
Universidade Federal de Minas Gerais – UFMG
Instituto de Ciências Exatas – ICEX
Departamento de Estatística
Programa de Pós-graduação em Estatística

*MODELOS DE REGRESSÃO NORMAL
INDEPENDENTE COM ERROS DE
MEDIDA E DADOS CENSURADOS[†]*

Tese de Doutorado

Alejandro Guillermo Monzón Montoya

Julho de 2018
Belo Horizonte - MG

[†] O presente trabalho foi realizado com apoio do Programa Estudantes-Convênio de Pós-Graduação – PEC-PG, da CAPES/CNPq - Brasil.

MODELOS DE REGRESSÃO NORMAL INDEPENDENTE COM ERROS DE MEDIDA E DADOS CENSURADOS

Alejandro Guillermo Monzón Montoya

Orientadora: Lourdes Coral Contreras Montenegro

Tese apresentada ao Programa de Pós-graduação em Estatística do Instituto de Ciências Exatas da Universidade Federal de Minas Gerais como requisito parcial para obtenção do título de Doutor em Estatística.

Departamento de Estatística
Instituto de Ciências Exatas
Universidade Federal de Minas Gerais

Belo Horizonte, MG - Brasil

Julho de 2018

Modelos de regressão normal independente com erros de medida e dados censurados

Esta versão da tese contém as correções e alterações sugeridas pela banca durante a defesa do trabalho realizada em 12 de julho de 2018.

Banca Examinadora:

Profa. Dra. Lourdes Coral Contreras Montenegro (Orientadora) – UFMG

Prof. Dr. Marcos Oliveira Prates – UFMG

Prof. Dr. Cristiano de Carvalho Santos – UFMG

Profa. Dra. Camila Borelli Zeller – UFJF

Prof. Dr. Gustavo Henrique Mitraud Assis Rocha – ENCE, Br

DEDICATÓRIA

*A **DEUS**, pela presença constante na minha vida e por ter me dado força e esperança nos momentos difíceis desta caminhada.*

A minha família, que está sempre comigo, sendo minha força e inspiração constante.

*Aos meus pais (in Memoriam), **OLINDA** e **LEANDRO**, meus melhores exemplos de luta e perseverança na vida.*

Agradecimentos

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio financeiro concedido durante todo o período de estudos, sem o qual não teria sido possível a realização de meus estudos de Doutorado.

A minha orientadora e amiga, Lourdes Coral Contreras Montenegro, pelo tempo, dedicação e ajuda constante.

Aos professores Camila Borelli Zeller, Gustavo Henrique Mitraud Assis Rocha, Marcos Oliveira Prates e Cristiano de Carvalho Santos, pelas correções e sugestões.

Aos professores do departamento de Estatística do ICEX e aos funcionários pela atenção, o carinho e a amizade.

Aos colegas e amigos do DEST pelas conversas, amizade, incentivo e por me fazer sentir como em casa, em especial ao Rumenick, Erick, Fernando, Gabriela Oliveira, Juliana, Fernanda, Zaida, Wagner Pinheiro, Guilherme Oliveira, Uriel e Victor.

Ao Brasil, minha gratidão eterna.

Resumo

Modelos com erros de medida (*MEM*) são úteis para descrever diferentes fenômenos em diversas áreas do conhecimento. São utilizados para comparar dispositivos de medição que variam em custo, tempo e eficiência. Embora vários modelos considerem a existência de covariáveis mal medidas, muitos deles não consideram observações censuradas para a variável resposta. Por outro lado, isto é fundamental uma vez que em vários estudos a resposta observada está sujeita a limites de detecção máximos e/ou mínimos. Neste contexto, estendemos o trabalho de Matos et al. (2016), que desenvolveram a estimação dos parâmetros do modelo com erros de medida multivariado usando a distribuição *t*-Student com observações censuradas, a uma classe mais geral de distribuições normal independente (*t*-Student multivariado e slash multivariado). Além de desenvolvermos os procedimentos de estimação e inferência robusta, no sentido de utilizar uma distribuição que acomode observações *outliers* de forma mais eficiente do que a distribuição normal, também realizamos um estudo de diagnóstico de influência global e local utilizando a metodologia proposta por Zhu e Lee (2001).

Palavras-chave: Algoritmo EM, dados censurados, distribuição normal independente, modelos com erros de medida.

Abstract

Measurement error models (*MEM*) are useful for describing different phenomena in several areas of knowledge. They are used to compare measuring devices that vary in cost, time and efficiency. Although several models consider the existence of poorly measured covariates, many of them do not consider censored observations for the response variable. On the other hand, this is fundamental since in several studies the observed response is subject to maximum and/or minimum detection limits. In this context, we extend the work of Matos et al. (2016), who developed the estimation of parameters of the model with a multivariate measurement error by using the Student-*t* distribution with censored observations, to a more general class of independent normal distributions (multivariate Student-*t* and multivariate slash). In addition to developing robust estimation and inference procedures in order to use a distribution that more efficiently accommodates outliers observations than the normal distribution, we also carry out a diagnostic study of global influence and local influence using the methodology proposed by Zhu e Lee (2001).

Keywords: Censored data, EM algorithm, independent normal distribution, measurement error models.

Sumário

1	Introdução	1
1.1	Introdução	1
1.2	Proposta da tese e objetivos	3
1.3	Organização do trabalho	4
1.4	Aspectos preliminares	4
1.4.1	Distribuições normal independente	4
1.4.1.1	Distribuição t -Student multivariada	5
1.4.1.2	Distribuição slash multivariada	6
1.4.2	Critérios de seleção de modelos	7
2	Modelo com erros de medida e dados censurados baseados nas distri-	
	buições normal independente multivariadas	8
2.1	Especificação do modelo	8
2.2	Distribuição slash multivariada	11
2.2.1	Função de verossimilhança	14
2.2.2	Algoritmo MCECM	16
2.2.2.1	Passo E	18
2.2.2.2	Passo CM	23
2.2.3	Matriz de informação observada	24
2.3	Distribuição t -Student multivariada	25
2.3.1	Função de verossimilhança	28
2.3.2	Algoritmo ECM	29
2.3.2.1	Passo E	30

2.3.2.2	Passo CM	32
2.3.3	Matriz de informação observada	32
3	Análise de diagnóstico	33
3.1	Influência global	34
3.2	Influência local	36
3.2.1	Matriz hessiana	38
3.2.2	Esquemas de perturbação	38
3.2.2.1	Perturbação de ponderação de casos	38
3.2.2.2	Perturbação na covariável	40
3.2.3	Caso modelo com erros de medida estrutural e dados censurados baseado na distribuição normal (<i>MEMC-N</i>)	40
3.2.3.1	Matriz hessiana	43
3.2.3.2	Perturbação de ponderação de casos	44
3.2.3.3	Perturbação na covariável	46
3.2.4	Caso <i>MEMC-t</i>	48
3.2.4.1	Matriz hessiana	48
3.2.4.2	Perturbação de ponderação de casos	50
3.2.4.3	Perturbação na covariável	53
3.2.5	Caso <i>MEMC-SI</i>	56
3.2.5.1	Matriz hessiana	58
3.2.5.2	Esquemas de perturbação	58
4	Estudo de simulação e aplicação	59
4.1	Estudo de simulação	59
4.1.1	Propriedades assintóticas	59
4.1.2	Estimação das medidas de influência	60
4.2	Aplicação	67
5	Considerações finais	80
5.1	Perspectivas futuras	80

Lista de Figuras

1.1	Gráfico das distribuições normal padrão, t -Student e slash para diferentes valores de ν	6
4.1	Estudo de simulação. EQM das estimativas dos parâmetros sob o $MEMC-t$ considerando 10% de censura.	61
4.2	Estudo de simulação. $VIÉS$ das estimativas dos parâmetros sob o $MEMC-t$ considerando 10% de censura.	62
4.3	Estudo de simulação. EQM das estimativas dos parâmetros sob o $MEMC-Sl$ considerando 10% de censura.	63
4.4	Estudo de simulação. $VIÉS$ das estimativas dos parâmetros sob o $MEMC-Sl$ considerando 10% de censura.	64
4.5	Logaritmo da função de verossimilhança do $MEMC-t$ e $MEMC-Sl$ para diferentes graus de liberdade.	69
4.6	Análise global de diagnóstico para os modelos $MEMC-N$, $MEMC-t$ e $MEMC-Sl$ segundo o afastamento da função Q considerando 10%, 30%, 50% e 70% de censura.	72
4.7	Análise global de diagnóstico para os modelos $MEMC-N$, $MEMC-t$ e $MEMC-Sl$ segundo a distância de Cook generalizada considerando 10%, 30%, 50% e 70% de censura.	73
4.8	Gráficos de índices de $M(0)$ para os modelos $MEMC-N$, $MEMC-t$ e $MEMC-Sl$ segundo o esquema de perturbação de ponderação de casos considerando 10%, 30%, 50% e 70% de censura.	78

4.9	Gráficos de índices de $M(0)$ para os modelos $MEMC-N$, $MEMC-t$ e $MEMC-Sl$ segundo o esquema de perturbação da covariável considerando 10%, 30%, 50% e 70% de censura.	79
-----	---	----

Lista de Tabelas

4.1	Estudo de simulação. Análise de influência via estudo de Monte Carlo para as observações #1 e #50 por distribuição e medida de diagnóstico: AQ (afastamento da função Q), DC (Distância generalizada de Cook), PPC (perturbação de ponderação de casos) e PC (perturbação da covariável) .	66
4.2	Dados de Chipkevitch <i>et al.</i> Dados do volume testicular (em ml) de 42 adolescentes	68
4.3	Dados de Chipkevitch <i>et al.</i> EMV e EP para os parâmetros estimados .	69
4.4	Dados de Chipkevitch <i>et al.</i> Critérios de comparação dos modelos	70
4.5	EMV , EP e MR dos parâmetros do $MEMC-N$ com a amostra completa e tirando as observações influentes.	74
4.6	EMV , EP e MR dos parâmetros do $MEMC-t$ com a amostra completa e tirando as observações influentes.	75
4.7	EMV , EP e MR dos parâmetros do $MEMC-SI$ com a amostra completa e tirando as observações influentes.	76
4.8	Mudanças (em %) nas estimativas dos parâmetros dos modelos ajustados depois de excluídas as observações 31 e 32.	77
4.9	Dados de Chipkevitch <i>et al.</i> Comparação das mudanças relativas nos EMV segundo a MRT para os três modelos considerados.	77

Capítulo 1

Introdução

1.1 Introdução

Modelos com erros de medida (*MEM*), também conhecido como modelos com erros nas variáveis, são úteis para descrever diferentes fenômenos em várias áreas de interesse como engenharia, saúde, sociais, entre outros. As principais bibliografias sobre *MEM* podem ser encontrados em Fuller (1987), Cheng e Van-Ness (1999), Carroll et al. (2006), Buonaccorsi (2010), entre outros. Os *MEM* também são usados em problemas de comparar dispositivos de medição (ver, Barnett, 1969; Bolfarine e Galea-Rojas, 1996), que podem variar em preço, tempo de medição e outras características, tais como a eficiência. Vários outros exemplos de *MEM*, por exemplo, na área médica são reportados na literatura, especificamente em Kelly (1984), Chipkevitch et al. (1996) e Lu et al. (1997). Exemplos em psicologia e educação foram considerados por Dunn (1992). No entanto, em todos estes estudos supõe-se que a distribuição dos erros aleatórios, assim como as covariáveis não observadas são Gaussianas. Mas a suposição de normalidade é muito restritiva e sofre com a falta de robustez, o que pode ter um efeito importante sobre as inferências. Alguns resultados de *MEM* utilizando a distribuição *t*-Student pode ser encontrado em Bolfarine e Galea-Rojas (1996) e Galea-Rojas et al. (2005). Arellano-Valle et al. (2005) mostraram a vantagem de utilizar a distribuição skew-normal no contexto de *MEM*. Assim, um estudo de suas propriedades sob supostos não padrão, tais como normalidade é muito pertinente.

A classe de distribuições normal independente é uma extensão das distribuições normais multivariadas. Esta extensão resulta em uma classe flexível de modelos para estimação robusta em *MEM* que contém as distribuições normal (*MEM-N*), *t*-Student (*MEM-t*), slash (*MEM-Sl*) e a normal contaminada (*MEM-NC*). Todas estas distribuições tem caudas mais pesadas que uma normal, e podem ser usadas para inferência robusta em vários tipos de modelos.

Embora muitos dos modelos para dados multivariados considerem a existência de covariáveis mal medidas, muitos deles não consideram observações censuradas ou limites de detecção para a variável resposta. Este aspecto é relevante, uma vez que em vários estudos a resposta observada está sujeita a limites de detecção máximos e/ou mínimos. Por essa razão, é evidente a necessidade de uma nova metodologia que leve em conta as respostas censuradas em dados multivariados e covariáveis mal medidas ao mesmo tempo, assim como metodologias que permitam realizar análise de diagnóstico.

A análise de diagnóstico tem o objetivo de verificar possíveis afastamentos das suposições feitas para o modelo, verificar a existência de observações extremas com interferência desproporcional no ajuste e detectar observações influentes nas estimativas do modelo.

Ainda que a característica de robustez esteja associada às distribuições normal independente, elas podem estar vulneráveis a observações influentes. Assim, é essencial avaliar a sensibilidade dos resultados obtidos em um processo de estimação e modelagem. Entre as alternativas usuais, temos a análise de influência global e local. Na análise de influência global avalia-se o impacto de uma observação sobre o processo de estimação, testes de hipóteses e ajuste de modelos, quando esta é eliminada do conjunto de dados (Cook, 1977). Já na análise de influência local, avalia-se o efeito de pequenos ruídos ao inserir um vetor de perturbação no conjunto de dados ou sobre as suposições do modelo, sem necessidade de eliminar observações (Cook, 1986).

No entanto, para as distribuições normal independente (e outras distribuições), a função de log-verossimilhança marginal é complexa e uma aplicação direta da abordagem de Cook pode ser muito difícil, pois essas medidas envolvem a primeira e segunda derivadas parciais desta função. Inspirados pela ideia básica do algoritmo EM (Espe-

rança-Maximização), Zhu e Lee (2001) propuseram um método unificado para análise de influência local em modelos estatísticos com dados faltantes, utilizando a função de afastamento da verossimilhança completa. Esta abordagem produz resultados muito semelhantes aos obtidos com o método de Cook.

Neste contexto, estudamos e desenvolvemos procedimentos para análise em modelos de regressão normal independente com erros de medida e dados censurados, sendo desenvolvida estimação e inferência robusta, no sentido de utilizar distribuições que acomodem de forma mais eficiente observações “outliers” do que a distribuição normal. Além disso, realizamos um estudo de diagnóstico de influência utilizando a metodologia proposta por Zhu e Lee (2001).

1.2 Proposta da tese e objetivos

A proposta desta tese é estender o trabalho de Matos et al. (2016) no sentido de desenvolver análise de diagnóstico. Além disso, pretendemos estudar e desenvolver para modelos de regressão com erros de medida e dados censurados, estimação e inferência, no sentido de utilizar distribuições normal e slash multivariada, assim como também a realização de análise de diagnóstico para esses modelos.

Nossos objetivos específicos para este trabalho são os seguintes:

- i) Desenvolver a estimação por máxima verossimilhança através da implementação de alguma das extensões do algoritmo EM, no modelo de regressão linear multivariado com erros de medida e respostas censuradas na abordagem estrutural, baseado na distribuição slash multivariada;
- ii) aplicar o método de influência global e local aos modelos com erros de medida estrutural e dados censurados baseados nas distribuições normal, t -Student e slash;
- iii) fazer simulações considerando 10% de censura e diferentes tamanhos de amostra para avaliar o comportamento assintótico das estimativas do algoritmo EM nos modelos propostos e, finalmente, fazer uma aplicação da teoria desenvolvida utilizando as distribuições normal, t -Student e slash a um conjunto de dados.

1.3 Organização do trabalho

O trabalho está organizado da seguinte forma. No Capítulo 2 apresentamos o modelo com erros de medida para respostas multivariadas censuradas sob a distribuição normal independente, na sua forma geral, e em seguida particularizamos aos modelos slash e t -Student multivariado, assim como as estimativas baseadas na verossimilhança e os erros padrão das estimativas dos parâmetros nos dois modelos por meio de algoritmos de tipo MCECM e ECM, respectivamente. No Capítulo 3 estendemos a análise de diagnóstico aos modelos com erros de medida para respostas multivariadas censuradas sob a classe normal independente, com foco nas distribuições normal, t -Student e slash multivariada, efetuando um estudo de diagnóstico de influência global e local usando a metodologia proposta por Zhu e Lee (2001). No Capítulo 4 apresentamos os resultados de um estudo de simulação realizado para examinar o desempenho do método proposto com relação às propriedades assintóticas das estimativas de máxima verossimilhança nos modelos t -Student e slash e também aplicamos a um conjunto de dados reais. Finalmente, as considerações finais do trabalho podem ser observados no Capítulo 5.

1.4 Aspectos preliminares

O objetivo desta seção é introduzir alguns conceitos básicos para o desenvolvimento de nosso trabalho, assim como definir a notação e terminologia pertinentes.

1.4.1 Distribuições normal independente

A família de distribuições normal independente tem sido investigada por vários autores, dentre eles Andrews e Mallows (1974) e Lange e Sinsheimer (1993).

Uma distribuição normal independente (Lange e Sinsheimer, 1993) ou simplesmente distribuição NI é definida como o vetor aleatório p -dimensional

$$\mathbf{Y} = \boldsymbol{\mu} + U^{-1/2}\mathbf{Z}, \quad (1.1)$$

em que $\boldsymbol{\mu} \in \mathbb{R}^p$ é um vetor de locação (constante), \mathbf{Z} é um vetor aleatório normal com vetor de médias $\mathbf{0}$, matriz de covariâncias $\boldsymbol{\Sigma}$ e U é uma variável aleatória positiva com

função de distribuição acumulada (fda) $H(u; \nu)$ e função de densidade de probabilidade (fdp) $h(u; \nu)$, indexado pelo parâmetro ν , independente de \mathbf{Z} . Dado U , \mathbf{Y} segue uma distribuição normal multivariada com vetor de médias $\boldsymbol{\mu}$ e matriz de covariâncias $u^{-1}\boldsymbol{\Sigma}$, ou seja, $\mathbf{Y}|U = u \sim N_p(\boldsymbol{\mu}, u^{-1}\boldsymbol{\Sigma})$. Conseqüentemente, a fdp de \mathbf{Y} é dada por

$$f(\mathbf{y}) = \int_0^\infty \phi_p(\mathbf{y}; \boldsymbol{\mu}, u^{-1}\boldsymbol{\Sigma}) dH(u; \nu), \quad (1.2)$$

em que $\phi_p(\cdot; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ representa a fdp da distribuição normal p -variada com vetor de médias $\boldsymbol{\mu}$ e matriz de covariâncias $\boldsymbol{\Sigma}$. Um caso particular desta distribuição é a distribuição normal, para o qual $U = 1$.

A família de distribuições normal independente inclui modelos tais como as distribuições t -Student, slash, normal contaminada, entre outros. Todas estas distribuições têm caudas mais pesadas do que uma normal e podem ser usadas para inferência robusta. Na seguinte subseção, apresentamos alguns casos especiais de distribuições normal independente. Outros membros de distribuições normal independente podem ser encontrados em Lange e Sinsheimer (1993).

1.4.1.1 Distribuição t -Student multivariada

A distribuição t -Student multivariada com ν graus de liberdade, $t_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$, pode ser derivada a partir do modelo de mistura (1.1), em que U é distribuída como $Gama(\nu/2, \nu/2)$, com $u > 0$ e $\nu > 0$, e em que $Gama(a, b)$ denota a distribuição gama com média a/b . A fdp de \mathbf{Y} toma a seguinte forma:

$$f(\mathbf{y}_i) = t_p(\mathbf{y}_i | \boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu) = \frac{\Gamma\left(\frac{p+\nu}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right) \pi^{p/2}} \nu^{-p/2} |\boldsymbol{\Sigma}|^{-1/2} \left(1 + \frac{\delta_i}{\nu}\right)^{-(p+\nu)/2}, \quad \mathbf{y}_i \in \mathbb{R}^p, \quad (1.3)$$

em que $\Gamma(\cdot)$ é a função gama padrão e

$$\delta_i = (\mathbf{y}_i - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_i - \boldsymbol{\mu}) \quad (1.4)$$

é a distância de Mahalanobis. Um caso particular da distribuição t -Student é a Cauchy, quando $\nu = 1$. Também, quando $\nu \uparrow \infty$, obtemos a distribuição normal.

Aplicações da distribuição t -Student para estimação robusta em *MEM* podem ser encontrados em Galea-Rojas et al. (2005).

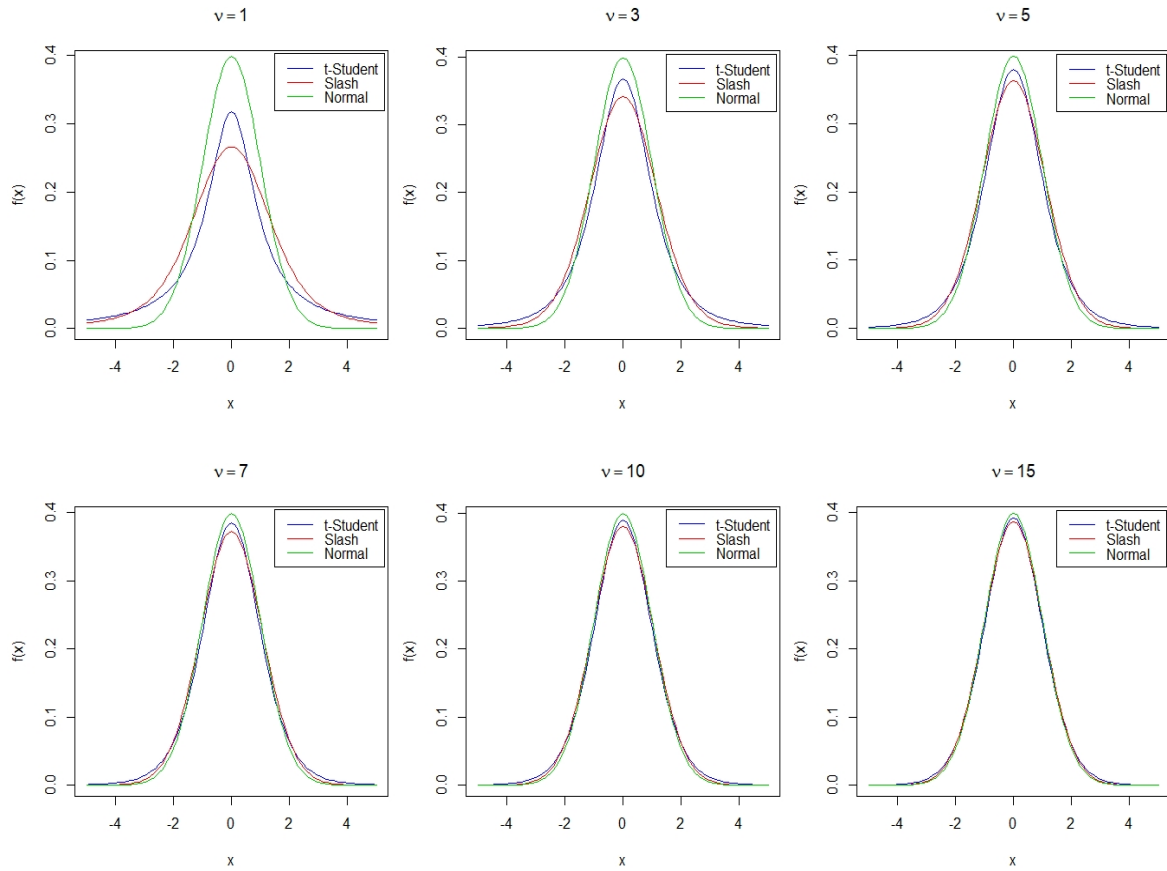


Figura 1.1: Gráfico das distribuições normal padrão, t -Student e slash para diferentes valores de ν .

1.4.1.2 Distribuição slash multivariada

Outra distribuição da classe normal independente, denominada distribuição slash multivariada e denotada por $Sl_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$, surge quando a distribuição de U é $Beta(\nu, 1)$, com $0 < u < 1$ e $\nu > 0$. Sua fdp é dada por

$$f(\mathbf{y}) = \nu \int_0^1 u^{\nu-1} \phi_p(\mathbf{y}; \boldsymbol{\mu}, u^{-1}\boldsymbol{\Sigma}) du, \quad \mathbf{y} \in \mathbb{R}^p.$$

A distribuição slash se reduz à distribuição normal quando $\nu \uparrow \infty$.

Na Figura 1.1 podemos observar a variação das curvas das distribuições t -Student e slash para diferentes valores de ν e vemos como elas aproximam-se à normal quando o valor de ν cresce.

1.4.2 Critérios de seleção de modelos

Alguns critérios comumente utilizados para a seleção de modelos são o critério de informação de Akaike (*AIC*) e o critério de informação bayesiano (*BIC*) propostos por Akaike (1973) e Schwarz (1978) respectivamente. Estes critérios são definidos por

$$AIC = -2\ell(\hat{\boldsymbol{\theta}}) + 2t,$$

e

$$BIC = -2\ell(\hat{\boldsymbol{\theta}}) + t \log(n),$$

onde $\hat{\boldsymbol{\theta}}$ é o estimador de máxima verossimilhança, ℓ é a função de log-verossimilhança, t é o número de parâmetros livres do modelo e n é o número de observações. A escolha do melhor modelo se faz considerando aquele que apresenta o menor valor dos critérios utilizados (*AIC* ou *BIC*).

Capítulo 2

Modelo com erros de medida e dados censurados baseados nas distribuições normal independente multivariadas

2.1 Especificação do modelo

Seja $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{ir})^\top$ o vetor de respostas para a i -ésima unidade experimental, onde Y_{ij} é a j -ésima resposta observada da unidade i (para $i = 1, \dots, n$ e $j = 1, \dots, r$). Seja X_i o i -ésimo valor observado e x_i o valor não observado (verdadeiro) da covariável para a unidade i . Seguindo Barnett (1969), o *MEM* multivariado é formulado como

$$X_i = x_i + \xi_i \quad (2.1)$$

e

$$\mathbf{Y}_i = \boldsymbol{\alpha} + \boldsymbol{\beta}x_i + \mathbf{e}_i \quad (2.2)$$

em que $\mathbf{e}_i = (e_{i1}, \dots, e_{ir})^\top$ é um vetor de erros de medição, $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_r)^\top$ e $\boldsymbol{\beta} = (\beta_1, \dots, \beta_r)^\top$ são vetores com parâmetros de regressão. Seja $\boldsymbol{\epsilon}_i = (\xi_i, \mathbf{e}_i^\top)^\top$ e $\mathbf{Z}_i = (X_i, \mathbf{Y}_i^\top)^\top = (Z_{i1}, \dots, Z_{ip})^\top$. Então, as equações (2.1) e (2.2) implicam

$$\mathbf{Z}_i = \mathbf{a} + \mathbf{b}x_i + \boldsymbol{\epsilon}_i = \mathbf{a} + \mathbf{B}\mathbf{r}_i, \quad i = 1, \dots, n, \quad (2.3)$$

onde $\mathbf{a} = (0, \boldsymbol{\alpha}^\top)^\top$ e $\mathbf{b} = (1, \boldsymbol{\beta}^\top)^\top$ são vetores $p \times 1$, com $p = r + 1$, $\mathbf{B} = [\mathbf{b}; \mathbf{I}_p]$ é uma matriz $p \times (p + 1)$, sendo \mathbf{I}_p a matriz identidade de ordem p e $\mathbf{r}_i = (x_i, \boldsymbol{\epsilon}_i^\top)^\top$. Assim, da Equação (2.3), a distribuição de \mathbf{Z}_i torna-se especificada uma vez que a distribuição de \mathbf{r}_i é especificada. Usualmente, é feita uma suposição de normalidade, tal que

$$\mathbf{r}_i = \begin{bmatrix} x_i \\ \boldsymbol{\epsilon}_i \end{bmatrix} \stackrel{\text{iid}}{\sim} N_{1+p} \left(\begin{bmatrix} \mu_x \\ \mathbf{0}_p \end{bmatrix}, \begin{bmatrix} \sigma_x^2 & \mathbf{0}_p^\top \\ \mathbf{0}_p & \boldsymbol{\Omega} \end{bmatrix} \right), \quad i = 1, \dots, n \quad (2.4)$$

em que $\mathbf{0}_p = (0, \dots, 0)^\top$ é um vetor $p \times 1$, $\boldsymbol{\Omega} = \text{diag}(\phi_1^2, \dots, \phi_p^2)$, e $\stackrel{\text{iid}}{\sim}$ denota vetores aleatórios independentes e identicamente distribuídos. Marginalmente, temos que $x_i \stackrel{\text{iid}}{\sim} N(\mu_x, \sigma_x^2)$ e $\boldsymbol{\epsilon}_i \stackrel{\text{iid}}{\sim} N_p(\mathbf{0}, \boldsymbol{\Omega})$ são independentes para todo $i = 1, \dots, n$. Para mais detalhes veja, por exemplo, Fuller (1987, Seção 4.1).

Para obter uma estimativa robusta dos parâmetros no modelo, consideramos a distribuição normal independente, dado por

$$\mathbf{r}_i = \begin{bmatrix} x_i \\ \boldsymbol{\epsilon}_i \end{bmatrix} \stackrel{\text{iid}}{\sim} NI_{1+p} \left(\begin{bmatrix} \mu_x \\ \mathbf{0}_p \end{bmatrix}, \begin{bmatrix} \sigma_x^2 & \mathbf{0}_p^\top \\ \mathbf{0}_p & \boldsymbol{\Omega} \end{bmatrix}; H(u_i; \nu) \right), \quad i = 1, \dots, n. \quad (2.5)$$

Utilizando a equação (1.1), esta formulação implica que

$$\begin{aligned} \begin{bmatrix} x_i \\ \boldsymbol{\epsilon}_i \end{bmatrix} | U_i = u_i &\sim N_{1+p} \left(\begin{bmatrix} \mu_x \\ \mathbf{0}_p \end{bmatrix}, u_i^{-1} \begin{bmatrix} \sigma_x^2 & \mathbf{0}_p^\top \\ \mathbf{0}_p & \boldsymbol{\Omega} \end{bmatrix} \right), \\ U_i &\sim H(u_i; \nu), \end{aligned}$$

para $i = 1, \dots, n$. Consequentemente,

$$x_i | U_i = u_i \stackrel{\text{ind}}{\sim} N(\mu_x, u_i^{-1} \sigma_x^2) \quad \text{e}, \quad (2.6)$$

$$\boldsymbol{\epsilon}_i | U_i = u_i \stackrel{\text{ind}}{\sim} N_p(\mathbf{0}_p, u_i^{-1} \boldsymbol{\Omega}). \quad (2.7)$$

Além disso, $\boldsymbol{\epsilon}_i$ e x_i tem distribuições com marginais normal independente, sendo $\boldsymbol{\epsilon}_i \sim NI_p(\mathbf{0}, \boldsymbol{\Omega}; H(u_i; \nu))$ e $x_i \sim NI(\mu_x, \sigma_x^2; H(u_i; \nu))$.

Desde que para cada i , $\boldsymbol{\epsilon}_i$ e x_i são indexados pelo mesmo fator de mistura de escala U_i , eles não são independentes em geral. Sua independência corresponde ao caso em que $U_i = 1$ (caso normal). No entanto, condicionado em U_i , $\boldsymbol{\epsilon}_i$ e x_i são independentes

para cada $i = 1, \dots, n$, o que implica que $\boldsymbol{\epsilon}_i$ e x_i são não correlacionados, desde que $Cov(\boldsymbol{\epsilon}_i, x_i) = E[\boldsymbol{\epsilon}_i x_i | U_i] = 0$. Por (2.3), \mathbf{Z}_i é uma transformação afim de \mathbf{r}_i . Assim, a distribuição é dada por

$$\mathbf{Z}_i \sim NI_p(\boldsymbol{\mu}_z, \boldsymbol{\Sigma}_z; H(u_i; \nu)), \quad i = 1, \dots, n, \quad (2.8)$$

em que

$$\boldsymbol{\mu}_z = \mathbf{a} + \mathbf{b}\mu_x \quad \text{e} \quad \boldsymbol{\Sigma}_z = \sigma_x^2 \mathbf{b}\mathbf{b}^\top + \boldsymbol{\Omega}. \quad (2.9)$$

Considerando agora o modelo com observações censuradas, temos o caso em que a resposta Z_{ij} não é totalmente observada para todo i, j . O que observamos na verdade, para cada $i = 1, \dots, n$, é o vetor aleatório $\mathbf{V}_i = (V_{i1}, \dots, V_{ip})^\top$, tal que $V_{ij} = \max\{Z_{ij}, \kappa_{ij}\}$, onde κ_{ij} é um nível de censura, isto é,

$$V_{ij} = \begin{cases} Z_{ij}, & \text{se } Z_{ij} > \kappa_{ij} \\ \kappa_{ij}, & \text{se } Z_{ij} \leq \kappa_{ij}. \end{cases} \quad (2.10)$$

O modelo definido pelas equações (2.1), (2.2), conjuntamente com (2.5) e (2.10) é denominado modelo com erros de medida estrutural e respostas censuradas baseados nas distribuições normal independente (*MEMC-NI*). Por conveniência, escolhemos trabalhar com o caso de censura à esquerda, mas os resultados são facilmente estendidos para outros tipos de censura.

O *MEMC-NI* pode ser formulado numa representação hierárquica flexível que é útil para a obtenção das derivadas. É obtida através das equações (2.3), (2.6) e (2.7) e é dado por

$$\mathbf{Z}_i | x_i, U_i = u_i \stackrel{\text{ind}}{\sim} N_p(\mathbf{a} + \mathbf{b}x_i, u_i^{-1}\boldsymbol{\Omega}), \quad (2.11)$$

$$x_i | U_i = u_i \stackrel{\text{ind}}{\sim} N(\mu_x, u_i^{-1}\sigma_x^2), \quad (2.12)$$

$$U_i \stackrel{\text{iid}}{\sim} H(u_i; \nu), \quad i = 1, \dots, n. \quad (2.13)$$

Para obter (2.11), sabemos por (2.8) que $\mathbf{Z}_i \sim NI_p(\boldsymbol{\mu}_z, \boldsymbol{\Sigma}_z; H(u_i; \nu))$, em que $\boldsymbol{\mu}_z = \mathbf{a} + \mathbf{b}\mu_x$ e $\boldsymbol{\Sigma}_z = \sigma_x^2 \mathbf{b}\mathbf{b}^\top + \boldsymbol{\Omega}$. Então \mathbf{Z}_i pode ser expresso como

$$\mathbf{Z}_i = \boldsymbol{\mu}_z + U_i^{-1/2} \mathbf{W}, \quad \text{sendo } \mathbf{W} \sim N_p(\mathbf{0}, \boldsymbol{\Sigma}_z) \quad \text{e} \quad U_i \sim H(u_i; \nu).$$

Então, $\mathbf{Z}_i|U_i = u_i \sim N_p(\boldsymbol{\mu}_z, u_i^{-1}\boldsymbol{\Sigma}_z)$.

Portanto,

$$\begin{aligned}\mathbf{Z}_i|x_i, U_i = u_i &\stackrel{\text{ind}}{\sim} N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \quad \text{sendo } \boldsymbol{\mu} = \mathbf{a} + \mathbf{b}x_i \\ \boldsymbol{\Sigma} &= u_i^{-1}\boldsymbol{\Sigma}_z = u_i^{-1}(\sigma_x^2\mathbf{b}\mathbf{b}^\top + \boldsymbol{\Omega}) = u_i^{-1}\boldsymbol{\Omega}\end{aligned}$$

ou seja $\mathbf{Z}_i|x_i, U_i = u_i \stackrel{\text{ind}}{\sim} N_p(\mathbf{a} + \mathbf{b}x_i, u_i^{-1}\boldsymbol{\Omega})$.

Proposição 1. *Considere a representação hierárquica do MEMC–NI dado em (2.11)–(2.13). Então,*

$$x_i|U_i = u_i, \mathbf{Z}_i = \mathbf{z}_i \sim N\left(\frac{\mu_x + \sigma_x^2\mathbf{b}^\top\boldsymbol{\Omega}^{-1}(\mathbf{z}_i - \mathbf{a})}{1 + \sigma_x^2\mathbf{b}^\top\boldsymbol{\Omega}^{-1}\mathbf{b}}, \frac{\sigma_x^2}{u_i(1 + \sigma_x^2\mathbf{b}^\top\boldsymbol{\Omega}^{-1}\mathbf{b})}\right).$$

A prova segue da relação $f(x_i|u_i, \mathbf{z}_i) \propto f(\mathbf{z}_i|x_i, u_i)f(x_i|u_i)$, em que $f(\cdot)$ denota uma fdp genérica. Uma prova desta proposição é apresentada no Apêndice A.1.

2.2 Distribuição slash multivariada

A distribuição slash multivariada, $Sl_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$, pode ser derivada a partir do modelo de mistura (1.1), em que U é distribuída como $Beta(\nu, 1)$, com $0 < u < 1$ e $\nu > 0$. A fdp é dada por

$$f(\mathbf{y}_i) = Sl_p(\mathbf{y}_i; \boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu) = \nu \int_0^1 u^{\nu-1} \phi_p(\mathbf{y}_i; \boldsymbol{\mu}, u^{-1}\boldsymbol{\Sigma}) du = \frac{\nu}{(2\pi)^{\frac{p}{2}}|\boldsymbol{\Sigma}|^{\frac{1}{2}}} \int_0^1 u^{\frac{p}{2}+\nu-1} e^{-\frac{u\delta_i}{2}} du,$$

em que ν é parâmetro de forma e δ_i é dado em (1.4). A fda de \mathbf{Y} é denotada por $SL_p(\cdot|\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$ e quando $\nu \uparrow \infty$, a distribuição slash se reduz à distribuição normal.

O vetor aleatório \mathbf{Y} admite a representação estocástica

$$\mathbf{Y} = \boldsymbol{\mu} + U^{-1/2}\mathbf{Z}, \quad \mathbf{Z} \sim N_p(\mathbf{0}, \boldsymbol{\Sigma}), \quad U \sim Beta(\nu, 1) \quad (2.14)$$

em que \mathbf{Z} e U são independentes, e $Beta(a, b)$ denota a distribuição beta.

A seguir, provamos as seguintes proposições que são importantes na implementação do algoritmo EM, utilizado no cálculo das *EMV* dos parâmetros neste modelo.

Proposição 2. *Os momentos recíprocos*

$$E[U^{-m}] = \frac{\nu}{\nu - m}$$

existem para $\nu > m$.

Prova:

Desde que $U \sim \text{Beta}(\nu, 1)$, então $f(u) = \nu u^{\nu-1}$, $0 < u < 1$ e fazendo $w = u^{-1}$ temos que $u = w^{-1}$ e

$$f_W(w) = f_U(u) \left| \frac{du}{dw} \right| = f_U(w^{-1}) \left| \frac{d(w^{-1})}{dw} \right| = \nu (w^{-1})^{\nu-1} w^{-2} = \nu w^{-\nu-1}, \quad 1 < w < \infty$$

$$\begin{aligned} E[U^{-m}] &= E[W^m] = \int_1^\infty w^m \nu w^{-\nu-1} dw = \nu \int_1^\infty w^{-\nu-1+m} dw = \frac{\nu}{-\nu+m} w^{-\nu+m} \Big|_1^\infty \\ &= \frac{\nu}{\nu-m}, \quad \nu > m. \end{aligned}$$

Proposição 3. *Suponha que $\mathbf{Y} \sim Sl_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$. Assim,*

i) $E[\mathbf{Y}] = \boldsymbol{\mu}, \quad \nu > \frac{1}{2};$

ii) $\text{Cov}(\mathbf{Y}) = \frac{\nu}{\nu-1} \boldsymbol{\Sigma}, \quad \nu > 1.$

Prova:

Desde que $\mathbf{Y} \sim Sl_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$, e utilizando a representação estocástica (2.14),

i) $E[\mathbf{Y}] = \boldsymbol{\mu} + E[U^{-1/2}]E[\mathbf{Z}] = \boldsymbol{\mu}, \quad \nu > \frac{1}{2}$

ii) Por (2.14), $\boldsymbol{\Sigma} = \text{Cov}(\mathbf{Z}) = E[\mathbf{Z}\mathbf{Z}^\top]$, pelo que

$$\text{Cov}(\mathbf{Y}) = \text{Cov}(U^{-1/2} \mathbf{Z}) = E[U^{-1}]E[\mathbf{Z}\mathbf{Z}^\top] = \frac{\nu}{\nu-1} \boldsymbol{\Sigma}, \quad \nu > 1.$$

Proposição 4. *Seja $\mathbf{X} \sim Sl_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$. Se $\mathbf{a} \in \mathbb{R}^q$ e \mathbf{B} é uma matriz $q \times p$ com $r(\mathbf{B}) = q$, então*

$$\mathbf{Y} = \mathbf{a} + \mathbf{B}\mathbf{X} \sim Sl_q(\mathbf{a} + \mathbf{B}\boldsymbol{\mu}, \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}^\top, \nu).$$

A prova pode ser encontrada em Fang et al. (1990) para a classe de distribuições elípticas, classe mais geral que contém às distribuições NI.

Proposição 5. *Seja $\mathbf{Y} \sim Sl_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$. Considere a partição de \mathbf{Y} , $\boldsymbol{\mu}$ e $\boldsymbol{\Sigma}$ como*

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{bmatrix}, \quad \boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix} \quad e \quad \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}$$

em que \mathbf{Y}_1 e $\boldsymbol{\mu}_1$ são vetores $p_1 \times 1$ e $\boldsymbol{\Sigma}_{11}$ é uma matriz $p_1 \times p_1$. Então

i) $\mathbf{Y}_1 \sim Sl_{p_1}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11}, \nu)$;

ii) $\mathbf{Y}_2 | \mathbf{Y}_1 = \mathbf{y}_1 \sim Sl_{p_2}(\boldsymbol{\mu}_{2.1}, \boldsymbol{\Sigma}_{22.1}, \nu + p_1)$, em que

$$\boldsymbol{\mu}_{2.1} = \boldsymbol{\mu}_2 + \boldsymbol{\Sigma}_{21} \boldsymbol{\Sigma}_{11}^{-1} (\mathbf{y}_1 - \boldsymbol{\mu}_1)$$

$$\boldsymbol{\Sigma}_{22.1} = \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21} \boldsymbol{\Sigma}_{11}^{-1} \boldsymbol{\Sigma}_{12}$$

Prova:

i) Fazendo $\mathbf{a} = \mathbf{0}_{p_1}$ e $\mathbf{B} = [\mathbf{I}_{p_1}, \mathbf{0}_{p_1 \times p_2}]$ na Proposição 4, temos que

$$\mathbf{a} + \mathbf{B}\mathbf{Y} = \mathbf{B}\mathbf{Y} = [\mathbb{I}, \mathbf{0}] \begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{bmatrix} = \mathbf{Y}_1$$

Então

$$E[\mathbf{Y}_1] = E[\mathbf{B}\mathbf{Y}] = \mathbf{B} E[\mathbf{Y}] = \mathbf{B}\boldsymbol{\mu} = [\mathbb{I}, \mathbf{0}] \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix} = \boldsymbol{\mu}_1;$$

$$Cov[\mathbf{Y}_1] = Cov[\mathbf{B}\mathbf{Y}] = \mathbf{B} \left(\frac{\nu}{\nu - 1} \boldsymbol{\Sigma} \right) \mathbf{B}^\top = \frac{\nu}{\nu - 1} [\mathbb{I}, \mathbf{0}] \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix} \begin{bmatrix} \mathbb{I} \\ \mathbf{0} \end{bmatrix} = \frac{\nu}{\nu - 1} \boldsymbol{\Sigma}_{11}$$

$$\therefore \mathbf{Y}_1 \sim Sl_{p_1}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11}, \nu).$$

A seguinte definição é importante no cálculo da função de verossimilhança.

