim N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, onde

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix} \text{ e } \boldsymbol{\Sigma} = \boldsymbol{\lambda}\boldsymbol{\lambda}^T + D(\phi) = \begin{pmatrix} \lambda_1^2 + \sigma_1^2 & \lambda_1\lambda_2 & \lambda_1\lambda_3 \\ \lambda_1\lambda_2 & \lambda_2^2 + \sigma_2^2 & \lambda_2\lambda_3 \\ \lambda_1\lambda_3 & \lambda_2\lambda_3 & \lambda_3^2 + \sigma_3^2 \end{pmatrix}.$$

Assim, sob esta nova parametrização de análise de fatores, o parâmetro do modelo é dado por

$$\boldsymbol{\theta} = (\boldsymbol{\mu}^T, \boldsymbol{\lambda}^T, \boldsymbol{\phi}^T)^T, \quad (3.10)$$

onde $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3)^T$, $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3)^T$ e $\boldsymbol{\phi} = (\phi_1, \phi_2, \phi_3)^T$.

3.4.1 Estimadores de Máxima Verossimilhança

Sob esta parametrização, Nice (1996) determinaram sob (a hipótese de) normalidade os estimadores de máxima verossimilhança dos parâmetros do modelo definido em (3.9), que são dados por $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\mu}}^T, \hat{\boldsymbol{\phi}}^T, \hat{\boldsymbol{\lambda}}^T)^T$, onde

$$\begin{aligned} \hat{\mu}_i &= \bar{Y}_i, \quad i = 1, 2, 3 \\ \hat{\phi}_i &= \hat{\sigma}_i^2 = S_{ii} - \hat{\lambda}_i^2, \quad i = 1, 2, 3 \\ \hat{\lambda}_1 &= \sqrt{\frac{S_{12}S_{13}}{S_{23}}}, \quad \hat{\lambda}_2 = \hat{\lambda}_1 \frac{S_{23}}{S_{13}}, \quad \hat{\lambda}_3 = \hat{\lambda}_1 \frac{S_{23}}{S_{12}}, \end{aligned}$$

com

$$S_{12} = \frac{1}{n} \sum_{j=1}^n (Y_{1j} - \bar{Y}_1)(Y_{2j} - \bar{Y}_2),$$

$$S_{13} = \frac{1}{n} \sum_{j=1}^n (Y_{1j} - \bar{Y}_1) (Y_{3j} - \bar{Y}_3),$$

$$S_{23} = \frac{1}{n} \sum_{j=1}^n (Y_{2j} - \bar{Y}_2) (Y_{3j} - \bar{Y}_3),$$

e

$$\bar{Y}_1 = \frac{1}{n} \sum_{j=1}^n Y_{1j},$$

$$\bar{Y}_2 = \frac{1}{n} \sum_{j=1}^n Y_{2j},$$

$$\bar{Y}_3 = \frac{1}{n} \sum_{j=1}^n Y_{3j}.$$

Observação 3.1: Sob o modelo de calibração comparativa e a parametrização de análise de fatores, as definições de confiabilidade e precisão do i -ésimo instrumento são dadas por

$$\rho_i = \frac{\lambda_i^2}{\lambda_i^2 + \sigma_i^2},$$

$$\pi_i = \frac{\lambda_i^2}{\sigma_x^2 \sigma_i^2}, \quad \tau_i = \frac{\lambda_i^2}{\sigma_i^2} \quad \text{e} \quad \delta_i = \frac{\sigma_i^2}{\lambda_i^2}, \quad i = 1, 2, 3.$$

Além disso, as hipóteses de interesse H_{01} , H_{02} e H_{03} definidas na seção 3.3.1 são equivalentes a

$$H_{01} : \mu_1 = \mu_2 = \mu_3, \quad \lambda_1 = \lambda_2 = \lambda_3, \quad \rho_1 = \rho_2 = \rho_3,$$

$$H_{02} : \mu_1 = \mu_2 = \mu_3, \quad \lambda_1 = \lambda_2 = \lambda_3,$$

e

$$H_{03} : \rho_1 = \rho_2 = \rho_3,$$

respectivamente.

Sob a parametrização de análise de fatores, os estimadores de máxima verossimilhança

de ρ_1, ρ_2 e ρ_3 são dados por

$$\hat{\rho}_i = \frac{\hat{\lambda}_i^2}{\hat{\lambda}_i^2 + \hat{\sigma}_i^2}, \quad i = 1, 2, 3.$$

respectivamente.

Similarmente, pelo princípio de invariância, podemos obter os estimadores para as precisões.

3.4.2 Matriz de Informação Observada

Nesta seção vamos apresentar a matriz de informação observada baseada no modelo de calibração comparativa proposto por Theobald e Mallison (1978), que é uma reparametrização do modelo proposto por Barnett (1969). Esta matriz de informação observada, será utilizada posteriormente para aplicar o método de influência local no modelo de calibração comparativa. Além disso, será utilizada para obter a matriz de informação esperada.

Sob normalidade, a função de densidade de probabilidade do modelo de calibração comparativa definido em (3.9) é dada por

$$f(\mathbf{y}_j; \boldsymbol{\theta}) = (2\pi)^{-3/2} |\boldsymbol{\Sigma}|^{-1/2} \exp \left\{ -\frac{1}{2} (\mathbf{y}_j - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{y}_j - \boldsymbol{\mu}) \right\}, \quad (3.11)$$

onde

$$\boldsymbol{\Sigma} = \boldsymbol{\lambda} \boldsymbol{\lambda}^T + D(\boldsymbol{\phi}), \quad \boldsymbol{\Sigma}^{-1} = D^{-1}(\boldsymbol{\phi}) - c^{-1} \mathbf{M} \quad \text{e} \quad |\boldsymbol{\Sigma}| = c |D(\boldsymbol{\phi})|,$$

com

$$c = 1 + \boldsymbol{\lambda}^T D^{-1}(\boldsymbol{\phi}) \boldsymbol{\lambda} \quad \text{e} \quad \mathbf{M} = D^{-1}(\boldsymbol{\phi}) \boldsymbol{\lambda} \boldsymbol{\lambda}^T D^{-1}(\boldsymbol{\phi}). \quad (3.12)$$

Logo, para uma amostra aleatória $\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_n$ a função de verossimilhança é dada por

$$L(\boldsymbol{\theta}; \mathbf{y}_1, \dots, \mathbf{y}_n) = \prod_{j=1}^n (2\pi)^{-3/2} |\boldsymbol{\Sigma}|^{-1/2} \exp \left\{ -\frac{1}{2} (\mathbf{y}_j - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{y}_j - \boldsymbol{\mu}) \right\}$$

e a função de log-verossimilhança é

$$l(\boldsymbol{\theta}) = \sum_{j=1}^n l_j(\boldsymbol{\theta}) = \sum_{j=1}^n \log f(\mathbf{y}_j; \boldsymbol{\theta}), \quad (3.13)$$

onde

$$l_j(\boldsymbol{\theta}) = -\frac{3}{2} \log(2\pi) - \frac{1}{2} \log |\boldsymbol{\Sigma}| - \frac{1}{2} \|T_j\|^2,$$

com

$$\|T_j\|^2 = (\mathbf{y}_j - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{y}_j - \boldsymbol{\mu}).$$

A fim de usarmos os resultados do Capítulo 2, relacionado com o método de Influência Local proposto por Cook (1986), precisaremos das derivadas parciais de primeira e segunda ordem da função $l(\boldsymbol{\theta})$, em relação a $\boldsymbol{\theta}$.

a) Derivada parcial de primeira ordem de $l(\boldsymbol{\theta})$

Como

$$l(\boldsymbol{\theta}) = \sum_{j=1}^n l_j(\boldsymbol{\theta}),$$

então de (3.13)

$$\frac{\partial l}{\partial \boldsymbol{\theta}} = \sum_{j=1}^n \frac{\partial l_j(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}},$$

onde

$$\begin{aligned} \frac{\partial l_j(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} &= \frac{\partial l_j}{\partial \boldsymbol{\theta}} = \begin{pmatrix} \frac{\partial l_j}{\partial \boldsymbol{\mu}} \\ \frac{\partial l_j}{\partial \boldsymbol{\lambda}} \\ \frac{\partial l_j}{\partial \boldsymbol{\phi}} \end{pmatrix} = -\frac{1}{2} \frac{\partial}{\partial \boldsymbol{\theta}} \log |\boldsymbol{\Sigma}| - \frac{1}{2} \frac{\partial}{\partial \boldsymbol{\theta}} \|T_j\|^2 \\ &= -\frac{1}{2} \left[\frac{\partial}{\partial \boldsymbol{\theta}} \log |\boldsymbol{\Sigma}| + \frac{\partial}{\partial \boldsymbol{\theta}} \|T_j\|^2 \right], \end{aligned} \quad (3.14)$$

com

$$\frac{\partial l_j}{\partial \boldsymbol{\mu}} = -\frac{1}{2} \left[\frac{\partial}{\partial \boldsymbol{\mu}} \log |\boldsymbol{\Sigma}| + \frac{\partial}{\partial \boldsymbol{\mu}} \|T_j\|^2 \right]$$

$$\frac{\partial l_j}{\partial \boldsymbol{\lambda}} = -\frac{1}{2} \left[\frac{\partial}{\partial \boldsymbol{\lambda}} \log |\boldsymbol{\Sigma}| + \frac{\partial}{\partial \boldsymbol{\lambda}} \|T_j\|^2 \right]$$

$$\frac{\partial l_j}{\partial \boldsymbol{\phi}} = -\frac{1}{2} \left[\frac{\partial}{\partial \boldsymbol{\phi}} \log |\boldsymbol{\Sigma}| + \frac{\partial}{\partial \boldsymbol{\phi}} \|T_j\|^2 \right].$$

Assim temos que as derivadas de primeira ordem de $l(\boldsymbol{\theta})$ dependem das derivadas de $\log |\boldsymbol{\Sigma}|$ e $\|T_j\|^2$ com respeito a $\boldsymbol{\mu}$, $\boldsymbol{\lambda}$, e $\boldsymbol{\phi}$, respectivamente.

Consideremos as seguintes notações: seja $\mathbf{a} = (a_1, a_2, a_3)^T$ um vetor em $\mathbf{R}^3$, então definimos $D(\mathbf{a}) = \text{Diag}(a_1, a_2, a_3)$ e $D^{-k}(\mathbf{a}) = \text{Diag}(a_1^{-k}, a_2^{-k}, a_3^{-k})$, $k = 1, 2, 3$.

A notação anterior também é válida se $\mathbf{a}$ for um vetor aleatório. Considerando as notações acima e após de manipulações algébricas (veja apêndice B) temos que

$$\frac{\partial \|T_j\|^2}{\partial \boldsymbol{\mu}} = -2\boldsymbol{\Sigma}^{-1}\mathbf{W}_j = -2(D^{-1}(\boldsymbol{\phi}) - c^{-1}\mathbf{M})\mathbf{W}_j,$$

$$\frac{\partial \|T_j\|^2}{\partial \boldsymbol{\lambda}} = 2c^{-2}\mathbf{W}_j^T\mathbf{M}\mathbf{W}_jD^{-1}(\boldsymbol{\phi})\boldsymbol{\lambda} - 2c^{-1}D^{-1}(\boldsymbol{\phi})\mathbf{W}_j\mathbf{W}_j^TD^{-1}(\boldsymbol{\phi})\boldsymbol{\lambda},$$

$$\begin{aligned} \frac{\partial \|T_j\|^2}{\partial \boldsymbol{\phi}} &= -D^{-2}(\boldsymbol{\phi})D(\mathbf{W}_j)\mathbf{W}_j - c^{-2}\mathbf{W}_j^T\mathbf{M}\mathbf{W}_jD(\boldsymbol{\lambda})D^{-2}(\boldsymbol{\phi})\boldsymbol{\lambda} \\ &\quad + 2c^{-1}D(\boldsymbol{\lambda})D^{-2}(\boldsymbol{\phi})\mathbf{W}_j\mathbf{W}_j^TD^{-1}(\boldsymbol{\phi})\boldsymbol{\lambda}, \end{aligned}$$

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

$$\frac{\partial \log |\boldsymbol{\Sigma}|}{\partial \boldsymbol{\lambda}} = 2c^{-1}D^{-1}(\boldsymbol{\phi})\boldsymbol{\lambda},$$

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

com $\mathbf{W}_j = \mathbf{Y}_j - \boldsymbol{\mu}$, c e $\mathbf{M}$ como em (3.10). Logo, o vetor de derivadas parciais de

primeira ordem, $\frac{\partial l_j}{\partial \theta}$ fica como

$$\frac{\partial l_j}{\partial \theta} = \begin{pmatrix} \Sigma^{-1} \mathbf{W}_j \\ c^{-1} (D^{-1}(\phi) \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi) \boldsymbol{\lambda} - c^{-1} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j D^{-1}(\phi) \boldsymbol{\lambda} - D^{-1}(\phi) \boldsymbol{\lambda}) \\ -\frac{1}{2} \{-c^{-1} D^{-2}(\phi) D(\boldsymbol{\lambda}) \boldsymbol{\lambda} + D^{-1}(\phi) \mathbf{1} - D^{-2}(\phi) D(\mathbf{W}_j) \mathbf{W}_j \\ -c^{-2} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j D(\boldsymbol{\lambda}) D^{-2}(\phi) \boldsymbol{\lambda} + 2c^{-1} D(\boldsymbol{\lambda}) D^{-2}(\phi) \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi) \boldsymbol{\lambda}\} \end{pmatrix}.$$

b) Derivadas de Segunda ordem de $l(\theta)$

Seja $l(\theta)$ a função de log-verossimilhança definida em (3.11). Então a matriz de derivadas parciais de segunda ordem com respeito a θ é dada por

$$\ddot{L} = \frac{\partial^2 l(\theta)}{\partial \theta \partial \theta^T} = \sum_{j=1}^n \frac{\partial^2 l_j}{\partial \theta \partial \theta^T}, \quad (3.15)$$

onde

$$\frac{\partial^2 l_j}{\partial \theta \partial \theta^T} = -\frac{1}{2} \left[\frac{\partial^2 \log |\Sigma|}{\partial \theta \partial \theta^T} + \frac{\partial^2 \|T_j\|^2}{\partial \theta \partial \theta^T} \right],$$

$$\frac{\partial^2 \log |\Sigma|}{\partial \theta \partial \theta^T} = \begin{pmatrix} \frac{\partial^2 \log |\Sigma|}{\partial \mu \partial \mu^T} & \frac{\partial^2 \log |\Sigma|}{\partial \mu \partial \lambda^T} & \frac{\partial^2 \log |\Sigma|}{\partial \mu \partial \phi^T} \\ \frac{\partial^2 \log |\Sigma|}{\partial \lambda \partial \lambda^T} & \frac{\partial^2 \log |\Sigma|}{\partial \lambda \partial \phi^T} \\ \text{simetria} & \frac{\partial^2 \log |\Sigma|}{\partial \phi \partial \phi^T} \end{pmatrix}$$

e

$$\frac{\partial^2 \|T_j\|^2}{\partial \theta \partial \theta^T} = \begin{pmatrix} \frac{\partial^2 \|T_j\|^2}{\partial \mu \partial \mu^T} & \frac{\partial^2 \|T_j\|^2}{\partial \mu \partial \lambda^T} & \frac{\partial^2 \|T_j\|^2}{\partial \mu \partial \phi^T} \\ \frac{\partial^2 \|T_j\|^2}{\partial \lambda \partial \lambda^T} & \frac{\partial^2 \|T_j\|^2}{\partial \lambda \partial \phi^T} \\ \text{simetria} & \frac{\partial^2 \|T_j\|^2}{\partial \phi \partial \phi^T} \end{pmatrix},$$

com

$$\frac{\partial^2 \log|\Sigma|}{\partial \mu \partial \mu^T} = \frac{\partial^2 \log|\Sigma|}{\partial \mu \partial \lambda^T} = \frac{\partial^2 \log|\Sigma|}{\partial \mu \partial \phi^T} = \mathbf{0}_{3 \times 3},$$

$$\frac{\partial^2 \log|\Sigma|}{\partial \lambda \partial \lambda^T} = 2c^{-1} (\Sigma^{-1} - c^{-1} \mathbf{M}) = 2c^{-1} (D^{-1}(\phi) - 2c^{-1} \mathbf{M}),$$

$$\frac{\partial^2 \log|\Sigma|}{\partial \lambda \partial \phi^T} = -2c^{-1} \Sigma^{-1} D^{-1}(\phi) D(\lambda),$$

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

$$\frac{\partial \|T_j\|^2}{\partial \mu \partial \mu^T} = 2 (D^{-1}(\phi) - c^{-1} \mathbf{M}),$$

$$\frac{\partial^2 \|T_j\|^2}{\partial \mu \partial \lambda^T} = 2c^{-1} \{ D^{-1}(\phi) \lambda \mathbf{W}_j^T D^{-1}(\phi) + \lambda^T D^{-1}(\phi) \mathbf{W}_j [\Sigma^{-1} - c^{-1} \mathbf{M}] \},$$

$$\begin{aligned} \frac{\partial \|T_j\|^2}{\partial \mu \partial \phi^T} &= 2D(\mathbf{W}_j) D^{-2}(\phi) + 2c^{-2} \mathbf{M} \mathbf{W}_j \lambda^T D(\lambda) D^{-2}(\phi) - 2c^{-1} D^{-1}(\phi) \lambda \mathbf{W}_j^T D^{-2}(\phi) D(\lambda) \\ &\quad - 2c^{-1} D^{-2}(\phi) D(\lambda) \lambda^T D^{-1}(\phi) \mathbf{W}_j, \end{aligned}$$

$$\begin{aligned} \frac{\partial \|T_j\|^2}{\partial \lambda \partial \lambda^T} &= -2c^{-1} D^{-1}(\phi) \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi) + 4c^{-2} D^{-1}(\phi) \mathbf{W}_j \mathbf{W}_j^T \mathbf{M} + 2c^{-2} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j \Sigma^{-1} \\ &\quad - 6c^{-3} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j \mathbf{M} + 4c^{-2} \mathbf{M} \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi), \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 \|T_j\|^2}{\partial \lambda \partial \phi^T} &= 4c^{-3} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j \mathbf{M} D^{-1}(\phi) D(\lambda) - 4c^{-2} \mathbf{M} \mathbf{W}_j \mathbf{W}_j^T D(\lambda) D^{-2}(\phi) \\ &\quad - 2c^{-2} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j D(\lambda) D^{-2}(\phi) - 2c^{-2} D^{-1}(\phi) \mathbf{W}_j \mathbf{W}_j^T \mathbf{M} D(\lambda) D^{-1}(\phi) \\ &\quad + 2c^{-1} D^{-1}(\phi) \mathbf{W}_j \mathbf{W}_j^T D(\lambda) D^{-2}(\phi) + 2c^{-1} \mathbf{W}_j^T D^{-1}(\phi) \lambda D(\mathbf{W}_j) D^{-2}(\phi), \end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 \|T_j\|^2}{\partial \phi \partial \phi^T} = & 2D(\mathbf{W}_j) D^{-3}(\phi) D(\mathbf{W}_j) - 2c^{-3} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j D(\lambda) D^{-1}(\phi) \mathbf{M} D^{-1}(\phi) D(\lambda) \\
& + 2c^{-2} D(\lambda) D^{-2}(\phi) \mathbf{W}_j \mathbf{W}_j^T \mathbf{M} D^{-1}(\phi) D(\lambda) \\
& + 2c^{-2} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j D(\lambda) D^{-3}(\phi) D(\lambda) \\
& + 2c^{-2} D(\lambda) D^{-1}(\phi) \mathbf{M} \mathbf{W}_j \mathbf{W}_j^T D(\lambda) D^{-2}(\phi) \\
& - 2c^{-1} D(\lambda) D^{-2}(\phi) \mathbf{W}_j \mathbf{W}_j^T D(\lambda) D^{-2}(\phi) \\
& - 4c^{-1} D^{-3}(\phi) D(\lambda) \mathbf{W}_j^T D^{-1}(\phi) \lambda D(\mathbf{W}_j).
\end{aligned}$$

Logo a matriz $\frac{\partial^2 l_j}{\partial \theta \partial \theta^T}$ é dada por

$$\begin{aligned}
\frac{\partial^2 l_j(\theta)}{\partial \theta \partial \theta^T} &= \frac{\partial}{\partial \theta} \left(\frac{\partial l_j(\theta)}{\partial \theta^T} \right) = \frac{\partial}{\partial \theta} \left[\frac{\partial}{\partial \mu^T} l_j, \frac{\partial}{\partial \lambda^T} l_j, \frac{\partial}{\partial \phi^T} l_j \right] \\
&= \left[\frac{\partial}{\partial \theta} \left(\frac{\partial}{\partial \mu^T} l_j \right), \frac{\partial}{\partial \theta} \left(\frac{\partial}{\partial \lambda^T} l_j \right), \frac{\partial}{\partial \theta} \left(\frac{\partial}{\partial \phi^T} l_j \right) \right] \\
&= \begin{pmatrix} \frac{\partial}{\partial \mu} \left(\frac{\partial}{\partial \mu^T} l_j \right) & \frac{\partial}{\partial \mu} \left(\frac{\partial}{\partial \lambda^T} l_j \right) & \frac{\partial}{\partial \mu} \left(\frac{\partial}{\partial \phi^T} l_j \right) \\ \frac{\partial}{\partial \lambda} \left(\frac{\partial}{\partial \mu^T} l_j \right) & \frac{\partial}{\partial \lambda} \left(\frac{\partial}{\partial \lambda^T} l_j \right) & \frac{\partial}{\partial \lambda} \left(\frac{\partial}{\partial \phi^T} l_j \right) \\ \frac{\partial}{\partial \phi} \left(\frac{\partial}{\partial \mu^T} l_j \right) & \frac{\partial}{\partial \phi} \left(\frac{\partial}{\partial \lambda^T} l_j \right) & \frac{\partial}{\partial \phi} \left(\frac{\partial}{\partial \phi^T} l_j \right) \end{pmatrix} \quad (3.16) \\
&= \begin{pmatrix} \frac{\partial^2 l_j}{\partial \mu \partial \mu^T} & \frac{\partial^2 l_j}{\partial \mu \partial \lambda^T} & \frac{\partial^2 l_j}{\partial \mu \partial \phi^T} \\ \frac{\partial^2 l_j}{\partial \lambda \partial \mu^T} & \frac{\partial^2 l_j}{\partial \lambda \partial \lambda^T} & \frac{\partial^2 l_j}{\partial \lambda \partial \phi^T} \\ \frac{\partial^2 l_j}{\partial \phi \partial \mu^T} & \frac{\partial^2 l_j}{\partial \phi \partial \lambda^T} & \frac{\partial^2 l_j}{\partial \phi \partial \phi^T} \end{pmatrix},
\end{aligned}$$

onde

$$\frac{\partial^2 l_j}{\partial \mu \partial \mu^T} = -D^{-1}(\phi) + c^{-1} \mathbf{M} = (\lambda \lambda^T + D(\phi))^{-1} = \Sigma^{-1},$$

$$\frac{\partial^2 l_j}{\partial \mu \partial \lambda^T} = -c^{-1} \{ D^{-1}(\phi) \lambda \mathbf{W}_j^T D^{-1}(\phi) + \lambda^T D^{-1}(\phi) \mathbf{W}_j [\Sigma^{-1} - c^{-1} \mathbf{M}] \},$$

$$\begin{aligned}
\frac{\partial^2 l_j}{\partial \mu \partial \phi^T} = & -D(\mathbf{W}_j) D^{-2}(\phi) - c^{-2} \mathbf{M} \mathbf{W}_j \lambda^T D(\lambda) D^{-2}(\phi) + c^{-1} D^{-1}(\phi) \lambda \mathbf{W}_j^T D^{-2}(\phi) D(\lambda) \\
& + c^{-1} D^{-2}(\phi) D(\lambda) \lambda^T D^{-1}(\phi) \mathbf{W}_j
\end{aligned}$$

$$\frac{\partial^2 l_j}{\partial \lambda \partial \lambda^T} = - \left\{ c^{-1} (\Sigma^{-1} - c^{-1} \mathbf{M}) + c^{-1} D^{-1}(\phi) \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi) - 2c^{-2} D^{-1}(\phi) \mathbf{W}_j \mathbf{W}_j^T \mathbf{M} \right.$$

$$\left. - c^{-2} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j (D^{-1}(\phi) - c^{-1} \mathbf{M} - 3c^{-1} \mathbf{M}) - 2c^{-2} \mathbf{M} \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi) \right\},$$

$$\begin{aligned} \frac{\partial^2 l_j}{\partial \lambda \partial \phi^T} &= c^{-1} D^{-1}(\phi) D(\lambda) \Sigma^{-1} - 2c^{-3} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j \mathbf{M} D^{-1}(\phi) D(\lambda) + 2c^{-2} \mathbf{M} \mathbf{W}_j \mathbf{W}_j^T D(\lambda) D^{-2}(\phi) \\ &+ c^{-2} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j D(\lambda) D^{-2}(\phi) + c^{-2} D^{-1}(\phi) \mathbf{W}_j \mathbf{W}_j^T \mathbf{M} D(\lambda) D^{-1}(\phi) \\ &- c^{-1} D^{-1}(\phi) \mathbf{W}_j \mathbf{W}_j^T D(\lambda) D^{-2}(\phi) - c^{-1} \mathbf{W}_j^T D^{-1}(\phi) \lambda D(\mathbf{W}_j) D^{-2}(\phi), \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 l_j}{\partial \phi \partial \phi^T} &= -\frac{1}{2} [-D^{-2}(\phi) + 2c^{-1} D^{-3}(\phi) D^2(\lambda) - c^{-2} D^{-1}(\phi) D(\lambda) \mathbf{M} D(\lambda) D^{-1}(\phi) \\ &+ 2D(\mathbf{W}_j) D^{-3}(\phi) D(\mathbf{W}_j) \\ &- 2c^{-3} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j D(\lambda) D^{-1}(\phi) \mathbf{M} D^{-1}(\phi) D(\lambda) \\ &+ 2c^{-2} D(\lambda) D^{-2}(\phi) \mathbf{W}_j \mathbf{W}_j^T \mathbf{M} D^{-1}(\phi) D(\lambda) \\ &+ 2c^{-2} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j D(\lambda) D^{-3}(\phi) D(\lambda) \\ &+ 2c^{-2} D(\lambda) D^{-1}(\phi) \mathbf{M} \mathbf{W}_j \mathbf{W}_j^T D(\lambda) D^{-2}(\phi) \\ &- 2c^{-1} D(\lambda) D^{-2}(\phi) \mathbf{W}_j \mathbf{W}_j^T D(\lambda) D^{-2}(\phi) \\ &- 4c^{-1} D^{-3}(\phi) D(\lambda) \mathbf{W}_j^T D^{-1}(\phi) \lambda D(\mathbf{W}_j)]. \end{aligned}$$

Portanto, dos resultados acima, temos que a matriz de informação observada, é obtida ao considerar

$$\begin{aligned} I_F(\theta) &= -\ddot{L} = -\frac{\partial^2 l(\theta)}{\partial \theta \partial \theta^T} \\ &= -\sum_{j=1}^n \frac{\partial^2 l_j(\theta)}{\partial \theta \partial \theta^T}, \end{aligned}$$

onde $\frac{\partial^2 l_j(\theta)}{\partial \theta \partial \theta^T}$ é como em (3.16), $j = 1, \dots, n$.

3.4.3 Matriz de Informação Esperada

Nesta seção vamos obter a matriz de informação esperada sob a parametrização de análise de fatores do modelo. A obtenção da matriz de informação esperada será importante para tratar hipóteses de interesse apresentadas em (3.3.1) e para estudar comportamento assintótico do estimador de máxima verossimilhança de θ ou de uma função diferenciável de θ .

A matriz de informação esperada que denotamos por $J_F(\theta)$, é definida como sendo

$$J_F(\theta) = E \left(-\frac{\partial^2 l}{\partial \theta \partial \theta^T} \right) = - \sum_{j=1}^n E \left(\frac{\partial^2 l_j}{\partial \theta \partial \theta^T} \right).$$

Para obter o valor esperado da matriz de informação observada, vamos utilizar os seguintes lemas:

Lema 3.1: *Sejam $\mathbf{W}$ e $\mathbf{T}$ matrizes aleatórias de ordem $k_1 \times k_2$, $\mathbf{A}_1$ uma matriz constante $m_1 \times k_1$ e $\mathbf{A}_2$ uma matriz constante de ordem $k_2 \times m_2$. Então*

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

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

Lema 3.2: *Se $\mathbf{W} \sim N_3(\mathbf{0}, \Sigma)$, onde $\Sigma = \lambda \lambda^T + D(\phi)$ e $c = 1 + \lambda^T D^{-1}(\phi) \lambda$, então:*

- 1) $E(\mathbf{W}^T \mathbf{M} \mathbf{W}) = \text{tr}(\mathbf{M} \Sigma) = c(c-1)$,
- 2) $E(D(\mathbf{W}) D(\mathbf{W})) = D(\lambda) D(\lambda) + D(\phi)$,
- 3) $D(\Sigma D^{-1}(\phi) \lambda) = c D(\lambda)$.

A prova do Lema 3.2, segue do Lema 3.1 e do Teorema 10.4.4 dado em Graybill (1983).

