UNIVERSIDADE ESTADUAL DO OESTE DO PARANÁ
CENTRO DE CIÊNCIAS EXATAS E TECNOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA AGRÍCOLA

MODELO ESPACIAL BIRNBAUM-SAUNDERS APLICADO A DADOS AGRÍCOLAS

FABIANA MAGDA GARCIA PAPANI

CASCAVEL - PR
FEVEREIRO - 2016

FABIANA MAGDA GARCIA PAPANI

MODELO ESPACIAL BIRNBAUM-SAUNDERS APLICADO A DADOS AGRÍCOLAS

Tese apresentada ao Programa de Pós-Graduação em Engenharia Agrícola da Universidade Estadual do Oeste do Paraná, em cumprimento parcial aos requisitos para a obtenção do título de Doutor em Engenharia Agrícola, área de concentração Sistemas Biológicos e Agroindustriais.

Orientador: Prof. Dr. Miguel Angel Uribe Opazo

Co-orientador: Prof. Dr. Victor Leiva

CASCAVEL - PR
FEVEREIRO - 2016

Dados Internacionais de Catalogação-na-Publicação (CIP)

P23m

Papani, Fabiana Magda Garcia
Modelo espacial birnbaum – saunders aplicado a dados agrícolas./Fabiana Magda Garcia Papani. Cascavel, 2016.
135 p.

Orientador: Prof. Dr. Miguel Angel Uribe Opazo
Coorientador: Prof. Dr. Victor Leiva

Tese (Doutorado) – Universidade Estadual do Oeste do Paraná, Campus de Cascavel, 2016
Programa de Pós-Graduação Stricto Sensu em Engenharia Agrícola

1. Análise de dados espaciais. 2. Diagnósticos - Análise. 3. Distribuições assimétricas. 4. Geoestatistica. 5. Variabilidade espacial. I. Uribe Opazo, Miguel Angel. II. Leiva, Victor. III. Universidade Estadual do Oeste do Paraná. IV. Titulo.

CDD 21.ed. 630.285
CIP-NBR 12899

Ficha catalográfica elaborada por Helena Soterio Bejio – CRB 9ª/965

Revisões de português, inglês e normas realizadas por Ana Maria Martins Alves Vasconcelos em 19 de abril de 2016.

FABIANA MAGDA GARCIAPAPANI

"MODELO ESPACIAL BIRNBAUM - SAUNDERS APLICADO A DADOS AGRÍCOLAS"

Tese apresentada ao Programa de Pós-Graduação "*Stricto Sensu*" em Engenharia Agrícola em cumprimento parcial aos requisitos para obtenção do título de doutora em Engenharia Agrícola, área de concentração Sistemas Biológicos e Agroindustriais, **aprovada** pela seguinte banca examinadora:

Prof.ª Dra. Luciana Pagliosa Carvalho Guedes
Universidade Estadual do Oeste do Paraná – *Campus* de Cascavel

Prof. Dr. Cristian Marcelo Villegas Lobos
Escola Superior de Agricultura "Luiz de Queiroz" Universidade de São Paulo

Prof. Dr. Joelmir André Borssoi
Universidade Tecnológica Federal do Paraná

Prof.ª Dra. Rosangela Aparecida Botinha Assumpção
Universidade Tecnológica Federal do Paraná

Prof. Dr. Jerry Adriani Johann
Universidade Estadual do Oeste do Paraná – *Campus* de Cascavel

Cascavel, 02 de fevereiro de 2016.

BIOGRAFIA

Nome: Fabiana Magda Garcia Papani

Ano de nascimento: 1973

Naturalidade: Olímpia-SP

Licenciada em Matemática pelo Instituto de Biociências Letras e Ciências Exatas da Universidade Estadual Júlio de Mesquita Filho (IBILCE/UNESP) no ano de 1994.

Mestre em Ciências Matemática também pelo IBILCE/UNESP no ano de 1997.

Professora assistente do Centro de Ciências Exatas e Tecnológicas da Universidade Estadual do Oeste do Paraná (CCET/UNIOESTE) de 1998 até a presente data.

Desistir? Eu já pensei seriamente nisso, mas nunca me levei realmente a sério, é que tem mais chão nos meus olhos do que o cansaço nas minhas pernas, mais esperanças nos meus passos do que tristeza nos meus ombros, mais estrada no meu coração do que medo na minha cabeça.

[Cora Coralina]

Ao meu amado esposo, Carlos, e aos meus incondicionalmente amados filhos, Gabriel e Felipe,

dedico.

AGRADECIMENTOS

À Universidade Estadual do Oeste do Paraná (UNIOESTE)/Campus Cascavel, pela oportunidade oferecida em fazer o doutorado com afastamento integral. Agradeço, em especial, aos colegas do Colegiado de Matemática que absorveram minhas atribuições durante o perído de afastamento;

Aos professores do Programa de Pós-Graduação em Engenharia Agrícola da UNIOESTE por contribuírem para minha formação;

Ao meu orientador, Professor Dr. Miguel Angel Uribe Opazo, pela confiança em mim depositada e pela orientação ao longo do desenvolvimento deste trabalho;

Ao meu co-orientador, Professor Dr. Victor Leiva, pela disponibilidade em ajudar-me e pelas contribuições a este trabalho;

Aos membros da banca avaliadora, pelas considerações que contribuíram para melhorar o texto final;

Aos colegas do grupo de estatística espacial: Daniela, Denise (tia Dê), Elisabeth, Gustavo, Regiane e Rosângela e à amiga Fernanda por compartilharem as alegrias e tristesas durante a elaboração deste trabalho, pelas constantes discussões e pela ajuda no uso do software R e com Latex. Ao amigo Petterson, pelas inúmeras ajudas com o Latex;

Aos meus amigos e vizinhos, Simone, Paulo, Sandra e Sandro, pela disposição em ajudar-me com o Latex e pelo convívio que, por várias vezes, trouxeram momentos de descontração;

À amiga Mariângela, pela disposição em ler o texto, pelos conselhos que por diversas vezes sossegaram minha alma, e pelas inúmeras vezes que me ajudou com os meninos, seja levando-os e trazendo-os, seja completando o carinho de mãe;

À amiga Nelci, que assim como a Mariângela, me ajudou muitas vezes sendo mãe para meus filhos;

Aos pequenos João Victor, Gabriel, Mariana, José Lucas (já não tão pequenos) e Paulinha, pelo simples fato de, juntamente com meus filhos, levaram alegria para minha casa;

A todos meus amigos que, embora não estivessem presente, diariamente, tenho certeza, torceram e oraram por mim;

Aos meus pais, irmão, minhas avós e demais familiares que, embora espacialmente distantes, contribuíram com palavras e pensamentos positivos para a conclusão deste trabalho;

Ao meu esposo Carlos e aos meus filhos Gabriel e Felipe, com meus pedidos de desculpas pelos inúmeros momentos de ausência, ainda que fisicamente presente, pela compreensão, apoio, palavras de incentivo e principalmente pela paciência e amor;

A Deus por tudo!

RESUMO

MODELO ESPACIAL BIRNBAUM-SAUNDERS APLICADO A DADOS AGRÍCOLAS

O conhecimento da distribuição espacial de dados georrefenciados é de interesse de diversas áreas do conhecimento, incluindo a área agrícola. Neste sentido, diversos trabalhos já foram realizados; no entanto, a maioria destes trabalhos assumem que o processo estocástico subjacente é gaussiano. Quando os dados associados com este processo não apresentam normalidade, transformações de dados são usadas. E ainda que o uso dessas transformações tenha apresentado resultados satisfatórios, considerar modelos que levem em conta as características do fenômeno pode ser mais adequado do que a utilização do modelo normal. O objetivo deste trabalho é propor um modelo espacial baseado na distribuição Birnbaum-Saunders (BS). Esta distribuição tem se mostrado eficiente para modelar conjuntos de dados formados por valores estritamente positivos e cujo comportamento apresenta assimetria positiva e unimodalidade. A metodologia proposta neste trabalho inclui a formulação do modelo espacial Birnbaum-Saunders, a estimação de seus parâmetros utilizando o método de máxima verossimilhança (ML), a aplicação de técnicas de diagnóstico que permitem detectar a sensibilidade do modelo a dados atípicos, a avaliação do modelo proposto por um estudo de simulação e aplicação da metodologia desenvolvida em análise de dados reais da área agrícola. Os dados utilizados para validação do modelo estudado foram obtidos em uma área comercial de produção de grãos de 167,35 ha de Cascavel. No estudo com dados simulados, para amostras grandes, a estimação dos parâmetros e a análise de diagnóstico apresentaram boa performance. No estudo com dados reais, os cálculos dos índices AIC, BIC e fator Bayes bem como a construção de Q-Q *plots* mostraram que o modelo proposto é adequado para ajustar os dados. Casos influentes foram detectados e suas retiradas do conjunto de dados causaram uma mudança considerável nos mapas de contorno. Conclui-se portanto, que o modelo espacial Birnbaum-Saunders é adequado para realização de estudos com dados espacialmente correlacionados, e é um modelo alternativo ao modelo normal quando o conjunto de dados apresenta distribuição assimétrica positiva.

Palavras-chave: análise de dados espaciais, análise de diagnósticos, distribuições assimétricas, geoestatística, variabilidade espacial.

ABSTRACT

BIRNBAUM-SAUNDERS SPATIAL MODEL APPLIED FOR AGRICULTURAL DATA

Understanding the spatial distribution knowledge regarding georeferenced data has been essencial to various areas including agriculture. Thus, several trials have been carried out. However, most of these studies assume that the underlying stochastic process is Gaussian. When the data associated with this process do not present normality, data transformations are applied. And though the use of these transformations has presented satisfactory results, it is important to consider models which take into account the characteristics of such phenomenon. It may be more appropriate than using a normal model. So, this trial aimed at proposing a spatial model based on the Birnbaum-Saunders distribution (BS). This distribution has been shown effective to model data that take positive values and whose behavior presents positive asymmetry and unimodality. Thefore, this trial has proposed a methodology that includes the formulation of the spatial Birnbaum-Saunders model , estimation of its parameters using maximum likelihood (ML), and application of diagnostic techniques which can detect the sensitivity of the model to atypical data and evaluate the proposed model through a simulation study and studies using real data sets of agricultural engineering. These data were obtadined in a 167.35-ha commercial area for grain production, in Cascavel city, to validate the studied model. In the study with simulated data and large samples, estimation parameters and diagnostic analysis showed a good performance. According to the study with real data, calculations of AIC (Akaike's information criterion) and BIC (Bayesian information criterion) indexes, Bayes factor as well as Q-Q *plots* constrution have shown that the proposed model is appropriate to fit the obtained data. Influential cases were detected, and their removal from data set caused a considerable change in contour maps. It is therefore concluided that Birnbaum-Saunders spatial model is adequate to carry out studies with spatially correlated data. Is is also an alternative model to the normal model when the data set present positive asymmetrical distribution.

Key-words: spatial data analysis, diagnostic analysis, asymmetric distributions, geostatistics, spatial variability.

SUMÁRIO

LISTA DE FIGURAS ..... **xii**

LISTA DE TABELAS ..... **xiv**

1 INTRODUÇÃO/JUSTIFICATIVA **1**

2 OBJETIVOS **3**

2.1 Objetivo geral 3

2.2 Objetivos específicos 3

3 REVISÃO BIBLIOGRÁFICA **4**

3.1 Distribuições Birnbaum-Saunders e log-Birnbaum-Saunders 4

3.1.1 Distribuição BS univariada 4

3.1.1.1 Propriedades da distribuição BS univariada 6

3.1.2 Distribuição BS bivariada 8

3.1.2.1 Propriedades do vetor aleatório BS bivariado 9

3.1.3 Distribuição BS multivariada 13

3.1.3.1 Propriedades do vetor aleatório BS multivariado 14

3.1.4 Distribuição log-BS univariada 17

3.1.4.1 Propriedades da variável aleatória log-BS univariada 19

3.1.5 Distribuição log-BS bivariada 23

3.1.5.1 Propriedades do vetor aleatório log-BS bivariado 24

3.1.6 Distribuição log-BS Multivariada 26

3.1.6.1 Propriedades do vetor aleatório log-BS multivariado 27

3.2 Estatística espacial 30

3.2.1 Modelos geoestatísticos 31

3.2.1.1 Ferramentas estatísticas para o estudo da variabilidade espacial 31

3.2.1.2 Modelos teóricos de semivariogramas 35

3.2.2 Krigagem 39

3.2.2.1 Krigagem ordinária 40

3.2.2.2 Krigagem com drift externo 41

3.2.3 Métodos de validação de modelos 43

3.2.3.1 Valor máximo da função verossimilhança 43

3.2.3.2 Validação cruzada 43

3.2.3.3 *Q-Q plots* 44

3.2.4 Medidas de similaridade entre mapas 45

3.2.5 Análise de diagnóstico 46
3.2.5.1 Influência global 46
3.2.5.2 Influência local 47
3.2.5.3 Alavanca generalizada 48
3.3 Atributos químicos do solo 49
3.3.1 Fósforo (P) 50
3.3.2 Cálcio (Ca) 50
3.3.3 Magnésio (Mg) 50
3.3.4 A relação Ca:Mg 51

4 REFERÊNCIAS BIBLIOGRÁFICAS 52

5 ARTIGOS 57
5.1 ARTIGO 1: Modelagem espacial Birnbaum-Saunders e diagnósticos aplicados a dados de engenharia agrícola 57
5.1.1 Introdução 57
5.1.2 Preliminares 59
5.1.2.1 A distribuição Birnbaum-Saunders 59
5.1.2.2 A distribuição log-Birnbaum-Saunders 60
5.1.2.3 Modelos espaciais 61
5.1.2.4 Covariância e modelos de variogramas 62
5.1.2.5 Interpolação Krigagem 63
5.1.2.6 Medidas de acurácia 63
5.1.3 O modelo espacial Birnbaum-Saunders 64
5.1.3.1 Formulação do modelo 64
5.1.3.2 Estimação dos parâmetros 64
5.1.3.3 Influência local 66
5.1.3.4 Seleção do esquema de perturbação apropriado 67
5.1.4 Estudos de simulação 68
5.1.4.1 Estudo I: Estimação ML 68
5.1.4.2 Estudo II: diagnósticos de influência 70
5.1.5 Aplicações numéricas 72
5.1.5.1 Descrição da área experimental e estudo das variáveis 72
5.1.5.2 Análise exploratória e de dependência espacial 73
5.1.5.3 Estimação dos parâmetros e krigagem 73
5.1.5.4 Seleção do modelo 74
5.1.5.5 Diagnóstico de Influência 75
5.1.6 Conclusões e trabalhos futuros 78
5.1.7 Referências 78
5.1.8 Apêndices 83
5.1.8.1 Apêndice I: O vetor escore do modelo BS log-linear espacial 83

5.1.8.2 Apêndice II: A matriz de informação observada para o modelo BS log-linear espacial 83
5.1.8.3 Apêndice III: Matriz de informação esperada para o modelo BS log-linear espacial 85
5.1.8.4 Apêndice IV: Vetor escore $U(\omega)$ e matriz $G(\omega)$ 86
5.1.8.5 Apêndice V: A matriz perturbação 86
5.2 ARTIGO 2: Krigagem com *drift* externo em um modelo geoestatístico Birnbaum-Saunders 88
5.2.1 Introdução 88
5.2.2 Modelo espacial Birnbaum-Saunders 90
5.2.2.1 Distribuições BS e log-BS multivariadas 90
5.2.2.2 Formulação do modelo de regressão 91
5.2.2.3 Formulação do modelo espacial 92
5.2.3 Estimação, validação e predição 93
5.2.3.1 Estimação dos parâmetros 93
5.2.3.2 Validação do modelo 94
5.2.3.3 Predição espacial: Krigagem com *drift* externo 95
5.2.4 Aplicação 97
5.2.4.1 Dados 97
5.2.4.2 Análise exploratória dos dados 98
5.2.4.3 Análise espacial do modelo 99
5.2.4.4 Bondade de ajuste 101
5.2.4.5 Predição 101
5.2.5 Conclusões 102
5.2.6 Trabalhos futuros 102
5.2.7 Referências 103
5.2.8 Apêndices 106
5.2.8.1 Apêndice I: Covariância da distribuição BS multivariada 106
5.2.8.2 Apêndice II: Matriz de informação observada de Fisher 107
5.2.8.3 Apêndice III: Matriz de informação esperada de Fisher 108
5.2.8.4 Apêndice IV: Matriz projeção para o modelo log-linear espacial BS . 109
5.3 ARTIGO 3: Diagnósticos de influência em um modelo log-linear Birnbaum-Saunders 111
5.3.1 Introdução 111
5.3.2 Modelo espacial Birnbaum-Saunders 113
5.3.2.1 Formulação do modelo 113
5.3.2.2 Estimação dos parâmetros 114
5.3.3 Diagnósticos de Influência 116
5.3.3.1 Influência global 116
5.3.3.2 Influência local 117

5.3.3.3 Alavanca generalizada 120
5.3.4 Aplicação 121
5.3.4.1 Descrição, análise exploratória e análise espacial dos dados 121
5.3.4.2 Diagnósticos 122
5.3.5 Conclusões 126
5.3.6 Trabalhos futuros 126
5.3.7 Referências 126
5.3.8 Apêndices 129
5.3.8.1 Apêndice I: Matrizes de informação observada e esperada para o modelo espacial log-linear BS 129
5.3.8.2 Apêndice II: Vetor escore $U(\omega)$ 130
5.3.8.3 Apêndice III: Matriz de pertubação para o modelo espacial log-linear BS 131

6 CONSIDERAÇÕES FINAIS 134

LISTA DE FIGURAS

Figura 1 Funções distribuições de probabilidade das distribuições $\mathrm{BS}(\alpha, 1)$ (a) e $\mathrm{BS}(1, \gamma)$ (b). 6
Figura 2 Funções distribuições de probabilidade da distribuição $\mathrm{BS}_2(\boldsymbol{\alpha}, \mathbf{1}; 0, 5)$ nos casos $\boldsymbol{\alpha} = (1,5; 1,5)$ (a) e $\boldsymbol{\alpha} = (0,5; 0,5)$ (b). 9
Figura 3 Funções densidades de probabilidade da distribuição $\mathrm{SHN}(\alpha, \mu, \sigma)$. 18
Figura 4 Funções densidades de probabilidade da distribuição $\log\text{-}\mathrm{BS}(\alpha, \mu)$. 19
Figura 5 Funções distribuição de probabilidade da distribuição $\log\text{-}\mathrm{BS}(\boldsymbol{\alpha}, \mathbf{0}; 0, 5)$ nos casos $\boldsymbol{\alpha} = (1,5; 1,5)$ (a); $\boldsymbol{\alpha} = (0,5; 0,5)$ (b) e $\boldsymbol{\alpha} = (1,5; 3)$ (c). 24
Figura 6 Malha bidimensional regular. 30
Figura 7 Ilustração de um h-*scatterplot*. 32
Figura 8 Principais direções para o semivariograma direcional. 33
Figura 9 Ilustração de um semivariograma experimental. 34
Figura 10 Ilustração dos parâmetros espaciais do modelo. 35
Figura 11 Semivariograma teórico: modelo esférico. 36
Figura 12 Semivariograma teórico: modelo exponencial. 37
Figura 13 Semivariograma teórico: modelo gaussiano. 37
Figura 14 Semivariogramas teóricos: modelo exponencial potência $p = 0,5$ (a) e $p = 1,5$ (b). 38
Figura 15 Semivariograma teórico: modelo família Matérn $\delta = 0,5$ (a) e $\delta = 2$ (b). 39
Figura 16 Funções distribuição de probabilidade das distribuições BS$(\alpha, 1)$ (a) e log-BS$(\alpha, 0)$ (b) para o valor do parâmetro de forma α indicado. 61
Figura 17 *Box plot* (a) e *post plot* (b) dos dados simulados para o modelo Matérn com $\delta = 0,5$. 70
Figura 18 Gráficos C_i *versus* i (a) e $|d_{\max_i}|$*versus* i (b) para os dados simulados para o modelo Matérn com $\delta = 0,5$. 71
Figura 19 *Box plot* (a) e *Post plot* (b) dos dados simulados para o modelo Matérn com $\delta = 1$. 71
Figura 20 Gráficos usando C_i *versus* i (a) e $|d_{\max}|$ *versus* i (b) dos dados simulados para o modelo Matérn com $\delta = 1$. 72
Figura 21 *Box plot* (a) e *post plot* (b) dos dados da concentração de Fósforo. 73
Figura 22 Gráficos C_i *versus* i (a) e $|d_{\max}|$ *versus* i (b) para os dados de concentração de Fósforo. 75
Figura 23 Mapas de contorno com os efeitos da remoção do(s) caso(s) indicados para os dados de concentração de Fósforo. 77
Figura 24 Grade amostral e pontos para serem preditos (a), *box plot* da concentração de Mg (b), pontos amostrais divididos em quartis e localização dos *outliers* (c) e semivariogramas direcionais (d). 99
Figura 25 Gráfico da validação cruzada em função de δ para os dados de Mg. 100
Figura 26 *Q-Q plots* usando a aproximação de Wilson-Hilferty (a) e resíduos quantis

randonizados (b) para os dados de Mg. 101
Figura 27 Grade amostral (a) e semivariogramas direcionais (b). 122
Figura 28 *Box plots* da concentração de Magnésio (a) e do logaritmo da concentração de Magnésio (b), e localização dos *outliers* na grade amostral (c). 122
Figura 29 Distância de Cook para o vetor de parâmetros $\boldsymbol{\theta}$ (a), para o vetor $\boldsymbol{\beta}$ (b) e para o vetor $\boldsymbol{\psi}$ (c). 123
Figura 30 Gráficos de influência local, B_i *versus* ordem (a), $B_{d_{max}}$ *versus* ordem (b), e S_{max} *versus*ordem (c), pertubando a variável resposta. 124
Figura 31 Gráficos de influência local, B_i *versus* ordem (a), $B_{d_{max}}$ *versus* ordem (b) e S_{max} *versus* ordem (c), pertubando a covariável. 125
Figura 32 Alavanca generalizada para os dados de concentração de magnésio. 125
Figura 33 Grade amostral com destaque para a localização da observação #27. 126

LISTA DE TABELAS

Tabela 1 Interpretação de $2\log(B_{12})$ associado com o BF. 44
Tabela 2 Matriz de erros genéricos $m \times m$. 46
Tabela 3 VRA (em $\%$) e REQM do estimador ML dos parâmetros e n indicados para dados simulados com o modelo espacial BS. 69
Tabela 4 Interpretação de $2\log(B_{12})$ associado com o BF. 74
Tabela 5 Comparação dos modelos BS e normal 74
Tabela 6 Estimativas ML dos parâmetros estimados (com SE estimados assintóticos entre parênteses e p-valores entre colchetes), e valores dos índices GA e κ para o conjunto de dados de concentração de Fósforo com caso(s) excluídos indicados. 76
Tabela 7 Valores da validação cruzada para o modelo de variabilidade espacial da família Matérn com δ indicado. 100
Tabela 8 Concentrações de Mg, amostradas e estimadas (cmolc/dm^{-3}). 102

1 INTRODUÇÃO/JUSTIFICATIVA

A teoria das variáveis regionalizadas foi formalizada por Matheron (1963), motivado pelo trabalho de Krige (1951). Matheron (1963) denominou essa teoria de geoestatística por considerar se tratar de uma contribuição da estatística para a geologia. O princípio fundamental da geoestatística é o fato da variabilidade (ou dependência) de um atributo estar associada ao espaço físico, isto é, no estudo de atributos do solo tais como resistência do solo à penetração, umidade do solo, dentre outros, ou à produtividade. Os valores encontrados para esses estão relacionados com a posição espacial do local onde eles foram coletados. Diante da impossibilidade de se obterem valores de determinada característica em todos os pontos do espaço, a geoestatística tem como objetivo caracterizar a dependência espacial de variáveis que dependem da posição no espaço, a partir de um conjunto discreto de observações. Modelos espaciais são bastante utilizados na agricultura para caracterizar a dependência espacial de diversos atributos, por exemplo, Borssoi, Uribe-Opazo e Galea (2011) utilizaram modelos espaciais lineares gaussiano para estudar a variabilidade espacial da produtividade. Em um estudo da variablidade espacial da produtividade, Assumpção, Uribe-Opazo e Galea (2014) usaram modelos espaciais lineares baseados na distribuição t-Student, uma distribuição de cauda mais pesada, com o objetivo de reduzir a influência de dados típicos na estimação dos parâmetros e nos mapas temáticos.