Definição 1. *Seja $\mathbf{Y} \sim Sl_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$ e \mathbb{D} um conjunto de Borel em \mathbb{R}^p . Dizemos que o vetor aleatório \mathbf{Z} tem uma distribuição slash truncada em \mathbb{D} quando \mathbf{Z} tem a mesma distribuição que $\mathbf{Y} | (\mathbf{Y} \in \mathbb{D})$. Neste caso, a fdp de \mathbf{Z} é dada por*

$$TSl_p(\mathbf{z} | \boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu; \mathbb{D}) = \frac{Sl_p(\mathbf{z} | \boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)}{P(\mathbf{Y} \in \mathbb{D})} \mathbb{I}_{\mathbb{D}}(\mathbf{z}),$$

onde $\mathbb{I}_{\mathbb{D}}(\cdot)$ é a função indicadora de \mathbb{D} , ou seja, $\mathbb{I}_{\mathbb{D}}(\mathbf{z}) = 1$ se $\mathbf{z} \in \mathbb{D}$ e $\mathbb{I}_{\mathbb{D}}(\mathbf{z}) = 0$ em outro caso. Nós usamos a notação $\mathbf{Z} \sim TSl_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu; \mathbb{D})$. Se \mathbb{D} tem a forma

$$\mathbb{D} = \{(x_1, \dots, x_p) \in \mathbb{R}^p; \ x_1 \leq d_1, \dots, x_p \leq d_p\}, \quad (2.15)$$

então usamos a notação $(\mathbf{Y} \in \mathbb{D}) = (\mathbf{Y} \leq \mathbf{d})$, em que $\mathbf{d} = (d_1, \dots, d_p)^\top$. Neste caso, $P(\mathbf{Y} \leq \mathbf{d}) = Sl_p(\mathbf{d}|\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$. Note que podemos ter $d_i = +\infty$, $i = 1, \dots, p$.

A fim de obter uma estimativa robusta dos parâmetros no modelo, substituímos a suposição (2.4) por

$$\mathbf{r}_i = \begin{bmatrix} x_i \\ \boldsymbol{\epsilon}_i \end{bmatrix} \stackrel{\text{iid}}{\sim} Sl_{1+p} \left(\begin{bmatrix} \mu_x \\ \mathbf{0}_p \end{bmatrix}, \begin{bmatrix} \sigma_x^2 & \mathbf{0}_p^\top \\ \mathbf{0}_p & \boldsymbol{\Omega} \end{bmatrix}, \nu \right), \quad i = 1, \dots, n, \quad (2.16)$$

e pela representação estocástica (2.14), pode ser expressa por

$$\begin{aligned} \begin{bmatrix} x_i \\ \boldsymbol{\epsilon}_i \end{bmatrix} | U_i = u_i &\sim N_{1+p} \left(\begin{bmatrix} \mu_x \\ \mathbf{0}_p \end{bmatrix}, u_i^{-1} \begin{bmatrix} \sigma_x^2 & \mathbf{0}_p^\top \\ \mathbf{0}_p & \boldsymbol{\Omega} \end{bmatrix} \right), \\ U_i &\sim Beta(\nu, 1), \end{aligned}$$

para $i = 1, \dots, n$.

Por (2.3), \mathbf{Z}_i é uma transformação afim de \mathbf{r}_i cuja distribuição é dada por

$$\mathbf{Z}_i \sim Sl_p(\boldsymbol{\mu}_z, \boldsymbol{\Sigma}_z, \nu), \quad i = 1, \dots, n, \quad (2.17)$$

com $\boldsymbol{\mu}_z$ e $\boldsymbol{\Sigma}_z$ dados em (2.9).

Denotaremos de *MEMC-Sl* ao modelo com erros de medida estrutural e respostas censuradas baseados na distribuição slash multivariada.

2.2.1 Função de verossimilhança

Primeiro, particionamos \mathbf{Z}_i nas componentes observadas e censuradas, $\mathbf{Z}_i = \text{vec}(\mathbf{Z}_i^o, \mathbf{Z}_i^c)$, em que $\mathbf{Z}_i^o \in \mathbb{R}^{p_o}$ corresponde ao primeiro caso, $\mathbf{Z}_i^c \in \mathbb{R}^{p_c}$ corresponde ao último caso e $\text{vec}(\cdot)$ denota a função que empilha vetores. De forma correspondente, consideremos

$$\mathbf{V}_i = \text{vec}(\mathbf{V}_i^o, \mathbf{V}_i^c) \text{ e desde que } \mathbf{Z}_i \sim Sl_p(\boldsymbol{\mu}_z, \boldsymbol{\Sigma}_z, \nu), \boldsymbol{\mu}_z = \text{vec}(\boldsymbol{\mu}_z^o, \boldsymbol{\mu}_z^c) \text{ e } \boldsymbol{\Sigma}_z = \begin{bmatrix} \boldsymbol{\Sigma}_z^{oo} & \boldsymbol{\Sigma}_z^{oc} \\ \boldsymbol{\Sigma}_z^{co} & \boldsymbol{\Sigma}_z^{cc} \end{bmatrix}$$

sendo $\boldsymbol{\kappa}_i^c$ o vetor com os correspondentes níveis de censura para \mathbf{Z}_i^c . Pela Proposição 5, temos que

$$\mathbf{Z}_i^o \sim Sl_{p_o}(\boldsymbol{\mu}_z^o, \boldsymbol{\Sigma}_z^{oo}, \nu) \quad \text{e} \quad \mathbf{Z}_i^c | \mathbf{Z}_i^o = \mathbf{z}_i^o \sim Sl_{p_c}(\boldsymbol{\mu}_z^{co}, \boldsymbol{\Sigma}_z^{cc.o}, \nu + p_o), \quad (2.18)$$

em que

$$\boldsymbol{\mu}_z^{co} = \boldsymbol{\mu}_z^c + \boldsymbol{\Sigma}_z^{co}(\boldsymbol{\Sigma}_z^{oo})^{-1}(\mathbf{z}_i^o - \boldsymbol{\mu}_z^o), \quad \text{e} \quad (2.19)$$

$$\boldsymbol{\Sigma}_z^{cc.o} = \boldsymbol{\Sigma}_z^{cc} - \boldsymbol{\Sigma}_z^{co}(\boldsymbol{\Sigma}_z^{oo})^{-1}\boldsymbol{\Sigma}_z^{oc} \quad (2.20)$$

A amostra observada para a i -ésima unidade experimental é $\{\mathbf{z}_i^o, \boldsymbol{\kappa}_i^c\}$ e a verossimilhança associada é

$$L_i(\boldsymbol{\theta}) = P(\mathbf{V}_i^c = \boldsymbol{\kappa}_i^c | \mathbf{Z}_i^o = \mathbf{z}_i^o) f(\mathbf{z}_i^o),$$

em que $f(\cdot)$ é a densidade marginal de \mathbf{Z}_i^o . Mas $\mathbf{V}_i^c = \boldsymbol{\kappa}_i^c$ se e somente se $\mathbf{Z}_i^c \leq \boldsymbol{\kappa}_i^c$. Por (2.18), obtemos

$$L_i(\boldsymbol{\theta}) = SL_{p_c}(\boldsymbol{\kappa}_i^c | \boldsymbol{\mu}_z^{co}, \boldsymbol{\Sigma}_z^{cc.o}, \nu + p_o) Sl_{p_o}(\mathbf{z}_i^o | \boldsymbol{\mu}_z^o, \boldsymbol{\Sigma}_z^{oo}, \nu),$$

e que pode ser obtido considerando três diferentes casos:

i) Se o i -ésimo indivíduo não tem componentes censurados,

$$L_i(\boldsymbol{\theta}) = Sl_p(\mathbf{z}_i | \boldsymbol{\mu}_z, \boldsymbol{\Sigma}_z, \nu) = \int_0^1 \nu u^{\nu-1} \phi_p(\mathbf{z}_i | \boldsymbol{\mu}_z, u^{-1}\boldsymbol{\Sigma}_z) du = \frac{\nu \Gamma(\frac{p}{2} + \nu) P_1(\frac{p}{2} + \nu, \frac{\delta_i}{2})}{(2\pi)^{\frac{p}{2}} |\boldsymbol{\Sigma}_z|^{\frac{1}{2}} (\frac{\delta_i}{2})^{\frac{p}{2} + \nu}},$$

que é a fdp da distribuição slash multivariada no ponto \mathbf{z}_i , δ_i é dado na equação (1.4) e $P_x(a, b)$ é a fda da distribuição Gama(a,b), com média $\frac{a}{b}$, avaliada em x .

ii) Se o i -ésimo indivíduo tem apenas componentes censurados:

$$L_i(\boldsymbol{\theta}) = SL_p(\boldsymbol{\kappa}_i | \boldsymbol{\mu}_z, \boldsymbol{\Sigma}_z, \nu) = \int_{-\infty}^{\kappa_{i1}} \dots \int_{-\infty}^{\kappa_{ip}} \int_0^1 \nu u^{\nu-1} \phi_p(\mathbf{z}_i | \boldsymbol{\mu}_z, u^{-1}\boldsymbol{\Sigma}_z) du dz_{ip} \dots dz_{i1}.$$

iii) Se o i -ésimo indivíduo tem componentes censurados e não censurados:

$$\begin{aligned} L_i(\boldsymbol{\theta}) &= SL_{p_c}(\boldsymbol{\kappa}_i^c | \boldsymbol{\mu}_z^{co}, \boldsymbol{\Sigma}_z^{cc.o}, \nu + p_o) Sl_{p_o}(\mathbf{z}_i^o | \boldsymbol{\mu}_z^o, \boldsymbol{\Sigma}_z^{oo}, \nu) \\ &= \int_{-\infty}^{\kappa_{i1}^c} \dots \int_{-\infty}^{\kappa_{ip_c}^c} \int_0^1 (\nu + p_o) u^{\nu+p_o-1} \phi_{p_c}(\mathbf{z}_i^c | \boldsymbol{\mu}_z^{co}, u^{-1}\boldsymbol{\Sigma}_z^{cc.o}) du dz_{ip_c}^c \dots dz_{i1}^c \\ &= \frac{\nu \Gamma(\frac{p_o}{2} + \nu) P_1(\frac{p_o}{2} + \nu, \frac{\delta_i^o}{2})}{(2\pi)^{\frac{p_o}{2}} |\boldsymbol{\Sigma}_z^{oo}|^{\frac{1}{2}} (\frac{\delta_i^o}{2})^{\frac{p_o}{2} + \nu}}, \end{aligned}$$

em que $\delta_i^o = (\mathbf{z}_i^o - \boldsymbol{\mu}_z^o)^\top (\boldsymbol{\Sigma}_z^{oo})^{-1} (\mathbf{z}_i^o - \boldsymbol{\mu}_z^o)$ e $\boldsymbol{\mu}_z^{co}$ e $\boldsymbol{\Sigma}_z^{cc.o}$ são dados pelas equações (2.19) e (2.20), respectivamente.

Desde que o cálculo da verossimilhança não tem expressões de forma fechada, ela é aproximada utilizando métodos computacionais (neste trabalho, por sua simplicidade, utilizaremos a regra do trapézio), baseados no fato que dado que a fdp da distribuição slash é $f(\mathbf{z}_i) = \int_0^1 \nu u^{\nu-1} \phi_p(\mathbf{z}_i | \boldsymbol{\mu}_z, u^{-1} \boldsymbol{\Sigma}_z) du$, com $\mathbf{z}_i = (z_{i1}, \dots, z_{ip})^\top$, temos que

$$\begin{aligned} P(\mathbf{Z}_i \leq \mathbf{z}_i) &= \int_{-\infty}^{z_{i1}} \dots \int_{-\infty}^{z_{ip}} \int_0^1 \nu u^{\nu-1} \phi_p(\mathbf{z}_i | \boldsymbol{\mu}_z, u^{-1} \boldsymbol{\Sigma}_z) du dz_{ip} \dots dz_{i1} \\ &= \int_0^1 \nu u^{\nu-1} \int_{-\infty}^{z_{i1}} \dots \int_{-\infty}^{z_{ip}} \phi_p(\mathbf{z}_i | \boldsymbol{\mu}_z, u^{-1} \boldsymbol{\Sigma}_z) dz_{ip} \dots dz_{i1} du \\ &= \int_0^1 \nu u^{\nu-1} \Phi_p(\mathbf{z}_i | \boldsymbol{\mu}_z, u^{-1} \boldsymbol{\Sigma}_z) du, \end{aligned}$$

onde $\Phi_p(\cdot | \boldsymbol{\mu}, \boldsymbol{\Sigma})$ é a fda da distribuição normal p-variada. Neste trabalho, a regra do trapézio foi usada com $m = 1000$ partições do intervalo $(0, 1)$, e para calcular a fda da distribuição normal p-variada foi utilizado o pacote do R *mvtnorm* disponível no CRAN.

A log-verossimilhança associada com a amostra completa é

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^n \log L_i(\boldsymbol{\theta}) \quad (2.21)$$

2.2.2 Algoritmo MCECM

O algoritmo EM (Dempster et al., 1977) é um popular algoritmo iterativo para calcular estimativas de parâmetros via máxima verossimilhança em modelos com dados faltantes ou em modelos que podem ser formulados como tal. Em circunstâncias como as que prevalecem aqui, a maximização da função de log-verossimilhança com base nos dados observados é difícil de executar devido à presença de integrais sem solução analítica. Ainda mais, algumas vezes a maximização tem que ser realizada por blocos do parâmetro $\boldsymbol{\theta}$ utilizando o algoritmo ECM (Meng e Rubin, 1993). No entanto, em algumas aplicações do algoritmo EM (ou ECM), o passo E não pode ser obtido analiticamente e deve ser calculado por simulação. Wei e Tanner (1990) propuseram o algoritmo EM Monte Carlo (MCEM), no qual o passo E é substituído por uma aproximação de Monte Carlo baseado

em um grande número de simulações independentes dos dados faltantes. O algoritmo MCEM pode ser resumido nos seguintes passos:

Passo E: Calcule a esperança condicional (na log-verossimilhança) dos dados faltantes condicionado as variáveis observadas e à estimativa $\hat{\boldsymbol{\theta}}^{(k)}$ na k -ésima etapa do algoritmo, via integração por Monte Carlo, ou seja, simule M conjuntos de valores para $x, U | \mathbf{Z}_o, \hat{\boldsymbol{\theta}}^{(k)}$ e calcule

$$Q(\boldsymbol{\theta} | \hat{\boldsymbol{\theta}}^{(k)}) = E \left[\ell_c(\boldsymbol{\theta} | \mathbf{Z}_c) | \mathbf{Z}_o, \hat{\boldsymbol{\theta}}^{(k)} \right] \approx \frac{1}{M} \sum_{i=1}^M \ell_c(\boldsymbol{\theta} | \mathbf{Z}_c)$$

Passo M: Maximizar $Q(\cdot | \hat{\boldsymbol{\theta}}^{(k)})$ em relação a $\boldsymbol{\theta}$, ou seja, obter

$$\hat{\boldsymbol{\theta}}^{(k+1)} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} Q(\boldsymbol{\theta} | \hat{\boldsymbol{\theta}}^{(k)})$$

O *MEMC-SI* pode ser formulado numa representação hierárquica flexível que é útil para a obtenção das derivadas. É obtida através das equações (2.3), (2.6) e (2.7) e é dado por

$$\mathbf{Z}_i | x_i, U_i = u_i \stackrel{\text{ind}}{\sim} N_p(\mathbf{a} + \mathbf{b}x_i, u_i^{-1}\boldsymbol{\Omega}), \quad (2.22)$$

$$x_i | U_i = u_i \stackrel{\text{ind}}{\sim} N(\mu_x, u_i^{-1}\sigma_x^2), \quad (2.23)$$

$$U_i \stackrel{\text{iid}}{\sim} \text{Beta}(\nu, 1), \quad i = 1, \dots, n. \quad (2.24)$$

Para obter (2.22), sabemos por (2.8) que $\mathbf{Z}_i \sim NI_p(\boldsymbol{\mu}_z, \boldsymbol{\Sigma}_z; H(u_i; \nu))$, em que $\boldsymbol{\mu}_z = \mathbf{a} + \mathbf{b}\mu_x$ e $\boldsymbol{\Sigma}_z = \sigma_x^2 \mathbf{b}\mathbf{b}^\top + \boldsymbol{\Omega}$. Então \mathbf{Z}_i pode ser expresso como

$$\mathbf{Z}_i = \boldsymbol{\mu}_z + U_i^{-1/2} \mathbf{W}, \quad \text{sendo } \mathbf{W} \sim N_p(\mathbf{0}, \boldsymbol{\Sigma}_z) \text{ e } U_i \sim H(u_i; \nu).$$

Então, $\mathbf{Z}_i | U_i = u_i \sim N_p(\boldsymbol{\mu}_z, u_i^{-1}\boldsymbol{\Sigma}_z)$.

Portanto,

$$\begin{aligned} \mathbf{Z}_i | x_i, U_i = u_i &\stackrel{\text{ind}}{\sim} N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \quad \text{sendo } \boldsymbol{\mu} = \mathbf{a} + \mathbf{b}x_i \\ \boldsymbol{\Sigma} &= u_i^{-1}\boldsymbol{\Sigma}_z = u_i^{-1}(\sigma_x^2 \mathbf{b}\mathbf{b}^\top + \boldsymbol{\Omega}) = u_i^{-1}\boldsymbol{\Omega} \end{aligned}$$

ou seja $\mathbf{Z}_i | x_i, U_i = u_i \stackrel{\text{ind}}{\sim} N_p(\mathbf{a} + \mathbf{b}x_i, u_i^{-1}\boldsymbol{\Omega})$.

A proposição seguinte é útil para obter as esperanças no passo E do algoritmo MCECM, que será utilizado para calcular as estimativas de máxima verossimilhança dos parâmetros no *MEMC-SI*.

Proposição 6. Para o MEMC-SI,

$$E[U_i | \mathbf{Z}_i = \mathbf{z}_i] = \frac{p + 2\nu}{\delta_i} \frac{P_1(\frac{p}{2} + \nu + 1, \frac{\delta_i}{2})}{P_1(\frac{p}{2} + \nu, \frac{\delta_i}{2})},$$

em que δ_i é dado na equação (1.4) e $P_x(a, b)$ denota a fda da distribuição Gama(a, b), com média $\frac{a}{b}$, avaliada em x .

Prova:

Considerando $\mathbf{Z}_i \sim Sl_p(\boldsymbol{\mu}_z, \boldsymbol{\Sigma}_z, \nu)$, o que implica que $\mathbf{Z}_i | U_i = u_i \sim N_p(\boldsymbol{\mu}_z, u_i^{-1} \boldsymbol{\Sigma}_z)$ e pela representação estocástica (2.14), $U_i \sim Beta(\nu, 1)$.

Também

$$\begin{aligned} f(u_i | \mathbf{z}_i) &= \frac{f(\mathbf{z}_i | u_i) f(u_i)}{f(\mathbf{z}_i)} = \frac{\frac{e^{-\frac{1}{2}(\mathbf{z}_i - \boldsymbol{\mu}_z)^\top [u_i^{-1} \boldsymbol{\Sigma}_z]^{-1} (\mathbf{z}_i - \boldsymbol{\mu}_z)}}{(2\pi)^{\frac{p}{2}} |u_i^{-1} \boldsymbol{\Sigma}_z|^{\frac{1}{2}}} \nu u_i^{\nu-1}}{\frac{\nu}{(2\pi)^{\frac{p}{2}} |\boldsymbol{\Sigma}_z|^{\frac{1}{2}}} \int_0^1 u_i^{\frac{p}{2} + \nu - 1} e^{-u_i \frac{\delta_i}{2}} du_i} \\ &= \frac{e^{-u_i \frac{\delta_i}{2}} |\boldsymbol{\Sigma}_z|^{\frac{1}{2}} u_i^{\nu-1} (\frac{\delta_i}{2})^{\frac{p}{2} + \nu}}{u_i^{-\frac{p}{2}} |\boldsymbol{\Sigma}_z|^{\frac{1}{2}} \Gamma(\frac{p}{2} + \nu) \int_0^1 \frac{(\frac{\delta_i}{2})^{\frac{p}{2} + \nu}}{\Gamma(\frac{p}{2} + \nu)} u_i^{\frac{p}{2} + \nu - 1} e^{-u_i \frac{\delta_i}{2}} du_i} = \frac{u_i^{\frac{p}{2} + \nu - 1} e^{-u_i \frac{\delta_i}{2}} (\frac{\delta_i}{2})^{\frac{p}{2} + \nu}}{\Gamma(\frac{p}{2} + \nu) P_1(\frac{p}{2} + \nu, \frac{\delta_i}{2})}. \end{aligned}$$

Portanto,

$$E[U_i | \mathbf{Z}_i] = \int_0^1 u_i f(u_i | \mathbf{z}_i) du_i = \frac{p + 2\nu}{\delta_i} \frac{P_1(\frac{p}{2} + \nu + 1, \frac{\delta_i}{2})}{P_1(\frac{p}{2} + \nu, \frac{\delta_i}{2})}.$$

2.2.2.1 Passo E

Seja $\mathbf{Z} = (\mathbf{Z}_1^\top, \dots, \mathbf{Z}_n^\top)^\top$, $\mathbf{x} = (x_1, \dots, x_n)^\top$, $\mathbf{u} = (u_1, \dots, u_n)^\top$ e $\boldsymbol{\theta} = (\boldsymbol{\alpha}^\top, \boldsymbol{\beta}^\top, \mu_x, \sigma_x^2, \boldsymbol{\phi}^\top)^\top$ o vetor com todos os parâmetros no modelo. Além de constantes que não dependem de $\boldsymbol{\theta}$, a log-verossimilhança completa associada aos dados completos $\mathbf{Z}_c = \{\mathbf{Z}, \mathbf{x}, \mathbf{u}\}$ é dada por

$$\begin{aligned} \ell_c(\boldsymbol{\theta} | \mathbf{Z}_c) &\propto -\frac{n}{2} \sum_{j=1}^p \log \phi_j^2 - \frac{1}{2} \sum_{i=1}^n u_i (\mathbf{Z}_i - \mathbf{a} - \mathbf{b}x_i)^\top \boldsymbol{\Omega}^{-1} (\mathbf{Z}_i - \mathbf{a} - \mathbf{b}x_i) \\ &\quad - \frac{n}{2} \log \sigma_x^2 - \frac{1}{2\sigma_x^2} \sum_{i=1}^n u_i (x_i - \mu_x)^2. \end{aligned} \tag{2.25}$$

Suponha que na k -ésima etapa do algoritmo obtemos uma estimativa $\hat{\boldsymbol{\theta}}^{(k)}$ de $\boldsymbol{\theta}$. O passo E consiste do cálculo da esperança condicional

$$Q(\boldsymbol{\theta} | \hat{\boldsymbol{\theta}}^{(k)}) = E_{\hat{\boldsymbol{\theta}}^{(k)}} [\ell_c(\boldsymbol{\theta} | \mathbf{Z}_c) | \mathbf{V}],$$

em que $E_{\hat{\boldsymbol{\theta}}^{(k)}}$ significa que a esperança está sendo afetada usando $\hat{\boldsymbol{\theta}}^{(k)}$ como o valor verdadeiro do parâmetro e $\mathbf{V} = (\mathbf{V}_1^\top, \dots, \mathbf{V}_n^\top)^\top$. O passo M consiste em maximizar $Q(\cdot|\hat{\boldsymbol{\theta}}^{(k)})$ em $\boldsymbol{\theta}$. Para fazer isso, observe que a função $Q(\cdot|\hat{\boldsymbol{\theta}}^{(k)})$ pode ser decomposto em

$$Q(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}}^{(k)}) = Q_1(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\phi}|\hat{\boldsymbol{\theta}}^{(k)}) + Q_2(\mu_x, \sigma_x^2|\hat{\boldsymbol{\theta}}^{(k)}), \quad (2.26)$$

em que $\boldsymbol{\phi} = (\phi_1^2, \dots, \phi_p^2)$,

$$\begin{aligned} Q_1(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\phi}|\hat{\boldsymbol{\theta}}^{(k)}) &= E_{\hat{\boldsymbol{\theta}}^{(k)}} \left[-\frac{n}{2} \sum_{j=1}^p \log \phi_j^2 - \frac{1}{2} \sum_{i=1}^n u_i (\mathbf{Z}_i - \mathbf{a} - \mathbf{b}x_i)^\top \boldsymbol{\Omega}^{-1} (\mathbf{Z}_i - \mathbf{a} - \mathbf{b}x_i) | \mathbf{V} \right] \\ Q_2(\mu_x, \sigma_x^2|\hat{\boldsymbol{\theta}}^{(k)}) &= E_{\hat{\boldsymbol{\theta}}^{(k)}} \left[-\frac{n}{2} \log \sigma_x^2 - \frac{1}{2\sigma_x^2} \sum_{i=1}^n u_i (x_i - \mu_x)^2 | \mathbf{V} \right]. \end{aligned} \quad (2.27)$$

Dada esta decomposição, podemos reduzir o problema à maximização de duas funções independentes, procurando por pontos críticos de $Q_1(\cdot|\hat{\boldsymbol{\theta}}^{(k)})$ e $Q_2(\cdot|\hat{\boldsymbol{\theta}}^{(k)})$ separadamente. Expandindo as expressões de $Q_1(\cdot|\hat{\boldsymbol{\theta}}^{(k)})$ e $Q_2(\cdot|\hat{\boldsymbol{\theta}}^{(k)})$ e tomando esperanças, segue que

$$\begin{aligned} Q_1(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\phi}|\hat{\boldsymbol{\theta}}^{(k)}) &= -\frac{n}{2} \sum_{j=1}^p \log \phi_j^2 - \frac{1}{2} \sum_{i=1}^n \left[\text{tr} \left(\boldsymbol{\Omega}^{-1} \widehat{u\mathbf{z}_i^2} \right) - 2\mathbf{a}^\top \boldsymbol{\Omega}^{-1} \widehat{u\mathbf{z}_i} - 2\widehat{u\mathbf{x}\mathbf{z}_i} \boldsymbol{\Omega}^{-1} \mathbf{b} \right. \\ &\quad \left. + \mathbf{a}^\top \boldsymbol{\Omega}^{-1} \mathbf{a} \widehat{u_i} + 2\mathbf{a}^\top \boldsymbol{\Omega}^{-1} \mathbf{b} \widehat{u\mathbf{x}_i} + \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b} \widehat{u\mathbf{x}_i^2} \right], \end{aligned} \quad (2.28)$$

$$Q_2(\mu_x, \sigma_x^2|\hat{\boldsymbol{\theta}}^{(k)}) = -\frac{n}{2} \log \sigma_x^2 - \frac{1}{2\sigma_x^2} \sum_{i=1}^n \left(\widehat{u\mathbf{x}_i^2} - 2\mu_x \widehat{u\mathbf{x}_i} + \mu_x^2 \widehat{u_i} \right),$$

onde $\text{tr}(\cdot)$ denota o traço de uma matriz,

$$\begin{aligned} \widehat{u\mathbf{z}_i^2} &= E[U_i \mathbf{Z}_i \mathbf{Z}_i^\top | \mathbf{V}_i], & \widehat{u\mathbf{z}_i} &= E[U_i \mathbf{Z}_i | \mathbf{V}_i], \\ \widehat{u_i} &= E[U_i | \mathbf{V}_i], & \widehat{u\mathbf{x}\mathbf{z}_i} &= E[U_i x_i \mathbf{Z}_i^\top | \mathbf{V}_i], \\ \widehat{u\mathbf{x}_i} &= E[U_i x_i | \mathbf{V}_i], & \widehat{u\mathbf{x}_i^2} &= E[U_i x_i^2 | \mathbf{V}_i], \end{aligned}$$

sendo omitido $\hat{\boldsymbol{\theta}}^{(k)}$ para simplificar a notação.

Para obter expressões para estas esperanças, usaremos uma propriedade da esperança condicional: Se X e Y são vetores aleatórios arbitrários e $f(\cdot)$ é uma função mensurável, então

$$E[E(X|Y)|f(Y)] = E[X|f(Y)] \quad (2.29)$$

Para uma prova, ver Ash (2000, Teorema 5.5.10). Agora, observe que por (2.10), \mathbf{V}_i é uma função de \mathbf{Z}_i . Então, pela propriedade da equação (2.29), podemos escrever

$$\begin{aligned}\widehat{u\mathbf{z}_i^2} &= E[U_i \mathbf{Z}_i \mathbf{Z}_i^\top | \mathbf{V}_i] = E[E[U_i \mathbf{Z}_i \mathbf{Z}_i^\top | \mathbf{Z}_i] | \mathbf{V}_i], \\ \widehat{u\mathbf{z}_i} &= E[U_i \mathbf{Z}_i | \mathbf{V}_i] = E[E[U_i \mathbf{Z}_i | \mathbf{Z}_i] | \mathbf{V}_i] \quad \text{e} \\ \widehat{u_i} &= E[U_i | \mathbf{V}_i] = E[E[U_i | \mathbf{Z}_i] | \mathbf{V}_i].\end{aligned}\tag{2.30}$$

Utilizando a Proposição 6 sobre esperança condicional $E[U_i | \mathbf{Z}_i]$ obtemos os resultados das seguintes expressões para $\widehat{u_i}$, $\widehat{u\mathbf{z}_i}$ e $\widehat{u\mathbf{z}_i^2}$, considerando três diferentes casos:

i) O i -ésimo indivíduo não tem componentes censurados. Aqui, $\mathbf{V}_i = \mathbf{Z}_i$, então

$$\begin{aligned}\widehat{u_i} &= E[U_i | \mathbf{V}_i] = E[U_i | \mathbf{Z}_i] = \frac{p + 2\nu P_1(\frac{p}{2} + \nu + 1, \frac{\delta_i}{2})}{\delta_i P_1(\frac{p}{2} + \nu, \frac{\delta_i}{2})}, \\ \widehat{u\mathbf{z}_i} &= E[U_i \mathbf{Z}_i | \mathbf{V}_i] = E[U_i \mathbf{Z}_i | \mathbf{Z}_i] = E[U_i | \mathbf{Z}_i] \mathbf{z}_i = \widehat{u_i} \mathbf{z}_i, \quad \text{e} \\ \widehat{u\mathbf{z}_i^2} &= E[U_i \mathbf{Z}_i \mathbf{Z}_i^\top | \mathbf{V}_i] = E[U_i \mathbf{Z}_i \mathbf{Z}_i^\top | \mathbf{Z}_i] = E[U_i | \mathbf{Z}_i] \mathbf{z}_i \mathbf{z}_i^\top = \widehat{u_i} \mathbf{z}_i \mathbf{z}_i^\top,\end{aligned}$$

com δ_i dado em (1.4) e $P_x(a, b)$ é a fda da distribuição Gama(a,b), com média $\frac{a}{b}$, avaliada no ponto x .

ii) O i -ésimo indivíduo tem apenas componentes censurados. Por (2.10), $\mathbf{Z}_i \leq \boldsymbol{\kappa}_i$, em que $\boldsymbol{\kappa}_i$ é o vetor com os níveis de censura para o indivíduo i . Assim, por (2.17) e a definição da distribuição slash truncada, temos que $\mathbf{Z}_i | (\mathbf{Z}_i \leq \boldsymbol{\kappa}_i) \sim TSl_p(\boldsymbol{\mu}_z, \boldsymbol{\Sigma}_z, \nu; \mathbb{D}_i)$, onde \mathbb{D}_i é como em (2.15) com $\mathbf{d} = \boldsymbol{\kappa}_i$. Então temos

$$\begin{aligned}\widehat{u_i} &= E[E[U_i | \mathbf{Z}_i] | \mathbf{V}_i] = E[E[U_i | \mathbf{Z}_i] | \mathbf{Z}_i \leq \boldsymbol{\kappa}_i] = E\left[\frac{p + 2\nu P_1(\frac{p}{2} + \nu + 1, \frac{\delta_i}{2})}{\delta_i P_1(\frac{p}{2} + \nu, \frac{\delta_i}{2})} | \mathbf{Z}_i \leq \boldsymbol{\kappa}_i\right], \\ \widehat{u\mathbf{z}_i} &= E[E[U_i \mathbf{Z}_i | \mathbf{Z}_i] | \mathbf{V}_i] = E\left[\frac{p + 2\nu P_1(\frac{p}{2} + \nu + 1, \frac{\delta_i}{2})}{\delta_i P_1(\frac{p}{2} + \nu, \frac{\delta_i}{2})} \mathbf{Z}_i | \mathbf{Z}_i \leq \boldsymbol{\kappa}_i\right], \\ \widehat{u\mathbf{z}_i^2} &= E[E[U_i \mathbf{Z}_i \mathbf{Z}_i^\top | \mathbf{Z}_i] | \mathbf{V}_i] = E\left[\frac{p + 2\nu P_1(\frac{p}{2} + \nu + 1, \frac{\delta_i}{2})}{\delta_i P_1(\frac{p}{2} + \nu, \frac{\delta_i}{2})} \mathbf{Z}_i \mathbf{Z}_i^\top | \mathbf{Z}_i \leq \boldsymbol{\kappa}_i\right].\end{aligned}$$

iii) O i -ésimo indivíduo tem componentes censurados e não censurados. Decompomos o vetor \mathbf{V}_i em dois subvetores, \mathbf{Z}_i^o e $\boldsymbol{\kappa}_i^c$, correspondentes às observações não censuradas e aos níveis de censura, respectivamente. Particionamos o vetor \mathbf{Z}_i como $\mathbf{Z}_i = \text{vec}(\mathbf{Z}_i^o, \mathbf{Z}_i^c)$. Os componentes são censurados se e somente se $\mathbf{Z}_i^c \leq \boldsymbol{\kappa}_i^c$.

Também temos que $\mathbf{Z}_i | (\mathbf{Z}_i^c \leq \boldsymbol{\kappa}_i^c) \sim TSl_p(\boldsymbol{\mu}_z, \boldsymbol{\Sigma}_z, \nu; \mathbb{D}_i^c)$, com

$$\mathbb{D}_i^c = \{(x_1, \dots, x_p) \in \mathbb{R}^p; \ x_i \leq \boldsymbol{\kappa}_i^c, i \in \mathcal{C}\}, \quad (2.31)$$

em que \mathcal{C} é o conjunto de índices para os componentes censurados - consequentemente, fazemos $d_i = +\infty$ para $i \notin \mathcal{C}$ em (2.15).

Utilizando a Proposição 5, com \mathbf{Z}_i^o e \mathbf{Z}_i^c desempenhando o papel de \mathbf{Y}_1 e \mathbf{Y}_2 , respectivamente, e a Definição 1, temos que

$$\mathbf{Z}_i^c | \mathbf{Z}_i^o = \mathbf{z}_i^o, \mathbf{Z}_i^c \leq \boldsymbol{\kappa}_i^c \sim TSl_{p_c}(\boldsymbol{\mu}_z^{co}, \boldsymbol{\Sigma}_z^{cc.o}, \nu + p_o; \mathbb{D}_i^c),$$

em que p_o e p_c são as dimensões dos vetores \mathbf{Z}_i^o e \mathbf{Z}_i^c , respectivamente, e $\boldsymbol{\mu}_z^{co}$ e $\boldsymbol{\Sigma}_z^{cc.o}$ são dados em (2.19) e (2.20), respectivamente. Assim,

$$\begin{aligned} \widehat{u}_i &= E[U_i | \mathbf{V}_i] = E[E[U_i | \mathbf{Z}_i] | \mathbf{V}_i] = E \left[\frac{p + 2\nu}{\delta_i} \frac{P_1(\frac{p}{2} + \nu + 1, \frac{\delta_i}{2})}{P_1(\frac{p}{2} + \nu, \frac{\delta_i}{2})} \middle| \mathbf{Z}_i^o = \mathbf{z}_i^o, \mathbf{Z}_i^c \leq \boldsymbol{\kappa}_i^c \right], \\ \widehat{u\mathbf{z}}_i &= E[E[U_i \mathbf{Z}_i | \mathbf{Z}_i] | \mathbf{V}_i] = E \left[\text{vec}(E[U_i | \mathbf{Z}_i] \mathbf{Z}_i^o, E[U_i | \mathbf{Z}_i] \mathbf{Z}_i^c) \middle| \mathbf{Z}_i^o = \mathbf{z}_i^o, \mathbf{Z}_i^c \leq \boldsymbol{\kappa}_i^c \right] \\ &= \text{vec} \left(\widehat{u}_i \mathbf{z}_i^o, \widehat{u\mathbf{z}}_i^c \right), \\ \widehat{u\mathbf{z}}_i^2 &= E[U_i \mathbf{Z}_i \mathbf{Z}_i^\top | \mathbf{V}_i] = E \left[E[U_i | \mathbf{Z}_i] \begin{bmatrix} \mathbf{Z}_i^o \mathbf{Z}_i^{o\top} & \mathbf{Z}_i^o \mathbf{Z}_i^{c\top} \\ \mathbf{Z}_i^c \mathbf{Z}_i^{o\top} & \mathbf{Z}_i^c \mathbf{Z}_i^{c\top} \end{bmatrix} \middle| \mathbf{Z}_i^o = \mathbf{z}_i^o, \mathbf{Z}_i^c \leq \boldsymbol{\kappa}_i^c \right] \\ &= \begin{bmatrix} \widehat{u}_i \mathbf{z}_i^o \mathbf{z}_i^{o\top} & \mathbf{z}_i^o \widehat{u\mathbf{z}}_i^{c\top} \\ \widehat{u\mathbf{z}}_i^c \mathbf{z}_i^{o\top} & \widehat{u\mathbf{z}}_i^c \mathbf{z}_i^{c\top} \end{bmatrix}, \end{aligned}$$

em que

$$\begin{aligned} \widehat{u\mathbf{z}}_i^c &= E \left[\frac{p + 2\nu}{\delta_i} \frac{P_1(\frac{p}{2} + \nu + 1, \frac{\delta_i}{2})}{P_1(\frac{p}{2} + \nu, \frac{\delta_i}{2})} \mathbf{Z}_i^c \middle| \mathbf{Z}_i^o = \mathbf{z}_i^o, \mathbf{Z}_i^c \leq \boldsymbol{\kappa}_i^c \right], \text{ e} \\ \widehat{u\mathbf{z}}_i^c \mathbf{z}_i^{c\top} &= E \left[\frac{p + 2\nu}{\delta_i} \frac{P_1(\frac{p}{2} + \nu + 1, \frac{\delta_i}{2})}{P_1(\frac{p}{2} + \nu, \frac{\delta_i}{2})} \mathbf{Z}_i^c \mathbf{Z}_i^{c\top} \middle| \mathbf{Z}_i^o = \mathbf{z}_i^o, \mathbf{Z}_i^c \leq \boldsymbol{\kappa}_i^c \right]. \end{aligned}$$

Em relação aos restantes valores esperados, temos

$$\begin{aligned} E[x_i U_i | \mathbf{V}_i = \mathbf{v}_i] &= \iint x_i u_i \pi(x_i, u_i | \mathbf{v}_i) dx_i du_i \\ &= E[x_i | U_i = u_i, \mathbf{V}_i = \mathbf{v}_i] E[U_i | \mathbf{V}_i = \mathbf{v}_i] \end{aligned} \quad (2.32)$$

em que $\pi(\cdot)$ denota uma função de densidade de probabilidade genérica.