Agora de (3.15) temos que,

$$\frac{\partial^2 l_j(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} = \begin{pmatrix} \frac{\partial^2}{\partial \boldsymbol{\mu}} \frac{l_j}{\partial \boldsymbol{\mu}^T} & \frac{\partial^2}{\partial \boldsymbol{\mu}} \frac{l_j}{\partial \boldsymbol{\lambda}^T} & \frac{\partial^2}{\partial \boldsymbol{\mu}} \frac{l_j}{\partial \boldsymbol{\phi}^T} \\ \frac{\partial^2}{\partial \boldsymbol{\lambda}} \frac{l_j}{\partial \boldsymbol{\mu}^T} & \frac{\partial^2}{\partial \boldsymbol{\lambda}} \frac{l_j}{\partial \boldsymbol{\lambda}^T} & \frac{\partial^2}{\partial \boldsymbol{\lambda}} \frac{l_j}{\partial \boldsymbol{\phi}^T} \\ \frac{\partial^2}{\partial \boldsymbol{\phi}} \frac{l_j}{\partial \boldsymbol{\mu}^T} & \frac{\partial^2}{\partial \boldsymbol{\phi}} \frac{l_j}{\partial \boldsymbol{\lambda}^T} & \frac{\partial^2}{\partial \boldsymbol{\phi}} \frac{l_j}{\partial \boldsymbol{\phi}^T} \end{pmatrix} = -\frac{1}{2} \left[\frac{\partial^2 \log |\boldsymbol{\Sigma}|}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} + \frac{\partial^2 \|T_j\|^2}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \right].$$

Agora vamos obter o valor esperado das submatrizes da matriz acima, utilizando os resultados dos Lemas 3.1 e 3.2, com $\mathbf{W} = \mathbf{W}_j \sim N_3(\mathbf{0}, \boldsymbol{\Sigma})$,

$$\begin{aligned} E \left(\frac{\partial^2 l_j}{\partial \boldsymbol{\mu} \partial \boldsymbol{\mu}^T} \right) &= -\frac{1}{2} E \left[\frac{\partial^2 \log |\boldsymbol{\Sigma}|}{\partial \boldsymbol{\mu} \partial \boldsymbol{\mu}^T} + \frac{\partial^2 \|T_j\|^2}{\partial \boldsymbol{\mu} \partial \boldsymbol{\mu}^T} \right] \\ &= -\boldsymbol{\Sigma}^{-1}. \end{aligned}$$

Como $E(\mathbf{W}_j) = \mathbf{0}$, então dos Lemas 3.1 e 3.2,

$$E \left(\frac{\partial^2 l_j}{\partial \boldsymbol{\mu} \partial \boldsymbol{\lambda}^T} \right) = E \left(\frac{\partial^2 l_j}{\partial \boldsymbol{\mu} \partial \boldsymbol{\phi}^T} \right) = 0$$

$$\begin{aligned} E \left(\frac{\partial^2 l_j}{\partial \boldsymbol{\lambda} \partial \boldsymbol{\lambda}^T} \right) &= -\frac{1}{2} E \left[\frac{\partial^2 \log |\boldsymbol{\Sigma}|}{\partial \boldsymbol{\lambda} \partial \boldsymbol{\lambda}^T} + \frac{\partial^2 \|T_j\|^2}{\partial \boldsymbol{\lambda} \partial \boldsymbol{\lambda}^T} \right] \\ &= -\frac{1}{2} \left[\frac{\partial^2 \log |\boldsymbol{\Sigma}|}{\partial \boldsymbol{\lambda} \partial \boldsymbol{\lambda}^T} - 2c^{-1} D^{-1}(\boldsymbol{\phi}) E(\mathbf{W}_j \mathbf{W}_j^T) D^{-1}(\boldsymbol{\phi}) \right. \\ &\quad \left. + 4c^{-2} D^{-1}(\boldsymbol{\phi}) E(\mathbf{W}_j \mathbf{W}_j^T) \mathbf{M} + 2c^{-2} E(\mathbf{W}_j^T \mathbf{M} \mathbf{W}_j) \boldsymbol{\Sigma}^{-1} \right. \\ &\quad \left. - 6c^{-3} E(\mathbf{W}_j^T \mathbf{M} \mathbf{W}_j) \mathbf{M} + 4c^{-2} \mathbf{M} E(\mathbf{W}_j \mathbf{W}_j^T) D^{-1}(\boldsymbol{\phi}) \right] \\ &= -\frac{1}{2} \left[2c^{-1} D^{-1}(\boldsymbol{\phi}) - 4c^{-2} \mathbf{M} - 2c^{-1} D^{-1}(\boldsymbol{\phi}) E(\mathbf{W}_j \mathbf{W}_j^T) D^{-1}(\boldsymbol{\phi}) \right. \\ &\quad \left. + 4c^{-2} D^{-1}(\boldsymbol{\phi}) \boldsymbol{\Sigma} \mathbf{M} + 2c^{-2} c(c-1) \boldsymbol{\Sigma}^{-1} \right. \\ &\quad \left. - 6c^{-3} c(c-1) \mathbf{M} + 4c^{-2} \mathbf{M} \boldsymbol{\Sigma} D^{-1}(\boldsymbol{\phi}) \right]. \end{aligned}$$

Após manipulações algébricas, obtemos que

$$E \left(\frac{\partial^2 l_j}{\partial \boldsymbol{\lambda} \partial \boldsymbol{\lambda}^T} \right) = - (1 - c^{-1}) D^{-1}(\boldsymbol{\phi}) - c^{-1} (2c^{-1} - 1) \mathbf{M},$$

$$\begin{aligned} E \left(\frac{\partial^2 l_j}{\partial \boldsymbol{\lambda} \partial \boldsymbol{\phi}^T} \right) &= -\frac{1}{2} E \left[\frac{\partial^2 \log |\boldsymbol{\Sigma}|}{\partial \boldsymbol{\lambda} \partial \boldsymbol{\phi}^T} + \frac{\partial^2 \|T_j\|^2}{\partial \boldsymbol{\lambda} \partial \boldsymbol{\phi}^T} \right] \\ &= -\frac{1}{2} \left[\frac{\partial^2 \log |\boldsymbol{\Sigma}|}{\partial \boldsymbol{\lambda} \partial \boldsymbol{\phi}^T} + 4c^{-3} E(\mathbf{W}_j^T \mathbf{M} \mathbf{W}_j) \mathbf{M} D^{-1}(\boldsymbol{\phi}) D(\boldsymbol{\lambda}) \right. \\ &\quad - 4c^{-2} \mathbf{M} E(\mathbf{W}_j \mathbf{W}_j^T) D(\boldsymbol{\lambda}) D^{-2}(\boldsymbol{\phi}) - 2c^{-2} E(\mathbf{W}_j^T \mathbf{M} \mathbf{W}_j) D(\boldsymbol{\lambda}) D^{-2}(\boldsymbol{\phi}) \\ &\quad - 2c^{-2} D^{-1}(\boldsymbol{\phi}) E(\mathbf{W}_j \mathbf{W}_j^T) \mathbf{M} D(\boldsymbol{\lambda}) D^{-1}(\boldsymbol{\phi}) \\ &\quad + 2c^{-1} D^{-1}(\boldsymbol{\phi}) E(\mathbf{W}_j \mathbf{W}_j^T) D(\boldsymbol{\lambda}) D^{-2}(\boldsymbol{\phi}) \\ &\quad \left. + 2c^{-1} D^{-2}(\boldsymbol{\phi}) E(D(\mathbf{W}_j) \mathbf{W}_j^T D^{-1}(\boldsymbol{\phi}) \boldsymbol{\lambda}) \right] \\ &= -\frac{1}{2} \left[-2c^{-1} \boldsymbol{\Sigma}^{-1} D^{-1}(\boldsymbol{\phi}) D(\boldsymbol{\lambda}) + 4c^{-3} c(c-1) \mathbf{M} D^{-1}(\boldsymbol{\phi}) D(\boldsymbol{\lambda}) \right. \\ &\quad - 4c^{-2} \mathbf{M} \boldsymbol{\Sigma} D(\boldsymbol{\lambda}) D^{-2}(\boldsymbol{\phi}) - 2c^{-2} c(c-1) D(\boldsymbol{\lambda}) D^{-2}(\boldsymbol{\phi}) \\ &\quad - 2c^{-2} D^{-1}(\boldsymbol{\phi}) E(\mathbf{W}_j \mathbf{W}_j^T) \mathbf{M} D(\boldsymbol{\lambda}) D^{-1}(\boldsymbol{\phi}) \\ &\quad \left. + 2c^{-1} D^{-1}(\boldsymbol{\phi}) \boldsymbol{\Sigma} D(\boldsymbol{\lambda}) D^{-2}(\boldsymbol{\phi}) + 2c^{-1} D^{-2} c D(\boldsymbol{\lambda}) \right]. \end{aligned}$$

Após manipulações algébricas, obtemos que

$$E \left(\frac{\partial^2 l_j}{\partial \boldsymbol{\lambda} \partial \boldsymbol{\phi}^T} \right) = -c^{-1} D(\boldsymbol{\lambda}) D^{-2}(\boldsymbol{\phi}) + c^{-2} \mathbf{M} D^{-1}(\boldsymbol{\phi}) D(\boldsymbol{\lambda}),$$

$$\begin{aligned} E \left(\frac{\partial^2 l_j}{\partial \boldsymbol{\phi} \partial \boldsymbol{\phi}^T} \right) &= -\frac{1}{2} E \left[\frac{\partial^2 \log |\boldsymbol{\Sigma}|}{\partial \boldsymbol{\phi} \partial \boldsymbol{\phi}^T} + \frac{\partial^2 \|T_j\|^2}{\partial \boldsymbol{\phi} \partial \boldsymbol{\phi}^T} \right] \\ &= -\frac{1}{2} \left[\frac{\partial^2 \log |\boldsymbol{\Sigma}|}{\partial \boldsymbol{\phi} \partial \boldsymbol{\phi}^T} + 2D^{-3}(\boldsymbol{\phi}) E(D(\mathbf{W}_j) D(\mathbf{W}_j)) \right. \\ &\quad - 2c^{-3} E(\mathbf{W}_j^T \mathbf{M} \mathbf{W}_j) D(\boldsymbol{\lambda}) D^{-1}(\boldsymbol{\phi}) \mathbf{M} D^{-1}(\boldsymbol{\phi}) D(\boldsymbol{\lambda}) \\ &\quad \left. + 2c^{-2} D(\boldsymbol{\lambda}) D^{-2}(\boldsymbol{\phi}) E(\mathbf{W}_j \mathbf{W}_j^T) \mathbf{M} D^{-1}(\boldsymbol{\phi}) D(\boldsymbol{\lambda}) \right] \end{aligned}$$

$$\begin{aligned}
& +2c^{-2}E\left(\mathbf{W}_j^T\mathbf{M}\mathbf{W}_j\right)D\left(\boldsymbol{\lambda}\right)D^{-3}\left(\boldsymbol{\phi}\right)D\left(\boldsymbol{\lambda}\right) \\
& +2c^{-2}D\left(\boldsymbol{\lambda}\right)D^{-1}\left(\boldsymbol{\phi}\right)\mathbf{M}E\left(\mathbf{W}_j\mathbf{W}_j^T\right)D\left(\boldsymbol{\lambda}\right)D^{-2}\left(\boldsymbol{\phi}\right) \\
& -2c^{-1}D\left(\boldsymbol{\lambda}\right)D^{-2}\left(\boldsymbol{\phi}\right)\mathbf{W}_j\mathbf{W}_j^TD\left(\boldsymbol{\lambda}\right)D^{-2}\left(\boldsymbol{\phi}\right) \\
& -2c^{-1}D^{-3}\left(\boldsymbol{\phi}\right)D\left(\boldsymbol{\lambda}\right)E\left(D\left(\mathbf{W}_j\right)\mathbf{W}_j^TD^{-1}\left(\boldsymbol{\phi}\right)\boldsymbol{\lambda}\right)] \\
= & -\frac{1}{2}\left[-D^{-2}\left(\boldsymbol{\phi}\right)+2c^{-1}D^{-3}\left(\boldsymbol{\phi}\right)D^{-2}\left(\boldsymbol{\lambda}\right)\right. \\
& -c^{-2}D^{-1}\left(\boldsymbol{\phi}\right)D\left(\boldsymbol{\lambda}\right)\mathbf{M}D\left(\boldsymbol{\lambda}\right)D^{-1}\left(\boldsymbol{\phi}\right) \\
& +2D^{-3}\left(\boldsymbol{\phi}\right)\left(D\left(\boldsymbol{\lambda}\right)D\left(\boldsymbol{\lambda}\right)+D\left(\boldsymbol{\phi}\right)\right) \\
& -2c^{-3}c\left(c-1\right)D\left(\boldsymbol{\lambda}\right)D^{-1}\left(\boldsymbol{\phi}\right)\mathbf{M}D^{-1}\left(\boldsymbol{\phi}\right)D\left(\boldsymbol{\lambda}\right) \\
& +2c^{-2}D\left(\boldsymbol{\lambda}\right)D^{-2}\left(\boldsymbol{\phi}\right)\boldsymbol{\Sigma}\mathbf{M}D^{-1}\left(\boldsymbol{\phi}\right)D\left(\boldsymbol{\lambda}\right) \\
& +2c^{-2}c\left(c-1\right)D\left(\boldsymbol{\lambda}\right)D^{-3}\left(\boldsymbol{\phi}\right)D\left(\boldsymbol{\lambda}\right) \\
& +2c^{-2}D\left(\boldsymbol{\lambda}\right)D^{-1}\left(\boldsymbol{\phi}\right)\mathbf{M}\boldsymbol{\Sigma}D\left(\boldsymbol{\lambda}\right)D^{-2}\left(\boldsymbol{\phi}\right) \\
& -2c^{-1}D\left(\boldsymbol{\lambda}\right)D^{-2}\left(\boldsymbol{\phi}\right)\boldsymbol{\Sigma}D\left(\boldsymbol{\lambda}\right)D^{-2}\left(\boldsymbol{\phi}\right) \\
& \left.-4c^{-1}D^{-3}\left(\boldsymbol{\phi}\right)D\left(\boldsymbol{\lambda}\right)D\left(c\boldsymbol{\lambda}\right)\right].
\end{aligned}$$

Após manipulações algébricas, obtemos que

$$E\left(\frac{\partial^2 l_j}{\partial \boldsymbol{\phi} \partial \boldsymbol{\phi}^T}\right) = -\frac{D^{-2}(\boldsymbol{\phi})}{2} + c^{-1}D^2(\boldsymbol{\lambda})D^{-3}(\boldsymbol{\phi}) - \frac{c}{2}D(\boldsymbol{\lambda})D^{-1}(\boldsymbol{\phi})\mathbf{M}D^{-1}(\boldsymbol{\phi})D(\boldsymbol{\lambda}).$$

Logo, a matriz esperada de $\frac{\partial^2 l_j}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T}$ é dada por

$$E\left(\frac{\partial^2 l_j}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T}\right) = \begin{pmatrix} E\left(\frac{\partial^2 l_j}{\partial \boldsymbol{\mu} \partial \boldsymbol{\mu}^T}\right) & E\left(\frac{\partial^2 l_j}{\partial \boldsymbol{\mu} \partial \boldsymbol{\lambda}^T}\right) & E\left(\frac{\partial^2 l_j}{\partial \boldsymbol{\mu} \partial \boldsymbol{\phi}^T}\right) \\ E\left(\frac{\partial^2 l_j}{\partial \boldsymbol{\lambda} \partial \boldsymbol{\lambda}^T}\right) & E\left(\frac{\partial^2 l_j}{\partial \boldsymbol{\lambda} \partial \boldsymbol{\phi}^T}\right) & \\ \text{Simetria} & & E\left(\frac{\partial^2 l_j}{\partial \boldsymbol{\phi} \partial \boldsymbol{\phi}^T}\right) \end{pmatrix},$$

onde

$$E\left(\frac{\partial^2 l_j}{\partial \boldsymbol{\mu} \partial \boldsymbol{\mu}^T}\right) = -\boldsymbol{\Sigma}^{-1},$$

$$E \left(\frac{\partial^2 l_j}{\partial \mu \partial \lambda^T} \right) = E \left(\frac{\partial^2 l_j}{\partial \mu \partial \phi^T} \right) = 0,$$

$$E \left(\frac{\partial^2 l_j}{\partial \lambda \partial \lambda^T} \right) = -(1 - c^{-1}) D^{-1}(\phi) - c^{-1} (2c^{-1} - 1) \mathbf{M},$$

$$E \left(\frac{\partial^2 l_j}{\partial \lambda \partial \phi^T} \right) = -c^{-1} D(\lambda) D^{-2}(\phi) + c^{-2} \mathbf{M} D^{-1}(\phi) D(\lambda),$$

$$E \left(\frac{\partial^2 l_j}{\partial \phi \partial \phi^T} \right) = -\frac{D^{-2}(\phi)}{2} + c^{-1} D^2(\lambda) D^{-3}(\phi) - \frac{c^{-2}}{2} D(\lambda) D^{-1}(\phi) \mathbf{M} D^{-1}(\phi) D(\lambda).$$

Dos resultados apresentados anteriormente, provamos o seguinte Teorema.

Teorema 3.1: *Sob o modelo de calibração estrutural definido em (3.9). A matriz de informação é dada por*

$$J_F(\theta) = \begin{pmatrix} J_{\mu\mu} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & J_{\lambda\lambda} & J_{\lambda\phi} \\ \mathbf{0} & J_{\phi\lambda} & J_{\phi\phi} \end{pmatrix},$$

onde

$$J_{\mu\mu} = \Sigma^{-1},$$

$$J_{\lambda\lambda} = (1 - c^{-1}) D^{-1}(\phi) + c^{-1} (2c^{-1} - 1) \mathbf{M},$$

$$J_{\lambda\phi} = c^{-1} D(\lambda) D^{-2}(\phi) - c^{-2} \mathbf{M} D^{-1}(\phi) D(\lambda),$$

$$J_{\phi\lambda} = (J_{\lambda\phi})^T \text{ e}$$

$$J_{\phi\phi} = \frac{D^{-2}(\phi)}{2} - c^{-1} D^2(\lambda) D^{-3}(\phi) + \frac{c^{-2}}{2} D(\lambda) D^{-1}(\phi) \mathbf{M} D^{-1}(\phi) D(\lambda).$$

Seja

$$\boldsymbol{\theta}_* = (\boldsymbol{\mu}^T, \boldsymbol{\lambda}^T, \boldsymbol{\rho}^T)^T, \quad (3.17)$$

onde $\boldsymbol{\rho} = (\rho_1, \rho_2, \rho_3)^T$ (que é função de $\boldsymbol{\lambda}$ e $\boldsymbol{\phi}$). Então pelo princípio de invariância do EMV, o EMV de $\boldsymbol{\theta}_*$ é dado por

$$\tilde{\boldsymbol{\theta}}_* = \left(\tilde{\boldsymbol{\mu}}^T, \tilde{\boldsymbol{\lambda}}^T, \tilde{\boldsymbol{\rho}}^T \right)^T,$$

onde $\tilde{\boldsymbol{\mu}}$, $\tilde{\boldsymbol{\lambda}}$ e $\tilde{\boldsymbol{\rho}}$ são definidos na seção 3.4.1.

Com o objetivo de testar se os instrumentos medem a característica sem vício e com a mesma confiabilidade, vamos apresentar a distribuição assintótica de $\hat{\boldsymbol{\theta}}_*$.

Teorema 3.2: *Sob o modelo de calibração definido em (3.9), temos que*

a) $\hat{\boldsymbol{\theta}}_*$ é um estimador consistente de $\boldsymbol{\theta}_*$ e

b) $\sqrt{n}(\tilde{\boldsymbol{\theta}}_* - \boldsymbol{\theta}_*) \xrightarrow{D} N_9(0, \boldsymbol{\Omega}(\boldsymbol{\theta}_*)),$

onde

$$\boldsymbol{\Omega}(\boldsymbol{\theta}_*) = \mathbf{G} \mathbf{J}_F^{-1}(\boldsymbol{\theta}) \mathbf{G}^T \quad (3.18)$$

com

$$\mathbf{G} = \begin{pmatrix} \mathbf{I}_3 & 0 & 0 \\ 0 & \mathbf{I}_3 & 0 \\ 0 & \mathbf{G}_1 & \mathbf{G}_2 \end{pmatrix}, \quad (3.19)$$

$$\mathbf{G}_1 = 2D^{-1}(\boldsymbol{\lambda}) D(\boldsymbol{\rho}) [\mathbf{I}_3 - D(\boldsymbol{\rho})] \quad e \quad \mathbf{G}_2 = -D^{-2}(\boldsymbol{\lambda}) D^2(\boldsymbol{\rho}).$$

Prova: Como $\boldsymbol{\theta}_* = g(\boldsymbol{\theta}) = (g_1(\boldsymbol{\theta}), \dots, g_9(\boldsymbol{\theta})) = (\boldsymbol{\mu}^T, \boldsymbol{\lambda}^T, \boldsymbol{\rho}^T)^T$ é uma função diferen-

ciável de θ , então $\tilde{\theta}_* = g(\tilde{\theta})$ converge em probabilidade para $g(\theta) = \theta_*$, o que prova o item (a).

Se $\tilde{\theta}$ é o EMV de θ , então utilizando o Teorema 5.2.1 dado em Sen e Singer (1993), (Capítulo 5) temos que

$$\sqrt{n}(\tilde{\theta} - \theta) \xrightarrow{D} N_9(0, J_F^{-1}(\theta)),$$

logo, usando o método delta temos que

$$\sqrt{n}(\tilde{\theta}_* - \theta_*) \xrightarrow{D} N_9(0, \Omega(\theta_*)),$$

onde $\Omega(\theta_*) = \mathbf{G} J_F^{-1}(\theta) \mathbf{G}^T$, com

$$\mathbf{G} = \left[\frac{\partial(g_i(\theta))}{\partial \theta_j} \right] = \begin{bmatrix} \mathbf{I}_3 & 0 & 0 \\ 0 & \mathbf{I}_3 & 0 \\ 0 & \mathbf{G}_1 & \mathbf{G}_2 \end{bmatrix}$$

sendo uma matriz 9×9 e $\mathbf{G}_1$ e $\mathbf{G}_2$ dados como acima.

Corolário 3.1 : *Sob o modelo de calibração definido em (3.9).*

$$\sqrt{n}(\tilde{\rho} - \rho) \xrightarrow{D} N_3(0, \mathbf{Q} J_F^{-1}(\theta) \mathbf{Q}^T),$$

onde $\mathbf{Q} = (0, \mathbf{G}_1, \mathbf{G}_2)$, com $\mathbf{G}_1$ e $\mathbf{G}_2$ como no Teorema 3.2.

No que segue desta seção propomos intervalo de confiança, para o parâmetro do modelo sob a parametrização de análise de fatores, baseado nos resultados apresentados acima.

Se $\tilde{\theta}$ é o EMV de θ , então

$$\sqrt{n}(\tilde{\theta} - \theta) \xrightarrow{D} N_9(0, J_F^{-1}(\theta)),$$

é uma matriz 6×9 de posto 6, com

$$\mathbf{C}_1 = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \quad (3.20)$$

e $\mathbf{q}_0 = \mathbf{0}$ vetor nulo de dimensão 6.

Logo, a estatística de Wald para testar H_{01} , denotada por W_{01} , é dada por

$$W_{01} = n \left(\mathbf{C} \tilde{\boldsymbol{\theta}}_* - \mathbf{q}_0 \right)^T \left(\mathbf{C} \tilde{\boldsymbol{\Omega}}(\boldsymbol{\theta}_*) \mathbf{C}^T \right)^{-1} \left(\mathbf{C} \tilde{\boldsymbol{\theta}}_* - \mathbf{q}_0 \right),$$

que converge em distribuição a uma variável aleatória $\chi^2_{(6)}$ sob H_{01} ; quando $n \rightarrow \infty$, $\tilde{\boldsymbol{\Omega}}(\boldsymbol{\theta}_*)$ é a matriz dada em $\boldsymbol{\Omega}(\boldsymbol{\theta}_*)$ (3.18) avaliada no EMV de $\boldsymbol{\theta}_*, \tilde{\boldsymbol{\theta}}_*$.

Portanto, H_{01} é rejeitada ao nível α se $W_{01} > \chi^2_{1-\alpha}(6)$, onde $\chi^2_{1-\alpha}(6)$ denota o percentil do 100 $(1 - \alpha)$ % da distribuição $\chi^2(6)$.

Para testar se os instrumentos medem sem vício, a hipótese poderia ser

$$H_{02} : \mu_1 = \mu_2 = \mu_3, \quad \lambda_1 = \lambda_2 = \lambda_3.$$

Note que, H_{02} pode ser escrita como

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

onde

$$\mathbf{C} = \begin{pmatrix} \mathbf{C}_1 & 0 & 0 \\ 0 & \mathbf{C}_1 & 0 \end{pmatrix}$$

é uma matriz 4×9 de posto 4 com $\mathbf{C}_1$ como em (3.20) e $\mathbf{q}_0 = \mathbf{0}$ um vetor nulo de dimensão 4. Assim, a estatística de Wald para testar H_{02} , denotada por W_{02} , é dada por

$$W_{02} = n \left(\mathbf{C} \tilde{\boldsymbol{\theta}}_* \right)^T \left(\mathbf{C} \tilde{\boldsymbol{\Omega}}(\boldsymbol{\theta}_*) \mathbf{C}^T \right)^{-1} \left(\mathbf{C} \tilde{\boldsymbol{\theta}}_* \right),$$

que converge em distribuição a uma variável aleatória $\chi^2_{(4)}$ sob H_{02} .

Portanto, H_{02} é rejeitada ao nível α se $W_{02} > \chi^2_{1-\alpha}(4)$, onde $\chi^2_{1-\alpha}(4)$ denota o percentil do 100(1 - α) % da distribuição $\chi^2(4)$.

Finalmente, para testar se os instrumentos são igualmente confiáveis, a hipótese de interesse é

$$H_{03} : \rho_1 = \rho_2 = \rho_3 \text{ ou } H_{03} : C\rho = 0,$$

C_1 como em (3.20). Assim, a estatística de Wald para testar H_{03} , que denotamos por W_{03} , é obtida do corolário 3.1 e é dada por

$$W_{03} = n(C_1\rho)^T \left(C_1 Q J_F^{-1}(\tilde{\theta}) Q^T C_1^T \right)^{-1} (C_1\rho),$$

que converge em distribuição a uma variável aleatória $\chi^2_{(2)}$ sob H_{03} .

Portanto, H_{03} é rejeitada ao nível α se $W_{03} > \chi^2_{1-\alpha}(2)$, onde $\chi^2_{1-\alpha}(2)$ denota o percentil do 100(1 - α) % da distribuição $\chi^2(2)$.

Usando o princípio de invariância do EMV, também pode ser estudado o comportamento assintótico das precisões e de maneira similar podemos testar a igualdade de precisões ou precisões relativas.

Embora o estudo apresentado anteriormente tenha sido desenvolvido sob a parametrização de análise de fatores, aplicando o método delta podem ser obtidos resultados baseada na parametrização original em função dos parâmetros α , β e ϕ .

3.5 Aplicação

Conjunto de dados Chipkevitch et al. (1996)

A estimação do tamanho testicular em adolescentes é importante para a avaliação do desenvolvimento normal da puberdade, para diagnósticos de distúrbios e para a valiação dos efeitos das doenças ou a terapia da função gonadal. Desde que semisferios tubu-

lares e elementos germinais compreendem aproximadamente o 90% da masa testicular, o tamanho do testículo de um adulto reflete aproximadamente a função espermagênico. A medição clínica, do tamanho testicular na puberdade, é um método mais efetivo e avaliável para a estimação da função testicular que outros métodos existentes, tais como espermioigramas ou a biopsia, os quais podem ser invasivos especialmente em adolescentes.

Desde que os testículos são aproximadamente elipsoides rotacionais, o volume (V) pode ser calculado para as medições de largura (L_3) e comprimento (C_3) usando um calibrador. (veja, Rundel, 1962). A comparação dos testículos com uma leitura feita a modelos de elipsoides graduados é outro caminho para avaliar o volume testicular (veja, Schonfeld, 1943). Os autores propõem um novo método simple para a medição do tamanho testicular baseado em uma comparação com modelo gráficos. O principal propósito deste estudo foi comparar estatisticamente a confiabilidade deste método com outros dois métodos bem conhecidos (calibrador e ultra-som). No artigo consideram cinco métodos, nós vamos considerar três e a medição do testículo direito para efeitos da aplicação.

Indivíduos: Todos os pacientes vistos na divisão da medicina de adolescentes do Hospital da Criança Darcy Vargas em uma consulta da saúde com ultra-som foi feito para a suspeitada patologia escrotal durante o ano 1993. Estes pacientes foram consecutivamente associados neste estudo.

Os pacientes com uma grande hidrocele (derrame líquido entre as túnica da vaginal do testículo) ou outras condições que poderiam prejudicar as medições clínicas dos testículos foram excluídos do estudo. Acredita-se que varicocele (tumor produzido pela dilatação varicosa das veias do cordão espermático) não prejudica a medição com o paciente supino. O protocolo do estudo foi aprovado pelo comite de bioética do Hospital e um informe do consentimento de todos os pacientes foi obtido.

A amostra inclui 42 adolescentes de 11 e 17 anos de idade (a idade média com mas ou menos um desvio padrão de $13,9 \pm 1,3$). O desenvolvimento da puberdade varia do estágio genital de Tanner de I a V (estágios III e IV em 69% dos casos). A examinação

clínica e o ultra-som revelam uma varicocele em 29 pacientes (69%), o qual foi sobre o lado esquerdo em 22 e o lado direito em 7, o grau I em 15, grau II em 9 e o grau III em 5. Três pacientes tiveram uma grande hidrocele e três um cisto no epididimo (na parte superior do testículo e que é o conduto formado pela reunião dos canais seminíferos). Em 10 indivíduos (24%) se observou a patologia escrotal.

Medições: Todos os indivíduos foram examinados por um dos autores do artigo. As genitais foram observadas com os pacientes deitados e sentados. Para a manobra Valsalva, a avaliação foi realizada com o paciente em pé. Desenvolvimento da puberdade foi avaliada de acordo à classificação de Tanner (veja, Tanner, 1962). Medições dos testículos foram feitas com o indivíduo supino e o testículo esquerdo foi sempre medido primeiro.

Três métodos foram usados:

Método 1: método já existente no mercado e de alta precisão, que consiste em um sistema de ultra-som para se obter as dimensões do testículo: largura (L_1), comprimento (C_1) e profundidade (P) e a partir daí calcular seu volume (V_1) através da fórmula $V_1 = \pi/6 * L_1 * C_1 * P$.

Método 2: método gráfico que consiste em 6 anéis elípticos que representam volumes de 2 ml, 5 ml, 10 ml, 15 ml e 20 ml e através desses anéis é feita de forma visual a estimativa do volume do testículo (ver figura 3.1). No caso de um volume estar entre dois anéis, por exemplo entre 10 ml e 15 ml, o volume estimado passa ser a média aritmética entre esses dois, (ou seja, 12.5 ml, no exemplo).

Método 3: método de medidas dimensionais, onde o comprimento (C_3) e a largura (L_3) do testículo são obtidos através de um calibrador específico e o volume é calculado pela fórmula $V_3 = \pi/6 * L_3 * C_3^2$.

Seja Y_{ij} a medida do volume do testículo em ml fornecida pelo método i para o paciente j , com $i = 1, 2, 3$ e $j = 1, \dots, 42$. O conjunto das observações $\{Y_{ij}\}$ é apresentado na Tabela 3.1.

Por ser um método conhecido e comumente usado, adotaremos método 1, como sendo de referência. Sendo assim, de acordo com a seção 3.3, assumimos que $\alpha_1 = 0$ e $\beta_1 = 1$.

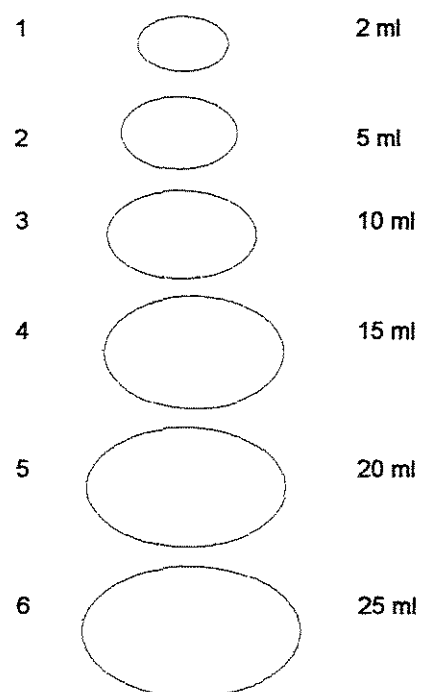


Figura 3-1: Método gráfico para a medição do volume testicular

Tabela 3.1: Medidas do volume do testículo direito humano para 42 adolescentes entre 11 e 17 anos de idade, feitas por três métodos de medição. Chipkevitch et al. (1996).