Apesar dos estudos geoestatísticos não exigirem que os dados tenham distribuição normal, esta é uma distribuição bastante utilizada. Em casos em que a suposição de normalidade não é satisfeita, muitos pesquisadores utilizam transformações dos dados para que se obtenha a normalidade. O uso de transformações tem apresentado resultados satisfatórios, porém considerar as características do fenômeno sob estudo pode ser mais adequado do que a utilização do modelo normal. Várias distribuições assimétricas têm sido propostas e discutidas na literatura para modelar fenômenos que fornecem dados assimétricos (Exponencial, Gamma, Weibull, Birnbaum-Saunders (BS)).

A distribuição BS tem se mostrado eficiente para modelar dados que assumem valores positivos e cujo comportamento apresenta assimetria positiva e unimodalidade (LEIVA, 2015). No entanto, além dessas características empíricas que alguns dados possuem, é importante considerar os argumentos teóricos que explicam o fenômeno que gera os dados sob análise. A distribuição BS possui argumentos teóricos que permitem descrever adequadamente processos acumulativos no tempo. Esses são induzidos por estresse ou outros fatores, processos que podem ocorrer no estudo de fenômenos agrícolas. Birnbaum e Saunders (1969a, 1969b) desenvolveram tal distribuição motivados por problemas de vibração em aeronaves e, embora a distribuição BS tenha sido proposta para resolver um problema de engenharia, as aplicações da mesma não ficaram restrita a esta área; veja, por exemplo, Leiva, Sanhueza e Angulo (2009)

utilizaram a distribuição BS para estudar a qualidade da água; Leiva et al. (2010) utilizaram a distribuição para desenvolver um estudo sobre poluição do ar e Vilca et al. (2010) desenvolveram um estudo sobre qualidade ambiental na cidade de Santiago a partir da distribuição BS.

Muitos fenômenos podem, ou necessitam (como no caso da teoria das variáveis regionalizáveis), ser descritos por variáveis aleatórias múltiplas, correlacionadas, fazendo-se necessário o uso de distribuições multivariadas. Uma extensão da distribuição BS para o caso bivariado foi proposta por Kundu, Balakrishnan e Jamalizadeh (2010). Kundu, Balakrishnan e Jamalizadeh (2013) definiram e estudaram as principais propriedades de uma versão multivariada da distribuição BS generalizada proposta por Díaz-García e Leiva (2005), da qual a distribuição BS multivariada é um caso particular.

O presente texto foi organizado da seguinte forma: a primeira parte aborda introdução, justificativa e objetivos da tese; na segunda parte apresenta-se uma revisão da literatura utilizada para o desenvolvimento do trabalho. Os resultados são apresentados a partir de três artigos científicos resultantes do projeto de pesquisa desenvolvido como tese de doutorado. Cada artigo pode ser lido de forma independente e têm em comum o uso de modelos espaciais baseados na distribuição assimétrica Birnbaum-Saunders. O Artigo 1 trata de um modelo log-linear Birnbaum-Saunders, cuja média é constante. O Artigo 2 trata da definição de um modelo log-linear Birnbaum-Saunders, cuja média é explicada por um número finito de covariáveis bem como da predição espacial a partir deste modelo. O Artigo 3 aborda a análise de diagnósticos do modelo BS definido no Artigo 2. Por fim, apresenta-se um capítulo com considerações finais sobre toda a tese. É importante destacar que o Artigo 1 foi aceito para publicação na revista Stochastic Environmental Research and Risk Assessment (SERRA), e sua versão *on line* já foi publicada (DOI 10.1007/s00477-015-1204-4).

2 OBJETIVOS

2.1 Objetivo geral

O principal objetivo deste trabalho foi formular um modelo espacial para análise de dados espacialmente correlacionados e que possuam distribuição amostral assimétrica positiva. Objetivou-se ainda utilizar o modelo proposto para estudos relacionados a dados agrícolas.

2.2 Objetivos específicos

Especificamente, os objetivos deste trabalho foram:

- desenvolver uma metodologia que inclui a formulação de um modelo espacial BS e a estimação dos seus parâmetros utilizando o método de máxima verossimilhança (ML);
- definir a distribuição log-BS multivariada, utilizada para a elaboração do modelo, abordando suas principais propriedades;
- desenvolver técnicas de diagnósticos para o modelo espacial BS;
- aplicar técnicas de diagnóstico de influência, que permitem detectar a influência de dados atípicos na estimação dos parâmetros, em conjuntos de dados simulados e em conjuntos de dados reais referentes a experimentos agrícolas;
- a partir da modelagem da estrutura de dependência espacial, construir mapas temáticos, utilizando a krigagem como método de interpolação, para um conjuntos de dados reais;
- verificar se os pontos apontados como influentes exercem influência sobre os parâmetros estimados bem como sobre os mapas temáticos;
- realizar predição espacial com o modelo proposto.

3 REVISÃO BIBLIOGRÁFICA

3.1 Distribuições Birnbaum-Saunders e log-Birnbaum-Saunders

A distribuição normal é a mais utilizada em estudos estatísticos. No entanto, em diversas situações, o modelo que melhor representa a distribuição dos dados não é o modelo normal e a utilização de outros modelos que levem em consideração as características do fenômeno sob estudo é mais adequada. Entre as várias distribuições existentes, a distribuição Birnbaum-Saunders (BS) tem se mostrado bastante eficiente para modelar dados que assumem valores positivos e cujo comportamento apresenta assimetria positiva e unimodalidade. Nesta Seção, o objetivo é apresentar as distribuição Birnbaum-Saunders (BS) e log-Birnbaum-Saunders (log-BS) nos casos univariado, bivariado e multivariado bem como suas principais propriedades.

3.1.1 Distribuição BS univariada

Birnbaum e Saunders (1968), motivados por problemas de vibração em aeronaves, apresentaram um modelo probabilístico para descrever o tempo de vida de materiais expostos a um processo de dano cumulativo. No ano seguinte, Birnbaum e Saunders (1969a, 1969b) formalizaram a distribuição BS e apresentaram um método de estimação para seus dois parâmetros. O modelo BS original considera as seguintes suposições:

- Um material é exposto a uma sequência cíclica de cargas, as quais produzem uma rachadura no mesmo;
- A carga imposta ao material é a mesma a cada ciclo;
- A cada carga C_i o tamanho da rachadura é incrementado por um valor aleatório X_i, cuja distribuição depende da extensão da rachadura no ciclo anterior e das cargas dos ciclos que precedem C_i;
- O tamanho total da rachadura, devido ao i-ésimo ciclo, Y_i, é uma variável aleatória que segue uma distribuição com média μ e variância σ^2;
- Quando o tamanho da rachadura ultrapassa um tamanho limite, ω, ocorre uma falha no material;
- Os tamanhos da rachadura em ciclos diferentes são independentes.

Diante dessas suposições, o tamanho total da rachadura após n ciclos é $S_n = \sum_{i=1}^{n} Y_i$. O modelo BS está interessado em encontrar o menor n, tal que S_n exceda ω. Isto é, está interessado em encontrar o número de ciclos até a ocorrência de uma falha.

Seja N o número de ciclos até que uma falha ocorra. Tem-se que:

$$\mathrm{P}(N \leq n) = \mathrm{P}(N \leq n, S_n > \omega) + \mathrm{P}(N \leq n, S_n \leq \omega) = \mathrm{P}(S_n > \omega) + \mathrm{P}(N \leq n, S_n \leq \omega).$$

Apesar de $\mathrm{P}(N \leq n, S_n \leq \omega) > 0$, Birnbaum & Saunders consideram essa probabilidade desprezível, isto é, consideraram $\mathrm{P}(N \leq n, S_n \leq \omega) = 0$. Além disso, pelo Teorema Central do Limite, quando $n \to \infty, S_n$ tem distribuição normal com média $n\mu$ e variância $n\sigma^2$. E assim:

$$\begin{aligned} \mathrm{P}(N \leq n) \approx \mathrm{P}(S_n > \omega) &= \mathrm{P}\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} > \frac{\omega - n\mu}{\sigma\sqrt{n}}\right) = \mathrm{P}\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq \frac{n\mu - \omega}{\sigma\sqrt{n}}\right) \\ &= \Phi\left(\frac{n\mu - \omega}{\sigma\sqrt{n}}\right) = \Phi\left(\frac{\sqrt{\omega\mu}}{\sigma}\left(\sqrt{\frac{n}{\omega/\mu}} - \sqrt{\frac{\omega/\mu}{n}}\right)\right), \end{aligned} \tag{3.1}$$

em que $\Phi(\cdot)$ denota a função distribuição acumulada (f.d.a.) da distribuição normal padrão.

Birnbaum e Saunders substituíram na equação (3.1) o número de ciclos até a falha (N) pelo tempo total até a falha (T), e o n-ésimo ciclo pelo tempo t e definiram assim uma distribuição contínua de vida com dois parâmetros da seguinte forma:

$$F_T(t; \alpha, \gamma) = \Phi\left(\frac{1}{\alpha}\left(\sqrt{\frac{t}{\gamma}} - \sqrt{\frac{\gamma}{t}}\right)\right),$$

em que $F_T(\cdot)$ representa a f.d.a. da distribuição BS, $\alpha = \frac{\sigma}{\sqrt{\omega\mu}}$ é um parâmetro de forma, $\gamma = \frac{\omega}{\mu}$ é um parâmetro de escala e $\Phi(\cdot)$, como já mencionado anteriormente, é f.d.a. de uma variável aleatória contínua com distribuição normal padrão. Assim, diz-se que uma variável aleatória contínua T tem distribuição BS com parâmetro de forma $\alpha > 0$ e parâmetro de escala $\gamma > 0$, $T \sim \mathrm{BS}(\alpha, \gamma)$, se, e somente se,

$$Z = \frac{1}{\alpha}\left(\sqrt{\frac{t}{\gamma}} - \sqrt{\frac{\gamma}{t}}\right) \sim \mathrm{N}(0, 1).$$

Da maneira como foi definida a distribuição $\mathrm{BS}(\alpha, \gamma)$, tem-se que a função densidade de probabilidade (f.d.p.) de uma variável aleatória contínua com distribuição BS é dada por,

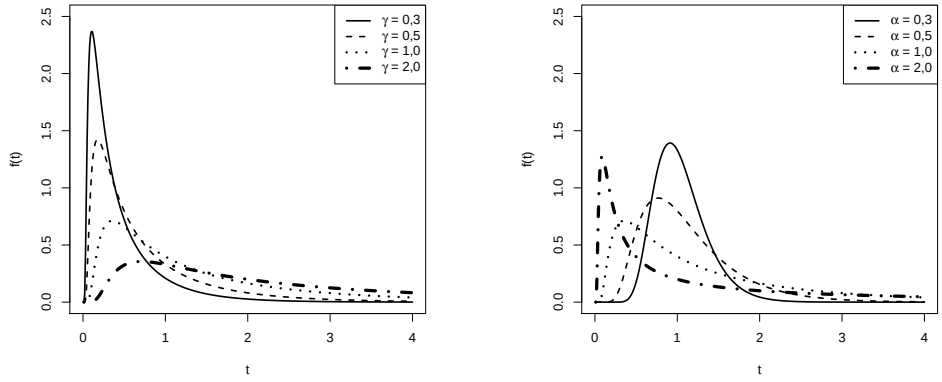
$$f_T(t; \alpha, \gamma) = \frac{dF_T(t; \alpha, \gamma)}{dt} = \phi(u(t; \alpha, \gamma))\frac{d}{dt}u(t; \alpha, \gamma); t > 0, \alpha > 0, \gamma > 0.$$

em que $\phi(\cdot)$ denota a f.d.p. da distribuição normal padrão, e $u(t; \alpha, \gamma) = \frac{1}{\alpha}\left(\sqrt{\frac{t}{\gamma}} - \sqrt{\frac{\gamma}{t}}\right)$.

Portanto,

$$f_T(t; \alpha, \gamma) = \frac{\exp(\alpha^{-2})}{2\alpha\sqrt{2\pi\gamma}}\exp\left\{-\frac{1}{2\alpha^2}\left(\frac{t}{\gamma} + \frac{\gamma}{t}\right)\right\} t^{\frac{-3}{2}}(t + \gamma); t > 0, \alpha > 0, \gamma > 0.$$

Na Figura 1 são apresentadas algumas formas da f.d.p. da distribuição $\mathrm{BS}(\alpha, \gamma)$. Na Figura 1(a), o parâmetro de forma assume o valor fixo $\alpha = 1$ e diferentes valores do parâmetro de escala $\gamma \in \{0,3; 0,5; 1; 2\}$. Na Figura 1(b) tem-se o parâmetro de escala fixo, $\gamma = 1$, e diferentes valores do parâmetro de forma $\alpha \in \{0,3; 0,5; 1; 2\}$. Observa-se que, à medida que α se aproxima de zero, a distribuição torna-se simétrica.



(a) (b)

Figura 1 Funções distribuições de probabilidade das distribuições $\mathrm{BS}(\alpha, 1)$ (a) e $\mathrm{BS}(1, \gamma)$ (b).

3.1.1.1 Propriedades da distribuição BS univariada

A distribuição BS possui várias propriedades interessantes. Muitas dessas propriedades podem ser facilmente demostradas quando se a estreita relação entre as distribuições BS e normal padrão. Algumas dessas propriedades são apresentadas a seguir.

1. Se $T \sim \mathrm{BS}(\alpha, \gamma)$, então, $cT \sim \mathrm{BS}(\alpha, c\gamma), \forall c \in \mathbb{R}_+^*$.

   Demonstração: Seja $T \sim \mathrm{BS}(\alpha, \gamma)$. Então,

   $$\mathrm{P}(cT \leq t) = \mathrm{P}\left(T \leq \frac{1}{c}t\right) = \Phi\left(\frac{1}{\alpha}\left(\sqrt{\frac{\frac{t}{c}}{\gamma}} - \sqrt{\frac{\gamma}{\frac{t}{c}}}\right)\right) = \Phi\left(\frac{1}{\alpha}\left(\sqrt{\frac{t}{c\gamma}} - \sqrt{\frac{c\gamma}{t}}\right)\right),$$

   que é a f.d.a. de uma variável aleatória contínua com distribuição $\mathrm{BS}(\alpha, c\gamma)$.

2. Se $T \sim \mathrm{BS}(\alpha, \gamma)$ então $T^{-1} \sim \mathrm{BS}\left(\alpha, \gamma^{-1}\right)$.

   Demonstração: Seja $T \sim \mathrm{BS}(\alpha, \gamma)$. Então,

   $$\begin{aligned}\mathrm{P}\left(T^{-1} \leq t\right) &= \mathrm{P}\left(T \geq \frac{1}{t}\right) = 1 - P\left(T \leq \frac{1}{t}\right) = 1 - \Phi\left(\frac{1}{\alpha}\left(\sqrt{\frac{\frac{1}{t}}{\gamma}} - \sqrt{\frac{\gamma}{\frac{1}{t}}}\right)\right) \\ &= 1 - \Phi\left(\frac{1}{\alpha}\left(\sqrt{\frac{1}{\gamma t}} - \sqrt{\gamma t}\right)\right) = \Phi\left(\frac{1}{\alpha}\left(\sqrt{\gamma t} - \sqrt{\frac{1}{\gamma t}}\right)\right) \\ &= \Phi\left(\frac{1}{\alpha}\left(\sqrt{\frac{t}{\frac{1}{\gamma}}} - \sqrt{\frac{\frac{1}{\gamma}}{t}}\right)\right),\end{aligned}$$

   que é a f.d.a. de uma variável aleatória com distribuição $\mathrm{BS}\left(\alpha, \gamma^{-1}\right)$.

3. Se $T \sim \mathrm{BS}(\alpha, \gamma)$, então:

   (a) $T = \gamma\left(\frac{\alpha Z}{2} + \sqrt{\left(\frac{\alpha Z}{2}\right)^2 + 1}\right)^2$, em que $Z \sim \mathrm{N}(0, 1)$.

Demonstração: Sabe-se da equação (3.1.1) que $Z = \frac{1}{\alpha}\left(\sqrt{\frac{T}{\gamma}} - \sqrt{\frac{\gamma}{T}}\right) = \frac{1}{\alpha}\left(\frac{T-\gamma}{\sqrt{\gamma T}}\right)$. Logo, $T - \alpha\sqrt{\gamma T}Z = \gamma$, de onde segue que $\left(\sqrt{T} - \frac{\alpha}{2}\sqrt{\gamma}Z\right)^2 = \gamma + \frac{\alpha^2}{4}\gamma Z^2$. E portanto, $T = \gamma\left(\frac{\alpha Z}{2} + \sqrt{\left(\frac{\alpha Z}{2}\right)^2 + 1}\right)^2$.

(b) T é uma função monótona crescente de Z.

Demonstração: Do item (a) tem-se que

$$T'(Z) = 2\gamma\left(\frac{\alpha Z}{2} + \sqrt{\left(\frac{\alpha Z}{2}\right)^2 + 1}\right)\left(\frac{\alpha}{2} + \left(\frac{\alpha}{2}\right)^2 \frac{Z}{\sqrt{\left(\frac{\alpha}{2}Z\right)^2 + 1}}\right).$$

Agora, $\frac{\alpha Z}{2} + \sqrt{\left(\frac{\alpha Z}{2}\right)^2 + 1} > 0$ e $\frac{\alpha}{2} + \left(\frac{\alpha}{2}\right)^2 \frac{Z}{\sqrt{\left(\frac{\alpha}{2}Z\right)^2+1}} > 0$, pois se $\frac{\alpha Z}{2} + \sqrt{\left(\frac{\alpha Z}{2}\right)^2 + 1} \leq 0$ ou $\frac{\alpha}{2} + \left(\frac{\alpha}{2}\right)^2 \frac{Z}{\sqrt{\left(\frac{\alpha}{2}Z\right)^2+1}} \leq 0$ tem-se que $\left(\frac{\alpha Z}{2}\right)^2 \geq \left(\frac{\alpha Z}{2}\right)^2 + 1$. Portanto, $T'(Z) > 0$ e T é uma função estritamente crescente.

4. (Momentos) Se $T \sim \mathrm{BS}(\alpha, \gamma)$ e $r \in \mathbb{N}^*$ então,

(a) $\mathrm{E}(T^r) = \gamma^r \sum_{j=0}^{r} \binom{2r}{2j} \sum_{i=0}^{j} \binom{j}{i} \left(\frac{\alpha}{2}\right)^{2(r-j+i)} \frac{[2(r-j+i)]!}{2^{r-j+i}(r-j+i)!}$

Demonstração: Pela propriedade 3(a) $T^r = \gamma^r\left(\frac{\alpha Z}{2} + \sqrt{\left(\frac{\alpha Z}{2}\right)^2 + 1}\right)^{2r}$. Ao se utilizar o binômio de Newton, tem-se que,

$$T^r = \gamma^r\left(\sum_{i=0}^{2r} \binom{2r}{i} \left(\frac{\alpha Z}{2}\right)^{2r-i} \left(\sqrt{\left(\frac{\alpha Z}{2}\right)^2 + 1}\right)^i\right).$$

Fazendo $i = 2j$ tem-se, $T^r = \gamma^r \sum_{j=0}^{r} \binom{2r}{2j} \left(\frac{\alpha Z}{2}\right)^{2r-2j} \left(\left(\frac{\alpha Z}{2}\right)^2 + 1\right)^j$. Portanto,

$$\begin{aligned}
\mathrm{E}(T^r) &= \gamma^r \sum_{j=0}^{r} \binom{2r}{2j} E\left(\left(\frac{\alpha Z}{2}\right)^{2r-2j} \left(\left(\frac{\alpha Z}{2}\right)^2 + 1\right)^j\right) \\
&= \gamma^r \sum_{j=0}^{r} \binom{2r}{2j} E\left(\left(\frac{\alpha Z}{2}\right)^{2(r-j)} \sum_{i=0}^{j} \binom{j}{i} 1^{j-i} \left(\frac{\alpha Z}{2}\right)^{2i}\right) \\
&= \gamma^r \sum_{j=0}^{r} \binom{2r}{2j} \sum_{i=0}^{j} \binom{j}{i} \left(\frac{\alpha}{2}\right)^{2(r-j+i)} E\left(Z^{2(r-j+i)}\right).
\end{aligned}$$

Como Z tem distribuição normal padrão, prova-se facilmente, por indução sobre n, que para n par, $E(Z^n) = \frac{n!}{2^{\frac{n}{2}}\left(\frac{n}{2}\right)!}$. E, portanto,

$$\mathrm{E}(T^r) = \gamma^r \sum_{j=0}^{r} \binom{2r}{2j} \sum_{i=0}^{j} \binom{j}{i} \left(\frac{\alpha}{2}\right)^{2(r-j+i)} \frac{(2(r-j+i))!}{2^{r-j+i}(r-j+i)!}.$$

(b) $\mathrm{E}(T) = \gamma\left(1 + \frac{\alpha^2}{2}\right)$;

Demonstração: O resultado segue diretamente do item (a), quando $r = 1$.

(c) A variância de T é dada por $\mathrm{Var}(T) = (\alpha\gamma)^2 \left(\frac{5}{4}\alpha^2 + 1\right)$;

Demonstração: O resultado também segue do item (a), lembrando-se que $\mathrm{Var}(T) = \mathrm{E}\left(T^2\right) - \mathrm{E}^2(T) = \gamma^2\left(1 + 2\alpha^2 + \frac{3}{2}\alpha^4\right) - \left(\gamma\left(1 + \frac{\alpha^2}{2}\right)\right)^2 = (\alpha\gamma)^2\left(\frac{5}{4}\alpha^2 + 1\right)$;

(d) O coeficiente de simetria da variável aleatória T é dado por $CS(T) = \frac{4\alpha(6+11\alpha^2)}{\left(\sqrt{5\alpha^2+4}\right)^3}$.

Demonstração: Quando $r = 3$ na equação do item (a), após algumas manipulações algébricas tem-se que $\mathrm{E}\left(T^3\right) = \gamma^3\left(1 + \frac{9}{2}\alpha^2 + 9\alpha^4 + \frac{15}{2}\alpha^6\right)$. Assim, $m_2 = \mathrm{E}\left(\left(T - \mathrm{E}(T)\right)^2\right) = \mathrm{Var}(T) = (\alpha\gamma)^2\left(\frac{5}{4}\alpha^2 + 1\right)$ e $m_3 = \mathrm{E}\left(\left(T - \mathrm{E}(T)\right)^3\right) = \gamma^3\alpha^4\left(\frac{6+11\alpha^2}{2}\right)$. Ao se substituírem as expressões de m_2 e m_3 na equação $CS(T) = \frac{m_3(T)}{m_2(T)^{3/2}}$ chega-se ao resultado;

(e) O coeficiente de curtose da variável aleatória T é dado por $CK(T) = 3 + \frac{2\alpha^2(40+93\alpha^2)}{(4+5\alpha^2)^2}$.

Demonstração: Existem várias formas de cálculo para o coeficiente de curtose. A adotada neste trabalho é $CK(T) = \frac{m_4(T)}{m_2(T)^2}$. Com base no item (a), tem-se que $\mathrm{E}\left(T^4\right) = \gamma^4\left(\frac{105}{2}\alpha^8 + 60\alpha^6 + 30\alpha^4 + 8\alpha^2 + 1\right)$. Portanto, o momento centrado de ordem 4, $m_4(T) = \mathrm{E}\left(T^4\right) - 4\mathrm{E}\left(T^3\right)\mathrm{E}(T) + 6\mathrm{E}^2(T)\mathrm{E}\left(T^2\right) - 3\mathrm{E}^4(T) = 3\gamma^4\alpha^4\left(\frac{16+120\alpha^2+211\alpha}{16}\right)$. Ao se substituírem $m_2(T)$ e $m_4(T)$ na equação $CK(T) = \frac{m_4(T)}{m_2(T)^2}$ segue o resultado.

Como $\lim_{\alpha\to 0} CS(T) = 0$ e $\lim_{\alpha\to 0} CK(T) = 3$, pode-se dizer que quando α se aproxima de zero a distribuição $\mathrm{BS}(\alpha, \gamma)$ se aproxima de uma distribuição normal com média γ, confirmando o observado na Figura 1.

5. A distribuição $\mathrm{BS}(\alpha, \gamma)$ é unimodal para quaisquer $\alpha > 0$ e $\gamma > 0$.

   Demonstração: Para se obter a moda de uma distribuição, deve-se derivar sua função densidade de probabilidade e calcular os zeros dessa derivada, pois a moda são o(s) ponto(s) de máximo da função densidade. Quando se deriva $f_T(t; \alpha, \gamma)$ e iguala-se a derivada a zero chegamos à equação cúbica $-t^3 - (\gamma + \gamma\alpha^2)t^2 + (\gamma^3 - 3\gamma^2\alpha^2)t + \gamma^3 = 0$. A resolução desta equação leva a uma única raiz real. Logo, a distribuição é unimodal.