Pela propriedade de esperança condicional dada na equação (2.29), temos

$$E[x_i|U_i, \mathbf{V}_i] = E[E[x_i|U_i, \mathbf{Z}_i]|U_i, \mathbf{V}_i].$$

Consequentemente,

$$\begin{aligned} \widehat{u}x_i &= E[E[x_i|U_i, \mathbf{Z}_i]|U_i, \mathbf{V}_i] E[U_i|\mathbf{V}_i] = E \left[\frac{\mu_x + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} (\mathbf{Z}_i - \mathbf{a})}{1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b}} \middle| U_i, \mathbf{V}_i \right] E[U_i|\mathbf{V}_i] \\ &= \frac{\mu_x E[U_i|\mathbf{V}_i] + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} E[\mathbf{Z}_i|U_i, \mathbf{V}_i] E[U_i|\mathbf{V}_i] - \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{a} E[U_i|\mathbf{V}_i]}{1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b}} \end{aligned}$$

mas

$$E[\mathbf{Z}_i|U_i, \mathbf{V}_i] E[U_i|\mathbf{V}_i] = E[U_i \mathbf{Z}_i|\mathbf{V}_i] \equiv \widehat{u}z_i.$$

Por (2.9) temos que $\boldsymbol{\mu}_z = \mathbf{a} + \mathbf{b}\mu_x$; então, $\mathbf{a} = \boldsymbol{\mu}_z - \mathbf{b}\mu_x$. Assim

$$\begin{aligned} \widehat{u}x_i &= \frac{\mu_x \widehat{u}i + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \widehat{u}z_i - \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} (\boldsymbol{\mu}_z - \mathbf{b}\mu_x) \widehat{u}i}{1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b}} \\ &= \frac{\mu_x \widehat{u}i + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \widehat{u}z_i - \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \boldsymbol{\mu}_z \widehat{u}i + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b} \mu_x \widehat{u}i}{1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b}} \\ &= \frac{(1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b}) \mu_x \widehat{u}i + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} (\widehat{u}z_i - \boldsymbol{\mu}_z \widehat{u}i)}{1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b}} = \mu_x \widehat{u}i + \varphi (\widehat{u}z_i - \boldsymbol{\mu}_z \widehat{u}i). \end{aligned} \quad (2.33)$$

em que

$$\varphi = \frac{\sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1}}{1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b}}. \quad (2.34)$$

De forma semelhante, obtemos

$$\widehat{u}x_i^2 = E[U_i x_i^2 | \mathbf{V}_i] = E[U_i | \mathbf{V}_i] E[x_i^2 | U_i, \mathbf{V}_i] = E[U_i | \mathbf{V}_i] E[E[x_i^2 | U_i, \mathbf{Z}_i] | U_i, \mathbf{V}_i]$$

Por propriedade de variância e pela Proposição 1, temos

$$\begin{aligned} E[x_i^2 | U_i, \mathbf{Z}_i] &= \text{Var}[x_i | U_i, \mathbf{Z}_i] + (E[x_i | U_i, \mathbf{Z}_i])^2 \\ &= \frac{\sigma_x^2}{u_i (1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b})} + \left(\frac{\mu_x + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} (\mathbf{z}_i - \mathbf{a})}{1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b}} \right)^2 \end{aligned} \quad (2.35)$$

Por (2.9), $\mathbf{a} = \boldsymbol{\mu}_z - \mathbf{b}\mu_x$; então o segundo termo de (2.35)

$$\begin{aligned}
\left(\frac{\mu_x + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} (\mathbf{z}_i - \mathbf{a})}{1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b}} \right)^2 &= \left(\frac{\mu_x + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} (\mathbf{z}_i - \boldsymbol{\mu}_z + \mathbf{b}\mu_x)}{1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b}} \right)^2 \\
&= \left(\frac{(1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b})\mu_x}{1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b}} + \frac{\sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{z}_i}{1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b}} - \frac{\sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \boldsymbol{\mu}_z}{1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b}} \right)^2 \\
&= (\mu_x + \boldsymbol{\varphi} \mathbf{z}_i - \boldsymbol{\varphi} \boldsymbol{\mu}_z)^2 = (\mu_x + \boldsymbol{\varphi} \mathbf{z}_i - \boldsymbol{\varphi} \boldsymbol{\mu}_z)(\mu_x + \boldsymbol{\varphi} \mathbf{z}_i - \boldsymbol{\varphi} \boldsymbol{\mu}_z)^\top \\
&= \mu_x^2 + 2\mu_x \boldsymbol{\varphi} \mathbf{z}_i - 2\mu_x \boldsymbol{\varphi} \boldsymbol{\mu}_z + \boldsymbol{\varphi} \mathbf{z}_i \mathbf{z}_i^\top \boldsymbol{\varphi}^\top - \boldsymbol{\varphi} \mathbf{z}_i \boldsymbol{\mu}_z^\top \boldsymbol{\varphi}^\top - \boldsymbol{\varphi} \boldsymbol{\mu}_z \mathbf{z}_i^\top \boldsymbol{\varphi}^\top + \boldsymbol{\varphi} \boldsymbol{\mu}_z \boldsymbol{\mu}_z^\top \boldsymbol{\varphi}^\top
\end{aligned}$$

pelo que

$$\widehat{ux_i^2} = \Lambda + \mu_x^2 \widehat{u_i} + 2\mu_x \boldsymbol{\varphi} (\widehat{uz_i} - \boldsymbol{\mu}_z \widehat{u_i}) + \boldsymbol{\varphi} (\widehat{uz_i^2} - \widehat{uz_i} \boldsymbol{\mu}_z^\top - \boldsymbol{\mu}_z \widehat{uz_i}^\top + \boldsymbol{\mu}_z \boldsymbol{\mu}_z^\top \widehat{u_i}) \boldsymbol{\varphi}^\top. \quad (2.36)$$

$$\widehat{uz_i} = E[U_i x_i \mathbf{Z}_i^\top | \mathbf{V}_i] = \mu_x \widehat{u_i}^\top + \boldsymbol{\varphi} (\widehat{uz_i^2} - \boldsymbol{\mu}_z \widehat{uz_i}^\top), \quad (2.37)$$

em que

$$\Lambda = \frac{\sigma_x^2}{1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b}}. \quad (2.38)$$

2.2.2.2 Passo CM

Quando o passo M do algoritmo EM é complicado, este pode ser amenizado realizando o processo de maximização condicional a alguma função dos parâmetros que estão sendo estimados. Dada a estimativa atual $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}^{(k)}$ na k -ésima etapa, o passo CM do algoritmo ECM (Meng e Rubin, 1993) consiste na maximização condicional da função Q dada em (2.26). O ECM substitui cada passo M do algoritmo EM de Dempster et al. (1977) por uma sequência de S passos de maximização condicional, chamados passos CM, cada um dos quais maximiza a função Q sobre $\boldsymbol{\theta}$ mas com alguma função vetorial de $\boldsymbol{\theta}$, digamos $(g_1(\boldsymbol{\theta}), \dots, g_s(\boldsymbol{\theta}))$, fixado em seu valor anterior. Em nosso caso, por exemplo, primeiro maximizamos condicionalmente a função $Q_1(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\phi} | \widehat{\boldsymbol{\theta}}^{(k)})$ em (2.28) sobre $\boldsymbol{\alpha}$ fixando os valores $\boldsymbol{\beta} = \widehat{\boldsymbol{\beta}}^{(k)}$ e $\boldsymbol{\phi} = \widehat{\boldsymbol{\phi}}^{(k)}$. Então maximizamos $Q_1(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\phi} | \widehat{\boldsymbol{\theta}}^{(k)})$ sobre $\boldsymbol{\beta}$ fixando os valores $\boldsymbol{\alpha} = \widehat{\boldsymbol{\alpha}}^{(k+1)}$ e $\boldsymbol{\phi} = \widehat{\boldsymbol{\phi}}^{(k)}$ e assim por diante. Obtemos as seguintes expressões

fechadas (cujos cálculos são dados no Apêndice B):

$$\begin{aligned}\widehat{\boldsymbol{\alpha}}^{(k+1)} &= \overline{\mathbf{z}}_u^{(k)} - \overline{x}_u^{(k)} \widehat{\boldsymbol{\beta}}^{(k)}, \\ \widehat{\boldsymbol{\beta}}^{(k+1)} &= \frac{n \overline{\widehat{u}}^{(k)} \sum_{i=1}^n \widehat{u} \widehat{x} \widehat{\mathbf{z}}_i^{*\top(k)} - \sum_{i=1}^n \widehat{u} \widehat{\mathbf{z}}_i^{*(k)} \sum_{i=1}^n \widehat{u} \widehat{x}_i^{(k)}}{n \overline{\widehat{u}}^{(k)} \sum_{i=1}^n \widehat{u} \widehat{x}_i^2{}^{(k)} - \left(\sum_{i=1}^n \widehat{u} \widehat{x}_i^{(k)} \right)^2}, \\ \widehat{\phi}_1^{2(k+1)} &= \frac{1}{n} \sum_{i=1}^n \left(\widehat{u} \widehat{\mathbf{z}}_{i11}^2{}^{(k)} - 2 \widehat{u} \widehat{x} \widehat{\mathbf{z}}_{i1}^{(k)} + \widehat{u} \widehat{x}_i^2{}^{(k)} \right), \\ \widehat{\phi}_{j+1}^2{}^{(k+1)} &= \frac{1}{n} \sum_{i=1}^n \left(\widehat{u} \widehat{\mathbf{z}}_{i(j+1)(j+1)}^2{}^{(k)} + \widehat{u}_i^{(k)} \widehat{\boldsymbol{\alpha}}_j^{2(k+1)} + \widehat{u} \widehat{x}_i^2{}^{(k)} \widehat{\boldsymbol{\beta}}_j^{2(k+1)} + 2 \widehat{u} \widehat{x}_i^{(k)} \widehat{\boldsymbol{\alpha}}_j^{(k+1)} \widehat{\boldsymbol{\beta}}_j^{(k+1)} \right. \\ &\quad \left. - 2 \widehat{u} \widehat{x} \widehat{\mathbf{z}}_{i(j+1)}^{(k)} \widehat{\boldsymbol{\beta}}_j^{(k+1)} - 2 \widehat{u} \widehat{\mathbf{z}}_{i(j+1)}^{(k)} \widehat{\boldsymbol{\alpha}}_j^{(k+1)} \right), \quad j = 1, \dots, r, \\ \widehat{\mu}_x^{(k+1)} &= \overline{x}_u^{(k)},\end{aligned}$$

$$\widehat{\sigma}_x^2{}^{(k+1)} = \frac{1}{n} \sum_{i=1}^n \left(\widehat{u} \widehat{x}_i^2{}^{(k)} - 2 \widehat{u} \widehat{x}_i^{(k)} \widehat{\mu}_x^{(k+1)} + \widehat{u}_i^{(k)} \widehat{\mu}_x^2{}^{(k+1)} \right),$$

$$\begin{aligned}\text{em que } \overline{\mathbf{z}}_u^{(k)} &= \frac{\sum_{i=1}^n \widehat{u} \widehat{\mathbf{z}}_i^{*(k)}}{\sum_{i=1}^n \widehat{u}_i^{(k)}}, \quad \overline{x}_u^{(k)} = \frac{\sum_{i=1}^n \widehat{u} \widehat{x}_i^{(k)}}{\sum_{i=1}^n \widehat{u}_i^{(k)}} \quad \text{e} \quad \overline{\widehat{u}}^{(k)} = \frac{1}{n} \sum_{i=1}^n \widehat{u}_i^{(k)} \quad \text{com} \quad \widehat{u} \widehat{\mathbf{z}}_i^{*(k)} = \\ &(\widehat{u} \widehat{z}_{i2}, \dots, \widehat{u} \widehat{z}_{ip})^\top \text{ e } \widehat{u} \widehat{x} \widehat{\mathbf{z}}_i^{*(k)} = (\widehat{u} \widehat{x} \widehat{z}_{i2}, \dots, \widehat{u} \widehat{x} \widehat{z}_{ip})^\top.\end{aligned}$$

2.2.3 Matriz de informação observada

A fim de obter as estimativas dos erros padrão para o vetor de parâmetros $\boldsymbol{\theta}$, calculamos antes a matriz de informação observada. Sob algumas condições de regularidade, seguimos Lin (2010) para fornecer um método baseado em informação para obter a covariância assintótica dos estimadores de máxima verossimilhança dos parâmetros do *MEMC-SL*. Como definido por Meilijson (1989), a matriz de informação empírica pode ser calculada como

$$\mathbf{I}_e(\boldsymbol{\theta}|\mathbf{Z}) = \sum_{i=1}^n \mathbf{s}(\mathbf{Z}_i|\boldsymbol{\theta}) \mathbf{s}^\top(\mathbf{Z}_i|\boldsymbol{\theta}) - \frac{1}{n} \mathbf{S}(\mathbf{Z}|\boldsymbol{\theta}) \mathbf{S}^\top(\mathbf{Z}|\boldsymbol{\theta}), \quad (2.39)$$

em que $\mathbf{S}(\mathbf{Z}|\boldsymbol{\theta}) = \sum_{i=1}^n \mathbf{s}(\mathbf{Z}_i|\boldsymbol{\theta})$ e $\mathbf{s}(\mathbf{Z}_i|\boldsymbol{\theta})$ é a função escore empírica para a unidade i . De acordo com Louis (1982) é possível relacionar a função escore da log-verossimilhança dos dados incompletos com a esperança condicional da função de log-verossimilhança dos dados completos. Portanto, o escore individual pode ser determinado como

$$\mathbf{s}(\mathbf{Z}_i|\boldsymbol{\theta}) = \frac{\partial \log f(\mathbf{Z}_i|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = E \left[\frac{\partial \ell_{ic}(\boldsymbol{\theta}|\mathbf{Z}_i^c)}{\partial \boldsymbol{\theta}} \middle| \mathbf{V}_i, \boldsymbol{\theta} \right],$$

em que $\ell_{ic}(\boldsymbol{\theta}|\mathbf{Z}_i^c)$ é a log-verossimilhança dos dados completos formada a partir da única observação \mathbf{Z}_i , $i = 1, \dots, n$. Usando as estimativas $\hat{\boldsymbol{\theta}}$ do algoritmo EM, $\mathbf{S}(\mathbf{Z}_i|\hat{\boldsymbol{\theta}}) = 0$, e então (2.39) é dado por

$$\mathbf{I}_e(\hat{\boldsymbol{\theta}}|\mathbf{Z}) = \sum_{i=1}^n \hat{\mathbf{s}}_i \hat{\mathbf{s}}_i^\top, \quad (2.40)$$

em que $\hat{\mathbf{s}}_i = (\hat{\mathbf{s}}_{i,\alpha}, \hat{\mathbf{s}}_{i,\beta}, \hat{\mathbf{s}}_{i,\phi}, \hat{\mathbf{s}}_{i,\mu_x}, \hat{\mathbf{s}}_{i,\sigma_x^2})^\top$ é um vetor $3p$ -dimensional, com componentes dadas por

$$\begin{aligned} \hat{\mathbf{s}}_{i,\alpha} &= (\hat{\mathbf{s}}_{i,\alpha_1}, \dots, \hat{\mathbf{s}}_{i,\alpha_r})^\top = \mathbb{I}_{(p)} \hat{\boldsymbol{\Omega}}^{-1} (\widehat{u\mathbf{z}}_i - \widehat{u}_i \widehat{\mathbf{a}} - \widehat{u x}_i \widehat{\mathbf{b}}), \\ \hat{\mathbf{s}}_{i,\beta} &= (\hat{\mathbf{s}}_{i,\beta_1}, \dots, \hat{\mathbf{s}}_{i,\beta_r})^\top = \mathbb{I}_{(p)} \hat{\boldsymbol{\Omega}}^{-1} (\widehat{u x \mathbf{z}}_i^\top - \widehat{u x}_i \widehat{\mathbf{a}} - \widehat{u x_i^2} \widehat{\mathbf{b}}), \\ \hat{\mathbf{s}}_{i,\phi} &= (\hat{\mathbf{s}}_{i,\phi_1^2}, \dots, \hat{\mathbf{s}}_{i,\phi_p^2})^\top = -\frac{1}{2} \hat{\boldsymbol{\Omega}}^{-1} \mathbf{1}_p + \frac{1}{2} \hat{\boldsymbol{\Omega}}^{-2} \text{diag}(\widehat{a}_i), \\ \hat{\mathbf{s}}_{i,\mu_x} &= \frac{1}{\widehat{\sigma}_x^2} (\widehat{u x}_i - \widehat{u}_i \widehat{\mu}_x), \\ \hat{\mathbf{s}}_{i,\sigma_x^2} &= -\frac{1}{2 \widehat{\sigma}_x^2} + \frac{1}{2 \widehat{\sigma}_x^4} (\widehat{u x_i^2} - 2 \widehat{u x}_i \widehat{\mu}_x + \widehat{u}_i \widehat{\mu}_x^2), \end{aligned}$$

com $\mathbb{I}_{(p)} = [\mathbf{0}, \mathbf{I}_{p-1}]_{(p-1) \times p}$, $\mathbf{1}_p = (1, \dots, 1)^\top$ vetor $p \times 1$ e $\widehat{a}_i = \widehat{u \mathbf{z}_i^2} - 2 \widehat{u \mathbf{z}_i} \widehat{\mathbf{a}}^\top - 2 \widehat{u x \mathbf{z}_i} \widehat{\mathbf{b}}^\top + 2 \widehat{u x}_i \widehat{\mathbf{a}} \widehat{\mathbf{b}}^\top + \widehat{u}_i \widehat{\mathbf{a}} \widehat{\mathbf{a}}^\top + \widehat{u x_i^2} \widehat{\mathbf{b}} \widehat{\mathbf{b}}^\top$. Os cálculos realizados para obter estas expressões são dados no Apêndice B. Após o cálculo da matriz de informação de Fisher, estimamos a matriz de covariâncias para o vetor de parâmetros $\boldsymbol{\theta}$, que é dada pela inversa da matriz de informação de Fisher.

2.3 Distribuição t -Student multivariada

Dizemos que o vetor aleatório $\mathbf{Y} \in \mathbb{R}^p$ tem uma distribuição t -Student com vetor de locação $\boldsymbol{\mu}$, matriz de dispersão $\boldsymbol{\Sigma}$ e ν graus de liberdade, quando sua fdp é dada por

(1.3). A fda de \mathbf{Y} é denotada por $T_p(\cdot|\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$. Se $\nu > 1$, $\boldsymbol{\mu}$ é a média de \mathbf{Y} , e se $\nu > 2$, $\nu(\nu - 2)^{-1}\boldsymbol{\Sigma}$ é a matriz de covariâncias. Se \mathbf{Y} tem distribuição t -Student com fdp dada por (1.3), usamos a notação $\mathbf{Y} \sim t_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$.

O vetor aleatório \mathbf{Y} admite a representação estocástica

$$\mathbf{Y} = \boldsymbol{\mu} + U^{-1/2}\mathbf{Z}, \quad \mathbf{Z} \sim N_p(\mathbf{0}, \boldsymbol{\Sigma}), \quad U \sim \text{Gama}(\nu/2, \nu/2), \quad (2.41)$$

em que \mathbf{Z} e U são independentes, e $\text{Gama}(a, b)$ denota a distribuição gama com média a/b . À medida que ν tende ao infinito, U converge para um com probabilidade um e \mathbf{Y} é aproximadamente distribuído como uma distribuição $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. A partir desta representação podemos deduzir que uma transformação afim $\mathbf{A}\mathbf{Y} + \mathbf{b}$ tem uma distribuição $t_q(\mathbf{A}\boldsymbol{\mu} + \mathbf{b}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^\top, \nu)$, em que \mathbf{A} é uma matriz com dimensão $(q \times p)$ e \mathbf{b} é um vetor q -dimensional. Para uma referência sobre a distribuição t -Student multivariada, ver Kotz e Nadarajah (2004).

Segundo Matos et al. (2013), a família de distribuições t -Student é fechada sob marginalização e condicionamento e é dado na seguinte proposição. Este resultado é útil na implementação do algoritmo EM.

Proposição 7. *Seja $\mathbf{Y} \sim t_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$. Considere a partição de \mathbf{Y} , $\boldsymbol{\mu}$ e $\boldsymbol{\Sigma}$ como*

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{bmatrix}, \quad \boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix} \quad e \quad \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}$$

em que \mathbf{Y}_1 e $\boldsymbol{\mu}_1$ são vetores $p_1 \times 1$ e $\boldsymbol{\Sigma}_{11}$ é uma matriz $p_1 \times p_1$. Então

i) $\mathbf{Y}_1 \sim t_{p_1}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11}, \nu)$;

ii) $\mathbf{Y}_2|\mathbf{Y}_1 = \mathbf{y}_1 \sim t_{p_2}(\boldsymbol{\mu}_{2,1}, \tilde{\boldsymbol{\Sigma}}_{22,1}, \nu + p_1)$, em que

$$\begin{aligned} \boldsymbol{\mu}_{2,1} &= \boldsymbol{\mu}_2 + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}(\mathbf{y}_1 - \boldsymbol{\mu}_1), & \tilde{\boldsymbol{\Sigma}}_{22,1} &= \frac{\nu + \delta_1}{\nu + p_1}\boldsymbol{\Sigma}_{22,1}, \\ \delta_1 &= (\mathbf{y}_1 - \boldsymbol{\mu}_1)^\top \boldsymbol{\Sigma}_{11}^{-1}(\mathbf{y}_1 - \boldsymbol{\mu}_1), & \boldsymbol{\Sigma}_{22,1} &= \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}. \end{aligned}$$

Definição 2. *Seja $\mathbf{Y} \sim t_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$ e \mathbb{D} um conjunto de Borel em \mathbb{R}^p . Dizemos que o vetor aleatório \mathbf{Z} tem uma distribuição t -Student truncada em \mathbb{D} , denotada por $\mathbf{Z} \sim$*

$Tt_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu; \mathbb{D})$, quando \mathbf{Z} tem a mesma distribuição que $\mathbf{Y} | (\mathbf{Y} \in \mathbb{D})$. Neste caso, a fdp de \mathbf{Z} é dada por

$$Tt_p(\mathbf{z} | \boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu; \mathbb{D}) = \frac{t_p(\mathbf{z} | \boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)}{P(\mathbf{Y} \in \mathbb{D})} \mathbb{I}_{\mathbb{D}}(\mathbf{z}),$$

sendo $\mathbb{I}_{\mathbb{D}}(\cdot)$ a função indicadora de \mathbb{D} . Se \mathbb{D} tem a forma dada em (2.15), então usamos a notação $(\mathbf{Y} \in \mathbb{D}) = (\mathbf{Y} \leq \mathbf{d})$, em que $\mathbf{d} = (d_1, \dots, d_p)^\top$.

Seguindo Matos et al. (2016), as seguintes proposições são cruciais para obter as esperanças no passo E do algoritmo EM, que será utilizado para calcular as estimativas de máxima verossimilhança dos parâmetros no *MEMC-t*. As provas são encontradas em Matos et al. (2013). Usaremos as notações $\mathbf{Z}^{(0)} = 1$, $\mathbf{Z}^{(1)} = \mathbf{Z}$ e $\mathbf{Z}^{(2)} = \mathbf{Z}\mathbf{Z}^\top$.

Proposição 8. *Seja $\mathbf{Z} \sim Tt_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu; \mathbb{D})$, onde \mathbb{D} é como em (2.15). Então, para $k = 0, 1, 2$,*

$$E \left[\left(\frac{\nu + p}{\nu + \delta} \right)^r \mathbf{Z}^{(k)} \right] = C_p(\nu, r) \frac{T_p(\mathbf{d} | \boldsymbol{\mu}, \boldsymbol{\Sigma}^*, \nu + 2r)}{T_p(\mathbf{d} | \boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)} E[\mathbf{Y}^{(k)}], \quad (2.42)$$

onde $\nu + 2r > 0$ e

$$\begin{aligned} \mathbf{Y} &\sim Tt_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}^*, \nu + 2r; \mathbb{D}), \\ \boldsymbol{\Sigma}^* &= \frac{\nu}{\nu + 2r} \boldsymbol{\Sigma}, \\ C_p(\nu, r) &= \left(\frac{\nu + p}{\nu} \right)^r \left(\frac{\Gamma((p + \nu)/2) \Gamma((\nu + 2r)/2)}{\Gamma(\nu/2) \Gamma((p + \nu + 2r)/2)} \right). \end{aligned} \quad (2.43)$$

Observe que o cálculo da esperança no lado esquerdo de (2.42) se reduz ao cálculo dos momentos da distribuição *t-Student* truncada em (2.43). Um estudo mais detalhado sobre esses momentos pode ser encontrado em Ho et al. (2012).

Proposição 9. *Seja $\mathbf{Z} \sim Tt_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu; \mathbb{D})$, onde \mathbb{D} é como em (2.15). Considere a partição $\mathbf{Z} = (\mathbf{Z}_1^\top, \mathbf{Z}_2^\top)^\top$ com $\mathbf{Z}_1 : p_1 \times 1$ e $\mathbf{Z}_2 : p_2 \times 1$. Assim, considere as partições $\boldsymbol{\mu} = (\boldsymbol{\mu}_1^\top, \boldsymbol{\mu}_2^\top)^\top$ e $\boldsymbol{\Sigma} = (\boldsymbol{\Sigma}_{ij}), i, j = 1, 2$. Então,*

$$E \left[\left(\frac{\nu + p}{\nu + \delta} \right)^r \mathbf{Z}_2^{(k)} | \mathbf{Z}_1 = \mathbf{z}_1 \right] = \frac{h_p(p_1, \nu, r)}{(\nu + \delta_1)^r} \frac{T_{p_2}(\mathbf{d}_2 | \boldsymbol{\mu}_{2.1}, \tilde{\boldsymbol{\Sigma}}_{22.1}^*, \nu + p_1 + 2r)}{T_{p_2}(\mathbf{d}_2 | \boldsymbol{\mu}_{2.1}, \tilde{\boldsymbol{\Sigma}}_{22.1}, \nu + p_1)} E[\mathbf{Y}^{(k)}],$$

em que $\nu + p_1 + 2r > 0$, $\mathbf{d}_2 = (d_{p_1+1}, \dots, d_p)^\top$, $\delta = (\mathbf{Z} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{Z} - \boldsymbol{\mu})$, $\delta_1 = (\mathbf{Z}_1 - \boldsymbol{\mu}_1)^\top \boldsymbol{\Sigma}_{11}^{-1} (\mathbf{Z}_1 - \boldsymbol{\mu}_1)$,

$$\begin{aligned}
\mathbf{Y} &\sim Tt_{p_2}(\boldsymbol{\mu}_{2.1}, \tilde{\boldsymbol{\Sigma}}_{22.1}^*, \nu + p_1 + 2r; \mathbb{D}_2), \\
\mathbb{D}_2 &= \{(x_{p_1+1}, \dots, x_p) \in \mathbb{R}^{p_2}; x_{p_1+1} \leq d_{p_1+1}, \dots, x_p \leq d_p\}, \\
\tilde{\boldsymbol{\Sigma}}_{22.1}^* &= \frac{\nu + \delta_1}{\nu + p_1 + 2r} \boldsymbol{\Sigma}_{22.1}, \\
h_p(p_1, \nu, r) &= (\nu + p)^r \left(\frac{\Gamma((p + \nu)/2) \Gamma((p_1 + \nu + 2r)/2)}{\Gamma((p_1 + \nu)/2) \Gamma((p + \nu + 2r)/2)} \right),
\end{aligned}$$

$\boldsymbol{\mu}_{2.1}$, $\boldsymbol{\Sigma}_{22.1}$ e $\tilde{\boldsymbol{\Sigma}}_{22.1}$ são dados na Proposição 7.

2.3.1 Função de verossimilhança

Primeiro, particionamos \mathbf{Z}_i , \mathbf{V}_i , $\boldsymbol{\mu}_z$ e $\boldsymbol{\Sigma}_z$ como na subseção 2.2.1, lembrando que $\mathbf{Z}_i \sim t_p(\boldsymbol{\mu}_z, \boldsymbol{\Sigma}_z, \nu)$, ver (2.8), e utilizando a Proposição 7, temos que

$$\mathbf{Z}_i^o \sim t_{p_o}(\boldsymbol{\mu}_z^o, \boldsymbol{\Sigma}_z^{oo}, \nu) \quad \text{e} \quad \mathbf{Z}_i^c | \mathbf{Z}_i^o = \mathbf{z}_i^o \sim t_{p_c}(\boldsymbol{\mu}_z^{co}, \mathbf{S}_z^{co}, \nu + p_o), \quad (2.44)$$

em que

$$\boldsymbol{\mu}_z^{co} = \boldsymbol{\mu}_z^c + \boldsymbol{\Sigma}_z^{co} (\boldsymbol{\Sigma}_z^{oo})^{-1} (\mathbf{z}_i^o - \boldsymbol{\mu}_z^o), \quad (2.45)$$

$$\mathbf{S}_z^{co} = \left(\frac{\nu + \delta_i^1}{\nu + p_o} \right) \boldsymbol{\Sigma}_z^{cc.o}, \quad (2.46)$$

$$\boldsymbol{\Sigma}_z^{cc.o} = \boldsymbol{\Sigma}_z^{cc} - \boldsymbol{\Sigma}_z^{co} (\boldsymbol{\Sigma}_z^{oo})^{-1} \boldsymbol{\Sigma}_z^{oc} \quad (2.47)$$

$$\delta_i^1 = (\mathbf{z}_i^o - \boldsymbol{\mu}_z^o)^\top (\boldsymbol{\Sigma}_z^{oo})^{-1} (\mathbf{z}_i^o - \boldsymbol{\mu}_z^o) \quad (2.48)$$

A amostra observada para a i -ésima unidade experimental é $\{\mathbf{z}_i^o, \boldsymbol{\kappa}_i^c\}$. A verossimilhança associada é

$$L_i(\boldsymbol{\theta}) = P(\mathbf{V}_i^c = \boldsymbol{\kappa}_i^c | \mathbf{Z}_i^o = \mathbf{z}_i^o) f(\mathbf{z}_i^o),$$

em que $f(\cdot)$ é a função de densidade marginal de \mathbf{Z}_i^o . Mas $\mathbf{V}_i^c = \boldsymbol{\kappa}_i^c$, se e somente se, $\mathbf{Z}_i^c \leq \boldsymbol{\kappa}_i^c$. Por (2.44), obtemos

$$L_i(\boldsymbol{\theta}) = T_{p_c}(\boldsymbol{\kappa}_i^c | \boldsymbol{\mu}_z^{co}, \mathbf{S}_z^{co}, \nu + p_o) t_{p_o}(\mathbf{z}_i^o | \boldsymbol{\mu}_z^o, \boldsymbol{\Sigma}_z^{oo}, \nu).$$

2.3.2 Algoritmo ECM

Nesta seção, como em Matos et al. (2016), utilizamos o algoritmo ECM (Meng e Rubin, 1993) para a estimação dos parâmetros do *MEMC-t*. Este algoritmo considera uma modificação simples ao tradicional algoritmo EM inicialmente proposto por Dempster et al. (1977) e é uma ferramenta eficiente para obter as estimativas de máxima verossimilhança na estrutura de dados faltantes.

O *MEMC-t* pode ser formulado numa representação hierárquica flexível que é útil para a obtenção das derivadas. É obtida através das equações (2.3), (2.6) e (2.7) e é dado por

$$\mathbf{Z}_i|x_i, U_i = u_i \stackrel{\text{ind}}{\sim} N_p(\mathbf{a} + \mathbf{b}x_i, u_i^{-1}\mathbf{\Omega}), \quad (2.49)$$

$$x_i|U_i = u_i \stackrel{\text{ind}}{\sim} N(\mu_x, u_i^{-1}\sigma_x^2), \quad (2.50)$$

$$U_i \stackrel{\text{iid}}{\sim} \text{Gama}(\nu/2, \nu/2), \quad i = 1, \dots, n. \quad (2.51)$$

Para obter (2.49), sabemos (por (2.8)) que $\mathbf{Z}_i \sim t_p(\boldsymbol{\mu}_z, \boldsymbol{\Sigma}_z, \nu)$, em que $\boldsymbol{\mu}_z = \mathbf{a} + \mathbf{b}\mu_x$ e $\boldsymbol{\Sigma}_z = \sigma_x^2\mathbf{b}\mathbf{b}^\top + \mathbf{\Omega}$. Então \mathbf{Z}_i pode ser expresso como

$$\mathbf{Z}_i = \boldsymbol{\mu}_z + U_i^{-1/2}\mathbf{W}, \quad \text{sendo } \mathbf{W} \sim N_p(\mathbf{0}, \boldsymbol{\Sigma}_z) \text{ e } U_i \sim \text{Gama}(\nu/2, \nu/2).$$

Então, $\mathbf{Z}_i|U_i = u_i \sim N_p(\boldsymbol{\mu}_z, u_i^{-1}\boldsymbol{\Sigma}_z)$.

Portanto,

$$\begin{aligned} \mathbf{Z}_i|x_i, U_i = u_i \stackrel{\text{ind}}{\sim} N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \quad \text{sendo } \boldsymbol{\mu} = \mathbf{a} + \mathbf{b}x_i \\ \boldsymbol{\Sigma} = u_i^{-1}\boldsymbol{\Sigma}_z = u_i^{-1}(\sigma_x^2\mathbf{b}\mathbf{b}^\top + \mathbf{\Omega}) = u_i^{-1}\mathbf{\Omega} \end{aligned}$$

ou seja $\mathbf{Z}_i|x_i, U_i = u_i \stackrel{\text{ind}}{\sim} N_p(\mathbf{a} + \mathbf{b}x_i, u_i^{-1}\mathbf{\Omega})$.

Proposição 10. *Para o MEMC-t,*

$$E[U_i|\mathbf{Z}_i = \mathbf{z}_i] = \frac{p + \nu}{\delta_i + \nu},$$

em que δ_i é como em (1.4).

Prova:

Lembrando que $\mathbf{Z}_i \sim t_p(\boldsymbol{\mu}_z, \boldsymbol{\Sigma}_z, \nu)$, o que implica que $\mathbf{Z}_i|U_i = u_i \sim N_p(\boldsymbol{\mu}_z, u_i^{-1}\boldsymbol{\Sigma}_z)$ e pela equação (2.41), $U_i \sim \text{Gama}(\nu/2, \nu/2)$.

Também

$$\begin{aligned} f(u_i|\mathbf{z}_i) &= \frac{f(u_i, \mathbf{z}_i)}{f(\mathbf{z}_i)} \propto f(\mathbf{z}_i|u_i)f(u_i) \propto \frac{e^{-\frac{1}{2}(\mathbf{z}_i - \boldsymbol{\mu}_z)^\top [u_i^{-1}\boldsymbol{\Sigma}_z]^{-1}(\mathbf{z}_i - \boldsymbol{\mu}_z)}}{|u_i^{-1}\boldsymbol{\Sigma}_z|^{\frac{1}{2}}} \cdot u_i^{\frac{\nu}{2}-1} e^{-\frac{\nu}{2}u_i} \\ &= \frac{e^{-\frac{u_i}{2}[(\mathbf{z}_i - \boldsymbol{\mu}_z)^\top \boldsymbol{\Sigma}_z^{-1}(\mathbf{z}_i - \boldsymbol{\mu}_z) + \nu]}}{u_i^{-p/2}|\boldsymbol{\Sigma}_z|^{1/2}} \cdot u_i^{\frac{\nu}{2}-1} = u_i^{\frac{p+\nu}{2}-1} \cdot e^{(-\frac{\delta_i+\nu}{2})u_i} \end{aligned}$$

Então

$$U_i|\mathbf{Z}_i = \mathbf{z}_i \sim \text{Gama}\left(\frac{p+\nu}{2}, \frac{\delta_i+\nu}{2}\right).$$

Portanto,

$$E[U_i|\mathbf{Z}_i = \mathbf{z}_i] = \frac{p+\nu}{\delta_i+\nu}.$$

2.3.2.1 Passo E

Seja $\mathbf{Z} = (\mathbf{Z}_1^\top, \dots, \mathbf{Z}_n^\top)^\top$, $\mathbf{x} = (x_1, \dots, x_n)^\top$, $\mathbf{u} = (u_1, \dots, u_n)^\top$ e $\boldsymbol{\theta} = (\boldsymbol{\alpha}^\top, \boldsymbol{\beta}^\top, \mu_x, \sigma_x^2, \boldsymbol{\phi}^\top)^\top$ o vetor com todos os parâmetros no modelo. Além de constantes que não dependem de $\boldsymbol{\theta}$, a log-verossimilhança completa associada aos dados completos $\mathbf{Z}_c = \{\mathbf{Z}, \mathbf{x}, \mathbf{u}\}$ é dada pela equação (2.25) e desde que a função $Q(\cdot|\widehat{\boldsymbol{\theta}}^{(k)})$ pode ser descomposta como em (2.26), ela pode ser expressa como em (2.28) sendo

$$\begin{aligned} \widehat{u\mathbf{z}}_i^2 &= E[U_i\mathbf{Z}_i\mathbf{Z}_i^\top|\mathbf{V}_i], & \widehat{u\mathbf{z}}_i &= E[U_i\mathbf{Z}_i|\mathbf{V}_i], \\ \widehat{u}_i &= E[U_i|\mathbf{V}_i], & \widehat{ux\mathbf{z}}_i &= E[U_ix_i\mathbf{Z}_i^\top|\mathbf{V}_i], \\ \widehat{ux}_i &= E[U_ix_i|\mathbf{V}_i], & \widehat{ux}_i^2 &= E[U_ix_i^2|\mathbf{V}_i]. \end{aligned}$$

Novamente, utilizando a equação (2.29), podemos escrever

$$\begin{aligned} \widehat{u\mathbf{z}}_i^2 &= E[E[U_i\mathbf{Z}_i\mathbf{Z}_i^\top|\mathbf{Z}_i]|\mathbf{V}_i], \\ \widehat{u\mathbf{z}}_i &= E[E[U_i\mathbf{Z}_i|\mathbf{Z}_i]|\mathbf{V}_i] \quad \text{e} \\ \widehat{u}_i &= E[E[U_i|\mathbf{Z}_i]|\mathbf{V}_i]. \end{aligned} \tag{2.52}$$

As esperanças condicionais citadas acima, obtidas por Matos et al. (2016), considerando três diferentes casos, são dadas pelas seguintes expressões:

i) o i -ésimo indivíduo não tem componentes censurados.

$$\widehat{u}_i = \frac{p + \nu}{\delta_i + \nu}, \quad \widehat{u\mathbf{z}}_i = \frac{p + \nu}{\delta_i + \nu} \mathbf{z}_i \quad \text{e} \quad \widehat{u\mathbf{z}}_i^2 = \frac{p + \nu}{\delta_i + \nu} \mathbf{z}_i \mathbf{z}_i^\top,$$

em que δ_i é dado na equação (1.4).

ii) o i -ésimo indivíduo tem apenas componentes censurados.

$$\widehat{u}_i = \frac{T_p(\boldsymbol{\kappa}_i | \boldsymbol{\mu}_z, \boldsymbol{\Sigma}_z^*, \nu + 2)}{T_p(\boldsymbol{\kappa}_i | \boldsymbol{\mu}_z, \boldsymbol{\Sigma}_z, \nu)}, \quad \widehat{u\mathbf{z}}_i = \frac{T_p(\boldsymbol{\kappa}_i | \boldsymbol{\mu}_z, \boldsymbol{\Sigma}_z^*, \nu + 2)}{T_p(\boldsymbol{\kappa}_i | \boldsymbol{\mu}_z, \boldsymbol{\Sigma}_z, \nu)} E[\mathbf{Y}_i],$$

$$\widehat{u\mathbf{z}}_i^2 = \frac{T_p(\boldsymbol{\kappa}_i | \boldsymbol{\mu}_z, \boldsymbol{\Sigma}_z^*, \nu + 2)}{T_p(\boldsymbol{\kappa}_i | \boldsymbol{\mu}_z, \boldsymbol{\Sigma}_z, \nu)} E[\mathbf{Y}_i \mathbf{Y}_i^\top],$$

em que $\boldsymbol{\kappa}_i$ é o vetor com os níveis de censura para o indivíduo i , $\boldsymbol{\Sigma}_z^* = \left(\frac{\nu}{\nu + 2} \right) \boldsymbol{\Sigma}_z$, $\boldsymbol{\mu}_z$ e $\boldsymbol{\Sigma}_z$ são dados em (2.9), e $\mathbf{Y}_i \sim Tt_p(\boldsymbol{\mu}_z, \boldsymbol{\Sigma}_z^*, \nu + 2; \mathbb{D}_i)$, em que \mathbb{D}_i é como em (2.15) com $\mathbf{d} = \boldsymbol{\kappa}_i$.

iii) o i -ésimo indivíduo tem componentes censurados e não censurados.

$$\widehat{u}_i = \frac{p_o + \nu}{\nu + \delta_i^1} \frac{T_{p_c}(\boldsymbol{\kappa}_i^c | \boldsymbol{\mu}_z^{co}, \widetilde{\mathbf{S}}_z^{co}, \nu + p_o + 2)}{T_{p_c}(\boldsymbol{\kappa}_i^c | \boldsymbol{\mu}_z^{co}, \mathbf{S}_z^{co}, \nu + p_o)},$$

$$\widehat{u\mathbf{z}}_i = \text{vec}(\widehat{u}_i \mathbf{z}_i^o, \widehat{u}_i E[\mathbf{Y}_i]),$$

$$\widehat{u\mathbf{z}}_i^2 = \begin{bmatrix} \widehat{u}_i \mathbf{z}_i^o \mathbf{z}_i^{o\top} & \widehat{u}_i \mathbf{z}_i^o E[\mathbf{Y}_i]^\top \\ \widehat{u}_i E[\mathbf{Y}_i] \mathbf{z}_i^{o\top} & \widehat{u}_i E[\mathbf{Y}_i \mathbf{Y}_i^\top] \end{bmatrix},$$

em que p_o e p_c são as dimensões dos vetores \mathbf{Z}_i^o e $\boldsymbol{\kappa}_i^c$, da partição $\mathbf{V}_i = \text{vec}(\mathbf{Z}_i^o, \boldsymbol{\kappa}_i^c)$, que correspondem às observações não censuradas e aos níveis de censura, respectivamente,

$$\widetilde{\mathbf{S}}_z^{co} = \frac{\nu + \delta_i^1}{\nu + p_o + 2} \boldsymbol{\Sigma}_z^{cc.o}, \quad \mathbf{Y}_i \sim Tt_{p_c}(\boldsymbol{\mu}_z^{co}, \widetilde{\mathbf{S}}_z^{co}, \nu + p_o + 2; \mathbb{D}_i^c)$$

$\nu + p_o + 2 > 0$, $\boldsymbol{\mu}_z^{co}$, \mathbf{S}_z^{co} , $\boldsymbol{\Sigma}_z^{cc.o}$, δ_i^1 e \mathbb{D}_i^c , são dados em (2.45), (2.46), (2.47), (2.48) e (2.31), respectivamente.

Os restantes valores esperados $\widehat{u\mathbf{x}}_i$, $\widehat{u\mathbf{x}}_i^2$ e $\widehat{u\mathbf{x}\mathbf{z}}_i$ são dados nas equações (2.33), (2.36) e (2.37), respectivamente.

2.3.2.2 Passo CM

Desde que a log-verossimilhança dos dados completos da equação (2.25) é a mesma nos modelos $MEMC-Sl$ e $MEMC-t$, eles compartilham também a mesma função $Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}^{(k)})$ da equação (2.28). Portanto, as expressões fechadas no passo CM dos algoritmos ECM e MCECM são as mesmas (ver Seção 2.2.2.2). A diferença fica no cálculo dos valores esperados \widehat{u}_i , \widehat{uz}_i , \widehat{uz}_i^2 , \widehat{ux}_i , \widehat{ux}_i^2 e \widehat{uxz}_i .

2.3.3 Matriz de informação observada

Como foi mencionado antes, a log-verossimilhança dos dados completos da equação (2.25) é a mesma dos modelos $MEMC-Sl$ e $MEMC-t$. Seguindo Lin (2010) encontramos a covariância assintótica dos estimadores de máxima verossimilhança dos parâmetros do $MEMC-t$ cujos resultados são dados na Seção 2.2.3.

Capítulo 3

Análise de diagnóstico

Em estudos de modelagem estatística, uma etapa importante corresponde à validação das suposições do modelo mediante estudos de sensibilidade. A análise de diagnóstico tem o objetivo de verificar possíveis afastamentos das suposições feitas para o modelo, verificar a existência de observações extremas com interferência desproporcional no ajuste e detectar observações influentes nas estimativas do modelo.

Pontos influentes são aqueles com influência desproporcional nas estimativas dos coeficientes, isto é, quando retirados do modelo mudam de forma substancial as estimativas ou mesmo a significância dos coeficientes. O método mais conhecido para detectar tais pontos é o de deleção de pontos, que consiste em retirar um ponto e verificar as variações nas estimativas e outros resultados inferenciais. As técnicas gráficas auxiliam na busca e em detectar pontos extremos na distribuição dos dados. Por último, análise de influência busca localizar observações influentes nas estimativas do modelo, feita através dos métodos de influência global e local.

A análise de influência global, via exclusão de casos, que mede o impacto de deletar uma (ou várias) observação na estimativa dos parâmetros é diretamente avaliada por métricas como a distância de Cook (Cook, 1977). Eliminação de casos é provavelmente a técnica mais utilizada para detectar observações influentes. No entanto, pesquisas sobre a influência de pequenas perturbações no modelo ou nos dados sobre as estimativas dos parâmetros receberam atenção crescente nos últimos anos, e podem ser obtidas através da análise de influência local.

A análise de influência local, baseada em geometria diferencial, é efetuada comparando estimativas de parâmetros antes e depois de perturbar os dados ou as hipóteses do modelo (Cook, 1986). Esta área de pesquisa recebeu atenção considerável na literatura estatística em modelos de regressão linear sendo útil para verificar as suposições do modelo, assim como a identificação de dados aberrantes e/ou influentes, por meio de estudar o efeito de introduzir pequenas perturbações no modelo (ou dados) usando uma medida de influência apropriada. No entanto, para as distribuições normal independente (e outras distribuições), a função de log-verossimilhança marginal é complexa e uma aplicação direta da abordagem de Cook pode ser muito difícil, pois essas medidas envolvem a primeira e segunda derivadas parciais desta função. Inspirados pela idéia básica do algoritmo EM, Zhu e Lee (2001) propuseram um método unificado para análise de influência local em modelos estatísticos com dados faltantes, utilizando a função de afastamento da verossimilhança completa (função Q). Esta abordagem produz resultados muito semelhantes aos obtidos com o método de Cook. Além disso, a eliminação de casos pode ser estudada pela função Q seguindo a abordagem de Zhu et al. (2001) e Zhu et al. (2009). Assim, neste capítulo desenvolvemos métodos para obter medidas de eliminação de casos e medidas de influência local usando o método de Zhu et al. (2001) e Zhu e Lee (2001) (ver também Lee e Xu, 2004) no contexto de modelos de regressão com erros de medida e dados censurados. Esta abordagem foi aplicada com sucesso para realizar análises de influência em vários modelos de regressão, ver, por exemplo, Bolfarine et al. (2007), Zeller et al. (2010), Lachos et al. (2011), Matos et al. (2013), entre outros. Usando este método geral desenvolvemos uma abordagem de influência local para os modelos $MEMC-N$, $MEMC-t$ e $MEMC-SI$ e mostramos que ele leva a medidas de influência simples.

3.1 Influência global

Um dos métodos de diagnóstico utilizados em modelos de regressão utiliza a exclusão de casos que consiste em comparar as estimativas de máxima verossimilhança $\hat{\theta}$ e $\hat{\theta}_{[i]}$, em que $\hat{\theta}$ é a EMV de θ com todos os dados da amostra e $\hat{\theta}_{[i]}$ a EMV de θ com a exclusão

da i -ésima observação; se $\hat{\boldsymbol{\theta}}_{[i]}$ está longe de $\hat{\boldsymbol{\theta}}$ em algum sentido, então há evidência de que a i -ésima observação é influente. Por exemplo, LD_i utiliza o afastamento da log-verossimilhança, dado por

$$LD_i(\boldsymbol{\theta}) = 2[\ell(\hat{\boldsymbol{\theta}}) - \ell(\hat{\boldsymbol{\theta}}_{[i]})], \quad i = 1, \dots, n \quad (3.1)$$

sendo $\ell(\boldsymbol{\theta})$ a função de log-verossimilhança.

Esses métodos são utilizados com frequência para diagnosticar globalmente possíveis observações influentes e podem ser facilmente adaptados aos modelos de regressão com erros de medida e dados censurados, dado que podemos calcular os estimadores de máxima verossimilhança com e sem a i -ésima observação.