Obs.	Y_1^1	Y_2^2	Y_3^3	Obs.	Y_1	Y_2	Y_3
1	5,00	7,50	5,90	22	16,50	10,00	15,30
2	5,70	5,00	4,80	23	4,50	3,50	3,90
3	7,40	5,00	6,80	24	5,60	5,00	4,50
4	2,60	3,50	3,10	25	11,00	7,50	9,70
5	5,70	5,00	5,00	26	9,20	10,00	11,30
6	6,10	5,00	4,40	27	8,50	7,50	8,80
7	6,20	5,00	6,00	28	5,40	5,00	6,10
8	10,40	10,00	8,80	29	6,70	7,50	7,20
9	9,10	7,50	7,90	30	5,30	5,00	5,90
10	14,80	10,00	13,00	31	20,00	20,00	16,30
11	16,40	12,50	10,30	32	18,80	15,00	16,30
12	9,60	7,50	8,20	33	13,90	12,50	12,20
13	15,70	15,00	19,80	34	9,40	10,00	10,30
14	3,00	2,00	2,00	35	9,10	7,50	10,80
15	16,40	15,00	17,30	36	14,10	15,00	13,00
16	17,60	15,00	17,30	37	9,30	10,00	8,40
17	10,00	7,50	7,90	38	20,90	20,00	22,10
18	4,10	3,50	4,40	39	11,50	10,00	10,60
19	12,70	10,00	11,40	40	9,70	10,00	9,70
20	2,70	3,50	4,10	41	13,70	12,50	11,60
21	10,20	10,00	11,10	42	8,90	10,00	8,10

¹ medida do volume do testículo em ml fornecida pelo método 1 para o paciente j .

² medida do volume do testículo em ml fornecida pelo método 2 para o paciente j .

³ medida do volume do testículo em ml fornecida pelo método 3 para o paciente j .

Como uma forma de avaliar o ajuste do modelo estrutural normal, construiremos o gráfico da distância transformada (ver Little, 1988 e Lange et al., 1989).

Seja

$$d_j^2 = (Y_j - \mu)^T \Sigma^{-1} (Y_j - \mu), \quad j = 1, \dots, n.$$

a distância quadrática de Mahalanobis. Sob normalidade, sabemos que d_j^2 tem uma distribuição χ^2 com 3 graus de liberdade. Ademais, $\tilde{d}_j^2$ tem assintoticamente a mesma distribuição χ^2 , onde $\tilde{d}_j^2 = d_j^2(\bar{\theta})$, $j = 1, \dots, n$. Usando a aproximação de Wilson e Hilferty (Johnson e Kotz, 1970: 176) temos que

$$z_j = \frac{\left\{ \left(\tilde{d}_j^2/3 \right)^{1/3} - \left(1 - \frac{2}{27} \right) \right\}}{\{2/27\}^{1/2}},$$

$j = 1, \dots, n$, é aproximadamente distribuído normal padrão. Um Q-Q plot das distâncias transformadas z_j , $j = 1, \dots, n$, dadas pela equação de acima, pode ser utilizada para avaliar o ajuste do modelo normal.

A Figura 3.1 mostra o gráfico da distância transformada para o modelo normal. Pode-se observar um ajuste razoável, apesar de que algumas observações poderiam ser consideradas “outliers”.

Para detectar estes possíveis “outliers” utilizamos a distância de Mahalanobis (ver Johnson e Wichern, 1998: pag. 203). Para isto é necessário obter os valores de z_j e $\tilde{d}_j^2$ que são dados na Tabela 3.2.

Os valores na coluna $\tilde{d}_j^2$ na tabela 3.2 revelam que as medições dos testículos dos adolescentes 11, 13, 22, 31 e 38 são “outliers” multivariados, desde que $\chi_3^2(0,05) = 7,81$.

Segundo o modelo (3.3) as variáveis observadas seriam:

Y_{1j} : medida fornecida pelo método 1 associado ao adolescente j ,

Y_{2j} : medida fornecida pelo método 2 associado ao adolescente j ,

Y_{3j} : medida fornecida pelo método 3 associado ao adolescente j , $j = 1, \dots, 42$.

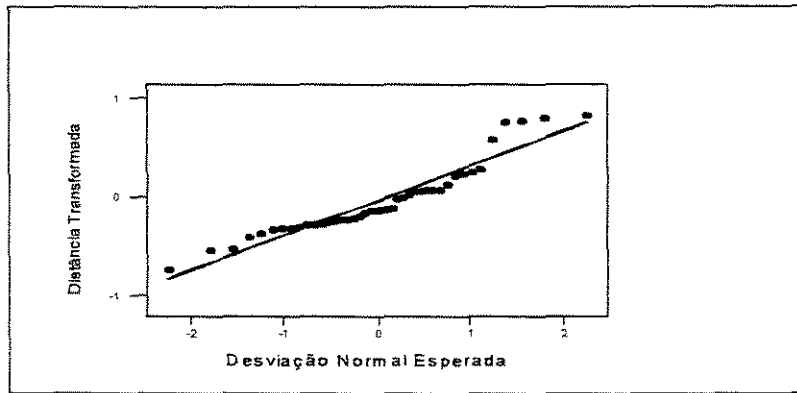


Figura 3-2: Gráfico da distância transformada para o modelo normal

$$\bar{Y}_1 = 10,081, \bar{Y}_2 = 9,012, \bar{Y}_3 = 9,562$$

$$S_{11} = 23,588, S_{22} = 18,708, S_{33} = 21,731$$

$$S_{12} = 19,679, S_{13} = 21,296, S_{23} = 18,674$$

Tabela 3.2: Valores de z_j e $\tilde{d}_j^2$ do volume do testículo humano para 42 adolescentes entre 11 e 17 anos de idade, feitas por três métodos de medição

Observ.	z_j	$\tilde{d}_j^2$	Observ.	z_j	$\tilde{d}_j^2$
1	0.241	4,248	22	0.761	11,106
2	-0.271	1,050	23	-0.133	1,638
3	-0.121	1,700	24	-0.222	1,238
4	0.025	2,543	25	-0.018	2,268
5	-0.300	0,947	26	0.071	2,856
6	-0.136	1,621	27	-0.525	0,369
7	-0.312	0,907	28	-0.230	1,208
8	-0.274	1,039	29	-0.228	1,216
9	-0.530	0,359	30	-0.246	1,147
10	0.287	4,682	31	0.818	12,140
11	0.770	11,267	32	0.217	4,028
12	-0.401	0,643	33	-0.275	1,036
13	0.803	11,866	34	-0.208	1,298
14	0.068	2,838	35	0.072	2,863
15	0.126	3,264	36	0.249	4,321
16	0.075	2,888	37	-0.115	1,728
17	-0.193	1,359	38	0.585	8,292
18	-0.123	1,687	39	-0.725	0,112
19	-0.316	0,896	40	-0.368	0,733
20	0.069	2,839	41	-0.152	1,547
21	-0.325	0,866	42	-0.005	2,348

Cálculo das estimativas para o modelo estrutural de regressão de calibração

De acordo com a subseção 3.3.2., calculamos os estimadores de máxima verossimilhança para o conjunto de dados em estudo

$$\tilde{\alpha}_2 = 0,171, \tilde{\alpha}_3 = -0,0047, \mu_x = 10,081$$

$$\tilde{\beta}_2 = 0,877, \tilde{\beta}_3 = 0,949$$

$$\sigma_x^2 = 22,442, \sigma_1^2 = 1,149$$

$$\sigma_2^2 = 1,452, \sigma_3^2 = 1,526$$

$$\tilde{\rho}_1 = 0,951, \tilde{\rho}_2 = 0,923, \tilde{\rho}_3 = 0,930.$$

Cálculo das estimativas para o modelo estrutural de calibração de análise de fatores

De acordo com a subseção 3.4.1., calculamos os estimadores de máxima verossimilhança para o conjunto de dados em estudo,

$$\tilde{\mu}_1 = 10,081, \tilde{\mu}_2 = 9,012, \tilde{\mu}_3 = 9,562$$

$$\tilde{\lambda}_1 = 4,737, \tilde{\lambda}_2 = 4,154, \tilde{\lambda}_3 = 4,495$$

$$\tilde{\phi}_1 = 1,149, \tilde{\phi}_2 = 1,452, \tilde{\phi}_3 = 1,526$$

$$\tilde{\rho}_1 = 0,951, \tilde{\rho}_2 = 0,923, \tilde{\rho}_3 = 0,930.$$

Matriz Estimada de Informação de Fisher

A obtenção da matriz esperada de Fisher é de grande utilidade para se obter a distribuição assintótica dos estimadores. Portanto, de acordo com o Teorema 3.1, a matriz esperada

de informação de Fisher é dada por

$$J_F(\tilde{\theta}) = \begin{pmatrix} \tilde{J}_{\mu\mu} & 0 & 0 \\ 0 & \tilde{J}_{\lambda\lambda} & \tilde{J}_{\lambda\phi} \\ 0 & \tilde{J}_{\phi\lambda} & \tilde{J}_{\phi\phi} \end{pmatrix},$$

onde

$$\tilde{J}_{\mu\mu} = \begin{pmatrix} 0,498 & -0,258 & -0,266 \\ -0,258 & 0,509 & -0,185 \\ -0,266 & -0,185 & 0,465 \end{pmatrix}, \quad \tilde{J}_{\lambda\lambda} = \begin{pmatrix} 0,495 & -0,247 & -0,254 \\ -0,247 & 0,502 & -0,177 \\ -0,254 & -0,177 & 0,459 \end{pmatrix},$$

$$\tilde{J}_{\lambda\phi} = \begin{pmatrix} 0,045 & -0,016 & -0,017 \\ -0,023 & 0,032 & -0,012 \\ -0,024 & -0,012 & 0,030 \end{pmatrix}, \quad \tilde{J}_{\phi\lambda} = \begin{pmatrix} 0,045 & -0,023 & -0,024 \\ -0,016 & 0,032 & -0,012 \\ -0,017 & -0,012 & 0,030 \end{pmatrix},$$

$$\tilde{J}_{\phi\phi} = \begin{pmatrix} 0,124 & 0,033 & 0,035 \\ 0,033 & 0,130 & 0,017 \\ 0,035 & 0,017 & 0,108 \end{pmatrix}.$$

Inversa da matriz esperada de informação de Fisher

$$J_F^{-1}(\tilde{\theta}) = \begin{pmatrix} 23,588 & 19,678 & 21,293 & 0 & 0 & 0 & 0 & 0 & 0 \\ 19,678 & 18,708 & 18,672 & 0 & 0 & 0 & 0 & 0 & 0 \\ 21,293 & 18,672 & 21,731 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 12,450 & 9,810 & 10,610 & -0,772 & 0,242 & 0,325 \\ 0 & 0 & 0 & 9,810 & 10,143 & 9,328 & 0,276 & -0,521 & 0,076 \\ 0 & 0 & 0 & 10,610 & 9,328 & 11,702 & 0,343 & 0,071 & -0,659 \\ 0 & 0 & 0 & -0,772 & 0,276 & 0,343 & 9,957 & -2,289 & -3,082 \\ 0 & 0 & 0 & 0,242 & -0,521 & 0,071 & -2,289 & 8,543 & -0,635 \\ 0 & 0 & 0 & 0,325 & 0,076 & -0,660 & -3,082 & -0,635 & 10,589 \end{pmatrix}$$

Intervalos de confiança para os θ_i com um nível do 95%

$$\mu_1 : [8,612; 11,550], \mu_2 : [7,704; 10,320], \mu_3 : [8,152; 10,972]$$

$$\lambda_1 : [3,670; 5,804], \lambda_2 : [3,191; 5,117], \lambda_3 : [3,460; 5,530]$$

$$\phi_1 : [0,195; 2,103], \phi_2 : [0,568; 2,336], \phi_3 : [0,542; 2,510]$$

$$\rho_1 : [0,906; 0,996], \rho_2 : [0,866; 0,980], \rho_3 : [0,877; 0,983]$$

Teste de Hipótese

Para provar se os três instrumentos medem a característica sem vício e com a mesma confiabilidade, a hipótese é

$$H_{01} : \mu_1 = \mu_2 = \mu_3, \lambda_1 = \lambda_2 = \lambda_3, \rho_1 = \rho_2 = \rho_3.$$

Para realizarmos esse teste de hipótese, foi calculada a matriz $\tilde{\Omega}(\theta_*)$ que é a matriz $\Omega(\theta_*)$ dada em (3.18) e avaliada no EMV de θ_* , $\tilde{\theta}_*$.

$$\tilde{\Omega} \theta_* = \begin{pmatrix} 23,588 & 19,678 & 21,293 & 0 & 0 & 0 & 0 & 0 & 0 \\ 19,678 & 18,708 & 18,672 & 0 & 0 & 0 & 0 & 0 & 0 \\ 21,293 & 18,672 & 21,731 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 12,450 & 9,810 & 10,610 & 0,276 & 0,324 & 0,293 \\ 0 & 0 & 0 & 9,810 & 10,143 & 9,328 & 0,182 & 0,373 & 0,267 \\ 0 & 0 & 0 & 10,610 & 9,328 & 11,702 & 0,195 & 0,316 & 0,367 \\ 0 & 0 & 0 & 0,276 & 0,182 & 0,195 & 0,022 & 0,001 & 0,000 \\ 0 & 0 & 0 & 0,324 & 0,373 & 0,316 & 0,001 & 0,035 & 0,008 \\ 0 & 0 & 0 & 0,293 & 0,267 & 0,367 & 0,000 & 0,008 & 0,031 \end{pmatrix},$$

$$\tilde{\theta}_* = (10,081 \ 9,012 \ 9,562 \ 4,737 \ 4,154 \ 4,495 \ 0,951 \ 0,923 \ 0,930)^T$$

Usando estes resultados e a estatística de Wald, podemos testar a hipótese mencionada acima. Com efeito, com $n = 42$, temos $W_{01} = 30,997$, comparando com a $\chi^2_{5\%}(6) = 12,59$, se rejeita a hipótese nula, isto é os métodos são viciados, ou seja medem com vício o volume do testículo direito e/ou não são igualmente confiáveis. Para obter informação respeito à rejeição da hipótese H_{01} , vamos testar em forma separada se os instrumentos medem com vício e se tem a mesma confiabilidade.

Para testar se os três instrumentos medem a característica sem vício a hipótese é,

$$H_{02} : \mu_1 = \mu_2 = \mu_3, \quad \lambda_1 = \lambda_2 = \lambda_3.$$

Usando os resultados de $\tilde{\Omega}(\theta_*)$ e a estatística de Wald, podemos testar a hipótese mencionada acima. Com efeito, com $n = 42$, temos $W_{02} = 31,215$, comparando com a $\chi^2_{5\%}(4) = 9,49$, rejeita-se a hipótese nula, isto é, os métodos são viciados, ou seja, medem com vício o volume do testículo direito. Não podemos dizer que o método i fornece a medida “exata” do volume do testículo direito.

Para provar se os três instrumentos são igualmente confiáveis, a hipótese é

$$H_{03} : \rho_1 = \rho_2 = \rho_3.$$

Usando estes resultados e a estatística de Wald, podemos testar a hipótese mencionada acima. Com efeito, com $n = 42$, temos $W_{03} = 0,655$, comparando com a $\chi^2_{5\%}(2) = 5,99$, não rejeita-se a hipótese nula, isto é, os métodos são igualmente confiáveis.

Dos resultados acima, concluímos que de uma análise estatística de nossos dados indicam que os dois métodos são tão bons como o método de ultra-som, ou seja são igualmente confiáveis para a medição do volume do testículo direito. O uso de diferentes tamanhos dos anéis para o método gráfico permite uma rápida comparação visual para os testículos. O tempo de manipulação parece ser reduzido comparando com outros métodos de medição, decrescendo o natural embaraçamento dos adolescentes.

Capítulo 4

Influência Local no Modelo de Calibração Comparativa

4.1 Introdução

Neste capítulo vamos desenvolver o estudo de influência e diagnóstico no modelo de calibração comparativa, onde é de interesse comparar métodos ou instrumentos de medição da mesma quantidade desconhecida x em um grupo de n unidades experimentais, supondo as medições dos instrumentos seguindo uma distribuição multivariada.

Especificamente, vamos aplicar o método de influência local proposto por Cook (1986) no modelo de calibração comparativa para três instrumentos de medição, baseado na parametrização de análise de fatores considerada por Theobald e Mallison (1978).

Na literatura são poucos os trabalhos sobre influência e diagnóstico no modelo com erros nas variáveis, em particular no modelo de calibração comparativa. Alguns trabalhos precursores em influência e diagnóstico no modelo com erros nas variáveis 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 na influência de uma observação individual em análise

de estruturas de covariância. Abdullah (1995) apresenta vários métodos para detectar observações influentes no modelo de regressão funcionais com erros 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 no tempo, para todas as observações 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. Resultados adicionais sob influência local e aplicações no modelo linear misto e de regressão podem ser encontrados em Beckman, Natchtsheim e Cook (1987), Lawrence (1988), Thomas e Cook (1990), Tsai e Wu (1992), Paula (1993), Galea-Rojas, Paula e Bolfarine (1997) e Lasafre e Verbeke (1998). Zhao e Lee (1996) utilizam o método de influência local em alguns modelos com erro linear e não-linear.

Recentemente, Zhao e Lee (1998) e Kwan e Fung (1998) aplicam o método de influência local em modelos de equações simultâneas e em análise de fator, respectivamente.

Contudo, este método não tem sido desenvolvido no modelo de calibração comparativa baseado na parametrização de análise de fatores. Neste capítulo, aplicamos o método de influência local apresentado no Capítulo II, considerando dois esquemas de perturbação: ponderação de casos e perturbação na resposta. No esquema de perturbação na resposta colocamos bastante atenção nas medidas obtidas pelo instrumento de referência.

No que segue, vamos introduzir alguns tipos de perturbações no modelo de calibração comparativa sob a parametrização de análise de fatores, a fim de estudarmos a influência local. Especificamente:

1. Ponderação de Casos

Consideremos o vetor de pesos $\omega = (\omega_1, \dots, \omega_n)^T$, para ponderar a contribuição de cada caso na função de log-verossimilhança. Por exemplo, se $\omega = (1, 0.5, 0, \dots, 0)^T$, a con-

tribuição do primeiro caso é o duplo que a contribuição do segundo caso na log-verossimilhança. Neste caso, o modelo não perturbado é obtido ao considerar $\omega_0 = (1, \dots, 1)^T$, isto é, a contribuição de todos os casos é igual na log-verossimilhança. A função de log-verossimilhança do modelo perturbado é dado por

$$l(\theta/\omega) = \sum_{j=1}^n \omega_j l(\theta, \mathbf{y}_j).$$

2. Perturbação na observação

Seja Y_{wj} a perturbação na observação Y_j , onde $j = 1, \dots, n$. Algumas situações que serão de interesse estudar são:

2.1. Perturbação Aditiva: Neste caso o modelo perturbado é dado por

$$\mathbf{Y}_{wj}^{(k)} = \mathbf{Y}_j + w_{kj} \mathbf{e}_k, \quad j = 1, \dots, n, \quad (4.1)$$

onde $\mathbf{e}_k$ é o k -ésimo vetor unitário de $\mathbf{R}^3$ e w_{kj} é a perturbação aditiva correspondente à medida do k -ésimo instrumento no método de medição. Sob esta perturbação, o modelo perturbado reduz-se ao modelo postulado quando $\omega_0 = (0, \dots, 0)^T$.

2.2. Perturbação Multiplicativa: Neste caso o modelo perturbado é dado por

$$\mathbf{Y}_{wj}^{(k)} = \mathbf{Y}_j * \mathbf{1}_3(w_{kj}), \quad j = 1, \dots, n, \quad (4.2)$$

onde $\mathbf{1}_3(w_{kj})$ é o vetor em $\mathbf{R}^3$, com w_{kj} na k -ésima posição, $k = 1, 2, 3$ e 1s em outro caso, isto é, $\mathbf{Y}_j * \mathbf{1}_3(w_{kj})$ representa a perturbação multiplicativa correspondente à medida do k -ésimo instrumento de medição. A notação $*$ denota o produto de Hadamard. Por exemplo, o modelo perturbado para $k = 2$ é dado por

$$\mathbf{Y}_{wj} = \mathbf{Y}_j * \begin{pmatrix} 1 \\ w_{2j} \\ 1 \end{pmatrix} = \begin{pmatrix} Y_{1j} \\ w_{2j} Y_{2j} \\ Y_{3j} \end{pmatrix}$$

ou seja, estamos perturbando somente as medidas do segundo instrumento. Para simplificar a notação, no que segue deste trabalho, representaremos $\mathbf{Y}_{wj}^{(k)}$ como sendo $\mathbf{Y}_{wj}$. Agora, seja $\mathbf{Y}_{wj}$ a perturbação na observação $\mathbf{Y}_j$ (aditiva ou multiplicativa). Então a função de log-verossimilhança sob o modelo perturbado para $\boldsymbol{\omega} = (\omega_{k1}, \dots, \omega_{kn})^T$ é dada por

$$l(\boldsymbol{\theta}/\boldsymbol{\omega}) = \sum_{j=1}^n l_j(\boldsymbol{\theta}/\omega_{kj}), \quad (4.3)$$

onde

$$l_j(\boldsymbol{\theta}/\omega_{kj}) = \log(2\pi)^{-3/2} - \frac{1}{2} \log |\Sigma| - \frac{1}{2} \|T_{wj}\|^2$$

e

$$\|T_{wj}\|^2 = (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{Y}_{wj} - \boldsymbol{\mu}), \quad j = 1, \dots, n.$$

Derivando $l(\boldsymbol{\theta}/\boldsymbol{\omega})$ com respeito a $\boldsymbol{\omega}^T$, obtemos

$$\frac{\partial l(\boldsymbol{\theta}/\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^T} = (A_1(\boldsymbol{\theta}, w_{k1}), \dots, A_n(\boldsymbol{\theta}, w_{kn})), \quad (4.4)$$

onde

$$A_j(\boldsymbol{\theta}, w_{kj}) = -\frac{1}{2} \frac{\partial}{\partial w_{kj}} \|T_{wj}\|^2 = -\frac{1}{2} \frac{\partial (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T}{\partial w_{kj}} \frac{\partial (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{Y}_{wj} - \boldsymbol{\mu})}{\partial (\mathbf{Y}_{wj} - \boldsymbol{\mu})},$$

$j = 1, \dots, n$.

Então derivando $\frac{\partial l(\boldsymbol{\theta}/\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^T}$ com respeito a $\boldsymbol{\theta}$, temos que a matriz Δ dada por

$$\begin{aligned} \Delta &= \frac{\partial^2 l(\boldsymbol{\theta}/\boldsymbol{\omega})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\omega}^T} = \frac{\partial}{\partial \boldsymbol{\theta}} \left(\frac{\partial l(\boldsymbol{\theta}/\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^T} \right) = \frac{\partial}{\partial \boldsymbol{\theta}} (A_1(\boldsymbol{\theta}, w_{k1}), \dots, A_n(\boldsymbol{\theta}, w_{kn})) \quad (4.5) \\ &= \left(\frac{\partial}{\partial \boldsymbol{\theta}} A_1(\boldsymbol{\theta}, w_{k1}), \dots, \frac{\partial}{\partial \boldsymbol{\theta}} A_n(\boldsymbol{\theta}, w_{kn}) \right) \end{aligned}$$

$$\begin{aligned}
&= \begin{pmatrix} \frac{\partial}{\partial \mu} A_1(\boldsymbol{\theta}, w_{k1}) & \cdots & \frac{\partial}{\partial \mu} A_n(\boldsymbol{\theta}, w_{kn}) \\ \frac{\partial}{\partial \lambda} A_1(\boldsymbol{\theta}, w_{k1}) & \cdots & \frac{\partial}{\partial \lambda} A_n(\boldsymbol{\theta}, w_{k1}) \\ \frac{\partial}{\partial \phi} A_1(\boldsymbol{\theta}, w_{k1}) & \cdots & \frac{\partial}{\partial \phi} A_n(\boldsymbol{\theta}, w_{kn}) \end{pmatrix} \\
&= \begin{pmatrix} \Delta_{\mu 1} & \cdots & \Delta_{\mu n} \\ \Delta_{\lambda 1} & \cdots & \Delta_{\lambda n} \\ \Delta_{\phi 1} & \cdots & \Delta_{\phi n} \end{pmatrix}_{9 \times n},
\end{aligned}$$

onde $A_j(\boldsymbol{\theta}, w_{kj})$ é como em (4.4) e $\Delta_{\mu_j} = \frac{\partial}{\partial \mu} A_j(\boldsymbol{\theta}, w_{kj})$, $\Delta_{\lambda_j} = \frac{\partial}{\partial \lambda} A_j(\boldsymbol{\theta}, w_{kj})$ e $\Delta_{\phi_j} = \frac{\partial}{\partial \phi} A_j(\boldsymbol{\theta}, w_{kj})$, $j = 1, \dots, n$.

A matriz Δ definida acima tem que ser avaliada em $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ e $\boldsymbol{\omega} = \boldsymbol{\omega}_0$.

Na seguinte seção vamos desenvolver o método de influência local definido no capítulo II para o modelo de calibração comparativa, considerando os esquemas de perturbação definidos acima.

4.2 Perturbação de Ponderação de Casos

Considerando o vetor de pesos $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n)^T$. A função de verossimilhança do modelo perturbado é dada por

$$l(\boldsymbol{\theta} / \boldsymbol{\omega}) = \sum_{j=1}^n \omega_j l_j(\boldsymbol{\theta}), \quad (4.6)$$

onde

$$l_j(\boldsymbol{\theta}) = \log(2\pi)^{-3/2} - \frac{1}{2} \log |\Sigma| - \frac{1}{2} \|T_j\|^2$$

com $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n)^T$ e $\boldsymbol{\theta} = (\boldsymbol{\mu}^T, \boldsymbol{\lambda}^T, \boldsymbol{\phi}^T)^T$. Derivando (4.6) com respeito a $\boldsymbol{\omega}^T$ temos que

$$\frac{\partial l(\boldsymbol{\theta} / \boldsymbol{\omega})}{\partial \boldsymbol{\omega}^T} = (l_1(\boldsymbol{\theta}), \dots, l_n(\boldsymbol{\theta})),$$

onde $l_j(\boldsymbol{\theta})$, $j = 1, \dots, n$, são como em (4.6). Logo derivando, com respeito a $\boldsymbol{\theta}$ a expressão anterior, temos

$$\begin{aligned}\Delta &= \frac{\partial^2 l(\boldsymbol{\theta}/\boldsymbol{\omega})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\omega}^T} = \frac{\partial}{\partial \boldsymbol{\theta}} \left(\frac{\partial l(\boldsymbol{\theta}/\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^T} \right) = \frac{\partial}{\partial \boldsymbol{\theta}} (l_1(\boldsymbol{\theta}), \dots, l_n(\boldsymbol{\theta})) \\ &= \left(\frac{\partial l_1(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}, \dots, \frac{\partial l_n(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) = \begin{pmatrix} \frac{\partial l_1(\boldsymbol{\theta})}{\partial \boldsymbol{\mu}} & \dots & \frac{\partial l_n(\boldsymbol{\theta})}{\partial \boldsymbol{\mu}} \\ \frac{\partial l_1(\boldsymbol{\theta})}{\partial \boldsymbol{\lambda}} & \dots & \frac{\partial l_n(\boldsymbol{\theta})}{\partial \boldsymbol{\lambda}} \\ \frac{\partial l_1(\boldsymbol{\theta})}{\partial \boldsymbol{\phi}} & \dots & \frac{\partial l_n(\boldsymbol{\theta})}{\partial \boldsymbol{\phi}} \end{pmatrix} \\ &= \begin{pmatrix} \Delta_{\mu_1} & \dots & \Delta_{\mu_n} \\ \Delta_{\lambda_1} & \dots & \Delta_{\lambda_n} \\ \Delta_{\phi_1} & \dots & \Delta_{\phi_n} \end{pmatrix}_{9 \times n},\end{aligned}\tag{4.7}$$

onde

$$\Delta_{\mu_j} = \frac{\partial}{\partial \boldsymbol{\mu}} l_j(\boldsymbol{\theta}) = -\frac{1}{2} \left[\frac{\partial}{\partial \boldsymbol{\mu}} \log |\boldsymbol{\Sigma}| + \frac{\partial}{\partial \boldsymbol{\mu}} \|T_j\|^2 \right]$$

$$\Delta_{\lambda_j} = \frac{\partial l_j}{\partial \boldsymbol{\lambda}} = -\frac{1}{2} \left[\frac{\partial}{\partial \boldsymbol{\lambda}} \log |\boldsymbol{\Sigma}| + \frac{\partial}{\partial \boldsymbol{\lambda}} \|T_j\|^2 \right]$$

$$\Delta_{\phi_j} = \frac{\partial l_j}{\partial \boldsymbol{\phi}} = -\frac{1}{2} \left[\frac{\partial}{\partial \boldsymbol{\phi}} \log |\boldsymbol{\Sigma}| + \frac{\partial}{\partial \boldsymbol{\phi}} \|T_j\|^2 \right], \quad j = 1, \dots, n.$$

Assim, temos que Δ_{μ_j} , Δ_{λ_j} e Δ_{ϕ_j} , dependem das derivadas de $\log |\boldsymbol{\Sigma}|$ e $\|T_j\|^2$ com respeito a $\boldsymbol{\mu}$, $\boldsymbol{\lambda}$ e $\boldsymbol{\phi}$, respectivamente. Estas derivadas foram dadas na seção 3.4.2 e são:

$$\begin{aligned}\Delta_{\mu_j} &= \boldsymbol{\Sigma}^{-1} \mathbf{W}_j \\ \Delta_{\lambda_j} &= -c^{-1} \left[D^{-1}(\boldsymbol{\phi}) \boldsymbol{\lambda} - c^{-1} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j D^{-1}(\boldsymbol{\phi}) \boldsymbol{\lambda} + D^{-1}(\boldsymbol{\phi}) \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\boldsymbol{\phi}) \boldsymbol{\lambda} \right] \\ \Delta_{\phi_j} &= -\frac{1}{2} \left\{ -c^{-1} D^{-2}(\boldsymbol{\phi}) D(\boldsymbol{\lambda}) \boldsymbol{\lambda} + D^{-1}(\boldsymbol{\phi}) \mathbf{1} - D^{-2}(\boldsymbol{\phi}) D(\mathbf{W}_j) \mathbf{W}_j \right. \\ &\quad \left. - c^{-2} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j D(\boldsymbol{\lambda}) D^{-2}(\boldsymbol{\phi}) \boldsymbol{\lambda} + 2c^{-1} D(\boldsymbol{\lambda}) D^{-2}(\boldsymbol{\phi}) \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\boldsymbol{\phi}) \boldsymbol{\lambda} \right\},\end{aligned}$$

onde as derivadas são avaliadas em $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\mu}}^T, \hat{\boldsymbol{\lambda}}^T, \hat{\boldsymbol{\phi}}^T)^T$ e $\boldsymbol{\omega}_0 = (1, \dots, 1)^T$. Observamos que, para esse esquema de perturbação, a matriz Δ não depende do vetor $\boldsymbol{\omega}$.

O vetor $\mathbf{d}_{\max}$ é então obtido, é o autovetor normalizado correspondente ao maior autovalor da matriz $\Delta^T \tilde{\mathbf{L}}^{-1} \Delta$. Um gráfico de $|\mathbf{d}_{\max}|$ versus a ordem das observações pode revelar quais são os casos que exercem mais influência sobre as estimativas dos parâmetros.