6. A mediana da distribuição $\mathrm{BS}(\alpha, \gamma)$ é γ.

   Demonstração: Deseja-se encontrar um valor de $t \in \mathbb{R}$ tal que $\mathrm{P}(T \leq t) = 0,5$. Isto é, $t \in \mathbb{R}$ tal que $F_T(t; \alpha, \gamma) = 0,5$. Equivalentemente, deve-se encontrar que $\Phi\left(\frac{1}{\alpha}\left(\sqrt{\frac{t}{\gamma}} - \sqrt{\frac{\gamma}{t}}\right)\right) = 0,5$. Como o percentil 50 da distribuição normal padrão é zero, tem-se que $\sqrt{\frac{t}{\gamma}} - \sqrt{\frac{\gamma}{t}} = 0$, de onde segue que $t = \gamma$.

3.1.2 Distribuição BS bivariada

Kundu, Balakrishnan e Jamalizadeh (2010) definiram a distribuição Birnbaum-Saunders bivariada com cinco parâmetros, $\alpha_1, \gamma_1, \alpha_2, \gamma_2$ e ρ, $\mathrm{BS}_2(\alpha_1, \gamma_1, \alpha_2, \gamma_2; \rho)$, estendendo a distribuição BS definida por Birnbaum e Saunders(1969a). Isto é, diz-se que um vetor

aleatório bivariado contínuo $\boldsymbol{T} = (T_1, T_2)^\top$ tem distribuição Birnbaum-Saunders bivariada $\mathrm{BS}_2(\alpha_1, \gamma_1, \alpha_2, \gamma_2; \rho)$, se e somente se a função distribuição acumulada de $\boldsymbol{T}$ for expressa como

$$F_{\boldsymbol{T}}(t_1, t_2) = P\left(T_1 \leq t_1, T_2 \leq t_2\right) = \Phi_2\left(\frac{1}{\alpha_1}\left(\sqrt{\frac{t_1}{\gamma_1}} - \sqrt{\frac{\gamma_1}{t_1}}\right), \frac{1}{\alpha_2}\left(\sqrt{\frac{t_2}{\gamma_2}} - \sqrt{\frac{\gamma_2}{t_2}}\right); \rho\right),$$

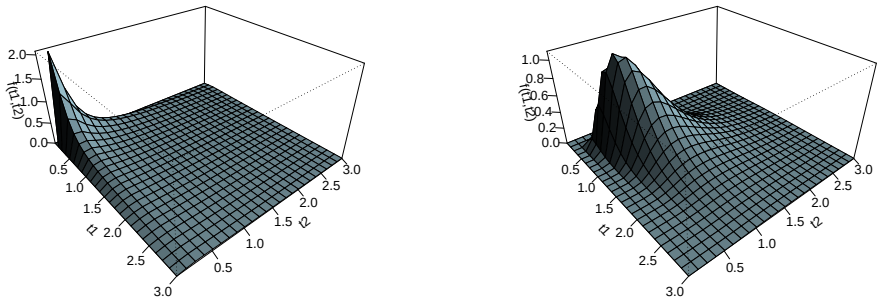
em que $t_1 > 0; t_2 > 0$ e $\Phi_2(u, v; \rho)$ é a f.d.a. de um vetor aleatório normal padrão bivariado com coeficiente de correlação ρ, $-1 < \rho < 1$. De forma similar ao caso univariado $\alpha_1, \alpha_2 > 0$ são parâmetros de forma e $\gamma_1, \gamma_2 > 0$ são parâmetros de escala.

A f.d.p. de um vetor aleatório $\boldsymbol{T} = (T_1, T_2)^\top$ com distribuição $\mathrm{BS}_2(\alpha_1, \alpha_2, \gamma_1, \gamma_2; \rho)$ é dada por

$$f_T(t_1, t_2) = \phi_2\left(\frac{1}{\alpha_1}\left(\sqrt{\frac{t_1}{\gamma_1}} - \sqrt{\frac{\gamma_1}{t_1}}\right), \frac{1}{\alpha_2}\left(\sqrt{\frac{t_2}{\gamma_2}} - \sqrt{\frac{\gamma_2}{t_2}}\right); \rho\right) \prod_{i=1}^{2} \frac{1}{2\alpha_i\gamma_i}\left(\left(\frac{\gamma_i}{t_i}\right)^{\frac{1}{2}} + \left(\frac{\gamma_i}{t_i}\right)^{\frac{3}{2}}\right),$$

em que $\phi_2(u, v; \rho)$ é a f.d.p. de um vetor aleatório normal padrão bivariado com coeficiente de correlação ρ, isto é, $\phi_2(u, v; \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}(u^2 + v^2 - 2\rho uv)\right\}$.

Na Figura 2 são apresentadas algumas formas da f.d.p. da distribuição $\mathrm{BS}_2(\alpha_1, \gamma_1, \alpha_2, \gamma_2; \rho)$.



(a) (b)

Figura 2 Funções distribuições de probabilidade da distribuição $\mathrm{BS}_2(\boldsymbol{\alpha}, \mathbf{1}; 0, 5)$ nos casos $\boldsymbol{\alpha} = (1,5; 1,5)$ (a) e $\boldsymbol{\alpha} = (0,5; 0,5)$ (b).

3.1.2.1 Propriedades do vetor aleatório BS bivariado

1. Se $\boldsymbol{T} = (T_1, T_2)^\top \sim \mathrm{BS}_2(\alpha_1, \alpha_2, \gamma_1, \gamma_2; \rho)$, então $T_i \sim \mathrm{BS}(\alpha_i, \gamma_i), i = 1, 2$.

   Demonstração: Se $\boldsymbol{T} = (T_1, T_2)^\top \sim \mathrm{BS}_2(\alpha_1, \alpha_2, \gamma_1, \gamma_2; \rho)$, então

   $$\left(\frac{1}{\alpha_1}\left(\sqrt{\frac{T_1}{\gamma_1}} - \sqrt{\frac{\gamma_1}{T_1}}\right), \frac{1}{\alpha_2}\left(\sqrt{\frac{T_2}{\gamma_2}} - \sqrt{\frac{\gamma_2}{T_2}}\right)\right)^\top \sim \mathrm{N}_2(\mathbf{0}, \boldsymbol{\rho}),$$

   em que $\boldsymbol{\rho}$ é a matriz de variância e covariância, dada por $\boldsymbol{\rho} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$.

Considere o vetor $\boldsymbol{v} = (1,0)^\top$. Tem-se que

$$\frac{1}{\alpha_1}\left(\sqrt{\frac{T_1}{\gamma_1}} - \sqrt{\frac{\gamma_1}{T_1}}\right) = \boldsymbol{v}^\top \begin{pmatrix} \frac{1}{\alpha_1}\left(\sqrt{\frac{T_1}{\gamma_1}} - \sqrt{\frac{\gamma_1}{T_1}}\right) \\ \\ \frac{1}{\alpha_2}\left(\sqrt{\frac{T_2}{\gamma_2}} - \sqrt{\frac{\gamma_2}{T_2}}\right) \end{pmatrix}$$

Portanto, $\frac{1}{\alpha_1}\left(\sqrt{\frac{T_1}{\gamma_1}} - \sqrt{\frac{\gamma_1}{T_1}}\right) \sim \mathrm{N}(\boldsymbol{v}^\top\boldsymbol{0}, \boldsymbol{v}^\top\boldsymbol{\rho}\boldsymbol{v}) = \mathrm{N}(0,1)$. Logo, por definição, $T_1 \sim \mathrm{BS}(\alpha_1, \gamma_1)$.

Considerando $\boldsymbol{v} = (0,1)^\top$ prova-se, de maneira análoga, que $T_2 \sim \mathrm{BS}(\alpha_2, \gamma_2)$.

2. Se $\boldsymbol{T} = (T_1, T_2)^\top \sim \mathrm{BS}_2(\alpha_1, \alpha_2, \gamma_1, \gamma_2; \rho)$ então:

   (a) $(c_1T_1, c_2T_2)^\top \sim \mathrm{BS}_2(\alpha_1, \alpha_2, c_1\gamma_1, c_2\gamma_2; \rho)$.
   Demonstração: Considere $\boldsymbol{T} = (T_1, T_2)^\top$ e $\boldsymbol{Y} = (c_1T_1, c_2T_2)^\top$. Ao se utilizar a propriedade de funções de vetores aleatórios tem-se que a f.d.p. de $\boldsymbol{Y}$ é da forma $f_{\boldsymbol{Y}}(t_1,t_2) = f_{\boldsymbol{T}}\left(c_1^{-1}t_1, c_2^{-1}t_2\right)\det(\boldsymbol{C})$, em que $\boldsymbol{C} = \begin{pmatrix} c_1^{-1} & 0 \\ 0 & c_2^{-1} \end{pmatrix}$. Assim,

$$\begin{aligned} f_{\boldsymbol{Y}}(t_1,t_2) &= \frac{1}{c_1c_2}\phi_2\left(\frac{1}{\alpha_1}\left(\sqrt{\frac{\frac{t_1}{c_1}}{\gamma_1}} - \sqrt{\frac{\gamma_1}{\frac{t_1}{c_1}}}\right), \frac{1}{\alpha_2}\left(\sqrt{\frac{\frac{t_2}{c_2}}{\gamma_2}} - \sqrt{\frac{\gamma_2}{\frac{t_2}{c_2}}}\right); \rho\right)\prod_{i=1}^{2}\frac{1}{2\alpha_i\gamma_i}\left(\left(\frac{\gamma_i}{\frac{t_i}{c_i}}\right)^{\frac{1}{2}} + \left(\frac{\gamma_i}{\frac{t_i}{c_i}}\right)^{\frac{3}{2}}\right) \\ &= \phi_2\left(\frac{1}{\alpha_1}\left(\sqrt{\frac{t_1}{c_1\gamma_1}} - \sqrt{\frac{c_1\gamma_1}{t_1}}\right), \frac{1}{\alpha_2}\left(\sqrt{\frac{t_2}{c_2\gamma_2}} - \sqrt{\frac{c_2\gamma_2}{t_2}}\right); \rho\right)\prod_{i=1}^{2}\frac{1}{2\alpha_ic_i\gamma_i}\left(\left(\frac{c_i\gamma_i}{t_i}\right)^{\frac{1}{2}} + \left(\frac{c_i\gamma_i}{t_i}\right)^{\frac{3}{2}}\right), \end{aligned}$$

   que é a f.d.p. de um vetor aleatório com distribuição $\mathrm{BS}_2(\alpha_1, \alpha_2, c_1\gamma_1, c_2\gamma_2; \rho)$.

   Seguem como casos particulares as seguintes propriedades:

   (b) $(cT_1, T_2)^\top \sim \mathrm{BS}_2(\alpha_1, \alpha_2, c\gamma_1, \gamma_2; \rho)$.

   (c) $(T_1, cT_2)^\top \sim \mathrm{BS}_2(\alpha_1, \alpha_2, \gamma_1, c\gamma_2; \rho)$.

3. Se $\boldsymbol{T} = (T_1, T_2)^\top \sim \mathrm{BS}_2(\alpha_1, \alpha_2, \gamma_1, \gamma_2; \rho)$ então:

   (a) $(T_1^{-1}, T_2^{-1})^\top \sim \mathrm{BS}_2(\alpha_1, \alpha_2, \gamma_1^{-1}, \gamma_2^{-1}; \rho)$.
   Demonstração: Considere $\boldsymbol{T} = (T_1, T_2)^\top$ e $\boldsymbol{Y} = (T_1^{-1}, T_2^{-1})^\top$. A f.d.p. de $\boldsymbol{Y}$ é dada por $f_{\boldsymbol{Y}}(t_1,t_2) = f_{\boldsymbol{T}}\left(\frac{1}{t_1}, \frac{1}{t_2}\right)\det(\boldsymbol{C})$, neste caso $\boldsymbol{C} = \begin{pmatrix} \frac{-1}{t_1^2} & 0 \\ 0 & \frac{-1}{t_2^2} \end{pmatrix}$. Logo,

$$\begin{aligned} f_{\boldsymbol{Y}}(t_1,t_2) &= \phi_2\left(\frac{1}{\alpha_1}\left(\sqrt{\frac{\frac{1}{t_1}}{\beta_1}} - \sqrt{\frac{\beta_1}{\frac{1}{t_1}}}\right), \frac{1}{\alpha_2}\left(\sqrt{\frac{\frac{1}{t_2}}{\beta_2}} - \sqrt{\frac{\beta_2}{\frac{1}{t_2}}}\right); \rho\right)\prod_{i=1}^{2}\frac{1}{2\alpha_i\beta_i}\left(\left(\frac{\beta_i}{\frac{1}{t_i}}\right)^{\frac{1}{2}} + \left(\frac{\beta_i}{\frac{1}{t_i}}\right)^{\frac{3}{2}}\right)\frac{1}{t_i^2} \\ &= \phi_2\left(\frac{-1}{\alpha_1}\left(\sqrt{\frac{t_1}{\frac{1}{\beta_1}}} - \sqrt{\frac{\frac{1}{\beta_1}}{t_1}}\right), \frac{-1}{\alpha_2}\left(\sqrt{\frac{t_2}{\frac{1}{\beta_2}}} - \sqrt{\frac{\frac{1}{\beta_2}}{t_2}}\right); \rho\right)\prod_{i=1}^{2}\frac{1}{2\alpha_i\frac{1}{\beta_i}}\left(\beta_i^{\frac{-3}{2}}t_i^{\frac{-3}{2}} + \beta_i^{\frac{-1}{2}}t_i^{\frac{-1}{2}}\right). \end{aligned}$$

   Observando que $\phi_2(-u,-v;\rho) = \phi_2(u,v;\rho)$, tem-se:

$$f_Y(t_1,t_2) = \phi_2\left(\frac{1}{\alpha_1}\left(\sqrt{\frac{t_1}{\frac{1}{\beta_1}}} - \sqrt{\frac{\frac{1}{\beta_1}}{t_1}}\right), \frac{1}{\alpha_2}\left(\sqrt{\frac{t_2}{\frac{1}{\beta_2}}} - \sqrt{\frac{\frac{1}{\beta_2}}{t_2}}\right); \rho\right)\prod_{i=1}^{2}\frac{1}{2\alpha_i\frac{1}{\beta_i}}\left(\left(\frac{\frac{1}{\beta_i}}{t_i}\right)^{\frac{3}{2}} + \left(\frac{\frac{1}{\beta_i}}{t_i}\right)^{\frac{1}{2}}\right),$$

   que é a f.d.p. de um vetor aleatório com distribuição $\mathrm{BS}_2(\alpha_1, \alpha_2, \beta_1^{-1}, \beta_2^{-1}; \rho)$.

(b) $\left(T_1^{-1}, T_2\right)^\top \sim \mathrm{BS}_2(\alpha_1, \alpha_2, \gamma_1^{-1}, \gamma_2; -\rho)$.

Demonstração: Neste caso, considere $\boldsymbol{Y} = \left(T_1^{-1}, T_2\right)^\top$. A f.d.p. de $\boldsymbol{Y}$ é

$$
\begin{aligned}
f_{\boldsymbol{Y}}(t_1,t_2) = f_{\boldsymbol{T}}\left(\frac{1}{t_1}, t_2\right)\frac{1}{t_1^2} &= \phi_2\left(\frac{1}{\alpha_1}\left(\sqrt{\frac{\frac{1}{t_1}}{\gamma_1}} - \sqrt{\frac{\gamma_1}{\frac{1}{t_1}}}\right), \frac{1}{\alpha_2}\left(\sqrt{\frac{t_2}{\gamma_2}} - \sqrt{\frac{\gamma_2}{t_2}}\right); \rho\right) \\
&\left(\frac{1}{2\alpha_1\gamma_1}\left(\left(\frac{\gamma_1}{\frac{1}{t_1}}\right)^{\frac{1}{2}} + \left(\frac{\gamma_1}{\frac{1}{t_1}}\right)^{\frac{3}{2}}\right)\right)\left(\frac{1}{2\alpha_2\gamma_2}\left(\left(\frac{\gamma_2}{t_2}\right)^{\frac{1}{2}} + \left(\frac{\gamma_2}{t_2}\right)^{\frac{3}{2}}\right)\right)\frac{1}{t_1^2} \\
&= \phi_2\left(\frac{-1}{\alpha_1}\left(\sqrt{\frac{t_1}{\frac{1}{\gamma_1}}} - \sqrt{\frac{\frac{1}{\gamma_1}}{t_1}}\right), \frac{1}{\alpha_2}\left(\sqrt{\frac{t_2}{\gamma_2}} - \sqrt{\frac{\gamma_2}{t_2}}\right); \rho\right)\left(\frac{1}{2\alpha_1\frac{1}{\gamma_1}}\left(\left(\frac{\frac{1}{\gamma_1}}{t_1}\right)^{\frac{1}{2}} + \left(\frac{\frac{1}{\gamma_1}}{t_1}\right)^{\frac{3}{2}}\right)\right) \\
&\left(\frac{1}{2\alpha_2\gamma_2}\left(\left(\frac{\gamma_2}{t_2}\right)^{\frac{1}{2}} + \left(\frac{\gamma_2}{t_1}\right)^{\frac{3}{2}}\right)\right).
\end{aligned}
$$

E quando se observa que $\phi_2(-u, v; \rho) = \phi_2(u, -v; \rho) = \phi_2(u, v; -\rho)$, tem-se:

$$
\begin{aligned}
f_Y(t_1,t_2) = \phi_2&\left(\frac{1}{\alpha_1}\left(\sqrt{\frac{t_1}{\frac{1}{\gamma_1}}} - \sqrt{\frac{\frac{1}{\gamma_1}}{t_1}}\right), \frac{1}{\alpha_2}\left(\sqrt{\frac{t_2}{\gamma_2}} - \sqrt{\frac{\gamma_2}{t_2}}\right); -\rho\right) \\
&\left(\frac{1}{2\alpha_1\frac{1}{\gamma_1}}\left(\left(\frac{\frac{1}{\gamma_1}}{t_1}\right)^{\frac{1}{2}} + \left(\frac{\frac{1}{\gamma_1}}{t_1}\right)^{\frac{3}{2}}\right)\right)\left(\frac{1}{2\alpha_2\gamma_2}\left(\left(\frac{\gamma_2}{t_2}\right)^{\frac{1}{2}} + \left(\frac{\gamma_2}{t_1}\right)^{\frac{3}{2}}\right)\right).
\end{aligned}
$$

que é a f.d.p. de um vetor aleatório com distribuição $\mathrm{BS}_2(\alpha_1, \alpha_2, \gamma_1^{-1}, \gamma_2; -\rho)$.

(c) $\left(T_1, T_2^{-1}\right)^\top \sim \mathrm{BS}_2(\alpha_1, \gamma_1, \alpha_2, \gamma_2^{-1}; -\rho)$.

Demonstração: A demonstração é análoga à fornecida no item (b) considerando $\boldsymbol{Y} = \left(T_1, T_2^{-1}\right)^\top$.

4. (Momentos)

(a) Se $\boldsymbol{T} = (T_1, T_2)^\top \sim \mathrm{BS}_2(\alpha_1, \alpha_2, \gamma_1, \gamma_2; \rho)$ então, $\boldsymbol{E}(\boldsymbol{T}) = \begin{pmatrix} \gamma_1\left(1 + \frac{\alpha_1^2}{2}\right) \\ \\ \gamma_2\left(1 + \frac{\alpha_2^2}{2}\right) \end{pmatrix}$.

Demonstração: $\boldsymbol{E}(\boldsymbol{T}) = (E(T_1), E(T_2))^\top$. Se $T = (T_1, T_2)^\top \sim \mathrm{BS}_2(\alpha_1, \alpha_2, \gamma_1, \gamma_2; \rho)$. Assim, a demonstração segue das propriedades 1 e 4(b) do caso univariado.

(b) A matriz de covariância de $\boldsymbol{T} = (T_1, T_2)^\top$ é dada por

$$
\boldsymbol{\Sigma} = \begin{pmatrix} \alpha_1^2\gamma_1^2\left(\frac{5}{4}\alpha_1^2 + 1\right) & \rho_{\boldsymbol{T}} \\ \rho_{\boldsymbol{T}} & \alpha_2^2\gamma_2^2\left(\frac{5}{4}\alpha_2^2 + 1\right) \end{pmatrix},
$$

em que $\rho_{\boldsymbol{T}} = \gamma_1\gamma_2\left(\left(2 + \frac{3\alpha_1^2}{2} + \frac{3\alpha_2^2}{2} + \frac{9\alpha_1^2\alpha_2^2}{2^5}\right)\rho + \frac{\alpha_1^2\alpha^2}{2}\rho^2 + \frac{3\alpha_1^2\alpha^2}{2^4}\rho^3 + r(Z_1, Z_2)\right)$ e

$r(Z_1, Z_2)$ é dado por

$$\begin{aligned} r(Z_1,Z_2) &= \sum_{k=2}^{\infty}(-1)^{k-1}\frac{1\cdot 3\cdots(2k-3)}{2^{3k}k!}\alpha_2^{2k}\mathrm{E}\left(Z_1 Z_2^{2k+1}\right) \\ &+ \sum_{k=2}^{\infty}(-1)^{k-1}\frac{1\cdot 3\cdots(2k-3)}{2^{3k}k!}\alpha_1^{2k}\mathrm{E}\left(Z_2 Z_1^{2k+1}\right) \\ &+ \frac{\alpha_1^2}{2^3}\sum_{k=2}^{\infty}(-1)^{k-1}\frac{1\cdot 3\cdots(2k-3)}{2^{3k}k!}\alpha_2^{2k}\mathrm{E}\left(Z_1^3 Z_2^{2k+1}\right) \\ &+ \frac{\alpha_2^2}{2^3}\sum_{k=2}^{\infty}(-1)^{k-1}\frac{1\cdot 3\cdots(2k-3)}{2^{3k}k!}\alpha_1^{2k}\mathrm{E}\left(Z_2^3 Z_1^{2k+1}\right) \\ &+ \sum_{k=2}^{\infty}\sum_{p=2}^{\infty}(-1)^{k+p-2}\frac{1\cdot 3\cdots(2k-3)}{2^{3k}k!}\frac{1\cdot 3\cdots(2p-3)}{2^{3p}p!}\alpha_1^{2k}\alpha_2^{2p}\mathrm{E}\left(Z_1^{2k+1}Z_1^{2p+1}\right). \end{aligned}$$

Demonstração: Sabe-se pela propriedade 1 que $T_i \sim \mathrm{BS}(\alpha_i, \gamma_i)$, $i = 1, 2$ e assim pela propriedade 4(b) da Seção 3.1.1.1 tem-se que $\mathrm{Var}(T_i) = (\alpha_i\gamma_i)^2\left(\frac{5}{4}\alpha_i^2 + 1\right), i = 1, 2$. A covariância entre T_1 e T_2 é dada por $\mathrm{Cov}(T_1, T_2) = \mathrm{E}(T_1T_2) - \mathrm{E}(T_1)\mathrm{E}(T_2)$. Tem-se ainda $T_i = \gamma_i\left(\frac{\alpha_i}{2}Z_i + \sqrt{\left(\frac{\alpha_i}{2}Z_i\right)^2 + 1}\right)^2,\ i = 1, 2$. Portanto,

$$\begin{aligned} \frac{\mathrm{E}(T_1T_2)}{\gamma_1\gamma_2} &= \mathrm{E}\left(\left(\frac{\alpha_1}{2}Z_1 + \sqrt{\left(\frac{\alpha_1}{2}Z_1\right)^2 + 1}\right)^2 \times \left(\frac{\alpha_2}{2}Z_2 + \sqrt{\left(\frac{\alpha_2}{2}Z_2\right)^2 + 1}^{\,2}\right)\right) \\ &= \mathrm{E}\left(\left(\frac{1}{2}\alpha_1^2Z_1^2 + 1 + \alpha_1Z_1\sqrt{\left(\frac{\alpha_1}{2}Z_1\right)^2 + 1}\right) \times \left(\frac{1}{2}\alpha_2^2Z_2^2 + 1 + \alpha_2Z_2\sqrt{\left(\frac{\alpha_2}{2}Z_2\right)^2 + 1}\right)\right) \\ &= 1 + \frac{\alpha_1^2\alpha_2^2}{4}\mathrm{E}(Z_1^2Z_2^2) + \frac{\alpha_1^2}{2}\mathrm{E}(Z_1^2) + \frac{\alpha_2^2}{2}\mathrm{E}(Z_2^2) + \alpha_2\mathrm{E}\left(Z_2\sqrt{\left(\frac{\alpha_2}{2}Z_2\right)^2 + 1}\right) \\ &\quad + \alpha_1\mathrm{E}\left(Z_1\sqrt{\left(\frac{\alpha_1}{2}Z_1\right)^2 + 1}\right) + \frac{\alpha_1^2\alpha_2}{2}\mathrm{E}\left(Z_1^2Z_2\sqrt{\left(\frac{\alpha_2}{2}Z_2\right)^2 + 1}\right) \\ &\quad + \frac{\alpha_2^2\alpha_1}{2}\mathrm{E}\left(Z_2^2Z_1\sqrt{\left(\frac{\alpha_1}{2}Z_1\right)^2 + 1}\right) + \alpha_1\alpha_2\mathrm{E}\left(Z_1Z_2\sqrt{\left(\frac{\alpha_2}{2}Z_2\right)^2 + 1}\sqrt{\left(\frac{\alpha_1}{2}Z_1\right)^2 + 1}\right) \end{aligned}$$