Neste trabalho utilizaremos duas medidas de diagnóstico de influência global, uma chamada de afastamento da função Q ou Q -afastamento QD_i (Zhu et al., 2001), em que

$$QD_i = 2[Q(\hat{\boldsymbol{\theta}}|\hat{\boldsymbol{\theta}}) - Q(\hat{\boldsymbol{\theta}}_{[i]}|\hat{\boldsymbol{\theta}})], \quad (3.2)$$

que está relacionado à mudança em valores da função Q , similar ao LD_i definido em (3.1), e outra chamada de distância de Cook generalizada, definida como

$$D_i = (\hat{\boldsymbol{\theta}}_{[i]} - \hat{\boldsymbol{\theta}})^\top [-\ddot{Q}(\hat{\boldsymbol{\theta}}|\hat{\boldsymbol{\theta}})](\hat{\boldsymbol{\theta}}_{[i]} - \hat{\boldsymbol{\theta}}), \quad i = 1, \dots, n \quad (3.3)$$

em que $\ddot{Q}(\hat{\boldsymbol{\theta}}|\hat{\boldsymbol{\theta}})$ é a matriz hessiana avaliada em $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$.

Desde que $\hat{\boldsymbol{\theta}}_{[i]}$ é necessário para cada caso, a estimação do modelo deve ser feita $n + 1$ vezes (em uma amostra de tamanho n) o que pode demandar um alto custo computacional. Para evitar isso, utilizamos a aproximação de um passo $\hat{\boldsymbol{\theta}}_{[i]}^1$ de $\hat{\boldsymbol{\theta}}_{[i]}$ (Zhu et al., 2001) definida por

$$\hat{\boldsymbol{\theta}}_{[i]}^1 = \hat{\boldsymbol{\theta}} + [-\ddot{Q}(\hat{\boldsymbol{\theta}}|\hat{\boldsymbol{\theta}})]^{-1} \dot{Q}_{[i]}(\hat{\boldsymbol{\theta}}|\hat{\boldsymbol{\theta}}), \quad i = 1, \dots, n \quad (3.4)$$

em que $\dot{Q}_{[i]}(\hat{\boldsymbol{\theta}}|\hat{\boldsymbol{\theta}}) = \left. \frac{\partial Q_{[i]}(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}} \right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}$ representa o vetor escore individual.

Ao substituir (3.4) em (3.2) e (3.3), obtemos as seguintes aproximações QD_i^1 e D_i^1 de QD_i e D_i , respectivamente:

$$QD_i^1 = 2[Q(\hat{\boldsymbol{\theta}}|\hat{\boldsymbol{\theta}}) - Q(\hat{\boldsymbol{\theta}}_{[i]}^1|\hat{\boldsymbol{\theta}})],$$

$$D_i^1 = \dot{Q}_{[i]}(\hat{\boldsymbol{\theta}}|\hat{\boldsymbol{\theta}})^\top [-\ddot{Q}(\hat{\boldsymbol{\theta}}|\hat{\boldsymbol{\theta}})]^{-1} \dot{Q}_{[i]}(\hat{\boldsymbol{\theta}}|\hat{\boldsymbol{\theta}}).$$

3.2 Influência local

A metodologia de influência local é útil para verificar as suposições do modelo, assim como a identificação de dados aberrantes e/ou influentes, por meio de estudar o efeito de introduzir pequenas perturbações no modelo (ou dados) usando uma medida de influência apropriada.

Para derivar medidas de influência local sob alguns esquemas de perturbação, utilizamos a metodologia de Zhu e Lee (2001), que inspirados pela ideia básica do algoritmo EM, propuseram uma abordagem para realizar diagnósticos de influência em modelos estatísticos com dados incompletos baseados na função Q -afastamento, representado por $f_Q(\boldsymbol{\omega})$.

Considere um vetor de perturbações $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n)^\top$ variando em uma região aberta $\boldsymbol{\Omega} \subset \mathbb{R}^n$. Seja $\ell_c(\boldsymbol{\theta}, \boldsymbol{\omega} | \mathbf{y}, \mathbf{u})$ a função de log-verossimilhança completa do modelo perturbado. Assume-se que existe um $\boldsymbol{\omega}_0 \in \boldsymbol{\Omega}$ tal que $\ell_c(\boldsymbol{\theta}, \boldsymbol{\omega}_0 | \mathbf{y}, \mathbf{u}) = \ell_c(\boldsymbol{\theta} | \mathbf{y}, \mathbf{u})$, para todo $\boldsymbol{\theta}$. Seja $\hat{\boldsymbol{\theta}}(\boldsymbol{\omega})$ a EMV de $\boldsymbol{\theta}$ para o modelo perturbado que maximiza a função $Q(\boldsymbol{\theta}, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}}) = E[\ell_c(\boldsymbol{\theta}, \boldsymbol{\omega} | \mathbf{y}, \mathbf{u}) | \mathbf{y}, \hat{\boldsymbol{\theta}}]$. Zhu e Lee (2001) propuseram a função Q -afastamento

$$f_Q(\boldsymbol{\omega}) = 2[Q(\hat{\boldsymbol{\theta}} | \hat{\boldsymbol{\theta}}) - Q(\hat{\boldsymbol{\theta}}(\boldsymbol{\omega}) | \hat{\boldsymbol{\theta}})],$$

definindo o gráfico de influência de $f_Q(\boldsymbol{\omega})$ como $\boldsymbol{\alpha}(\boldsymbol{\omega}) = (\boldsymbol{\omega}^\top, f_Q(\boldsymbol{\omega}))^\top$. Segundo Zhu e Lee (2001), a curvatura normal $C_{f_Q, \mathbf{h}}$ de $\boldsymbol{\alpha}(\boldsymbol{\omega})$ em $\boldsymbol{\omega}_0$ na direção de um vetor unitário $\mathbf{h} \in \mathbb{R}^n$, pode ser usada para resumir o comportamento local da função Q -afastamento.

Seguindo o procedimento adotado por Cook (1986), pode ser mostrado (veja Zhu e Lee, 2001) que a curvatura normal $C_{f_Q, \mathbf{h}}$ de $\boldsymbol{\alpha}(\boldsymbol{\omega})$ em $\boldsymbol{\omega}_0$ é

$$C_{f_Q, \mathbf{h}} = -2\mathbf{h}^\top \ddot{\mathbf{Q}}_{\boldsymbol{\omega}_0} \mathbf{h} = 2\mathbf{h}^\top \boldsymbol{\Delta}_{\boldsymbol{\omega}_0}^\top [-\ddot{\mathbf{Q}}_{\boldsymbol{\theta}}(\hat{\boldsymbol{\theta}} | \hat{\boldsymbol{\theta}})]^{-1} \boldsymbol{\Delta}_{\boldsymbol{\omega}_0} \mathbf{h},$$

em que $\ddot{\mathbf{Q}}_{\boldsymbol{\omega}_0} = \frac{\partial^2 Q(\hat{\boldsymbol{\theta}}(\boldsymbol{\omega}) | \hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\omega} \partial \boldsymbol{\omega}^\top} \Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0}$, $\ddot{\mathbf{Q}}_{\boldsymbol{\theta}}(\hat{\boldsymbol{\theta}} | \hat{\boldsymbol{\theta}}) = \frac{\partial^2 Q(\boldsymbol{\theta} | \hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}$ e $\boldsymbol{\Delta}_{\boldsymbol{\omega}} = \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\omega}^\top} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}(\boldsymbol{\omega})}$.

Analogamente a Cook (1986), a expressão $-\ddot{\mathbf{Q}}_{\boldsymbol{\omega}_0}$, ou equivalentemente $\boldsymbol{\Delta}_{\boldsymbol{\omega}_0}$ e $-\ddot{\mathbf{Q}}_{\boldsymbol{\theta}}(\hat{\boldsymbol{\theta}} | \hat{\boldsymbol{\theta}})$, é a matriz fundamental para detectar observações influentes. Mas utilizar a curvatura normal em sua forma original para avaliar a influência de uma determinada observação pode gerar alguns problemas, uma vez que $C_{f_Q, \mathbf{h}}$ pode assumir qualquer valor na reta e

não é invariante a mudanças de escala uniformes (veja Poon e Poon (1999) para discussão e exemplos). Desta forma, com base no trabalho de Poon e Poon (1999) e de Zhu e Lee (2001), a curvatura normal conforme $B_{f_Q, \mathbf{h}}$ em $\boldsymbol{\omega}_0$ na direção de um vetor unitário \mathbf{h} é definida por

$$B_{f_Q, \mathbf{h}} = \frac{-2\mathbf{h}^\top \ddot{\mathbf{Q}}_{\boldsymbol{\omega}_0} \mathbf{h}}{\text{tr}(-2\ddot{\mathbf{Q}}_{\boldsymbol{\omega}_0})}.$$

Seja $B = \frac{-2\ddot{\mathbf{Q}}_{\boldsymbol{\omega}_0}}{\text{tr}(-2\ddot{\mathbf{Q}}_{\boldsymbol{\omega}_0})}$, e $\lambda_1 \geq \dots \geq \lambda_r > 0$ os r autovalores de B diferentes de zero, e $\mathbf{e}_1, \dots, \mathbf{e}_r$ os autovetores ortogonais correspondentes. Segundo Lesaffre e Verbeke (1998), Poon e Poon (1999), e Zhu e Lee (2001), o seguinte vetor de contribuição agregado de todos os autovetores que estão associados com todos os autovalores diferentes de zero

$$M(0) = \sum_{i=1}^r \lambda_i \mathbf{e}_i^2$$

em que $\mathbf{e}_i^2 = (e_{i1}^2, \dots, e_{in}^2)^\top$, é usado para avaliar a influência local. Para $j = 1, \dots, n$, segue de Zhu e Lee (2001) que a j -ésima componente de $M(0)$, $M(0)_j = b_{jj}$ para $j = 1, \dots, n$, onde b_{jj} é o j -ésimo elemento diagonal da matriz B . Portanto, é muito simples calcular b_{jj} e reduzir em grande parte a carga computacional, porque não há autovetores e autovalores envolvidos. Portanto, nossas medidas de influência locais baseiam-se na curvatura normal conforme em vez da curvatura normal clássica, porque a curvatura normal conforme possui as propriedades mencionadas acima, assim como também que $0 \leq B_{f_Q, \mathbf{h}} \leq 1$.

Assim, a avaliação de casos influentes é baseada na inspeção visual do gráfico de $\{M(0)_j, j = 1, \dots, n\}$ plotado contra o índice j . O j -ésimo caso é então considerado influente se $M(0)_j$ é maior do que um ponto de referência adequado. Na análise de influência local, até agora não existem regras gerais para selecionar o valor de referência. No entanto, considerando $\overline{M(0)}$ e $DP[M(0)]$ como a média e o desvio padrão de $\{M(0)_j, j = 1, \dots, n\}$, respectivamente, Poon e Poon (1999) propuseram utilizar $2\overline{M(0)}$ como ponto de referência, Zhu e Lee (2001) propuseram $\overline{M(0)} + 2DP[M(0)]$, enquanto que Lee e Xu (2004) propuseram $\overline{M(0)} + c^*DP[M(0)]$ em que c^* é uma constante arbitrária apropriada maior ou igual que 2. Bolfarine et al. (2007) utilizaram $c^* = 2$, Lachos et al. (2011) utilizaram $c^* = 4$, Massuia et al. (2015) utilizaram $c^* = 3, 5$. Neste trabalho utilizaremos $c^* = 3$.

3.2.1 Matriz hessiana

A fim de obter as medidas de diagnóstico para a influência local, é necessário calcular a matriz hessiana, que, em geral, é expressa como

$$\ddot{Q}(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}}) = \frac{\partial^2 Q(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} = \begin{bmatrix} \ddot{Q}_{\alpha\alpha} & \ddot{Q}_{\alpha\beta} & \mathbf{0} & \mathbf{0} & \ddot{Q}_{\alpha\phi} \\ \ddot{Q}_{\beta\alpha} & \ddot{Q}_{\beta\beta} & \mathbf{0} & \mathbf{0} & \ddot{Q}_{\beta\phi} \\ \mathbf{0} & \mathbf{0} & \ddot{Q}_{\mu_x\mu_x} & \ddot{Q}_{\mu_x\sigma_x^2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \ddot{Q}_{\sigma_x^2\mu_x} & \ddot{Q}_{\sigma_x^2\sigma_x^2} & \mathbf{0} \\ \ddot{Q}_{\phi\alpha} & \ddot{Q}_{\phi\beta} & \mathbf{0} & \mathbf{0} & \ddot{Q}_{\phi\phi} \end{bmatrix}$$

em que $\ddot{Q}_{\tau\lambda} = \frac{\partial^2 Q(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}})}{\partial \tau \partial \lambda^\top}$, com $\tau, \lambda = \boldsymbol{\alpha}, \boldsymbol{\beta}, \mu_x, \sigma_x^2, \boldsymbol{\phi}$, é uma matriz de dimensão $(3p \times 3p)$.

3.2.2 Esquemas de perturbação

3.2.2.1 Perturbação de ponderação de casos

A ponderação de casos tem sido o esquema de perturbação mais amplamente difundido na análise de diagnóstico. Este esquema de perturbação permite avaliar a contribuição individual de cada observação sobre o processo de estimação. Seja $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n)^\top$ um vetor $n \times 1$ de ponderações. Sendo $\omega_i = 0$, temos que a i -ésima observação é eliminada e $\boldsymbol{\omega}_0 = (1, \dots, 1)^\top$ implica que todas as observações são consideradas.

Utilizando (2.26) temos que

$$Q(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}}) = \sum_{i=1}^n Q_i(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}}) = \sum_{i=1}^n \left[Q_{1i}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\phi}|\hat{\boldsymbol{\theta}}) + Q_{2i}(\mu_x, \sigma_x^2|\hat{\boldsymbol{\theta}}) \right].$$

Como vemos, a função $Q(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}})$ se decompõe em duas parcelas: uma que depende apenas de $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ e $\boldsymbol{\phi}$ e outra que depende apenas de μ_x e σ_x^2 .

A função $Q(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}})$ do modelo perturbado, considerando ponderação de casos, é dada por

$$Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\hat{\boldsymbol{\theta}}) = \sum_{i=1}^n \left\{ \omega_i Q_i(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}}) \right\} = \sum_{i=1}^n \left\{ \omega_i Q_{1i}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\phi}|\hat{\boldsymbol{\theta}}) \right\} + \sum_{i=1}^n \left\{ \omega_i Q_{2i}(\mu_x, \sigma_x^2|\hat{\boldsymbol{\theta}}) \right\} \quad (3.5)$$

em que $0 \leq \omega_i \leq 1$ e $\boldsymbol{\theta} = (\boldsymbol{\alpha}^\top, \boldsymbol{\beta}^\top, \mu_x, \sigma_x^2, \boldsymbol{\phi}^\top)^\top$.

Derivando (3.5) em relação a $\boldsymbol{\omega}^\top$,

$$\begin{aligned}\frac{\partial Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\omega}^\top} &= \frac{\partial \sum_{i=1}^n \{\omega_i Q_i(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}})\}}{\partial \boldsymbol{\omega}^\top} = \left(\frac{\partial}{\partial \omega_1} \sum_{i=1}^n \{\omega_i Q_i(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}})\}, \dots, \frac{\partial}{\partial \omega_n} \sum_{i=1}^n \{\omega_i Q_i(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}})\} \right) \\ &= \left(Q_1(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}}), \dots, Q_n(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}}) \right),\end{aligned}$$

logo, derivando em relação a $\boldsymbol{\theta}$ a expressão anterior, temos que

$$\boldsymbol{\Delta} = \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\omega}^\top} = \frac{\partial}{\partial \boldsymbol{\theta}} \left[\frac{\partial Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\omega}^\top} \right] = \left(\frac{\partial Q_1(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}}, \dots, \frac{\partial Q_n(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}} \right).$$

Observe que, para o esquema de perturbação de ponderação de casos, a matriz $\boldsymbol{\Delta}$ não depende do vetor $\boldsymbol{\omega}$.

Em geral, nos modelos aqui considerados temos que $\boldsymbol{\theta} = (\boldsymbol{\alpha}^\top, \boldsymbol{\beta}^\top, \mu_x, \sigma_x^2, \boldsymbol{\phi}^\top)^\top$, em que $\boldsymbol{\alpha}$ e $\boldsymbol{\beta}$ são de dimensão $(r \times 1)$, $\boldsymbol{\phi}$ é de dimensão $(p \times 1)$ e μ_x e σ_x^2 são escalares; então $\boldsymbol{\Delta}$ é uma matriz de dimensão $(3p \times n)$ dada por

$$\boldsymbol{\Delta} = \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\omega}^\top} = \begin{bmatrix} \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\hat{\boldsymbol{\theta}})}{\partial \alpha_1 \partial \omega_1} & \cdots & \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\hat{\boldsymbol{\theta}})}{\partial \alpha_1 \partial \omega_n} \\ \vdots & & \vdots \\ \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\hat{\boldsymbol{\theta}})}{\partial \alpha_r \partial \omega_1} & \cdots & \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\hat{\boldsymbol{\theta}})}{\partial \alpha_r \partial \omega_n} \\ \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\hat{\boldsymbol{\theta}})}{\partial \beta_1 \partial \omega_1} & \cdots & \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\hat{\boldsymbol{\theta}})}{\partial \beta_1 \partial \omega_n} \\ \vdots & & \vdots \\ \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\hat{\boldsymbol{\theta}})}{\partial \beta_r \partial \omega_1} & \cdots & \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\hat{\boldsymbol{\theta}})}{\partial \beta_r \partial \omega_n} \\ \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\hat{\boldsymbol{\theta}})}{\partial \mu_x \partial \omega_1} & \cdots & \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\hat{\boldsymbol{\theta}})}{\partial \mu_x \partial \omega_n} \\ \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\hat{\boldsymbol{\theta}})}{\partial \sigma_x^2 \partial \omega_1} & \cdots & \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\hat{\boldsymbol{\theta}})}{\partial \sigma_x^2 \partial \omega_n} \\ \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\hat{\boldsymbol{\theta}})}{\partial \phi_1 \partial \omega_1} & \cdots & \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\hat{\boldsymbol{\theta}})}{\partial \phi_1 \partial \omega_n} \\ \vdots & & \vdots \\ \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\hat{\boldsymbol{\theta}})}{\partial \phi_p \partial \omega_1} & \cdots & \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\hat{\boldsymbol{\theta}})}{\partial \phi_p \partial \omega_n} \end{bmatrix}$$

avaliada em $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\alpha}}^\top, \hat{\boldsymbol{\beta}}^\top, \hat{\mu}_x, \hat{\sigma}_x^2, \hat{\boldsymbol{\phi}}^\top)^\top$ e $\boldsymbol{\omega} = \boldsymbol{\omega}_0 = (1, \dots, 1)^\top$.

Assim, é possível dividir Δ na forma

$$\Delta = \begin{bmatrix} \Delta_\alpha \\ \Delta_\beta \\ \Delta_{\mu_x} \\ \Delta_{\sigma_x^2} \\ \Delta_\phi \end{bmatrix}$$

em que

$$\begin{aligned} \Delta_\alpha &= \left. \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})}{\partial \alpha \partial \boldsymbol{\omega}^\top} \right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\omega}=\boldsymbol{\omega}_0}, \\ \Delta_\beta &= \left. \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})}{\partial \beta \partial \boldsymbol{\omega}^\top} \right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\omega}=\boldsymbol{\omega}_0}, \\ \Delta_{\mu_x} &= \left. \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})}{\partial \mu_x \partial \boldsymbol{\omega}^\top} \right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\omega}=\boldsymbol{\omega}_0}, \\ \Delta_{\sigma_x^2} &= \left. \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})}{\partial \sigma_x^2 \partial \boldsymbol{\omega}^\top} \right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\omega}=\boldsymbol{\omega}_0}, \\ \Delta_\phi &= \left. \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})}{\partial \phi \partial \boldsymbol{\omega}^\top} \right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\omega}=\boldsymbol{\omega}_0}. \end{aligned}$$

3.2.2.2 Perturbação na covariável

Este esquema de perturbação pode ser utilizado se o objetivo for avaliar a sensibilidade das estimativas quando são introduzidas pequenas perturbações na covariável. Neste trabalho consideramos a perturbação aditiva $x_\omega = x + \omega$, pelo que $\boldsymbol{\omega}_0 = (0, \dots, 0)^\top$.

3.2.3 Caso modelo com erros de medida estrutural e dados censurados baseado na distribuição normal (*MEMC-N*)

No caso do *MEMC-N* (veja equação (2.8)), temos que $\mathbf{Z}_i \sim N_p(\boldsymbol{\mu}_z, \boldsymbol{\Sigma}_z)$ e a função de log-verossimilhança é dada por

$$\begin{aligned} \ell_c(\boldsymbol{\theta} | \mathbf{Z}_c) &= -\frac{n}{2} \sum_{j=1}^p \log \phi_j^2 - \frac{1}{2} \sum_{i=1}^n (\mathbf{Z}_i - \mathbf{a} - \mathbf{b}x_i)^\top \boldsymbol{\Omega}^{-1} (\mathbf{Z}_i - \mathbf{a} - \mathbf{b}x_i) - \frac{n}{2} \log \sigma_x^2 \\ &\quad - \frac{1}{2\sigma_x^2} \sum_{i=1}^n (x_i - \mu_x)^2 \end{aligned}$$

e desde que $Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}^{(k)}) = E_{\widehat{\boldsymbol{\theta}}^{(k)}}[\ell_c(\boldsymbol{\theta}|\mathbf{Z}_c)|\mathbf{V}]$, sendo

$$Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}^{(k)}) = Q_1(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\phi}|\widehat{\boldsymbol{\theta}}^{(k)}) + Q_2(\mu_x, \sigma_x^2|\widehat{\boldsymbol{\theta}}^{(k)})$$

temos que

$$Q_1(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\phi}|\widehat{\boldsymbol{\theta}}^{(k)}) = E_{\widehat{\boldsymbol{\theta}}^{(k)}} \left[-\frac{n}{2} \sum_{j=1}^p \log \phi_j^2 - \frac{1}{2} \sum_{i=1}^n (\mathbf{Z}_i - \mathbf{a} - \mathbf{b}x_i)^\top \boldsymbol{\Omega}^{-1} (\mathbf{Z}_i - \mathbf{a} - \mathbf{b}x_i) \middle| \mathbf{V} \right],$$

$$Q_2(\mu_x, \sigma_x^2|\widehat{\boldsymbol{\theta}}^{(k)}) = E_{\widehat{\boldsymbol{\theta}}^{(k)}} \left[-\frac{n}{2} \log \sigma_x^2 - \frac{1}{2\sigma_x^2} \sum_{i=1}^n (x_i - \mu_x)^2 \middle| \mathbf{V} \right].$$

Expandindo as expressões $Q_1(\cdot|\widehat{\boldsymbol{\theta}}^{(k)})$ e $Q_2(\cdot|\widehat{\boldsymbol{\theta}}^{(k)})$ e tomando esperanças, segue que

$$Q_1(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\phi}|\widehat{\boldsymbol{\theta}}^{(k)}) = -\frac{n}{2} \sum_{j=1}^p \log \phi_j^2 - \frac{1}{2} \sum_{i=1}^n \left[\text{tr} \left(\boldsymbol{\Omega}^{-1} \widehat{\mathbf{z}}_i^2 \right) - 2\mathbf{a}^\top \boldsymbol{\Omega}^{-1} \widehat{\mathbf{z}}_i - 2\widehat{x}_i \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{a} \right. \\ \left. + \mathbf{a}^\top \boldsymbol{\Omega}^{-1} \mathbf{a} + 2\mathbf{a}^\top \boldsymbol{\Omega}^{-1} \mathbf{b} \widehat{x}_i + \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b} \widehat{x}_i^2 \right],$$

$$Q_2(\mu_x, \sigma_x^2|\widehat{\boldsymbol{\theta}}^{(k)}) = -\frac{n}{2} \log \sigma_x^2 - \frac{1}{2\sigma_x^2} \sum_{i=1}^n \left(\widehat{x}_i^2 - 2\mu_x \widehat{x}_i + \mu_x^2 \right)$$

em que

$$\widehat{\mathbf{z}}_i^2 = E[\mathbf{Z}_i \mathbf{Z}_i^\top | \mathbf{V}_i], \quad \widehat{\mathbf{z}}_i = E[\mathbf{Z}_i | \mathbf{V}_i], \quad \widehat{x} \widehat{\mathbf{z}}_i = E[x_i \mathbf{Z}_i^\top | \mathbf{V}_i], \quad \widehat{x}_i = E[x_i | \mathbf{V}_i], \quad \text{e} \quad \widehat{x}_i^2 = E[x_i^2 | \mathbf{V}_i]$$

as quais tem as seguintes formas:

i) O i -ésimo indivíduo não tem componentes censurados. Aqui, $\mathbf{V}_i = \mathbf{Z}_i$, então

$$\widehat{\mathbf{z}}_i = E[\mathbf{Z}_i | \mathbf{V}_i] = E[\mathbf{Z}_i | \mathbf{Z}_i] = \mathbf{z}_i, \quad \text{e} \\ \widehat{\mathbf{z}}_i^2 = E[\mathbf{Z}_i \mathbf{Z}_i^\top | \mathbf{V}_i] = E[\mathbf{Z}_i \mathbf{Z}_i^\top | \mathbf{Z}_i] = \mathbf{z}_i \mathbf{z}_i^\top.$$

ii) O i -ésimo indivíduo tem apenas componentes censurados. Neste caso, $\mathbf{Z}_i \leq \boldsymbol{\kappa}_i$.

Assim, pela definição de uma distribuição normal truncada, temos que $\mathbf{Z}_i | (\mathbf{Z}_i \leq \boldsymbol{\kappa}_i) \sim NT_p(\boldsymbol{\mu}_z, \boldsymbol{\Sigma}_z; \mathbb{D}_i)$, em que \mathbb{D}_i é como em (2.15) com $\mathbf{d} = \boldsymbol{\kappa}_i$.

$$\widehat{\mathbf{z}}_i = E[\mathbf{Z}_i | \mathbf{Z}_i \leq \boldsymbol{\kappa}_i] = E[\mathbf{Y}_i], \quad \mathbf{Y}_i \sim NT_p(\boldsymbol{\mu}_z, \boldsymbol{\Sigma}_z; \mathbb{D}_i). \\ \widehat{\mathbf{z}}_i^2 = E[\mathbf{Z}_i \mathbf{Z}_i^\top | \mathbf{Z}_i \leq \boldsymbol{\kappa}_i] = E[\mathbf{Y}_i \mathbf{Y}_i^\top], \quad \mathbf{Y}_i \sim NT_p(\boldsymbol{\mu}_z, \boldsymbol{\Sigma}_z; \mathbb{D}_i).$$

iii) O i -ésimo indivíduo tem componentes censurados e não censurados. Particionamos o vetor $\mathbf{Z}_i = \text{vec}(\mathbf{Z}_i^o, \mathbf{Z}_i^c)$. Os componentes são censurados se e somente se $\mathbf{Z}_i^c \leq \boldsymbol{\kappa}_i^c$.

$$\begin{aligned}\widehat{\mathbf{z}}_i &= E[\mathbf{Z}_i | \mathbf{V}_i] = E \left[\text{vec}(\mathbf{Z}_i^o, \mathbf{Z}_i^c) \middle| \mathbf{Z}_i^o = \mathbf{z}_i^o, \mathbf{Z}_i^c \leq \boldsymbol{\kappa}_i^c \right] \\ &= \text{vec} \left(E \left[\mathbf{Z}_i^o \middle| \mathbf{Z}_i^o = \mathbf{z}_i^o, \mathbf{Z}_i^c \leq \boldsymbol{\kappa}_i^c \right], E \left[\mathbf{Z}_i^c \middle| \mathbf{Z}_i^o = \mathbf{z}_i^o, \mathbf{Z}_i^c \leq \boldsymbol{\kappa}_i^c \right] \right) = \text{vec}(\mathbf{z}_i^o, E[\mathbf{Y}_i])\end{aligned}$$

em que

$$\mathbf{Y}_i \sim NT_{p_c}(\boldsymbol{\mu}_z^{co}, \mathbf{S}_z^{co}; \mathbb{D}_i^c). \quad (3.6)$$

$$\begin{aligned}\widehat{\mathbf{z}}_i^2 &= E[\mathbf{Z}_i \mathbf{Z}_i^\top | \mathbf{V}_i] = E \left[\begin{bmatrix} \mathbf{Z}_i^o \mathbf{Z}_i^{o\top} & \mathbf{Z}_i^o \mathbf{Z}_i^{c\top} \\ \mathbf{Z}_i^c \mathbf{Z}_i^{o\top} & \mathbf{Z}_i^c \mathbf{Z}_i^{c\top} \end{bmatrix} \middle| \mathbf{Z}_i^o = \mathbf{z}_i^o, \mathbf{Z}_i^c \leq \boldsymbol{\kappa}_i^c \right] \\ &= \begin{bmatrix} \mathbf{z}_i^o \mathbf{z}_i^{o\top} & \mathbf{z}_i^o E[\mathbf{Y}_i]^\top \\ E[\mathbf{Y}_i] \mathbf{z}_i^{o\top} & E[\mathbf{Y}_i \mathbf{Y}_i^\top] \end{bmatrix},\end{aligned}$$

em que \mathbf{Y}_i é como em (3.6) e

$$\begin{aligned}\mathbf{Z}_i &\sim N_p(\boldsymbol{\mu}_z, \boldsymbol{\Sigma}_z), \quad \boldsymbol{\mu}_z^{co} = \boldsymbol{\mu}_z^c + \boldsymbol{\Sigma}_z^{co} \boldsymbol{\Sigma}_z^{oo-1} (\mathbf{Z}_i^o - \boldsymbol{\mu}_z^o), \quad \mathbf{S}_z^{co} = \boldsymbol{\Sigma}_z^{cc} - \boldsymbol{\Sigma}_z^{co} \boldsymbol{\Sigma}_z^{oo-1} \boldsymbol{\Sigma}_z^{oc}, \\ \mathbf{Z}_i &= \begin{bmatrix} \mathbf{Z}_i^o \\ \mathbf{Z}_i^c \end{bmatrix}, \quad \boldsymbol{\mu}_z = \begin{bmatrix} \boldsymbol{\mu}_z^o \\ \boldsymbol{\mu}_z^c \end{bmatrix} \quad \text{e} \quad \boldsymbol{\Sigma}_z = \begin{bmatrix} \boldsymbol{\Sigma}_z^{oo} & \boldsymbol{\Sigma}_z^{oc} \\ \boldsymbol{\Sigma}_z^{co} & \boldsymbol{\Sigma}_z^{cc} \end{bmatrix}.\end{aligned}$$

Finalmente,

$$\begin{aligned}\widehat{x}_i &= E[x_i | \mathbf{V}_i] = E[E[x_i | \mathbf{Z}_i] | \mathbf{V}_i] = E \left[\frac{\mu_x + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} (\mathbf{Z}_i - \mathbf{a})}{1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b}} \middle| \mathbf{V}_i \right] \\ &= \frac{\mu_x + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} E[\mathbf{Z}_i | \mathbf{V}_i] - \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{a}}{1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b}} = \mu_x + \frac{\sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} (\widehat{\mathbf{z}}_i - \boldsymbol{\mu}_z)}{1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b}} = \mu_x + \boldsymbol{\varphi}(\widehat{\mathbf{z}}_i - \boldsymbol{\mu}_z),\end{aligned}$$

$$\begin{aligned}\widehat{x}_i \widehat{\mathbf{z}}_i &= E[x_i \mathbf{Z}_i^\top | \mathbf{V}_i] = E[E[x_i \mathbf{Z}_i^\top | \mathbf{Z}_i] | \mathbf{V}_i] = E \left[\frac{\mu_x + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} (\mathbf{Z}_i - \mathbf{a})}{1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b}} \mathbf{Z}_i^\top \middle| \mathbf{V}_i \right] \\ &= \frac{\mu_x \widehat{\mathbf{z}}_i^\top + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \widehat{\mathbf{z}}_i^2 - \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \boldsymbol{\mu}_z \widehat{\mathbf{z}}_i^\top + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b} \mu_x \widehat{\mathbf{z}}_i^\top}{1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b}} = \mu_x \widehat{\mathbf{z}}_i^\top + \boldsymbol{\varphi}(\widehat{\mathbf{z}}_i^2 - \boldsymbol{\mu}_z \widehat{\mathbf{z}}_i^\top),\end{aligned}$$

$$\begin{aligned}\widehat{x}_i^2 &= E[x_i^2 | \mathbf{V}_i] = E[E[x_i^2 | \mathbf{Z}_i] | \mathbf{V}_i] = E \left[\frac{\sigma_x^2}{1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b}} + \left(\frac{\mu_x + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} (\mathbf{z}_i - \mathbf{a})}{1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b}} \right)^2 \middle| \mathbf{V}_i \right] \\ &= \Lambda + \mu_x^2 + 2\mu_x \boldsymbol{\varphi}(\widehat{\mathbf{z}}_i - \boldsymbol{\mu}_z) + \boldsymbol{\varphi}(\widehat{\mathbf{z}}_i^2 - \widehat{\mathbf{z}}_i \boldsymbol{\mu}_z^\top - \boldsymbol{\mu}_z \widehat{\mathbf{z}}_i^\top + \boldsymbol{\mu}_z \boldsymbol{\mu}_z^\top) \boldsymbol{\varphi}^\top,\end{aligned}$$

em que $\boldsymbol{\varphi}$ e Λ são como em (2.34) e (2.38), respectivamente.

3.2.3.1 Matriz hessiana

$$\begin{aligned}
\ddot{Q}_{\alpha\alpha} &= \frac{\partial}{\partial \boldsymbol{\alpha}} \left[\sum_{i=1}^n \left(\widehat{\mathbf{z}}_i^{*\top} - \boldsymbol{\alpha}^\top - \widehat{x}_i \boldsymbol{\beta}^\top \right) \boldsymbol{\Omega}_{22}^{-1} \right] = - \sum_{i=1}^n \boldsymbol{\Omega}_{22}^{-1} = -n \boldsymbol{\Omega}_{22}^{-1}, \\
\ddot{Q}_{\alpha\beta} &= \frac{\partial}{\partial \boldsymbol{\alpha}} \left[\sum_{i=1}^n \left(\widehat{x} \widehat{\mathbf{z}}_i^* - \widehat{x}_i \boldsymbol{\alpha}^\top - \widehat{x}_i^2 \boldsymbol{\beta}^\top \right) \boldsymbol{\Omega}_{22}^{-1} \right] = - \sum_{i=1}^n \widehat{x}_i \boldsymbol{\Omega}_{22}^{-1} = \ddot{Q}_{\beta\alpha}^\top, \\
\ddot{Q}_{\alpha\phi} &= \ddot{Q}_{\phi\alpha}^\top, \\
\ddot{Q}_{\beta\beta} &= \frac{\partial}{\partial \boldsymbol{\beta}} \left[\sum_{i=1}^n \left(\widehat{x} \widehat{\mathbf{z}}_i^* - \widehat{x}_i \boldsymbol{\alpha}^\top - \widehat{x}_i^2 \boldsymbol{\beta}^\top \right) \boldsymbol{\Omega}_{22}^{-1} \right] = - \sum_{i=1}^n \widehat{x}_i^2 \boldsymbol{\Omega}_{22}^{-1}, \\
\ddot{Q}_{\beta\phi} &= \ddot{Q}_{\phi\beta}^\top, \\
\ddot{Q}_{\mu_x \mu_x} &= \frac{\partial}{\partial \mu_x} \left[\frac{1}{\sigma_x^2} \sum_{i=1}^n (\widehat{x}_i - \mu_x) \right] = -\frac{n}{\sigma_x^2}, \\
\ddot{Q}_{\mu_x \sigma_x^2} &= \frac{\partial}{\partial \mu_x} \left[-\frac{n}{2\sigma_x^2} + \frac{1}{2\sigma_x^4} \sum_{i=1}^n (\widehat{x}_i^2 - 2\mu_x \widehat{x}_i + \mu_x^2) \right] = -\frac{1}{\sigma_x^4} \sum_{i=1}^n (\widehat{x}_i - \mu_x) = \ddot{Q}_{\sigma_x^2 \mu_x}, \\
\ddot{Q}_{\sigma_x^2 \sigma_x^2} &= \frac{\partial}{\partial \sigma_x^2} \left[-\frac{n}{2\sigma_x^2} + \frac{1}{2\sigma_x^4} \sum_{i=1}^n (\widehat{x}_i^2 - 2\mu_x \widehat{x}_i + \mu_x^2) \right] = \frac{n}{2\sigma_x^4} - \frac{1}{\sigma_x^6} \sum_{i=1}^n (\widehat{x}_i^2 - 2\mu_x \widehat{x}_i + \mu_x^2), \\
\ddot{Q}_{\phi\alpha} &= \frac{\partial}{\partial \boldsymbol{\phi}} \left[\sum_{i=1}^n \left(\widehat{\mathbf{z}}_i^{*\top} - \boldsymbol{\alpha}^\top - \widehat{x}_i \boldsymbol{\beta}^\top \right) \boldsymbol{\Omega}_{22}^{-1} \right] \\
&= \frac{\partial}{\partial \boldsymbol{\phi}} \left[\sum_{i=1}^n \frac{\widehat{z}_{i2} - \alpha_1 - \widehat{x}_i \beta_1}{\phi_2^2}, \dots, \sum_{i=1}^n \frac{\widehat{z}_{ip} - \alpha_r - \widehat{x}_i \beta_r}{\phi_p^2} \right] \\
&= \begin{bmatrix} 0 & \dots & 0 \\ -\frac{2}{\phi_2^3} \sum_{i=1}^n (\widehat{z}_{i2} - \alpha_1 - \widehat{x}_i \beta_1) & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & -\frac{2}{\phi_p^3} \sum_{i=1}^n (\widehat{z}_{ip} - \alpha_r - \widehat{x}_i \beta_r) \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
\ddot{Q}_{\phi\beta} &= \frac{\partial}{\partial \phi} \left[\sum_{i=1}^n \left(\widehat{xz}_i^* - \widehat{x}_i \boldsymbol{\alpha}^\top - \widehat{x}_i^2 \boldsymbol{\beta}^\top \right) \boldsymbol{\Omega}_{22}^{-1} \right] \\
&= \frac{\partial}{\partial \phi} \left[\sum_{i=1}^n \frac{\widehat{xz}_{i2} - \widehat{x}_i \alpha_1 - \widehat{x}_i^2 \beta_1}{\phi_2^2}, \dots, \sum_{i=1}^n \frac{\widehat{xz}_{ip} - \widehat{x}_i \alpha_r - \widehat{x}_i^2 \beta_r}{\phi_p^2} \right] \\
&= \begin{bmatrix} 0 & \dots & 0 \\ -\frac{2}{\phi_2^3} \sum_{i=1}^n \left(\widehat{xz}_{i2} - \widehat{x}_i \alpha_1 - \widehat{x}_i^2 \beta_1 \right) & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & -\frac{2}{\phi_p^3} \sum_{i=1}^n \left(\widehat{xz}_{ip} - \widehat{x}_i \alpha_r - \widehat{x}_i^2 \beta_r \right) \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
\ddot{Q}_{\phi\phi} &= \frac{\partial}{\partial \phi} \left[-\frac{n}{\phi_1} + \frac{1}{\phi_1^3} \sum_{i=1}^n \left(\widehat{z}_{i11}^2 - 2\widehat{xz}_{i1} + \widehat{x}_i^2 \right), -\frac{n}{\phi_2} + \frac{1}{\phi_2^3} \sum_{i=1}^n \left(\widehat{z}_{i22}^2 - 2\widehat{z}_{i2} \alpha_1 \right. \right. \\
&\quad \left. \left. - 2\widehat{xz}_{i2} \beta_1 + \alpha_1^2 + 2\widehat{x}_i \alpha_1 \beta_1 + \widehat{x}_i^2 \beta_1^2 \right), \dots, -\frac{n}{\phi_p} + \frac{1}{\phi_p^3} \sum_{i=1}^n \left(\widehat{z}_{ipp}^2 - 2\widehat{z}_{ip} \alpha_r \right. \right. \\
&\quad \left. \left. - 2\widehat{xz}_{ip} \beta_r + \alpha_r^2 + 2\widehat{x}_i \alpha_r \beta_r + \widehat{x}_i^2 \beta_r^2 \right) \right] = \text{diag}(H_{11}, \dots, H_{pp})
\end{aligned}$$

sendo que

$$\begin{aligned}
H_{11} &= \frac{n}{\phi_1^2} - \frac{3}{\phi_1^4} \sum_{i=1}^n \left(\widehat{z}_{i11}^2 - 2\widehat{xz}_{i1} + \widehat{x}_i^2 \right), \quad \text{e} \\
H_{jj} &= \frac{n}{\phi_j^2} - \frac{3}{\phi_j^4} \sum_{i=1}^n \left(\widehat{z}_{ijj}^2 - 2\widehat{z}_{ij} \alpha_{j-1} - 2\widehat{xz}_{ij} \beta_{j-1} + \alpha_{j-1}^2 + 2\widehat{x}_i \alpha_{j-1} \beta_{j-1} + \widehat{x}_i^2 \beta_{j-1}^2 \right)
\end{aligned}$$

para $j = 2, \dots, p$.

3.2.3.2 Perturbação de ponderação de casos

Consideramos uma atribuição arbitrária de pesos para o valor esperado da função de log-verossimilhança dos dados completos (função Q perturbada), que pode capturar observações com contribuição notável na função de log-verossimilhança e que pode exercer grande influência sobre as estimativas de máxima verossimilhança, representada por $Q(\boldsymbol{\theta}, \boldsymbol{\omega} | \widehat{\boldsymbol{\theta}})$. Neste caso a matriz $\boldsymbol{\Delta}$ tem elementos dados por

i)

$$\frac{\partial Q_i(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial \boldsymbol{\alpha}^\top} = \left(\widehat{\mathbf{z}}_i^{*\top} - \boldsymbol{\alpha}^\top - \widehat{x}_i \boldsymbol{\beta}^\top \right) \boldsymbol{\Omega}_{22}^{-1}, \quad i = 1, \dots, n.$$

Então,

$$\frac{\partial Q_i(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial \boldsymbol{\alpha}} = \boldsymbol{\Omega}_{22}^{-1} (\widehat{\mathbf{z}}_i^* - \boldsymbol{\alpha} - \widehat{x}_i \boldsymbol{\beta}), \quad i = 1, \dots, n.$$

Portanto,

$$\boldsymbol{\Delta}_\alpha = \boldsymbol{\Omega}_{22}^{-1} [\widehat{\mathbf{z}}_1^* - \boldsymbol{\alpha} - \widehat{x}_1 \boldsymbol{\beta}, \dots, \widehat{\mathbf{z}}_n^* - \boldsymbol{\alpha} - \widehat{x}_n \boldsymbol{\beta}]$$

em que $\widehat{\mathbf{z}}_i^* = [\widehat{z}_{i2}, \dots, \widehat{z}_{ip}]^\top$.

ii)

$$\frac{\partial Q_i(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial \boldsymbol{\beta}^\top} = \left(\widehat{x} \widehat{\mathbf{z}}_i^* - \widehat{x}_i \boldsymbol{\alpha}^\top - \widehat{x}_i^2 \boldsymbol{\beta}^\top \right) \boldsymbol{\Omega}_{22}^{-1}, \quad i = 1, \dots, n.$$

Então,

$$\frac{\partial Q_i(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial \boldsymbol{\beta}} = \boldsymbol{\Omega}_{22}^{-1} \left(\widehat{x} \widehat{\mathbf{z}}_i^{*\top} - \widehat{x}_i \boldsymbol{\alpha} - \widehat{x}_i^2 \boldsymbol{\beta} \right), \quad i = 1, \dots, n.$$

Portanto,

$$\boldsymbol{\Delta}_\beta = \boldsymbol{\Omega}_{22}^{-1} [\widehat{x} \widehat{\mathbf{z}}_1^{*\top} - \widehat{x}_1 \boldsymbol{\alpha} - \widehat{x}_1^2 \boldsymbol{\beta}, \dots, \widehat{x} \widehat{\mathbf{z}}_n^{*\top} - \widehat{x}_n \boldsymbol{\alpha} - \widehat{x}_n^2 \boldsymbol{\beta}]$$

em que $\widehat{x} \widehat{\mathbf{z}}_i^* = [\widehat{x} \widehat{z}_{i2}, \dots, \widehat{x} \widehat{z}_{ip}]$.

iii)

$$\begin{aligned} \frac{\partial Q_{1i}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\phi}|\widehat{\boldsymbol{\theta}})}{\partial \phi_1} &= -\frac{1}{\phi_1} + \frac{1}{\phi_1^3} \left(\widehat{z}_{i11}^2 - 2\widehat{x} \widehat{z}_{i1} + \widehat{x}_i^2 \right), \quad i = 1, \dots, n, \\ \frac{\partial Q_{1i}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\phi}|\widehat{\boldsymbol{\theta}})}{\partial \phi_j} &= -\frac{1}{\phi_j} + \frac{1}{\phi_j^3} \left(\widehat{z}_{ijj}^2 - 2\widehat{z}_{ij} \alpha_{j-1} - 2\widehat{x} \widehat{z}_{ij} \beta_{j-1} + \alpha_{j-1}^2 \right. \\ &\quad \left. + 2\widehat{x}_i \alpha_{j-1} \beta_{j-1} + \widehat{x}_i^2 \beta_{j-1}^2 \right), \quad i = 1, \dots, n, \quad j = 2, \dots, p. \end{aligned}$$

iv)

$$\frac{\partial Q_{2i}(\mu_x, \sigma_x^2|\widehat{\boldsymbol{\theta}})}{\partial \mu_x} = \frac{\widehat{x}_i - \mu_x}{\sigma_x^2}, \quad i = 1, \dots, n.$$

Portanto,

$$\boldsymbol{\Delta}_{\mu_x} = \frac{1}{\sigma_x^2} [\widehat{x}_1 - \mu_x, \dots, \widehat{x}_n - \mu_x].$$

v)

$$\frac{\partial Q_{2i}(\mu_x, \sigma_x^2 | \hat{\boldsymbol{\theta}})}{\partial \sigma_x^2} = -\frac{1}{2\sigma_x^2} + \frac{1}{2\sigma_x^4} \left(\hat{x}_i^2 - 2\mu_x \hat{x}_i + \mu_x^2 \right), \quad i = 1, \dots, n.$$

Portanto,

$$\boldsymbol{\Delta}_{\sigma_x^2} = \left[-\frac{1}{2\sigma_x^2} + \frac{1}{2\sigma_x^4} \left(\hat{x}_1^2 - 2\mu_x \hat{x}_1 + \mu_x^2 \right), \dots, -\frac{1}{2\sigma_x^2} + \frac{1}{2\sigma_x^4} \left(\hat{x}_n^2 - 2\mu_x \hat{x}_n + \mu_x^2 \right) \right].$$

avaliada em $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}} = \left(\hat{\boldsymbol{\alpha}}^\top, \hat{\boldsymbol{\beta}}^\top, \hat{\mu}_x, \hat{\sigma}_x^2, \hat{\boldsymbol{\phi}}^\top \right)^\top$.