4.3 Perturbação Aditiva

Consideremos o modelo sob a perturbação aditiva dado por

$$\mathbf{Y}_{w_j} = \mathbf{Y}_j + w_{kj} \mathbf{e}_k, \quad (4.8)$$

onde $\mathbf{e}_k$ é o k -ésimo vetor unitário de $\mathbf{R}^3$, $k = 1, 2, 3$ e $j = 1, \dots, n$. Neste caso, o modelo não perturbado é obtido ao considerar $\boldsymbol{\omega}_0 = (0, \dots, 0)^T$, o vetor nulo n -dimensional. A função de log-verossimilhança para o modelo perturbado $l(\boldsymbol{\theta}/\boldsymbol{\omega})$ é dado por (4.3), com $\mathbf{Y}_{w_j} = \mathbf{Y}_j + w_{kj} \mathbf{e}_k$ e $\boldsymbol{\omega} = (\omega_{kj}, \dots, \omega_{kn})^T$. De (4.4), temos que a derivada $l(\boldsymbol{\theta}/\boldsymbol{\omega})$ depende das funções $A_j(\boldsymbol{\theta}, \omega_{kj})$, $j = 1, \dots, n$, isto é,

$$\frac{\partial l(\boldsymbol{\theta}/\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^T} = (A_1(\boldsymbol{\theta}, w_{k1}), \dots, A_n(\boldsymbol{\theta}, w_{kn})),$$

onde

$$A_j(\boldsymbol{\theta}, w_{kj}) = -\frac{1}{2} \frac{\partial}{\partial w_{kj}} \|T_{w_{kj}}\|^2 = -\frac{1}{2} \frac{\partial}{\partial w_{kj}} [(\mathbf{Y}_{w_j} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{Y}_{w_j} - \boldsymbol{\mu})]$$

$$= -\frac{1}{2} \frac{\partial}{\partial w_{kj}} (\mathbf{Y}_{w_j} - \boldsymbol{\mu})^T \frac{\partial [(\mathbf{Y}_{w_j} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{Y}_{w_j} - \boldsymbol{\mu})]}{\partial (\mathbf{Y}_{w_j} - \boldsymbol{\mu})} \quad \text{Pr op. A2}$$

$$= -\frac{1}{2} \frac{\partial}{\partial w_{kj}} (\mathbf{e}_k^T w_{kj}) 2\boldsymbol{\Sigma}^{-1} (\mathbf{Y}_{w_j} - \boldsymbol{\mu}) \quad \text{Pr op. A1 : b}$$

$$= -\underbrace{\mathbf{e}_k^T \boldsymbol{\Sigma}^{-1} (\mathbf{Y}_{w_j} - \boldsymbol{\mu})}_{\text{número}}$$

$$= -(\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} \mathbf{e}_k.$$

Por outro lado, a matriz Δ dada em (4.5), depende das matrizes

$$\begin{aligned}\Delta_{\mu_j} &= \frac{\partial}{\partial \mu} A_j(\boldsymbol{\theta}, w_{kj}), \\ \Delta_{\lambda_j} &= \frac{\partial}{\partial \lambda} A_j(\boldsymbol{\theta}, w_{kj})\end{aligned}$$

e

$$\Delta_{\phi_j} = \frac{\partial}{\partial \phi} A_j(\boldsymbol{\theta}, w_{kj}),$$

que serão dadas no seguinte Lema.

Lema 4.1: *Seja $l_j(\boldsymbol{\theta}/w_{kj})$ como em (4.3). Então*

$$1) \Delta_{\mu_j} = \boldsymbol{\Sigma}^{-1} \mathbf{e}_k$$

$$\begin{aligned}2) \Delta_{\lambda_j} = c^{-1} \Big\{ &-2c^{-1} D^{-1}(\phi) \boldsymbol{\lambda} (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \mathbf{M} + D^{-1}(\phi) (\mathbf{Y}_{wj} - \boldsymbol{\mu}) \boldsymbol{\lambda}^T D^{-1}(\phi) \\ &+ \boldsymbol{\lambda}^T D^{-1}(\phi) (\mathbf{Y}_{wj} - \boldsymbol{\mu}) D^{-1}(\phi) \Big\} \mathbf{e}_k\end{aligned}$$

$$\begin{aligned}3) \Delta_{\phi_j} = \Big\{ &D(\mathbf{Y}_{wj} - \boldsymbol{\mu}) D^{-2}(\phi) + c^{-2} D(\boldsymbol{\lambda}) D^{-2}(\phi) \boldsymbol{\lambda} (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \mathbf{M} \\ &- c^{-1} D(\mathbf{Y}_{wj} - \boldsymbol{\mu}) D^{-1}(\phi) \mathbf{M} - c^{-1} (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T D^{-1}(\phi) \boldsymbol{\lambda} D(\boldsymbol{\lambda}) D^{-2}(\phi) \Big\} \mathbf{e}_k.\end{aligned}$$

Prova:

(1) A prova é imediata.

(2)

$$\begin{aligned}\Delta_{\lambda_j} &= \frac{\partial}{\partial \boldsymbol{\lambda}} A_j(\boldsymbol{\theta}, w_{kj}) = \frac{\partial}{\partial \boldsymbol{\lambda}} \left(-(\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} \mathbf{e}_k \right) \\ &= -\frac{\partial}{\partial \boldsymbol{\lambda}} \left[(\mathbf{Y}_{wj} - \boldsymbol{\mu})^T (D^{-1}(\phi) - c^{-1} \mathbf{M}) \mathbf{e}_k \right]\end{aligned}$$

$$\begin{aligned}
&= -\frac{\partial}{\partial \boldsymbol{\lambda}} \left[(\mathbf{Y}_{wj} - \boldsymbol{\mu})^T D^{-1}(\boldsymbol{\phi}) \mathbf{e}_k - c^{-1} (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \mathbf{M} \mathbf{e}_k \right] \\
&= \frac{\partial}{\partial \boldsymbol{\lambda}} \left(c^{-1} (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \mathbf{M} \mathbf{e}_k \right) \\
&= -c^{-2} \frac{\partial}{\partial \boldsymbol{\lambda}} (c) (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \mathbf{M} \mathbf{e}_k + c^{-1} \frac{\partial}{\partial \boldsymbol{\lambda}} (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \mathbf{M} \mathbf{e}_k \\
&= -c^{-2} (2D^{-1}(\boldsymbol{\phi}) \boldsymbol{\lambda}) (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \mathbf{M} \mathbf{e}_k \quad \text{Lema A1, 1} \\
&\quad + c^{-1} [D^{-1}(\boldsymbol{\phi}) (\mathbf{Y}_{wj} - \boldsymbol{\mu}) \boldsymbol{\lambda}^T D^{-1}(\boldsymbol{\phi}) + \boldsymbol{\lambda}^T D^{-1}(\boldsymbol{\phi}) (\mathbf{Y}_{wj} - \boldsymbol{\mu}) D^{-1}(\boldsymbol{\phi}) \mathbf{e}_k] \quad \text{Lema A2, 4.} \\
&= c^{-1} \left\{ -2c^{-1} D^{-1}(\boldsymbol{\phi}) \boldsymbol{\lambda} (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \mathbf{M} + D^{-1}(\boldsymbol{\phi}) (\mathbf{Y}_{wj} - \boldsymbol{\mu}) \boldsymbol{\lambda}^T D^{-1}(\boldsymbol{\phi}) \right. \\
&\quad \left. + \boldsymbol{\lambda}^T D^{-1}(\boldsymbol{\phi}) (\mathbf{Y}_{wj} - \boldsymbol{\mu}) D^{-1}(\boldsymbol{\phi}) \right\} \mathbf{e}_k
\end{aligned}$$

(3)

$$\begin{aligned}
\Delta_{\phi_j} &= \frac{\partial}{\partial \boldsymbol{\phi}} A_j(\boldsymbol{\theta}, w_{kj}) = \frac{\partial}{\partial \boldsymbol{\phi}} \left(-(\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} \mathbf{e}_k \right) \\
&= -\frac{\partial}{\partial \boldsymbol{\phi}} \left[(\mathbf{Y}_{wj} - \boldsymbol{\mu})^T (D^{-1}(\boldsymbol{\phi}) - c^{-1} \mathbf{M}) \mathbf{e}_k \right] \\
&= -\frac{\partial}{\partial \boldsymbol{\phi}} \left[(\mathbf{Y}_{wj} - \boldsymbol{\mu})^T D^{-1}(\boldsymbol{\phi}) \mathbf{e}_k - c^{-1} (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \mathbf{M} \mathbf{e}_k \right] \\
&= -\left\{ \frac{\partial}{\partial \boldsymbol{\phi}} (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T D^{-1}(\boldsymbol{\phi}) \mathbf{e}_k + c^{-2} \frac{\partial}{\partial \boldsymbol{\phi}} (c) (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \mathbf{M} \mathbf{e}_k \right. \\
&\quad \left. - c^{-1} \frac{\partial}{\partial \boldsymbol{\phi}} (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \mathbf{M} \mathbf{e}_k \right\} \\
&= D(\mathbf{Y}_{wj} - \boldsymbol{\mu}) D^{-2}(\boldsymbol{\phi}) \mathbf{e}_k - c^{-2} (-D(\boldsymbol{\lambda}) D^{-2}(\boldsymbol{\phi}) \boldsymbol{\lambda}) (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \mathbf{M} \mathbf{e}_k \quad \text{Lema A1 e A2} \\
&\quad + c^{-1} \left[-D(\mathbf{Y}_{wj} - \boldsymbol{\mu}) D^{-2}(\boldsymbol{\phi}) \boldsymbol{\lambda} \boldsymbol{\lambda}^T D^{-1}(\boldsymbol{\phi}) \mathbf{e}_k - (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T D^{-1}(\boldsymbol{\phi}) \boldsymbol{\lambda} D(\boldsymbol{\lambda}) D^{-2}(\boldsymbol{\phi}) \mathbf{e}_k \right] \\
&= D(\mathbf{Y}_{wj} - \boldsymbol{\mu}) D^{-2}(\boldsymbol{\phi}) \mathbf{e}_k + c^{-2} D(\boldsymbol{\lambda}) D^{-2}(\boldsymbol{\phi}) \boldsymbol{\lambda} (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \mathbf{M} \mathbf{e}_k \\
&\quad - c^{-1} D(\mathbf{Y}_{wj} - \boldsymbol{\mu}) D^{-2}(\boldsymbol{\phi}) \boldsymbol{\lambda} \boldsymbol{\lambda}^T D^{-1}(\boldsymbol{\phi}) \mathbf{e}_k - c^{-1} (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T D^{-1}(\boldsymbol{\phi}) \boldsymbol{\lambda} D(\boldsymbol{\lambda}) D^{-2}(\boldsymbol{\phi}) \mathbf{e}_k \\
&= \left\{ D(\mathbf{Y}_{wj} - \boldsymbol{\mu}) D^{-2}(\boldsymbol{\phi}) + c^{-2} D(\boldsymbol{\lambda}) D^{-2}(\boldsymbol{\phi}) \boldsymbol{\lambda} (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \mathbf{M} \right. \\
&\quad \left. - c^{-1} D(\mathbf{Y}_{wj} - \boldsymbol{\mu}) D^{-1}(\boldsymbol{\phi}) \mathbf{M} - c^{-1} (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T D^{-1}(\boldsymbol{\phi}) \boldsymbol{\lambda} D(\boldsymbol{\lambda}) D^{-2}(\boldsymbol{\phi}) \right\} \mathbf{e}_k
\end{aligned}$$

Dos resultados do Lema 3.1, obtemos a matriz $\Delta = \frac{\partial l(\theta/\omega)}{\partial \theta \partial \sigma \omega^T}$, fica dada por

$$\Delta = \begin{pmatrix} \Delta_{\mu_1} & \cdots & \Delta_{\mu_n} \\ \Delta_{\lambda_1} & \cdots & \Delta_{\lambda_n} \\ \Delta_{\phi_1} & \cdots & \Delta_{\phi_n} \end{pmatrix}_{9 \times n},$$

onde as derivadas são avaliadas em $\theta = \hat{\theta} = (\hat{\mu}^T, \hat{\lambda}^T, \hat{\phi}^T)^T$ e em $\omega = \omega_0$.

Segundo o método de Cook(1986), o vetor normalizado correspondente ao maior autovalor da matriz $\Delta^T \tilde{L}^{-1} \Delta$ fornece a direção de maior influência local. Este vetor é denotado por $d_{\max}$. Esse vetor contém informação sobre quais são as observações que exercem maior influência sobre as estimativas dos parâmetros.

4.4 Perturbação Multiplicativa

Consideremos o modelo sob a perturbação multiplicativa dado por

$$Y_{w_j} = Y_j * \mathbf{1}_3(w_{kj}),$$

onde e_k é o k -ésimo vetor unitário de $\mathbf{R}^3$ e $\mathbf{1}_3(w_{kj})$ é o vetor em $\mathbf{R}^3$ com w_{kj} na k -ésima posição, $k = 1, 2, 3$ e 1's em outros casos $j = 1, \dots, n$. Neste caso, o vetor correspondente à não perturbação é dado por $\omega_0 = 1$. Sob este esquema de perturbação, a função de log-verossimilhança $l(\theta/\omega)$ é dada como em (4.3), com $Y_{\omega_j} = Y_j * \mathbf{1}_3(w_{kj})$. Derivando $l(\theta/\omega)$ com respeito a ω^T , obtemos

$$\frac{\partial l(\theta/\omega)}{\partial \omega^T} = (A_1(\theta, w_{k1}), \dots, A_n(\theta, w_{kn})),$$

onde

$$\begin{aligned}
A_j(\boldsymbol{\theta}, w_{kj}) &= -\frac{1}{2} \frac{\partial}{\partial w_{kj}} \|T_{w_{kj}}\|^2 = -\frac{1}{2} \frac{\partial}{\partial w_{kj}} \left[(\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{Y}_{wj} - \boldsymbol{\mu}) \right] \\
&= -\frac{1}{2} \frac{\partial}{\partial w_{kj}} (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \frac{\partial \left[(\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{Y}_{wj} - \boldsymbol{\mu}) \right]}{\partial (\mathbf{Y}_{wj} - \boldsymbol{\mu})}, \quad \text{Prop. A2} \\
&= -\frac{1}{2} \frac{\partial}{\partial w_{kj}} (\mathbf{Y}_j * \mathbf{1}_3(w_{kj}))^T 2\boldsymbol{\Sigma}^{-1} (\mathbf{Y}_{wj} - \boldsymbol{\mu}), \quad \text{Prop. A1} \\
&= -Y_{kj} (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} \mathbf{e}_k.
\end{aligned}$$

De (4.5), temos que a matriz Δ dependem das matrizes

$$\begin{aligned}
\Delta_{\mu_j} &= \frac{\partial}{\partial \boldsymbol{\mu}} A_j(\boldsymbol{\theta}, w_{kj}) \\
\Delta_{\lambda_j} &= \frac{\partial}{\partial \boldsymbol{\lambda}} A_j(\boldsymbol{\theta}, w_{kj}) \\
\Delta_{\phi_j} &= \frac{\partial}{\partial \boldsymbol{\phi}} A_j(\boldsymbol{\theta}, w_{kj}),
\end{aligned}$$

que serão dadas no seguinte Lema:

Lema 4.2: *Seja $l_j(\boldsymbol{\theta}/\omega_j)$ como em (4.3). Então*

$$\begin{aligned}
1) \quad \Delta_{\mu_j} &= Y_{kj} \boldsymbol{\Sigma}^{-1} \mathbf{e}_k \\
2) \quad \Delta_{\lambda_j} &= Y_{kj} c^{-1} \left\{ -2c^{-1} D^{-1}(\boldsymbol{\phi}) \boldsymbol{\lambda} (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \mathbf{M} + D^{-1}(\boldsymbol{\phi}) (\mathbf{Y}_{wj} - \boldsymbol{\mu}) \boldsymbol{\lambda}^T D^{-1}(\boldsymbol{\phi}) \right. \\
&\quad \left. + \boldsymbol{\lambda}^T D^{-1}(\boldsymbol{\phi}) (\mathbf{Y}_{wj} - \boldsymbol{\mu}) D^{-1}(\boldsymbol{\phi}) \right\} \mathbf{e}_k \\
3) \quad \Delta_{\phi_j} &= Y_{kj} \left[D(\mathbf{Y}_{wj} - \boldsymbol{\mu}) D^{-2}(\boldsymbol{\phi}) + c^{-2} D(\boldsymbol{\lambda}) D^{-2}(\boldsymbol{\phi}) \boldsymbol{\lambda} (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \mathbf{M} \right. \\
&\quad \left. - c^{-1} D(\mathbf{Y}_{wj} - \boldsymbol{\mu}) D^{-1}(\boldsymbol{\phi}) \mathbf{M} - c^{-1} (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T D^{-1}(\boldsymbol{\phi}) \boldsymbol{\lambda} D(\boldsymbol{\lambda}) D^{-2}(\boldsymbol{\phi}) \right] \mathbf{e}_k.
\end{aligned}$$

Prova:

(1) A prova é imediata.

(2)

$$\begin{aligned}
\Delta_{\lambda_j} &= \frac{\partial}{\partial \lambda} A_j(\boldsymbol{\theta}, w_{kj}) = \frac{\partial}{\partial \lambda} \left(-Y_{kj} (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} \mathbf{e}_k \right) \\
&= \frac{\partial}{\partial \lambda} \left[-Y_{kj} (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T (D^{-1}(\phi) - c^{-1} \mathbf{M}) \mathbf{e}_k \right] \\
&= -\frac{\partial}{\partial \lambda} \left[Y_{kj} (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T D^{-1}(\phi) \mathbf{e}_k - c^{-1} Y_{kj} (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \mathbf{M} \mathbf{e}_k \right] \\
&= Y_{kj} \frac{\partial}{\partial \lambda} \left(c^{-1} (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \mathbf{M} \mathbf{e}_k \right) \\
&= Y_{kj} (-c^{-2}) \frac{\partial}{\partial \lambda} (c) (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \mathbf{M} \mathbf{e}_k + Y_{kj} c^{-1} \frac{\partial}{\partial \lambda} (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \mathbf{M} \mathbf{e}_k \\
&= -c^{-2} (2D^{-1}(\phi) \boldsymbol{\lambda}) Y_{kj} (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \mathbf{M} \mathbf{e}_k + Y_{kj} c^{-1} \left[\frac{\partial}{\partial \lambda} (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \mathbf{M} \mathbf{e}_k \right] \quad \text{Lema B2 : 2} \\
&= -2c^{-2} Y_{kj} D^{-1}(\phi) \boldsymbol{\lambda} (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \mathbf{M} \mathbf{e}_k \quad \text{Lema B2 : 4} \\
&\quad + c^{-1} Y_{kj} \left[D^{-1}(\phi) (\mathbf{Y}_{wj} - \boldsymbol{\mu}) \boldsymbol{\lambda}^T D^{-1}(\phi) + \boldsymbol{\lambda}^T D^{-1}(\phi) (\mathbf{Y}_{wj} - \boldsymbol{\mu}) D^{-1}(\phi) \right] \mathbf{e}_k \\
&= Y_{kj} c^{-1} \left\{ -2c^{-1} D^{-1}(\phi) \boldsymbol{\lambda} (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \mathbf{M} + D^{-1}(\phi) (\mathbf{Y}_{wj} - \boldsymbol{\mu}) \boldsymbol{\lambda}^T D^{-1}(\phi) \right. \\
&\quad \left. + \boldsymbol{\lambda}^T D^{-1}(\phi) (\mathbf{Y}_{wj} - \boldsymbol{\mu}) D^{-1}(\phi) \right\} \mathbf{e}_k
\end{aligned}$$

(3)

$$\begin{aligned}
\Delta_{\phi_j} &= \frac{\partial}{\partial \phi} A_j(\boldsymbol{\theta}, w_{kj}) = \frac{\partial}{\partial \phi} \left(-Y_{kj} (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} \mathbf{e}_k \right) \\
&= -\frac{\partial}{\partial \phi} \left[Y_{kj} (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T (D^{-1}(\phi) - c^{-1} \mathbf{M}) \mathbf{e}_k \right] \\
&= -Y_{kj} \frac{\partial}{\partial \phi} \left[(\mathbf{Y}_{wj} - \boldsymbol{\mu})^T D^{-1}(\phi) \mathbf{e}_k \right] + Y_{kj} \frac{\partial}{\partial \phi} \left[(\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \mathbf{M} \mathbf{e}_k \right] \\
&= -Y_{kj} \left[-D(\mathbf{Y}_{wj} - \boldsymbol{\mu}) D^{-2}(\phi) \right] \mathbf{e}_k + Y_{kj} \frac{\partial}{\partial \phi} (c^{-1}) (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \mathbf{M} \mathbf{e}_k \\
&\quad + Y_{kj} c^{-1} \frac{\partial}{\partial \phi} \left[(\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \mathbf{M} \mathbf{e}_k \right], \quad \text{Lema B2, 2} \\
&= Y_{kj} \left[D(\mathbf{Y}_{wj} - \boldsymbol{\mu}) D^{-2}(\phi) \mathbf{e}_k + c^{-2} D(\boldsymbol{\lambda}) D^{-2}(\phi) \boldsymbol{\lambda} (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \mathbf{M} \mathbf{e}_k, \quad \text{Lema B1, B2} \right. \\
&\quad \left. + c^{-1} \left(-D(\mathbf{Y}_{wj} - \boldsymbol{\mu}) D^{-2}(\phi) \boldsymbol{\lambda} \boldsymbol{\lambda}^T D^{-1}(\phi) - (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T D^{-1}(\phi) \boldsymbol{\lambda} D(\boldsymbol{\lambda}) D^{-2}(\phi) \right) \mathbf{e}_k \right] \\
&= Y_{kj} \left[D(\mathbf{Y}_{wj} - \boldsymbol{\mu}) D^{-2}(\phi) + c^{-2} D(\boldsymbol{\lambda}) D^{-2}(\phi) \boldsymbol{\lambda} (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T \mathbf{M} \right. \\
&\quad \left. - c^{-1} D(\mathbf{Y}_{wj} - \boldsymbol{\mu}) D^{-1}(\phi) \mathbf{M} - c^{-1} (\mathbf{Y}_{wj} - \boldsymbol{\mu})^T D^{-1}(\phi) \boldsymbol{\lambda} D(\boldsymbol{\lambda}) D^{-2}(\phi) \right] \mathbf{e}_k
\end{aligned}$$

Dos resultados no Lema anterior, obtemos a matriz $\Delta = \frac{\partial^2 l(\theta/\omega)}{\partial \theta \partial \omega^T}$, dada por

$$\Delta = \begin{pmatrix} \Delta_{\mu_1} & \cdots & \Delta_{\mu_n} \\ \Delta_{\lambda_1} & \cdots & \Delta_{\lambda_n} \\ \Delta_{\phi_1} & \cdots & \Delta_{\phi_n} \end{pmatrix}_{9 \times n}$$

as derivadas são avaliadas em $\theta = \hat{\theta} = (\hat{\mu}^T, \hat{\lambda}^T, \hat{\phi}^T)^T$ e em $\omega = \omega_0$.

Segundo o método de Cook(1986) o autovetor normalizado $\mathbf{d}_{\max}$ (que corresponde ao maior autovalor da matriz $\Delta^T \tilde{\mathbf{L}}^{-1} \Delta$) indica quais são as observações que exercem maior influência sobre as estimativas dos parâmetros.

4.5 Aplicação 1

Dados de Chipkevitch et al. (1996)

A primeira aplicação que faremos consiste dos dados das medições dos testículos em adolescentes, que foram utilizados na seção 3.5 do capítulo III.

Baseado no modelo de calibração comparativa proposto por Theobald e Mallison (1978), será desenvolvida a aplicação do método de influência local apresentado nas seções anteriores, considerando alguns tipos de perturbações.

Para iniciar esta seção, apresentamos na Figura 4.1 um gráfico das medições dos três métodos nos 42 adolescentes e na Figura 4.2 um gráfico tridimensional dos três métodos. Na Figura 4.1, observa-se que o ponto 38 se destaca do resto nas três medições; além disso, algumas das medidas dos adolescentes 13, 31 e 32 apresentam valores que se destacam das outras medidas. Na Figura 4.2, pode-se observar que o ponto 38 também se destaca do resto das observações.(veja também a Tabela 3.2).

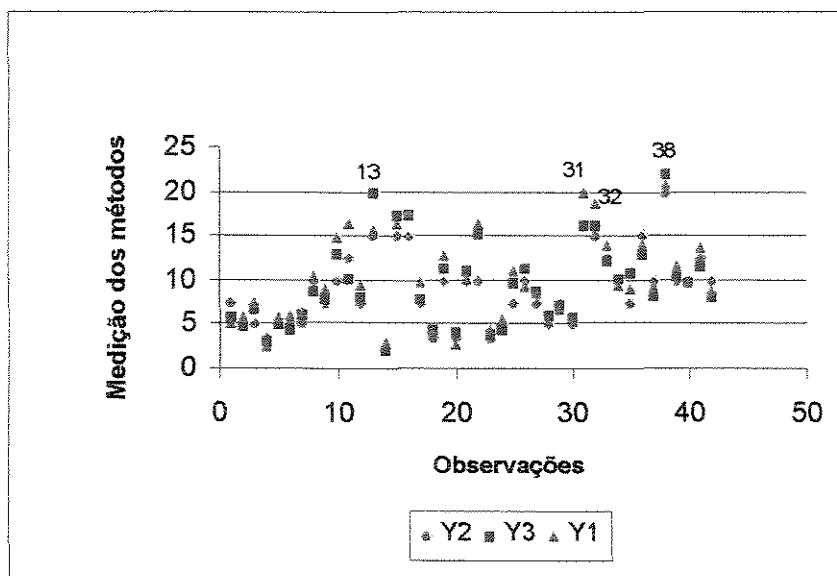


Figura 4-1: Gráfico dos três métodos vs observações

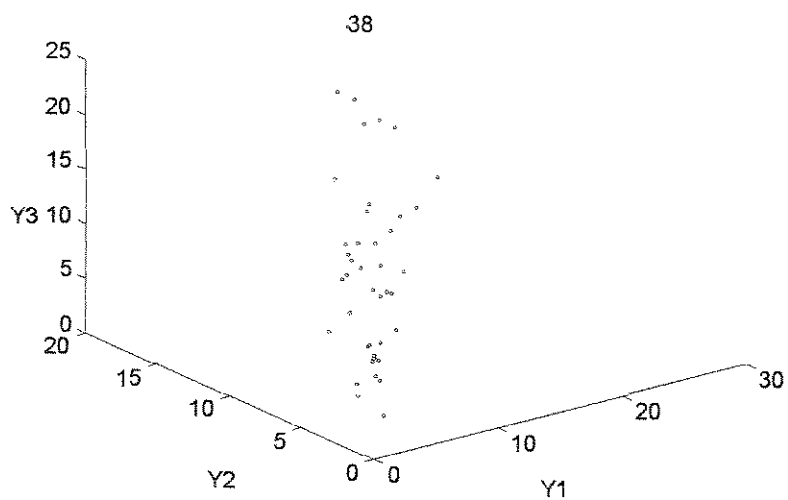


Figura 4-2: Gráfico das medições do testículo direito

4.5.1 Perturbação de Ponderação de Casos

Assumindo o modelo de calibração comparativa baseado na parametrização de análise de fatores proposto por Theobald e Mallison (1978), consideramos o esquema de perturbação de ponderação de casos. Aqui estamos olhando a contribuição de todos os adolescentes na função de log-verossimilhança e observamos, na Figura 4.3, que dois pontos se destacam correspondendo aos adolescentes 31 e 38.

Na Tabela 4.1 apresentamos os elementos em valor absoluto do autovetor, correspondente

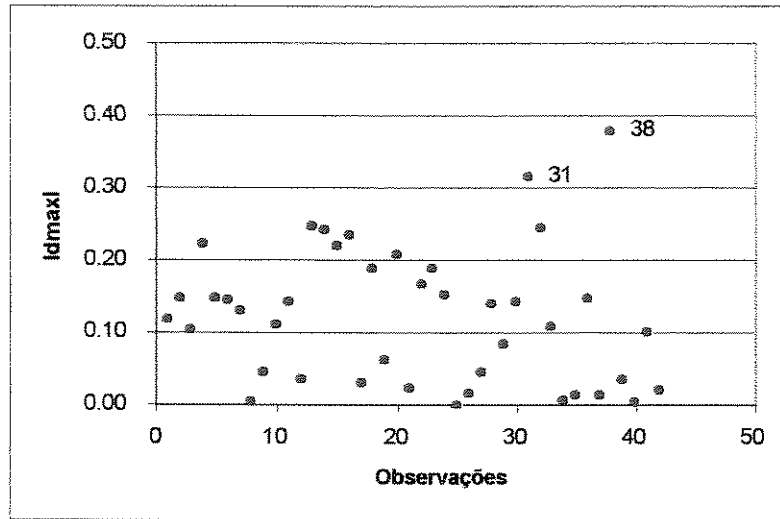


Figura 4-3: Perturbação de ponderação de casos

ao maior autovalor de $\Delta^T (\ddot{L})^{-1} \Delta$, isto é, os valores absolutos dos componentes do vetor $\mathbf{d}_{\max}$.

Na Tabela 4.2 apresentamos a estimativa dos parâmetros; a primeira linha mostra a estimativa dos parâmetros considerando o total dos dados, a segunda linha mostra a estimativa sem considerar as observações 31 e 38.

Tabela 4.1: Elementos de $d_{\max}$ em valor absoluto correspondente às medidas do volume do testículo direito humano em 42 adolescentes, feitas por três métodos de medição.

Obs.	$ d_{\max} $	Obs.	$ d_{\max} $	Obs.	$ d_{\max} $
1	0,1185	15	0,2211	29	0,0863
2	0,1498	16	0,2369	30	0,1439
3	0,1049	17	0,0329	31	0,3176
4	0,2232	18	0,1903	32	0,2465
5	0,1479	19	0,0625	33	0,1108
6	0,1473	20	0,2108	34	0,0067
7	0,1308	21	0,025	35	0,0152
8	0,0047	22	0,1672	36	0,1488
9	0,0473	23	0,1899	37	0,0137
10	0,1131	24	0,1540	38	0,3810
11	0,1448	25	0,0000	39	0,0367
12	0,0368	26	0,0167	40	0,0037
13	0,2487	27	0,0467	41	0,1028
14	0,2438	28	0,1404	42	0,0215

Tabela 4.2: Estimativa dos parâmetros

μ_1	μ_2	μ_3	λ_1	λ_2	λ_3	ϕ_1	ϕ_2	ϕ_3	ρ_1	ρ_2	ρ_3
10,081	9,012	9,562	4,737	4,154	4,495	1,149	1,452	1,526	0,951	0,923	0,930
9,563	8,463	9,080	4,220	3,480	4,040	1,296	1,208	1,207	0,932	0,909	0,931

Intervalos de confiança para os θ_i com um nível do 95% da segunda linha da Tabela 4.2

$$\mu_1 : [8,241; 10,885], \mu_2 : [7,359; 9,567], \mu_3 : [7,813; 10,346]$$

$$\lambda_1 : [3,250; 5,190], \lambda_2 : [2,662; 4,298], \lambda_3 : [3,110; 4,970]$$

$$\phi_1 : [0,354; 2,238], \phi_2 : [0,476; 1,940], \phi_3 : [0,339; 2,075]$$

$$\rho_1 : [0,754; 1,110], \rho_2 : [0,843; 0,975], \rho_3 : [0,874; 0,988]$$

4.5.2 Perturbação Aditiva

a) Perturbação Aditiva na medição do método 1

Assumindo o modelo de calibração comparativa para três instrumentos baseado no modelo proposto por Theobald e Mallison (1978), consideramos a perturbação aditiva na medição do método 1; neste caso estamos perturbando a medição fornecida pelo método 1. Na Figura 4.4 apresentamos um gráfico mostrando a influência dos valores observados sobre a estimativa dos parâmetros. Nesse gráfico pode-se observar que os pontos 4, 14 e 20 se destacam do resto. Na Tabela 4.3 apresentamos o autovetor (em valor absoluto),

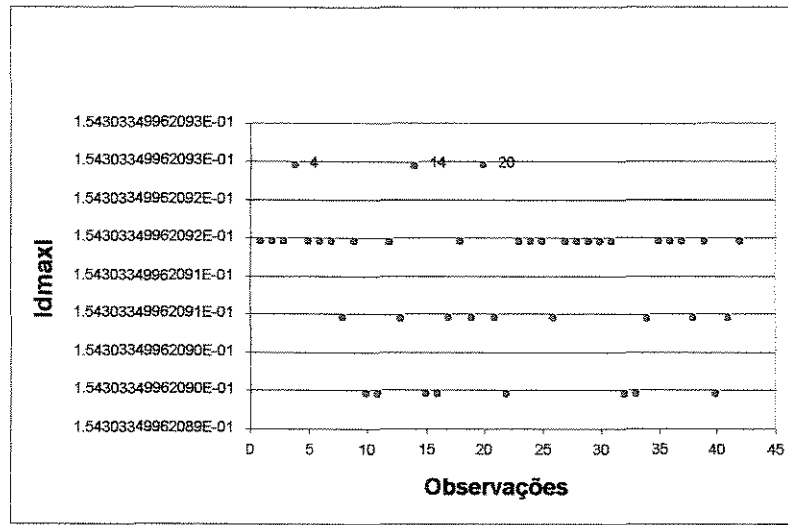


Figura 4-4: Perturbação aditiva no método 1

correspondente ao maior autovalor de $\Delta^T (\hat{\mathbf{L}})^{-1} \Delta$, isto é, os valores absolutos dos componentes do vetor $\mathbf{d}_{\max}$.