Sabe-se que a função geradora de momentos da distribuição normal padrão bivariada $m(t_1, t_2) = \exp\left\{\frac{1}{2}\left(t_1^2 + 2t_1t_2\rho + t_2^2\right)\right\}$, e que para se obter o momento conjunto $\mathrm{E}(T_1^rT_2^s)$ diferencia-se $m(t_1, t_2)$ r vezes em relação a t_1 e s vezes em relação a t_2 e faz-se t_1 e t_2 iguais a zero (MOOD; GRAYBILL; BOES, 1974). Os resultados desses cálculos monstram que $\mathrm{E}(Z_1^2) = \mathrm{E}(Z_2^2) = 1$ e $\mathrm{E}(Z_1^2Z_2^2) = 1 + 2\rho^2$. Além disso, uma vez que $Z_i\sqrt{\left(\frac{\alpha_i}{2}Z_i\right)^2 + 1}$ é uma função ímpar, sabe-se que $\mathrm{E}\left(Z_i\sqrt{\left(\frac{\alpha_i}{2}Z_i\right)^2 + 1}\right) = \mathrm{E}\left(Z_j^2Z_i\sqrt{\left(\frac{\alpha_i}{2}Z_i\right)^2 + 1}\right) = 0$ para $i, j = 1, 2$ (KUNDU; BALAKRISHNAN; JAMALIZADEH, 2010). Portanto,

$$\frac{\mathrm{E}(T_1T_2)}{\gamma_1\gamma_2} = \left(1 + \frac{\alpha_1^2\alpha_2^2}{4}(1 + 2\rho^2) + \frac{1}{2}\left(\alpha_1^2 + \alpha_2^2\right) + \alpha_1\alpha_2\mathrm{E}\left(Z_1Z_2\sqrt{\left(\frac{\alpha_2}{2}Z_2\right)^2 + 1}\sqrt{\left(\frac{\alpha_1}{2}Z_1\right)^2 + 1}\right)\right).$$

Utilizando o Teorema Binomial pode-se escrever,

$$Z_i\sqrt{\left(\frac{\alpha_i}{2}Z_i\right)^2 + 1} = Z_i + \frac{\alpha_i^2}{2^3}Z_i^3 + \sum_{k=2}^{\infty}(-1)^{k-1}\frac{1\cdot 3\cdots(2k-3)}{2^{3k}k!}\alpha_i^{2k}Z_i^{2k+1},\ i = 1, 2.$$

E ainda quando se utiliza

$$\mathrm{E}\left(Z_i^{2m+1}Z_j^{2n+1}\right)=\frac{(2m+1)!(2n+1)!}{2^m n}\sum_{p=0}^{\min(m,n)}\frac{(2\rho)^{2p+1}}{(m-p)!(n-p)!(2p+1)!},\ i=1,2.$$

tem-se que,

$$\mathrm{E}\left(Z_1Z_2\sqrt{\left(\frac{\alpha_2}{2}Z_2\right)^2+1}\sqrt{\left(\frac{\alpha_1}{2}Z_1\right)^2+1}\right)=\left(2+\frac{3\alpha_2^2}{4}+\frac{3\alpha_1^2}{4}+\frac{9\alpha_1^2\alpha_2^2}{2^5}\right)\rho+\frac{3\alpha_1^2\alpha_2^2}{2^4}\rho^3+r(Z_1,Z_2),$$

em que $r(Z_1,Z_2)$ é dado como no enunciado.

Desta forma, tem-se

$$\mathrm{Cov}(T_1,T_2)=\gamma_1\gamma_2\left(\left(2+\frac{3\alpha_1^2}{2}+\frac{3\alpha_2^2}{2}+\frac{9\alpha_1^2\alpha_2^2}{2^5}\right)\rho+\frac{\alpha_1^2\alpha^2}{2}\rho^2+\frac{3\alpha_1^2\alpha^2}{2^4}\rho^3+r(Z_1,Z_2)\right).$$

(c) Se $\boldsymbol{T}=(T_1,T_2)^\top\sim\mathrm{BS}_2(\alpha_1,\alpha_2,\gamma_1,\gamma_2;\rho)$ então, os momentos de ordem 3 e de ordem 4 são dados por $\boldsymbol{E(T^3)}=\left(\gamma_1^3\left(\frac{15}{2}\alpha_1^6+9\alpha_1^4+\frac{9}{2}\alpha_1^2+1\right),\gamma_2^3\left(\frac{15}{2}\alpha_2^6+9\alpha_2^4+\frac{9}{2}\alpha_2^2+1\right)\right)^\top$ e $\boldsymbol{E(T^4)}=\left(\gamma_1^4\left(\frac{105}{2}\alpha_1^8+60\alpha_1^6+30\alpha_1^4+8\alpha_1^2+1\right),\gamma_2^4\left(\frac{105}{2}\alpha_2^8+60\alpha_2^6+30\alpha_2^4+8\alpha_2^2+1\right)\right)^\top$. Demonstração: Os momentos de ordem 3 e de ordem 4 podem ser obtidos observando-se que $\boldsymbol{E(T^r)}=(E(T_1)^r,E(T_2)^r)^\top$ e utilizando-se as propriedades 1, 4 ((d) e 4 (e)) da Seção 3.1.1.1.

3.1.3 Distribuição BS multivariada

Kundu, Balakrishnan e Jamalizadeh (2013) definiram e mostraram algumas propriedades da distribuição BS multivariada generalizada, ao considerarem as distribuições da família de distribuições elípticas. Nesta Seção, apresenta-se a distribuição BS multivariada, um caso particular do estudo de Kundu, Balakrishnan e Jamalizadeh (2013) quando se considera a distribuição normal.

Um vetor aleatório multivariado $\boldsymbol{T}=(T_1,\ldots,T_n)^\top$ tem distribuição BS multivariada com parâmetros $\boldsymbol{\alpha}=(\alpha_1,\ldots,\alpha_n)^\top$ e $\boldsymbol{\gamma}=(\gamma_1,\ldots,\gamma_n)^\top$, se a f.d.a. de $\boldsymbol{T}$ for expressa por

$$F_{\boldsymbol{T}}(\boldsymbol{t})=P\left(T_1\leq t_1,\cdots,T_n\leq t_n\right)=\Phi_n\left(\frac{1}{\alpha_1}\left(\sqrt{\frac{t_1}{\gamma_1}}-\sqrt{\frac{\gamma_1}{t_1}}\right),\cdots,\frac{1}{\alpha_n}\left(\sqrt{\frac{t_n}{\gamma_n}}-\sqrt{\frac{\gamma_n}{t_n}}\right);\boldsymbol{\Sigma}\right),$$

em que $\boldsymbol{t}=(t_1,\ldots,t_n)^\top,t_i>0;\alpha_i>0;\gamma_i>0,1\leq i\leq n$ e $\Phi_n\left(\boldsymbol{u};\boldsymbol{\Sigma}\right)$, $\boldsymbol{u}=(u_1,\ldots,u_n)^\top$ é a f.d.a. do vetor normal padrão multivariado com matriz de variância e covariância $\boldsymbol{\Sigma}$, isto é, o vetor de médias da distribuição normal é o vetor nulo e $\boldsymbol{\Sigma}=\begin{cases}1 & \text{se}\quad i=j\\ \rho_{ij} & \text{se}i\neq j\end{cases}$

A notação utilizada neste caso é $\boldsymbol{T}\sim\mathrm{BS}_n\left(\boldsymbol{\alpha},\boldsymbol{\gamma};\boldsymbol{\Sigma}\right)$.

A f.d.p. de um vetor aleatório $\boldsymbol{T}=(T_1,\ldots,T_n)^\top$ com distribuição $\mathrm{BS}_n\left(\boldsymbol{\alpha},\boldsymbol{\gamma};\boldsymbol{\Sigma}\right)$ é dada por

$$\phi_n\left(\frac{1}{\alpha_1}\left(\sqrt{\frac{t_1}{\gamma_1}}-\sqrt{\frac{\gamma_1}{t_1}}\right),\ldots,\frac{1}{\alpha_n}\left(\sqrt{\frac{t_n}{\gamma_n}}-\sqrt{\frac{\gamma_n}{t_n}}\right);\boldsymbol{\Sigma}\right)\prod_{i=1}^{n}\frac{1}{2\alpha_i\gamma_i}\left(\left(\frac{\gamma_i}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\gamma_i}{t_i}\right)^{\frac{3}{2}}\right).$$

em que $\phi_n(\boldsymbol{u};\boldsymbol{\Sigma})$ é a f.d.p. do vetor normal padrão n-variado com matriz de covariância $\boldsymbol{\Sigma}$, isto

é, $\phi_n(u;\boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{n/2}(det(\boldsymbol{\Sigma}))^{1/2}} \exp\left\{-\frac{\boldsymbol{u}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{u}}{2}\right\}$.

3.1.3.1 Propriedades do vetor aleatório BS multivariado

1. Se $\boldsymbol{T} \sim \mathrm{BS}_n(\boldsymbol{\alpha},\boldsymbol{\gamma};\Sigma)$, então $T_i \sim \mathrm{BS}(\alpha_i,\gamma_i),\ \forall\, 1 \leq i \leq n$.

   Demonstração: A demonstração é análoga a da propriedade 1 da Seção 3.1.2.1 considerando $\boldsymbol{v} = (v_1,\ldots,v_n)^\top$, com $v_k = \begin{cases} 1 & \text{se} \quad k = i \\ 0 & \text{se} \quad k \neq i \end{cases}$.

2. Sejam $p = (p_1,\ldots,p_n)$ uma permutação de $(1,\ldots,n)$ e $\boldsymbol{b_p} = (b_{p_1},\cdots,b_{p_n})^\top$ o vetor obtido permutando o vetor $\boldsymbol{b} = (b_1,\ldots,b_n)^\top$ segundo a permutação p. Seja ainda para toda matriz $\boldsymbol{B}$, $\boldsymbol{B_{pp}}$ a matriz cujas linhas e colunas foram permutadas segundo a permutação p. Se $\boldsymbol{T} \sim \mathrm{BS}_n(\boldsymbol{\alpha},\boldsymbol{\gamma};\Sigma)$, então $\boldsymbol{T_p} \sim \mathrm{BS}_n(\boldsymbol{\alpha_p},\boldsymbol{\gamma_p};\boldsymbol{\Sigma_{pp}})$.

   Demonstração: Seja $\boldsymbol{I_n}$ a matriz identidade $n \times n$ e $p = (p_1,\ldots,p_n)$ uma permutação de $(1,\ldots,n)$. Seja $\boldsymbol{A}$ a matriz obtida de $\boldsymbol{I_n}$ por meio da permutação p, isto é, $\boldsymbol{A}$ é matriz obtida ao se aplicar a permutação p nas linhas de $\boldsymbol{I_n}$.Tem-se que $\boldsymbol{AT} = \boldsymbol{T_p}$. Como $\boldsymbol{T} \sim \mathrm{BS}_n(\boldsymbol{\alpha},\boldsymbol{\gamma};\boldsymbol{\Sigma})$ tem-se por definição que

$$\boldsymbol{u} = \left(\frac{1}{\alpha_1}\left(\sqrt{\frac{T_1}{\gamma_1}} - \sqrt{\frac{\gamma_1}{T_1}}\right),\ldots,\frac{1}{\alpha_n}\left(\sqrt{\frac{T_n}{\gamma_n}} - \sqrt{\frac{\gamma_n}{T_n}}\right)\right) \sim \mathrm{N}_n(\boldsymbol{0},\boldsymbol{\Sigma}).$$

   Assim, $\boldsymbol{u_p} = \boldsymbol{Au} \sim \mathrm{N}_n\left(\boldsymbol{A0},\boldsymbol{A\Sigma A}^\top\right) = \mathrm{N}_n(\boldsymbol{0},\boldsymbol{\Sigma}_{pp})$, de onde segue que $\boldsymbol{T_p} \sim \mathrm{BS}_n(\boldsymbol{\alpha_p},\boldsymbol{\gamma_p},\boldsymbol{\Sigma_{pp}})$.

   Para as próximas demonstrações, consideram-se $\boldsymbol{T} \sim \mathrm{BS}_n(\boldsymbol{\alpha},\boldsymbol{\gamma};\boldsymbol{\Sigma})$ e as seguintes partições em $\boldsymbol{T},\boldsymbol{\alpha},\boldsymbol{\gamma}$ e $\boldsymbol{\Sigma}$.

   $\boldsymbol{T} = \left(\boldsymbol{T_1}^\top,\boldsymbol{T_2}^\top\right)^\top$, $\boldsymbol{\alpha} = \left(\boldsymbol{\alpha_1}^\top,\boldsymbol{\alpha_2}^\top\right)^\top$, $\boldsymbol{\gamma} = \left(\boldsymbol{\gamma_1}^\top,\boldsymbol{\gamma_2}^\top\right)^\top$ e $\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma_{11}} & \boldsymbol{\Sigma_{12}} \\ \boldsymbol{\Sigma_{21}} & \boldsymbol{\Sigma_{22}} \end{pmatrix}$

   em que $\boldsymbol{T_1},\boldsymbol{\alpha_1},\boldsymbol{\gamma_1}$ são vetores $q \times 1$; $\boldsymbol{T_2},\boldsymbol{\alpha_2},\boldsymbol{\gamma_2}$ são vetores $p \times 1$; $\boldsymbol{\Sigma_{11}}$ é uma matriz $q \times q$; $\boldsymbol{\Sigma_{22}}$ é uma matriz $p \times p$; $\boldsymbol{\Sigma_{12}}$ é uma matriz $q \times p$ e $\boldsymbol{\Sigma_{21}} = \boldsymbol{\Sigma_{12}}^\top$, com $p + q = n$.

3. Se $\boldsymbol{T} = \left(\boldsymbol{T_1}^\top,\boldsymbol{T_2}^\top\right)^\top \sim \mathrm{BS}_n(\boldsymbol{\alpha},\boldsymbol{\gamma};\boldsymbol{\Sigma})$, então:

   (a) $\boldsymbol{T_1} \sim \mathrm{BS}_q(\boldsymbol{\alpha_1},\boldsymbol{\gamma_1};\boldsymbol{\Sigma_{11}})$.

   Demonstração: Considere $\boldsymbol{A} = [a_{ij}]$ uma matriz $q \times n$ em que $a_{ij} = \begin{cases} 1 & \text{se} \quad i = j \\ 0 & \text{se} \quad i \neq j \end{cases}$.
   Observa-se que $\boldsymbol{AT} = \boldsymbol{T_1}$. Agora, como $\boldsymbol{T} \sim \mathrm{BS}_n(\boldsymbol{\alpha},\boldsymbol{\gamma};\boldsymbol{\Sigma})$ tem-se que

$$\left(\frac{1}{\alpha_1}\left(\sqrt{\frac{t_1}{\gamma_1}} - \sqrt{\frac{\gamma_1}{t_1}}\right),\cdots,\frac{1}{\alpha_n}\left(\sqrt{\frac{t_n}{\gamma_n}} - \sqrt{\frac{\gamma_n}{t_n}}\right)\right) \sim \mathrm{N}_n(\boldsymbol{0},\boldsymbol{\Sigma}).$$

   Assim,

$$\left(\frac{1}{\alpha_1}\left(\sqrt{\frac{t_1}{\gamma_1}} - \sqrt{\frac{\gamma_1}{t_1}}\right),\cdots,\frac{1}{\alpha_q}\left(\sqrt{\frac{t_q}{\gamma_q}} - \sqrt{\frac{\gamma_q}{t_q}}\right)\right) \sim N_q\left(\boldsymbol{A0},\boldsymbol{A\Sigma A}^\top\right) = \mathrm{N}_q(\boldsymbol{0},\boldsymbol{\Sigma_{11}}).$$

   Portanto, $\boldsymbol{T_1} \sim \mathrm{BS}_q(\boldsymbol{\alpha_1},\boldsymbol{\gamma_1};\boldsymbol{\Sigma_{11}})$.

(b) $\boldsymbol{T_2} \sim \mathrm{BS}_p\left(\boldsymbol{\alpha_2}, \boldsymbol{\gamma_2}; \boldsymbol{\Sigma_{22}}\right)$.

Demonstração: A demonstração, neste caso, é análoga à do item (a) considerando a matriz $\boldsymbol{A} = [a_{ij}],\ p \times n$, cujos elementos são $a_{ij} = \begin{cases} 1 & \text{se} \quad j = q + i \\ 0 & \text{caso contrário} \end{cases}$.

Observa-se que $\boldsymbol{AT} = \boldsymbol{T_2}$ e que $\boldsymbol{A\Sigma A}^\top = \boldsymbol{\Sigma_{22}}$. Logo,

$$\left(\frac{1}{\alpha_{q+1}}\left(\sqrt{\frac{t_{q+1}}{\gamma_{q+1}}} - \sqrt{\frac{\gamma_{q+1}}{t_{q+1}}}\right), \cdots, \frac{1}{\alpha_n}\left(\sqrt{\frac{t_n}{\gamma_n}} - \sqrt{\frac{\gamma_n}{t_n}}\right)\right) \sim \mathrm{N}_p\left(\boldsymbol{A0}, \boldsymbol{A\Sigma A}^\top\right) = \mathrm{N}_p\left(\boldsymbol{0}, \boldsymbol{\Sigma_{22}}\right),$$

e, portanto, $\boldsymbol{T_2} \sim \mathrm{BS}_p\left(\boldsymbol{\alpha_2}, \boldsymbol{\gamma_2}; \boldsymbol{\Sigma_{22}}\right)$.

Esta propriedade em conjunto com a propriedade 2 assegura que qualquer subvetor de $\boldsymbol{T}$ tem distribuição marginal BS.

4. Se $\boldsymbol{T} \sim \mathrm{BS}_n\left(\boldsymbol{\alpha}, \boldsymbol{\gamma}; \Sigma\right)$, então:

(a) $\left(c\boldsymbol{T_1}, \boldsymbol{T_2}\right)^\top \sim \mathrm{BS}_n\left(\left(\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}\right)^\top, \left(c\boldsymbol{\gamma_1}, \boldsymbol{\gamma_2}\right)^\top; \boldsymbol{\Sigma}\right)$.

Demonstração: Considerem $\boldsymbol{T} = \left(\boldsymbol{T_1}, \boldsymbol{T_2}\right)^\top$, $\boldsymbol{Y} = \left(c\boldsymbol{T_1}, \boldsymbol{T_2}\right)^\top$ e $u_i(t, \gamma) = \frac{1}{\alpha_i}\left(\sqrt{\frac{t}{\gamma}} - \sqrt{\frac{\gamma}{t}}\right)$. Assim, quando se utiliza a propriedade de funções de vetores aleatórios tem-se que $f_{\boldsymbol{Y}}(t_1, \ldots, t_n) = f_{\boldsymbol{T}}\left(\frac{1}{c}t_1, \ldots, \frac{1}{c}t_q, t_{q+1}, \ldots, t_n\right)\left|\det\left(\boldsymbol{A}\right)\right|$, em que a matriz $\boldsymbol{A} = [a_{ij}], n \times n$, é tal que $a_{ij} = \begin{cases} \frac{1}{c}, & \text{se} \quad i = j; 1 \leq i \leq q \\ 1, & \text{se} \quad i = j; q+1 \leq i \leq n \\ 0, & \text{caso contrário} \end{cases}$

Assim,

$$\begin{aligned} f_{\boldsymbol{Y}}(t_1, \cdots, t_n) = \phi_n\left(u_1\left(\frac{t_1}{c}, \gamma_1\right), \cdots, u_q\left(\frac{t_q}{c}, \gamma_q\right), u_{q+1}\left(t_{q+1}, \gamma_{q+1}\right), \cdots, u_n\left(t_n, \gamma_n\right); \boldsymbol{\Sigma}\right) \\ \prod_{i=1}^{q} \frac{1}{2\alpha_i\gamma_i}\left(\left(\frac{\gamma_i}{\frac{t_i}{c}}\right)^{\frac{1}{2}} + \left(\frac{\gamma_i}{\frac{t_i}{c}}\right)^{\frac{3}{2}}\right)\left(\prod_{i=q+1}^{n} \frac{1}{2\alpha_i\gamma_i}\left(\left(\frac{\gamma_i}{t_i}\right)^{\frac{1}{2}} + \left(\frac{\gamma_i}{t_i}\right)^{\frac{3}{2}}\right)\right)\frac{1}{c^q} = \\ = \phi_n\left(u_1\left(t_1, c\gamma_1\right), \cdots, u_q\left(t_q, , c\gamma_q\right), u_{q+1}\left(t_{q+1}, \gamma_{q+1}\right), \cdots, u_n\left(t_n, \gamma_n\right); \boldsymbol{\Sigma}\right) \\ \prod_{i=1}^{q} \frac{1}{2\alpha_i c\gamma_i}\left(\left(\frac{c\gamma_i}{t_i}\right)^{\frac{1}{2}} + \left(\frac{\gamma_i}{t_i}\right)^{\frac{3}{2}}\right)\prod_{i=q+1}^{n} \frac{1}{2\alpha_i\gamma_i}\left(\left(\frac{\gamma_i}{t_i}\right)^{\frac{1}{2}} + \left(\frac{\gamma_i}{t_i}\right)^{\frac{3}{2}}\right), \end{aligned}$$

que é a f.d.p. de uma variável aleatória com distribuição $\mathrm{BS}_n\left(\left(\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}\right)^\top, \left(c\boldsymbol{\gamma_1}, \boldsymbol{\gamma_2}\right)^\top; \boldsymbol{\Sigma}\right)$.

(b) $\left(\boldsymbol{T_1}, c\boldsymbol{T_2}\right)^\top \sim \mathrm{BS}_n\left(\left(\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}\right)^\top, \left(\boldsymbol{\gamma_1}, c\boldsymbol{\gamma_2}\right)^\top; \boldsymbol{\Sigma}\right)$.

Demonstração: Análoga a do item (a) quando $\boldsymbol{Y} = \left(\boldsymbol{T_1}, c\boldsymbol{T_2}\right)^\top$.

(c) $\left(c\boldsymbol{T_1}, c\boldsymbol{T_2}\right)^\top \sim \mathrm{BS}_n\left(\left(\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}\right)^\top, \left(c\boldsymbol{\gamma_1}, c\boldsymbol{\gamma_2}\right)^\top; \boldsymbol{\Sigma}\right)$.

Demonstração: Análoga a do item (a) quando $\boldsymbol{Y} = \left(c\boldsymbol{T_1}, c\boldsymbol{T_2}\right)^\top$.

5. Se $\left(\boldsymbol{T_1}, \boldsymbol{T_2}\right)^\top \sim \mathrm{BS}_n\left(\left(\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}\right)^\top, \left(\boldsymbol{\gamma_1}, \boldsymbol{\gamma_2}\right)^\top; \boldsymbol{\Sigma}\right)$ então:

(a) $\left(\boldsymbol{T_1^{-1}}, \boldsymbol{T_2}\right)^\top \sim \mathrm{BS}_n\left(\left(\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}\right)^\top, \left(\boldsymbol{\gamma_1^{-1}}, \boldsymbol{\gamma_2}\right)^\top; \boldsymbol{\Sigma_1}\right)$, em que $\boldsymbol{\Sigma_1} = \begin{pmatrix} \boldsymbol{\Sigma_{11}} & -\boldsymbol{\Sigma_{12}} \\ -\boldsymbol{\Sigma_{21}} & \boldsymbol{\Sigma_{22}} \end{pmatrix}$.