3.2.3.3 Perturbação na covariável

Para avaliar a sensibilidade das estimativas quando são introduzidas pequenas perturbações na covariável, consideramos a perturbação aditiva $x_\omega = x + \omega$, pelo que $\boldsymbol{\omega}_0 = (0, \dots, 0)^\top$ e conseqüentemente,

$$\widehat{x_\omega \mathbf{z}_i} = E[x_\omega \mathbf{Z}_i^\top | \mathbf{V}_i] = E[(x_i + \omega_i) \mathbf{Z}_i^\top | \mathbf{V}_i] = \widehat{x \mathbf{z}_i} + \omega_i \widehat{\mathbf{z}_i}^\top,$$

$$\widehat{x_\omega} = E[x_\omega | \mathbf{V}_i] = \widehat{x}_i + \omega_i, \quad e$$

$$\widehat{x_\omega^2} = E[x_\omega^2 | \mathbf{V}_i] = E[(x_i + \omega_i)^2 | \mathbf{V}_i] = \widehat{x}_i^2 + \omega_i^2 + 2\omega_i \widehat{x}_i$$

Derivando a função $Q(\boldsymbol{\theta}, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})$ em relação a ω_i , temos

$$\frac{\partial Q_{1i}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})}{\partial \omega_i} = \widehat{\mathbf{z}}_i^\top \boldsymbol{\Omega}^{-1} \mathbf{b} - \mathbf{a}^\top \boldsymbol{\Omega}^{-1} \mathbf{b} - \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b} \omega_i - \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b} \widehat{x}_i, \quad (3.7)$$

$$\frac{\partial Q_{2i}(\mu_x, \sigma_x^2, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})}{\partial \omega_i} = -\frac{1}{\sigma_x^2} (\omega_i + \widehat{x}_i - \mu_x). \quad (3.8)$$

Então

i)

$$\begin{aligned} \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\omega}_i} &= \left(\frac{\partial}{\partial \boldsymbol{\alpha}^\top} \left[\frac{\partial Q(\boldsymbol{\theta}, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})}{\partial \omega_i} \right] \right)^\top = \left(\frac{\partial}{\partial \boldsymbol{\alpha}^\top} \left[\frac{\partial Q(\boldsymbol{\theta}, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})}{\partial \omega_i} \right] \frac{\partial \boldsymbol{\alpha}}{\partial \boldsymbol{\alpha}^\top} \right)^\top \\ &= \left(-\mathbf{b}^\top \boldsymbol{\Omega}^{-1} \begin{bmatrix} \mathbf{0} \\ \mathbf{I}_r \end{bmatrix} \right)^\top = \left(-[1, \boldsymbol{\beta}^\top] \begin{bmatrix} \frac{1}{\phi_1^2} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Omega}_{22}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ \mathbf{I}_r \end{bmatrix} \right)^\top = -\boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\beta}, \end{aligned}$$

pelo que

$$\boldsymbol{\Delta}_\alpha = \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\omega}^\top} = [-\boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\beta}, \dots, -\boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\beta}] = -\boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\beta} \mathbf{1}_n^\top, \quad \text{em que } \mathbf{1}_n = [1, \dots, 1]_{n \times 1}^\top.$$

ii)

$$\begin{aligned}
\frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega} | \widehat{\boldsymbol{\theta}})}{\partial \boldsymbol{\beta} \partial \omega_i} &= \left(\frac{\partial}{\partial \boldsymbol{\beta}^\top} \left[\frac{\partial Q(\boldsymbol{\theta}, \boldsymbol{\omega} | \widehat{\boldsymbol{\theta}})}{\partial \omega_i} \right] \right)^\top = \left(\frac{\partial}{\partial \boldsymbol{b}^\top} \left[\frac{\partial Q(\boldsymbol{\theta}, \boldsymbol{\omega} | \widehat{\boldsymbol{\theta}})}{\partial \omega_i} \right] \frac{\partial \boldsymbol{b}}{\partial \boldsymbol{\beta}^\top} \right)^\top \\
&= \left(\left[\widehat{\boldsymbol{z}}_i^\top \boldsymbol{\Omega}^{-1} - \boldsymbol{a}^\top \boldsymbol{\Omega}^{-1} - 2\boldsymbol{b}^\top \boldsymbol{\Omega}^{-1} \omega_i - 2\boldsymbol{b}^\top \boldsymbol{\Omega}^{-1} \widehat{\boldsymbol{x}}_i \right] \begin{bmatrix} \mathbf{0} \\ \mathbf{I}_r \end{bmatrix} \right)^\top \\
&= \left(\left(\left[\widehat{\boldsymbol{z}}_{i1}, \widehat{\boldsymbol{z}}_i^{*\top} \right] - [0, \boldsymbol{\alpha}^\top] - 2[1, \boldsymbol{\beta}^\top] \omega_i - 2[1, \boldsymbol{\beta}^\top] \widehat{\boldsymbol{x}}_i \right) \begin{bmatrix} \mathbf{0} \\ \boldsymbol{\Omega}_{22}^{-1} \end{bmatrix} \right)^\top \\
&= \boldsymbol{\Omega}_{22}^{-1} (\widehat{\boldsymbol{z}}_i^* - \boldsymbol{\alpha} - 2\omega_i \boldsymbol{\beta} - 2\widehat{\boldsymbol{x}}_i \boldsymbol{\beta})
\end{aligned}$$

pele que

$$\boldsymbol{\Delta}_\beta = \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega} | \widehat{\boldsymbol{\theta}})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\omega}^\top} = \boldsymbol{\Omega}_{22}^{-1} [\widehat{\boldsymbol{z}}_1^* - \boldsymbol{\alpha} - 2\omega_1 \boldsymbol{\beta} - 2\widehat{\boldsymbol{x}}_1 \boldsymbol{\beta}, \dots, \widehat{\boldsymbol{z}}_n^* - \boldsymbol{\alpha} - 2\omega_n \boldsymbol{\beta} - 2\widehat{\boldsymbol{x}}_n \boldsymbol{\beta}],$$

sendo $\widehat{\boldsymbol{z}}_i^* = [\widehat{z}_{i2}, \dots, \widehat{z}_{ip}]^\top$

iii) Multiplicando as matrizes da equação (3.7), temos que

$$\begin{aligned}
\frac{\partial Q_{1i}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\omega} | \widehat{\boldsymbol{\theta}})}{\partial \omega_i} &= \frac{\widehat{z}_{i1}}{\phi_1^2} + \sum_{j=2}^p \frac{\widehat{z}_{ij} \beta_{j-1}}{\phi_j^2} - \sum_{j=2}^p \frac{\alpha_{j-1} \beta_{j-1}}{\phi_j^2} - \frac{\omega_i}{\phi_1^2} - \sum_{j=2}^p \frac{\omega_i \beta_{j-1}^2}{\phi_j^2} \\
&\quad - \frac{\widehat{x}_i}{\phi_1^2} - \sum_{j=2}^p \frac{\widehat{x}_i \beta_{j-1}^2}{\phi_j^2}
\end{aligned}$$

Então

$$\frac{\partial^2 Q_{1i}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\omega} | \widehat{\boldsymbol{\theta}})}{\partial \phi_1 \partial \omega_i} = -\frac{2\widehat{z}_{i1}}{\phi_1^3} + \frac{2\omega_i}{\phi_1^3} + \frac{2\widehat{x}_i}{\phi_1^3}$$

pele que

$$\boldsymbol{\Delta}_{\phi_1} = \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega} | \widehat{\boldsymbol{\theta}})}{\partial \phi_1 \partial \boldsymbol{\omega}^\top} = -\frac{2}{\phi_1^3} [\widehat{z}_{11} - \omega_1 - \widehat{x}_1, \dots, \widehat{z}_{n1} - \omega_n - \widehat{x}_n], \quad e$$

$$\frac{\partial^2 Q_{1i}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\omega} | \widehat{\boldsymbol{\theta}})}{\partial \phi_j \partial \omega_i} = -\frac{2\widehat{z}_{ij} \beta_{j-1}}{\phi_j^3} + \frac{2\alpha_{j-1} \beta_{j-1}}{\phi_j^3} + \frac{2\omega_i \beta_{j-1}^2}{\phi_j^3} + \frac{2\widehat{x}_i \beta_{j-1}^2}{\phi_j^3}.$$

Portanto, para $j = 2, \dots, p$

$$\begin{aligned}
\boldsymbol{\Delta}_{\phi_j} &= -\frac{2}{\phi_j^3} [\widehat{z}_{1j} \beta_{j-1} - \alpha_{j-1} \beta_{j-1} - \omega_1 \beta_{j-1}^2 - \widehat{x}_1 \beta_{j-1}^2, \dots, \\
&\quad \widehat{z}_{nj} \beta_{j-1} - \alpha_{j-1} \beta_{j-1} - \omega_n \beta_{j-1}^2 - \widehat{x}_n \beta_{j-1}^2].
\end{aligned}$$

iv)

$$\frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega} | \widehat{\boldsymbol{\theta}})}{\partial \mu_x \partial \omega_i} = \frac{\partial}{\partial \mu_x} \left[-\frac{1}{\sigma_x^2} (\omega_i + \widehat{x}_i - \mu_x) \right] = \frac{1}{\sigma_x^2}, \quad i = 1, \dots, n.$$

Portanto,

$$\boldsymbol{\Delta}_{\mu_x} = \frac{1}{\sigma_x^2} [1, \dots, 1] = \frac{\mathbf{1}_n^\top}{\sigma_x^2}, \quad \text{sendo } \mathbf{1}_n = [1, \dots, 1]_{n \times 1}^\top.$$

v)

$$\frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega} | \widehat{\boldsymbol{\theta}})}{\partial \sigma_x^2 \partial \omega_i} = \frac{\partial}{\partial \sigma_x^2} \left[-\frac{1}{\sigma_x^2} (\omega_i + \widehat{x}_i - \mu_x) \right] = \frac{1}{\sigma_x^4} (\omega_i + \widehat{x}_i - \mu_x), \quad i = 1, \dots, n.$$

Portanto,

$$\boldsymbol{\Delta}_{\sigma_x^2} = \frac{1}{\sigma_x^4} [\omega_1 + \widehat{x}_1 - \mu_x, \dots, \omega_n + \widehat{x}_n - \mu_x].$$

3.2.4 Caso *MEMC-t*

3.2.4.1 Matriz hessiana

Para obter as medidas de diagnóstico para a influência local de um esquema de perturbação em particular, é necessário calcular a matriz hessiana.

Da Seção 2.3.2.1, temos que

$$\begin{aligned} Q_1(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\phi} | \widehat{\boldsymbol{\theta}}) = & -\frac{n}{2} \sum_{j=1}^p \log \phi_j^2 - \frac{1}{2} \sum_{i=1}^n \left[\text{tr} \left(\boldsymbol{\Omega}^{-1} \widehat{u} \widehat{z}_i^2 \right) - 2\mathbf{a}^\top \boldsymbol{\Omega}^{-1} \widehat{u} \widehat{z}_i - 2\widehat{u} \widehat{z}_i^\top \boldsymbol{\Omega}^{-1} \mathbf{b} \right. \\ & \left. + \mathbf{a}^\top \boldsymbol{\Omega}^{-1} \mathbf{a} \widehat{u}_i + 2\mathbf{a}^\top \boldsymbol{\Omega}^{-1} \mathbf{b} \widehat{u} \widehat{x}_i + \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b} \widehat{u} \widehat{x}_i^2 \right], \quad \text{e} \end{aligned} \quad (3.9)$$

$$Q_2(\mu_x, \sigma_x^2 | \widehat{\boldsymbol{\theta}}) = -\frac{n}{2} \log \sigma_x^2 - \frac{1}{2\sigma_x^2} \sum_{i=1}^n \left(\widehat{u} \widehat{x}_i^2 - 2\mu_x \widehat{u} \widehat{x}_i + \mu_x^2 \widehat{u}_i \right).$$

Daí, temos

$$\begin{aligned}
\ddot{Q}_{\alpha\alpha} &= \frac{\partial}{\partial \boldsymbol{\alpha}} \left[\sum_{i=1}^n \left(\widehat{u\mathbf{z}}_i^{*\top} - \widehat{u}_i \boldsymbol{\alpha}^\top - \widehat{ux}_i \boldsymbol{\beta}^\top \right) \boldsymbol{\Omega}_{22}^{-1} \right] = - \sum_{i=1}^n \widehat{u}_i \boldsymbol{\Omega}_{22}^{-1}, \\
\ddot{Q}_{\alpha\beta} &= \frac{\partial}{\partial \boldsymbol{\alpha}} \left[\sum_{i=1}^n \left(\widehat{ux\mathbf{z}}_i^* - \widehat{ux}_i \boldsymbol{\alpha}^\top - \widehat{ux}_i^2 \boldsymbol{\beta}^\top \right) \boldsymbol{\Omega}_{22}^{-1} \right] = - \sum_{i=1}^n \widehat{ux}_i \boldsymbol{\Omega}_{22}^{-1} = \ddot{Q}_{\beta\alpha}^\top, \\
\ddot{Q}_{\alpha\phi} &= \ddot{Q}_{\phi\alpha}^\top, \\
\ddot{Q}_{\beta\beta} &= \frac{\partial}{\partial \boldsymbol{\beta}} \left[\sum_{i=1}^n \left(\widehat{ux\mathbf{z}}_i^* - \widehat{ux}_i \boldsymbol{\alpha}^\top - \widehat{ux}_i^2 \boldsymbol{\beta}^\top \right) \boldsymbol{\Omega}_{22}^{-1} \right] = - \sum_{i=1}^n \widehat{ux}_i^2 \boldsymbol{\Omega}_{22}^{-1}, \\
\ddot{Q}_{\beta\phi} &= \ddot{Q}_{\phi\beta}^\top, \\
\ddot{Q}_{\mu_x \mu_x} &= \frac{\partial}{\partial \mu_x} \left[\frac{1}{\sigma_x^2} \sum_{i=1}^n (\widehat{ux}_i - \mu_x \widehat{u}_i) \right] = - \frac{1}{\sigma_x^2} \sum_{i=1}^n \widehat{u}_i, \\
\ddot{Q}_{\mu_x \sigma_x^2} &= \frac{\partial}{\partial \mu_x} \left[- \frac{n}{2\sigma_x^2} + \frac{1}{2\sigma_x^4} \sum_{i=1}^n (\widehat{ux}_i^2 - 2\mu_x \widehat{ux}_i + \mu_x^2 \widehat{u}_i) \right] = - \frac{1}{\sigma_x^4} \sum_{i=1}^n (\widehat{ux}_i - \mu_x \widehat{u}_i) = \ddot{Q}_{\sigma_x^2 \mu_x}, \\
\ddot{Q}_{\sigma_x^2 \sigma_x^2} &= \frac{\partial}{\partial \sigma_x^2} \left[- \frac{n}{2\sigma_x^2} + \frac{1}{2\sigma_x^4} \sum_{i=1}^n (\widehat{ux}_i^2 - 2\mu_x \widehat{ux}_i + \mu_x^2 \widehat{u}_i) \right] = \frac{n}{2\sigma_x^4} - \frac{1}{\sigma_x^6} \sum_{i=1}^n (\widehat{ux}_i^2 - 2\mu_x \widehat{ux}_i + \mu_x^2 \widehat{u}_i), \\
\ddot{Q}_{\phi\alpha} &= \frac{\partial}{\partial \boldsymbol{\phi}} \left[\sum_{i=1}^n \left(\widehat{u\mathbf{z}}_i^{*\top} - \widehat{u}_i \boldsymbol{\alpha}^\top - \widehat{ux}_i \boldsymbol{\beta}^\top \right) \boldsymbol{\Omega}_{22}^{-1} \right] \\
&= \frac{\partial}{\partial \boldsymbol{\phi}} \left[\sum_{i=1}^n \frac{\widehat{uz}_{i2} - \widehat{u}_i \alpha_1 - \widehat{ux}_i \beta_1}{\phi_2^2}, \dots, \sum_{i=1}^n \frac{\widehat{uz}_{ip} - \widehat{u}_i \alpha_r - \widehat{ux}_i \beta_r}{\phi_p^2} \right] \\
&= \begin{bmatrix} 0 & \dots & 0 \\ -\frac{2}{\phi_2^3} \sum_{i=1}^n (\widehat{uz}_{i2} - \widehat{u}_i \alpha_1 - \widehat{ux}_i \beta_1) & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & -\frac{2}{\phi_p^3} \sum_{i=1}^n (\widehat{uz}_{ip} - \widehat{u}_i \alpha_r - \widehat{ux}_i \beta_r) \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
\ddot{Q}_{\phi\beta} &= \frac{\partial}{\partial \phi} \left[\sum_{i=1}^n \left(\widehat{uxz}_i^* - \widehat{ux}_i \alpha^\top - \widehat{ux}_i^2 \beta^\top \right) \Omega_{22}^{-1} \right] \\
&= \frac{\partial}{\partial \phi} \left[\sum_{i=1}^n \frac{\widehat{uxz}_{i2} - \widehat{ux}_i \alpha_1 - \widehat{ux}_i^2 \beta_1}{\phi_2^2}, \dots, \sum_{i=1}^n \frac{\widehat{uxz}_{ip} - \widehat{ux}_i \alpha_r - \widehat{ux}_i^2 \beta_r}{\phi_p^2} \right] \\
&= \begin{bmatrix} 0 & \dots & 0 \\ -\frac{2}{\phi_2^3} \sum_{i=1}^n \left(\widehat{uxz}_{i2} - \widehat{ux}_i \alpha_1 - \widehat{ux}_i^2 \beta_1 \right) & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & -\frac{2}{\phi_p^3} \sum_{i=1}^n \left(\widehat{uxz}_{ip} - \widehat{ux}_i \alpha_r - \widehat{ux}_i^2 \beta_r \right) \end{bmatrix}
\end{aligned}$$

Por fim, temos

$$\begin{aligned}
\ddot{Q}_{\phi\phi} &= \frac{\partial}{\partial \phi} \left[-\frac{n}{\phi_1} + \frac{1}{\phi_1^3} \sum_{i=1}^n \left(\widehat{uz}_{i11}^2 - 2\widehat{uxz}_{i1} + \widehat{ux}_i^2 \right), -\frac{n}{\phi_2} + \frac{1}{\phi_2^3} \sum_{i=1}^n \left(\widehat{uz}_{i22}^2 - 2\widehat{uz}_{i2} \alpha_1 \right. \right. \\
&\quad \left. \left. - 2\widehat{uxz}_{i2} \beta_1 + \widehat{u}_i \alpha_1^2 + 2\widehat{ux}_i \alpha_1 \beta_1 + \widehat{ux}_i^2 \beta_1^2 \right), \dots, -\frac{n}{\phi_p} + \frac{1}{\phi_p^3} \sum_{i=1}^n \left(\widehat{uz}_{ipp}^2 - 2\widehat{uz}_{ip} \alpha_r \right. \right. \\
&\quad \left. \left. - 2\widehat{uxz}_{ip} \beta_r + \widehat{u}_i \alpha_r^2 + 2\widehat{ux}_i \alpha_r \beta_r + \widehat{ux}_i^2 \beta_r^2 \right) \right] = \text{diag}(H_{11}, \dots, H_{pp})
\end{aligned}$$

sendo que

$$\begin{aligned}
H_{11} &= \frac{n}{\phi_1^2} - \frac{3}{\phi_1^4} \sum_{i=1}^n \left(\widehat{uz}_{i11}^2 - 2\widehat{uxz}_{i1} + \widehat{ux}_i^2 \right), \quad \text{e} \\
H_{jj} &= \frac{n}{\phi_j^2} - \frac{3}{\phi_j^4} \sum_{i=1}^n \left(\widehat{uz}_{ijj}^2 - 2\widehat{uz}_{ij} \alpha_{j-1} - 2\widehat{uxz}_{ij} \beta_{j-1} + \widehat{u}_i \alpha_{j-1}^2 + 2\widehat{ux}_i \alpha_{j-1} \beta_{j-1} + \widehat{ux}_i^2 \beta_{j-1}^2 \right)
\end{aligned}$$

para $j = 2, \dots, p$.

3.2.4.2 Perturbação de ponderação de casos

Consideramos uma atribuição arbitrária de pesos para a função Q perturbada, que pode capturar observações com contribuição notável na função de log-verossimilhança e que pode exercer grande influência sobre as estimativas de máxima verossimilhança.

Considerando que $Q_i(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}) = Q_{1i}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \phi|\widehat{\boldsymbol{\theta}}) + Q_{2i}(\mu_x, \sigma_x^2|\widehat{\boldsymbol{\theta}})$, em que

$$\begin{aligned}
Q_{1i}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \phi|\widehat{\boldsymbol{\theta}}) &= -\frac{1}{2} \sum_{j=1}^p \log \phi_j^2 - \frac{1}{2} \text{tr}(\boldsymbol{\Omega}^{-1} \widehat{\boldsymbol{u}} \widehat{\boldsymbol{z}}_i^2) + \boldsymbol{a}^\top \boldsymbol{\Omega}^{-1} \widehat{\boldsymbol{u}} \widehat{\boldsymbol{z}}_i + \widehat{\boldsymbol{u}} \widehat{\boldsymbol{x}}_i \boldsymbol{\Omega}^{-1} \boldsymbol{b} - \frac{1}{2} \boldsymbol{a}^\top \boldsymbol{\Omega}^{-1} \boldsymbol{a} \widehat{u}_i \\
&\quad - \boldsymbol{a}^\top \boldsymbol{\Omega}^{-1} \boldsymbol{b} \widehat{u}_i - \frac{1}{2} \boldsymbol{b}^\top \boldsymbol{\Omega}^{-1} \boldsymbol{b} \widehat{u}_i^2, \quad \text{e} \\
Q_{2i}(\mu_x, \sigma_x^2|\widehat{\boldsymbol{\theta}}) &= -\frac{1}{2} \log \sigma_x^2 - \frac{1}{2\sigma_x^2} (\widehat{u}_i^2 - 2\mu_x \widehat{u}_i + \mu_x^2 \widehat{u}_i),
\end{aligned} \tag{3.10}$$

os elementos da matriz $\boldsymbol{\Delta}$ sao dados por

i)

$$\begin{aligned}
\frac{\partial Q_i(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial \boldsymbol{\alpha}^\top} &= \frac{\partial Q_{1i}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \phi|\widehat{\boldsymbol{\theta}})}{\partial \boldsymbol{\alpha}^\top} \frac{\partial \boldsymbol{\alpha}}{\partial \boldsymbol{\alpha}^\top} = (\widehat{\boldsymbol{u}} \widehat{\boldsymbol{z}}_i^\top \boldsymbol{\Omega}^{-1} - \boldsymbol{a}^\top \boldsymbol{\Omega}^{-1} \widehat{u}_i - \boldsymbol{b}^\top \boldsymbol{\Omega}^{-1} \widehat{u}_i) \begin{bmatrix} \mathbf{0} \\ \mathbf{I}_r \end{bmatrix} \\
&= (\widehat{\boldsymbol{u}} \widehat{\boldsymbol{z}}_i^\top - \boldsymbol{a}^\top \widehat{u}_i - \boldsymbol{b}^\top \widehat{u}_i) \begin{bmatrix} \frac{1}{\phi_1^2} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Omega}_{22}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ \mathbf{I}_r \end{bmatrix} \\
&= \left([\widehat{u}_{z_{i1}}, \widehat{\boldsymbol{u}} \widehat{\boldsymbol{z}}_i^{*\top}] - [0, \boldsymbol{\alpha}^\top \widehat{u}_i] - [\widehat{u}_i, \boldsymbol{\beta}^\top \widehat{u}_i] \right) \begin{bmatrix} \mathbf{0} \\ \boldsymbol{\Omega}_{22}^{-1} \end{bmatrix} \\
&= \left(\widehat{\boldsymbol{u}} \widehat{\boldsymbol{z}}_i^{*\top} - \boldsymbol{\alpha}^\top \widehat{u}_i - \boldsymbol{\beta}^\top \widehat{u}_i \right) \boldsymbol{\Omega}_{22}^{-1}.
\end{aligned}$$

Então,

$$\frac{\partial Q_i(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial \boldsymbol{\alpha}} = \left(\frac{\partial Q_i(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial \boldsymbol{\alpha}^\top} \right)^\top = \boldsymbol{\Omega}_{22}^{-1} (\widehat{\boldsymbol{u}} \widehat{\boldsymbol{z}}_i^* - \widehat{u}_i \boldsymbol{\alpha} - \widehat{u}_i \boldsymbol{\beta}), \quad \text{para } i = 1, \dots, n.$$

Portanto, a matriz $\boldsymbol{\Delta}_\alpha$ é dada por

$$\boldsymbol{\Delta}_\alpha = \boldsymbol{\Omega}_{22}^{-1} [\widehat{\boldsymbol{u}} \widehat{\boldsymbol{z}}_1^* - \widehat{u}_1 \boldsymbol{\alpha} - \widehat{u}_1 \boldsymbol{\beta}, \dots, \widehat{\boldsymbol{u}} \widehat{\boldsymbol{z}}_n^* - \widehat{u}_n \boldsymbol{\alpha} - \widehat{u}_n \boldsymbol{\beta}]$$

em que $\widehat{\boldsymbol{u}} \widehat{\boldsymbol{z}}_i^* = [\widehat{u}_{z_{i2}}, \dots, \widehat{u}_{z_{ip}}]^\top$.

ii)

$$\begin{aligned}
\frac{\partial Q_i(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial \boldsymbol{\beta}^\top} &= \frac{\partial Q_{1i}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\phi}|\widehat{\boldsymbol{\theta}})}{\partial \mathbf{b}^\top} \frac{\partial \mathbf{b}}{\partial \boldsymbol{\beta}^\top} = (\widehat{uxz}_i \boldsymbol{\Omega}^{-1} - \boldsymbol{\alpha}^\top \boldsymbol{\Omega}^{-1} \widehat{ux}_i - \boldsymbol{\beta}^\top \boldsymbol{\Omega}^{-1} \widehat{ux}_i^2) \begin{bmatrix} \mathbf{0} \\ \mathbf{I}_r \end{bmatrix} \\
&= (\widehat{uxz}_i - \boldsymbol{\alpha}^\top \widehat{ux}_i - \boldsymbol{\beta}^\top \widehat{ux}_i^2) \begin{bmatrix} \frac{1}{\phi_1^2} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Omega}_{22}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ \mathbf{I}_r \end{bmatrix} \\
&= \left([\widehat{uxz}_{i1}, \widehat{uxz}_i^*] - [0, \boldsymbol{\alpha}^\top \widehat{ux}_i] - [\widehat{ux}_i^2, \boldsymbol{\beta}^\top \widehat{ux}_i^2] \right) \begin{bmatrix} \mathbf{0} \\ \boldsymbol{\Omega}_{22}^{-1} \end{bmatrix} \\
&= \left(\widehat{uxz}_i^* - \boldsymbol{\alpha}^\top \widehat{ux}_i - \boldsymbol{\beta}^\top \widehat{ux}_i^2 \right) \boldsymbol{\Omega}_{22}^{-1}.
\end{aligned}$$

Então,

$$\frac{\partial Q_i(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial \boldsymbol{\beta}} = \left(\frac{\partial Q_i(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial \boldsymbol{\beta}^\top} \right)^\top = \boldsymbol{\Omega}_{22}^{-1} \left(\widehat{uxz}_i^{*\top} - \widehat{ux}_i \boldsymbol{\alpha} - \widehat{ux}_i^2 \boldsymbol{\beta} \right), \quad \text{para } i = 1, \dots, n.$$

Portanto, a matriz $\boldsymbol{\Delta}_\beta$ é dada por

$$\boldsymbol{\Delta}_\beta = \boldsymbol{\Omega}_{22}^{-1} \left[\widehat{uxz}_1^{*\top} - \widehat{ux}_1 \boldsymbol{\alpha} - \widehat{ux}_1^2 \boldsymbol{\beta}, \dots, \widehat{uxz}_n^{*\top} - \widehat{ux}_n \boldsymbol{\alpha} - \widehat{ux}_n^2 \boldsymbol{\beta} \right]$$

em que $\widehat{uxz}_i^* = [\widehat{uxz}_{i2}, \dots, \widehat{uxz}_{ip}]$.

iii) A partir de (3.10) e após alguma manipulação algébrica temos

$$\begin{aligned}
Q_{1i}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\phi}|\widehat{\boldsymbol{\theta}}) &= -\frac{1}{2} \sum_{j=1}^p \log \phi_j^2 - \frac{1}{2} \sum_{j=1}^p \frac{\widehat{uz}_{ij}^2}{\phi_j^2} + \sum_{j=2}^p \frac{\widehat{uz}_{ij} \alpha_{j-1}}{\phi_j^2} + \frac{\widehat{uxz}_{i1}}{\phi_1^2} + \sum_{j=2}^p \frac{\widehat{uxz}_{ij} \beta_{j-1}}{\phi_j^2} \\
&\quad - \frac{1}{2} \sum_{j=2}^p \frac{\widehat{u}_i \alpha_{j-1}^2}{\phi_j^2} - \sum_{j=2}^p \frac{\widehat{ux}_i \alpha_{j-1} \beta_{j-1}}{\phi_j^2} - \frac{\widehat{ux}_i^2}{2\phi_1^2} - \frac{1}{2} \sum_{j=2}^p \frac{\widehat{ux}_i^2 \beta_{j-1}^2}{\phi_j^2}
\end{aligned}$$

pelo que

$$\begin{aligned}
\frac{\partial Q_{1i}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\phi}|\widehat{\boldsymbol{\theta}})}{\partial \phi_1} &= -\frac{1}{\phi_1} + \frac{1}{\phi_1^3} \left(\widehat{uz}_{i1}^2 - 2\widehat{uxz}_{i1} + \widehat{ux}_i^2 \right), \quad i = 1, \dots, n, \quad \text{e} \\
\frac{\partial Q_{1i}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\phi}|\widehat{\boldsymbol{\theta}})}{\partial \phi_j} &= -\frac{1}{\phi_j} + \frac{1}{\phi_j^3} \left(\widehat{uz}_{ij}^2 - 2\widehat{uz}_{ij} \alpha_{j-1} - 2\widehat{uxz}_{ij} \beta_{j-1} + \widehat{u}_i \alpha_{j-1}^2 \right. \\
&\quad \left. + 2\widehat{ux}_i \alpha_{j-1} \beta_{j-1} + \widehat{ux}_i^2 \beta_{j-1}^2 \right), \quad i = 1, \dots, n, \quad j = 2, \dots, p.
\end{aligned}$$

iv)

$$\frac{\partial Q_i(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial \mu_x} = \frac{\partial Q_{2i}(\mu_x, \sigma_x^2|\widehat{\boldsymbol{\theta}})}{\partial \mu_x} = -\frac{1}{2\sigma_x^2}(-2\widehat{u}x_i + 2\mu_x\widehat{u}_i) = \frac{\widehat{u}x_i - \mu_x\widehat{u}_i}{\sigma_x^2}, \quad i = 1, \dots, n.$$

Portanto,

$$\boldsymbol{\Delta}_{\mu_x} = \frac{1}{\sigma_x^2} [\widehat{u}x_1 - \mu_x\widehat{u}_1, \dots, \widehat{u}x_n - \mu_x\widehat{u}_n].$$

v)

$$\frac{\partial Q_i(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial \sigma_x^2} = \frac{\partial Q_{2i}(\mu_x, \sigma_x^2|\widehat{\boldsymbol{\theta}})}{\partial \sigma_x^2} = -\frac{1}{2\sigma_x^2} + \frac{1}{2\sigma_x^4} (\widehat{u}x_i^2 - 2\mu_x\widehat{u}x_i + \mu_x^2\widehat{u}_i), \quad i = 1, \dots, n.$$

Portanto,

$$\boldsymbol{\Delta}_{\sigma_x^2} = \left[-\frac{1}{2\sigma_x^2} + \frac{1}{2\sigma_x^4} (\widehat{u}x_1^2 - 2\mu_x\widehat{u}x_1 + \mu_x^2\widehat{u}_1), \dots, -\frac{1}{2\sigma_x^2} + \frac{1}{2\sigma_x^4} (\widehat{u}x_n^2 - 2\mu_x\widehat{u}x_n + \mu_x^2\widehat{u}_n) \right].$$

avaliada em $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}} = (\widehat{\boldsymbol{\alpha}}^\top, \widehat{\boldsymbol{\beta}}^\top, \widehat{\mu}_x, \widehat{\sigma}_x^2, \widehat{\boldsymbol{\phi}}^\top)^\top$.

3.2.4.3 Perturbação na covariável

No intuito de avaliar a sensibilidade das estimativas quando são introduzidas pequenas perturbações na covariável, consideramos a perturbação aditiva $x_\omega = x + \omega$, pelo que $\boldsymbol{\omega}_0 = (0, \dots, 0)^\top$.

Utilizando (2.26) e a perturbação na covariável temos

$$Q_1(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\omega}|\widehat{\boldsymbol{\theta}}) = -\frac{n}{2} \sum_{j=1}^p \log \phi_j^2 - \frac{1}{2} \sum_{i=1}^n \left[\text{tr}(\boldsymbol{\Omega}^{-1} \widehat{u}z_i^2) - 2\mathbf{a}^\top \boldsymbol{\Omega}^{-1} \widehat{u}z_i - 2\widehat{u}x_\omega z_i \boldsymbol{\Omega}^{-1} \mathbf{b} + \mathbf{a}^\top \boldsymbol{\Omega}^{-1} \mathbf{a} \widehat{u}_i + 2\mathbf{a}^\top \boldsymbol{\Omega}^{-1} \mathbf{b} \widehat{u}x_{\omega i} + \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b} \widehat{u}x_{\omega i}^2 \right], \quad \text{e} \quad (3.11)$$

$$Q_2(\mu_x, \sigma_x^2, \boldsymbol{\omega}|\widehat{\boldsymbol{\theta}}) = -\frac{n}{2} \log \sigma_x^2 - \frac{1}{2\sigma_x^2} \sum_{i=1}^n (\widehat{u}x_{\omega i}^2 - 2\mu_x \widehat{u}x_{\omega i} + \mu_x^2 \widehat{u}_i).$$

em que $\widehat{u}z_i^2$, $\widehat{u}z_i$ e \widehat{u}_i são como em (2.30) e

$$\widehat{u}x_\omega z_i = E[U_i x_{\omega i} \mathbf{Z}_i^\top | \mathbf{V}_i] = E[U_i (x_i + \omega_i) \mathbf{Z}_i^\top | \mathbf{V}_i] = \widehat{u}x z_i + \omega_i \widehat{u}z_i^\top,$$

$$\widehat{u}x_{\omega i} = E[U_i x_{\omega i} | \mathbf{V}_i] = \widehat{u}x_i + \omega_i \widehat{u}_i, \quad \text{e}$$

$$\widehat{u}x_{\omega i}^2 = E[U_i x_{\omega i}^2 | \mathbf{V}_i] = E[U_i (x_i + \omega_i)^2 | \mathbf{V}_i] = \widehat{u}x_i^2 + \omega_i^2 \widehat{u}_i + 2\omega_i \widehat{u}x_i$$

Derivando (3.11) em relação a ω_i para obter $\frac{\partial Q(\boldsymbol{\theta}, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\omega}^\top}$, temos

$$\frac{\partial Q_{1i}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \phi, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})}{\partial \omega_i} = \widehat{u} \widehat{\mathbf{z}}_i^\top \boldsymbol{\Omega}^{-1} \mathbf{b} - \boldsymbol{\alpha}^\top \boldsymbol{\Omega}^{-1} \mathbf{b} \widehat{u}_i - \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b} \widehat{u}_i \omega_i - \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b} \widehat{u} \widehat{x}_i, \quad (3.12)$$

$$\frac{\partial Q_{2i}(\mu_x, \sigma_x^2, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})}{\partial \omega_i} = -\frac{1}{\sigma_x^2} (\omega_i \widehat{u}_i + \widehat{u} \widehat{x}_i - \mu_x \widehat{u}_i) \quad (3.13)$$

A partir destas duas ultimas equações, obtemos os elementos da matriz $\boldsymbol{\Delta}$ para este esquema de perturbação.

i)

$$\begin{aligned} \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\alpha} \partial \omega_i} &= \left(\frac{\partial}{\partial \boldsymbol{\alpha}^\top} \left[\frac{\partial Q_{1i}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \phi, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})}{\partial \omega_i} \right] \right)^\top = \left(\frac{\partial}{\partial \boldsymbol{\alpha}^\top} \left[\frac{\partial Q_{1i}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \phi, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})}{\partial \omega_i} \right] \frac{\partial \boldsymbol{\alpha}}{\partial \boldsymbol{\alpha}^\top} \right)^\top \\ &= \left(-\widehat{u}_i \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \begin{bmatrix} \mathbf{0} \\ \mathbf{I}_r \end{bmatrix} \right)^\top = \left(-\widehat{u}_i [1, \boldsymbol{\beta}^\top] \begin{bmatrix} \frac{1}{\phi_1^2} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Omega}_{22}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ \mathbf{I}_r \end{bmatrix} \right)^\top \\ &= \left(-[\widehat{u}_i, \widehat{u}_i \boldsymbol{\beta}^\top] \begin{bmatrix} \mathbf{0} \\ \boldsymbol{\Omega}_{22}^{-1} \end{bmatrix} \right)^\top = (-\widehat{u}_i \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1})^\top = -\boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\beta} \widehat{u}_i, \end{aligned}$$

pelo que

$$\boldsymbol{\Delta}_\alpha = \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\omega}^\top} = [-\boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\beta} \widehat{u}_1, \dots, -\boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\beta} \widehat{u}_n] = -\boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\beta} \widehat{\mathbf{u}}^\top, \text{ em que } \widehat{\mathbf{u}} = [\widehat{u}_1, \dots, \widehat{u}_n]^\top.$$

ii)

$$\begin{aligned} \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\beta} \partial \omega_i} &= \left(\frac{\partial}{\partial \boldsymbol{\beta}^\top} \left[\frac{\partial Q_{1i}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \phi, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})}{\partial \omega_i} \right] \right)^\top = \left(\frac{\partial}{\partial \boldsymbol{\beta}^\top} \left[\frac{\partial Q_{1i}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \phi, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})}{\partial \omega_i} \right] \frac{\partial \boldsymbol{\beta}}{\partial \boldsymbol{\beta}^\top} \right)^\top \\ &= \left(\left[\widehat{u} \widehat{\mathbf{z}}_i^\top \boldsymbol{\Omega}^{-1} - \boldsymbol{\alpha}^\top \boldsymbol{\Omega}^{-1} \widehat{u}_i - 2\mathbf{b}^\top \boldsymbol{\Omega}^{-1} \widehat{u}_i \omega_i - 2\mathbf{b}^\top \boldsymbol{\Omega}^{-1} \widehat{u} \widehat{x}_i \right] \begin{bmatrix} \mathbf{0} \\ \mathbf{I}_r \end{bmatrix} \right)^\top \\ &= \left(\left(\left[\widehat{u} \widehat{\mathbf{z}}_{i1}, \widehat{u} \widehat{\mathbf{z}}_i^{*\top} \right] - [0, \boldsymbol{\alpha}^\top] \widehat{u}_i - 2 [1, \boldsymbol{\beta}^\top] \widehat{u}_i \omega_i - 2 [1, \boldsymbol{\beta}^\top] \widehat{u} \widehat{x}_i \right) \begin{bmatrix} \mathbf{0} \\ \boldsymbol{\Omega}_{22}^{-1} \end{bmatrix} \right)^\top \\ &= \boldsymbol{\Omega}_{22}^{-1} (\widehat{u} \widehat{\mathbf{z}}_i^* - \widehat{u}_i \boldsymbol{\alpha} - 2\widehat{u}_i \omega_i \boldsymbol{\beta} - 2\widehat{u} \widehat{x}_i \boldsymbol{\beta}) \end{aligned}$$

pelo que

$$\begin{aligned} \boldsymbol{\Delta}_\beta &= \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\omega}^\top} \\ &= \boldsymbol{\Omega}_{22}^{-1} [\widehat{u} \widehat{\mathbf{z}}_1^* - \widehat{u}_1 \boldsymbol{\alpha} - 2\widehat{u}_1 \omega_1 \boldsymbol{\beta} - 2\widehat{u} \widehat{x}_1 \boldsymbol{\beta}, \dots, \widehat{u} \widehat{\mathbf{z}}_n^* - \widehat{u}_n \boldsymbol{\alpha} - 2\widehat{u}_n \omega_n \boldsymbol{\beta} - 2\widehat{u} \widehat{x}_n \boldsymbol{\beta}], \end{aligned}$$

sendo $\widehat{u\mathbf{z}}_i^* = [\widehat{uz}_{i2}, \dots, \widehat{uz}_{ip}]^\top$.

iii) Fazendo a multiplicação das matrizes da equação (3.12), temos que

$$\begin{aligned} \frac{\partial Q_{1i}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\omega} | \widehat{\boldsymbol{\theta}})}{\partial \omega_i} &= \frac{\widehat{uz}_{i1}}{\phi_1^2} + \sum_{j=2}^p \frac{\widehat{uz}_{ij}\beta_{j-1}}{\phi_j^2} - \sum_{j=2}^p \frac{\widehat{u}_i\alpha_{j-1}\beta_{j-1}}{\phi_j^2} - \frac{\widehat{u}_i\omega_i}{\phi_1^2} - \sum_{j=2}^p \frac{\widehat{u}_i\omega_i\beta_{j-1}^2}{\phi_j^2} \\ &\quad - \frac{\widehat{ux}_i}{\phi_1^2} - \sum_{j=2}^p \frac{\widehat{ux}_i\beta_{j-1}^2}{\phi_j^2} \end{aligned}$$