Na Tabela 4.4., apresentamos a estimativa dos parâmetros, a primeira linha mostra as estimativas dos parâmetros considerando o total dos dados, a segunda linha mostra as estimativas dos parâmetros sem considerar as medições das observações 4, 14 e 20.

Tabela 4.3: Elementos de $d_{\max}$ em valor absoluto correspondente às medidas do volume do testículo direito humano em 42 adolescentes, feitas por três métodos de medição em perturbação aditiva no método 1.

Obs.	$ d_{\max} $	Obs.	$ d_{\max} $	Obs.	$ d_{\max} $
1	$1,543033499620926e^{-1}$	15	$1,543033499620908e^{-1}$	29	$1,543033499620926e^{-1}$
2	$1,543033499620928e^{-1}$	16	$1,543033499620904e^{-1}$	30	$1,543033499620928e^{-1}$
3	$1,543033499620923e^{-1}$	17	$1,543033499620918e^{-1}$	31	$1,543033499620900e^{-1}$
4	$1,543033499620934e^{-1}$	18	$1,543033499620929e^{-1}$	32	$1,543033499620901e^{-1}$
5	$1,543033499620928e^{-1}$	19	$1,543033499620913e^{-1}$	33	$1,543033499620912e^{-1}$
6	$1,543033499620926e^{-1}$	20	$1,543033499620932e^{-1}$	34	$1,543033499620921e^{-1}$
7	$1,543033499620926e^{-1}$	21	$1,543033499620919e^{-1}$	35	$1,543033499620921e^{-1}$
8	$1,543033499620918e^{-1}$	22	$1,543033499620906e^{-1}$	36	$1,543033499620912e^{-1}$
9	$1,543033499620920e^{-1}$	23	$1,543033499620929e^{-1}$	37	$1,543033499620921e^{-1}$
10	$1,543033499620908e^{-1}$	24	$1,543033499620927e^{-1}$	38	$1,543033499620900e^{-1}$
11	$1,543033499620906e^{-1}$	25	$1,543033499620916e^{-1}$	39	$1,543033499620916e^{-1}$
12	$1,543033499620920e^{-1}$	26	$1,543033499620921e^{-1}$	40	$1,543033499620920e^{-1}$
13	$1,543033499620910e^{-1}$	27	$1,543033499620922e^{-1}$	41	$1,543033499620912e^{-1}$
14	$1,543033499620932e^{-1}$	28	$1,543033499620928e^{-1}$	42	$1,543033499620921e^{-1}$

Tabela 4.4: Estimativa dos parâmetros

μ_1	μ_2	μ_3	λ_1	λ_2	λ_3	ϕ_1	ϕ_2	ϕ_3	ρ_1	ρ_2	ρ_3
10,081	9,012	9,562	4,737	4,154	4,495	1,149	1,452	1,526	0,951	0,923	0,930
10,644	9,474	10,062	4,452	3,944	4,269	1,145	1,562	1,628	0,945	0,909	0,918

Intervalos de confiança para os θ_i com um nível do 95% da segunda linha da Tabela 4.4

$$\mu_1 : [9,259; 12,029], \mu_2 : [8,222; 10,725], \mu_3 : [8,715; 11,409]$$

$$\lambda_1 : [3,442; 5,462], \lambda_2 : [3,016; 4,872], \lambda_3 : [3,274; 5,264]$$

$$\phi_1 : [0,167; 2,123], \phi_2 : [0,620; 2,504], \phi_3 : [0,584; 2,672]$$

$$\rho_1 : [0,893; 0,997], \rho_2 : [0,843; 0,975], \rho_3 : [0,856; 0,980]$$

b) Perturbação Aditiva na medição do método 3

Aqui estamos considerando a perturbação aditiva na medição do método 3; neste caso estamos perturbando a medição fornecida pelo método 3. Na Figura 4.5 apresentamos um gráfico mostrando a influência dos valores observados sobre as estimativas dos parâmetros. Nesse gráfico pode-se observar que os pontos 13, 14 e 38 se destacam do resto.

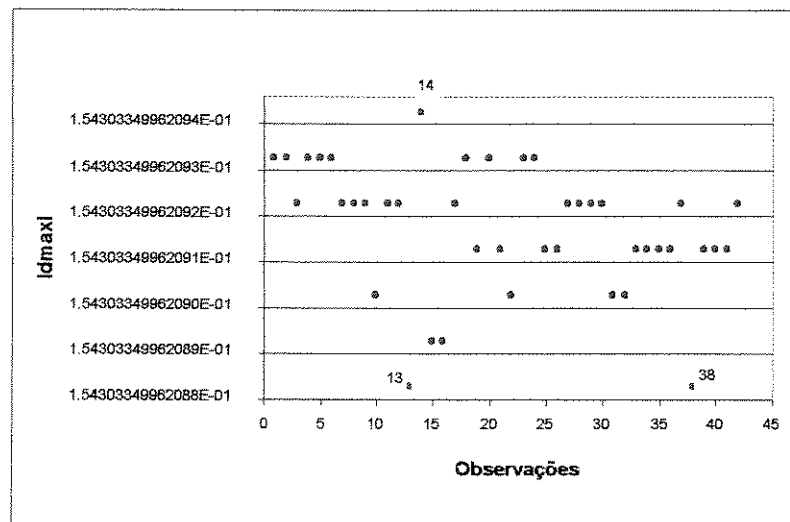


Figura 4-5: Perturbação aditiva no método 3

Na tabela 4.5 apresentamos o autovetor (em valor absoluto), correspondente ao maior autovalor de $\Delta^T \left(\overset{\circ\circ}{\mathbf{L}} \right)^{-1} \Delta$, isto é, os valores absolutos dos componentes do vetor $\mathbf{d}_{\max}$. Na Tabela 4.6, apresentamos as estimativas dos parâmetros, a primeira linha mostra as estimativas dos parâmetros considerando o total dos dados, a segunda linha mostra as estimativas dos parâmetros sem considerar as medições das observações 13, 14 e 38.

Tabela 4.5: Elementos de $\mathbf{d}_{\max}$ em valor absoluto correspondente às medidas do volume do testículo direito humano em 42 adolescentes, feitas por três métodos de medição em perturbação aditiva no método 3.

Obs.	$ \mathbf{d}_{\max} $	Obs.	$ \mathbf{d}_{\max} $	Obs.	$ \mathbf{d}_{\max} $
1	$1,543033499620931\text{e}^{-1}$	15	$1,543033499620895\text{e}^{-1}$	29	$1,543033499620925\text{e}^{-1}$
2	$1,543033499620933\text{e}^{-1}$	16	$1,543033499620895\text{e}^{-1}$	30	$1,543033499620929\text{e}^{-1}$
3	$1,543033499620926\text{e}^{-1}$	17	$1,543033499620926\text{e}^{-1}$	31	$1,543033499620903\text{e}^{-1}$
4	$1,543033499620938\text{e}^{-1}$	18	$1,543033499620933\text{e}^{-1}$	32	$1,543033499620900\text{e}^{-1}$
5	$1,543033499620932\text{e}^{-1}$	19	$1,543033499620913\text{e}^{-1}$	33	$1,543033499620913\text{e}^{-1}$
6	$1,543033499620935\text{e}^{-1}$	20	$1,543033499620933\text{e}^{-1}$	34	$1,543033499620916\text{e}^{-1}$
7	$1,543033499620929\text{e}^{-1}$	21	$1,543033499620913\text{e}^{-1}$	35	$1,543033499620912\text{e}^{-1}$
8	$1,543033499620923\text{e}^{-1}$	22	$1,543033499620901\text{e}^{-1}$	36	$1,543033499620910\text{e}^{-1}$
9	$1,543033499620924\text{e}^{-1}$	23	$1,543033499620935\text{e}^{-1}$	37	$1,543033499620924\text{e}^{-1}$
10	$1,543033499620908\text{e}^{-1}$	24	$1,543033499620935\text{e}^{-1}$	38	$1,543033499620880\text{e}^{-1}$
11	$1,543033499620921\text{e}^{-1}$	25	$1,543033499620918\text{e}^{-1}$	39	$1,543033499620916\text{e}^{-1}$
12	$1,543033499620923\text{e}^{-1}$	26	$1,543033499620911\text{e}^{-1}$	40	$1,543033499620919\text{e}^{-1}$
13	$1,543033499620884\text{e}^{-1}$	27	$1,543033499620920\text{e}^{-1}$	41	$1,543033499620915\text{e}^{-1}$
14	$1,543033499620942\text{e}^{-1}$	28	$1,543033499620928\text{e}^{-1}$	42	$1,543033499620925\text{e}^{-1}$

Tabela 4.6: Estimativa dos parâmetros

μ_1	μ_2	μ_3	λ_1	λ_2	λ_3	ϕ_1	ϕ_2	ϕ_3	ρ_1	ρ_2	ρ_3
10,081	9,012	9,562	4,737	4,154	4,495	1,149	1,452	1,526	0,951	0,923	0,930
9,841	8,756	9,172	4,249	3,485	3,609	0,751	1,601	0,963	0,960	0,883	0,931

Intervalos de confiança para os θ_i com um nível do 95% da segunda linha da Tabela 4.6

$$\mu_1 : [8,530; 11,152], \mu_2 : [7,635; 9,877], \mu_3 : [8,041; 10,303]$$

$$\lambda_1 : [3,299; 5,199], \lambda_2 : [2,644; 4,326], \lambda_3 : [2,779; 4,439]$$

$$\phi_1 : [0,000; 1,557], \phi_2 : [0,754; 2,447], \phi_3 : [0,289; 1,637]$$

$$\rho_1 : [0,914; 1,006], \rho_2 : [0,808; 0,958], \rho_3 : [0,875; 0,987]$$

4.5.3 Perturbação Multiplicativa

a) Perturbação Multiplicativa no método 1

Assumindo o modelo de calibração comparativa para três instrumentos baseado no modelo proposto por Theobald e Mallison (1978), consideramos nesta seção a perturbação multiplicativa, neste caso vamos perturbar a medição fornecida pelo método 1. Na Figura 4.6 apresentamos um gráfico mostrando a influência dos valores observados sobre as estimativas dos parâmetros. Nesse gráfico, observa-se que a influência maior sobre as medições do método 1 está nos adolescentes 31 e 38. Na Tabela 4.7 apresentamos o

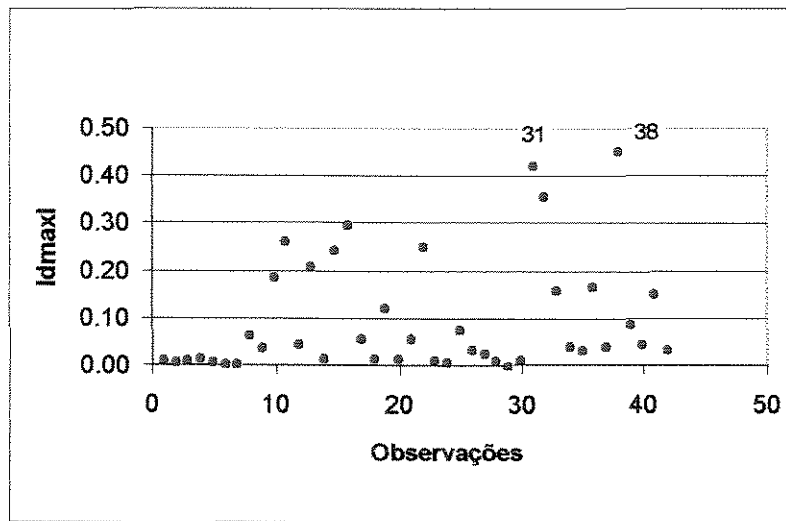


Figura 4-6: Perturbação multiplicativa no método 1

autovetor (em valor absoluto), correspondente ao maior autovalor de $\Delta^T \left(\ddot{\mathbf{L}} \right)^{-1} \Delta$, isto é, os valores absolutos dos componentes do vetor $\mathbf{d}_{\max}$.

Na Tabela 4.8 apresentamos as estimativas dos parâmetros, a primeira linha mostra as estimativas dos parâmetros estimados considerando o total dos dados, a segunda linha mostra as estimativas dos parâmetros estimados sem considerar as medições das observações 31 e 38.

Tabela 4.9: Elementos de $d_{\max}$ em valor absoluto correspondente às medidas do volume do testículo direito humano em 42 adolescentes, feitas por três métodos de medição em perturbação multiplicativa no método 3.

Obs.	$ d_{\max} $	Obs.	$ d_{\max} $	Obs.	$ d_{\max} $
1	0,0033	15	0,3175	29	0,0110
2	0,0169	16	0,2963	30	0,0007
3	0,0043	17	0,0017	31	0,1591
4	0,0147	18	0,0108	32	0,2113
5	0,0146	19	0,0771	33	0,0842
6	0,0224	20	0,0079	34	0,0710
7	0,003	21	0,0936	35	0,1043
8	0,0129	22	0,2196	36	0,1048
9	0,009	23	0,0169	37	0,0119
10	0,1212	24	0,0193	38	0,5721
11	0,0127	25	0,0436	39	0,0597
12	0,0124	26	0,1134	40	0,0468
13	0,5229	27	0,0387	41	0,0620
14	0,0156	28	0,0037	42	0,0076

Tabela 4.10: Estimativas dos parâmetros

μ_1	μ_2	μ_3	λ_1	λ_2	λ_3	ϕ_1	ϕ_2	ϕ_3	ρ_1	ρ_2	ρ_3
10,081	9,012	9,562	4,737	4,154	4,495	1,149	1,452	1,526	0,951	0,923	0,930
9,670	8,590	8,990	4,370	3,630	3,770	0,798	1,596	0,968	0,959	0,892	0,936

Intervalos de confiança para os θ_i com um nível do 95% da segunda linha da Tabela 4.10

$$\mu_1 : [8,321; 11,019], \mu_2 : [7,428; 9,752], \mu_3 : [7,812; 10,168]$$

$$\lambda_1 : [3,394; 5,346], \lambda_2 : [2,762; 4,498], \lambda_3 : [2,908; 4,632]$$

$$\phi_1 : [0,000; 1,608], \phi_2 : [0,746; 2,446], \phi_3 : [0,282; 1,654]$$

$$\rho_1 : [0,915; 1,003], \rho_2 : [0,821; 0,963], \rho_3 : [0,884; 0,987]$$

4.5.4 Outros Resultados de Influência Local

Nesta seção vamos investigar a influência que algumas observações exercem na estimação de um subconjunto do parâmetro $\theta = (\mu^T, \lambda^T, \phi^T)^T$, baseado no esquema de perturbação de ponderação de casos.

a) **Parâmetro de interesse** $\theta_1 = (\mu^T, \lambda^T)^T$.

Vamos supor agora uma partição $\theta = (\theta_1^T, \theta_2^T)^T = ((\mu^T, \lambda^T), \phi^T)^T$, onde $\theta_1 = (\mu^T, \lambda^T)^T$ é de dimensão 6×1 e $\theta_2 = \phi$ é de dimensão 3×1 ; neste caso nosso interesse reside em θ_1 .

Na Figura 4.8 foi empregado o esquema de perturbação que consiste na ponderação de casos. Aqui estamos olhando a contribuição de todos os adolescentes na função de log-verossimilhança e observamos que os adolescentes 11, 31 e 32 se revelam como os mais influentes sobre as estimativas do parâmetro θ_1 .

Na Tabela 4.11 apresentamos o autovetor (em valor absoluto), correspondente ao maior autovalor de $\Delta^T(\hat{\mathbf{L}}^{\bullet\bullet-1} - \mathbf{B}_{22})\Delta$, com $\mathbf{B}_{22}$ definido em (2.23), isto é, os valores absolutos dos componentes do vetor $\mathbf{d}_{\max}$. Na Tabela 4.12 apresentamos as estimativas

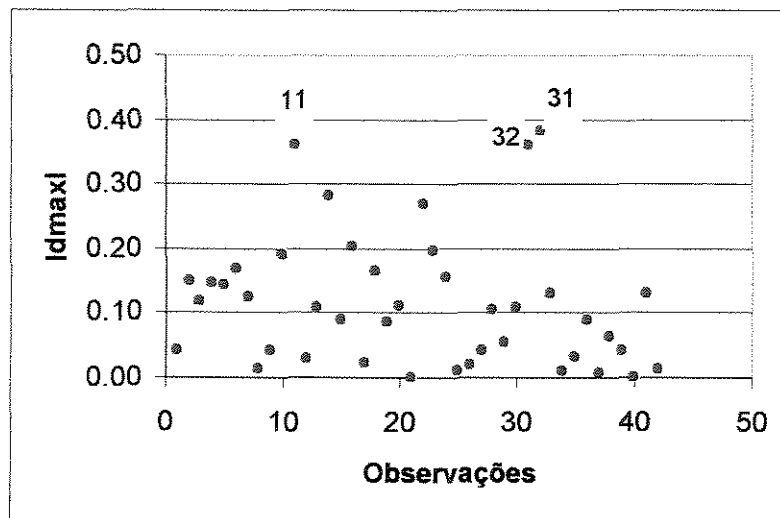


Figura 4-8: Perturbação de ponderação de casos quando o é interesse é $\theta_1 = (\mu^T, \lambda^T)^T$.

dos parâmetros; a primeira linha mostra as estimativas dos parâmetros considerando o total dos dados, a segunda linha mostra as estimativas sem considerar as medições das observações 11, 31 e 32.

Tabela 4.11: Elementos de $\mathbf{d}_{\max}$ em valor absoluto correspondente às medidas do volume do testículo direito humano em 42 adolescentes, feitas por três métodos de medição, quando o parâmetro de interesse é $\boldsymbol{\theta}_1 = (\boldsymbol{\mu}^T, \boldsymbol{\lambda}^T)^T$.

Obs.	d max	Obs.	d max	Obs.	d max
1	0,0438	15	0,0919	29	0,0556
2	0,1518	16	0,2062	30	0,1113
3	0,1212	17	0,0263	31	0,365
4	0,1492	18	0,1666	32	0,3871
5	0,1457	19	0,0897	33	0,1343
6	0,172	20	0,1138	34	0,0117
7	0,1277	21	0,0021	35	0,0349
8	0,0145	22	0,2716	36	0,0913
9	0,0437	23	0,1996	37	0,0102
10	0,1941	24	0,1591	38	0,0658
11	0,3644	25	0,0119	39	0,0439
12	0,0325	26	0,0232	40	0,0035
13	0,1112	27	0,0431	41	0,1339
14	0,2835	28	0,1087	42	0,0169

Tabela 4.12: Estimativas dos parâmetros

μ_1	μ_2	μ_3	λ_1	λ_2	λ_3	ϕ_1	ϕ_2	ϕ_3	ρ_1	ρ_2	ρ_3
10,081	9,012	9,562	4,737	4,154	4,495	1,149	1,452	1,526	0,951	0,923	0,930
9,441	8,487	9,197	4,143	3,621	4,339	0,936	1,322	0,607	0,948	0,908	0,969

Intervalos de confiança para os θ_i com um nível do 95% da segunda linha da Tabela 4.12

$$\mu_1 : [8,154; 10,728], \mu_2 : [7,338; 9,636], \mu_3 : [7,864; 10,530]$$

$$\lambda_1 : [3,208; 5,078], \lambda_2 : [2,770; 4,472], \lambda_3 : [3,379; 5,299]$$

$$\phi_1 : [0,270; 1,602], \phi_2 : [0,626; 2,018], \phi_3 : [0,000; 1,246]$$

$$\rho_1 : [0,905; 0,991], \rho_2 : [0,852; 0,964], \rho_3 : [0,933; 1,005]$$

b) **Parâmetro de interesse $\theta_2 = \phi$.**

Vamos supor agora uma partição $\theta = (\theta_1^T, \theta_2^T)^T = ((\mu^T, \lambda^T), \phi^T)^T$, onde $\theta_1 = (\mu^T, \lambda^T)^T$ é de dimensão 6×1 e $\theta_2 = \phi$ é de dimensão 3×1 ; neste caso nosso interesse reside em θ_2 .

Na Figura 4.9 foi empregado o esquema de perturbação que consiste na ponderação de casos. Aqui estamos olhando a contribuição de todos os adolescentes na função de log-verossimilhança e observamos que os adolescentes 31 e 38 se revelam como os mais influentes sobre as estimativas do parâmetro θ_2 .

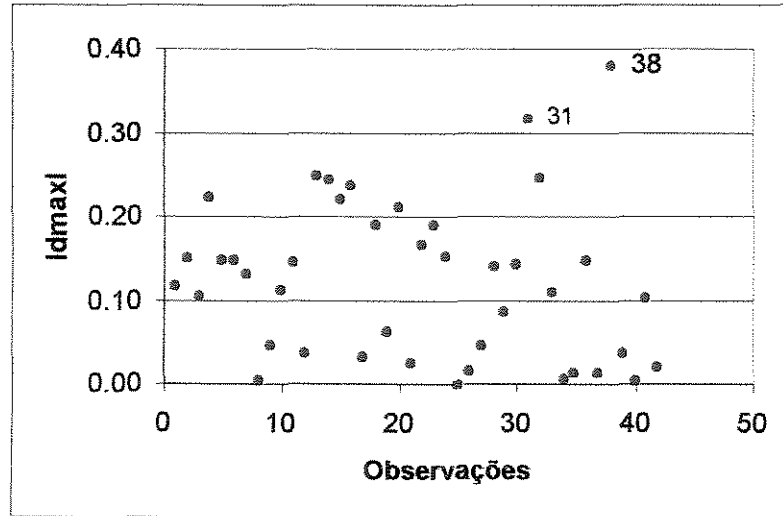


Figura 4-9: Perturbação de ponderação de casos quanto o interesse é $\theta_2 = \phi$.

Tabela 4.13: Elementos de $d_{\max}$ em valor absoluto correspondente às medidas do volume do testículo direito humano em 42 adolescentes, feitas por três métodos de medição, quando o parâmetro de interesse é $\theta_2 = \phi$.

Obs.	$ d_{\max} $	Obs.	$ d_{\max} $	Obs.	$ d_{\max} $
1	0,1184	15	0,2211	29	0,0863
2	0,1498	16	0,2370	30	0,1439
3	0,1049	17	0,0329	31	0,3176
4	0,2231	18	0,1903	32	0,2466
5	0,1479	19	0,0625	33	0,1108
6	0,1473	20	0,2107	34	0,0067
7	0,1308	21	0,0249	35	0,0152
8	0,0047	22	0,1673	36	0,1488
9	0,0473	23	0,1899	37	0,0137
10	0,1131	24	0,154	38	0,3809
11	0,1449	25	0,000	39	0,0367
12	0,0368	26	0,0167	40	0,0037
13	0,2486	27	0,0467	41	0,1028
14	0,2438	28	0,1404	42	0,0215

4.6 Aplicação 2

Dados de Barnett (1969):

A segunda aplicação que apresentamos consiste dos dados tomados de Barnett (1969), relativos a um estudo médico para avaliar a qualidade de dois instrumentos usados para medir a capacidade vital pulmonar x em um grupo de pacientes. Um instrumento é de tipo padrão e outro novo, portátil e mais fácil de operar. Ademais, o instrumento padrão requer um operador especializado. Os dois instrumentos foram comparados em

um grupo comum de $n=72$ pacientes, quando operados por um operador especializado e por um não especializado. As três combinações instrumento-operador serão referidas como (Theobald e Mallison, 1978):

Instrumento 1: Instrumento estandar e operador especializado;

Instrumento 2: Instrumento estandar e operador não-especializado.

Instrumento 3: Instrumento novo e operador especializado.

Seja Y_{ij} a medida da capacidade vital fornecida pelo instrumento i para o paciente j , com $i = 1, 2, 3$ e $j = 1, 2, \dots, 72$. O conjunto de observações $\{Y_{ij}\}$ é apresentado na Tabela 4.14.

Para avaliar o ajuste do modelo normal pode-se considerar um Q-Q plot das distâncias transformadas, como fizemos na primeira aplicação.

$$\begin{aligned}\bar{Y}_1 &= 2246,11, \bar{Y}_2 = 2175,69, \bar{Y}_3 = 2148,61 \\ S_{11} &= 584290,43, S_{22} = 618838,41, S_{33} = 787936,96 \\ S_{12} &= 573377,70, S_{13} = 631330,71, S_{23} = 672753,74\end{aligned}$$

Tabela 4.14 : Medidas da capacidade vital para 72 pacientes, feitas por 3 instrumentos

Obs.	Y ₁	Y ₂	Y ₃	Obs.	Y ₁	Y ₂	Y ₃	Obs.	Y ₁	Y ₂	Y ₃
1	3450	3530	4030	25	2220	1400	1840	49	1000	1130	650
2	1310	1320	1610	26	1880	1820	1900	50	1400	1400	1350
3	3820	3720	4150	27	940	960	1060	51	1880	1710	1600
4	2110	2880	2740	28	2480	2220	2150	52	1280	1260	1160
5	1860	1420	1540	29	1660	1780	1760	53	3120	3000	3110
6	1940	1780	2020	30	4040	4180	4000	54	3770	3340	3900
7	2360	2260	2430	31	2540	2560	2080	55	3420	3220	3120
8	2880	2920	2650	32	1780	1700	1390	56	2740	2880	2850
9	1980	1720	1800	33	1280	1300	800	57	2840	2920	2710
10	3120	3180	3250	34	1940	2060	2030	58	3800	3740	3440
11	1760	1630	1390	35	1760	2000	1860	59	2100	1680	1650
12	1480	1760	1700	36	2040	1660	1470	60	1820	1400	1060
13	1840	1660	1400	37	1060	1000	850	61	1400	1320	1350
14	3580	3480	3680	38	2000	1800	1270	62	2200	1680	1640
15	1880	2000	2090	39	2280	2280	2380	63	1940	1900	1820
16	2400	2320	2550	40	1940	1800	1670	64	3260	3200	3250
17	2220	2120	2290	41	2580	2700	2850	65	1960	1940	1890
18	2540	2500	2620	42	1400	1440	1680	66	1320	1260	1140
19	920	1200	640	43	1260	1100	1000	67	2840	3060	3650
20	2240	2160	2300	44	2320	2420	2360	68	2060	1840	1720
21	2240	2130	2030	45	2000	1940	1980	69	2200	1970	1900
22	2260	2510	2400	46	2400	1900	1470	70	1260	1150	860
23	3860	4180	3980	47	2880	2980	3240	71	3040	2840	2850
24	2780	2100	1890	48	3420	3150	3200	72	2140	2180	2560

Cálculo das estimativas dos parâmetros sob o modelo de calibração estrutural de regressão

De acordo com a subseção 3.3.2., as estimativas de máxima verossimilhança dos parâmetros para o conjunto de dados em estudo, são dados por

$$\hat{\alpha}_2 = -217,792, \hat{\alpha}_3 = -486,788, \hat{\mu}_x = 2246,11$$

$$\hat{\beta}_2 = 1,066, \hat{\beta}_3 = 1,173$$

$$\hat{\sigma}_x^2 = 538073,49, \hat{\sigma}_1^2 = 46216,943$$

$$\hat{\sigma}_2^2 = 7840,107, \hat{\sigma}_3^2 = 47185,975$$

$$\hat{\rho}_1 = 0,921, \hat{\rho}_2 = 0,987, \hat{\rho}_3 = 0,940.$$

Adotamos o instrumento 1 como sendo o instrumento de referência. Isto é, assumimos que $\alpha_1 = 0$ e $\beta_1 = 1$.

Cálculo das estimativas dos parâmetros sob o modelo de calibração estrutural de análise de fatores

Sob a parametrização de análise de fatores, a estimativa de máxima verossimilhança dos parâmetros baseado nos dados de Barnett(1969) são:

$$\hat{\mu}_1 = 2246,11, \hat{\mu}_2 = 2175,69, \hat{\mu}_3 = 2148,61$$

$$\hat{\lambda}_1 = 733,535, \hat{\lambda}_2 = 781,664, \hat{\lambda}_3 = 860,669$$

$$\hat{\phi}_1 = 46216,943, \hat{\phi}_2 = 7840,107, \hat{\phi}_3 = 47185,975$$

$$\hat{\rho}_1 = 0,921, \hat{\rho}_2 = 0,987, \hat{\rho}_3 = 0,940.$$

Na Figura 4.10 apresentamos um gráfico das medições dos instrumentos nos pacientes e na Figura 4.11, é apresentado um gráfico tridimensional dos três métodos; neste gráfico não está claro a existência de “outliers”.

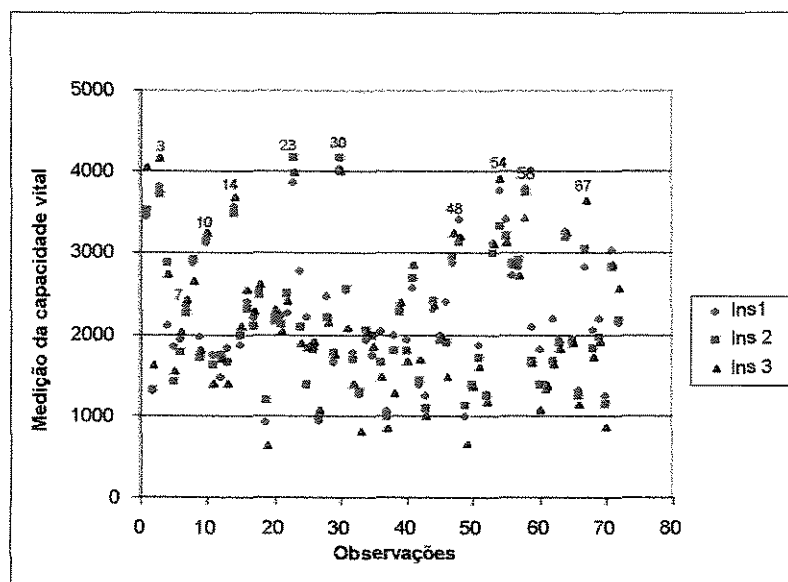


Figura 4-10: Gráfico da capacidade vital para 72 pacientes, feitas por três instrumentos

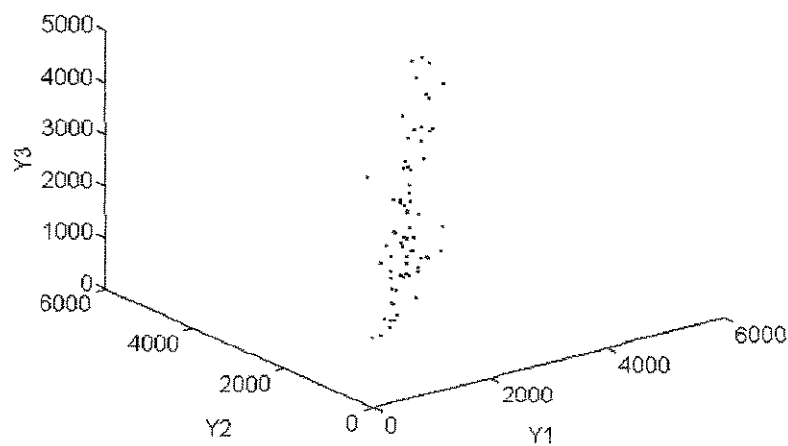


Figura 4-11: Gráfico das medições da capacidade vital para 72 pacientes

4.6.1 Perturbação de Ponderação de Casos

Considerando o modelo de calibração comparativa baseado na parametrização de análise de fatores proposto por Theobald e Mallison (1978), e o esquema de perturbação de ponderação de casos, apresentamos na Figura 4.12 a influência local das observações sobre a estimativa dos parâmetros, observamos que os pontos 23 e 30 se destacam do resto. Na Tabela 4.15 apresentamos as estimativas dos parâmetros; a primeira linha mostra as estimativas dos parâmetros considerando o total dos dados, a segunda linha mostra as estimativas sem considerar as medições das observações 23 e 30. Se observa uma mudança considerável no parâmetro ϕ .