Demonstração: Considere $\boldsymbol{T} = (\boldsymbol{T_1}, \boldsymbol{T_2})^\top$, $\boldsymbol{Y} = (\boldsymbol{T_1^{-1}}, \boldsymbol{T_2})^\top$ e $u_i(t, \gamma) = \frac{1}{\alpha_i}\left(\sqrt{\frac{t}{\gamma}} - \sqrt{\frac{\gamma}{t}}\right)$. Assim, $f_{\boldsymbol{Y}}(t_1, \cdots, t_n) = f_{\boldsymbol{T}}\left(\frac{1}{t_1}, \cdots, \frac{1}{t_q}, t_{q+1}, \cdots, t_n\right)|\det(\boldsymbol{M})|$, em que a matriz $\boldsymbol{M} = [m_{ij}], n \times n$, é tal que $m_{ij} = \begin{cases} \frac{-1}{t^2}, & \text{se} \quad i = j; 1 \leq i \leq q \\ 1, & \text{se} \quad i = j; q+1 \leq i \leq n \\ 0, & \text{caso contrário} \end{cases}$

Logo,

$$\begin{aligned} f_{\boldsymbol{Y}}(t_1, \cdots, t_n) &= \phi_n\left(u_1\left(\frac{1}{t_1}, \gamma_1\right), \cdots, u_q\left(\frac{1}{t_q}, \gamma_q\right), u_{q+1}(t_{q+1}, \gamma_{q+1}), \cdots, u_n(t_n, \gamma_n); \boldsymbol{\Sigma}\right) \\ &\prod_{i=1}^{q} \frac{1}{2\alpha_i\gamma_i}\left(\left(\frac{\gamma_i}{\frac{1}{t_i}}\right)^{\frac{1}{2}} + \left(\frac{\gamma_i}{\frac{1}{t_i}}\right)^{\frac{3}{2}}\right) \prod_{i=q+1}^{n} \frac{1}{2\alpha_i\gamma_i}\left(\left(\frac{\gamma_i}{t_i}\right)^{\frac{1}{2}} + \left(\frac{\gamma_i}{t_i}\right)^{\frac{3}{2}} \prod_{i=1}^{q}\right) \frac{1}{t_i^2} \\ &= \phi_n\left(-u_1\left(t_1, \frac{1}{\gamma_1}\right), \cdots, -u_q\left(t_q, \frac{1}{\gamma_q}\right), u_{q+1}(t_{q+1}, \gamma_{q+1}), \cdots, u_n(t_n, \gamma_n); \boldsymbol{\Sigma}\right) \\ &\prod_{i=1}^{q} \frac{1}{2\alpha_i\frac{1}{\gamma_i}}\left(\left(\frac{\frac{1}{\gamma_i}}{t_i}\right)^{\frac{1}{2}} + \left(\frac{\frac{1}{\gamma_i}}{t_i}\right)^{\frac{3}{2}}\right) \prod_{i=q+1}^{n} \frac{1}{2\alpha_i\gamma_i}\left(\left(\frac{\gamma_i}{t_i}\right)^{\frac{1}{2}} + \left(\frac{\gamma_i}{t_i}\right)^{\frac{3}{2}}\right). \end{aligned}$$

Agora, considerando a matriz $\boldsymbol{A} = [a_{ij}], n \times n$, tal que,

$$a_{ij} = \begin{cases} -1, & \text{se} \quad i = j; 1 \leq i \leq q \\ 1, & \text{se} \quad i = j; q+1 \leq i \leq n \\ 0, & \text{se} \quad i \neq j \end{cases}$$

Tem-se que $\boldsymbol{AT} = (-\boldsymbol{T_1}, \boldsymbol{T_2})^\top$. Como $\boldsymbol{Z} \sim \mathrm{N}_n(\boldsymbol{0}, \boldsymbol{\Sigma})$ tem-se que $(-\boldsymbol{T_1}, \boldsymbol{T_2})^\top \sim \mathrm{N}_n\left(\boldsymbol{A0}, \boldsymbol{A\Sigma A}^\top\right) = \mathrm{N}_n(\boldsymbol{0}, \boldsymbol{\Sigma_1})$. E portanto,

$$\begin{aligned} f_{\boldsymbol{Y}}(t_1, \cdots, t_n) =& \phi_n\left(u_1\left(t_1, \frac{1}{\gamma_1}\right), \cdots, u_q\left(t_q, \frac{1}{\gamma_q}\right), u_{q+1}(t_{q+1}, \gamma_{q+1}), \cdots, u_n(t_n, \gamma_n); \boldsymbol{\Sigma_1}\right) \\ &\prod_{i=1}^{q} \frac{1}{2\alpha_i\frac{1}{\gamma_i}}\left(\left(\frac{\frac{1}{\gamma_i}}{t_i}\right)^{\frac{1}{2}} + \left(\frac{\frac{1}{\gamma_i}}{t_i}\right)^{\frac{3}{2}}\right) \prod_{i=q+1}^{n} \frac{1}{2\alpha_i\gamma_i}\left(\left(\frac{\gamma_i}{t_i}\right)^{\frac{1}{2}} + \left(\frac{\gamma_i}{t_i}\right)^{\frac{3}{2}}\right), \end{aligned}$$

que é a função densidade de probabilidade de uma variável aleatória com distribuição $\mathrm{BS}_n\left((\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2})^\top, (\boldsymbol{\gamma_1^{-1}}, \boldsymbol{\gamma_2})^\top; \boldsymbol{\Sigma_1}\right)$.

(b) $\left(\boldsymbol{T_1}, \boldsymbol{T_2^{-1}}\right)^\top \sim \mathrm{BS}_n\left((\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2})^\top, (\boldsymbol{\gamma_1}, \boldsymbol{\gamma_2^{-1}})^\top; \boldsymbol{\Sigma_1}\right)$, em que $\boldsymbol{\Sigma_1}$ é a matriz descrita no item (a).

Demonstração: Análoga à do item (a) considerando a matriz $\boldsymbol{A} = [a_{ij}], n \times n$, tal que

$$a_{ij} = \begin{cases} 1, & \text{se} \quad i = j; 1 \leq i \leq q \\ -1, & \text{se} \quad i = j; q+1 \leq i \leq n \\ 0, & \text{se} \quad i \neq j \end{cases}$$

e observando que $\boldsymbol{AT} = (\boldsymbol{T_1}, -\boldsymbol{T_2})^\top$ e que $\boldsymbol{A\Sigma A}^\top = \boldsymbol{\Sigma_1}$.

(c) $\left(\boldsymbol{T_1^{-1}}, \boldsymbol{T_2^{-1}}\right)^\top \sim \mathrm{BS}_n\left((\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2})^\top, (\boldsymbol{\gamma_1^{-1}}, \boldsymbol{\gamma_2^{-1}})^\top; \boldsymbol{\Sigma}\right)$.

Demonstração: Análoga a dos itens anteriores quando $\boldsymbol{A} = [a_{ij}], n \times n$, com $a_{ij} = \begin{cases} -1, & \text{se} \quad i = j \\ 0, & se \quad i \neq j \end{cases}$, e ao observar que $\boldsymbol{AT} = (-\boldsymbol{T_1}, -\boldsymbol{T_2})^\top$ e que $\boldsymbol{A\Sigma A}^\top = \boldsymbol{\Sigma}$.

6. (Momentos) Se $\boldsymbol{T} = (T_1, \ldots, T_n) \sim \mathrm{BS}_n(\boldsymbol{\alpha}, \boldsymbol{\gamma}; \boldsymbol{\Sigma})$ então:

   (a) $\boldsymbol{E}(\boldsymbol{T}) = \left(\gamma_1\left(1+\frac{\alpha_1^2}{2}\right), \cdots, \gamma_n\left(1+\frac{\alpha_n^2}{2}\right)\right)^\top$.

   Demonstração: Segue diretamente de $\boldsymbol{E}(\boldsymbol{T}) = (E(T_1), \cdots, E(T_n))^\top$ e das propriedades 4(b) da seção 3.1.1.1 e propriedade 1 desta seção.

   (b) A matriz de covariância do vetor aleatório $\boldsymbol{T}$ é dada por $\mathrm{Cov}(\boldsymbol{T}) = [a_{ij}]$, com

$$a_{ij} = \begin{cases} \alpha_i^2\gamma_i^2(\frac{5}{4}\alpha_i^2 + 1) & \text{se} \quad i = j \\ \rho_{\boldsymbol{T}_{ij}} & \text{se} \quad i \neq j \end{cases},$$

   em que $\rho_{\boldsymbol{T}_{ij}} = \gamma_i\gamma_j\left(\left(2 + \frac{3\alpha_i^2}{2} + \frac{3\alpha_j^2}{2} + \frac{9\alpha_i^2\alpha_j^2}{2^5}\right)\rho_{ij} + \frac{\alpha_i^2\alpha^j}{2}\rho_{ij}^2 + \frac{3\alpha_i^2\alpha_j^2}{2^4}\rho_{ij}^3 + r(Z_i, Z_j)\right)$, com ρ_{ij}, a covariância entre Z_i e Z_j e $r(Z_i, Z_j)$ dado de forma análoga ao realizado na propriedade 4(b) da Seção 3.1.1.1.

   (c) Os momentos de ordem 3 e 4 são dados por

$$\begin{aligned} \boldsymbol{E}(\boldsymbol{T^3}) &= \left(\gamma_1^3\left(\frac{15}{2}\alpha_1^6 + 9\alpha_1^4 + \frac{9}{2}\alpha_1^2 + 1\right), \cdots, \gamma_n^3\left(\frac{15}{2}\alpha_n^6 + 9\alpha_n^4 + \frac{9}{2}\alpha_n^2 + 1\right)\right)^\top \text{ e} \\ \boldsymbol{E}(\boldsymbol{T^4}) &= \left(\gamma_1^4\left(\frac{105}{2}\alpha_1^8 + 60\alpha_1^6 + 30\alpha_1^4 + 8\alpha_1^2 + 1\right), \cdots, \gamma_n^4\left(\frac{105}{2}\alpha_n^8 + 60\alpha_n^6 + 30\alpha_n^4 + 8\alpha_n^2 + 1\right)\right)^\top. \end{aligned}$$

   Demonstração: O resultado segue diretamente de $\boldsymbol{E}(\boldsymbol{T^r}) = (E(T_1^r), \cdots, E(T_n^r))^\top$ e dos resultados obtidos nas propriedades 4(d) e 4(e) da seção 3.1.1.1.

3.1.4 Distribuição log-BS univariada

Seja Y uma variável aleatória contínua. Diz-se que Y tem distribuição seno hiperbólico normal com parâmetros $\alpha > 0,\ \sigma > 0$ e $\mu \in \mathbb{R}$, a qual denota-se por $\mathrm{SHN}(\alpha, \mu, \sigma)$ se, e somente se,

$$Z = \frac{2}{\alpha}\mathrm{senh}\left(\frac{y-\mu}{\sigma}\right) \sim \mathrm{N}(0, 1).$$

Desta forma, se $Y \sim \mathrm{SHN}(\alpha, \mu, \sigma)$ a f.d.a. de Y é dada por

$$F(y) = \Phi\left(\frac{2}{\alpha}\mathrm{senh}\left(\frac{y-\mu}{\sigma}\right)\right), y \in \mathbb{R}$$

em que $\Phi(\cdot)$ é a f.d.a. da distribuição normal padrão.

Consequentemente a f.d.p. de Y com distribuição seno hiperbólico normal é dada por

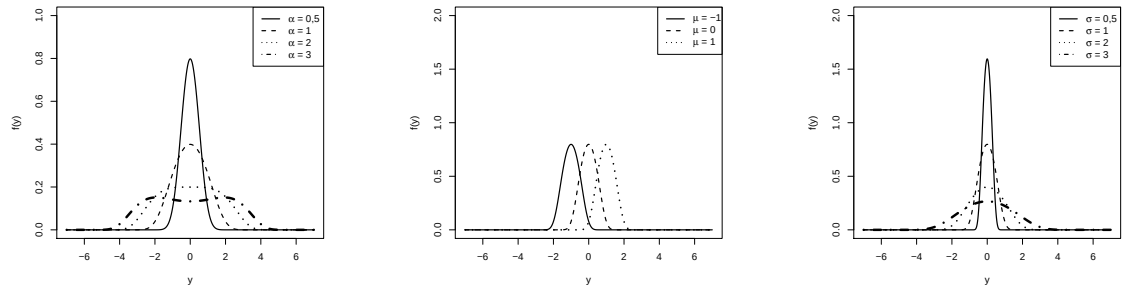
$$f(y) = \phi\left(\frac{2}{\alpha}\mathrm{senh}\left(\frac{y-\mu}{\sigma}\right)\right)\frac{d}{dy}\left(\frac{2}{\alpha}\mathrm{senh}\left(\frac{y-\mu}{\sigma}\right)\right)$$

em que $\phi(\cdot)$ é a f.d.p. da distribuição normal padrão. Isto é,

$$f(y) = \frac{2}{\alpha\sigma\sqrt{2\pi}}\cosh\left(\frac{y-\mu}{\sigma}\right)\exp\left(-\frac{2}{\alpha^2}\mathrm{senh}^2\left(\frac{y-\mu}{\sigma}\right)\right).$$

A Figura 3 mostra gráficos de f.d.p.'s da distribuição $\mathrm{SHN}(\alpha, \mu, \sigma)$. Pode-se observar que os parâmetros α, μ e σ são, respectivamente, parâmetros de forma, posição e escala e que o gráfico da função densidade é simétrico em torno de μ. Na Figura 3(a), têm-se $\mu = 0$, $\sigma = 1$, e diferentes valores do parâmetro de forma $\alpha \in \{0,5; 1; 2; 3\}$. Na Figura 3(b), têm-se os

parâmetros de forma e escala fixos, $\alpha = \sigma = 1$, e diferentes valores do parâmetro de posição μ, $\mu \in \{-1; 0; 1\}$. Finalmente, na Figura 3(c), têm-se os parâmetros de forma e posição fixos, $\alpha = 1$ e $\mu = 0$, e o parâmetro de escala, σ assume diferentes valores, $\sigma \in \{0,5; 1; 2; 3\}$.



(a) (b) (c)

Figura 3 Funções densidades de probabilidade da distribuição $\mathrm{SHN}(\alpha, \mu, \sigma)$.

A propriedade apresentada a seguir estabelece uma importante relação entre as distribuições seno hiperbólico normal e Birnbaum-Saunders:

Se uma variável aleatória contínua $T \sim \mathrm{BS}(\alpha, \gamma)$, então $Y = \log(T) \sim \mathrm{SHN}(\alpha, \log(\gamma), 2)$.

Demonstração:

$$\begin{aligned} F_Y(y) &= \mathrm{P}(Y \leq y) = \mathrm{P}(\log(T) \leq y) = \mathrm{P}(T \leq \exp(y)) \\ &= \Phi\left(\alpha^{-1}\left(\left(\frac{\exp(y)}{\gamma}\right)^{\frac{1}{2}} - \left(\frac{\exp(y)}{\gamma}\right)^{-\frac{1}{2}}\right)\right) = \Phi\left(\alpha^{-1}\left(\exp(y-\log(\gamma))^{\frac{1}{2}} - \exp(y-\log(\gamma))^{-\frac{1}{2}}\right)\right) \\ &= \Phi\left(\alpha^{-1}\left(\exp\left(\frac{y-\log(\gamma)}{2}\right) - \exp\left(-\frac{y-\log(\gamma)}{2}\right)\right)\right) = \Phi\left(\frac{2}{\alpha}\mathrm{senh}\left(\frac{y-\log(\gamma)}{2}\right)\right), \end{aligned}$$

que é a f.d.a. de uma variável aleatória com distribuição $\mathrm{SHN}(\alpha, \mu, \sigma)$.

Devido à relação estabelecida na propriedade anterior, a distribuição $\mathrm{SHN}(\alpha, \mu, 2)$ é chamada distribuição log-Birnbaum-Saunders (log-BS) com parâmetros α e μ ($\text{log-BS}(\alpha, \mu)$), isto é, a distribuição log-BS é um caso particular da distribuição seno hiperbólico normal, quando o parâmetro de escala $\sigma = 2$. Assim, diz-se que uma variável aleatória $Y \sim \text{log-BS}(\alpha, \mu)$ se, e somente se,

$$Z = \frac{2}{\alpha}\mathrm{senh}\left(\frac{y-\mu}{2}\right) \sim \mathrm{N}(0, 1).$$

A f.d.a. de uma variável aleatória Y como distribuição $\text{log-BS}(\alpha, \mu)$ é dada por

$$F(y) = \Phi\left(\frac{2}{\alpha}\mathrm{senh}\left(\frac{y-\mu}{2}\right)\right), -\infty < y < \infty$$

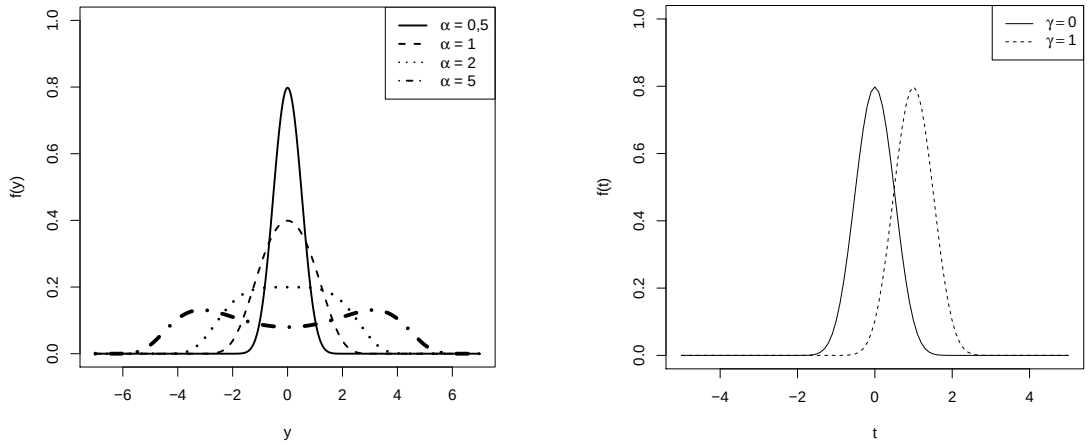
em que $\Phi(\cdot)$ é a f.d.a. da distribuição normal padrão. E a f.d.p. de Y é dada por

$$\begin{aligned} f(y) &= \phi\left(\frac{2}{\alpha}\mathrm{senh}\left(\frac{y-\mu}{2}\right)\right)\frac{\partial}{\partial y}\left(\frac{2}{\alpha}\mathrm{senh}\left(\frac{y-\mu}{2}\right)\right) \\ &= \frac{1}{\alpha\sqrt{2\pi}}\cosh\left(\frac{y-\mu}{2}\right)\exp\left(-\frac{2}{\alpha^2}\mathrm{senh}^2\left(\frac{y-\mu}{2}\right)\right). \end{aligned}$$

em que $\phi(\cdot)$ é a f.d.p. da distribuição normal padrão.

Na Figura 4, são apresentados exemplos da distribuição $\text{log-BS}(\alpha, \mu)$. Na Figura 4(a), o

parâmetro de posição assume o valor fixo $\mu = 0$ e o parâmetro de forma assume diferentes valores, $\alpha \in \{0,5; 1; 2; 5\}$. Podemos observar a simetria em torno de μ em todos os casos, a unimodalidade para os casos $\alpha = 0,5, \alpha = 1$, e $\alpha = 2$ e a bimodalidade para o caso $\alpha = 5$. Na Figura 4(b), têm-se os parâmetros de forma fixo, $\alpha = 0,5$, e μ assumindo os valores 0 e 1. Pode-se, novamente, observar a simetria em torno do parâmetro μ e o deslocamento do gráfico (mudança de posição) no eixo das abcissas.



(a) (b)

Figura 4 Funções densidades de probabilidade da distribuição $\log\text{-BS}(\alpha, \mu)$.

3.1.4.1 Propriedades da variável aleatória log-BS univariada

1. A distribuição $\log\text{-BS}(\alpha, \mu)$ é simétrica com respeito ao parâmetro μ.

   Demonstração: Uma distribuição é simétrica com respeito a μ se, e somente se, a distribuição de $Y - \mu$ é igual a de $-(Y - \mu)$. Seja $Y \sim \log\text{-BS}(\alpha, \mu)$ e $X = Y - \mu$,

   $$\mathrm{P}(X \leq x) = \mathrm{P}(Y - \mu \leq x) = \mathrm{P}(Y \leq x + \mu) = \Phi\left(\frac{2}{\alpha}\mathrm{senh}\left(\frac{x + \mu - \mu}{2}\right)\right) = \Phi\left(\frac{2}{\alpha}\mathrm{senh}\left(\frac{x}{2}\right)\right).$$

   Por outro lado, se considerarmos $X = -(Y - \mu) = -Y + \mu$, tem-se que

   $$\begin{aligned} \mathrm{P}(X \leq x) &= \mathrm{P}(-Y + \mu \leq x) = \mathrm{P}(-Y \leq x - \mu) = 1 - \mathrm{P}(Y \leq -x + \mu) \\ &= 1 - \Phi\left(\frac{2}{\alpha}\mathrm{senh}\left(\frac{-x + \mu - \mu}{2}\right)\right) = \Phi\left(\frac{-2}{\alpha}\mathrm{senh}\left(\frac{-x}{2}\right)\right) \\ &= \Phi\left(\frac{-2}{\alpha}\left(-\mathrm{senh}\left(\frac{x}{2}\right)\right)\right) = \Phi\left(\frac{2}{\alpha}\mathrm{senh}\left(\frac{x}{2}\right)\right). \end{aligned}$$

   Assim, as distribuições de $(Y - \mu)$ e $-(Y - \mu)$ são iguais e pode-se concluir que Y é uma distribuição simétrica em relação a μ.

2. Se $Y \sim \log\text{-BS}(\alpha, \mu)$, então $cY + a \sim \mathrm{SHN}(\alpha, c\mu + a, 2|c|)$. Em particular, $(\pm Y + a) \sim \log\text{-BS}(\alpha, \pm\mu + a)$.

Demonstração: Primeiramente, considera-se o caso $X = cY + a, c > 0$.

$$\mathrm{P}(X \leq x) \quad = \mathrm{P}(Y \leq \tfrac{x-a}{c}) = \Phi\left(\tfrac{2}{\alpha}\mathrm{senh}\left(\tfrac{\frac{x-a}{c}-\mu}{2}\right)\right) = \Phi\left(\tfrac{2}{\alpha}\mathrm{senh}\left(\tfrac{x-(a+c\mu)}{2c}\right)\right)$$

que é a f.d.a. da distribuição $\mathrm{SHN}(\alpha, c\mu + a, 2c)$.

Agora, se $X = cY + a$, com $c < 0$, tem-se

$$\begin{aligned}\mathrm{P}(X \leq x) &= \mathrm{P}(Y \geq \frac{x-a}{c}) = 1 - \mathrm{P}(Y \leq \frac{x-a}{c}) = 1 - \Phi\left(\frac{2}{\alpha}\mathrm{senh}\left(\frac{x-(a+c\mu)}{2c}\right)\right) \\ &= \Phi\left(-\frac{2}{\alpha}\mathrm{senh}\left(-\frac{x-(a+c\mu)}{-2c}\right)\right) = \Phi\left(\frac{2}{\alpha}\mathrm{senh}\left(\frac{x-(a+c\mu)}{-2c}\right)\right)\end{aligned}$$

que é a f.d.a. da distribuição $\mathrm{SHN}(\alpha, c\mu + a, -2c)$.

Assim, de um modo geral, tem-se $X \sim \mathrm{SHN}(\alpha, c\mu + a, 2|c|)$.

3. A distribuição $\log\text{-BS}(\alpha, \mu)$ é unimodal para $\alpha \leq 2$ e bimodal para $\alpha > 2$.

   Demonstração: Para que se encontre(m) a(s) moda(s) de uma distribuição, deve-se encontrar $y \in \mathbb{R}$ tal que $f'(y) = 0$, em que $f(\cdot)$ é a f.d.p. da distribuição. No caso da distribuição $\log\text{-BS}(\alpha, \mu)$, tem-se

   $$f(y) = \frac{1}{\alpha\sqrt{2\pi}}\cosh\left(\frac{y-\mu}{2}\right)\exp\left(-\frac{2}{\alpha^2}\mathrm{senh}^2\left(\frac{y-\mu}{2}\right)\right), \text{ e consequentemente,}$$

   $$f^{'}(y) = \frac{1}{2\alpha\sqrt{2\pi}}\exp\left(-\frac{2}{\alpha^2}\mathrm{senh}^2\left(\frac{y-\mu}{2}\right)\right)\mathrm{senh}\left(\frac{y-\mu}{2}\right)\left(1 - \frac{4}{\alpha^2}\cosh^2\left(\frac{y-\mu}{2}\right)\right).$$

   Assim, $f^{'}(y) = 0 \Leftrightarrow \mathrm{senh}\left(\frac{y-\mu}{2}\right) = 0$ ou $\left(1 - \frac{4}{\alpha^2}\cosh^2\left(\frac{y-\mu}{2}\right)\right) = 0$. Agora, têm-se que $\mathrm{senh}\left(\frac{y-\mu}{2}\right) = 0 \Leftrightarrow \frac{y-\mu}{2} = 0 \Leftrightarrow y = \mu$ e que $\cosh^2\left(\frac{y-\mu}{2}\right) \geq 1$, sendo que $\cosh^2\left(\frac{y-\mu}{2}\right) = 1 \Leftrightarrow y = \mu$.