Então

$$\frac{\partial^2 Q_{1i}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\omega} | \widehat{\boldsymbol{\theta}})}{\partial \phi_1 \partial \omega_i} = -\frac{2\widehat{uz}_{i1}}{\phi_1^3} + \frac{2\widehat{u}_i\omega_i}{\phi_1^3} + \frac{2\widehat{ux}_i}{\phi_1^3}$$

pelo que

$$\Delta_{\phi_1} = \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega} | \widehat{\boldsymbol{\theta}})}{\partial \phi_1 \partial \boldsymbol{\omega}^\top} = -\frac{2}{\phi_1^3} [\widehat{uz}_{11} - \widehat{u}_1\omega_1 - \widehat{ux}_1, \dots, \widehat{uz}_{n1} - \widehat{u}_n\omega_n - \widehat{ux}_n], \quad e$$

$$\frac{\partial^2 Q_{1i}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\omega} | \widehat{\boldsymbol{\theta}})}{\partial \phi_j \partial \omega_i} = -\frac{2\widehat{uz}_{ij}\beta_{j-1}}{\phi_j^3} + \frac{2\widehat{u}_i\alpha_{j-1}\beta_{j-1}}{\phi_j^3} + \frac{2\widehat{u}_i\omega_i\beta_{j-1}^2}{\phi_j^3} + \frac{2\widehat{ux}_i\beta_{j-1}^2}{\phi_j^3}.$$

Portanto, para $j = 2, \dots, p$

$$\begin{aligned} \Delta_{\phi_j} &= -\frac{2}{\phi_j^3} [\widehat{uz}_{1j}\beta_{j-1} - \widehat{u}_1\alpha_{j-1}\beta_{j-1} - \widehat{u}_1\omega_1\beta_{j-1}^2 - \widehat{ux}_1\beta_{j-1}^2, \dots, \\ &\quad \widehat{uz}_{nj}\beta_{j-1} - \widehat{u}_n\alpha_{j-1}\beta_{j-1} - \widehat{u}_n\omega_n\beta_{j-1}^2 - \widehat{ux}_n\beta_{j-1}^2]. \end{aligned}$$

iv)

$$\frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega} | \widehat{\boldsymbol{\theta}})}{\partial \mu_x \partial \omega_i} = \frac{\partial}{\partial \mu_x} \left[\frac{\partial Q_{2i}(\mu_x, \sigma_x^2, \boldsymbol{\omega} | \widehat{\boldsymbol{\theta}})}{\partial \omega_i} \right] = \frac{\widehat{u}_i}{\sigma_x^2}, \quad i = 1, \dots, n.$$

Portanto,

$$\Delta_{\mu_x} = \frac{1}{\sigma_x^2} [\widehat{u}_1, \dots, \widehat{u}_n] = \frac{\widehat{\mathbf{u}}^\top}{\sigma_x^2}, \quad \text{sendo } \widehat{\mathbf{u}} = [\widehat{u}_1, \dots, \widehat{u}_n]^\top.$$

v)

$$\frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega} | \widehat{\boldsymbol{\theta}})}{\partial \sigma_x^2 \partial \omega_i} = \frac{\partial}{\partial \sigma_x^2} \left[\frac{\partial Q_{2i}(\mu_x, \sigma_x^2, \boldsymbol{\omega} | \widehat{\boldsymbol{\theta}})}{\partial \omega_i} \right] = \frac{1}{\sigma_x^4} (\omega_i \widehat{u}_i + \widehat{ux}_i - \mu_x \widehat{u}_i), \quad i = 1, \dots, n.$$

Portanto,

$$\Delta_{\sigma_x^2} = \frac{1}{\sigma_x^4} [\omega_1 \widehat{u}_1 + \widehat{ux}_1 - \mu_x \widehat{u}_1, \dots, \omega_n \widehat{u}_n + \widehat{ux}_n - \mu_x \widehat{u}_n].$$

3.2.5 Caso MEMC-SI

No caso do modelo com erros de medida e respostas censuradas baseados na distribuição slash (equação (2.16)), temos que $\mathbf{Z}_i \sim Sl_p(\boldsymbol{\mu}_z, \boldsymbol{\Sigma}_z, \nu)$ e a função de log-verossimilhança dos dados completos é dada por

$$\begin{aligned} \ell_c(\boldsymbol{\theta}|\mathbf{Z}_c) &= -\frac{n}{2} \sum_{j=1}^p \log \phi_j^2 - \frac{1}{2} \sum_{i=1}^n u_i (\mathbf{Z}_i - \mathbf{a} - \mathbf{b}x_i)^\top \boldsymbol{\Omega}^{-1} (\mathbf{Z}_i - \mathbf{a} - \mathbf{b}x_i) - \frac{n}{2} \log \sigma_x^2 \\ &\quad - \frac{1}{2\sigma_x^2} \sum_{i=1}^n u_i (x_i - \mu_x)^2 \end{aligned}$$

e desde que $Q(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}}^{(k)}) = E_{\hat{\boldsymbol{\theta}}^{(k)}}[\ell_c(\boldsymbol{\theta}|\mathbf{Z}_c)|\mathbf{V}]$, sendo

$$Q(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}}^{(k)}) = Q_1(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\phi}|\hat{\boldsymbol{\theta}}^{(k)}) + Q_2(\mu_x, \sigma_x^2|\hat{\boldsymbol{\theta}}^{(k)})$$

temos que

$$\begin{aligned} Q_1(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\phi}|\hat{\boldsymbol{\theta}}^{(k)}) &= E_{\hat{\boldsymbol{\theta}}^{(k)}} \left[-\frac{n}{2} \sum_{j=1}^p \log \phi_j^2 - \frac{1}{2} \sum_{i=1}^n u_i (\mathbf{Z}_i - \mathbf{a} - \mathbf{b}x_i)^\top \boldsymbol{\Omega}^{-1} (\mathbf{Z}_i - \mathbf{a} - \mathbf{b}x_i) \middle| \mathbf{V} \right], \text{ e} \\ Q_2(\mu_x, \sigma_x^2|\hat{\boldsymbol{\theta}}^{(k)}) &= E_{\hat{\boldsymbol{\theta}}^{(k)}} \left[-\frac{n}{2} \log \sigma_x^2 - \frac{1}{2\sigma_x^2} \sum_{i=1}^n u_i (x_i - \mu_x)^2 \middle| \mathbf{V} \right]. \end{aligned}$$

Novamente expandindo as expressões de $Q_1(\cdot|\hat{\boldsymbol{\theta}}^{(k)})$ e $Q_2(\cdot|\hat{\boldsymbol{\theta}}^{(k)})$ e tomando esperanças, segue que

$$\begin{aligned} Q_1(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\phi}|\hat{\boldsymbol{\theta}}^{(k)}) &= -\frac{n}{2} \sum_{j=1}^p \log \phi_j^2 - \frac{1}{2} \sum_{i=1}^n \left[\text{tr} \left(\boldsymbol{\Omega}^{-1} \widehat{u\mathbf{z}_i^2} \right) - 2\mathbf{a}^\top \boldsymbol{\Omega}^{-1} \widehat{u\mathbf{z}_i} - 2\widehat{ux_i} \mathbf{z}_i^\top \boldsymbol{\Omega}^{-1} \mathbf{b} \right. \\ &\quad \left. + \mathbf{a}^\top \boldsymbol{\Omega}^{-1} \mathbf{a} \widehat{u_i} + 2\mathbf{a}^\top \boldsymbol{\Omega}^{-1} \mathbf{b} \widehat{ux_i} + \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b} \widehat{ux_i^2} \right], \\ Q_2(\mu_x, \sigma_x^2|\hat{\boldsymbol{\theta}}^{(k)}) &= -\frac{n}{2} \log \sigma_x^2 - \frac{1}{2\sigma_x^2} \sum_{i=1}^n \left(\widehat{ux_i^2} - 2\mu_x \widehat{ux_i} + \mu_x^2 \widehat{u_i} \right) \end{aligned}$$

em que $\widehat{u\mathbf{z}_i^2} = E[U_i \mathbf{Z}_i \mathbf{Z}_i^\top | \mathbf{V}_i]$, $\widehat{u\mathbf{z}_i} = E[U_i \mathbf{Z}_i | \mathbf{V}_i]$, $\widehat{u_i} = E[U_i | \mathbf{V}_i]$, $\widehat{ux_i} = E[U_i x_i \mathbf{Z}_i^\top | \mathbf{V}_i]$, $\widehat{ux_i} = E[U_i x_i | \mathbf{V}_i]$, e $\widehat{ux_i^2} = E[U_i x_i^2 | \mathbf{V}_i]$, as quais tem as seguintes formas:

i) O i -ésimo indivíduo não tem componentes censurados. Aqui, $\mathbf{V}_i = \mathbf{Z}_i$, então

$$\begin{aligned} \widehat{u_i} &= E[U_i | \mathbf{V}_i] = E[U_i | \mathbf{Z}_i] = \frac{p + 2\nu}{\delta_i} \frac{P_1\left(\frac{p}{2} + \nu + 1, \frac{\delta_i}{2}\right)}{P_1\left(\frac{p}{2} + \nu, \frac{\delta_i}{2}\right)}, \\ \widehat{u\mathbf{z}_i} &= E[U_i \mathbf{Z}_i | \mathbf{V}_i] = E[U_i \mathbf{Z}_i | \mathbf{Z}_i] = E[U_i | \mathbf{Z}_i] \mathbf{z}_i = \widehat{u_i} \mathbf{z}_i, \text{ e} \\ \widehat{u\mathbf{z}_i^2} &= E[U_i \mathbf{Z}_i \mathbf{Z}_i^\top | \mathbf{V}_i] = E[U_i \mathbf{Z}_i \mathbf{Z}_i^\top | \mathbf{Z}_i] = E[U_i | \mathbf{Z}_i] \mathbf{z}_i \mathbf{z}_i^\top = \widehat{u_i} \mathbf{z}_i \mathbf{z}_i^\top. \end{aligned}$$

em que δ_i é dado na equação (1.4) e $P_x(a, b)$ denota a fda da distribuição Gama(a,b), com média $\frac{a}{b}$, avaliada em x .

- ii) O i -ésimo indivíduo tem apenas componentes censurados. Neste caso, $\mathbf{Z}_i \leq \boldsymbol{\kappa}_i$. Assim, pela definição de uma distribuição slash truncada, temos que $\mathbf{Z}_i | (\mathbf{Z}_i \leq \boldsymbol{\kappa}_i) \sim TSl_p(\boldsymbol{\mu}_z, \boldsymbol{\Sigma}_z, \nu; \mathbb{D}_i)$, em que \mathbb{D}_i é como em (2.15) com $\mathbf{d} = \boldsymbol{\kappa}_i$. Assim,

$$\begin{aligned}\widehat{u}_i &= E[E[U_i | \mathbf{Z}_i] | \mathbf{V}_i] = E[E[U_i | \mathbf{Z}_i] | \mathbf{Z}_i \leq \boldsymbol{\kappa}_i] = E \left[\frac{p + 2\nu}{\delta_i} \frac{P_1(\frac{p}{2} + \nu + 1, \frac{\delta_i}{2})}{P_1(\frac{p}{2} + \nu, \frac{\delta_i}{2})} \middle| \mathbf{Z}_i \leq \boldsymbol{\kappa}_i \right], \\ \widehat{u_i \mathbf{z}_i} &= E[E[U_i | \mathbf{Z}_i] \mathbf{Z}_i | \mathbf{Z}_i \leq \boldsymbol{\kappa}_i] = E \left[\frac{p + 2\nu}{\delta_i} \frac{P_1(\frac{p}{2} + \nu + 1, \frac{\delta_i}{2})}{P_1(\frac{p}{2} + \nu, \frac{\delta_i}{2})} \mathbf{Z}_i \middle| \mathbf{Z}_i \leq \boldsymbol{\kappa}_i \right], \\ \widehat{u_i \mathbf{z}_i^2} &= E[E[U_i | \mathbf{Z}_i] \mathbf{Z}_i \mathbf{Z}_i^\top | \mathbf{Z}_i \leq \boldsymbol{\kappa}_i] = E \left[\frac{p + 2\nu}{\delta_i} \frac{P_1(\frac{p}{2} + \nu + 1, \frac{\delta_i}{2})}{P_1(\frac{p}{2} + \nu, \frac{\delta_i}{2})} \mathbf{Z}_i \mathbf{Z}_i^\top \middle| \mathbf{Z}_i \leq \boldsymbol{\kappa}_i \right].\end{aligned}$$

- iii) O i -ésimo indivíduo tem componentes censurados e não censurados. Particionamos o vetor $\mathbf{Z}_i = \text{vec}(\mathbf{Z}_i^o, \mathbf{Z}_i^c)$. Os componentes são censurados se e somente se $\mathbf{Z}_i^c \leq \boldsymbol{\kappa}_i^c$.

$$\begin{aligned}\widehat{u}_i &= E[U_i | \mathbf{V}_i] = E[E[U_i | \mathbf{Z}_i] | \mathbf{V}_i] = E[E[U_i | \mathbf{Z}_i] | \mathbf{Z}_i^o = \mathbf{z}_i^o, \mathbf{Z}_i^c \leq \boldsymbol{\kappa}_i^c] \\ &= E \left[\frac{p + 2\nu}{\delta_i} \frac{P_1(\frac{p}{2} + \nu + 1, \frac{\delta_i}{2})}{P_1(\frac{p}{2} + \nu, \frac{\delta_i}{2})} \middle| \mathbf{Z}_i^o = \mathbf{z}_i^o, \mathbf{Z}_i^c \leq \boldsymbol{\kappa}_i^c \right] \\ \widehat{u \mathbf{z}_i} &= E[E[U_i \mathbf{Z}_i | \mathbf{Z}_i] | \mathbf{V}_i] = E \left[\text{vec}(E[U_i | \mathbf{Z}_i] \mathbf{Z}_i^o, E[U_i | \mathbf{Z}_i] \mathbf{Z}_i^c) \middle| \mathbf{Z}_i^o = \mathbf{z}_i^o, \mathbf{Z}_i^c \leq \boldsymbol{\kappa}_i^c \right] \\ &= \text{vec} \left(\widehat{u_i \mathbf{z}_i^o}, \widehat{u \mathbf{z}_i^c} \right), \\ \widehat{u \mathbf{z}_i^2} &= E[U_i \mathbf{Z}_i \mathbf{Z}_i^\top | \mathbf{V}_i] = E \left[E[U_i | \mathbf{Z}_i] \begin{bmatrix} \mathbf{Z}_i^o \mathbf{Z}_i^{o\top} & \mathbf{Z}_i^o \mathbf{Z}_i^{c\top} \\ \mathbf{Z}_i^c \mathbf{Z}_i^{o\top} & \mathbf{Z}_i^c \mathbf{Z}_i^{c\top} \end{bmatrix} \middle| \mathbf{Z}_i^o = \mathbf{z}_i^o, \mathbf{Z}_i^c \leq \boldsymbol{\kappa}_i^c \right] \\ &= \begin{bmatrix} \widehat{u_i \mathbf{z}_i^o \mathbf{z}_i^{o\top}} & \mathbf{z}_i^o \widehat{u \mathbf{z}_i^c}^\top \\ \widehat{u \mathbf{z}_i^c \mathbf{z}_i^{o\top}} & \widehat{u \mathbf{z}_i^c \mathbf{z}_i^{c\top}} \end{bmatrix},\end{aligned}$$

em que

$$\begin{aligned}\widehat{u \mathbf{z}_i^c} &= E \left[\frac{p + 2\nu}{\delta_i} \frac{P_1(\frac{p}{2} + \nu + 1, \frac{\delta_i}{2})}{P_1(\frac{p}{2} + \nu, \frac{\delta_i}{2})} \mathbf{Z}_i^c \middle| \mathbf{Z}_i^o = \mathbf{z}_i^o, \mathbf{Z}_i^c \leq \boldsymbol{\kappa}_i^c \right], \text{ e} \\ \widehat{u \mathbf{z}_i^c \mathbf{z}_i^{c\top}} &= E \left[\frac{p + 2\nu}{\delta_i} \frac{P_1(\frac{p}{2} + \nu + 1, \frac{\delta_i}{2})}{P_1(\frac{p}{2} + \nu, \frac{\delta_i}{2})} \mathbf{Z}_i^c \mathbf{Z}_i^{c\top} \middle| \mathbf{Z}_i^o = \mathbf{z}_i^o, \mathbf{Z}_i^c \leq \boldsymbol{\kappa}_i^c \right].\end{aligned}$$

Os restantes valores esperados $\widehat{u x_i}$, $\widehat{u x_i^2}$ e $\widehat{u x_i \mathbf{z}_i}$ são dados nas equações (2.33), (2.36) e (2.37), respectivamente.

3.2.5.1 Matriz hessiana

A matriz hessiana para o *MEMC-Sl* coincide com a matriz hessiana do *MEMC-t*, com a diferença de que os valores esperados \widehat{u}_i , \widehat{uz}_i , \widehat{uz}_i^2 , \widehat{ux}_i , \widehat{ux}_i^2 e \widehat{uxz}_i utilizados no cálculo da matriz hessiana correspondem aos do *MEMC-Sl*.

3.2.5.2 Esquemas de perturbação

Da mesma forma que a matriz hessiana, as matrizes Δ correspondentes aos esquemas de perturbação de ponderação de casos e perturbação na covariável, utilizadas neste trabalho, são equivalentes às do *MEMC-t*, exceto pelos seus valores esperados.

Capítulo 4

Estudo de simulação e aplicação

4.1 Estudo de simulação

Para estudar o desempenho do método proposto (modelo de regressão com erro de medida e dados censurados), apresentamos um estudo de simulação. Ele mostra o comportamento assintótico das estimativas do EM para os modelos propostos sendo que para avaliar o comportamento assintótico utilizamos o viés absoluto e o erro quadrático médio.

4.1.1 Propriedades assintóticas

Apresentamos a seguir alguns resultados de simulações para avaliar a consistência dos estimadores analisando o viés absoluto (*VIES*) e o erro quadrático médio (*EQM*) das estimativas dos coeficientes da regressão obtidas a partir do *MEMC-t* e o *MEMC-SI* para seis diferentes tamanhos de amostra n , tais como 50, 100, 200, 300, 400 e 500. Estas medidas são definidas por

$$VIES_k = \frac{1}{M} \sum_{i=1}^M |\hat{\theta}_k^{(i)} - \theta_k| \quad (4.1)$$

e

$$EQM_k = \frac{1}{M} \sum_{i=1}^M \left(\hat{\theta}_k^{(i)} - \theta_k \right)^2, \quad (4.2)$$

em que $\hat{\theta}_k^{(i)}$ é a estimativa do parâmetro θ_k , $k = 1, \dots, 3p$, para a i -ésima amostra, sendo p a dimensão da distribuição multivariada utilizada. A ideia chave desta simulação é

fornecer evidência empírica sobre a consistência dos estimadores sob os modelos propostos. Para cada tamanho de amostra, geramos $M = 100$ conjuntos de dados dos modelos $MEMC-t$ e $MEMC-Sl$ com 10% de censura, respectivamente. Usando os correspondentes algoritmos EM de cada modelo e as equações (4.1) e (4.2), foi calculado o viés absoluto e o erro quadrático médio para cada parâmetro ao longo dos 100 conjuntos de dados com cada um dos dois modelos. Assumimos os seguintes valores dos parâmetros

$$\begin{aligned} \boldsymbol{\alpha} &= (3, 2, 1, 2)^\top, \quad \boldsymbol{\beta} = (1, 5; 1; 1, 5; 1)^\top, \quad \mu_x = 4, \quad \sigma_x^2 = 2 \quad \text{e} \\ \boldsymbol{\Omega} &= \text{diag}(0, 5; 0, 5; 0, 5; 0, 5; 0, 5), \end{aligned} \quad (4.3)$$

e o valor do grau de liberdade considerado foi $\nu = 5$ nos dois modelos.

Na Figura 4.1, apresentamos um gráfico com os erros quadráticos médios das estimativas dos parâmetros sob o $MEMC-t$ considerando 10% de censura. A partir desta figura, podemos observar que o EQM decresce à medida que o tamanho da amostra aumenta. Resultados similares foram obtidos após a análise do viés absoluto como pode ser visto na Figura 4.2. O mesmo procedimento e análise foi realizado para o $MEMC-Sl$ e os gráficos são apresentados nas Figuras 4.3 e 4.4. Como esperado, os algoritmos ECM e MCECM propostos fornecem estimativas de máxima verossimilhança com boas propriedades assintóticas para os $MEMC-t$ e $MEMC-Sl$, respectivamente.

4.1.2 Estimação das medidas de influência

Neste estudo o $MEMC-N$ da equação (2.8) foi gerado considerando-se 10% de censura e valores dos parâmetros dados por (4.3) para $i = 2, 3, \dots, 49$, $\mathbf{Z}_1 = \boldsymbol{\mu}_z - 5 \times \text{diag}(\boldsymbol{\Sigma}_z)$ e $\mathbf{Z}_{50} = \boldsymbol{\mu}_z + 5 \times \text{diag}(\boldsymbol{\Sigma}_z)$. Esta definição determinou a perturbação sobre os casos #1 e #50, e dessa forma não prejudicou a simetria da distribuição dos erros. O objetivo foi verificar se a metodologia proposta consegue identificar corretamente as observações influentes e se os modelos de caudas pesadas são menos influenciados por estas observações que o $MEMC-N$.

Um estudo de Monte Carlo com 400 réplicas dos modelos propostos foi realizado para avaliar o percentual de réplicas em que as observações contaminadas foram influentes, e calcular a média e o desvio-padrão das medidas de influência.

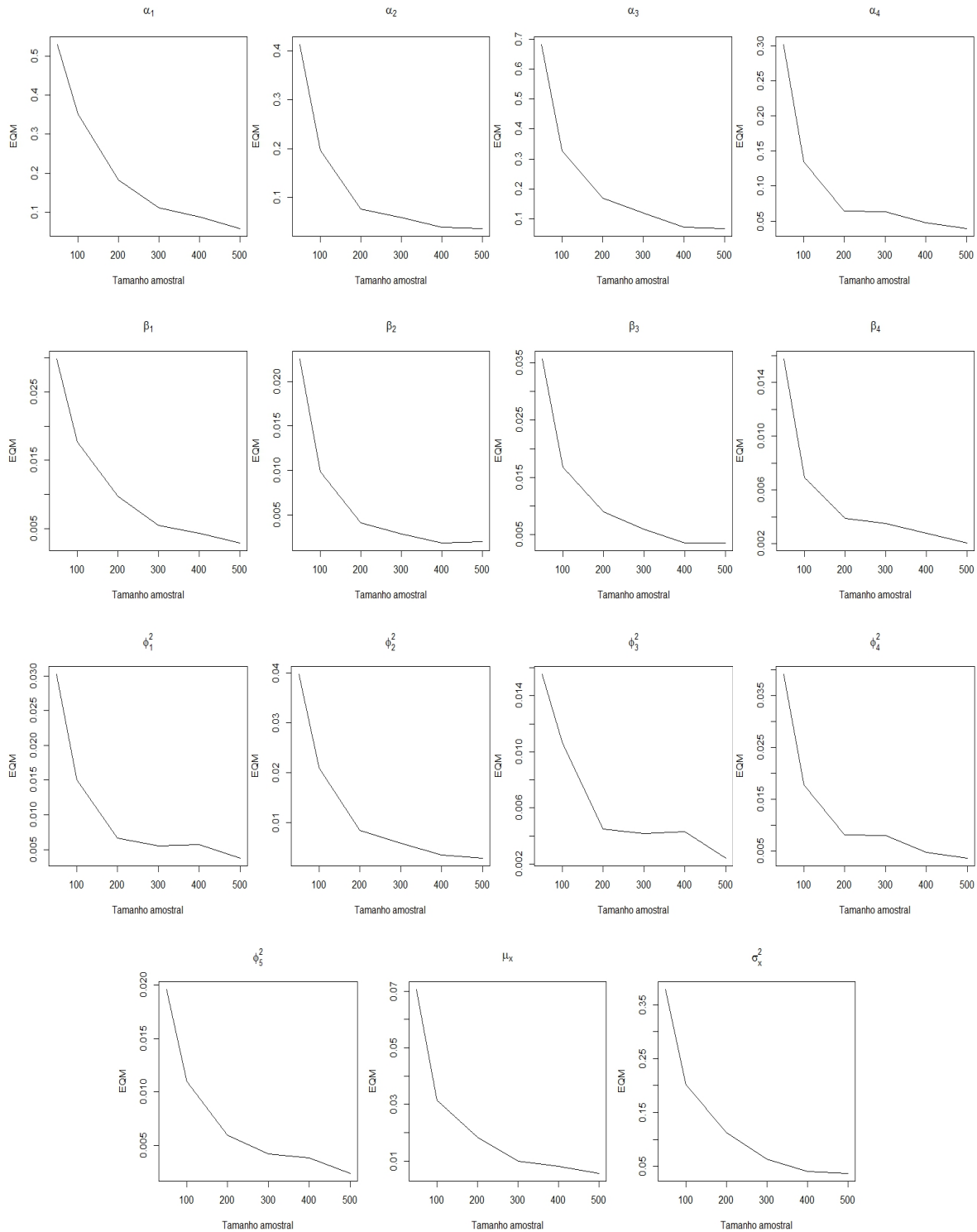


Figura 4.1: Estudo de simulação. *EQM* das estimativas dos parâmetros sob o *MEMC-t* considerando 10% de censura.

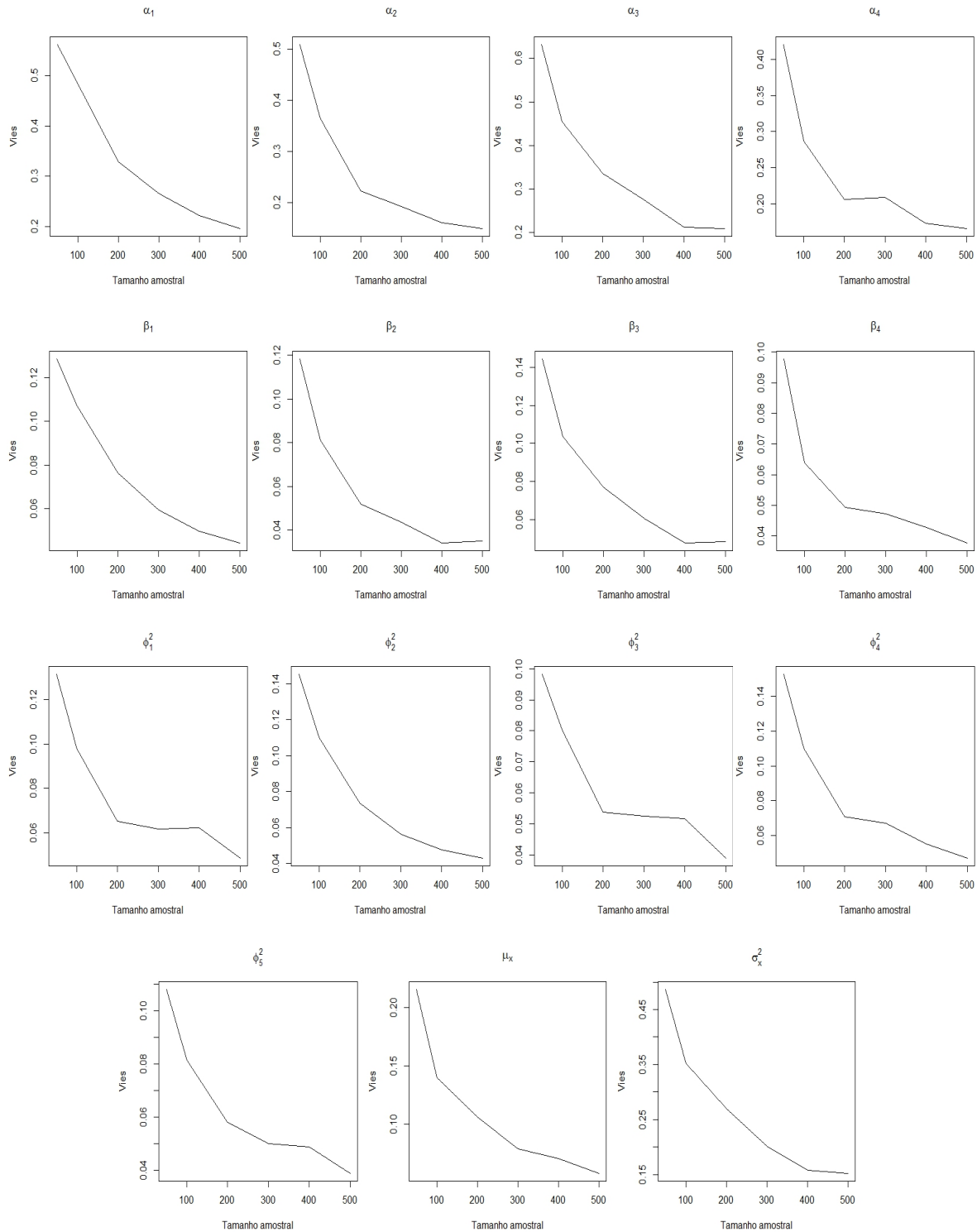


Figura 4.2: Estudo de simulação. *VIES* das estimativas dos parâmetros sob o *MEMC-t* considerando 10% de censura.

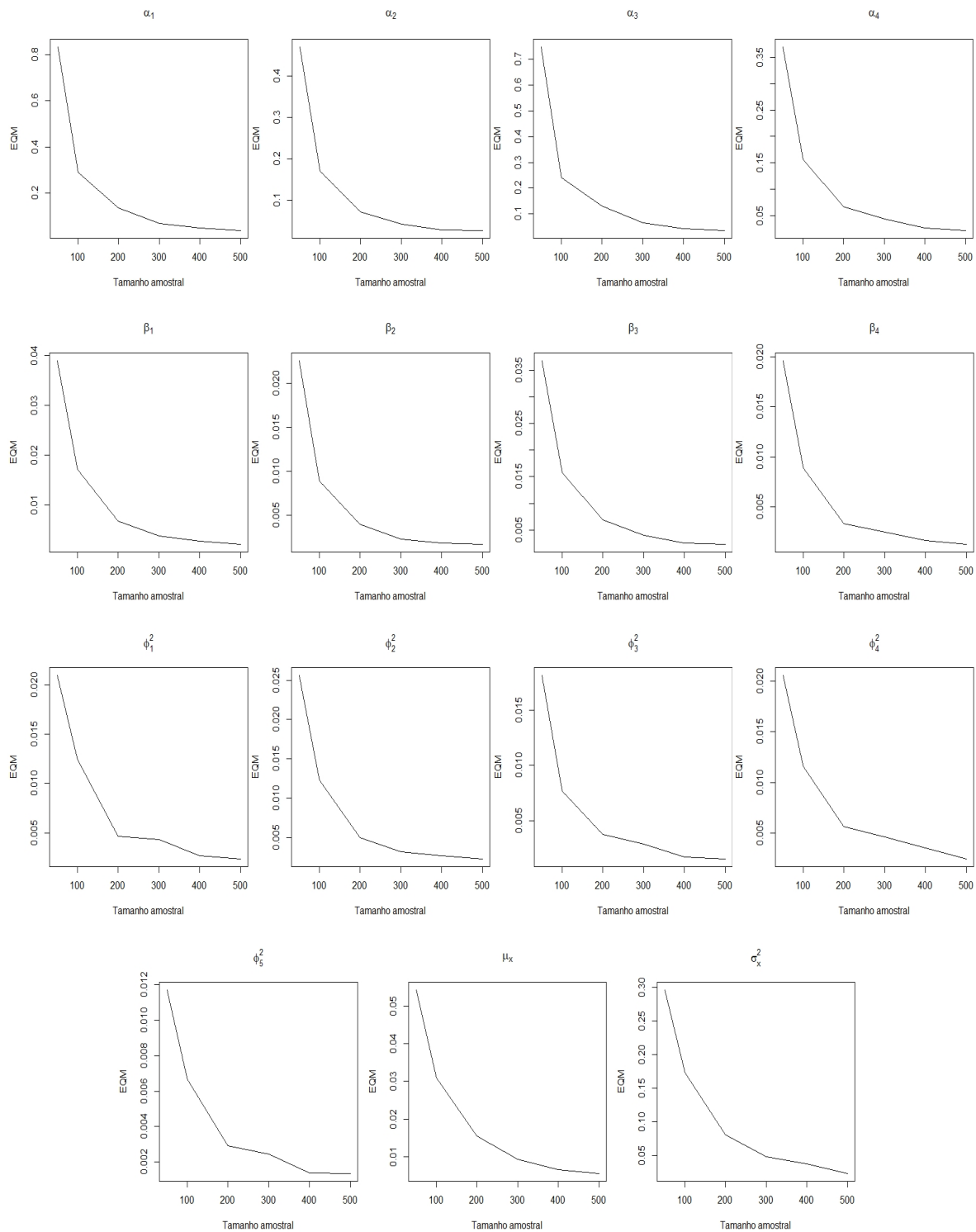


Figura 4.3: Estudo de simulação. *EQM* das estimativas dos parâmetros sob o *MEMC-SI* considerando 10% de censura.

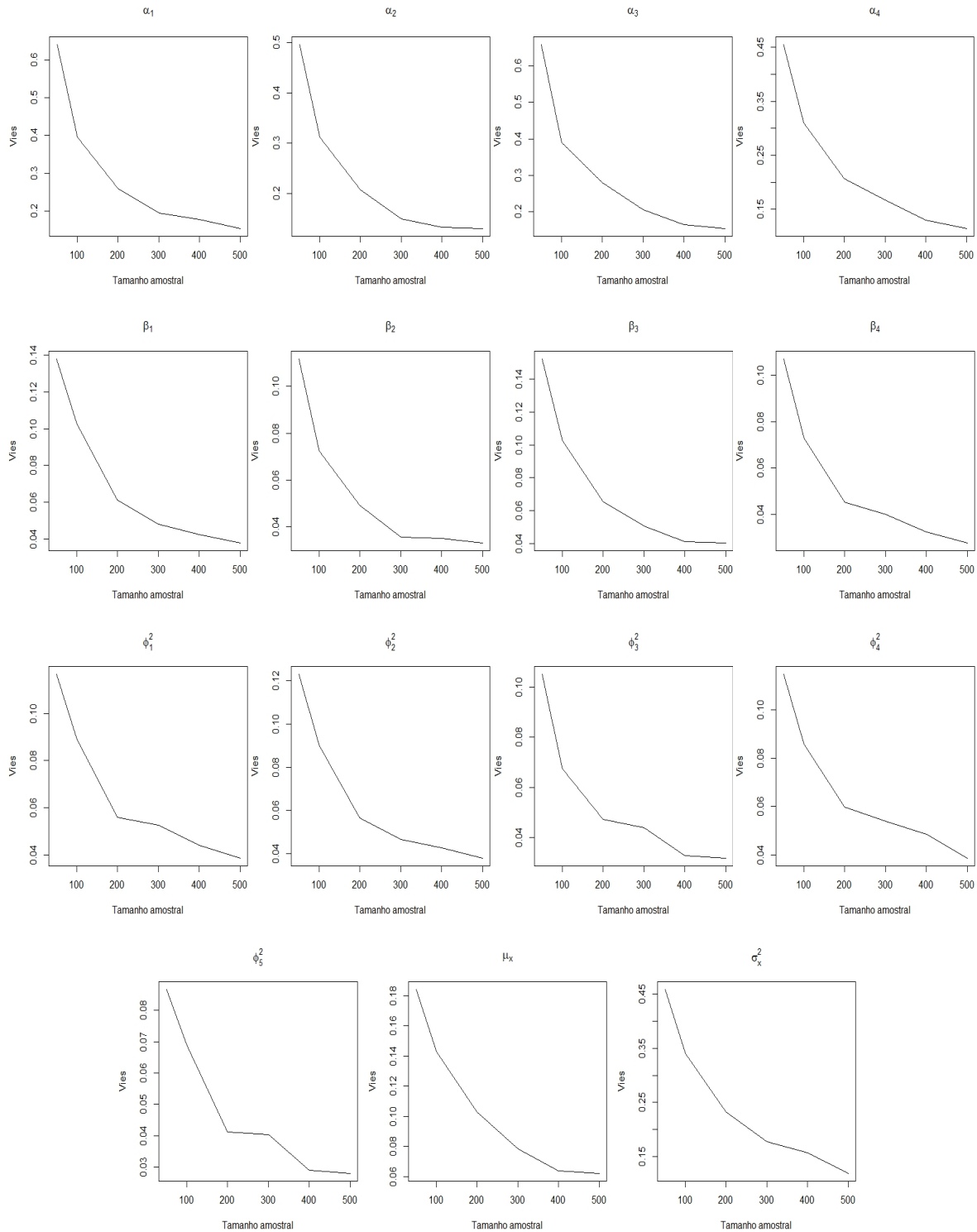


Figura 4.4: Estudo de simulação. *VIES* das estimativas dos parâmetros sob o *MEMC-SI* considerando 10% de censura.

Foi observada uma diferença considerável entre o *MEMC-N* e os modelos *MEMC-t* e *MEMC-SI* para todas as medidas de diagnóstico avaliadas (Tabela 4.1). A observação #50 foi classificada como influente na maioria das réplicas para o esquema de perturbação de ponderação de casos para todos os modelos estudados, sendo que o valor médio das medidas de influência desta observação para os modelos *MEMC-t* e *MEMC-SI* ficaram mais próximos dos valores de referência, enquanto que para o *MEMC-N* as medidas dessa observação são bem maiores que o valor de referência. No caso do esquema de perturbação da covariável esta observação foi influente apenas para o *MEMC-N*. Já no caso do afastamento da função Q , a observação #50 foi classificada como influente em várias réplicas de todos os modelos estudados e na distância generalizada de Cook foi classificada como influente na maioria das réplicas mas somente nos modelos *MEMC-N* e *MEMC-SI*.

A observação #1 não foi identificada como influente em nenhum dos casos. Isso se deve ao fato de que esta observação representa um “outlier” à esquerda, sendo, portanto, censurada em todas as réplicas.

Os resultados deste estudo sugerem que as observações contaminadas exerceram forte influência sobre a estimação construída mediante o *MEMC-N*. No esquema de perturbação da covariável a observação #50 foi influente apenas para o *MEMC-N* e na distância generalizada de Cook não tivemos pontos influentes em nenhuma das réplicas do *MEMC-t*. Conclui-se então que a influência exercida pelas observações contaminadas foi substancialmente menor para os modelos de caudas pesadas.

Tabela 4.1: Estudo de simulação. Análise de influência via estudo de Monte Carlo para as observações #1 e #50 por distribuição e medida de diagnóstico: *AQ* (afastamento da função Q), *DC* (Distância generalizada de Cook), *PPC* (perturbação de ponderação de casos) e *PC* (perturbação da covariável)

Medida de diagnóstico	Estatística	<i>Normal</i>		<i>t-Student</i>		<i>Slash</i>	
		#1	#50	#1	#50	#1	#50
<i>AQ</i>	% Inf ¹	0,0%	100,0%	0,0%	27,8%	0,0%	81,6%
	M ²	0,1111	39,2905	0,1353	0,7258	0,2305	0,9497
	DP ³	0,0096	10,0560	0,0162	0,1099	0,0489	0,1517
	Ref ⁴	0,8000		0,8000		0,8000	
<i>DC</i>	% Inf	0,0%	100,0%	0,0%	0,0%	0,0%	84,6%
	M	0,1060	12,8516	0,1268	0,6780	0,2467	0,9102
	DP	0,0088	1,4929	0,0147	0,0834	0,0410	0,1046
	Ref	0,8000		0,8000		0,8000	
<i>PPC</i>	% Inf	0,0%	100,0%	0,0%	86,6%	0,0%	56,2%
	M	0,0055	0,6605	0,0161	0,0860	0,0258	0,0988
	DP	0,0005	0,0309	0,0018	0,0104	0,0041	0,0121
	M (DP) Ref	0,4371 (0,0196)		0,0773 (0,0079)		0,0976 (0,0057)	
<i>PC</i>	% Inf	0,0%	100,0%	0,0%	0,0%	0,0%	0,0%
	M	0,0235	0,2076	0,0121	0,0029	0,0266	0,0043
	DP	0,0019	0,0179	0,0011	0,0009	0,0019	0,0013
	M (DP) Ref	0,1320 (0,0099)		0,0429 (0,0024)		0,0464 (0,0024)	

¹ Percentual de réplicas de Monte Carlo em que a observação foi considerada influente (maior que o valor de referência).

² Média das medidas de influência.

³ Desvio padrão das medidas de influência.

⁴ Valor de referência para considerar uma observação influente.

4.2 Aplicação

A seguir aplicamos a metodologia desenvolvida nos capítulos anteriores em um conjunto de dados reais. Consideramos o conjunto de dados de Chipkevitch et al. (1996). Os dados consistem de medidas do volume testicular de 42 adolescentes usando cinco técnicas diferentes: ultra-som (US), método gráfico proposto pelos autores (I), medição dimensional (II), orquidômetro de Prader (III), e orquidômetro de anel (IV). A técnica de ultra-som foi assumida como o dispositivo de medida de referência (variável medida com erro). Galea-Rojas et al. (2002) analisaram o mesmo conjunto de dados ajustando o modelo usual com erro de medida normal e recomendaram considerar uma transformação dos dados para obter a normalidade. Lachos et al. (2010) também analisaram este conjunto de dados com o objetivo de proporcionar um melhor ajuste, tentando evitar uma possível transformação desnecessária de dados. De fato, eles consideraram um modelo conjunto da variável latente e erros de observação usando a classe de distribuições de mistura de escala Skew-normal (SMSN). Eles também mostraram evidência do comportamento de caudas pesadas dos dados.

Para aplicar o método proposto nesta tese a este conjunto de dados, censuramos 10% (21 observações) dos dados. Como consequência, o limite de detecção κ_{ij} foi fixado em 4,49 para todo i e j . A Tabela 4.2 mostra os dados do volume testicular com o valor verdadeiro entre parênteses para as observações censuradas. Nós ajustamos os modelos $MEMC-N$, $MEMC-t$ (com $\nu = 6$) e $MEMC-Sl$ (com $\nu = 2$). Para estimar ν usamos o método proposto por Lange et al. (1989). Em particular, nos modelos $MEMC-t$ e $MEMC-Sl$, estimamos os graus de liberdade ν como um problema de seleção de modelos. Ou seja, para uma grade de valores aceitáveis de ν , realizamos a estimação de θ usando o algoritmo EM apropriado. Em seguida, estimamos a função de log-verossimilhança do modelo. Finalmente, escolhemos ν tais que a função de log-verossimilhança é maximizada. A Figura 4.5 apresenta o máximo do logaritmo da função de verossimilhança para vários valores de ν , mostrando que $\nu = 6$ no $MEMC-t$ e $\nu = 2$ no $MEMC-Sl$ são valores razoáveis a considerar.