Tabela 4.15: Estimativas dos parâmetros

μ_1	μ_2	μ_3	λ_1	λ_2	λ_3	ϕ_1	ϕ_2	ϕ_3
2246,11	2175,69	2148,61	733,535	781,664	860,669	46216,943	7840,107	47185,975
2197,428	2118,428	2096,00	683,576	715,937	815,035	48157,272	5895,915	46519,235

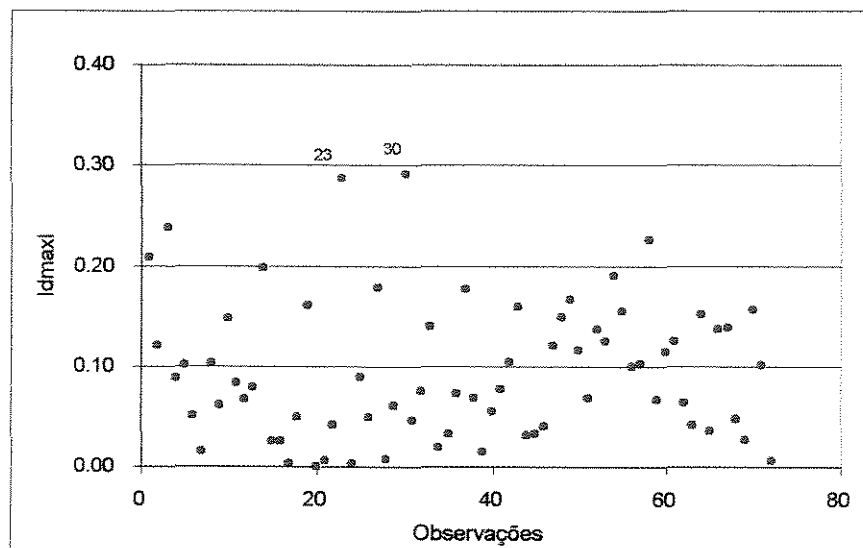


Figura 4-12: Perturbação de ponderação de casos

Tabela 4.16: Elementos de $d_{\max}$ em valor absoluto correspondente aos dados Barnett(1969), em perturbação de casos.

Obs.	$ d_{\max} $	Obs.	$ d_{\max} $	Obs.	$ d_{\max} $
1	0,2097	25	0,0908	49	0,1686
2	0,1224	26	0,0512	50	0,1175
3	0,2397	27	0,1803	51	0,0695
4	0,0891	28	0,0089	52	0,139
5	0,1028	29	0,0624	53	0,1263
6	0,0519	30	0,2924	54	0,1922
7	0,0170	31	0,0469	55	0,1563
8	0,1042	32	0,0766	56	0,1011
9	0,0624	33	0,1417	57	0,1047
10	0,1492	34	0,0208	58	0,2277
11	0,0848	35	0,034	59	0,0677
12	0,0689	36	0,0748	60	0,1159
13	0,0798	37	0,1781	61	0,1263
14	0,1993	38	0,0693	62	0,066
15	0,0273	39	0,0167	63	0,0429
16	0,0268	40	0,0570	64	0,1538
17	0,0039	41	0,0782	65	0,0367
18	0,0506	42	0,1062	66	0,1387
19	0,1621	43	0,1604	67	0,1392
20	0,0011	44	0,0326	68	0,0494
21	0,0078	45	0,0341	69	0,0289
22	0,0425	46	0,0417	70	0,1577
23	0,2892	47	0,1225	71	0,1019
24	0,0041	48	0,1501	72	0,0070

4.6.2 Perturbação Aditiva

Considerando a perturbação aditiva na medição do método 1. Neste caso estamos perturbando a medição fornecida pelo método 1. Na Figura 4.13 temos um gráfico mostrando a influência dos valores observados sobre a estimativa dos parâmetros. Nesse gráfico pode-se observar que os adolescentes 3, 30, 54 e 58 exercem uma influência moderada na estimação dos parâmetros, pois não há observação que se destaque fortemente em relação as outras observações.

Na Tabela 4.17 apresentamos o autovetor (em valor absoluto), correspondente ao maior autovalor de $\Delta^T (\ddot{\mathbf{L}})^{-1} \Delta$, isto é, os valores absolutos dos componentes do vetor $\mathbf{d}_{\max}$.

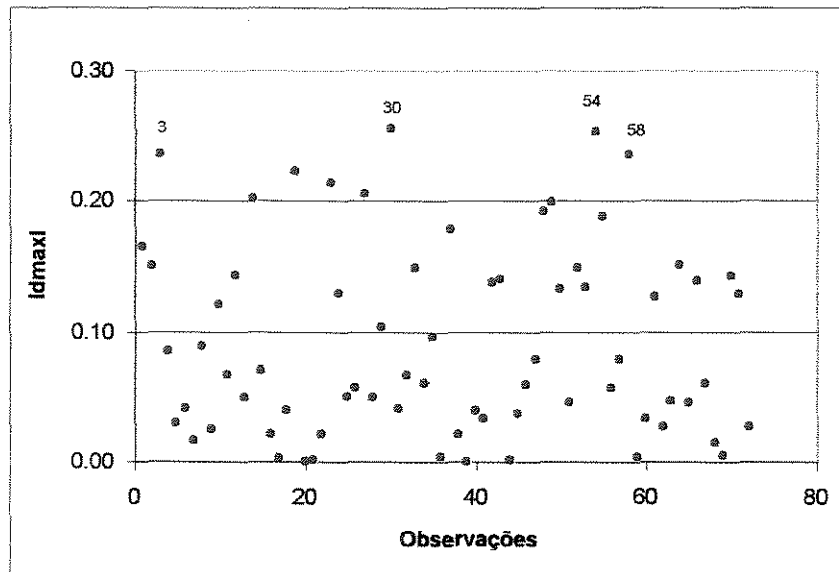


Figura 4-13: Perturbação aditiva no instrumento 1

Tabela 4.17: Elementos de $d_{\max}$ em valor absoluto correspondente aos dados Barnett(1969), em perturbação aditiva.

Obs.	$ d_{\max} $	Obs.	$ d_{\max} $	Obs.	$ d_{\max} $
1	0,1660	25	0,0510	49	0,2007
2	0,1511	26	0,0573	50	0,1337
3	0,2369	27	0,2063	51	0,0469
4	0,0859	28	0,0508	52	0,1498
5	0,0307	29	0,1040	53	0,1352
6	0,0418	30	0,2568	54	0,2542
7	0,0178	31	0,0415	55	0,1889
8	0,0896	32	0,0674	56	0,0583
9	0,0264	33	0,1494	57	0,0798
10	0,1214	34	0,0614	58	0,2373
11	0,0671	35	0,0972	59	0,0054
12	0,1435	36	0,0054	60	0,0342
13	0,0509	37	0,1796	61	0,1281
14	0,2026	38	0,0226	62	0,0286
15	0,0716	39	0,0016	63	0,0483
16	0,0218	40	0,0400	64	0,1522
17	0,0034	41	0,0340	65	0,0472
18	0,0408	42	0,1394	66	0,1404
19	0,2239	43	0,1418	67	0,0616
20	0,0017	44	0,0019	68	0,0156
21	0,0028	45	0,0387	69	0,006
22	0,0224	46	0,0608	70	0,1441
23	0,2155	47	0,0801	71	0,1303
24	0,1307	48	0,1931	72	0,0284

4.6.3 Perturbação Multiplicativa

Nesta seção vamos considerar a perturbação multiplicativa. Neste caso vamos perturbar a medição fornecida pelo método 1. Na Figura 4.14 temos um gráfico mostrando a influência dos valores observados sobre a estimativa dos parâmetros. Nesse gráfico, observa-se que os adolescentes 23 e 30 se destacam do resto.

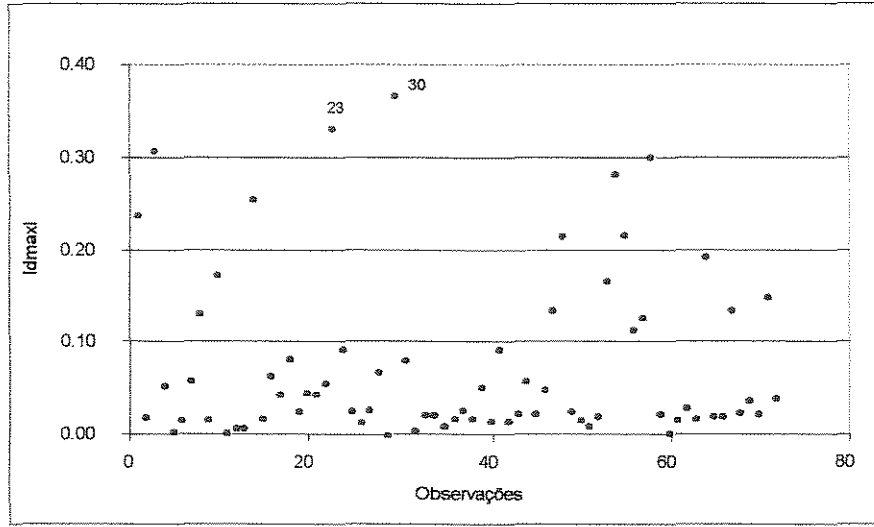


Figura 4-14: Perturbação multiplicativa no instrumento 1

Na Tabela 4.19, apresentamos o autovetor (em valor absoluto), correspondente ao maior autovalor de $\Delta^T (\ddot{\mathbf{L}})^{-1} \Delta$, isto é, os valores absolutos dos componentes do vetor $\mathbf{d}_{\max}$. Na Tabela 4.18 apresentamos as estimativas dos parâmetros; a primeira linha mostra as estimativas dos parâmetros considerando o total dos dados, a segunda linha mostra as estimativas sem considerar as medições das observações 23 e 30.

Tabela 4.18: Estimativas dos parâmetros

μ_1	μ_2	μ_3	λ_1	λ_2	λ_3	ϕ_1	ϕ_2	ϕ_3
2246,11	2175,69	2148,61	733,535	781,664	860,669	46216,943	7840,107	47185,975
2197,428	2118,428	2096,00	683,576	715,937	815,035	48157,272	5895,915	46519,235

$$\begin{aligned}
&= \frac{\partial}{\partial \phi} \left\{ (\lambda_1, \dots, \lambda_p) \begin{bmatrix} \frac{1}{\phi_1} & & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{1}{\phi_p} \end{bmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_p \end{pmatrix} \right\} \\
&= \frac{\partial}{\partial \phi} \left\{ \left(\frac{\lambda_1}{\phi_1}, \dots, \frac{\lambda_p}{\phi_p} \right) \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_p \end{pmatrix} \right\} \\
&= \left(\frac{\partial}{\partial \phi} \left(\frac{\lambda_1}{\phi_1} \right), \dots, \frac{\partial}{\partial \phi} \left(\frac{\lambda_p}{\phi_p} \right) \right) \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_p \end{pmatrix} \\
&= \begin{pmatrix} \frac{\partial}{\partial \phi_1} \left(\frac{\lambda_1}{\phi_1} \right) & \dots & \frac{\partial \lambda_p}{\partial \phi_1} \left(\frac{\lambda_p}{\phi_p} \right) \\ \dots & \ddots & \dots \\ \frac{\partial \lambda_p}{\partial \phi_p} \left(\frac{\lambda_p}{\phi_p} \right) & \dots & \frac{\partial \lambda_p}{\partial \phi_p} \left(\frac{\lambda_p}{\phi_p} \right) \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_p \end{pmatrix} \\
&= \begin{pmatrix} -\frac{\lambda_1}{\phi_1^2} & \dots & 0 \\ \dots & \ddots & \dots \\ 0 & \dots & -\frac{\lambda_p}{\phi_p^2} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_p \end{pmatrix} \\
&= -D(\lambda) D^{-2}(\phi) \lambda.
\end{aligned}$$

(3) Da definição A2, segue que

$$\begin{aligned}
\frac{\partial \log |D(\phi)|}{\partial \phi} &= \frac{\partial \lg \Pi_{j=1}^p \phi_j}{\partial \phi} = \frac{1}{\Pi_{j=1}^p \phi_j} \left(\frac{\partial}{\partial \phi} (\Pi_{j=1}^p \phi_j) \right) \\
&= \frac{1}{\Pi_{j=1}^p \phi_j} \begin{pmatrix} \frac{\partial}{\partial \phi_1} (\Pi_{j=1}^p \phi_j) \\ \vdots \\ \frac{\partial}{\partial \phi_p} (\Pi_{j=1}^p \phi_j) \end{pmatrix} = \frac{1}{\Pi_{j=1}^p \phi_j} \begin{pmatrix} \Pi_{j \neq 1}^p \phi_j \\ \vdots \\ \Pi_{j \neq p}^p \phi_j \end{pmatrix} = \begin{pmatrix} \frac{1}{\phi_1} \\ \vdots \\ \frac{1}{\phi_p} \end{pmatrix} \\
&= \begin{pmatrix} \frac{1}{\phi_1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{1}{\phi_p} \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = D^{-1}(\phi) \mathbf{1}.
\end{aligned}$$

Lema A2: Sejam $\mathbf{a}$ e $\mathbf{b}$ vetores e $\mathbf{M} = D^{-1}(\phi) \boldsymbol{\lambda} \boldsymbol{\lambda}^T D^{-1}(\phi)$. Então

- 1) $\frac{\partial}{\partial \boldsymbol{\lambda}} \{ \boldsymbol{\lambda}^T \mathbf{M} \boldsymbol{\lambda} \} = 2 \boldsymbol{\lambda}^T D^{-1}(\phi) \boldsymbol{\lambda} D^{-1}(\phi) \boldsymbol{\lambda}$
- 2) $\frac{\partial}{\partial \phi} (\mathbf{a}^T D^{-k}(\phi)) = -k D(\mathbf{a}) D^{-(k+1)}(\phi)$
- 3) $\frac{\partial}{\partial \phi} \{ \mathbf{a}^T \mathbf{M} \mathbf{b} \} = -\boldsymbol{\lambda}^T D^{-1}(\phi) \mathbf{b} D(\boldsymbol{\lambda}) D^{-2}(\phi) \mathbf{a} - \boldsymbol{\lambda}^T D^{-1}(\phi) \mathbf{a} D(\boldsymbol{\lambda}) D^{-2}(\phi) \mathbf{b}$
- 4) $\frac{\partial}{\partial \boldsymbol{\lambda}} \mathbf{a}^T \mathbf{M} = D^{-1}(\phi) \mathbf{a} \boldsymbol{\lambda}^T D^{-1}(\phi) + \boldsymbol{\lambda}^T D^{-1}(\phi) \mathbf{a} D^{-1}(\phi)$
- 5) $\frac{\partial}{\partial \phi} (\mathbf{a}^T \mathbf{M}) = -D(\boldsymbol{\lambda}) D^{-2}(\phi) \mathbf{a} \boldsymbol{\lambda}^T D^{-1}(\phi) - \mathbf{a}^T D^{-1}(\phi) \boldsymbol{\lambda} D(\boldsymbol{\lambda}) D^{-2}(\phi)$
 $= -D(\mathbf{a}) D^{-2}(\phi) \boldsymbol{\lambda} \boldsymbol{\lambda}^T D^{-1}(\phi) - \mathbf{a}^T D^{-1}(\phi) \boldsymbol{\lambda} D(\boldsymbol{\lambda}) D^{-2}(\phi).$

Prova: (1) Da propriedade A1 (a), segue que

$$\begin{aligned}
 \frac{\partial}{\partial \boldsymbol{\lambda}} \{ \boldsymbol{\lambda}^T \mathbf{M} \boldsymbol{\lambda} \} &= \frac{\partial}{\partial \boldsymbol{\lambda}} \left\{ \underbrace{\boldsymbol{\lambda}^T D^{-1}(\phi)}_{\text{número}} \underbrace{\boldsymbol{\lambda} \boldsymbol{\lambda}^T D^{-1}(\phi)}_{\text{número}} \boldsymbol{\lambda} \right\} \\
 &= \frac{\partial}{\partial \boldsymbol{\lambda}} (\boldsymbol{\lambda}^T D^{-1}(\phi) \boldsymbol{\lambda})^2 \\
 &= 2 (\boldsymbol{\lambda}^T D^{-1}(\phi) \boldsymbol{\lambda}) \frac{\partial}{\partial \boldsymbol{\lambda}} (\boldsymbol{\lambda}^T D^{-1}(\phi) \boldsymbol{\lambda}) \\
 &= 2 (\boldsymbol{\lambda}^T D^{-1}(\phi) \boldsymbol{\lambda}) (D^{-1}(\phi) \boldsymbol{\lambda}) \\
 &= 2 \boldsymbol{\lambda}^T D^{-1}(\phi) \boldsymbol{\lambda} D^{-1}(\phi) \boldsymbol{\lambda}.
 \end{aligned}$$

(2) Da definição A2, vamos ter que

$$\begin{aligned}
 \frac{\partial}{\partial \phi} (\mathbf{a}^T D^{-k}(\phi)) &= \frac{\partial}{\partial \phi} \left\{ (a_1, \dots, a_p) \begin{pmatrix} \frac{1}{\phi_1^k} & 0 & 0 \\ 0 & \ddots & \\ 0 & \dots & \frac{1}{\phi_p^k} \end{pmatrix} \right\} \\
 &= \left[\frac{\partial}{\partial \phi} \left(\frac{a_1}{\phi_1^k} \right), \dots, \frac{\partial}{\partial \phi} \left(\frac{a_p}{\phi_p^k} \right) \right] = \begin{pmatrix} \frac{\partial}{\partial \phi_1} \left(\frac{a_1}{\phi_1^k} \right) & \dots & \frac{\partial}{\partial \phi_1} \left(\frac{a_1}{\phi_p^k} \right) \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial \phi_p} \left(\frac{a_1}{\phi_1^k} \right) & \dots & \frac{\partial}{\partial \phi_p} \left(\frac{a_1}{\phi_p^k} \right) \end{pmatrix}
 \end{aligned}$$

$$\begin{aligned}
&= \begin{pmatrix} -\frac{k a_1}{\phi_1^{k+1}} & \dots & 0 \\ \vdots & \ddots & 0 \\ 0 & \dots & -\frac{k a_p}{\phi_p^{k+1}} \end{pmatrix} \\
&= -k D(\mathbf{a}) D^{-(k+1)}(\boldsymbol{\phi}).
\end{aligned}$$

(3) Aplicando a definição A2 e a propriedade A2, segue que

$$\begin{aligned}
\frac{\partial}{\partial \boldsymbol{\phi}} \{ \mathbf{a}^T \mathbf{M} \mathbf{b} \} &= \frac{\partial}{\partial \boldsymbol{\phi}} \left\{ \underbrace{\mathbf{a}^T D^{-1}(\boldsymbol{\phi}) \boldsymbol{\lambda}}_{\text{número}} \underbrace{\boldsymbol{\lambda}^T D^{-1}(\boldsymbol{\phi}) \mathbf{b}}_{\text{número}} \right\} \\
&= \frac{\partial}{\partial \boldsymbol{\phi}} (\boldsymbol{\lambda}^T D^{-1}(\boldsymbol{\phi}) \mathbf{a}) (\boldsymbol{\lambda}^T D^{-1}(\boldsymbol{\phi}) \mathbf{b}) \\
&= \frac{\partial}{\partial \boldsymbol{\phi}} (\boldsymbol{\lambda}^T D^{-1}(\boldsymbol{\phi}) \mathbf{a}) \boldsymbol{\lambda}^T D^{-1}(\boldsymbol{\phi}) \mathbf{b} + \boldsymbol{\lambda}^T D^{-1}(\boldsymbol{\phi}) \mathbf{a} \frac{\partial}{\partial \boldsymbol{\phi}} (\boldsymbol{\lambda}^T D^{-1}(\boldsymbol{\phi}) \mathbf{b}) \\
&= \left[\frac{\partial}{\partial \boldsymbol{\phi}} \left(\frac{\lambda_1}{\phi_1} \right), \dots, \frac{\partial}{\partial \boldsymbol{\phi}} \left(\frac{\lambda_p}{\phi_p} \right) \right] \begin{pmatrix} a_1 \\ \vdots \\ a_p \end{pmatrix} \boldsymbol{\lambda}^T D^{-1}(\boldsymbol{\phi}) \mathbf{b} \\
&\quad + \boldsymbol{\lambda}^T D^{-1}(\boldsymbol{\phi}) \mathbf{a} \begin{bmatrix} \frac{\partial}{\partial \phi_1} \left(\frac{\lambda_1}{\phi_1} \right) & \dots & \frac{\partial}{\partial \phi_1} \left(\frac{\lambda_p}{\phi_p} \right) \\ \vdots & \dots & \vdots \\ \frac{\partial}{\partial \phi_p} \left(\frac{\lambda_1}{\phi_1} \right) & \dots & \frac{\partial}{\partial \phi_p} \left(\frac{\lambda_p}{\phi_p} \right) \end{bmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix} \\
&= \begin{bmatrix} \frac{\partial}{\partial \phi_1} \left(\frac{\lambda_1}{\phi_1} \right) & \dots & \frac{\partial}{\partial \phi_1} \left(\frac{\lambda_p}{\phi_p} \right) \\ \vdots & \dots & \vdots \\ \frac{\partial}{\partial \phi_p} \left(\frac{\lambda_1}{\phi_1} \right) & \dots & \frac{\partial}{\partial \phi_p} \left(\frac{\lambda_p}{\phi_p} \right) \end{bmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_p \end{pmatrix} \boldsymbol{\lambda}^T D^{-1}(\boldsymbol{\phi}) \mathbf{b} \\
&\quad + \boldsymbol{\lambda}^T D^{-1}(\boldsymbol{\phi}) \mathbf{a} \begin{bmatrix} \frac{\partial}{\partial \phi_1} \left(\frac{\lambda_1}{\phi_1} \right) & \dots & \frac{\partial}{\partial \phi_1} \left(\frac{\lambda_p}{\phi_p} \right) \\ \vdots & \dots & \vdots \\ \frac{\partial}{\partial \phi_p} \left(\frac{\lambda_1}{\phi_1} \right) & \dots & \frac{\partial}{\partial \phi_p} \left(\frac{\lambda_p}{\phi_p} \right) \end{bmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix} \\
&= \begin{bmatrix} -\frac{\lambda_1}{\phi_1^2} & \dots & 0 \\ \vdots & \dots & \vdots \\ 0 & \dots & -\frac{\lambda_p}{\phi_p^2} \end{bmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_p \end{pmatrix} \boldsymbol{\lambda}^T D^{-1}(\boldsymbol{\phi}) \mathbf{b}
\end{aligned}$$

$$\begin{aligned}
& + \boldsymbol{\lambda}^T D^{-1}(\phi) \mathbf{a} \begin{bmatrix} -\frac{\lambda_1}{\partial \phi_1^2} & \dots & 0 \\ \vdots & \dots & \vdots \\ 0 & \dots & -\frac{\lambda_p}{\partial \phi_p^2} \end{bmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix} \\
& = -D(\boldsymbol{\lambda}) D^{-2}(\phi) \mathbf{a} \underbrace{\boldsymbol{\lambda}^T D^{-1}(\phi) \mathbf{b}}_{\text{número}} - \boldsymbol{\lambda}^T D^{-1}(\phi) \mathbf{a} D(\boldsymbol{\lambda}) D^{-2}(\phi) \mathbf{b} \\
& = -\boldsymbol{\lambda}^T D^{-1}(\phi) \mathbf{b} D(\boldsymbol{\lambda}) D^{-2}(\phi) \mathbf{a} - \boldsymbol{\lambda}^T D^{-1}(\phi) \mathbf{a} D(\boldsymbol{\lambda}) D^{-2}(\phi) \mathbf{b}.
\end{aligned}$$

(4) Aplicando a Definição A2 e a Propriedade A1(a), obtemos

$$\begin{aligned}
\frac{\partial}{\partial \boldsymbol{\lambda}} \{ \mathbf{a}^T \mathbf{M} \} &= \frac{\partial}{\partial \boldsymbol{\lambda}} \{ \mathbf{a}^T D^{-1}(\phi) \boldsymbol{\lambda} \boldsymbol{\lambda}^T D^{-1}(\phi) \} \\
&= \frac{\partial}{\partial \boldsymbol{\lambda}} \{ \boldsymbol{\lambda}^T D^{-1}(\phi) \mathbf{a} \boldsymbol{\lambda}^T D^{-1}(\phi) \} \\
&= \frac{\partial}{\partial \boldsymbol{\lambda}} (\boldsymbol{\lambda}^T D^{-1}(\phi) \mathbf{a}) \boldsymbol{\lambda}^T D^{-1}(\phi) + \boldsymbol{\lambda}^T D^{-1}(\phi) \mathbf{a} \frac{\partial}{\partial \boldsymbol{\lambda}} (\boldsymbol{\lambda}^T D^{-1}(\phi)) \\
&= D^{-1}(\phi) \mathbf{a} \boldsymbol{\lambda}^T D^{-1}(\phi) + \boldsymbol{\lambda}^T D^{-1}(\phi) \mathbf{a} D^{-1}(\phi).
\end{aligned}$$

(5) Aplicando a definição A2 e o Lema A2(2), obtemos que

$$\begin{aligned}
\frac{\partial}{\partial \phi} \{ \mathbf{a}^T \mathbf{M} \} &= \frac{\partial}{\partial \phi} \left(\underbrace{\mathbf{a}^T D^{-1}(\phi) \boldsymbol{\lambda} \boldsymbol{\lambda}^T D^{-1}(\phi)}_{\text{número}} \right) \\
&= \frac{\partial}{\partial \phi} (\boldsymbol{\lambda}^T D^{-1}(\phi) \mathbf{a}) \boldsymbol{\lambda}^T D^{-1}(\phi) + \mathbf{a}^T D^{-1}(\phi) \boldsymbol{\lambda} \frac{\partial}{\partial \phi} (\boldsymbol{\lambda}^T D^{-1}(\phi)) \\
&= -D(\boldsymbol{\lambda}) D^{-2}(\phi) \mathbf{a} \boldsymbol{\lambda}^T D^{-1}(\phi) + \mathbf{a}^T D^{-1}(\phi) \boldsymbol{\lambda} (-D(\boldsymbol{\lambda}) D^{-2}(\phi)) \\
&= -D(\boldsymbol{\lambda}) D^{-2}(\phi) \mathbf{a} \boldsymbol{\lambda}^T D^{-1}(\phi) - \mathbf{a}^T D^{-1}(\phi) \boldsymbol{\lambda} D(\boldsymbol{\lambda}) D^{-2}(\phi).
\end{aligned}$$

Lema A3: Sejam $\mathbf{a}$ e $\mathbf{b}$ vetores. Então

$$\frac{\partial}{\partial \mathbf{a}} (\mathbf{a}^T D(\mathbf{a}) \mathbf{b}) = 2D(\mathbf{a}) \mathbf{b}.$$

Prova: A prova segue aplicando a Definição A2

$$\begin{aligned}
\frac{\partial}{\partial \mathbf{a}} (\mathbf{a}^T D(\mathbf{a}) \mathbf{b}) &= \frac{\partial}{\partial \mathbf{a}} \left\{ (a_1, \dots, a_p) \begin{pmatrix} a_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & a_p \end{pmatrix} \right\} \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix} \\
&= \frac{\partial}{\partial \mathbf{a}} \left\{ (a_1^2, \dots, a_p^2) \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix} \right\} \\
&= \left[\frac{\partial}{\partial \mathbf{a}} (a_1^2), \dots, \frac{\partial}{\partial \mathbf{a}} (a_p^2) \right] \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix} \\
&= \begin{bmatrix} \frac{\partial}{\partial a_1} (a_1^2) & \dots & \frac{\partial}{\partial a_1} (a_p^2) \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial a_p} (a_1^2) & \dots & \frac{\partial}{\partial a_p} (a_p^2) \end{bmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix} \\
&= \begin{bmatrix} 2a_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 2a_p \end{bmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix} \\
&= 2D(\mathbf{a}) \mathbf{b}.
\end{aligned}$$

Apêndice B

Obtenção da Matriz de Informação de Fisher observada no Modelo de Calibração Estrutural

Nesta seção apresentamos os cálculos para obter a matriz de informação de Fisher observada.

Seja $l(\boldsymbol{\theta}) = \sum_{j=1}^n l_j(\boldsymbol{\theta})$ o log-verossimilhança associado ao modelo de calibração comparativa, onde

$$l_j(\boldsymbol{\theta}) = \log(2\pi)^{-3/2} - \frac{1}{2} \log |\boldsymbol{\Sigma}| - \frac{1}{2} \|T_j\|^2, \quad j = 1, \dots, n.$$

Então a matriz de informação de Fisher observada é definido por

$$-\frac{\partial^2 l}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} = -\sum_{j=1}^n \frac{\partial^2 l_j}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T}, \quad (\text{B.1})$$

onde $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^T, \boldsymbol{\theta}_2^T, \boldsymbol{\theta}_3^T)^T = (\boldsymbol{\mu}^T, \boldsymbol{\lambda}^T, \boldsymbol{\phi}^T)^T$ e

$$\begin{aligned} \frac{\partial^2 l_j}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} &= -\frac{1}{2} \frac{\partial^2 \log |\boldsymbol{\Sigma}|}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} - \frac{1}{2} \frac{\partial^2 \log \|T_j\|^2}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \\ &= \begin{pmatrix} \frac{\partial^2 l_j}{\partial \boldsymbol{\mu} \partial \boldsymbol{\mu}^T} & \frac{\partial^2 l_j}{\partial \boldsymbol{\mu} \partial \boldsymbol{\lambda}^T} & \frac{\partial^2 l_j}{\partial \boldsymbol{\mu} \partial \boldsymbol{\phi}^T} \\ \frac{\partial^2 l_j}{\partial \boldsymbol{\lambda} \partial \boldsymbol{\mu}^T} & \frac{\partial^2 l_j}{\partial \boldsymbol{\lambda} \partial \boldsymbol{\lambda}^T} & \frac{\partial^2 l_j}{\partial \boldsymbol{\lambda} \partial \boldsymbol{\phi}^T} \\ \frac{\partial^2 l_j}{\partial \boldsymbol{\phi} \partial \boldsymbol{\mu}^T} & \frac{\partial^2 l_j}{\partial \boldsymbol{\phi} \partial \boldsymbol{\lambda}^T} & \frac{\partial^2 l_j}{\partial \boldsymbol{\phi} \partial \boldsymbol{\phi}^T} \end{pmatrix} \end{aligned}$$

com

$$\frac{\partial^2 l_j}{\partial \theta_i \partial \theta_k^T} = -\frac{1}{2} \frac{\partial^2 \log |\Sigma|}{\partial \theta_i \partial \theta_k^T} - \frac{1}{2} \frac{\partial^2 \log \|T_j\|^2}{\partial \theta_i \partial \theta_k^T} \quad (\text{B.2})$$

$$i, k = 1, 2, 3, \quad j = 1, \dots, n.$$

Observe que

$$\|T_j\|^2 = \mathbf{W}_j^T D^{-1}(\phi) \mathbf{W}_j - c^{-1} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j, \quad (\text{B.3})$$

onde $\mathbf{W}_j = \mathbf{Y}_j - \boldsymbol{\mu}$, $\mathbf{M} = \mathbf{D}^{-1}(\phi) \boldsymbol{\lambda} \boldsymbol{\lambda}^T \mathbf{D}^{-1}(\phi)$ e $c = 1 + \boldsymbol{\lambda}^T \mathbf{D}^{-1}(\phi) \boldsymbol{\lambda}$.