   Se $\alpha = 2$ tem-se $\frac{4}{\alpha^2} = 1$ e $1 - \frac{4}{\alpha^2}\cosh^2\left(\frac{y-\mu}{2}\right) = 0 \Leftrightarrow \cosh^2\left(\frac{y-\mu}{2}\right) = 1 \Leftrightarrow y = \mu$. Neste caso, $y = \mu$ é a única solução, isto é, a distribuição é unimodal.

   Se $0 < \alpha < 2 \Rightarrow 0 < \alpha^2 < 4 \Rightarrow \frac{4}{\alpha^2} > 1 \Rightarrow \frac{4}{\alpha^2}\cosh^2\left(\frac{y-\mu}{2}\right) > 1$. Assim, $1 - \frac{4}{\alpha^2}\cosh^2\left(\frac{y-\mu}{2}\right) \neq 0$, e neste caso também a distribuição é unimodal.

   Finalmente, se $\alpha > 2 \Rightarrow \alpha^2 > 4 \Rightarrow 1 - \frac{4}{\alpha^2}\cosh^2\left(\frac{y-\mu}{2}\right) = 0 \Leftrightarrow \cosh^2\left(\frac{y-\mu}{2}\right) = \frac{\alpha^2}{4}$ há duas soluções diferentes de μ. Logo, neste caso, a distribuição é bimodal.

4. Se $T \sim \mathrm{BS}(\alpha, \gamma)$, então $\log(cT) \sim \log\text{-BS}(\alpha, \log(c) + \log(\gamma)), \forall c \in \mathbb{R}, c > 0$.

   Demonstração: $T \sim \mathrm{BS}(\alpha, \gamma) \Rightarrow cT \sim \mathrm{BS}(\alpha, c\gamma) \Rightarrow \log(cT) \sim \log\text{-BS}(\alpha, \log(c\gamma)) = \log\text{-BS}(\alpha, \log(c) + \log(\gamma)$.

5. Se $T \sim \mathrm{BS}(\alpha, \gamma)$, então $\log(T^{-1}) \sim \log\text{-BS}(\alpha, -\log(\gamma))$.

   Demonstração: $T \sim \mathrm{BS}(\alpha, \gamma) \Rightarrow T^{-1} \sim \mathrm{BS}(\alpha, \gamma^{-1}) \Rightarrow \log(T^{-1}) \sim \log\text{-BS}(\alpha, \log(\gamma^{-1})) = \log\text{-BS}(\alpha, -\log(\gamma)$.

6. (Momentos) Se $Y \sim \log\text{-BS}(\alpha, \mu)$, então

(a) A função geradora de momentos de Y é dada por

$$M_Y(t) = \frac{\exp(t\mu)}{2K_{1/2}(\alpha^{-2})}\left(K_{\frac{2t+1}{2}}(\alpha^{-2}) + K_{\frac{2t-1}{2}}(\alpha^{-2})\right).$$

Demonstração: Primeiramente considera-se uma variável aleatória S tal que $S \sim \mathrm{SHN}(\alpha, 0, 1)$. Utilizando-se a f.d.p. da distribuição SHN, a definição de $\cosh$ e a relação $\cosh(2s) = 2\mathrm{senh}^2(s) + 1$, tem-se que

$$\begin{aligned} f_S(s) &= \frac{2}{\alpha\sqrt{2\pi}}\cosh(s)\exp\left(\frac{-2}{\alpha^2}\mathrm{senh}^2(s)\right) \\ &= \frac{2}{\alpha\sqrt{2\pi}}\frac{\exp(s)+\exp(-s)}{2}\exp\left(\frac{-2}{\alpha^2}\left(\frac{\cosh(2s)-1}{2}\right)\right) \\ &= \frac{\exp(\alpha^{-2})}{\alpha\sqrt{2\pi}}\left(\exp(s)+\exp(-s)\right)\exp\left(-\alpha^{-2}\cosh(2s)\right). \end{aligned}$$

A função geradora de momentos de S é

$$\begin{aligned} M_S(t) &= \mathrm{E}(\exp(tS)) = \int_{-\infty}^{\infty}\exp(ts)f_S(s)ds \\ &= \frac{\exp(\alpha^{-2})}{\alpha\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp(ts)\left(\exp(s)+\exp(-s)\right)\exp\left(-\alpha^{-2}\cosh(2s)\right)ds \\ &= \frac{\exp(\alpha^{-2})}{\alpha\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\left((t+1)s-\alpha^{-2}\cosh(2s)\right)ds \\ &+ \frac{\exp(\alpha^{-2})}{\alpha\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\left((t-1)s-\alpha^{-2}\cosh(2s)\right)ds \end{aligned}$$

Ao se fazer a mudança de variável $x = 2s$ tem-se

$$\begin{aligned} M_S(t) &= \frac{\exp(\alpha^{-2})}{\alpha\sqrt{2\pi}}\left(\frac{1}{2}\int_{-\infty}^{\infty}\exp\left((t+1)\frac{x}{2}-\alpha^{-2}\cosh(x)\right)dx\right) \\ &+ \frac{\exp(\alpha^{-2})}{\alpha\sqrt{2\pi}}\left(\frac{1}{2}\int_{-\infty}^{\infty}\exp\left((t-1)\frac{x}{2}-\alpha^{-2}\cosh(x)\right)dx\right). \end{aligned}$$

De acordo com a definição da função de Bessel modificada do terceiro tipo de ordem υ, $K_\upsilon(z)$, e que $K_{1/2}(z) = \sqrt{\pi/2z}\exp(-z)$ (RIECK, 1989), tem-se que

$$M_S(t) = \frac{K_{\frac{t+1}{2}}(\alpha^{-2}) + K_{\frac{t-1}{2}}(\alpha^{-2})}{2K_{1/2}(\alpha^{-2})}. \tag{3.2}$$

Para a função geradora de momentos de uma variável aleatória Y com distribuição $\log(\alpha, \mu) = \mathrm{SHN}(\alpha, \mu, 2)$, considera-se a transformação $Y = \mu + 2S$.

$$M_Y(t) = \mathrm{E}(\exp(tY)) = \mathrm{E}(\exp(t(\mu+2S))) = \mathrm{E}(\exp(t\mu)\exp(2tS)) = \exp(t\mu)\mathrm{E}(\exp(2tS))$$

Pela propriedade 2, tem-se que $S \sim \mathrm{SHN}(\alpha, 0, 1)$, portanto, utiliza-se a equação (3.2) e tem-se que

$$M_Y(t) = \exp(t\mu)\frac{K_{\frac{2t+1}{2}}(\alpha^{-2}) + K_{\frac{2t-1}{2}}(\alpha^{-2})}{2K_{1/2}(\alpha^{-2})}.$$

(b) $\mathrm{E}(Y) = \mu$.

Demonstração: Segue diretamente da propriedade 6(a).

(c) Rieck e Nedelman (1991) observam que não existe forma fechada para a variância de uma variável com distribuição seno hiberbólico normal. Rieck (1989) fornece aproximações assintóticas para os casos α pequeno ($\alpha \to 0$) e α muito grande ($\alpha \to \infty$). As aproximações fornecidas por Rieck (1989) são baseadas em aproximações da função geradora de momentos da distribuição seno hiperbólico normal, fornecidas por Abramowitz e Stegun (1972). A seguir são apresentadas as aproximações para o caso particular em que o parâmetro de escala é $\sigma = 2$, isto é, para a distribuição log-Birnbaum-Saunders,

i. Se $\alpha \to 0$, então $\mathrm{Var}(Y) \approx \alpha^2 - \frac{\alpha^4}{4}$.
Demonstração: Neste caso, $M_Y(t) \approx e^{\mu t}\left(1 + \frac{\alpha^2 t^2}{2} + \frac{\alpha^4(t^4 - t^2)}{8}\right)$, consequentemente, $M_Y'(t) = \mu e^{\mu t}\left(1 + \frac{\alpha^2 t^2}{2} + \frac{\alpha^4(t^4 - t^2)}{8}\right) + e^{\mu t}\left(\alpha^2 t + \frac{\alpha^4(4t^3 - 2t)}{8}\right)$, e $M_Y''(t) = \mu^2 e^{\mu t}\left(1 + \frac{\alpha^2 t^2}{2} + \frac{\alpha^4(t^4 - t^2)}{8}\right) + 2\mu e^{\mu t}\left(\alpha^2 t + \frac{\alpha^4(2t^3 - t)}{4}\right) + e^{\mu t}\left(\alpha^2 + \frac{\alpha^4(12t^2 - 2)}{8}\right)$, acarretando $\mathrm{E}(Y^2) = M_Y''(0) = \mu^2 + \alpha^2 - \frac{\alpha^4}{4}$. Assim:

$$\mathrm{Var}(Y) = \mathrm{E}(Y^2) - \mathrm{E}^2(Y) \approx \mu^2 + \alpha^2 - \frac{\alpha^4}{4} - \mu^2 = \alpha^2 - \frac{\alpha^4}{4}.$$

ii. Se $\alpha \to \infty$, então $\mathrm{Var}(Y) \approx 4(-2\log(\delta) + \log^2(\delta) + 2)$.
Demonstração: Neste caso, $M_Y(t) \approx \frac{e^{\mu t}}{2\sqrt{\pi}}\left(\mu\left(\frac{1}{2} + t\right)\delta^{2t} + \mu\left(\frac{1}{2} - t\right)\delta^{-2t}\right); |t| \leq \frac{1}{2}$, em que $\mu(.)$ é a função Gamma e $\delta = \alpha\sqrt{2}$. Após alguns cálculos, pode-se concluir que $M_Y''(0) \approx \mu^2 + 4\left(-2\log(\delta) + \log^2(\delta) + 2\right)$. E, consequentemente, $\mathrm{Var}(Y) \approx 4(-2\log(\delta) + \log^2(\delta) + 2)$.

(d) O cálculo do coeficiente de assimetria, AS(Y), será realizado para os casos assintóticos, baseado nas aproximações da função geradora de momentos da distribuição log-Birnbaum-Saunders como no cálculo da variância.

i. Se $\alpha \to 0$, então $AS(Y) = \frac{\mu^3 + 3\mu\alpha^2\left(\frac{-\alpha^2}{4} + 1\right)}{\sqrt{\left(\mu^2 + \alpha^2\left(1 - \frac{\alpha^2}{4}\right)\right)^3}}$.
Demonstração:

$$\begin{aligned} M_Y^3(t) &= \mu^3 e^{\mu t}\left(1 + \frac{\alpha^2 t^2}{2} + \frac{\alpha^4(t^4 - t^2)}{8}\right) + 3\mu^2 e^{\mu t}\left(\alpha^2 t + \frac{\alpha^4(2t^3 - t)}{4}\right) \\ &+ 3\mu e^{\mu t}\left(\alpha^2 + \frac{\alpha^4(12t^2 - 2)}{8}\right) + 3t\alpha^4 e^{\mu t}. \end{aligned}$$

Consequentemente, $M_Y^3(0) = \mu^3 + 3\mu\left(-\frac{\alpha^4}{4} + \alpha^2\right)$ e, portanto,

$$AS(Y) = \frac{M_Y^{(3)}(0)}{\sqrt{M_Y''(0)^3}} = \frac{\mu^3 + 3\mu\alpha^2\left(\frac{-\alpha^2}{4} + 1\right)}{\sqrt{\left(\mu^2 + \alpha^2\left(1 - \frac{\alpha^2}{4}\right)\right)^3}}.$$

ii. Se $\alpha \to \infty$, então $\mathrm{AS}(Y) \approx \frac{\mu^3 + \mu\left(12\log^2(\delta) - 23\log(\delta) + 26\right)}{\sqrt{\left(\mu^2 + 4(-2\log(\delta) + \log^2(\delta) + 2)\right)^3}}$.
Demonstração: Segue diretamente da aproximação $M_Y^{(3)}(0) \approx \mu^3 +$

$$\mu\left(12\log^2(\delta) - 24\log(\delta) + 26\right).$$

(e) O coeficiente de curtose também será calculado para os casos assintóticos:

i. Se $\alpha \to 0$, então $\mathrm{K}(Y) = \frac{M_Y^{(4)}(0)}{M_Y''(0)^2} = \frac{\mu^4 - \frac{3}{2}\mu^2\alpha^4 + 6\mu^2\alpha^2 + 3\alpha^4}{\left(\mu^2 + \alpha^2 - \frac{\alpha^2}{4}\right)^2}$.

Demonstração:

$$\begin{aligned} M_Y^{(4)}(t) &= \mu^4 e^{\mu t}\left(1 + \frac{\alpha^2 t^2}{2} + \frac{\alpha^4(t^4 - t^2)}{8}\right) + 4\mu^3 e^{\mu t}\left(\alpha^2 t + \frac{\alpha^4(2t^3 - t)}{4}\right) \\ &+ 6\mu^2 e^{\mu t}\left(\alpha^2 + \frac{\alpha^4(12t^2 - 2)}{8}\right) + 9\mu t\alpha^4 e^{\mu t} + 3\alpha^4 e^{\mu t}. \end{aligned}$$

E assim, $M_Y^{(4)}(0) = \mu^4 - \frac{3}{2}\mu^2\alpha^4 + 6\mu^2\alpha^2 + 3\alpha^4$. Logo,

$$\mathrm{K}(Y) = \frac{M_Y^{(4)}(0)}{M_Y''(0)^2} = \frac{\mu^4 - \frac{3}{2}\mu^2\alpha^4 + 6\mu^2\alpha^2 + 3\alpha^4}{\left(\mu^2 + \alpha^2 - \frac{\alpha^2}{4}\right)^2}.$$

ii. Se $\alpha \to \infty$, então

$$\begin{aligned} K(Y) &= \frac{\mu^4 + \mu^2\left(24\log^2(\delta) - 47\log(\delta) + 53\right) + 16\log^4(\delta)}{(\mu^2 + 4(-2\log(\delta) + \log^2(\delta) + 2))^2} \\ &+ \frac{-63\log^3(\delta) + 211\log^2(\delta) - 428\log(\delta) + 432}{(\mu^2 + 4(-2\log(\delta) + \log^2(\delta) + 2))^2}. \end{aligned}$$

Demonstração: Neste caso, o resultado segue de

$$\begin{aligned} M_Y^{(4)}(0) &\approx \mu^4 + \mu^2\left(24\log^2(\delta) - 47\log(\delta) + 53\right) \\ &+ 16\log^4(\delta) - 63\log^3(\delta) + 211\log^2(\delta) - 428\log(\delta) + 432. \end{aligned}$$

3.1.5 Distribuição log-BS bivariada

Nesta Seção, define-se a distribuição log-Birnbaum-Saunders bivariada por meio de uma extensão da distribuição log-Birnbaum-Saunders.

Um vetor aleatório $\boldsymbol{Y} = (Y_1, Y_2)^\top$ tem distribuição log-Birnbaum-Saunders bivariada com parâmetros $\alpha_1, \mu_1, \alpha_2, \mu_2$ e ρ (Notação: $\log\text{-}\mathrm{BS}_2(\alpha_1, \mu_1, \alpha_2, \mu_2; \rho)$) se a f.d.a. de $\boldsymbol{Y}$ for expressa como

$$F_{\boldsymbol{Y}}(\boldsymbol{y}) = \mathrm{P}(Y_1 \leq y_1, Y_2 \leq y_2) = \Phi_2\left(\frac{2}{\alpha_1}\mathrm{senh}\left(\frac{y_1 - \mu_1}{2}\right), \frac{2}{\alpha_2}\mathrm{senh}\left(\frac{y_2 - \mu_2}{2}\right); \rho\right),$$

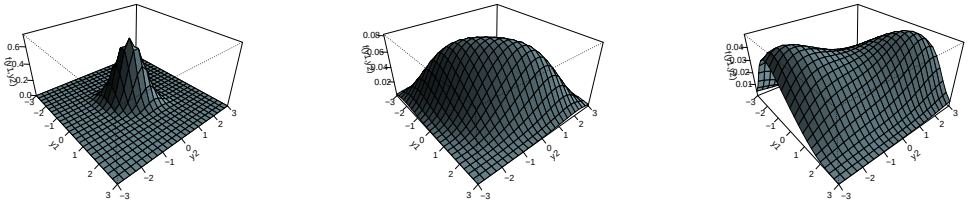
em que $\alpha_1, \alpha_2 > 0, -1 \leq \rho \leq 1$ e $\Phi_2(u, v; \rho)$ é a f.d.a. do vetor normal padrão bivariado com coeficiente de correlação ρ.

Decorre diretamente da definição que a f.d.p. da distribuição $\log\text{-}\mathrm{BS}(\alpha_1, \mu_1, \alpha_2, \mu_2; \rho)$ é dada por

$$f_Y(y) = \phi_2\left(\frac{2}{\alpha_1}\mathrm{senh}\left(\frac{y_1 - \mu_1}{2}\right), \frac{2}{\alpha_2}\mathrm{senh}\left(\frac{y_2 - \mu_2}{2}\right); \rho\right)\cosh\left(\frac{y_1 - \mu_1}{2}\right)\cosh\left(\frac{y_2 - \mu_2}{2}\right)\frac{1}{\alpha_1\alpha_2}.$$

Na Figura 5 são apresentadas algumas formas da f.d.p. da distribuição

log-BS$(\alpha_1, \mu_1, \alpha_2, \mu_2; \rho)$



(a) (b) (c)

Figura 5 Funções distribuição de probabilidade da distribuição $\text{log-BS}(\boldsymbol{\alpha}, \mathbf{0}; 0, 5)$ nos casos $\boldsymbol{\alpha} = (1,5; 1,5)$ (a); $\boldsymbol{\alpha} = (0,5; 0,5)$ (b) e $\boldsymbol{\alpha} = (1,5; 3)$ (c).

3.1.5.1 Propriedades do vetor aleatório log-BS bivariado

1. Se $\boldsymbol{Y} = (Y_1, Y_2)^\top \sim \text{log-BS}_2(\alpha_1, \mu_1, \alpha_2, \mu_2; \rho)$, então, $Y_i \sim \text{log-BS}(\alpha_i, \mu_i), i = 1, 2$.

   Demonstração: Se $\boldsymbol{Y} = (Y_1, Y_2)^\top \sim \text{log-BS}_2(\alpha_1, \mu_1, \alpha_2, \mu_2; \rho)$, então, $\left(\frac{2}{\alpha_1}\text{senh}\left(\frac{y_1-\mu_1}{2}\right), \frac{2}{\alpha_2}\text{senh}\left(\frac{y_2-\mu_2}{2}\right)\right)^\top \sim N_2(\mathbf{0}, \boldsymbol{\rho})$, logo, $\frac{2}{\alpha_i}\text{senh}\left(\frac{y_i-\mu_i}{2}\right) \sim \text{N}(0,1)$ e, portanto, $Y_i \sim \text{log-BS}(\alpha_i, \mu_i), i = 1, 2$.

2. Se $\boldsymbol{Y} = (Y_1, Y_2)^\top \sim \text{log-BS}_2(\alpha_1, \mu_1, \alpha_2, \mu_2; \rho)$, então:

   (a) $(cY_1, Y_2)^\top \sim \text{SHN}_2(\alpha_1, c\mu_1, 2|c|, \alpha_2, \mu_2, 2; \rho)$ se $c > 0$, e
   $(cY_1, Y_2)^\top \sim \text{SHN}_2(\alpha_1, c\mu_1, 2|c|, \alpha_2, \mu_2, 2; -\rho)$ se $c < 0$.
   Demonstração: Seja $\boldsymbol{Y} = (Y_1, Y_2)$ e $\boldsymbol{X} = (cY_1, Y_2)$. Assim, $f_{\boldsymbol{X}}(y_1, y_2) = f_{\boldsymbol{Y}}\left(\frac{1}{c}y_1, y_2\right) det(\boldsymbol{J})$, em que $\boldsymbol{J} = \begin{pmatrix} \frac{1}{c} & 0 \\ 0 & 1 \end{pmatrix}$. Então,

   $$f_{\boldsymbol{X}}(y_1, y_2) = \phi_2\left(\frac{2}{\alpha_1}\text{senh}\left(\frac{\frac{y_1}{c}-\mu_1}{2}\right), \frac{2}{\alpha_2}\,\text{senh}\left(\frac{y_2-\mu_2}{2}\right); \rho\right)\frac{1}{\alpha_1\alpha_2}\cosh\left(\frac{\frac{y_1}{c}-\mu_1}{2}\right)\cosh\left(\frac{y_2-\mu_2}{2}\right)\frac{1}{|c|}.$$

   Assim, para $c > 0$ tem-se

   $$\begin{aligned} f_{\boldsymbol{X}}(y_1, y_2) &= \phi_2\left(\frac{2}{\alpha_1}\text{senh}\left(\frac{y_1 - c\mu_1}{2|c|}\right), \frac{2}{\alpha_2}\text{senh}\left(\frac{y_2-\mu_2}{2}\right); \rho\right) \\ &\quad \frac{1}{|c|\alpha_1\alpha_2}\cosh\left(\frac{y_1 - c\mu_1}{2|c|}\right)\cosh\left(\frac{y_2-\mu_2}{2}\right), \end{aligned}$$

   que é a f.d.p. da distribuição $\text{SHN}_2(\alpha_1, c\mu_1, 2|c|, \alpha_2, \mu_2, 2; \rho)$.
   Para $c < 0$, vamos fazer a seguinte manipulação algébrica:

   $$\text{senh}\left(\frac{y_1 - c\mu_1}{2c}\right) = \text{senh}\left(\frac{-(y_1 - c\mu_1)}{-2c}\right) = \text{senh}\left(-\frac{y_1 - c\mu_1}{2|c|}\right) = -\text{senh}\left(\frac{y_1 - c\mu_1}{2|c|}\right).$$

Portanto,

$$\begin{aligned} f_{\boldsymbol{X}}(y_1,y_2) &= \phi_2\left(\frac{2}{\alpha_1}\text{senh}\left(\frac{y_1-c\mu_1}{2c}\right),\frac{2}{\alpha_2}\text{senh}\left(\frac{y_2-\mu_2}{2}\right);\rho\right) \\ &\quad\times\frac{1}{|c|\alpha_1\alpha_2}\cosh\left(\frac{y_1-c\mu_1}{2c}\right)\cosh\left(\frac{y_2-\mu_2}{2}\right) \\ &= \phi_2\left(\frac{2}{\alpha_1}\text{senh}\left(\frac{y_1-c\mu_1}{2|c|}\right),\frac{2}{\alpha_2}\text{senh}\left(\frac{y_2-\mu_2}{2}\right);-\rho\right) \\ &\quad\times\frac{1}{|c|\alpha_1\alpha_2}\cosh\left(\frac{y_1-c\mu_1}{2c}\right)\cosh\left(\frac{y_2-\mu_2}{2}\right). \end{aligned}$$

que é a f.d.p. da distribuição $\text{SHN}_2(\alpha_1, c\mu_1, 2|c|, \alpha_2, \mu_2, 2; -\rho)$.

(b) $(Y_1, cY_2)^\top \sim \text{SHN}_2(\alpha_1, \mu_1, 2, \alpha_2, c\mu_2, 2|c|; \rho)$ se $c > 0$, e $(Y_1, cY_2)^\top \sim \text{SHN}_2(\alpha_1, \mu_1, 2, \alpha_2, c\mu_2, 2|c|; -\rho)$ se $c < 0$.

Demonstração: A demonstração é análoga à do item (a).

(c) $(cY_1, cY_2)^\top \sim \text{SHN}_2(\alpha_1, c\mu_1, 2|c|, \alpha_2, c\mu_2, 2|c|; \rho)$.

Demonstração: Neste item, o raciocínio também é análogo ao empregado nos itens anteriores, porém, deve-se lembrar que $\phi_2(-u, -v, \rho) = \phi_2(u, v, \rho)$.

3. Se $\boldsymbol{Y} = (Y_1, Y_2)^\top \sim \text{log-BS}_2(\alpha_1, \mu_1, \alpha_2, \mu_2; \rho)$, então:

(a) $(-Y_1, Y_2)^\top \sim \text{log-BS}_2(\alpha_1, -\mu_1, \alpha_2, \mu_2; -\rho)$.

(b) $(Y_1, -Y_2)^\top \sim \text{log-BS}_2(\alpha_1, \mu_1, \alpha_2, -\mu_2; -\rho)$.

(c) $(-Y_1, -Y_2)^\top \sim \text{log-BS}_2(\alpha_1, -\mu_1, \alpha_2, -\mu_2; \rho)$.

Demonstração: O resultado segue da propriedade 2 considerando $c = -1$.