As estimativas do algoritmo EM para os parâmetros dos três modelos, assim como os

Tabela 4.2: Dados de Chipkevitch *et al.* Dados do volume testicular (em ml) de 42 adolescentes

i	Métodos					i	Métodos				
	US	I	II	III	IV		US	I	II	III	IV
1	5,0	7,5	5,9	8,0	9,0	22	16,5	10,0	15,3	15,0	15,0
2	5,7	5,0	4,8	6,0	10,0	23	4,5	4,5(3,5)	4,5(3,9)	6,0	7,0
3	7,4	5,0	6,8	9,0	12,0	24	5,6	5,0	4,5	4,5	6,0
4	4,5(2,6)	4,5(3,5)	4,5(3,1)	4,5(4,0)	4,5(4,0)	25	11,0	7,5	9,7	9,0	11,0
5	5,7	5,0	5,0	6,0	7,0	26	9,2	10,0	11,3	12,0	13,5
6	6,1	5,0	4,5(4,4)	7,0	8,0	27	8,5	7,5	8,8	12,0	12,0
7	6,2	5,0	6,0	8,0	9,0	28	5,4	5,0	6,1	8,0	8,0
8	10,4	10,0	8,8	10,0	10,0	29	6,7	7,5	7,2	10,0	8,0
9	9,1	7,5	7,9	10,0	11,0	30	5,3	5,0	5,9	8,0	10,0
10	14,8	10,0	13,0	12,0	15,0	31	20,0	20,0	16,3	25,0	22,5
11	16,4	12,5	10,3	17,5	17,5	32	18,8	15,0	16,3	20,0	25,0
12	9,6	7,5	8,2	10,0	11,0	33	13,9	12,5	12,2	15,0	17,5
13	15,7	15,0	19,8	20,0	20,0	34	9,4	10,0	10,3	12,0	13,5
14	4,5(3,0)	4,5(2,0)	4,5(2,0)	4,5(3,0)	4,5(4,0)	35	9,1	7,5	10,8	12,0	12,0
15	16,4	15,0	17,3	20,0	20,0	36	14,1	15,0	13,0	13,5	15,0
16	17,6	15,0	17,3	20,0	22,5	37	9,3	10,0	8,4	10,0	10,0
17	10,0	7,5	7,9	12,0	12,0	38	20,9	20,0	22,1	25,0	25,0
18	4,5(4,1)	4,5(3,5)	4,5(4,4)	4,5(4,0)	6,0	39	11,5	10,0	10,6	15,0	13,5
19	12,7	10,0	11,4	12,0	12,0	40	9,7	10,0	9,7	11,0	12,0
20	4,5(2,7)	4,5(3,5)	4,5(4,1)	4,5(2,5)	6,0	41	13,7	12,5	11,6	17,5	15,0
21	10,2	10,0	11,1	12,0	13,5	42	8,9	10,0	8,1	12,0	12,0

correspondentes erros padrões (*EP*) obtidos através da matriz de informação empírica são reportados na Tabela 4.3. Esta tabela mostra que as estimativas de α , β e ϕ para os modelos *MEMC-N*, *MEMC-t* e *MEMC-Sl* são parecidos. No entanto, os erros padrões dos modelos com caudas pesadas são menores do que o *MEMC-N*, indicando que estes modelos parecem produzir estimativas mais precisas.

Na Tabela 4.4 comparamos o ajuste dos três modelos utilizando os critérios de seleção de modelos *AIC* e *BIC*. Observemos que, como esperado, os modelos com caudas pesadas se ajustam melhor aos dados do que o *MEMC-N*.

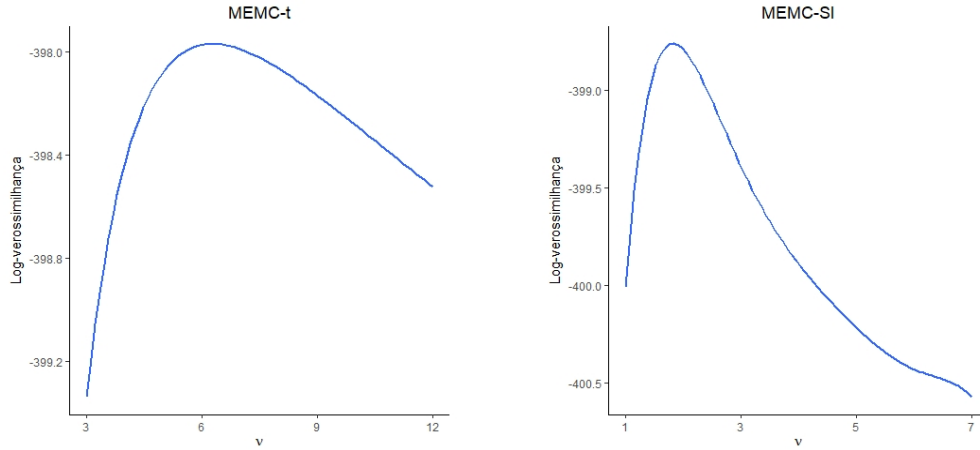


Figura 4.5: Logaritmo da função de verossimilhança do $MEMC-t$ e $MEMC-SI$ para diferentes graus de liberdade.

Tabela 4.3: Dados de Chipkevitch *et al.* EMV e EP para os parâmetros estimados

Parâmetro	$MEMC-N$		$MEMC-t$		$MEMC-SI$	
	Estimativa	EP	Estimativa	EP	Estimativa	EP
α_1	-0,0514	1,6008	-0,0408	1,4159	0,0166	1,4152
α_2	-0,4061	1,6782	-0,6479	1,1468	-0,6350	1,2576
α_3	0,1164	1,4889	0,2821	1,2189	0,2181	1,2707
α_4	1,7903	1,5882	1,8860	1,2858	1,8412	1,3487
β_1	0,8954	0,1520	0,9059	0,1452	0,8986	0,1426
β_2	0,9782	0,1292	1,0199	0,1071	1,0179	0,1097
β_3	1,1372	0,1528	1,1399	0,1343	1,1408	0,1377
β_4	1,0633	0,1608	1,0660	0,1399	1,0663	0,1441
μ_x	9,9295	1,1575	9,1044	0,8972	9,3349	0,9442
σ_x^2	24,9012	12,0535	18,3190	7,1342	13,4630	5,4669
ϕ_1	1,4450	1,0053	1,1051	0,6814	0,7885	0,4890
ϕ_2	1,4317	0,5650	1,1161	0,4459	0,7921	0,2968
ϕ_3	1,9075	0,6711	1,1203	0,3902	0,8192	0,2775
ϕ_4	1,1372	0,5464	0,9396	0,4678	0,6773	0,3193
ϕ_5	1,5399	0,5694	1,1426	0,4274	0,8415	0,2943

Tabela 4.4: Dados de Chipkevitch *et al.* Critérios de comparação dos modelos

	<i>MEMC-N</i>	<i>MEMC-t</i>	<i>MEMC-SI</i>
log-verossimilhança	-400,5370	-397,9718	-398,7905
<i>AIC</i>	831,0740	825,9435	827,5810
<i>BIC</i>	881,2806	876,1501	877,7876

Com o intuito de verificar a existência de possíveis pontos influentes nos dados, que possam influenciar de forma desproporcional o ajuste dos modelos, aplicamos as técnicas de influência global (Q -afastamento e distância de Cook generalizada) e influência local sob os esquemas de perturbação de ponderação de casos e perturbação da covariável (veja o Capítulo 3) ao conjunto de dados utilizado, considerando 4 níveis de censura (10%, 30%, 50% e 70%).

Na análise de influência global, as observações influentes na estimação do vetor de parâmetros θ detectadas pelo afastamento da função Q (veja Figura 4.6) e pela distância de Cook generalizada (Figura 4.7) no *MEMC-N* são os dados #31 e #32, não havendo pontos detectados pelos modelos *MEMC-t* e *MEMC-SI* com todos os percentuais de censura.

Na análise de influência local, as observações possivelmente influentes na estimação do vetor de parâmetros θ detectadas pelo esquema de perturbação de ponderação de casos (Figura 4.8) no *MEMC-N* são as observações #31 e #32, não havendo observações detectadas pelos modelos *MEMC-t* e *MEMC-SI*; já no caso do esquema de perturbação da covariável, é detectada a observação #31 quando considerado 10% de censura, enquanto que nas outras porcentagens de censura é detectada a observação #32 apenas para o *MEMC-N*. Note também que nenhuma observação é detectada no *MEMC-t* em cada uma das porcentagens de censura utilizadas, enquanto que no caso do *MEMC-SI* é detectada a observação #38 quando consideramos 10%, 30% e 50% como porcentagens de censura e não havendo nenhuma observação detectada quando censuramos 70% das observações. As Figuras 4.8 e 4.9 mostram os gráficos para as medidas de influência local, baseadas nas quantidades $M(0)$ sob perturbação de ponderação de casos e perturbação na covariável para os três modelos estudados, e utilizamos o critério $M(0)_j > \overline{M(0)} + 3DP[M(0)]$ para

classificar a j -ésima observação como influente (linha pontilhada).

Finalmente, analisando a influência global e local, podemos observar uma quantidade menor de observações influentes sob os modelos com caudas mais pesadas do que o *MEMC-N*.

Para verificar o impacto de cada observação possivelmente influente foram retiradas uma a uma e conjuntamente do modelo em estudo. Para avaliar a magnitude do impacto exercido pelas observações calculamos as mudanças relativas (MR), em porcentagem, de cada estimativa definido por

$$MR = \left| \frac{\hat{\theta}_j - \hat{\theta}_j^*}{\hat{\theta}_j} \right| \times 100\%, \quad j = 1, \dots, k, \quad (4.4)$$

sendo $\hat{\theta}_j^*$ o estimador do parâmetro θ_j estimado ao retirar uma ou várias (conforme o caso) observações atípicas e k é o número de parâmetros.

Nas Tabelas 4.5–4.7 encontram-se os valores das estimativas de máxima verossimilhança, os respectivos desvios padrões (entre parênteses) e as mudanças relativas das estimativas ao retirar uma a uma e coletivamente as observações apontadas como possivelmente influentes com cada um dos modelos. Note que a estimativa dos parâmetros α_1 , α_2 e α_3 são as que sofrem o maior impacto em todos os casos avaliados; além disso, as maiores variações percentuais ocorrem para as estimativas de quase todos os parâmetros quando a observação #31 é retirada individualmente e quando as observações #31 e #32 são retiradas coletivamente.

Na Tabela 4.8 fazemos uma comparação entre os três modelos, eliminando as observações influentes. Fazendo a comparação dos parâmetros, temos que a eliminação das observações influentes produz maiores mudanças nas estimativas do *MEMC-N* do que nas estimativas dos *MEMC-t* e *MEMC-SI*, sendo ainda melhor no caso do *MEMC-t*; ou seja, como esperado, os modelos com caudas pesadas parecem produzir estimativas mais robustas contra os pontos discrepantes do que o *MEMC-N* pelo fato de ter menores variações nas estimativas dos parâmetros.

Como em Lu e Song (2006), calculamos a medida de mudança relativa total (*MRT*)

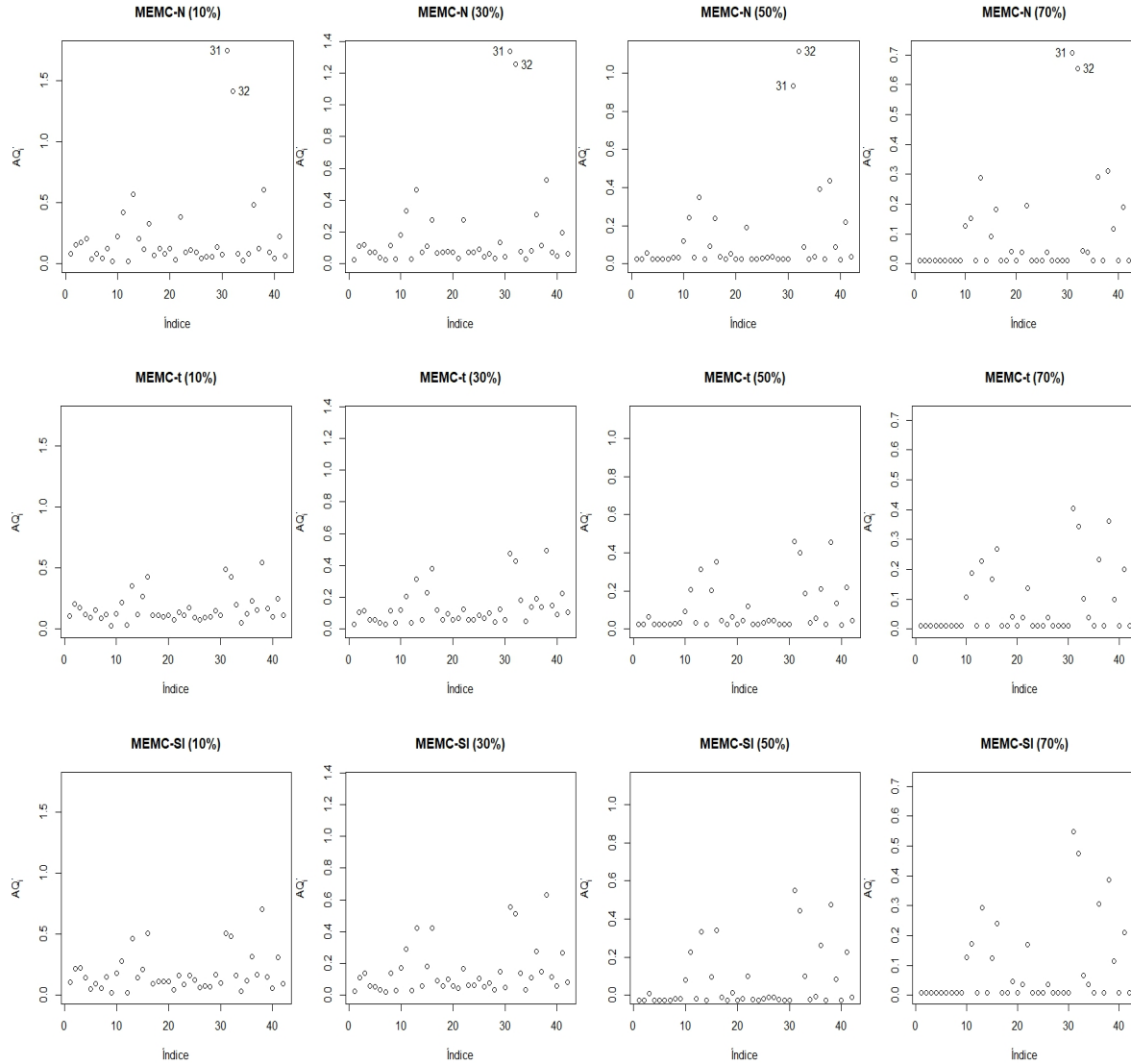


Figura 4.6: Análise global de diagnóstico para os modelos $MEMC-N$, $MEMC-t$ e $MEMC-SI$ segundo o afastamento da função Q considerando 10%, 30%, 50% e 70% de censura.

que é utilizada para medir a diferença entre a EMV original, $\hat{\theta}$, e $\hat{\theta}^*$.

$$MRT = \sum_{j=1}^k \left| \frac{\hat{\theta}_j - \hat{\theta}_j^*}{\hat{\theta}_j} \right|,$$

sendo k o número de parâmetros. Na Tabela 4.9 observamos que as maiores mudanças ocorrem quando consideramos a distribuição normal, isto é, os EMV são menos sensíveis na presença de dados atípicos quando utilizamos distribuições com caudas mais pesadas.

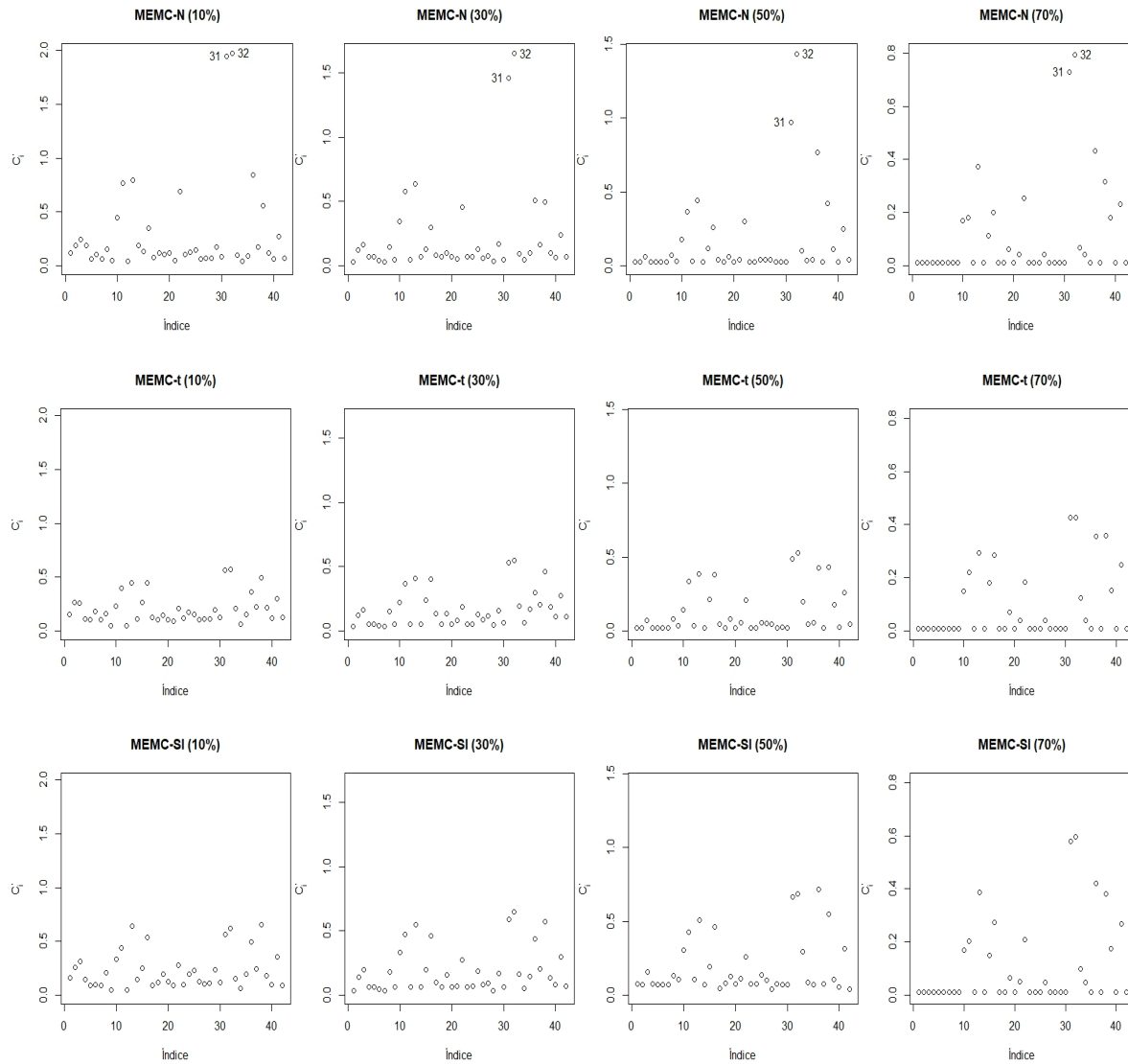


Figura 4.7: Análise global de diagnóstico para os modelos $MEMC-N$, $MEMC-t$ e $MEMC-SI$ segundo a distância de Cook generalizada considerando 10%, 30%, 50% e 70% de censura.

Tabela 4.5: *EMV*, *EP* e *MR* dos parâmetros do *MEMC-N* com a amostra completa e tirando as observações influentes.

Estimativas	Observações eliminadas						
	Nenhuma	31		32		31, 32	
$\hat{\alpha}_1^*$	-0,0514 (1,6008)	0,2309 (1,6252)	549,2%	-0,2120 (1,6972)	312,5%	0,0749 (1,7353)	245,7%
$\hat{\alpha}_2^*$	-0,4061 (1,6782)	-0,7793 (1,6228)	91,9%	-0,5700 (1,7608)	40,3%	-1,0212 (1,7000)	151,4%
$\hat{\alpha}_3^*$	0,1164 (1,4889)	0,3820 (1,5612)	228,3%	-0,0389 (1,5328)	133,4%	0,2397 (1,5930)	106,0%
$\hat{\alpha}_4^*$	1,7903 (1,5882)	1,7299 (1,6428)	3,4%	1,9838 (1,5153)	10,8%	1,9656 (1,5632)	9,8%
$\hat{\beta}_1^*$	0,8954 (0,1520)	0,8633 (0,1545)	3,6%	0,9153 (0,1639)	2,2%	0,8825 (0,1681)	1,4%
$\hat{\beta}_2^*$	0,9782 (0,1292)	1,0219 (0,1265)	4,5%	0,9982 (0,1384)	2,1%	1,0514 (0,1368)	7,5%
$\hat{\beta}_3^*$	1,1372 (0,1528)	1,1064 (0,1615)	2,7%	1,1561 (0,1587)	1,7%	1,1237 (0,1662)	1,2%
$\hat{\beta}_4^*$	1,0633 (0,1608)	1,0707 (0,1648)	0,7%	1,0361 (0,1559)	2,6%	1,0385 (0,1595)	2,3%
$\hat{\mu}_x^*$	9,9295 (1,1575)	9,6886 (1,1237)	2,4%	9,7181 (1,1648)	2,1%	9,4663 (1,1093)	4,7%
$\hat{\sigma}_x^{2*}$	24,9012 (12,0535)	22,8452 (11,4123)	8,3%	23,3940 (11,3502)	6,1%	21,1533 (10,5201)	15,1%
$\hat{\phi}_1^*$	1,4450 (1,0053)	1,4997 (1,1499)	3,8%	1,5399 (1,0505)	6,6%	1,5765 (1,1971)	9,1%
$\hat{\phi}_2^*$	1,4317 (0,5650)	1,3315 (0,5039)	7,0%	1,4601 (0,5450)	2,0%	1,3792 (0,4996)	3,7%
$\hat{\phi}_3^*$	1,9075 (0,6711)	1,5217 (0,5108)	20,2%	1,9193 (0,6696)	0,6%	1,4743 (0,4792)	22,7%
$\hat{\phi}_4^*$	1,1372 (0,5464)	1,0328 (0,4617)	9,2%	1,0840 (0,4697)	4,7%	1,0156 (0,4107)	10,7%
$\hat{\phi}_5^*$	1,5399 (0,5694)	1,5718 (0,5710)	2,1%	0,9967 (0,3830)	35,3%	1,0355 (0,3935)	32,8%

Tabela 4.6: *EMV*, *EP* e *MR* dos parâmetros do *MEMC-t* com a amostra completa e tirando as observações influentes.

Estimativas	Observações eliminadas						
	Nenhuma	31		32		31, 32	
$\hat{\alpha}_1^*$	-0,0408 (1,4159)	0,0688 (1,4572)	268,6%	-0,1483 (1,4976)	263,2%	-0,0394 (1,5472)	3,6%
$\hat{\alpha}_2^*$	-0,6479 (1,1467)	-0,8793 (1,1583)	35,7%	-0,7821 (1,2103)	20,7%	-1,0503 (1,2186)	62,1%
$\hat{\alpha}_3^*$	0,2821 (1,2188)	0,3938 (1,2832)	39,6%	0,1807 (1,2737)	36,0%	0,2989 (1,3406)	5,9%
$\hat{\alpha}_4^*$	1,8860 (1,2858)	1,8559 (1,3336)	1,6%	1,9946 (1,2668)	5,8%	1,9771 (1,3128)	4,8%
$\hat{\beta}_1^*$	0,9059 (0,1452)	0,8929 (0,1500)	1,4%	0,9198 (0,1548)	1,5%	0,9070 (0,1607)	0,1%
$\hat{\beta}_2^*$	1,0199 (0,1071)	1,0497 (0,1106)	2,9%	1,0374 (0,1148)	1,7%	1,0722 (0,1183)	5,1%
$\hat{\beta}_3^*$	1,1399 (0,1343)	1,1258 (0,1415)	1,2%	1,1532 (0,1409)	1,2%	1,1383 (0,1486)	0,1%
$\hat{\beta}_4^*$	1,0660 (0,1399)	1,0706 (0,1448)	0,4%	1,0503 (0,1390)	1,5%	1,0534 (0,1437)	1,2%
$\hat{\mu}_x^*$	9,1044 (0,8972)	8,9706 (0,8669)	1,5%	8,9813 (0,8919)	1,4%	8,8426 (0,8571)	2,9%
$\hat{\sigma}_x^{2*}$	18,3190 (7,1341)	17,2997 (6,8028)	5,6%	17,3249 (6,7104)	5,4%	16,2428 (6,3525)	11,3%
$\hat{\phi}_1^*$	1,1051 (0,6814)	1,1288 (0,7408)	2,1%	1,1477 (0,6953)	3,8%	1,1609 (0,7541)	5,0%
$\hat{\phi}_2^*$	1,1161 (0,4459)	1,0728 (0,4289)	3,9%	1,1319 (0,4507)	1,4%	1,0956 (0,4385)	1,8%
$\hat{\phi}_3^*$	1,1203 (0,3901)	0,9033 (0,3221)	19,4%	1,1104 (0,3877)	0,9%	0,8683 (0,3125)	22,5%
$\hat{\phi}_4^*$	0,9396 (0,4678)	0,8878 (0,4240)	5,5%	0,9167 (0,4426)	2,4%	0,8786 (0,4076)	6,5%
$\hat{\phi}_5^*$	1,1426 (0,4274)	1,1719 (0,4360)	2,6%	0,8794 (0,3358)	23,0%	0,9118 (0,3455)	20,2%

Tabela 4.7: *EMV*, *EP* e *MR* dos parâmetros do *MEMC-Sl* com a amostra completa e tirando as observações influentes.

Estimativas	Observações eliminadas						
	Nenhuma	31		32		31, 32	
$\hat{\alpha}_1^*$	0,0166 (1,4152)	0,1483 (1,4131)	791,0%	-0,0820 (1,4823)	593,0%	0,0642 (1,4833)	285,6%
$\hat{\alpha}_2^*$	-0,6350 (1,2576)	-0,8472 (1,2407)	33,4%	-0,7740 (1,3122)	21,9%	-1,0012 (1,2803)	57,7%
$\hat{\alpha}_3^*$	0,2181 (1,2707)	0,3548 (1,3064)	62,7%	0,1190 (1,3097)	45,4%	0,2958 (1,3389)	35,6%
$\hat{\alpha}_4^*$	1,8412 (1,3487)	1,8217 (1,3708)	1,1%	1,9614 (1,3098)	6,5%	1,9696 (1,3281)	7,0%
$\hat{\beta}_1^*$	0,8986 (0,1426)	0,8851 (0,1429)	1,5%	0,9122 (0,1509)	1,5%	0,8974 (0,1519)	0,1%
$\hat{\beta}_2^*$	1,0179 (0,1097)	1,0454 (0,1113)	2,7%	1,0353 (0,1162)	1,7%	1,0668 (0,1171)	4,8%
$\hat{\beta}_3^*$	1,1408 (0,1377)	1,1259 (0,1421)	1,3%	1,1532 (0,1426)	1,1%	1,1363 (0,1470)	0,4%
$\hat{\beta}_4^*$	1,0663 (0,1441)	1,0708 (0,1464)	0,4%	1,0492 (0,1413)	1,6%	1,0520 (0,1436)	1,3%
$\hat{\mu}_x^*$	9,3349 (0,9442)	9,1802 (0,9166)	1,7%	9,2074 (0,9551)	1,4%	9,0371 (0,9236)	3,2%
$\hat{\sigma}_x^{2*}$	13,4630 (5,4669)	12,7967 (5,2276)	4,9%	12,9291 (5,2296)	4,0%	12,2680 (4,9754)	8,9%
$\hat{\phi}_1^*$	0,7885 (0,4890)	0,8034 (0,5320)	1,9%	0,8231 (0,5066)	4,4%	0,8264 (0,5467)	4,8%
$\hat{\phi}_2^*$	0,7921 (0,2968)	0,7637 (0,2838)	3,6%	0,8112 (0,3005)	2,4%	0,7880 (0,2925)	0,5%
$\hat{\phi}_3^*$	0,8192 (0,2775)	0,6528 (0,2293)	20,3%	0,8133 (0,2761)	0,7%	0,6233 (0,2223)	23,9%
$\hat{\phi}_4^*$	0,6773 (0,3193)	0,6499 (0,2963)	4,0%	0,6678 (0,3022)	1,4%	0,6522 (0,2854)	3,7%
$\hat{\phi}_5^*$	0,8415 (0,2943)	0,8648 (0,2997)	2,8%	0,6651 (0,2352)	21,0%	0,6924 (0,2421)	17,7%

Tabela 4.8: Mudanças (em %) nas estimativas dos parâmetros dos modelos ajustados depois de excluídas as observações 31 e 32.

Parâmetro	<i>MEMC-N</i>	<i>MEMC-t</i>	<i>MEMC-Sl</i>
α_1	245,7%	3,6%	285,6%
α_2	151,4%	62,1%	57,7%
α_3	106,0%	5,9%	35,6%
α_4	9,8%	4,8%	7,0%
β_1	1,4%	0,1%	0,1%
β_2	7,5%	5,1%	4,8%
β_3	1,2%	0,1%	0,4%
β_4	2,3%	1,2%	1,3%
μ_x	4,7%	2,9%	3,2%
σ_x^2	15,1%	11,3%	8,9%
ϕ_1	9,1%	5,0%	4,8%
ϕ_2	3,7%	1,8%	0,5%
ϕ_3	22,7%	22,5%	23,9%
ϕ_4	10,7%	6,5%	3,7%
ϕ_5	32,8%	20,2%	17,7%

Tabela 4.9: Dados de Chipkevitch *et al.* Comparação das mudanças relativas nos *EMV* segundo a *MRT* para os três modelos considerados.

	<i>MRT</i>
<i>MEMC-N</i>	5,9647
<i>MEMC-t</i>	1,6931
<i>MEMC-Sl</i>	4,4531

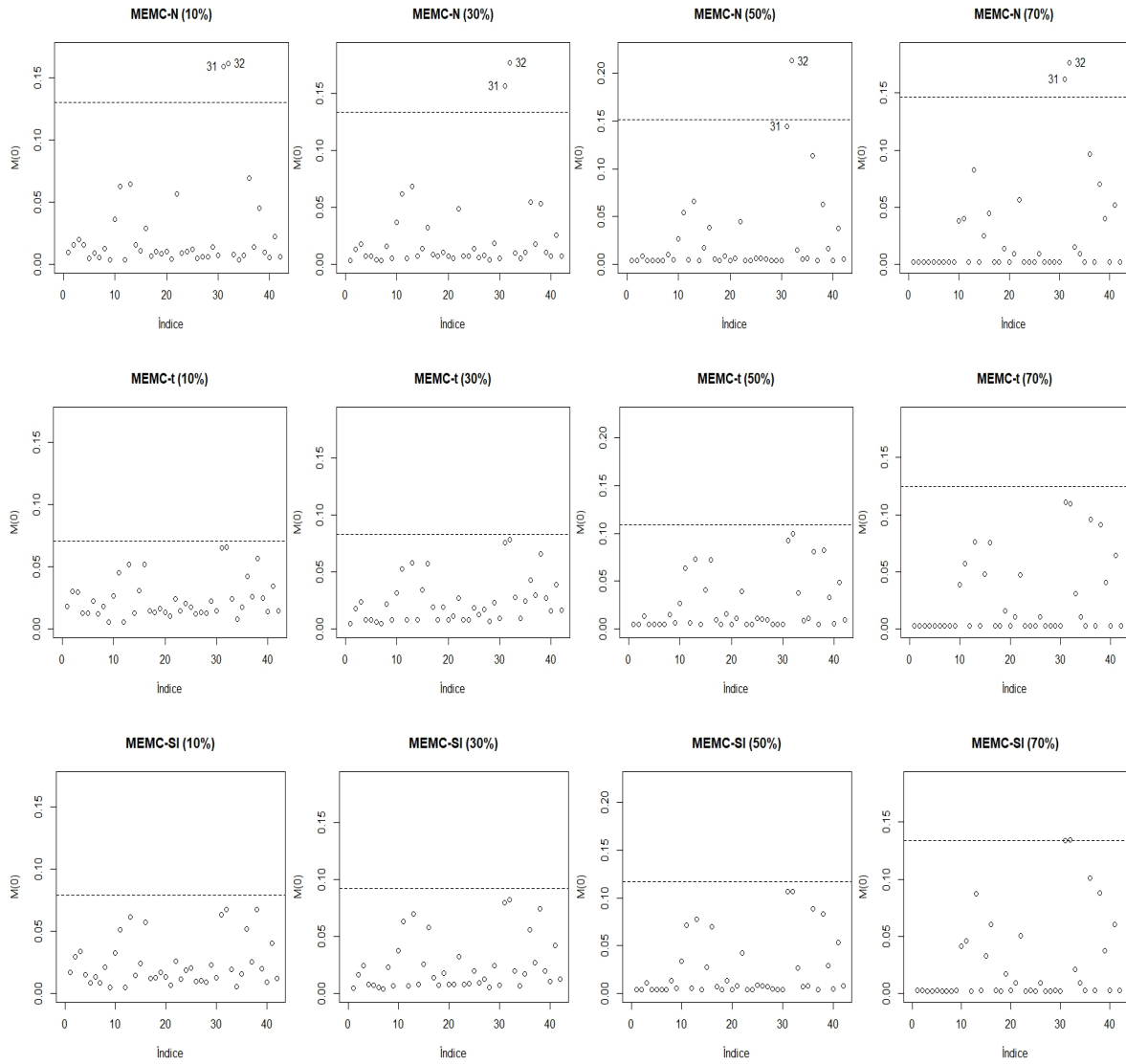


Figura 4.8: Gráficos de índices de $M(0)$ para os modelos $MEMC-N$, $MEMC-t$ e $MEMC-SI$ segundo o esquema de perturbação de ponderação de casos considerando 10%, 30%, 50% e 70% de censura.

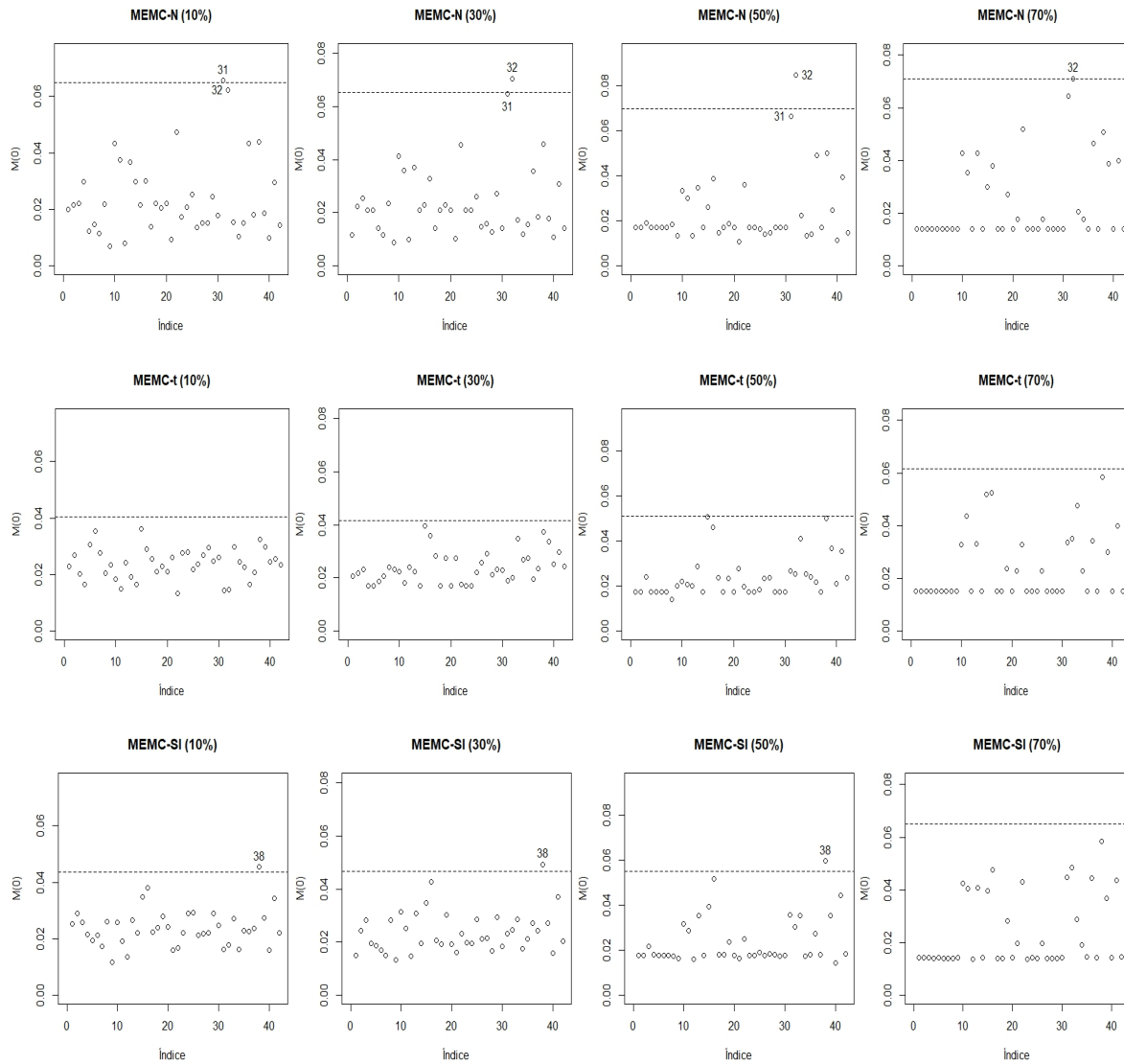


Figura 4.9: Gráficos de índices de $M(0)$ para os modelos $MEMC-N$, $MEMC-t$ e $MEMC-SI$ segundo o esquema de perturbação da covariável considerando 10%, 30%, 50% e 70% de censura.

Capítulo 5

Considerações finais

Neste trabalho foram estudados três modelos com erros de medida e respostas censuradas. Para cada modelo apresentado desenvolvemos estimação por máxima verossimilhança e implementamos o algoritmo EM, fazendo uso de alguma de suas extensões, para estimar os parâmetros de regressão dos três modelos estudados. Para validar as suposições e identificar observações influentes dos modelos de regressão, desenvolvemos métodos de diagnóstico. Para isto, desenvolvemos medidas de influência global e local baseadas na curvatura normal conforme e derivamos expressões matriciais sob o esquema de perturbação de ponderação de casos e perturbação na covariável. Em um exemplo com dados reais ilustramos a potencialidade da metodologia de influência local no sentido de detectar as observações que podem afetar os resultados inferenciais obtidos para o modelo ajustado aos dados.

5.1 Perspectivas futuras

Como foco de trabalhos futuros sugerimos os seguintes temas de pesquisa:

1. utilizar o enfoque funcional para contornar a presença de erros de medida;
2. estudar a influência local fazendo uso de outros tipos de perturbação;
3. desenvolver metodologias de inferência sob o enfoque Bayesiano para o modelo proposto nesta tese.

Apêndice

Apêndice A: Provas das proposições.