Lema B1: Seja $\|T_j\|^2$ como em B.3. Então

$$\begin{aligned} 1) \quad \frac{\partial \|T_j\|^2}{\partial \boldsymbol{\mu}} &= -2 \boldsymbol{\Sigma}^{-1} \mathbf{W}_j \\ 2) \quad \frac{\partial \|T_j\|^2}{\partial \boldsymbol{\lambda}} &= 2c^{-2} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j D^{-1}(\phi) \boldsymbol{\lambda} - 2c^{-1} D^{-1}(\phi) \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi) \boldsymbol{\lambda} \\ 3) \quad \frac{\partial \|T_j\|^2}{\partial \phi} &= -D^{-2}(\phi) D(\mathbf{W}_j) \mathbf{W}_j - c^{-2} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j D(\boldsymbol{\lambda}) D^{-2}(\phi) \boldsymbol{\lambda} \\ &\quad + 2c^{-1} D(\boldsymbol{\lambda}) D^{-2}(\phi) \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi) \boldsymbol{\lambda}. \end{aligned}$$

Prova: (1) Aplicando a Propriedade A2

$$\begin{aligned} \frac{\partial \|T_j\|^2}{\partial \boldsymbol{\mu}} &= \frac{\partial}{\partial \boldsymbol{\mu}} \mathbf{W}_j^T \boldsymbol{\Sigma}^{-1} \mathbf{W}_j \\ &= \frac{\partial \mathbf{W}_j^T}{\partial \boldsymbol{\mu}} \times \frac{\partial \mathbf{W}_j^T \boldsymbol{\Sigma}^{-1} \mathbf{W}_j}{\partial \mathbf{W}_j} \\ &= -2 \boldsymbol{\Sigma}^{-1} \mathbf{W}_j \end{aligned}$$

(2) Aplicando o Lema 1(1) e a Propriedade A1, segue que

$$\begin{aligned} \frac{\partial \|T_j\|^2}{\partial \boldsymbol{\lambda}} &= \frac{\partial}{\partial \boldsymbol{\lambda}} \left[\mathbf{W}_j^T D^{-1}(\phi) \mathbf{W}_j - c^{-1} (\boldsymbol{\lambda}^T D^{-1}(\phi) \mathbf{W}_j)^2 \right] \\ &= - \left[\frac{\partial}{\partial \boldsymbol{\lambda}} (c^{-1}) (\boldsymbol{\lambda}^T D^{-1}(\phi) (\mathbf{W}_j))^2 \right. \\ &\quad \left. + c^{-1} 2 \boldsymbol{\lambda}^T D^{-1}(\phi) \mathbf{W}_j \frac{\partial}{\partial \boldsymbol{\lambda}} (\boldsymbol{\lambda}^T D^{-1}(\phi) \mathbf{W}_j) \right] \\ &= - \left[-2c^{-2} D^{-1}(\phi) \boldsymbol{\lambda} (\boldsymbol{\lambda}^T D^{-1}(\phi) \mathbf{W}_j)^2 \right. \end{aligned}$$

$$\begin{aligned}
& +2c^{-1}\boldsymbol{\lambda}^T D^{-1}(\phi) \mathbf{W}_j \frac{\partial}{\partial \boldsymbol{\lambda}} \left(\boldsymbol{\lambda}^T D^{-1}(\phi) \mathbf{W}_j \right) \Big] \\
& = 2c^{-2} D^{-1}(\phi) \boldsymbol{\lambda} \left(\boldsymbol{\lambda}^T D^{-1}(\phi) \mathbf{W}_j \right)^2 \\
& \quad -2c^{-1} \boldsymbol{\lambda}^T D^{-1}(\phi) \mathbf{W}_j D^{-1}(\phi) \mathbf{W}_j \\
& = 2c^{-2} D^{-1}(\phi) \boldsymbol{\lambda} \underbrace{\left(\boldsymbol{\lambda}^T D^{-1}(\phi) \mathbf{W}_j \right)}_{\text{número}} \underbrace{\left(\boldsymbol{\lambda}^T D^{-1}(\phi) \mathbf{W}_j \right)}_{\text{número}} \\
& \quad -2c^{-1} \boldsymbol{\lambda}^T D^{-1}(\phi) \mathbf{W}_j D^{-1}(\phi) \mathbf{W}_j \\
& = 2c^{-2} \boldsymbol{\lambda}^T D^{-1}(\phi) \mathbf{W}_j \boldsymbol{\lambda}^T D^{-1}(\phi) \mathbf{W}_j D^{-1}(\phi) \boldsymbol{\lambda} \\
& \quad -2c^{-1} \underbrace{\boldsymbol{\lambda}^T D^{-1}(\phi) \mathbf{W}_j}_{\text{número}} D^{-1}(\phi) \mathbf{W}_j \\
& = 2c^{-2} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j D^{-1}(\phi) \boldsymbol{\lambda} - 2c^{-1} D^{-1}(\phi) \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi) \boldsymbol{\lambda}
\end{aligned}$$

(3) Aplicando os Lemas A1 e A2, segue que

$$\begin{aligned}
\frac{\partial \|T_j\|^2}{\partial \phi} & = \frac{\partial}{\partial \phi} [\mathbf{W}_j^T D^{-1}(\phi) \mathbf{W}_j - c^{-1} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j] \\
& = \frac{\partial}{\partial \phi} (\mathbf{W}_j^T D^{-1}(\phi) \mathbf{W}_j) - \frac{\partial}{\partial \phi} (c^{-1}) \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j - c^{-1} \frac{\partial}{\partial \phi} (\mathbf{W}_j^T \mathbf{M} \mathbf{W}_j) \\
& = -D(\mathbf{W}_j) D^{-2}(\phi) \mathbf{W}_j + c^{-2} (-D(\boldsymbol{\lambda}) D^{-2}(\phi) \boldsymbol{\lambda}) \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j \\
& \quad -c^{-1} (2\boldsymbol{\lambda}^T D^{-1}(\phi) \mathbf{W}_j D(\boldsymbol{\lambda}) D^{-2}(\phi) \mathbf{W}_j) \\
& = -D(\mathbf{W}_j) D^{-2}(\phi) \mathbf{W}_j - c^{-2} D(\boldsymbol{\lambda}) D^{-2}(\phi) \boldsymbol{\lambda} \underbrace{\mathbf{W}_j^T \mathbf{M} \mathbf{W}_j}_{\text{número}} \\
& \quad +2c^{-1} \underbrace{\boldsymbol{\lambda}^T D^{-1}(\phi) \mathbf{W}_j}_{\text{número}} D(\boldsymbol{\lambda}) D^{-2}(\phi) \mathbf{W}_j \\
& = -D^{-2}(\phi) D(\mathbf{W}_j) \mathbf{W}_j - c^{-2} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j D(\boldsymbol{\lambda}) D^{-2}(\phi) \boldsymbol{\lambda} \\
& \quad +2c^{-1} D(\boldsymbol{\lambda}) D^{-2}(\phi) \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi) \boldsymbol{\lambda}
\end{aligned}$$

Lema B2: Seja $\|T_j\|^2$ como em (B.3). Então

$$1) \frac{\partial \|T_j\|^2}{\partial \mu \partial \mu^T} = 2(D^{-1}(\phi) - c^{-1}\mathbf{M}) = 2\Sigma^{-1}.$$

$$2) \frac{\partial \|T_j\|}{\partial \lambda \partial \lambda^T} = -2c^{-1}D^{-1}(\phi) \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi) + 4c^{-2}D^{-1}(\phi) \mathbf{W}_j \mathbf{W}_j^T \mathbf{M} \\ + 2c^{-2}\mathbf{W}_j^T \mathbf{M} \mathbf{W}_j (\Sigma^{-1} - 3c^{-1}\mathbf{M}) + 4c^{-2}\mathbf{M} \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi)$$

$$3) \frac{\partial^2 \|T_j\|^2}{\partial \phi \partial \phi^T} = 2D(\mathbf{W}_j) D^{-3}(\phi) D(\mathbf{W}_j)$$

$$-2c^{-3}\mathbf{W}_j^T \mathbf{M} \mathbf{W}_j D(\lambda) D^{-1}(\phi) \mathbf{M} D^{-1}(\phi) D(\lambda) \\ + 2c^{-2}D(\lambda) D^{-2}(\phi) \mathbf{W}_j \mathbf{W}_j^T \mathbf{M} D^{-1}(\phi) D(\lambda) \\ + 2c^{-2}\mathbf{W}_j^T \mathbf{M} \mathbf{W}_j D(\lambda) D^{-3}(\phi) D(\lambda) \\ + 2c^{-2}D(\lambda) D^{-1}(\phi) \mathbf{M} \mathbf{W}_j \mathbf{W}_j^T D(\lambda) D^{-2}(\phi) \\ - 2c^{-1}D(\lambda) D^{-2}(\phi) \mathbf{W}_j \mathbf{W}_j^T D(\lambda) D^{-2}(\phi) \\ - 4c^{-1}D^{-3}(\phi) D(\lambda) \mathbf{W}_j^T D^{-1}(\phi) \lambda D(\mathbf{W}_j).$$

$$4) \frac{\partial^2 \|T_j\|^2}{\partial \lambda \partial \mu^T} = 2c^{-1} \{D^{-1}(\phi) \mathbf{W}_j \lambda^T D^{-1}(\phi) + \lambda^T D^{-1}(\phi) \mathbf{W}_j [\Sigma^{-1} - c^{-1}\mathbf{M}]\}$$

$$5) \frac{\partial^2 \|T_j\|^2}{\partial \phi \partial \mu^T} = 2D(\mathbf{W}_j) D^{-2}(\phi) + 2c^{-2}D(\lambda) D^{-2}(\phi) \lambda \mathbf{W}_j^T \mathbf{M} - 2c^{-1}D(\lambda) D^{-2}(\phi) \mathbf{W}_j \lambda^T D^{-1}(\phi) \\ - 2c^{-1}\mathbf{W}_j^T D^{-1}(\phi) \lambda D(\lambda) D^{-2}(\phi)$$

$$6) \frac{\partial^2 \|T_j\|^2}{\partial \phi \partial \lambda^T} = 4c^{-3}\mathbf{W}_j^T \mathbf{M} \mathbf{W}_j D^{-1}(\phi) D(\lambda) \mathbf{M} - 4c^{-2}D(\lambda) D^{-2}(\phi) \mathbf{W}_j \mathbf{W}_j^T \mathbf{M} \\ - 2c^{-2}\mathbf{W}_j^T \mathbf{M} \mathbf{W}_j D(\lambda) D^{-2}(\phi) - 2c^{-2}D(\lambda) D^{-1}(\phi) \mathbf{M} \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi) \\ + 2c^{-1}D(\lambda) D^{-2}(\phi) \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi) + 2c^{-1}\mathbf{W}_j^T D^{-1}(\phi) \lambda D(\mathbf{W}_j) D^{-2}(\phi).$$

Prova: (1) De (B.3) , do Lema B1 e aplicando a Propriedade A1, temos que

$$\frac{\partial^2}{\partial \mu} \left(\frac{\|T_j\|^2}{\mu^T} \right) = \frac{\partial}{\partial \mu} (-2\mathbf{W}_j^T D^{-1}(\phi) + 2c^{-1}\mathbf{W}_j^T \mathbf{M}) \\ = -2 \frac{\partial}{\partial \mu} (\mathbf{W}_j^T D^{-1}(\phi)) + 2c^{-1} \frac{\partial}{\partial \mu} (\mathbf{W}_j^T \mathbf{M}) \\ = -2(-D^{-1}(\phi)) - 2c^{-1}\mathbf{M} \\ = 2(D^{-1}(\phi) - c^{-1}\mathbf{M}) = 2\Sigma^{-1}.$$

(2) Do Lema B1, temos que

$$\frac{\partial \|T_j\|^2}{\partial \lambda^T} = 2c^{-2} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j \lambda^T D^{-1}(\phi) - 2c^{-1} \lambda^T D^{-1}(\phi) \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi).$$

Assim, aplicando os Lemas A1 e A2, temos que

$$\begin{aligned} \frac{\partial^2 \|T_j\|^2}{\partial \lambda \partial \lambda^T} &= \frac{\partial}{\partial \lambda} \left(\frac{\partial \|T_j\|}{\partial \lambda^T} \right) \\ &= 2 \frac{\partial}{\partial \lambda} (c^{-2}) \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j \lambda^T D^{-1}(\phi) + 2c^{-2} \frac{\partial}{\partial \lambda} (\mathbf{W}_j^T \mathbf{M} \mathbf{W}_j) \lambda^T D^{-1}(\phi) \\ &\quad + 2c^{-2} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j \frac{\partial}{\partial \lambda} (\lambda^T D^{-1}(\phi)) - 2 \frac{\partial}{\partial \lambda} (c^{-1}) \lambda^T D^{-1}(\phi) \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi) \\ &\quad - 2c^{-1} \frac{\partial}{\partial \lambda} (\lambda^T D^{-1}(\phi) \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi)) \\ &= 2 (-4c^{-3} D^{-1}(\phi) \lambda \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j \lambda^T D^{-1}(\phi)) \\ &\quad + c^{-2} [2\lambda^T D^{-1}(\phi) \mathbf{W}_j D^{-1}(\phi) \mathbf{W}_j] \lambda^T D^{-1}(\phi) + c^{-2} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j D^{-1}(\phi) \\ &\quad - 2 \{-2c^{-2} D^{-1}(\phi) \lambda \lambda^T D^{-1}(\phi) \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi) + c^{-1} D^{-1}(\phi) \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi)\} \\ &= -8c^{-3} D^{-1}(\phi) \lambda \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j \\ &\quad + 4\lambda^T D^{-1} \left(\phi + 4c^{-2} \underbrace{\lambda^T D^{-1}(\phi) \mathbf{W}_j}_{\text{número}} D^{-1}(\phi) \mathbf{W}_j \lambda^T D^{-1}(\phi) \right) \\ &\quad + 4c^{-2} \underbrace{\lambda^T D^{-1}(\phi) \mathbf{W}_j}_{\text{número}} D^{-1}(\phi) \mathbf{W}_j \lambda^T D^{-1}(\phi) + 2c^{-2} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j D^{-1}(\phi) \\ &\quad + 4c^{-2} D^{-1}(\phi) \lambda \lambda^T D^{-1}(\phi) \mathbf{W}_j \mathbf{W}_j D^{-1}(\phi) - 2c^{-1} D^{-1}(\phi) \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi) \\ &= -8c^{-3} D^{-1}(\phi) \lambda \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j \lambda^T D^{-1}(\phi) + 4c^{-2} D^{-1}(\phi) \lambda \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi) \lambda \lambda^T D^{-1}(\phi) (\phi) \\ &\quad + 2c^{-2} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j D^{-1}(\phi) + 4c^{-2} \mathbf{M} \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi) (\phi) \\ &\quad - 2c^{-1} D^{-1}(\phi) \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi) \\ &= -8c^{-3} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j D^{-1}(\phi) \lambda \lambda^T D^{-1}(\phi) + 4c^{-2} D^{-1}(\phi) \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi) \underbrace{\lambda \lambda^T D^{-1}(\phi)}_{\text{M}} (\phi) \\ &\quad + 2c^{-2} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j D^{-1}(\phi) + 4c^{-2} \underbrace{D^{-1}(\phi) \lambda \lambda^T D^{-1}(\phi)}_{\text{M}} \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi) (\phi) \\ &\quad - 2c^{-1} D^{-1}(\phi) \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi) \\ &= -8c^{-3} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j \mathbf{M} + 4c^{-2} D^{-1}(\phi) \mathbf{W}_j \mathbf{W}_j^T \mathbf{M} + 2c^{-2} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j D^{-1}(\phi) \\ &\quad + 4c^{-2} \mathbf{M} \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi) - 2c^{-1} D^{-1}(\phi) \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi) \\ &= -2c^{-1} D^{-1}(\phi) \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi) + 4c^{-2} D^{-1}(\phi) \mathbf{W}_j \mathbf{W}_j^T \mathbf{M} \end{aligned}$$

$$+2c^{-2}\mathbf{W}_j^T\mathbf{M}\mathbf{W}_j\left(\underbrace{D^{-1}(\phi)-c^{-1}\mathbf{M}}_{\Sigma^{-1}}-3c^{-1}\mathbf{M}\right)+4c^{-2}\mathbf{M}\mathbf{W}_j\mathbf{W}_j^TD^{-1}(\phi).$$

(3) Do Lema B1, obtemos que

$$\begin{aligned}\frac{\partial\|T_j\|^2}{\partial\phi^T} &= -\mathbf{W}_j^TD^{-2}(\phi)D(\mathbf{W}_j)-c^{-2}\mathbf{W}_j^T\mathbf{M}\mathbf{W}_j\boldsymbol{\lambda}^TD(\boldsymbol{\lambda})D^{-2}(\phi) \\ &\quad +2c^{-1}\boldsymbol{\lambda}^TD^{-1}(\phi)\mathbf{W}_j\mathbf{W}_j^TD(\boldsymbol{\lambda})D^{-2}(\phi)\end{aligned}$$

Assim, o resultado segue após a aplicação dos Lemas A1 e A2

$$\begin{aligned}\frac{\partial^2\|T_j\|^2}{\partial\phi\partial\phi^T} &= \frac{\partial}{\partial\phi}\left\{\frac{\partial\|T_j\|^2}{\partial\phi^T}\right\} = -\frac{\partial}{\partial\phi}\left\{\mathbf{W}_j^TD^{-2}(\phi)D(\mathbf{W}_j)\right\} \\ &\quad -\frac{\partial}{\partial\phi}\left\{\underbrace{c^{-2}\mathbf{W}_j^T\mathbf{M}\mathbf{W}_j\boldsymbol{\lambda}^TD(\boldsymbol{\lambda})D^{-2}(\phi)}_{(1)}\right\} + 2\frac{\partial}{\partial\phi}\left\{\underbrace{c^{-1}\boldsymbol{\lambda}^TD^{-1}(\phi)\mathbf{W}_j\mathbf{W}_j^TD(\boldsymbol{\lambda})D^{-2}(\phi)}_{(2)}\right\}.\end{aligned}$$

Acharemos,

$$\begin{aligned}\frac{\partial}{\partial\phi}(1) &= \frac{\partial}{\partial\phi}(c^{-2})\mathbf{W}_j^T\mathbf{M}\mathbf{W}_j\boldsymbol{\lambda}^TD(\boldsymbol{\lambda})D^{-2}(\phi) + c^{-2}\frac{\partial}{\partial\phi}(\mathbf{W}_j^T\mathbf{M}\mathbf{W}_j)\boldsymbol{\lambda}^TD(\boldsymbol{\lambda})D^{-2}(\phi) \\ &\quad + c^{-2}\mathbf{W}_j^T\mathbf{M}\mathbf{W}_j\frac{\partial}{\partial\phi}(\boldsymbol{\lambda}^TD(\boldsymbol{\lambda})D^{-2}(\phi)) \\ &= (2c^{-3}D(\boldsymbol{\lambda})D^{-2}(\phi)\boldsymbol{\lambda})\mathbf{W}_j^T\mathbf{M}\mathbf{W}_j\boldsymbol{\lambda}^TD(\boldsymbol{\lambda})D^{-2}(\phi) \\ &\quad + c^{-2}(-2\boldsymbol{\lambda}^TD^{-1}(\phi)\mathbf{W}_jD(\boldsymbol{\lambda})D^{-2}(\phi)\mathbf{W}_j)\boldsymbol{\lambda}^TD(\boldsymbol{\lambda})D^{-2}(\phi) \\ &\quad + c^{-2}\mathbf{W}_j^T\mathbf{M}\mathbf{W}_j(-2D(\boldsymbol{\lambda})D^{-3}(\phi)D(\boldsymbol{\lambda})) \\ &= 2c^{-3}D(\boldsymbol{\lambda})D^{-2}(\phi)\underbrace{\boldsymbol{\lambda}\mathbf{W}_j^T\mathbf{M}\mathbf{W}_j}_{\text{número}}\boldsymbol{\lambda}^TD(\boldsymbol{\lambda})D^{-2}(\phi) \\ &\quad - 2c^{-2}\boldsymbol{\lambda}^TD^{-1}(\phi)\mathbf{W}_jD(\boldsymbol{\lambda})D^{-2}(\phi)\mathbf{W}_j\boldsymbol{\lambda}^TD(\boldsymbol{\lambda})D^{-2}(\phi) \\ &\quad - 2c^{-2}\mathbf{W}_j^T\mathbf{M}\mathbf{W}_jD(\boldsymbol{\lambda})D^{-3}(\phi)D(\boldsymbol{\lambda}) \\ &= 2c^{-3}\mathbf{W}_j^T\mathbf{M}\mathbf{W}_jD(\boldsymbol{\lambda})D^{-2}(\phi)\boldsymbol{\lambda}\boldsymbol{\lambda}^TD(\boldsymbol{\lambda})D^{-2}(\phi) \\ &\quad - 2c^{-2}D(\boldsymbol{\lambda})D^{-2}(\phi)\mathbf{W}_j\mathbf{W}_j^TD^{-1}(\phi)\boldsymbol{\lambda}\boldsymbol{\lambda}^TD(\boldsymbol{\lambda})D^{-2}(\phi) \\ &\quad - 2c^{-2}\mathbf{W}_j^T\mathbf{M}\mathbf{W}_jD(\boldsymbol{\lambda})D^{-3}(\phi)D(\boldsymbol{\lambda}) \\ &= 2c^{-3}\mathbf{W}_j^T\mathbf{M}\mathbf{W}_jD(\boldsymbol{\lambda})D^{-1}(\phi)\mathbf{M}D^{-1}(\phi)D(\boldsymbol{\lambda}) \\ &\quad - 2c^{-2}D(\boldsymbol{\lambda})D^{-2}(\phi)\mathbf{W}_j\mathbf{W}_j^T\mathbf{M}D^{-1}(\phi)D(\boldsymbol{\lambda}) \\ &\quad - 2c^{-2}\mathbf{W}_j^T\mathbf{M}\mathbf{W}_jD(\boldsymbol{\lambda})D^{-3}(\phi)D(\boldsymbol{\lambda})\end{aligned}$$

$$\begin{aligned}
\frac{\partial}{\partial \phi} (2) &= \frac{\partial}{\partial \phi} \left(c^{-1} \left[\underbrace{\lambda^T D^{-1}(\phi) \mathbf{W}_j}_{a} \underbrace{\mathbf{W}_j^T D(\lambda) D^{-2}(\phi)}_b \right] \right) \\
&= \frac{\partial}{\partial \phi} (c^{-1}) a + c^{-1} \frac{\partial}{\partial \phi} (\lambda^T D^{-1}(\phi) \mathbf{W}_j) b \\
&\quad + c^{-1} \lambda^T D^{-1}(\phi) \mathbf{W}_j \frac{\partial}{\partial \phi} (\mathbf{W}_j^T D(\lambda) D^{-2}(\phi)) \\
&= (c^{-2} D(\lambda) D^{-2}(\phi) \lambda) a + c^{-1} (-D(\lambda) D^{-2}(\phi) \mathbf{W}_j) b \\
&\quad + c^{-1} \lambda^T D^{-1}(\phi) \mathbf{W}_j (-2D(\mathbf{W}_j) D^{-3}(\phi) D(\lambda)) \\
&= c^{-2} D(\lambda) D^{-1}(\phi) \mathbf{M} \mathbf{W}_j \mathbf{W}_j^T D(\lambda) D^{-2}(\phi) \\
&\quad - c^{-1} D(\lambda) D^{-2}(\phi) \mathbf{W}_j \mathbf{W}_j^T D(\lambda) D^{-2}(\phi) \\
&\quad - 2c^{-1} \lambda^T D^{-3}(\phi) D(\lambda) \mathbf{W}_j^T D^{-1}(\phi) \lambda D(\mathbf{W}_j).
\end{aligned}$$

Substituindo $\frac{\partial}{\partial \phi} \{1\}$, $\frac{\partial}{\partial \phi} \{2\}$ em $\frac{\partial}{\partial \phi} \left(\frac{\partial}{\partial \phi^T} \|T_j\|^2 \right)$ obtemos,

$$\frac{\partial^2 \|T_j\|^2}{\partial \phi \partial \phi^T} = \frac{\partial}{\partial \phi} \{ -\mathbf{W}_j^T D^{-2}(\phi) D(\mathbf{W}_j) \} - \frac{\partial}{\partial \phi} \{1\} + 2 \frac{\partial}{\partial \phi} \{2\}$$

logo,

$$\begin{aligned}
\frac{\partial^2 \|T_j\|^2}{\partial \phi \partial \phi^T} &= 2D(\mathbf{W}_j) D^{-3}(\phi) D(\mathbf{W}_j) \\
&\quad - 2c^{-3} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j D(\lambda) D^{-1}(\phi) \mathbf{M} D^{-1}(\phi) D(\lambda) \\
&\quad + 2c^{-2} D(\lambda) D^{-2}(\phi) \mathbf{W}_j \mathbf{W}_j^T \mathbf{M} D^{-1}(\phi) D(\lambda) \\
&\quad + 2c^{-2} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j D(\lambda) D^{-3}(\phi) D(\lambda) \\
&\quad + 2c^{-2} D(\lambda) D^{-1}(\phi) \mathbf{M} \mathbf{W}_j \mathbf{W}_j^T D(\lambda) D^{-2}(\phi) \\
&\quad - 2c^{-1} D(\lambda) D^{-2}(\phi) \mathbf{W}_j \mathbf{W}_j^T D(\lambda) D^{-2}(\phi) \\
&\quad - 4c^{-1} D^{-3}(\phi) D(\lambda) \mathbf{W}_j^T D^{-1}(\phi) \lambda D(\mathbf{W}_j).
\end{aligned}$$

(4) Aplicando os resultados nos Lemas B1, A1 e A2, temos que

$$\begin{aligned}
\frac{\partial^2 \|T_j\|^2}{\partial \lambda \partial \mu^T} &= \frac{\partial}{\partial \lambda} \left\{ \frac{\partial}{\partial \mu^T} \|T_j\|^2 \right\} = \frac{\partial}{\partial \lambda} \left\{ -2\mathbf{W}_j^T D^{-1}(\phi) + 2c^{-1} \underbrace{\mathbf{W}_j^T D^{-1}(\phi) \lambda \lambda^T D^{-1}(\phi)}_{\text{número}} \right\} \\
&= \frac{\partial}{\partial \lambda} \left\{ -2\mathbf{W}_j^T D^{-1}(\phi) + 2c^{-1} \overbrace{\lambda^T D^{-1}(\phi) \mathbf{W}_j}^a \underbrace{\lambda^T D^{-1}(\phi)}_b \right\} \\
&= 2 \frac{\partial}{\partial \lambda} (c^{-1}) a + 2c^{-1} \frac{\partial}{\partial \lambda} (\lambda^T D^{-1}(\phi) \mathbf{W}_j) b + 2c^{-1} \lambda^T D^{-1}(\phi) \mathbf{W}_j \frac{\partial}{\partial \lambda} (\lambda^T D^{-1}(\phi)) \\
&= 2(-2c^{-2} D^{-1}(\phi) \lambda) \lambda^T D^{-1}(\phi) \mathbf{W}_j \lambda^T D^{-1}(\phi) + 2c^{-1} (D^{-1}(\phi) \mathbf{W}_j) \lambda^T D^{-1}(\phi) \\
&\quad + 2c^{-1} \lambda^T D^{-1}(\phi) \mathbf{W}_j (D^{-1}(\phi)) \\
&= -4c^{-2} \lambda^T D^{-1}(\phi) \mathbf{W}_j \mathbf{M} + 2c^{-1} D^{-1}(\phi) \mathbf{W}_j \lambda^T D^{-1}(\phi) + 2c^{-1} \lambda^T D^{-1}(\phi) \mathbf{W}_j D^{-1}(\phi) \\
&= 2c^{-1} \{ D^{-1}(\phi) \mathbf{W}_j \lambda^T D^{-1}(\phi) + \lambda^T D^{-1}(\phi) \mathbf{W}_j [\Sigma^{-1} - c^{-1} \mathbf{M}] \}.
\end{aligned}$$

(5) O resultado segue dos Lemas B1, A1 e A2.

$$\begin{aligned}
\frac{\partial^2 \|T_j\|^2}{\partial \phi \partial \mu^T} &= \frac{\partial}{\partial \phi} \left(\frac{\partial \|T_j\|}{\partial \mu^T} \right) = \frac{\partial}{\partial \phi} (-2\mathbf{W}_j D^{-1}(\phi) + 2c^{-1} \mathbf{W}_j^T \mathbf{M}) \\
&= -2 \frac{\partial}{\partial \phi} (\mathbf{W}_j^T D^{-1}(\phi)) + 2 \frac{\partial}{\partial \phi} (c^{-1}) \mathbf{W}_j^T \mathbf{M} + 2c^{-1} \frac{\partial}{\partial \phi} (\mathbf{W}_j^T \mathbf{M}) \\
&= -2(-D(\mathbf{W}_j) D^{-2}(\phi)) + 2(-c^{-2} D(\lambda) D^{-2}(\phi) \lambda) \mathbf{W}_j^T \mathbf{M} \\
&\quad + 2c^{-1} (-D(\lambda) D^{-2}(\phi) \mathbf{W}_j \lambda^T D^{-1}(\phi) - \mathbf{W}_j^T D^{-1}(\phi) \lambda D(\lambda) D^{-2}(\phi)) \\
&= 2D(\mathbf{W}_j) D^{-2}(\phi) + 2c^{-2} D(\lambda) D^{-2}(\phi) \lambda \mathbf{W}_j^T \mathbf{M} \\
&\quad - 2c^{-1} D(\lambda) D^{-2}(\phi) \mathbf{W}_j \lambda^T D^{-1}(\phi) - 2c^{-1} \mathbf{W}_j^T D^{-1}(\phi) \lambda D(\lambda) D^{-2}(\phi).
\end{aligned}$$

(6) O resultado é obtido após a aplicação dos Lemas B1, A1 e A2.