4. Se $\boldsymbol{Y} = (Y_1, Y_2)^\top \sim \text{BS}_2(\alpha_1, \beta_1, \alpha_2, \beta_2; \rho)$, então $\log(\boldsymbol{Y}) = (\log(Y_1), \log(Y_2))^\top \sim \text{BS}_2(\alpha_1, \log(\beta_1), \alpha_2, \log(\beta_2); \rho)$.

Demonstração: Considere $(X_1, X_2) = (\log(Y_1), \log(Y_2))$. Assim,

$$\begin{aligned} F_{\boldsymbol{X}}(\boldsymbol{y}) &= \text{P}(X_1 \le y_1, X_2 \le y_2) = \text{P}(\log(Y_1) \le y_1, \log(Y_2) \le y_2) = \text{P}(Y_1 \le \exp(y_1), Y_2 \le \exp(y_2)) \\ &= \phi_2\left(\alpha_1^{-1}\left(\left(\frac{\exp(y_1)}{\beta_1}\right)^{\frac{1}{2}} - \left(\frac{\exp(y_1)}{\beta_1}\right)^{\frac{-1}{2}}\right), \alpha_2^{-1}\left(\left(\frac{\exp(y_2)}{\beta_2}\right)^{\frac{1}{2}} - \left(\frac{\exp(y_2)}{\beta_2}\right)^{\frac{-1}{2}}\right); \rho\right). \end{aligned}$$

Após alguns cálculos, tem-se que

$$F_{\boldsymbol{X}}(\boldsymbol{y}) = \phi_2\left(\frac{2}{\alpha_1}\text{senh}\left(\frac{y_1-\log(\beta_1)}{2}\right),\frac{2}{\alpha_2}\text{senh}\left(\frac{y_2-\log(\beta_2)}{2}\right);\rho\right),$$

que é a f.d.a. da distribuição $\text{log-BS}_2(\alpha_1, \log(\beta_1), \alpha_2, \log(\beta_2); \rho)$.

5. Se $\boldsymbol{Y} = (Y_1, Y_2)^\top \sim \text{BS}_2(\alpha_1, \beta_1, \alpha_2, \beta_2; \rho)$, então:

(a) $(\log(cY_1), \log(Y_2))^\top \sim \text{log-BS}_2(\alpha_1, \log(c) + \log(\beta_1), \alpha_2, \log(\beta_2); \rho)$.

Demonstração: Se $\boldsymbol{Y} = (Y_1, Y_2)^\top \sim \text{BS}_2(\alpha_1, \beta_1, \alpha_2, \beta_2; \rho)$, então, $(cY_1, Y_2)^\top \sim \text{BS}_2(\alpha_1, c\beta_1, \alpha_2, \beta_2; \rho)$, portanto, $(\log(cY_1), \log(Y_2))^\top \sim \text{log-BS}_2(\alpha_1, \log(c\beta_1), \alpha_2, \log(\beta_2); \rho)$. Isto é,

$$(\log(cY_1), \log(Y_2))^\top \sim \text{log-BS}_2(\alpha_1, \log(c) + \log(\beta_1), \alpha_2, \log(\beta_2); \rho).$$

(b) $(\log(Y_1), \log(cY_2))^\top \sim \text{log-BS}_2(\alpha_1, \log(\beta_1), \alpha_2, \log(c) + \log(\beta_2); \rho)$.

(c) $(\log(cY_1), \log(cY_2))^\top \sim \text{log-BS}_2(\alpha_1, \log(c) + \log(\beta_1), \alpha_2, \log(c) + \log(\beta_2); \rho)$.

Demonstração: As demonstrações dos itens (b) e (c) são análogas à do item (a).

6. Se $\boldsymbol{Y} = (Y_1, Y_2)^\top \sim \text{BS}_2(\alpha_1, \beta_1, \alpha_2, \beta_2; \rho$, então:

(a) $(\log(Y_1^{-1}), \log(Y_2))^\top \sim \text{log-BS}_2(\alpha_1, -\log(\beta_1), \alpha_2, \log(\beta_2); -\rho)$.

Demonstração: Se $\boldsymbol{Y} = (Y_1, Y_2)^\top \sim \text{BS}_2(\alpha_1, \beta_1, \alpha_2, \beta_2; \rho)$, então $(Y_1^{-1}, Y_2)^\top \sim \text{BS}_2(\alpha_1, \beta_1^{-1}, \alpha_2, \beta_2; -\rho)$. Deste modo, $(\log(Y_1^{-1}), \log(Y_2))^\top \sim \text{log-BS}_2(\alpha_1, \log(\beta_1^{-1}), \alpha_2, \log(\beta_2); -\rho)$. Isto é,

$$(\log(Y_1^{-1}), \log(Y_2))^\top \sim \text{log-BS}_2(\alpha_1, -\log(\beta_1), \alpha_2, \log(\beta_2); -\rho).$$

(b) $(\log(Y_1), \log(Y_2^{-1}))^\top \sim \text{log-BS}_2(\alpha_1, \log(\beta_1), \alpha_2, -\log(\beta_2); -\rho)$.

Demonstração: A Demonstração é análoga à do item (a).

(c) $(\log(Y_1^{-1}), \log(Y_2^{-1})^\top \sim \text{logBS}_2(\alpha_1, -\log(\beta_1), \alpha_2, -\log(\beta_2); \rho)$.

Demonstração: Se $\boldsymbol{Y} = (Y_1, Y_2)^\top \sim \text{BS}_2(\alpha_1, \beta_1, \alpha_2, \beta_2; \rho)$, então $(Y_1^{-1}, Y_2^{-1})^\top \sim \text{BS}_2(\alpha_1, \beta_1^{-1}, \alpha_2, \beta_2^{-1}; \rho)$. Assim, $(\log(Y_1^{-1}), \log(Y_2^{-1}))^\top \sim \text{log-BS}_2(\alpha_1, \log(\beta_1^{-1}), \alpha_2, \log(\beta_2^{-1}); \rho)$. Ou seja,

$$(\log(Y_1^{-1}), \log(Y_2^{-1}))^\top \sim \text{log-BS}_2(\alpha_1, -\log(\beta_1), \alpha_2, -\log(\beta_2); \rho).$$

7. Se $T = (Y_1, Y_2)^\top \sim \text{log-BS}_2(\alpha_1, \mu_1, \alpha_2, \mu_2; \rho)$, então $E(\boldsymbol{Y}) = (\mu_1, \mu_2)^\top$.

Demonstração: $E(\boldsymbol{Y}) = (E(Y_1), E(Y_2))^\top$. Se $(Y_1, Y_2)^\top \sim \text{log-BS}_2(\alpha_1, \mu_1, \alpha_2, \mu_2; \rho)$, então $Y_i \sim \text{log-BS}(\alpha_i, \mu_i)$ para $i = 1, 2$, portanto, $E(Y_i) = \mu_i$.

3.1.6 Distribuição log-BS Multivariada

Define-se a distribuição log-Birbaum-Saunders multivariada da mesma forma que definiu-se a distribuição log-Birbaum-Saunders bivariada, isto é, estende-se o conceito da distribuição log-Birnbaum-Saunders univariada.

Um vetor aleatório $\boldsymbol{Y} = (Y_1, \ldots, Y_n)^\top$ tem distribuição log-Birnbaum-Saunders multivariada com parâmetros $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)$, $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)$ e $\boldsymbol{\Sigma}$ se a f.d.a. de $\boldsymbol{Y}$ for expressa como

$$F_{\boldsymbol{Y}}(y_1, ..., y_n) = \Phi_n\left(\frac{2}{\alpha_1}\text{senh}\left(\frac{y_1 - \mu_1}{2}\right), \ldots, \frac{2}{\alpha_n}\,\text{senh}\left(\frac{y_n - \mu_n}{2}\right); \boldsymbol{\Sigma}\right),$$

em que $\alpha_1, \ldots, \alpha_n > 0$ e $\Phi_n(u_1, \ldots, u_n; \boldsymbol{\Sigma})$ é a f.d.a. do vetor normal padrão multivariado com matriz de correlação $\boldsymbol{\Sigma}$.

Assim, a f.d.p. da distribuição $\text{log-BS}_n(\boldsymbol{\alpha}, \boldsymbol{\mu}; \boldsymbol{\Sigma})$ é dada por

$$f_{\boldsymbol{Y}}(y_1, \ldots, y_n) = \phi_n\left(\frac{2}{\alpha_1}\text{senh}\left(\frac{y_1 - \mu_1}{2}\right), \ldots, \frac{2}{\alpha_n}\text{senh}\left(\frac{y_n - \mu_n}{2}\right); \Sigma\right)\prod_{i=1}^{n}\frac{1}{\alpha_i}\cosh\left(\frac{y_i - \mu_i}{2}\right),$$

em que $\phi_n(u_1, \ldots, u_n; \boldsymbol{\Sigma})$ é a f.d.p. do vetor normal padrão multivariado com matriz de correlação $\boldsymbol{\Sigma}$. Se $\boldsymbol{Y}$ é um vetor aleatório com distribuição log-BS n-variada, com parâmetros $\boldsymbol{\alpha}, \boldsymbol{\mu}$ e $\boldsymbol{\Sigma}$ utiliza-se a notação $\boldsymbol{Y} \sim \text{log-BS}_n(\boldsymbol{\alpha}, \boldsymbol{\mu}; \boldsymbol{\Sigma})$.

3.1.6.1 Propriedades do vetor aleatório log-BS multivariado

1. Se $\boldsymbol{Y} = (Y_1, ..., Y_n)^\top \sim \text{log-BS}_n(\boldsymbol{\alpha}, \boldsymbol{\mu}; \boldsymbol{\Sigma})$, então $Y_i \sim \text{log-BS}(\alpha_i, \mu_i), i = 1, \ldots, n$.

   Demonstração: Se $\boldsymbol{Y} = (Y_1, \ldots, Y_n) \sim \text{log-BS}_n(\boldsymbol{\alpha}, \boldsymbol{\mu}; \boldsymbol{\Sigma})$, então $\left(\frac{2}{\alpha_1}\text{senh}\left(\frac{y_1-\mu_1}{2}\right), \ldots \frac{2}{\alpha_n}\text{senh}\left(\frac{y_n-\mu_n}{2}\right)\right)^\top \sim \text{N}_n(\boldsymbol{0}, \boldsymbol{\Sigma})$.

   Considerando-se um vetor $n \times 1$ $\boldsymbol{v} = [v_k]$, com $v_k = \begin{cases} 0 & \text{se} \quad k \neq i \\ 1 & \text{se} \quad k = i \end{cases}$, tem-se que $\frac{2}{\alpha_i}\text{senh}\left(\frac{y_i-\mu_i}{2}\right) = \boldsymbol{v}^\top \left(\frac{2}{\alpha_1}\text{senh}\left(\frac{y_1-\mu_1}{2}\right), \ldots, \frac{2}{\alpha_n}\text{senh}\left(\frac{y_n-\mu_n}{2}\right)\right)^\top$, e, portanto, $\frac{2}{\alpha_i}\text{senh}\left(\frac{y_i-\mu_i}{2}\right) \sim \text{N}(\boldsymbol{v}^\top\boldsymbol{0}, \boldsymbol{v}^\top\boldsymbol{\Sigma}\boldsymbol{v}) = \text{N}(0, 1)$. O que implica, por definição, que $Y_i \sim \text{log-BS}(\alpha_i, \mu_i)$.

2. Sejam $\boldsymbol{p} = (p_1, \ldots, p_n)$ uma permutação de $(1, \ldots, n)$ e $\boldsymbol{b_p} = (b_{p_1}, \ldots, b_{p_n})^\top$ o vetor obtido permutando o vetor $\boldsymbol{b} = (b_1, \ldots, b_n)^\top$ segundo a permutação p. Seja ainda para toda matriz $\boldsymbol{B}$, $\boldsymbol{B_{pp}}$ a matriz cujas linhas e colunas foram permutadas segundo a permutação p. Assim, se $\boldsymbol{Y} \sim \text{log-BS}_n(\boldsymbol{\alpha}, \boldsymbol{\beta}; \boldsymbol{\Sigma})$ então $\boldsymbol{Y_p} \sim \text{log-BS}_n(\boldsymbol{\alpha_p}, \boldsymbol{\beta_p}; \boldsymbol{\Sigma_{pp}})$.

   Demontração: A demontração é análoga a da propriedade 2 da Seção 3.1.3.1 considerando

   $$\boldsymbol{u} = \left(\frac{2}{\alpha_1}\text{ senh}\left(\frac{Y_1-\mu_1}{2}\right), \ldots, \frac{2}{\alpha_n}\text{senh}\left(\frac{Y_n-\mu_n}{2}\right)\right).$$

   Assim, como no caso da distribuição BS multivariada, para as próximas demonstrações consideram-se $\boldsymbol{Y} \sim \text{log-BS}_n(\boldsymbol{\alpha}, \boldsymbol{\beta}; \boldsymbol{\Sigma})$ e as seguintes partições em $\boldsymbol{T}, \boldsymbol{\alpha}, \boldsymbol{\beta}$ e $\boldsymbol{\Sigma}$: $\boldsymbol{Y} = \left(\boldsymbol{Y_1}^\top, \boldsymbol{Y_2}^\top\right)^\top, \boldsymbol{\alpha} = \left(\boldsymbol{\alpha_1}^\top, \boldsymbol{\alpha_2}^\top\right)^\top, \boldsymbol{\beta} = \left(\boldsymbol{\beta_1}^\top, \boldsymbol{\beta_2}^\top\right)^\top$ e $\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma_{11}} & \boldsymbol{\Sigma_{12}} \\ \boldsymbol{\Sigma_{21}} & \boldsymbol{\Sigma_{22}} \end{pmatrix}$, em que $\boldsymbol{Y_1}, \boldsymbol{\alpha_1}, \boldsymbol{\beta_1}$ são vetores $q \times 1$; $\boldsymbol{Y_2}, \boldsymbol{\alpha_2}, \boldsymbol{\beta_2}$ são vetores $p \times 1$; $\boldsymbol{\Sigma_{11}}$ é uma matriz $q \times q$; $\boldsymbol{\Sigma_{22}}$ é uma matriz $p \times p$; $\boldsymbol{\Sigma_{12}}$ é uma matriz $q \times p$ e $\boldsymbol{\Sigma_{21}} = \boldsymbol{\Sigma_{12}}^\top$, com $p + q = n$.

3. Se $\boldsymbol{Y} = \left(\boldsymbol{Y_1}^\top, \boldsymbol{Y_2}^\top\right)^\top \sim \text{log-BS}_n(\boldsymbol{\alpha}, \boldsymbol{\mu}; \boldsymbol{\Sigma})$, então:

   (a) $\boldsymbol{Y_1} \sim \text{log-BS}_q(\boldsymbol{\alpha_1}, \boldsymbol{\mu_1}; \boldsymbol{\Sigma_{11}})$.

   Demonstração: Considere a matriz $\boldsymbol{A} = [a_{ij}], q \times n$, em que $a_{ij} = \begin{cases} 1 & \text{se} \quad i = j \\ 0 & \text{se} \quad i \neq j \end{cases}$

   Como $\boldsymbol{Y} \sim \text{log-BS}_n(\boldsymbol{\alpha}, \boldsymbol{\beta}; \boldsymbol{\Sigma})$, tem-se que

   $$\left(\frac{2}{\alpha_1}\text{senh}\left(\frac{y_1-\mu_1}{2}\right), \ldots, \frac{2}{\alpha_n}\text{senh}\left(\frac{y_n-\mu_n}{2}\right)\right) \sim \text{N}_n(\boldsymbol{0}, \boldsymbol{\Sigma}).$$

   Como $\boldsymbol{AY} = \boldsymbol{Y_1}$, segue que

   $$\left(\frac{2}{\alpha_1}\text{senh}\left(\frac{y_1-\mu_1}{2}\right), \ldots, \frac{2}{\alpha_q}\text{senh}\left(\frac{y_q-\mu_q}{2}\right)\right) \sim N_q\left(\boldsymbol{A0}, \boldsymbol{A\Sigma A}^\top\right) = N_q(\boldsymbol{0}, \boldsymbol{\Sigma_{11}}),$$

portanto, $\boldsymbol{Y_1} \sim \text{log-BS}_q\left(\boldsymbol{\alpha_1}, \boldsymbol{\beta_1}; \boldsymbol{\Sigma_{11}}\right)$.

(b) $\boldsymbol{Y_2} \sim \text{log-BS}_p\left(\boldsymbol{\alpha_2}, \boldsymbol{\mu_2}; \boldsymbol{\Sigma_{22}}\right)$.

Demonstração: A demonstração, neste caso, é análoga à do item (a), considerando a matriz $\boldsymbol{A} = [a_{ij}],\ p \times n$, em que $a_{ij} = \begin{cases} 1 & \text{se} \quad j = q + i \\ 0 & \text{caso contrário} \end{cases}$ e observando-se que $\boldsymbol{AY} = \boldsymbol{Y_2}$ e que $\boldsymbol{A\Sigma A}^\top = \boldsymbol{\Sigma_{22}}$.

Esta propriedade em conjunto com a propriedade 2 assegura que qualquer subvetor de $\boldsymbol{Y}$ tem distribuição marginal log-BS.

4. Se $\boldsymbol{Y} = \left(\boldsymbol{Y_1}^\top, \boldsymbol{Y_2}^\top\right)^\top \sim \text{log-BS}_n\left(\left(\boldsymbol{\alpha_1}^\top, \boldsymbol{\alpha_2}^\top\right)^\top, \left(\boldsymbol{\mu_1}^\top, \boldsymbol{\mu_2}^\top\right)^\top; \boldsymbol{\Sigma}\right)$, então:

(a) $\left(c\boldsymbol{Y_1}^\top, \boldsymbol{Y_2}^\top\right)^\top \sim \text{SHN}_n\left(\left(\boldsymbol{\alpha_1}^\top, \boldsymbol{\alpha_2}^\top\right)^\top, \left(c\boldsymbol{\mu_1}^\top, \boldsymbol{\mu_2}^\top\right)^\top, \left(2|c| \cdot \mathbf{1}_q^\top, 2 \cdot \mathbf{1}_p^\top\right)^\top; \boldsymbol{\Sigma}\right)$, se $c > 0$ e $\left(c\boldsymbol{Y_1}^\top, \boldsymbol{Y_2}^\top\right)^\top \sim \text{SHN}_n\left(\left(\boldsymbol{\alpha_1}^\top, \boldsymbol{\alpha_2}^\top\right)^\top, \left(c\boldsymbol{\mu_1}^\top, \boldsymbol{\mu_2}^\top\right)^\top, \left(2|c| \cdot \mathbf{1}_q^\top, 2 \cdot \mathbf{1}_p^\top\right)^\top; \boldsymbol{\Sigma_1}\right)$, se $c < 0$, em que $\boldsymbol{\Sigma_1}$ é dada como na propriedade 5 da Seção 3.1.3.1.
Demonstração: Seja $\boldsymbol{Y} = (Y_1, \ldots, Y_n)$ e $\boldsymbol{X} = (cY_1, \ldots, cY_q, Y_{q+1}, \ldots, Y_n)$. Assim, $f_{\boldsymbol{X}}(y_1, \ldots y_n) = f_{\boldsymbol{Y}}\left(\frac{1}{c}y_1, \ldots, \frac{1}{c}y_q, y_{q+1}, \ldots, y_n\right) det(\boldsymbol{A})$, em que $\boldsymbol{A} = \begin{pmatrix} \frac{1}{c}\boldsymbol{I}_q & \mathbf{0} \\ \mathbf{0} & \boldsymbol{I}_p \end{pmatrix}$.
Logo,

$$\begin{aligned} f_{\boldsymbol{X}}(y_1, \ldots y_n) &= \phi_n\left(u_1\left(\frac{\frac{y_1}{c} - \mu_1}{2}\right), \ldots, u_q\left(\frac{\frac{y_q}{c} - \mu_q}{2}\right), u_{q+1}\left(\frac{y_{q+1} - \mu_{q+1}}{2}\right), \ldots, u_n\left(\frac{y_n - \mu_n}{2}\right)\right) \cdot \\ &\quad \prod_{i=1}^{q} \frac{1}{\alpha_i}\cosh\left(\frac{\frac{y_i}{c} - \mu_i}{2}\right) \prod_{i=q+1}^{n} \frac{1}{\alpha_i}\cosh\left(\frac{y_i - \mu_i}{2}\right) \frac{1}{|c|^q}. \end{aligned}$$

em que $u_i(x) = \frac{2}{\alpha_i}\text{senh}(x)$.
Para $c > 0$ tem-se:

$$\begin{aligned} f_{\boldsymbol{X}}(y_1, \ldots y_n) &= \phi_n\left(u_1\left(\frac{y_1 - c\mu_1}{2|c|}\right), \ldots, u_q\left(\frac{y_q - c\mu_q}{2|c|}\right), u_{q+1}\left(\frac{y_{q+1} - \mu_{q+1}}{2}\right), \ldots, u_n\left(\frac{y_n - \mu_n}{2}\right)\right) \\ &\quad \prod_{i=1}^{q} \frac{1}{|c|\alpha_i}\cosh\left(\frac{y_i - c\mu_i}{2|c|}\right) \prod_{i=q+1}^{n} \frac{1}{\alpha_i}\cosh\left(\frac{y_i - \mu_i}{2}\right). \end{aligned}$$

que é a f.d.p. da distribuição $\text{SHN}_n\left(\left(\boldsymbol{\alpha_1}^\top, \boldsymbol{\alpha_2}^\top\right)^\top, \left(c\boldsymbol{\mu_1}^\top, \boldsymbol{\mu_2}^\top\right)^\top, \left(2|c| \cdot \mathbf{1}_q^\top, 2 \cdot \mathbf{1}_p^\top\right)^\top; \boldsymbol{\Sigma}\right)$.
Agora se $c < 0$, tem-se que

$$\begin{aligned} f_X(\boldsymbol{y}) &= f_Y\left(\frac{1}{c}y_1, \ldots, \frac{1}{c}y_q, y_{q+1}, \ldots, y_n\right) det(\boldsymbol{A}) \\ &= \phi_n\left(u_1\left(\frac{\frac{y_1}{c} - \mu_1}{2}\right), \ldots, u_q\left(\frac{\frac{y_q}{c} - \mu_q}{2}\right), u_{q+1}\left(\frac{y_{q+1} - \mu_{q+1}}{2}\right), \ldots, u_n\left(\frac{y_n - \mu_n}{2}\right)\right) \\ &\quad \prod_{i=1}^{q} \frac{1}{\alpha_i}\cosh\left(\frac{\frac{y_i}{c} - \mu_i}{2}\right) \prod_{i=q+1}^{n} \frac{1}{\alpha_i}\cosh\left(\frac{y_i - \mu_i}{2}\right) \frac{1}{|c|^q} \\ &= \phi_n\left(-u_1\left(\frac{y_1 - c\mu_1}{2|c|}\right), \ldots, -u_q\left(\frac{y_q - c\mu_q}{2|c|}\right), u_{q+1}\left(\frac{y_{q+1} - \mu_{q+1}}{2}\right), \ldots, u_n\left(\frac{y_n - \mu_n}{2}\right)\right) \\ &\quad \prod_{i=1}^{q} \frac{1}{|c|\alpha_i}\cosh\left(\frac{y_i - c\mu_i}{2|c|}\right) \prod_{i=q+1}^{n} \frac{1}{\alpha_i}\cosh\left(\frac{y_i - \mu_i}{2}\right). \end{aligned}$$

que é a f.d.p. da distribuição $\text{SHN}_n\left(\left(\boldsymbol{\alpha_1}^\top, \boldsymbol{\alpha_2}^\top\right)^\top, \left(c\boldsymbol{\mu_1}^\top, \boldsymbol{\mu_2}^\top\right)^\top, \left(2|c| \cdot \mathbf{1}_q^\top, 2 \cdot \mathbf{1}_p^\top\right)^\top; \boldsymbol{\Sigma_1}\right)$.

(b) $\left(\boldsymbol{Y_1}^\top, c\boldsymbol{Y_2}^\top\right)^\top \sim \text{SHN}_n\left(\left(\boldsymbol{\alpha_1}^\top, \boldsymbol{\alpha_2}^\top\right)^\top, \left(\boldsymbol{\mu_1}^\top, c\boldsymbol{\mu_2}^\top\right)^\top, \left(2\cdot\mathbf{1}_q^\top, 2|c|\cdot\mathbf{1}_p^\top\right)^\top; \boldsymbol{\Sigma}\right)$, se $c > 0$ e $\left(\boldsymbol{Y_1}^\top, c\boldsymbol{Y_2}^\top\right)^\top \sim \text{SHN}_n\left(\left(\boldsymbol{\alpha_1}^\top, \boldsymbol{\alpha_2}^\top\right)^\top, \left(\boldsymbol{\mu_1}^\top, c\boldsymbol{\mu_2}^\top\right)^\top, \left(2\cdot\mathbf{1}_q^\top, 2|c|\cdot\mathbf{1}_p^\top\right)^\top; \boldsymbol{\Sigma_1}\right)$, se $c < 0$.