A.1 Proposição 1: Considere a representação hierárquica do *MEMC-NI* dado em (2.11)-(2.13). Então,

$$x_i|U_i = u_i, \mathbf{Z}_i = \mathbf{z}_i \sim N\left(\frac{\mu_x + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1}(\mathbf{z}_i - \mathbf{a})}{1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b}}, \frac{\sigma_x^2}{u_i(1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b})}\right).$$

Prova:

Sabemos que

$$\begin{aligned} \mathbf{Z}_i|x_i, U_i = u_i &\stackrel{\text{ind}}{\sim} N_p(\mathbf{a} + \mathbf{b}x_i, u_i^{-1}\boldsymbol{\Omega}), \\ x_i|U_i = u_i &\stackrel{\text{ind}}{\sim} N(\mu_x, u_i^{-1}\sigma_x^2), \\ U_i &\stackrel{\text{iid}}{\sim} H(u_i; \nu), \quad i = 1, \dots, n. \end{aligned}$$

$$\begin{aligned} f(x_i|u_i, \mathbf{z}_i) &= \frac{f(x_i, u_i, \mathbf{z}_i)}{f(u_i, \mathbf{z}_i)} = \frac{f(u_i)f(x_i|u_i)f(\mathbf{z}_i|x_i, u_i)}{f(u_i)f(\mathbf{z}_i|u_i)} \propto f(\mathbf{z}_i|x_i, u_i)f(x_i|u_i) \\ &\propto e^{-\frac{1}{2}(\mathbf{z}_i - \mathbf{a} - \mathbf{b}x_i)^\top (u_i^{-1}\boldsymbol{\Omega})^{-1}(\mathbf{z}_i - \mathbf{a} - \mathbf{b}x_i)} e^{-\frac{1}{2}\frac{(x_i - \mu_x)^2}{u_i^{-1}\sigma_x^2}} \\ &= e^{-\frac{u_i}{2} \overbrace{\left[(\mathbf{z}_i - \mathbf{a} - \mathbf{b}x_i)^\top \boldsymbol{\Omega}^{-1}(\mathbf{z}_i - \mathbf{a} - \mathbf{b}x_i) + \frac{(x_i - \mu_x)^2}{\sigma_x^2} \right]}^*}, \end{aligned}$$

em que

$$\boldsymbol{\Omega} = \begin{bmatrix} \phi_1^2 & 0 & \cdots & 0 \\ 0 & \phi_2^2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \phi_p^2 \end{bmatrix} = \begin{bmatrix} \phi_1^2 & \mathbf{0} \\ \mathbf{0}^\top & \boldsymbol{\Omega}_{22} \end{bmatrix}, \quad \text{então } \boldsymbol{\Omega}^{-1} = \begin{bmatrix} \frac{1}{\phi_1^2} & \mathbf{0} \\ \mathbf{0}^\top & \boldsymbol{\Omega}_{22}^{-1} \end{bmatrix} \text{ e}$$

$$\mathbf{Z}_i - \mathbf{a} - \mathbf{b}x_i = \begin{bmatrix} X_i \\ \mathbf{Y}_i \end{bmatrix} - \begin{bmatrix} 0 \\ \boldsymbol{\alpha} \end{bmatrix} - \begin{bmatrix} 1 \\ \boldsymbol{\beta} \end{bmatrix} x_i = \begin{bmatrix} X_i - x_i \\ \mathbf{Y}_i - \boldsymbol{\alpha} - \boldsymbol{\beta}x_i \end{bmatrix}.$$

Assim, temos que

$$\begin{aligned}
& (\mathbf{Z}_i - \mathbf{a} - \mathbf{b}x_i)^\top \boldsymbol{\Omega}^{-1} (\mathbf{Z}_i - \mathbf{a} - \mathbf{b}x_i) \\
&= [X_i - x_i, (\mathbf{Y}_i - \boldsymbol{\alpha} - \boldsymbol{\beta}x_i)^\top] \begin{bmatrix} \frac{1}{\phi_1^2} & \mathbf{0} \\ \mathbf{0}^\top & \boldsymbol{\Omega}_{22}^{-1} \end{bmatrix} \begin{bmatrix} X_i - x_i \\ \mathbf{Y}_i - \boldsymbol{\alpha} - \boldsymbol{\beta}x_i \end{bmatrix} \\
&= \frac{(X_i - x_i)^2}{\phi_1^2} + (\mathbf{Y}_i - \boldsymbol{\alpha} - \boldsymbol{\beta}x_i)^\top \boldsymbol{\Omega}_{22}^{-1} (\mathbf{Y}_i - \boldsymbol{\alpha} - \boldsymbol{\beta}x_i),
\end{aligned}$$

e

$$\begin{aligned}
* &= \frac{(X_i - x_i)^2}{\phi_1^2} + (\mathbf{Y}_i - \boldsymbol{\alpha} - \boldsymbol{\beta}x_i)^\top \boldsymbol{\Omega}_{22}^{-1} (\mathbf{Y}_i - \boldsymbol{\alpha} - \boldsymbol{\beta}x_i) + \frac{(x_i - \mu_x)^2}{\sigma_x^2} \\
&= \frac{X_i^2 - 2x_iX_i + x_i^2}{\phi_1^2} + \frac{x_i^2 - 2x_i\mu_x + \mu_x^2}{\sigma_x^2} + \mathbf{Y}_i^\top \boldsymbol{\Omega}_{22}^{-1} \mathbf{Y}_i - \mathbf{Y}_i^\top \boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\alpha} - \mathbf{Y}_i^\top \boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\beta}x_i \\
&\quad - \boldsymbol{\alpha}^\top \boldsymbol{\Omega}_{22}^{-1} \mathbf{Y}_i + \boldsymbol{\alpha}^\top \boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\alpha} + \boldsymbol{\alpha}^\top \boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\beta}x_i - x_i \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} \mathbf{Y}_i + x_i \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\alpha} + x_i \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\beta}x_i \\
&= \frac{x_i^2 - 2x_iX_i}{\phi_1^2} + \frac{x_i^2 - 2x_i\mu_x}{\sigma_x^2} - 2x_i \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} \mathbf{Y}_i + 2x_i \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\alpha} + x_i^2 \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\beta} + C \\
&\propto \frac{x_i^2 \sigma_x^2 - 2x_i X_i \sigma_x^2 + x_i^2 \phi_1^2 - 2x_i \mu_x \phi_1^2 - 2x_i \phi_1^2 \sigma_x^2 \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} \mathbf{Y}_i + 2x_i \phi_1^2 \sigma_x^2 \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\alpha} + x_i^2 \phi_1^2 \sigma_x^2 \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\beta}}{\phi_1^2 \sigma_x^2} \\
&= x_i^2 \left(\frac{\sigma_x^2 + \phi_1^2 + \phi_1^2 \sigma_x^2 \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\beta}}{\phi_1^2 \sigma_x^2} \right) - 2x_i \left(\frac{X_i \sigma_x^2 + \mu_x \phi_1^2 + \phi_1^2 \sigma_x^2 \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} \mathbf{Y}_i - \phi_1^2 \sigma_x^2 \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\alpha}}{\phi_1^2 \sigma_x^2} \right) \\
&= \frac{\sigma_x^2 + \phi_1^2 + \phi_1^2 \sigma_x^2 \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\beta}}{\phi_1^2 \sigma_x^2} \left[x_i^2 - 2x_i \left(\frac{X_i \sigma_x^2 + \mu_x \phi_1^2 + \phi_1^2 \sigma_x^2 \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} \mathbf{Y}_i - \phi_1^2 \sigma_x^2 \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\alpha}}{\sigma_x^2 + \phi_1^2 + \phi_1^2 \sigma_x^2 \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\beta}} \right) \right. \\
&\quad \left. \pm \left(\frac{X_i \sigma_x^2 + \mu_x \phi_1^2 + \phi_1^2 \sigma_x^2 \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} \mathbf{Y}_i - \phi_1^2 \sigma_x^2 \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\alpha}}{\sigma_x^2 + \phi_1^2 + \phi_1^2 \sigma_x^2 \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\beta}} \right)^2 \right] \\
&= \underbrace{\frac{\sigma_x^2 + \phi_1^2 + \phi_1^2 \sigma_x^2 \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\beta}}{\phi_1^2 \sigma_x^2}}_I \left(x_i - \underbrace{\frac{X_i \sigma_x^2 + \mu_x \phi_1^2 + \phi_1^2 \sigma_x^2 \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} \mathbf{Y}_i - \phi_1^2 \sigma_x^2 \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\alpha}}{\sigma_x^2 + \phi_1^2 + \phi_1^2 \sigma_x^2 \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\beta}}}_{II} \right)^2 \\
I &= \frac{\sigma_x^2 + \phi_1^2 + \phi_1^2 \sigma_x^2 \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\beta}}{\phi_1^2 \sigma_x^2} = \frac{1}{\phi_1^2} + \frac{1}{\sigma_x^2} + \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\beta} = \frac{1}{\sigma_x^2} + \left[\frac{1}{\phi_1^2}, \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} \right] \begin{bmatrix} 1 \\ \boldsymbol{\beta} \end{bmatrix} \\
&= \frac{1}{\sigma_x^2} + [1, \boldsymbol{\beta}^\top] \begin{bmatrix} \frac{1}{\phi_1^2} & \mathbf{0} \\ \mathbf{0}^\top & \boldsymbol{\Omega}_{22}^{-1} \end{bmatrix} \begin{bmatrix} 1 \\ \boldsymbol{\beta} \end{bmatrix} = \frac{1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b}}{\sigma_x^2}
\end{aligned}$$

$$\begin{aligned}
II &= \frac{\frac{1}{\phi_1^2} (X_i \sigma_x^2 + \mu_x \phi_1^2 + \phi_1^2 \sigma_x^2 \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} \mathbf{Y}_i - \phi_1^2 \sigma_x^2 \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\alpha})}{\frac{1}{\phi_1^2} (\sigma_x^2 + \phi_1^2 + \phi_1^2 \sigma_x^2 \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\beta})} \\
&= \frac{\frac{X_i}{\phi_1^2} \sigma_x^2 + \mu_x + \sigma_x^2 \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} \mathbf{Y}_i - \sigma_x^2 \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\alpha}}{1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b}} = \frac{\mu_x + \sigma_x^2 \left(\frac{X_i}{\phi_1^2} + \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} \mathbf{Y}_i - \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\alpha} \right)}{1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b}} \\
&= \frac{\mu_x + \sigma_x^2 \left(\frac{X_i}{\phi_1^2} + \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} (\mathbf{Y}_i - \boldsymbol{\alpha}) \right)}{1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b}} = \frac{\mu_x + \sigma_x^2 \left(\begin{bmatrix} \frac{1}{\phi_1^2}, \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} \end{bmatrix} \begin{bmatrix} X_i \\ \mathbf{Y}_i - \boldsymbol{\alpha} \end{bmatrix} \right)}{1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b}} \\
&= \frac{\mu_x + \sigma_x^2 [1, \boldsymbol{\beta}^\top] \begin{bmatrix} \frac{1}{\phi_1^2} & \mathbf{0} \\ \mathbf{0}^\top & \boldsymbol{\Omega}_{22}^{-1} \end{bmatrix} \left(\begin{bmatrix} X_i \\ \mathbf{Y}_i \end{bmatrix} - \begin{bmatrix} 0 \\ \boldsymbol{\alpha} \end{bmatrix} \right)}{1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b}} = \frac{\mu_x + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} (\mathbf{Z}_i - \mathbf{a})}{1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b}}
\end{aligned}$$

Substituindo (I) e (II), temos que

$$\begin{aligned}
* &= \frac{1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b}}{\sigma_x^2} \left(x_i - \frac{\mu_x + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} (\mathbf{Z}_i - \mathbf{a})}{1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b}} \right)^2 + C \\
f(x_i | u_i, \mathbf{z}_i) &\propto e^{-\frac{u_i}{2} \left[\frac{1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b}}{\sigma_x^2} \left(x_i - \frac{\mu_x + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} (\mathbf{Z}_i - \mathbf{a})}{1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b}} \right)^2 \right]}
\end{aligned}$$

Portanto,

$$x_i | U_i = u_i, \mathbf{Z}_i = \mathbf{z}_i \sim N \left(\frac{\mu_x + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} (\mathbf{Z}_i - \mathbf{a})}{1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b}}, \frac{\sigma_x^2}{u_i (1 + \sigma_x^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b})} \right)$$

Apêndice B: Cálculo das estimativas no algoritmo ECM e dos elementos da matriz de informação observada.

B.1 Cálculo das estimativas no passo CM

Sabemos que

$$Q_1(\boldsymbol{\alpha}, \boldsymbol{\beta}, \phi | \widehat{\boldsymbol{\theta}}^{(k)}) = -\frac{n}{2} \sum_{j=1}^p \log \phi_j^2 - \frac{1}{2} \sum_{i=1}^n \left[\text{tr} \left(\boldsymbol{\Omega}^{-1} \widehat{u} \widehat{\mathbf{z}}_i^2 \right) - 2\mathbf{a}^\top \boldsymbol{\Omega}^{-1} \widehat{u} \widehat{\mathbf{z}}_i - 2\widehat{u} \widehat{\mathbf{x}}_i \boldsymbol{\Omega}^{-1} \mathbf{b} \right. \\ \left. + \mathbf{a}^\top \boldsymbol{\Omega}^{-1} \mathbf{a} \widehat{u}_i + 2\mathbf{a}^\top \boldsymbol{\Omega}^{-1} \mathbf{b} \widehat{u} \widehat{x}_i + \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b} \widehat{u} \widehat{x}_i^2 \right],$$

$$Q_2(\mu_x, \sigma_x^2 | \widehat{\boldsymbol{\theta}}^{(k)}) = -\frac{n}{2} \log \sigma_x^2 - \frac{1}{2\sigma_x^2} \sum_{i=1}^n \left(\widehat{u} \widehat{x}_i^2 - 2\mu_x \widehat{u} \widehat{x}_i + \mu_x^2 \widehat{u}_i \right),$$

i)

$$\frac{\partial Q_1}{\partial \boldsymbol{\alpha}^\top} = \frac{\partial Q_1}{\partial \mathbf{a}^\top} \frac{\partial \mathbf{a}}{\partial \boldsymbol{\alpha}^\top} = -\frac{1}{2} \sum_{i=1}^n \left(-2\widehat{u} \widehat{\mathbf{z}}_i^\top \boldsymbol{\Omega}^{-1} + 2\mathbf{a}^\top \boldsymbol{\Omega}^{-1} \widehat{u}_i + 2\widehat{u} \widehat{x}_i \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \right) \begin{bmatrix} \mathbf{0} \\ \mathbb{I}_r \end{bmatrix} = \mathbf{0}$$

$$\left(\sum_{i=1}^n \widehat{u} \widehat{\mathbf{z}}_i^\top \boldsymbol{\Omega}^{-1} - \sum_{i=1}^n \widehat{u}_i \mathbf{a}^\top \boldsymbol{\Omega}^{-1} - \sum_{i=1}^n \widehat{u} \widehat{x}_i \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \right) \begin{bmatrix} \mathbf{0} \\ \mathbb{I}_r \end{bmatrix} = \mathbf{0}$$

$$\left[\sum_{i=1}^n \widehat{u} \widehat{\mathbf{z}}_{i1}, \sum_{i=1}^n \widehat{u} \widehat{\mathbf{z}}_i^{*\top} \right] \begin{bmatrix} \frac{1}{\phi_1^2} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Omega}_{22}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ \mathbb{I}_r \end{bmatrix} - \left[\sum_{i=1}^n \widehat{u} \widehat{x}_{i1}, \sum_{i=1}^n \widehat{u} \widehat{\mathbf{z}}_i^{*\top} \right] \begin{bmatrix} \frac{1}{\phi_1^2} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Omega}_{22}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ \mathbb{I}_r \end{bmatrix} \\ - \sum_{i=1}^n \widehat{u} \widehat{x}_i [1, \boldsymbol{\beta}^\top] \begin{bmatrix} \frac{1}{\phi_1^2} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Omega}_{22}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ \mathbb{I}_r \end{bmatrix} = \mathbf{0}$$

$$\left[\frac{\sum_{i=1}^n \widehat{u} \widehat{x}_{i1}}{\phi_1^2}, \sum_{i=1}^n \widehat{u} \widehat{\mathbf{z}}_i^{*\top} \boldsymbol{\Omega}_{22}^{-1} \right] \begin{bmatrix} \mathbf{0} \\ \mathbb{I}_r \end{bmatrix} - \left[\frac{\sum_{i=1}^n \widehat{u} \widehat{x}_i}{\phi_1^2}, \sum_{i=1}^n \widehat{u} \widehat{x}_i \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} \right] \begin{bmatrix} \mathbf{0} \\ \mathbb{I}_r \end{bmatrix} \\ = \begin{bmatrix} 0, \sum_{i=1}^n \widehat{u}_i \boldsymbol{\alpha}^\top \boldsymbol{\Omega}_{22}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ \mathbb{I}_r \end{bmatrix}$$

$$\sum_{i=1}^n \widehat{u} \widehat{\mathbf{z}}_i^{*\top} \boldsymbol{\Omega}_{22}^{-1} - \sum_{i=1}^n \widehat{u} \widehat{x}_i \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} = \sum_{i=1}^n \widehat{u}_i \boldsymbol{\alpha}^\top \boldsymbol{\Omega}_{22}^{-1} \Rightarrow \boldsymbol{\alpha}^\top \sum_{i=1}^n \widehat{u}_i = \sum_{i=1}^n \widehat{u} \widehat{\mathbf{z}}_i^{*\top} - \sum_{i=1}^n \widehat{u} \widehat{x}_i \boldsymbol{\beta}^\top$$

$$\Rightarrow \boldsymbol{\alpha} = \frac{\sum_{i=1}^n \widehat{u} \widehat{\mathbf{z}}_i^* - \sum_{i=1}^n \widehat{u} \widehat{x}_i \boldsymbol{\beta}}{\sum_{i=1}^n \widehat{u}_i}$$

$$\therefore \boxed{\widehat{\boldsymbol{\alpha}}^{(k+1)} = \overline{\mathbf{Z}}_u^{(k)} - \overline{x}_u^{(k)} \widehat{\boldsymbol{\beta}}^{(k)}}$$

$$\text{sendo } \bar{\mathbf{z}}_u^{(k)} = \frac{\sum_{i=1}^n \widehat{u\mathbf{z}}_i^{*(k)}}{\sum_{i=1}^n \widehat{u}_i^{(k)}}, \quad \bar{x}_u^{(k)} = \frac{\sum_{i=1}^n \widehat{ux}_i^{(k)}}{\sum_{i=1}^n \widehat{u}_i^{(k)}} \quad \text{e } \widehat{u\mathbf{z}}_i^{*(k)} = (\widehat{uz}_{i2}, \dots, \widehat{uz}_{ip})^\top$$

ii)

$$\frac{\partial Q_1}{\partial \boldsymbol{\beta}^\top} = \frac{\partial Q_1}{\partial \mathbf{b}^\top} \frac{\partial \mathbf{b}}{\partial \boldsymbol{\beta}^\top} = \sum_{i=1}^n \left(\widehat{ux\mathbf{z}}_i \boldsymbol{\Omega}^{-1} - \mathbf{a}^\top \boldsymbol{\Omega}^{-1} \widehat{ux}_i - \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \widehat{ux}_i^2 \right) \begin{bmatrix} \mathbf{0} \\ \mathbb{I}_r \end{bmatrix} = \mathbf{0}$$

$$\mathbf{b}^\top \boldsymbol{\Omega}^{-1} \sum_{i=1}^n \widehat{ux}_i^2 \begin{bmatrix} \mathbf{0} \\ \mathbb{I}_r \end{bmatrix} = \sum_{i=1}^n \widehat{ux\mathbf{z}}_i \boldsymbol{\Omega}^{-1} \begin{bmatrix} \mathbf{0} \\ \mathbb{I}_r \end{bmatrix} - \mathbf{a}^\top \boldsymbol{\Omega}^{-1} \sum_{i=1}^n \widehat{ux}_i \begin{bmatrix} \mathbf{0} \\ \mathbb{I}_r \end{bmatrix}$$

$$\begin{aligned} \sum_{i=1}^n \widehat{ux}_i^2 [1, \boldsymbol{\beta}^\top] \begin{bmatrix} \frac{1}{\phi_1^2} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Omega}_{22}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ \mathbb{I}_r \end{bmatrix} &= \begin{bmatrix} \sum_{i=1}^n \widehat{ux\mathbf{z}}_{i1}, & \sum_{i=1}^n \widehat{ux\mathbf{z}}_i^* \end{bmatrix} \begin{bmatrix} \frac{1}{\phi_1^2} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Omega}_{22}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ \mathbb{I}_r \end{bmatrix} \\ &\quad - \sum_{i=1}^n \widehat{ux}_i [0, \boldsymbol{\alpha}^\top] \begin{bmatrix} \frac{1}{\phi_1^2} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Omega}_{22}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ \mathbb{I}_r \end{bmatrix} \end{aligned}$$

$$\begin{aligned} \begin{bmatrix} \sum_{i=1}^n \widehat{ux}_i^2 \\ \frac{1}{\phi_1^2}, \sum_{i=1}^n \widehat{ux}_i^2 \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ \mathbb{I}_r \end{bmatrix} &= \begin{bmatrix} \sum_{i=1}^n \widehat{ux\mathbf{z}}_{i1}, & \sum_{i=1}^n \widehat{ux\mathbf{z}}_i^* \boldsymbol{\Omega}_{22}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ \mathbb{I}_r \end{bmatrix} \\ &\quad - \begin{bmatrix} 0, & \sum_{i=1}^n \widehat{ux}_i \boldsymbol{\alpha}^\top \boldsymbol{\Omega}_{22}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ \mathbb{I}_r \end{bmatrix} \end{aligned}$$

$$\sum_{i=1}^n \widehat{ux}_i^2 \boldsymbol{\beta}^\top \boldsymbol{\Omega}_{22}^{-1} = \sum_{i=1}^n \widehat{ux\mathbf{z}}_i^* \boldsymbol{\Omega}_{22}^{-1} - \sum_{i=1}^n \widehat{ux}_i \boldsymbol{\alpha}^\top \boldsymbol{\Omega}_{22}^{-1}$$

$$\sum_{i=1}^n \widehat{ux}_i^2 \boldsymbol{\beta} = \sum_{i=1}^n \widehat{ux\mathbf{z}}_i^{*\top} - \sum_{i=1}^n \widehat{ux}_i \boldsymbol{\alpha} = \sum_{i=1}^n \widehat{ux\mathbf{z}}_i^{*\top} - \sum_{i=1}^n \widehat{ux}_i \left(\frac{\sum_{i=1}^n \widehat{uz}_i^* - \sum_{i=1}^n \widehat{ux}_i \boldsymbol{\beta}}{\sum_{i=1}^n \widehat{u}_i} \right)$$

$$\left(\sum_{i=1}^n \widehat{ux_i^2} - \frac{\left(\sum_{i=1}^n \widehat{ux_i} \right)^2}{\sum_{i=1}^n \widehat{u_i}} \right) \widehat{\beta} = \frac{\sum_{i=1}^n \widehat{u_i} \sum_{i=1}^n \widehat{uxz_i}^{*\top} - \sum_{i=1}^n \widehat{ux_i} \sum_{i=1}^n \widehat{uz_i}^*}{\sum_{i=1}^n \widehat{u_i}}$$

Então
$$\widehat{\beta} = \frac{\sum_{i=1}^n \widehat{u_i} \sum_{i=1}^n \widehat{uxz_i}^{*\top} - \sum_{i=1}^n \widehat{ux_i} \sum_{i=1}^n \widehat{uz_i}^*}{\sum_{i=1}^n \widehat{u_i} \sum_{i=1}^n \widehat{ux_i^2} - \left(\sum_{i=1}^n \widehat{ux_i} \right)^2}$$

Mas
$$\overline{\widehat{u}^{(k)}} = \frac{1}{n} \sum_{i=1}^n \widehat{u_i}^{(k)} \Rightarrow \sum_{i=1}^n \widehat{u_i}^{(k)} = n \overline{\widehat{u}^{(k)}}$$

$$\therefore \widehat{\beta}^{(k+1)} = \frac{n \overline{\widehat{u}^{(k)}} \sum_{i=1}^n \widehat{uxz_i}^{*\top(k)} - \sum_{i=1}^n \widehat{ux_i}^{(k)} \sum_{i=1}^n \widehat{uz_i}^{*(k)}}{n \overline{\widehat{u}^{(k)}} \sum_{i=1}^n \widehat{ux_i^2}^{(k)} - \left(\sum_{i=1}^n \widehat{ux_i}^{(k)} \right)^2}$$

em que $\widehat{uxz_i}^* = (\widehat{uxz_{i2}}, \dots, \widehat{uxz_{ip}})$ e $\widehat{uz_i}^* = (\widehat{uz_{i2}}, \dots, \widehat{uz_{ip}})^\top$.

iii)

$$Q_1 = -\frac{n}{2} \sum_{j=1}^p \log \phi_j^2 - \frac{1}{2} \sum_{i=1}^n \left\{ \text{tr} \left(\begin{bmatrix} \frac{1}{\phi_1^2} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{1}{\phi_p^2} \end{bmatrix} \begin{bmatrix} \widehat{uz_{i11}^2} & \dots & \widehat{uz_{i1p}^2} \\ \vdots & \ddots & \vdots \\ \widehat{uz_{ip1}^2} & \dots & \widehat{uz_{ipp}^2} \end{bmatrix} \right) \right.$$

$$- 2[0, \alpha_1, \dots, \alpha_r] \begin{bmatrix} \frac{1}{\phi_1^2} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{1}{\phi_p^2} \end{bmatrix} \begin{bmatrix} \widehat{uz_{i1}} \\ \vdots \\ \widehat{uz_{ip}} \end{bmatrix} - 2[\widehat{uxz_{i1}}, \dots, \widehat{uxz_{ip}}] \begin{bmatrix} \frac{1}{\phi_1^2} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{1}{\phi_p^2} \end{bmatrix} \begin{bmatrix} 1 \\ \beta_1 \\ \vdots \\ \beta_r \end{bmatrix}$$

$$+ [0, \alpha_1, \dots, \alpha_r] \begin{bmatrix} \frac{1}{\phi_1^2} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{1}{\phi_p^2} \end{bmatrix} \begin{bmatrix} 0 \\ \alpha_1 \\ \vdots \\ \alpha_r \end{bmatrix} \widehat{u_i} + 2[0, \alpha_1, \dots, \alpha_r] \begin{bmatrix} \frac{1}{\phi_1^2} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{1}{\phi_p^2} \end{bmatrix} \begin{bmatrix} 1 \\ \beta_1 \\ \vdots \\ \beta_r \end{bmatrix} \widehat{ux_i}$$

$$\begin{aligned}
& + [1, \beta_1, \dots, \beta_r] \left[\begin{array}{ccc} \frac{1}{\phi_1^2} & \dots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \dots & \frac{1}{\phi_p^2} \end{array} \right] \left[\begin{array}{c} 1 \\ \beta_1 \\ \vdots \\ \beta_r \end{array} \right] \widehat{ux_i^2} \Bigg\} \\
& = -\frac{n}{2} \sum_{j=1}^p \log \phi_j^2 - \frac{1}{2} \sum_{i=1}^n \left[\left(\frac{\widehat{uz_{i11}^2}}{\phi_1^2} + \dots + \frac{\widehat{uz_{ipp}^2}}{\phi_p^2} \right) - 2 \left(\frac{\widehat{uz_{i2}\alpha_1}}{\phi_2^2} + \dots + \frac{\widehat{uz_{ip}\alpha_r}}{\phi_p^2} \right) \right. \\
& \quad - 2 \left(\frac{\widehat{uxz_{i1}}}{\phi_1^2} + \frac{\widehat{uxz_{i2}\beta_1}}{\phi_2^2} + \dots + \frac{\widehat{uxz_{ip}\beta_r}}{\phi_p^2} \right) + \left(\frac{\widehat{u_i\alpha_1^2}}{\phi_2^2} + \dots + \frac{\widehat{u_i\alpha_r^2}}{\phi_p^2} \right) \\
& \quad \left. + 2 \left(\frac{\widehat{ux_i\alpha_1\beta_1}}{\phi_2^2} + \dots + \frac{\widehat{ux_i\alpha_r\beta_r}}{\phi_p^2} \right) + \left(\frac{\widehat{ux_i^2}}{\phi_1^2} + \frac{\widehat{ux_i^2\beta_1^2}}{\phi_2^2} + \dots + \frac{\widehat{ux_i^2\beta_p^2}}{\phi_p^2} \right) \right].
\end{aligned}$$

$$\begin{aligned}
\frac{\partial Q_1}{\partial \phi_1} &= -\frac{2n\phi_1}{2\phi_1^2} - \frac{1}{2} \sum_{i=1}^n \left[-\frac{2\widehat{uz_{i11}^2}}{\phi_1^3} - 2 \left(\frac{-2\widehat{uxz_{i1}}}{\phi_1^3} \right) + \left(\frac{-2\widehat{ux_i^2}}{\phi_1^3} \right) \right] = 0 \\
\Rightarrow \frac{n}{\phi_1} &= \frac{1}{\phi_1^3} \sum_{i=1}^n \left(\widehat{uz_{i11}^2} - 2\widehat{uxz_{i1}} + \widehat{ux_i^2} \right)
\end{aligned}$$

$$\therefore \boxed{\widehat{\phi_1^{(k+1)}} = \frac{1}{n} \sum_{i=1}^n \left(\widehat{uz_{i11}^{(k)}} - 2\widehat{uxz_{i1}^{(k)}} + \widehat{ux_i^{(k)}} \right)}$$

$$\begin{aligned}
\frac{\partial Q_1}{\partial \phi_j} &= -\frac{2n\phi_j}{2\phi_j^2} - \frac{1}{2} \sum_{i=1}^n \left[-\frac{2\widehat{uz_{ijj}^2}}{\phi_j^3} - 2 \left(\frac{-2\widehat{uz_{ij}\hat{\alpha}_{j-1}}}{\phi_j^3} \right) - 2 \left(\frac{-2\widehat{uxz_{ij}\hat{\beta}_{j-1}}}{\phi_j^3} \right) \right. \\
& \quad \left. + \left(\frac{-2\widehat{u_i\hat{\alpha}_{j-1}^2}}{\phi_j^3} \right) + 2 \left(\frac{-2\widehat{ux_i\hat{\alpha}_{j-1}\hat{\beta}_{j-1}}}{\phi_j^3} \right) + \left(\frac{-2\widehat{ux_i^2\hat{\beta}_{j-1}^2}}{\phi_j^3} \right) \right] = 0 \\
\Rightarrow \frac{n}{\phi_j} &= \frac{1}{\phi_j^3} \sum_{i=1}^n \left(\widehat{uz_{ijj}^2} - 2\widehat{uz_{ij}\hat{\alpha}_{j-1}} - 2\widehat{uxz_{ij}\hat{\beta}_{j-1}} + \widehat{u_i\hat{\alpha}_{j-1}^2} + 2\widehat{ux_i\hat{\alpha}_{j-1}\hat{\beta}_{j-1}} + \widehat{ux_i^2\hat{\beta}_{j-1}^2} \right)
\end{aligned}$$

Portanto,

$$\boxed{\widehat{\phi_{j+1}^{(k+1)}} = \frac{1}{n} \sum_{i=1}^n \left(\widehat{uz_{i(j+1)(j+1)}^{(k)}} + \widehat{u_i^{(k)}\hat{\alpha}_j^{2(k+1)}} + \widehat{ux_i^{(k)}\hat{\beta}_j^{2(k+1)}} + 2\widehat{ux_i^{(k)}\hat{\alpha}_j^{(k+1)}\hat{\beta}_j^{(k+1)}} \right.}$$

$$\left. - 2\widehat{uxz_{i(j+1)}^{(k)}\hat{\beta}_j^{(k+1)}} - 2\widehat{uz_{i(j+1)}^{(k)}\hat{\alpha}_j^{(k+1)}} \right), \quad j = 1, \dots, r.$$

iv)

$$\frac{\partial Q_2}{\partial \mu_x} = -\frac{1}{2\sigma_x^2} \sum_{i=1}^n (-2\widehat{u}x_i + 2\mu_x \widehat{u}_i) = 0 \quad \Rightarrow \quad \sum_{i=1}^n \widehat{u}x_i = \mu_x \sum_{i=1}^n \widehat{u}_i$$

$$\therefore \widehat{\mu}_x^{(k+1)} = \frac{\sum_{i=1}^n \widehat{u}x_i^{(k)}}{\sum_{i=1}^n \widehat{u}_i^{(k)}} = \bar{x}_u^{(k)}$$

v)

$$\frac{\partial Q_2}{\partial \sigma_x^2} = -\frac{n}{2\sigma_x^2} + \frac{1}{2\sigma_x^4} \sum_{i=1}^n (\widehat{u}x_i^2 - 2\mu_x \widehat{u}x_i + \mu_x^2 \widehat{u}_i) = 0$$

$$\therefore \widehat{\sigma}_x^{2(k+1)} = \frac{1}{n} \sum_{i=1}^n \left(\widehat{u}x_i^{2(k)} - 2\widehat{\mu}_x^{(k+1)} \widehat{u}x_i^{(k)} + \widehat{\mu}_x^{2(k+1)} \widehat{u}_i^{(k)} \right) \quad \blacksquare$$

B.2 Cálculo dos elementos da matriz de informação observada

Da Seção 2.3.2.1, temos que

$$\begin{aligned} \ell_{ic}(\boldsymbol{\theta} | \mathbf{Z}_i^c) &= -\frac{1}{2} \sum_{j=1}^p \log \phi_j^2 - \frac{1}{2} u_i (\mathbf{Z}_i - \mathbf{a} - \mathbf{b}x_i)^\top \boldsymbol{\Omega}^{-1} (\mathbf{Z}_i - \mathbf{a} - \mathbf{b}x_i) - \frac{1}{2} \log \sigma_x^2 - \frac{u_i}{2\sigma_x^2} (x_i - \mu_x)^2 \\ &= -\frac{1}{2} \sum_{j=1}^p \log \phi_j^2 - \frac{1}{2} u_i \mathbf{Z}_i^\top \boldsymbol{\Omega}^{-1} \mathbf{Z}_i + \mathbf{a}^\top \boldsymbol{\Omega}^{-1} u_i \mathbf{Z}_i + u_i x_i \mathbf{Z}_i^\top \boldsymbol{\Omega}^{-1} \mathbf{b} - \frac{1}{2} u_i \mathbf{a}^\top \boldsymbol{\Omega}^{-1} \mathbf{a} \\ &\quad - \mathbf{a}^\top \boldsymbol{\Omega}^{-1} \mathbf{b} u_i x_i - \frac{1}{2} u_i x_i^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1} \mathbf{b} - \frac{1}{2} \log \sigma_x^2 - \frac{u_i}{2\sigma_x^2} (x_i - \mu_x)^2 \end{aligned}$$

i)

$$\begin{aligned} \frac{\partial \ell_{ic}(\boldsymbol{\theta} | \mathbf{Z}_i^c)}{\partial \boldsymbol{\alpha}^\top} &= \frac{\partial \ell_{ic}(\boldsymbol{\theta} | \mathbf{Z}_i^c)}{\partial \mathbf{a}^\top} \frac{\partial \mathbf{a}}{\partial \boldsymbol{\alpha}^\top} = [(\boldsymbol{\Omega}^{-1} u_i \mathbf{Z}_i)^\top - u_i \mathbf{a}^\top \boldsymbol{\Omega}^{-1} - (\boldsymbol{\Omega}^{-1} \mathbf{b} u_i x_i)^\top] \begin{bmatrix} \mathbf{0}_{1 \times r} \\ \mathbb{I}_r \end{bmatrix} \\ \Rightarrow \frac{\partial \ell_{ic}(\boldsymbol{\theta} | \mathbf{Z}_i^c)}{\partial \boldsymbol{\alpha}} &= \left[\frac{\partial \ell_{ic}(\boldsymbol{\theta} | \mathbf{Z}_i^c)}{\partial \boldsymbol{\alpha}^\top} \right]^\top = [\mathbf{0}_{r \times 1}, \mathbb{I}_r] [\boldsymbol{\Omega}^{-1} u_i \mathbf{Z}_i - \boldsymbol{\Omega}^{-1} u_i \mathbf{a} - \boldsymbol{\Omega}^{-1} \mathbf{b} u_i x_i] \\ \widehat{\mathbf{s}}_{i,\boldsymbol{\alpha}} &= E \left[\frac{\partial \ell_{ic}(\boldsymbol{\theta} | \mathbf{Z}_i^c)}{\partial \boldsymbol{\alpha}} \middle| \mathbf{V}_i, \widehat{\boldsymbol{\theta}} \right] = [\mathbf{0}, \mathbb{I}_r] \widehat{\boldsymbol{\Omega}}^{-1} \left(E[U_i \mathbf{Z}_i | \mathbf{V}_i] - E[U_i | \mathbf{V}_i] \widehat{\mathbf{a}} - E[U_i x_i | \mathbf{V}_i] \widehat{\mathbf{b}} \right) \end{aligned}$$

$$\therefore \widehat{\mathbf{s}}_{i,\boldsymbol{\alpha}} = \mathbb{I}_{(p)} \widehat{\boldsymbol{\Omega}}^{-1} (\widehat{u} \widehat{\mathbf{z}}_i - \widehat{u}_i \widehat{\mathbf{a}} - \widehat{u} x_i \widehat{\mathbf{b}})$$

ii)

$$\begin{aligned} \frac{\partial \ell_{ic}(\boldsymbol{\theta}|\mathbf{Z}_i^c)}{\partial \boldsymbol{\beta}^\top} &= \frac{\partial \ell_{ic}(\boldsymbol{\theta}|\mathbf{Z}_i^c)}{\partial \mathbf{b}^\top} \frac{\partial \mathbf{b}}{\partial \boldsymbol{\beta}^\top} = [u_i x_i \mathbf{Z}_i^\top \boldsymbol{\Omega}^{-1} - \mathbf{a}^\top \boldsymbol{\Omega}^{-1} u_i x_i - u_i x_i^2 \mathbf{b}^\top \boldsymbol{\Omega}^{-1}] \begin{bmatrix} \mathbf{0} \\ \mathbb{I}_r \end{bmatrix}_{p \times r} \\ \Rightarrow \frac{\partial \ell_{ic}(\boldsymbol{\theta}|\mathbf{Z}_i^c)}{\partial \boldsymbol{\beta}} &= \left[\frac{\partial \ell_{ic}(\boldsymbol{\theta}|\mathbf{Z}_i^c)}{\partial \boldsymbol{\beta}^\top} \right]^\top = [\mathbf{0}, \mathbb{I}_r] \boldsymbol{\Omega}^{-1} [u_i x_i \mathbf{Z}_i - u_i x_i \mathbf{a} - u_i x_i^2 \mathbf{b}] \\ \Rightarrow \hat{\mathbf{s}}_{i,\boldsymbol{\beta}} &= E \left[\frac{\partial \ell_{ic}(\boldsymbol{\theta}|\mathbf{Z}_i^c)}{\partial \boldsymbol{\beta}} \middle| \mathbf{V}_i, \hat{\boldsymbol{\theta}} \right] \\ &= [\mathbf{0}, \mathbb{I}_r] \hat{\boldsymbol{\Omega}}^{-1} \left(E[U_i x_i \mathbf{Z}_i | \mathbf{V}_i] - E[U_i x_i | \mathbf{V}_i] \hat{\mathbf{a}} - E[U_i x_i^2 | \mathbf{V}_i] \hat{\mathbf{b}} \right) \end{aligned}$$

Mas $E[U_i x_i \mathbf{Z}_i | \mathbf{V}_i] = (E[U_i x_i \mathbf{Z}_i^\top | \mathbf{V}_i])^\top = \widehat{u x \mathbf{z}}_i^\top$

$$\therefore \boxed{\hat{\mathbf{s}}_{i,\boldsymbol{\beta}} = \mathbb{I}_{(p)} \hat{\boldsymbol{\Omega}}^{-1} (\widehat{u x \mathbf{z}}_i^\top - \widehat{u x}_i \hat{\mathbf{a}} - \widehat{u x^2}_i \hat{\mathbf{b}})}$$

iv)

$$\frac{\partial \ell_{ic}(\boldsymbol{\theta}|\mathbf{Z}_i^c)}{\partial \mu_x} = \frac{u_i}{\sigma_x^2} (x_i - \mu_x) \quad \Rightarrow \quad \hat{\mathbf{s}}_{i,\mu_x} = E \left[\frac{U_i x_i - U_i \mu_x}{\sigma_x^2} \middle| \mathbf{V}_i, \hat{\boldsymbol{\theta}} \right]$$

$$\therefore \boxed{\hat{\mathbf{s}}_{i,\mu_x} = \frac{1}{\hat{\sigma}_x^2} (\widehat{u x}_i - \widehat{u}_i \widehat{\mu}_x)}$$

v)

$$\frac{\partial \ell_{ic}(\boldsymbol{\theta}|\mathbf{Z}_i^c)}{\partial \sigma_x^2} = -\frac{1}{2\sigma_x^2} + \frac{u_i}{2\sigma_x^4} (x_i - \mu_x)^2 = -\frac{1}{2\sigma_x^2} + \frac{1}{2\sigma_x^4} (u_i x_i^2 - 2u_i x_i \mu_x + u_i \mu_x^2)$$

$$\Rightarrow \hat{\mathbf{s}}_{i,\sigma_x^2} = E \left[-\frac{1}{2\sigma_x^2} + \frac{1}{2\sigma_x^4} (U_i x_i^2 - 2U_i x_i \mu_x + U_i \mu_x^2) \middle| \mathbf{V}_i, \hat{\boldsymbol{\theta}} \right]$$

$$\therefore \boxed{\hat{\mathbf{s}}_{i,\sigma_x^2} = -\frac{1}{2\hat{\sigma}_x^2} + \frac{1}{2\hat{\sigma}_x^4} (\widehat{u x^2}_i - 2\widehat{u x}_i \widehat{\mu}_x + \widehat{u}_i \widehat{\mu}_x^2)} \quad \blacksquare$$

Referências Bibliográficas

- Andrews, D. F. e Mallows, S. L. (1974), “Scale mixtures of normal distributions,” *Journal of the Royal Statistical Society, Series B*, **36**, 99–102.
- Arellano-Valle, R. B., Ozan, S., Bolfarine, H., e Lachos, V. H. (2005), “Skew normal measurement error models,” *Journal of Multivariate Analysis*, **98**, 265–281.
- Ash, R. B. e Doleans-Dade, C. A. (2000), *Probability and Measure Theory*, Academic Press, San Diego.
- Barnett, V. D. (1969), “Simultaneous pairwise linear structural relationships,” *Biometrics*, **25**, 129–142.
- Bolfarine, H. e Galea-Rojas, M. (1996), “On structural comparative calibration under a t-model,” *Computational Statistics*, **11**, 63–85.
- Bolfarine, H., Montenegro, L. C., e Lachos, V. H. (2007), “Influence Diagnostics for Skew-Normal Linear Mixed Models,” *Sankhya: The Indian Journal of Statistics*, **69** (4), 648–670.
- Buonaccorsi, J. (2010), *Measurement Error: Models, Methods, and Applications*, Chapman and Hall/CRC, Boca Raton.
- Carroll, R. J., Ruppert, D., Stefanski, L. A., e Crainiceanu, C. M. (2006), *Measurement Error in Nonlinear Models*, Chapman & Hall/CRC, Boca Raton, second edition.
- Cheng, C. L. e Van-Ness, J. W. (1999), *Statistical regression with measurement error*, Arnold, London.
- Chipkevitch, E., Nishimura, R. T., Tu, D. G. S., e Galea-Rojas, M. (1996), “Clinical measurement of testicular volume in adolescents: Comparison of the reliability of 5 methods,” *The Journal of Urology*, **156**, 2050–2053.
- Cook, R. D. (1977), “Detection of Influential Observation in Linear Regression,” *Tech-*

- nometrics*, **19**, 15–18.
- Cook, R. D. (1986), “Assessment of local influence,” *Journal of the Royal Statistical Society, Series B*, **48**, 133–169.
- Dempster, A. P., Laird, N. M., e Rubin, D. B. (1977), “Maximum likelihood from incomplete data via the EM algorithm,” *Journal of the Royal Statistical Society, Series B*, **39**, 1–22.
- Dunn, G. (1992), *Design and Analysis of Reliability: The statistical evaluation of measurement errors*, Edward Arnold, New York.
- Fang, K. T., Kotz, S., e Ng, K. W. (1990), *Symmetric multivariate and related distribution*, Chapman and Hall, London.
- Fuller, W. A. (1987), *Measurement Error Models*, John Wiley and Sons, New York.
- Galea-Rojas, M., Bolfarine, H., e de Castro, M. (2002), “Local influence in comparative calibration models,” *Biometrical journal*, **44**, 59–81.
- Galea-Rojas, M., Bolfarine, H., e Vilca, L. F. (2005), “Local influence in Comparative calibration models under elliptical t-distribution,” *Biometrical Journal*, **47**, 691–706.
- Genz, A., Bretz, F., Miwa, T., Mi, X., Leisch, F., Scheipl, F., e Hothorn, T. (2018), *mvtnorm: Multivariate Normal and t Distributions*, R package version 1.0-8.
- Ho, H. J., Lin, T. I., Chen, H. Y., e Wang, W. L. (2012), “Some results on the truncated multivariate t distribution,” *Journal of Statistical Planning and Inference*, **142**, 25–40.
- Kelly, G. (1984), “The influence function in the errors in variables problem,” *Annals of Statistics*, **12**, 87–100.
- Kotz, S. e Nadarajah, S. (2004), *Multivariate t Distributions and Their Applications*, Cambridge University Press, Cambridge.
- Lachos, V. H., Labra, F. V., Bolfarine, H., e Ghosh, P. (2010), “Multivariate measurement error models based on scale mixtures of the skew-normal distribution,” *Statistics*, **44**, 541–556.
- Lachos, V. H., Angolini, T., e Abanto-Valle, C. A. (2011), “On estimation and local influence analysis for measurement errors models under heavy-tailed distributions,” *Statistical Papers*, **52**, 567–590.
- Lange, K. L. e Sinsheimer, J. S. (1993), “Normal/independent distributions and their

- applications in robust regression,” *J. Comput. Graph. Stat*, **2**, 175–198.
- Lange, K. L., Little, R. J. A., e Taylor, J. M. G. (1989), “Robust statistical modeling using the t distribution,” *Journal of the American Statistical Association*, **84**, 881–896.
- Lee, S. Y. e Xu, L. (2004), “Influence analysis of nonlinear mixed-effects models,” *Computational Statistics and Data Analysis*, **45**, 321–341.
- Lesaffre, E. e Verbeke, G. (1998), “Local Influence in Linear Mixed Models,” *Biometrics, Series B*, **54**, 570–582.
- Lin, T. I. (2010), “Robust mixture modeling using multivariate skew t distributions,” *Statistics and Computing*, **20**, 343–356.
- Louis, T. A. (1982), “Finding the observed information matrix when using the EM algorithm,” *Journal of the Royal Statistical Society, Series B*, **44**, 226–233.
- Lu, B. e Song, X. Y. (2006), “Local influence analysis of multivariate probit latent variable models,” *Journal of Multivariate Analysis*, **97**, 1783–1798.
- Lu, Y., Ye, K., Mathur, A., Hui, S., Fuerst, T., e Genant, H. (1997), “Comparative calibration without a gold standard,” *Statistics in Medicine*, **16**, 1889–1905.
- Massuia, M. B., Cabral, C. R. B., Matos, L. A., e Lachos, V. H. (2015), “Influence diagnostics for Student-t censored linear regression models,” *Statistics*, **49**, 1074–1094.
- Matos, L. A., Lachos, V. H., Balakrishnan, N., e Labra, F. V. (2013), “Influence diagnostics in linear and nonlinear mixed-effects models with censored data,” *Computational Statistics and Data Analysis*, **57**, 450–464.
- Matos, L. A., Castro, L. M., Cabral, C. R. B., e Lachos, V. H. (2016), “Multivariate measurement error models based on Student-t distribution under censored responses,” Tech. rep., UNICAMP.
- Meilijson, I. (1989), “A fast improvement to the EM algorithm on its own terms,” *Journal of the Royal Statistical Society, Series B*, **51**, 127–138.
- Meng, X. L. e Rubin, D. B. (1993), “Maximum likelihood estimation via the ECM algorithm: A general framework,” *Biometrika*, **80**, 267–278.
- Poon, W. Y. e Poon, Y. S. (1999), “Conformal normal curvature and assessment of local influence,” *Journal of the Royal Statistical Society, Series B*, **61**, 51–61.
- R Core Team (2018), *R: A Language and Environment for Statistical Computing*, R

Foundation for Statistical Computing, Vienna, Austria.

- Wei, G. e Tanner, M. (1990), “A Monte Carlo implementation of the EM algorithm and the poor man’s data augmentation algorithms,” *J. Am. Stat. Assoc.*, **85**, 699–704.
- Wu, C. F. J. (1983), “On the convergence properties of the EM algorithm,” *The Annals of Statistics*, **11**, 95–103.
- Zeller, C. B., Labra, F. V., Lachos, V. H., e Balakrishnan, N. (2010), “Influence analyses of skew-normal/independent linear mixed models,” *Computational Statistics and Data Analysis*, **54**, 1266–1280.
- Zhu, H. e Lee, S. (2001), “Local influence for incomplete-data models,” *Journal of the Royal Statistical Society, Series B*, **63**, 121–126.
- Zhu, H., Lee, S., Wei, B., e Zhou, J. (2001), “Case-deletion measures for models with incomplete data,” *Biometrika*, **88** (3), 727–737.
- Zhu, H., Ibrahim, J. G., e Shi, X. (2009), “Diagnostic measures for generalized linear models with missing covariates,” *Scandinavian Journal of Statistics*, **36**, 686–712.