$$\frac{\partial^2 \|T_j\|^2}{\partial \phi \partial \lambda^T} = \frac{\partial}{\partial \phi} \left(\frac{\partial \|T_j\|^2}{\partial \lambda^T} \right) = \frac{\partial}{\partial \phi} (2c^{-2} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j \lambda^T D^{-1} - 2c^{-1} \lambda^T D^{-1}(\phi) \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi))$$

$$= 2 \frac{\partial}{\partial \phi} \left(\underbrace{c^{-2} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j \boldsymbol{\lambda}^T D^{-1}(\phi)}_{(1)} \right) - 2 \frac{\partial}{\partial \phi} \left(\underbrace{c^{-1} \boldsymbol{\lambda}^T D^{-1}(\phi) \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi)}_{(2)} \right)$$

Achando $\frac{\partial(1)}{\partial \phi}$,

$$\begin{aligned} \frac{\partial(1)}{\partial \phi} &= \frac{\partial}{\partial \phi} (c^{-2}) \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j \boldsymbol{\lambda}^T D^{-1}(\phi) + c^{-2} \frac{\partial}{\partial \phi} (\mathbf{W}_j^T \mathbf{M} \mathbf{W}_j) \boldsymbol{\lambda}^T D^{-1}(\phi) \\ &\quad + c^{-2} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j \frac{\partial}{\partial \phi} (\boldsymbol{\lambda}^T D^{-1}(\phi)) \\ &= (2c^{-3} D(\boldsymbol{\lambda}) D^{-2}(\phi) \boldsymbol{\lambda}) \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j \boldsymbol{\lambda}^T D^{-1}(\phi) \\ &\quad + c^{-2} (-2 \boldsymbol{\lambda}^T D^{-1}(\phi) \mathbf{W}_j D(\boldsymbol{\lambda}) D^{-2}(\phi) \mathbf{W}_j) \boldsymbol{\lambda}^T D^{-1}(\phi) + c^{-2} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j (-D(\boldsymbol{\lambda}) D^{-2}(\phi)) \\ &= 2c^{-3} D(\boldsymbol{\lambda}) D^{-2}(\phi) \boldsymbol{\lambda} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j \boldsymbol{\lambda}^T D^{-1}(\phi) - 2c^{-2} \boldsymbol{\lambda}^T D^{-1}(\phi) \mathbf{W}_j D(\boldsymbol{\lambda}) D^{-2}(\phi) \mathbf{W}_j \boldsymbol{\lambda}^T D^{-1}(\phi) \\ &\quad - c^{-2} (\mathbf{W}_j^T \mathbf{M} \mathbf{W}_j) D(\boldsymbol{\lambda}) D^{-2}(\phi) \\ &= 2c^{-3} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j \mathbf{M} D^{-1}(\phi) D(\boldsymbol{\lambda}) - 2c^{-2} D(\boldsymbol{\lambda}) D^{-2}(\phi) \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi) \boldsymbol{\lambda} \boldsymbol{\lambda}^T D^{-1}(\phi) \\ &\quad - c^{-2} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j D^{-2}(\phi) D(\boldsymbol{\lambda}). \end{aligned}$$

Achando $\frac{\partial(2)}{\partial \phi}$,

$$\begin{aligned} \frac{\partial(2)}{\partial \phi} &= \frac{\partial}{\partial \phi} \left(c^{-1} \underbrace{\boldsymbol{\lambda}^T D^{-1}(\phi) \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi)}_A \right) \\ &= \frac{\partial(c^{-1})}{\partial \phi} A + c^{-1} \frac{\partial}{\partial \phi} (\boldsymbol{\lambda}^T D^{-1}(\phi) \mathbf{W}_j) \mathbf{W}_j^T D^{-1}(\phi) \\ &\quad + c^{-1} \boldsymbol{\lambda}^T D^{-1}(\phi) \mathbf{W}_j \frac{\partial}{\partial \phi} (\mathbf{W}_j^T D^{-1}(\phi)) \\ &= (c^{-2} D(\boldsymbol{\lambda}) D^{-2}(\phi) \boldsymbol{\lambda}) A + c^{-1} (-D(\boldsymbol{\lambda}) D^{-2}(\phi) \mathbf{W}_j) \mathbf{W}_j^T D^{-1}(\phi) \\ &\quad + c^{-1} \boldsymbol{\lambda}^T D^{-1}(\phi) \mathbf{W}_j (-D(\mathbf{W}_j) D^{-2}(\phi)) \\ &= c^{-2} D(\boldsymbol{\lambda}) D^{-1}(\phi) \mathbf{M} \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi) - c^{-1} D(\boldsymbol{\lambda}) D^{-2}(\phi) \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi) \\ &\quad - c^{-1} \mathbf{W}_j^T D^{-1}(\phi) \boldsymbol{\lambda} D(\mathbf{W}_j) D^{-2}(\phi). \end{aligned}$$

substituindo $\frac{\partial(1)}{\partial \phi}$, $\frac{\partial(2)}{\partial \phi}$ em $\frac{\partial \|T_j\|^2}{\partial \phi \partial \boldsymbol{\lambda}^T}$ obtemos, o resultado desejado,

$$\frac{\partial^2 \|T_j\|^2}{\partial \phi \partial \boldsymbol{\lambda}^T} = 4c^{-3} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j D^{-1}(\phi) D(\boldsymbol{\lambda}) \mathbf{M}$$

$$\begin{aligned}
& -4c^{-2}D(\lambda)D^{-2}(\phi)W_jW_j^TM \\
& -2c^{-2}W_j^TMW_jD(\lambda)D^{-2}(\phi) \\
& -2c^{-2}D(\lambda)D^{-1}(\phi)MW_jW_j^TD^{-1}(\phi) \\
& +2c^{-1}D(\lambda)D^{-2}(\phi)W_jW_j^TD^{-1}(\phi) \\
& +2c^{-1}W_j^TD^{-1}(\phi)\lambda D(W_j)D^{-2}(\phi).
\end{aligned}$$

No seguinte resultado apresentaremos a segunda derivada de $\log|\Sigma|$ com respeito $\theta = (\mu^T, \lambda^T, \phi^T)$, isto é

$$\begin{aligned}
\frac{\partial^2 \log|\Sigma|}{\partial \theta \partial \theta^T} &= \frac{\partial}{\partial \theta} \left(\frac{\partial \log|\Sigma|}{\partial \theta^T} \right) = \frac{\partial}{\partial \theta} \left[\frac{\partial \log|\Sigma|}{\partial \mu^T}, \frac{\partial \log|\Sigma|}{\partial \lambda^T}, \frac{\partial \log|\Sigma|}{\partial \phi^T} \right] \\
&= \begin{pmatrix} \frac{\partial^2 \log|\Sigma|}{\partial \mu \partial \mu^T} & \frac{\partial^2 \log|\Sigma|}{\partial \mu \partial \lambda^T} & \frac{\partial^2 \log|\Sigma|}{\partial \mu \partial \phi^T} \\ \frac{\partial^2 \log|\Sigma|}{\partial \lambda \partial \mu^T} & \frac{\partial^2 \log|\Sigma|}{\partial \lambda \partial \lambda^T} & \frac{\partial^2 \log|\Sigma|}{\partial \lambda \partial \phi^T} \\ \frac{\partial^2 \log|\Sigma|}{\partial \phi \partial \mu^T} & \frac{\partial^2 \log|\Sigma|}{\partial \phi \partial \lambda^T} & \frac{\partial^2 \log|\Sigma|}{\partial \phi \partial \phi^T} \end{pmatrix},
\end{aligned}$$

onde $|\Sigma| = c|D(\phi)|$, $c = 1 + \lambda^T D^{-1}(\phi)\lambda$, tomando logaritmo a $|\Sigma|$ obtemos que, $\log|\Sigma| = \log c|D(\phi)| = \log c + \log|D(\phi)| = \log|\Sigma| = \log c + \log(\prod_{j=1}^p \phi_j)$.

A seguir são apresentados os cálculos necessários para achar a matriz anterior,

Lema B3: *Seja $\Sigma = \lambda\lambda^T + D^{-1}(\phi)$. Então*

- 1) $\frac{\partial \log|\Sigma|}{\partial \mu} = 0$
- 2) $\frac{\partial \log|\Sigma|}{\partial \lambda} = 2c^{-1}D^{-1}(\phi)\lambda$
- 3) $\frac{\partial \log|\Sigma|}{\partial \phi} = -c^{-1}D^{-2}(\phi)D(\lambda)\lambda + D^{-1}(\phi)1$
- 4) $\frac{\partial^2 \log|\Sigma|}{\partial \mu \partial \mu^T} = \frac{\partial^2 \log|\Sigma|}{\partial \mu \partial \lambda^T} = \frac{\partial^2 \log|\Sigma|}{\partial \mu \partial \phi^T} = 0$
- 5) $\frac{\partial^2 \log|\Sigma|}{\partial \lambda \partial \lambda^T} = 2c^{-1}(\Sigma^{-1} - c^{-1}M)$
- 6) $\frac{\partial^2 \log|\Sigma|}{\partial \phi \partial \phi^T} = -D^{-2}(\phi) + 2c^{-1}D^{-3}(\phi)D(\lambda)D(\lambda) - c^{-2}D^{-1}(\phi)D(\lambda)MD(\lambda)D^{-1}(\phi)$

$$7) \frac{\partial^2 \log |\Sigma|}{\partial \lambda \partial \phi^T} = -2c^{-1} \Sigma^{-1} D^{-1}(\phi) D(\lambda).$$

Prova: Como Σ não depende de μ , a prova de (1) e (4) é imediata.

(2) aplicando o Lema A1, temos que

$$\begin{aligned} \frac{\partial \log |\Sigma|}{\partial \lambda} &= \frac{\partial}{\partial \lambda} (\log c + \log (\Pi_{j=1}^p \phi_j)) = c^{-1} \frac{\partial}{\partial \lambda} (c) = c^{-1} (2D^{-1}(\phi) \lambda) \\ &= 2c^{-1} D^{-1}(\phi) \lambda \end{aligned}$$

(3) Da definição A2 e do lema A1, temos que

$$\begin{aligned} \frac{\partial \log |\Sigma|}{\partial \phi} &= \frac{\partial}{\partial \phi} (\log c + \log (\Pi_{j=1}^p \phi_j)) \\ &= c^{-1} (-D(\lambda) D^{-2}(\phi) + D^{-1}(\phi) \mathbf{1}) \\ &= -c^{-1} D^{-2}(\phi) D(\lambda) \lambda + D^{-1}(\phi) \mathbf{1} \end{aligned}$$

(5) De (2) e dos Lemas A1 e A2, temos que

$$\begin{aligned} \frac{\partial^2 \log |\Sigma|}{\partial \lambda \partial \lambda^T} &= \frac{\partial}{\partial \lambda} \left(\frac{\partial \log |\Sigma|}{\partial \lambda^T} \right) = \frac{\partial}{\partial \lambda} (2c^{-1} \lambda^T D^{-1}(\phi)) \\ &= 2 \frac{\partial}{\partial \lambda} (c^{-1}) \lambda^T D^{-1}(\phi) + 2c^{-1} \frac{\partial}{\partial \lambda} (\lambda^T D^{-1}(\phi)) \\ &= 2 (-2c^{-2} D^{-1}(\phi) \lambda) \lambda^T D^{-1}(\phi) + 2c^{-1} (D^{-1}(\phi)) \\ &= -4c^{-2} \underbrace{D^{-1}(\phi) \lambda \lambda^T D^{-1}(\phi)}_{\mathbf{M}} + 2c^{-1} D^{-1}(\phi) \\ &= 2c^{-1} \left(\underbrace{-c^{-1} \mathbf{M} + D^{-1}(\phi)}_{\Sigma^{-1}} - c^{-1} \mathbf{M} \right) = 2c^{-1} (\Sigma^{-1} - c^{-1} \mathbf{M}). \end{aligned}$$

(6) De (3) e dos Lemas A1 e A2, segue que

$$\begin{aligned} \frac{\partial^2 \log |\Sigma|}{\partial \phi \partial \phi^T} &= \frac{\partial}{\partial \phi} \left(\frac{\log |\Sigma|}{\phi^T} \right) = -\frac{\partial}{\partial \phi} (c^{-1}) \lambda^T D^{-2}(\phi) D(\lambda) \\ &\quad - c^{-1} \frac{\partial}{\partial \phi} (\lambda^T D^{-2}(\phi) D(\lambda)) + \frac{\partial}{\partial \phi} (\mathbf{1}^T D^{-1}(\phi)) \\ &= - (c^{-2} D(\lambda) D^{-2}(\phi) \lambda) \lambda^T D^{-2}(\phi) D(\lambda) - c^{-1} (-2D(\lambda) D^{-3}(\phi) D(\lambda)) \end{aligned}$$

$$\begin{aligned}
& + \left(- \underbrace{D(\mathbf{1})}_{\mathbf{I}} D^{-2}(\phi) \right) \\
& = -c^{-2} D(\lambda) D^{-2}(\phi) \lambda \lambda^T D^{-2}(\phi) + 2c^{-1} D(\lambda) D^{-3}(\phi) D(\lambda) - D^{-2}(\phi) \\
& = -D^{-2}(\phi) + 2c^{-1} D^{-3}(\phi) D(\lambda) D(\lambda) - c^{-2} D^{-1}(\phi) D(\lambda) \mathbf{M} D(\lambda) D^{-1}(\phi).
\end{aligned}$$

(7) De (3) e dos Lemas A1 e da definição A2, segue que

$$\begin{aligned}
\frac{\partial^2 \log |\Sigma|}{\partial \lambda \partial \phi^T} &= \frac{\partial}{\partial \lambda} \left(\frac{\partial \log |\Sigma|}{\partial \phi^T} \right) = \frac{\partial}{\partial \lambda} (-c^{-1} \lambda^T D(\lambda) D^{-2}(\phi) + \mathbf{1}^T D^{-1}(\phi)) \\
&= -\frac{\partial}{\partial \lambda} (c^{-1}) \lambda^T D(\lambda) D^{-2}(\phi) - c^{-1} \frac{\partial}{\partial \lambda} (\lambda^T D(\lambda) D^{-2}(\phi)) \\
&= -(-2c^{-2} D^{-1}(\phi) \lambda) \lambda^T D(\lambda) D^{-2}(\phi) - c^{-1} (2D(\lambda) D^{-2}(\phi)) \\
&= 2c^{-2} \underbrace{D^{-1}(\phi) \lambda \lambda^T D^{-1}(\phi)}_{\mathbf{M}} D^{-1}(\phi) D(\lambda) - 2c^{-1} D(\lambda) D^{-2}(\phi) \\
&= -2c^{-1} \left(\underbrace{D^{-1}(\phi) - c^{-1} \mathbf{M}}_{\Sigma^{-1}} \right) D^{-1}(\phi) D(\lambda) = -2c^{-1} \Sigma^{-1} D^{-1}(\phi) D(\lambda).
\end{aligned}$$

Finalmente, de (B.1), (B.2) e dos Lemas B2 e B3, obtemos a matriz de informação observada, onde

$$\frac{\partial^2 l_j}{\partial \mu \partial \mu^T} = -D^{-1}(\phi) + c^{-1} \mathbf{M}$$

$$\frac{\partial^2 l_j}{\partial \lambda \partial \mu^T} = -c^{-1} \{ D^{-1}(\phi) \mathbf{W}_j \lambda^T D^{-1}(\phi) + \lambda^T D^{-1}(\phi) \mathbf{W}_j [\Sigma^{-1} - c^{-1} \mathbf{M}] \}$$

$$\begin{aligned}
\frac{\partial^2 l_j}{\partial \phi \partial \mu^T} &= -D(\mathbf{W}_j) D^{-2}(\phi) - c^{-2} D(\lambda) D^{-2}(\phi) \lambda \mathbf{W}_j^T \mathbf{M} \\
&\quad + c^{-1} D(\lambda) D^{-2}(\phi) \mathbf{W}_j \lambda^T D^{-1}(\phi) + c^{-1} \mathbf{W}_j^T D^{-1}(\phi) \lambda D(\lambda) D^{-2}(\phi)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 l_j}{\partial \lambda \partial \lambda^T} &= -c^{-1} (\Sigma^{-1} - c^{-1} \mathbf{M}) + c^{-1} D^{-1}(\phi) \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi) - 2c^{-2} D^{-1}(\phi) \mathbf{W}_j \mathbf{W}_j^T \mathbf{M} \\
&\quad - c^{-2} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j \left(\underbrace{D^{-1}(\phi) - c^{-1} \mathbf{M}}_{\Sigma^{-1}} - 3c^{-1} \mathbf{M} \right) - 2c^{-2} \mathbf{M} \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 l_j}{\partial \phi \partial \lambda^T} &= c^{-1} D^{-1}(\phi) D(\lambda) \Sigma^{-1} - 2c^{-3} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j D^{-1}(\phi) D(\lambda) \mathbf{M} \\
&\quad + 2c^{-2} D(\lambda) D^{-2}(\phi) \mathbf{W}_j \mathbf{W}_j^T \mathbf{M} + c^{-2} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j D(\lambda) D^{-2}(\phi) \\
&\quad + c^{-2} D(\lambda) D^{-1}(\phi) \mathbf{M} \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi) - c^{-1} D(\lambda) D^{-2}(\phi) \mathbf{W}_j \mathbf{W}_j^T D^{-1}(\phi) \\
&\quad - c^{-1} \mathbf{W}_j^T D^{-1}(\phi) \lambda D(\mathbf{W}_j) D^{-2}(\phi)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 l_j}{\partial \phi \partial \phi^T} &= -\frac{1}{2} \{ -D^{-2}(\phi) + 2c^{-1} D^{-3}(\phi) D(\lambda) D(\lambda) - c^{-2} D^{-1}(\phi) D(\lambda) \mathbf{M} D(\lambda) D^{-1}(\phi) \\
&\quad + 2D(\mathbf{W}_j) D^{-3}(\phi) D(\mathbf{W}_j) - 2c^{-3} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j D(\lambda) D^{-1}(\phi) \mathbf{M} D^{-1}(\phi) D(\lambda) \\
&\quad + 2c^{-2} D(\lambda) D^{-2}(\phi) \mathbf{W}_j \mathbf{W}_j^T \mathbf{M} D^{-1}(\phi) D(\lambda) + 2c^{-2} \mathbf{W}_j^T \mathbf{M} \mathbf{W}_j D(\lambda) D^{-3}(\phi) D(\lambda) \\
&\quad + 2c^{-2} D(\lambda) D^{-1}(\phi) \mathbf{M} \mathbf{W}_j \mathbf{W}_j^T D(\lambda) D^{-2}(\phi) - 2c^{-1} D(\lambda) D^{-2}(\phi) \mathbf{W}_j \mathbf{W}_j^T D(\lambda) D^{-2}(\phi) \\
&\quad - 4c^{-1} D^{-3}(\phi) D(\lambda) \mathbf{W}_j^T D^{-1}(\phi) \lambda D(\mathbf{W}_j) \}.
\end{aligned}$$

Bibliografia

- [1] Abdullah, M. B.(1995). Detection of influential observations in functional errors-in-variables model. *Communication in Statistics - Theory and Methods* 24, 6, 1585-1595.
- [2] Anscombe, F. J. (1973). Graphs in statistical analysis. *The American Statistician* 27, 17-21.
- [3] Barnett, V.D. (1969). Simultaneous pairwise linear structural relationships. *Biometrics* 25, 129-142.
- [4] Beckman, R.J., Natchtsheim, C. J. e Cook, R. D.(1987). Diagnostics for mixed-model analysis of variance. *Technometrics* 29, 4, 413-426.
- [5] Belsley, D. A., Kuh, E. e Welsch, R. E. (1980). *Regression Diagnostics*. John Wiley. New York.
- [6] Billor, N. e Loynes, R.M. (1993). Local influence: A new approach. *Communication in Statistics - Theory and Methods* 22, 6, 1595-1611.
- [7] Bolfarine, H. e Galea-Rojas, M.(1995). Functional comparative calibration using on EM algorithm - comment. *Biometrics*, 51, 4, 1579-1580.
- [8] Bolfarine, H. e Galea, M.(1995 b). Maximum likelihood estimation of simultaneous pairwise linear structural relationships. *Biometrical Journal*, 37, 673-689.

Functional comparative calibration using on EM algorithm - comment. *Biometrics*, 51, 4, 1579-1580.
- [9] Carter, R.(1981). Restricted maximum likelihood estimation of several bias and reliability in the comparison measuring methods. *Biometrics* 37, 733-741.
- [10] Chatterjee, S. e Hadi, A . S. (1988). *Sensitivity Analysis in Linear Regression*. John Wiley. New York.

- [11] Chipkevitch, E., Nishimura, R.T., Tu, D.G.S. e Galea-Rojas, M. Clinical measurement of testicular volume in adolescents: Comparison of the reliability of 5 methods. *Journal of Urology* 156, 6, 2050-2053.
- [12] Cochran, W. G. (1968). Errors of measurement in statistics. *Technometrics* 10, 637-666.
- [13] Cook, R.D. (1977). Detection of influential observation in linear regression. *Technometrics* 19, 15-18.
- [14] Cook, R.D. (1979). Influential observations in linear regression. *Journal of the American Statistical Association*, 74, 365,169-174.
- [15] Cook, R.D. (1986). Assessment of local influence. *Journal of the Royal Statistical Society B* 48, 133-169.
- [16] Cook, R.D. (1987). Influence assessment. *Journal of Applied Statistics* 14, 117-13.
- [17] Cook, R.D. e Weisberg, S.(1982). *Residuals and Influence in Regression*. Chapman and Hall. London.
- [18] Cook e Peña(1988). The likelihood displacement: A unifying principle for influence measures. *Communication in Statistics - Theory and Methods* 17, 3, 623-640.
- [19] Cook, R.D. e Wang, P.C.(1983). Transformations and influential cases in regression. *Technometrics* 25, 337-344.
- [20] Cook, R.D., Peña, D. e Weisberg, S. (1988). The Likelihood displacement: A unifying principle for influence measures. *Communications in Statistics - Theory and Methods* 17, 3, 623-640.
- [21] Cox, D.R. e Hinkley, D. V. (1974). *Theorical Statistics*: Chapman and Hal. Londonl.

- [22] Draper, N. e Guttman, I. (1975). Two simultaneous measurement procedures: A bayesian approach. *Journal of the American Statistical Association* 70, 43-46.
- [23] Draper, N. e Guttman, I. (1976). Simultaneous measurement procedures: A bayesian approach. *Journal of the American Statistical Association* 71, 605-607.
- [24] Dunn, G. (1992). Design and analysis of reliability: *The statistical evaluation of measurement errors*. Edward Arnold. New York.
- [25] Emerson, J.D., Hoaglin, D.C. e Kempthorne, P.J.(1984). Leverage in least squares additive-plus-multiplicative fits for two-way tables. *Journal of the American Statistical Association* 79, 386, 329-335.
- [26] Escobar, L. A. e Meeker, W.Q. (1992). Assessing influence in regression analysis with censored data. *Biometrics* 48, 507-528.
- [27] Fuller, W. A. (1987). *Measurement error models*. John Wiley. New York.
- [28] Fung, W.K. e Tang, M.K.(1997). Assessment of local influence in multivariate regression analysis. *Communications in Statistics - Theory and Methods* 26, 4, 821-837.
- [29] Galea, M., Paula, G.A. e Bolfarine, H. (1997). Local Influence in elliptical regression models. *The Statistician* 46, 1, 71-79.
- [30] Galea-Rojas, M.(1995). *Calibração Estrutural e Funcional*. Tese de doutorado. IME-USP.
- [31] Galea, M., Paula, G. A. e Bolfarine, H. (1997). Local influence in elliptical linear regression models. *The Statistician* 46, 71-79.
- [32] Gabrielsen, A. (1978). Consistency and identifiability. *Journal of Econometrics* 8, 261-263.

- [33] Graybill, F. A. (1983). *Matrices with Applications in Statistics*. Second Edition. Wadsworth. California.
- [34] Grubbs, F. E. (1948). On estimating precision of measuring instruments and product variability. *Journal of the American Statistical Association* 43, 243-264.
- [35] Grubbs, F.E. (1973). Errors of measurement precision, accuracy and the statistical comparison of measuring instruments. *Technometrics* 15, 53-66.
- [36] Hoaglin, D. C. e Welsch, R. E. (1978). The hat matrix regression and ANOVA. *The American Statistician* 32, 17-22.
- [37] Jaech, J.L. (1964). A program to estimate measurement errors in nondestructive avaluation of fuel element quality. *Technometrics* 6, 293-300.
- [38] Jaech, J.L. (1985). Statistical analysis of measurement errors. *Exxon Monographs*. John Wiley. New York.
- [39] Jiancheng, J. e Xizhi, W. (1998). Assessment of local influence on Cox's proportional Hazards model. *Acta Mathematicae Applicatae Sinica* 14, 4, 414-424.
- [40] Johnson, W. e Geisser, S.(1983). A predictive view of the detection and characterization of influential observations in regression analysis. *Journal of the American Statistical Association* 78, 137-144.
- [41] Johnson, N. e Kotz, S. (1970). Distributions in statistics: *Continuous Univariate Distributions* - Vol. 2. Houghton Mifflin. New York.
- [42] Jonhnsen, R. A. e Wichern, D. W. (1998). *Applied Multivariate Statistical Analysis*. Fourth Edition. Prentice Hall.New Jersey.
- [43] Jorgensen, B. (1997a). Proper dispersion models. *Brazilian Journal of Probability and Statistics* 11, 2, 89-140.

- [44] Jorgensen, B. (1997b). *The Theory of Dispersion Models*. Chapman and Hall. London
- [45] 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.
- [46] Kelly, G. (1984). The influence function in the errors in variables problem. *Annals of Statistics* 12, 87-100.
- [47] Kelly, G. (1985). Use of the structural equations model in assessing the reliability of a new measurement technique. *Applied Statistics* 34, 258-263.
- [48] Kendall, M.G. e Stuart, A. (1979). *The Advanced Theory of Statistics*, Vol. 2, 4th ed. Hafner. New York.
- [49] Kim, C. (1996). Local influence and replacement measure. *Communications in Statistics - Theory and Methods* 25, 1, 49-61.
- [50] Kim, M.G. (1995). Local influence in multivariate regression. *Communications in Statistics - Theory and Methods* 20, 1271-1278.
- [51] Kimura, D.K. (1992). Functional comparative calibration using an EM algorithm. *Biometrics* 48, 1263-1271.
- [52] Kwan, C. W. e Fung, W. K. (1998). Assessing local influence for specific restricted likelihood: Application to fator analysis. *Psychometrika* 63, 1, 35-46.
- [53] Lange, K.L., Little, R.J. and Taylor, J. (1989). Robust statistical modelling using the *t*-distribution. *Journal of the American Statistical Association* 84, 881-896.
- [54] Lasafre, E. e Verbeke, G.(1998). Local influence in linear mixed models. *Biometrics* 54, 570-582.

- [55] Laurent, R.T. e Cook, R.D. (1993). Leverage, local influence and curvature in nonlinear regression. *Biometrika* 80, 1, 99-106.
- [56] Lavine, M. (1992). Local predictive influence in bayesian linear models with conjugate priors. *Communications in Statistics - Simulation* 21, 269-283.
- [57] Lawrence, A. J. (1988). Regression transformation diagnostic using local influence. *Journal of the American Statistical Association* 83, 1067-1072.
- [58] Lee, A.H. e Zhao, Y. (1996). Assessing local influence in measurement error models. *Biometrical Journal* 38, 7, 829-841.
- [59] Lesaffre, E. e Verbeke G. (1998). Local influence in linear mixed models. *Biometrics* 54, 570-582.
- [60] Leurgans, S. (1980). Evaluating laboratory measurement techniques. In *Biostatistics Casebook*. Edited by Miller, R., Efron, B., Brown, B. and Moses, L. John Wiley. New York.
- [61] Little, R. J. (1988). Robust estimation of the mean and covariance matrix from data with missing values. *Applied Statistics* 37, 23-38.
- [62] Mak, T.K. (1982). Estimation in the presence of incidental parameters. *The Canadian Journal of Statistics* 10, 121-132.
- [63] Nice, C. (1996). *Comparação de instrumentos de medição usando calibração comparativa*. Tese de Mestrado. IME-USP.
- [64] Pan, J., Fang, K., e Liski, E. (1996). Bayesian local influence for the growth curve model with Rao's simple covariance structure. *Journal of Multivariate Analysis* 58, 55-81.
- [65] Paula, G. A. (1993). Assessing local influence in restricted regressions models. *Computational Statistics and Data Analysis* 16, 63-79.

- [66] Paula, G. A. (1995). Influence and residuals in restricted generalized linear models. *Journal of Statistical Computation and Simulation* 51, 315-331.
- [67] Paula, G. A. (1996). Influence diagnostics in proper dispersion models. *Australian Journal of Statistics* 38, 307-316.
- [68] Pregibon, D. (1981). Logistic regression diagnostics. *The Annals of Statistics* 9, 705-724.
- [69] Polasek, W. (1984). Regression diagnostics for general linear regression models. *Journal of the American Statistical Association* ,79 386, 336-340.
- [70] Rao, C. R. (1973). *Linear Statistical Inference and its applications*. 2nd. edition. New York: John Wiley.
- [71] Rundle, A. T. e Sylvester, P. E. (1962). Measurement of testicular volume. Its application to assessment of maturation, and its use in diagnosis of hypogonadism. *Arch. Dis. Child.*, 37:514.
- [72] Sen, P.K. e Singer, J. M. (1993). *Large Sample Methods in Statistics*. An Introduction with Applications. Chapman e Hall. New York.
- [73] Schonfeld, W. A. (1943). Primary and secondary sexual characteristics: study of their development in males from birth through maturity, with biometric study of penis and testes. *Amer. J. Dis. Child.*, 65: 535.
- [74] 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.
- [75] Stoker, J.J. (1969). *Differential Geometry*. John Wiley. New York.
- [76] Tanaka, Y., Watadani, S. e Moon, S. H.(1991). Influence in covariance structure analysis with an application to confirmatory factor analysis. *Communication in Statistic - Theory and Methods* 20, 12, 3805-3821.

- [77] Tanner, J. M. (1962). Growth at Adolescence, 2nd ed. Oxford: Blackwell Scientific Publications.
- [78] Theobald, C. M. e Mallison, J. R. (1978). Comparative calibration, linear structural relationship and congeneric measurements. *Biometrics* 34, 39-45.
- [79] Thomas, W. e Cook, R. D.(1989). Assessing influence on regression coefficients in generalized linear models. *Biometrika* 76, 741-749.
- [80] Thomas, W. e Cook, R.D. (1990). Assesing influence on predictions from generalized linear models. *Technometrics* 32, 1, 59-65.
- [81] Tsai, C. L. e Wu, X. (1992). Transformation-model diagnostics. *Technometrics* 34, 197-202.
- [82] Vilca-Labra, F., Arellano-Valle, R.B., Bolfarine, H.(1998). Elliptical functional models. *Journal of Multivariate Analysis* 65, 1, 36-57.
- [83] Weisberg, S. (1985). *Applied Linear Regression*. 2nd edition. John Wiley. New York: John
- [84] Williams, E. J. (1969). Regression methods in calibration problems. *Bulletin of the International Statistical Institute* 43, 17-28.
- [85] Zhao, Y. e Lee, H.(1996). Assessing local influence in measurement error models. *Biometrika* 38, 7, 829-841,
- [86] Zhao, Y. e Lee, H.(1998). Influence diagnostics for simultaneous equations models. *Australian Statistical Publishing Association Inc.* 345-357.