Demonstração: A demonstração é análoga à do item (a).

(c) $\left(c\boldsymbol{Y_1}^\top, c\boldsymbol{Y_2}^\top\right)^\top \sim \text{SHN}_n\left(\left(\boldsymbol{\alpha_1}^\top, \boldsymbol{\alpha_2}^\top\right)^\top, \left(c\boldsymbol{\mu_1}^\top, c\boldsymbol{\mu_2}^\top\right)^\top, \left(2|c|\mathbf{1}_q^\top, 2|c|\mathbf{1}_p^\top\right)^\top; \boldsymbol{\Sigma}\right)$.

Demonstração: A demonstração, neste caso, é análoga às dos itens anteriores, lembrando-se que $\Phi_n(-u_1, -u_2, \ldots, -u_n; \boldsymbol{\Sigma}) = \Phi_n(u_1, u_2, \ldots, u_n; \boldsymbol{\Sigma})$.

5. Se $\boldsymbol{Y} = (\boldsymbol{Y_1}, \boldsymbol{Y_2})^\top \sim \text{log-BS}_n\left(\left(\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}\right)^\top, \left(\boldsymbol{\mu_1}, \boldsymbol{\mu_2}\right)^\top; \boldsymbol{\Sigma}\right)^\top$, então:

(a) $(-\boldsymbol{Y_1}, \boldsymbol{Y_2})^\top \sim \text{log-BS}_n\left(\left(\left(\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}\right)^\top, \left(-\boldsymbol{\mu_1}, \boldsymbol{\mu_2}\right)^\top; \boldsymbol{\Sigma}_1\right)\right)^\top$;

(b) $(\boldsymbol{Y_1}, -\boldsymbol{Y_2})^\top \sim \text{log-BS}_n\left(\left(\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}\right)^\top, \left(\boldsymbol{\mu_1}, -\boldsymbol{\mu_2}\right)^\top; \boldsymbol{\Sigma}_1\right)^\top$;

(c) $(-\boldsymbol{Y_1}, -\boldsymbol{Y_2})^\top \sim \text{log-BS}_n\left(\left(\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}\right)^\top, \left(-\boldsymbol{\mu_1}, -\boldsymbol{\mu_2}\right)^\top; \boldsymbol{\Sigma}\right)^\top$.

Demonstração: O resultado segue diretamente da propriedade 4 ao se fazer $c = -1$.

6. Se $(Y_1, \ldots, Y_n) \sim \text{BS}_n(\boldsymbol{\alpha}, \boldsymbol{\gamma}; \boldsymbol{\Sigma})$, então $(\log(Y_1), \ldots, \log(Y_n)) \sim \text{log-BS}(\boldsymbol{\alpha}, \log(\boldsymbol{\gamma}); \boldsymbol{\Sigma})$.

Demonstração: Considere $\boldsymbol{X} = (X_1, \ldots, X_n) = (\log(Y_1), \ldots, \log(Y_n))$. Assim,

$$\begin{aligned} F_{\boldsymbol{X}} &= \text{P}(X_1 \leq y_1, \ldots, X_n \leq y_n) = \text{P}(Y_1 \leq \exp(y_1), \ldots, Y_n \leq \exp(y_n)) \\ &= \phi_n\left(\alpha_1^{-1}\left(\left(\frac{\exp(y_1)}{\beta_1}\right)^{\frac{1}{2}} - \left(\frac{\exp(y_1)}{\beta_1}\right)^{-\frac{1}{2}}\right), \ldots, \alpha_n^{-1}\left(\left(\frac{\exp(y_n)}{\beta_n}\right)^{\frac{1}{2}} - \left(\frac{\exp(y_n)}{\beta_n}\right)^{-\frac{1}{2}}\right); \boldsymbol{\Sigma}\right) \\ &= \phi_n\left(\frac{2}{\alpha_1}\text{senh}\left(\frac{y_1 - \log(\gamma_1)}{2}\right), \ldots, \frac{2}{\alpha_n}\text{senh}\left(\frac{y_n - \log(\gamma_n)}{2}\right); \boldsymbol{\Sigma}\right), \end{aligned}$$

que é a f.d.a. da distribuição $\text{log-BS}(\boldsymbol{\alpha}, \log(\boldsymbol{\gamma}); \boldsymbol{\Sigma})$.

7. Se $\boldsymbol{Y} = (\boldsymbol{Y_1}, \boldsymbol{Y_2})^\top \sim \text{BS}_n\left(\left(\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}\right)^\top, \left(\boldsymbol{\mu_1}, \boldsymbol{\mu_2}\right)^\top; \boldsymbol{\Sigma}\right)^\top$, então $(\log(c\boldsymbol{Y_1}), \log(\boldsymbol{Y_2}))^\top \sim \text{log-BS}_n\left(\left(\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}\right)^\top, \left(\log(c\boldsymbol{\mu_1}), \log(\boldsymbol{\mu_2})\right)^\top; \boldsymbol{\Sigma}\right)^\top$.

Demonstração: Se $(\boldsymbol{Y_1}, \boldsymbol{Y_2})^\top \sim \text{BS}_n\left(\left(\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}\right)^\top, \left(\boldsymbol{\mu_1}, \boldsymbol{\mu_2}\right)^\top; \boldsymbol{\Sigma}\right)^\top$, então $(c\boldsymbol{Y_1}, \boldsymbol{Y_2})^\top \sim \text{BS}_n\left(\left(\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}\right)^\top, \left(c\boldsymbol{\mu_1}, \boldsymbol{\mu_2}\right)^\top; \boldsymbol{\Sigma}\right)^\top$ e, consequentemente, $(\log(c\boldsymbol{Y_1}), \log(\boldsymbol{Y_2}))^\top \sim \text{log-BS}_n\left(\left(\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}\right)^\top, \left(\log(c\boldsymbol{\mu_1}), \log(\boldsymbol{\mu_2})\right)^\top; \boldsymbol{\Sigma}\right)^\top$.

8. Se $\boldsymbol{Y} = (\boldsymbol{Y_1}, \boldsymbol{Y_2})^\top \sim \text{BS}_n\left(\left(\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}\right)^\top, \left(\boldsymbol{\mu_1}, \boldsymbol{\mu_2}\right)^\top; \boldsymbol{\Sigma}\right)^\top$, então:

(a) $\left(\log(\boldsymbol{Y_1}^{-1}), \log(\boldsymbol{Y_2})\right)^\top \sim \text{log-BS}_n\left(\left(\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}\right)^\top, \left(-\log(\boldsymbol{\mu_1}), \log(\boldsymbol{\mu_2})\right)^\top; \boldsymbol{\Sigma}_1\right)$.

Demonstração: Se $(\boldsymbol{Y_1}, \boldsymbol{Y_2})^\top \sim \text{BS}_n\left(\left(\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}\right)^\top, \left(\boldsymbol{\mu_1}, \boldsymbol{\mu_2}\right)^\top; \boldsymbol{\Sigma}\right)^\top$, então $\left(\boldsymbol{Y_1}^{-1}, \boldsymbol{Y_2}\right)^\top \sim \text{BS}_n\left(\left(\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}\right)^\top, \left(\boldsymbol{\mu_1}^{-1}, \boldsymbol{\mu_2}\right)^\top; \boldsymbol{\Sigma}_1\right)$. Portanto, $\left(\log(\boldsymbol{Y_1}^{-1}), \log(\boldsymbol{Y_2})\right)^\top \sim \text{log-BS}_n\left(\left(\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}\right)^\top, \left(\log(\boldsymbol{\mu_1}^{-1}), \log(\boldsymbol{\mu_2})\right)^\top; \boldsymbol{\Sigma}_1\right)^\top = \text{log-BS}_n\left(\left(\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}\right)^\top, \left(-\log(\boldsymbol{\mu_1}), \log(\boldsymbol{\mu_2})\right)^\top; \boldsymbol{\Sigma}_1\right)^\top$.

(b) $\left(\log(\boldsymbol{Y_1}), \log(\boldsymbol{Y_2}^{-1})\right)^\top \sim \text{log-BS}_n\left(\left(\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}, \right)^\top, \left(\log(\boldsymbol{\mu_1}), -\log(\boldsymbol{\mu_2})\right)^\top; \boldsymbol{\Sigma}_1\right)^\top$.

(c) $\left(\log(\boldsymbol{Y_1}^{-1}), \log(\boldsymbol{Y_2}^{-1})\right)^\top \sim \text{log-BS}_n\left(\left(\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}\right)^\top, \left(-\log(\boldsymbol{\mu_1}), -\log(\boldsymbol{\mu_2})\right)^\top; \boldsymbol{\Sigma}\right)^\top$.

Demonstração: As demonstrações dos itens (b) e (c) são análogas à do item (a).

9. Se $\boldsymbol{Y} = (\boldsymbol{Y_1}, \boldsymbol{Y_2})^\top \sim \text{log-BS}_n \left((\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2})^\top, (\boldsymbol{\mu_1}, \boldsymbol{\mu_2})^\top ; \boldsymbol{\Sigma} \right)^\top$, então $E(\boldsymbol{Y}) = \boldsymbol{\mu}$.

   Demonstração: $\boldsymbol{Y} = (Y_1, \ldots, Y_n)^\top \sim \text{log-BS}_n \left((\alpha_1, \ldots, \alpha_n)^\top, (\mu_1, \ldots, \mu_n)^\top ; \boldsymbol{\Sigma} \right)^\top$, então $Y_i \sim \text{log-BS}(\alpha_i, \mu_i),\ i = 1, \ldots, n$, e, consequentemente $E(Y_i) = \mu_i$. Logo, $\boldsymbol{E(Y)} = (\mu_1, \ldots, \mu_n)^\top$.

3.2 Estatística espacial

O estudo da variabilidade espacial (caracterização e modelagem) é o principal objetivo da geoestatística. Esta teoria, criada por Matheron (MATHERON, 1963), tem como princípio fundamental a depedência ou continuidade espacial. Este princípio diz que a variável de interesse está associada ao espaço físico que ocupa, e, mais que isso, se a variável regionalizada é definida como uma função $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ a medida que um ponto $P \in \mathbb{R}^n$ se aproxima de $P_0 \in \mathbb{R}^n$, $f(P)$ se aproxima de $f(P_0)$, isto é, quando P se aproxima de P_0 o valor da variável em P se aproxima do valor da variável em P_0.

Diante da impossibilidade de se obterem valores de determinada variável em todos os pontos do espaço de interesse, a geoestatística tem como objetivo caracterizar a variabilidade espacial de variáveis correlaçionadas a partir de um conjunto discreto de observações. Por exemplo, no caso de estudo de característica do solo, é comum determinar uma malha de pontos georeferenciados na superfície, onde são efetuadas sondagens, de maneira a avaliar o teor de determinado minério. Deste modo, tem-se um conjunto de dados segundo uma malha bidimensional, que permitirão avaliar o teor do minério em toda área sob estudo. A Figura 6 ilustra uma malha bidimensional regular.

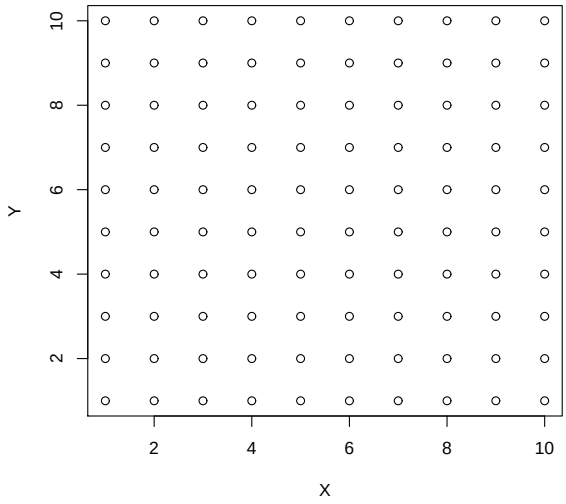


Figura 6 Malha bidimensional regular.

A exatidão da avaliação pode depender do espaçamento das amostras (GUEDES et al., 2011) e também da variabilidade espacial da variável.

3.2.1 Modelos geoestatísticos

Em modelos geoestatísticos, um valor localizado espacialmente em $\boldsymbol{s_i}$ é interpretado como uma realização $z(\boldsymbol{s_i})$ da variável aleatória $Z(\boldsymbol{s_i})$. No espaço $D \subset \mathbb{R}^n$, no qual se distribui o conjunto de amostras, têm-se as realizações de n variáveis aleatórias $Z(\boldsymbol{s_1}), \ldots, Z(\boldsymbol{s_n})$, correlacionadas entre si. Para cada uma das variáveis aleatórias, têm-se definidas a média de $Z(\boldsymbol{s_i})$, $\mu_i = \mathrm{E}(Z(\boldsymbol{s_i}))$, a variância de $Z(\boldsymbol{s_i})$, $\mathrm{Var}(Z(\boldsymbol{s_i})) = \mathrm{E}(Z(\boldsymbol{s_i})^2) - \mathrm{E}^2(Z(\boldsymbol{s_i}))$ e, para duas variáveis aleatórias, $Z(\boldsymbol{s_i})$ e $Z(\boldsymbol{s_j})$, a covariância $\mathrm{C}(Z(\boldsymbol{s_i}), Z(\boldsymbol{s_j})) = \mathrm{E}(Z(\boldsymbol{s_i})Z(\boldsymbol{s_j})) - \mathrm{E}(Z(\boldsymbol{s_i}))\mathrm{E}(Z(\boldsymbol{s_j}))$, a correlação $\rho(Z(\boldsymbol{s_i}), Z(\boldsymbol{s_j})) = \mathrm{C}(Z(\boldsymbol{s_i}), Z(\boldsymbol{s_j}))/\sqrt{\mathrm{Var}(Z(\boldsymbol{s_i}))}\sqrt{\mathrm{Var}(Z(\boldsymbol{s_i}))}$, e a semivariância, $\gamma(Z(\boldsymbol{s_i}), Z(\boldsymbol{s_j})) = \frac{1}{2}\mathrm{E}((Z(\boldsymbol{s_i}) - Z(\boldsymbol{s_j}))^2)$ entre $Z(\boldsymbol{s_i})$ e $Z(\boldsymbol{s_j})$, $i, j = 1, \ldots, n$.

Geralmente, o conjunto das variáveis aleatórias constitui um vetor de variáveis aleatórias n-variado $\boldsymbol{Z} = (Z(\boldsymbol{s_1}), \ldots, Z(\boldsymbol{s_n}))^T$, do qual só se conhece uma realização. Com apenas uma realização, teoricamente, é impossível estimar qualquer parâmetro estatístico. Assim, de modo a permitir a inferência sobre algumas estatísticas, a teoria supõe que $\{Z(\boldsymbol{s}), s \in D\}$ é um processo estocástico estacionário (MARDIA; MARSHALL, 1984), isto é, supõe que:

- Todas as variáveis aleatórias têm a mesma média, isto é,

$$\mathrm{E}(Z(s_1)) = \cdots = \mathrm{E}(Z(\boldsymbol{s_n})) = \mathrm{E}(Z(\boldsymbol{s})) = \mu.$$

- A correlação entre duas variáveis aleatórias depende somente da distância espacial que as separa, isto é,

$$\mathrm{C}(Z(\boldsymbol{s_i}), Z(\boldsymbol{s_j})) = \mathrm{C}(Z(\boldsymbol{s_i}), Z(\boldsymbol{s_i} + h_{ij})) = C(h_{ij})$$

  em que h_{ij} é a distância euclidiana entre $\boldsymbol{s_i}$ e $\boldsymbol{s_j}$.

  Consequentemente,

$$\rho(Z(\boldsymbol{s_i}), Z(\boldsymbol{s_j})) = \frac{C(h_{ij})}{C(0)} = \rho(h_{ij})$$

  E, admitindo as hipóteses de estacionaridade, têm-se

$$\gamma(h_{ij}) = \mathrm{C}(0) - \mathrm{C}(h_{ij}) \qquad \text{e} \qquad \rho(h_{ij}) = 1 - \frac{\gamma(h_{ij})}{\mathrm{C}(0)}. \tag{3.3}$$

3.2.1.1 Ferramentas estatísticas para o estudo da variabilidade espacial

Considere $s \in D \subset \mathbb{R}^n$ e uma variável quantitativa $Z(\boldsymbol{s})$. Os diagramas de representação dos pares de pontos $Z(\boldsymbol{s}) \times Z(\boldsymbol{s} + h)$, para diferentes valores de h, h-*scatterplots*, contêm informações ricas sobre a variabilidade espacial da variável (ISAAKS; SRIVASTAVA, 1989; SOARES, 2000). Por exemplo, em um 1-*scatterplot*, a coordenada x é o valor da variável em determinada localização e a coordenada y é o valor da variável em uma localização que está a uma unidade de medida da localização inicial, x. Assim, se os valores dos dados de locais separados por uma distância h são muito semelhantes, os pontos do h-*scatterplot* ficam perto da reta $y = x$ (Figura 7). Ainda é possível constuir h-*scatterplot* considerando-se diferentes

direções. Por exemplo, em um (0,1)-*scatterplot*a coordenada y é o valor de uma variável que está localizada a uma unidade de medida ao norte da localização x, enquanto em um (1,0)-*scatterplot* a coordenada y é o valor de uma variável que está localizada a uma unidade de medida ao sul da localização x.

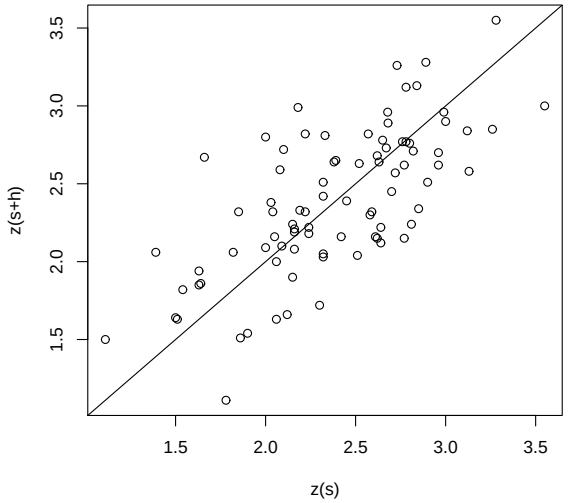


Figura 7 Ilustração de um h-*scatterplot*.

A forma das "nuvens de pontos"dos h-*scatterplots* fornece informações, tais como:

- se existe isotropia, isto é, se o comportamento da variável é o mesmo em todas as direções. Quando isso ocorre, o comportamento dos h-*scatterplots* é semelhante para diferentes direções;
- se existe ou não continuidade espacial, isto é, se os valores da variável em pontos mais próximos são mais semelhantes do que em pontos mais distantes. Quando existe continuidade espacial, os valores da variável tornam-se menos semelhantes à medida que h cresce. Portanto, a "nuvem de pontos"vai ficando mais "espalhada"conforme h cresce;
- se existem pontos com um comportamento não usual, isto é, o valor da variável é muito diferente, quando comparado com pontos em sua vizinhança. Um ponto discrepante é um valor solitário em um gráfico, pois apresenta uma relação não usual com os outros valores. Deve-se dispensar uma atenção especial a esses pontos, pois as estatísticas utilizadas para estudar a variabilidade espacial podem ser influenciadas consideravelmente por esses valores discrepantes.

Apesar dos h-*scatterplots* serem bastante eficientes no estudo da variabilidade espacial, o estudo desta por h-*scatterplots* é muito trabalhoso. Sendo assim, é usual utilizar instrumentos que os sintetizam. Por exemplo, um instrumento que quantifica a continuidade de $Z(\boldsymbol{s})$ é a

média do quadrado das diferenças entre $z(\boldsymbol{s})$ e $z(\boldsymbol{s}+h)$, isto é, o estimador do semivariância dado por

$$\widehat{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} \left(z(\boldsymbol{s_i}) - z(\boldsymbol{s_i}+h)\right)^2,$$

em que $N(h)$ é o número de pares de pontos para cada distância h.

Cada h-*scatterplot* dá origem a um valor de $\widehat{\gamma}(h)$. Representando num gráfico os valores de $\widehat{\gamma}(h)$, em função de valores possíveis para h, tem-se o semivariograma experimental, também denominado semivariograma clássico de Matheron. Os modelos espaciais utilizados para a estimação da estrutura de dependência espacial consideram que a variável é isotrópica, isto é, que a variável apresenta o mesmo padrão de variabilidade espacial em todas as direções (JOURNEL; HUIJBREGTS, 1978). Portanto, o estudo da dependência espacial pode ser feito por um único semivariograma, chamado de semivariograma omnidirecional. Se a variabilidade espacial apresenta comportamento diferente em diferentes direções, isto é, é anisotrópica, faz-se necessário corrigi-la antes de se utilizar o semivariograma omnidirecional. Para mais detalhes sobre anisotropia ver Isaaks e Srivastava (1989), Diggle e Ribeiro Júnior (2007), Guedes et al. (2008). As direções para a construção dos semivariogramas direcionais são infinitas, porém as direções apontadas como convencionais são: 0º, 45º, 90º e 135º (ISAAKS; SRIVASTAVA, 1989). Por convenção, esses ângulos são obtidos a partir da direção norte, no sentido horário, como pode ser observado na Figura 8.

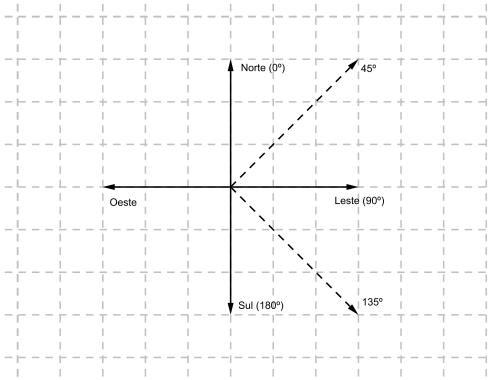


Figura 8 Principais direções para o semivariograma direcional.

Devido às relações estabelecidas em (3.3), estudar a variabilidade espacial por meio dos semivariogramas é equivalente a estudar a variabilidade utilizando as funções covariância ou correlação. A Figura 9 ilustra um semivariograma experimental.

Uma vez obtido o semivariograma experimental, a fim de conhecer a semivariância para qualquer valor de $h \in \mathbb{R}$, ajusta-se um modelo teórico a esse . Assume-se que tal modelo é dado por uma função de três parâmetros (MARDIA; MARSHALL, 1984), que serão estimados: efeitos pepita ou *nugget* (φ_1), contribuição ou *sill* (φ_2) e alcance $(a = g(\varphi_3))$. Supõe-se ainda que $Z(\boldsymbol{s_1}), \ldots, Z(\boldsymbol{s_n})$ são dados pelo modelo

$$Z(\boldsymbol{s_i}) = \mu(\boldsymbol{s_i}) + \epsilon(\boldsymbol{s_i}),$$

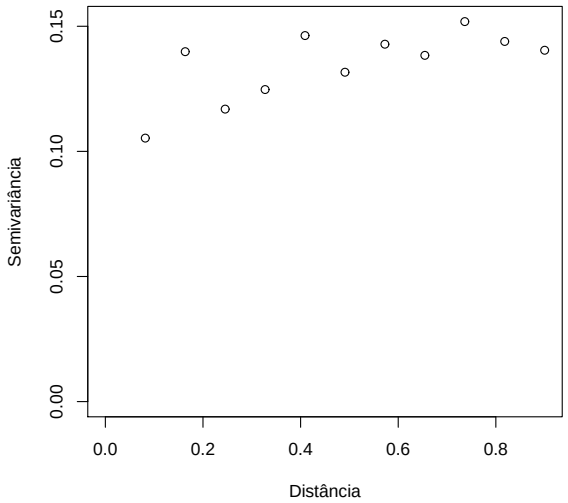


Figura 9 Ilustração de um semivariograma experimental.

em que $\mu(\cdot)$ é o termo determinístico e $\varepsilon(\cdot)$ é o erro do modelo. Assume-se também que $\mu(\boldsymbol{s_i}) = \sum_{j=0}^{p} \beta_j x_j(\boldsymbol{s_i})$, em que $x_j(\boldsymbol{s_i})$ são funções conhecidas de $\boldsymbol{s_i}$, β_i são parâmetros desconhecidos a serem estimados, e os erros têm média zero, $\mathrm{E}(\epsilon(\mathrm{s_i})) = 0$. Isso implica que $\mathrm{E}(Y(\boldsymbol{s_i})) = \mu(\boldsymbol{s_i})$.

Assume-se ainda que a dependência espacial é determinada por uma matriz escala $\boldsymbol{\Sigma}, n \times n$, simétrica, não singular e positiva definida, proporcional a $\mathrm{Cov}\,(\boldsymbol{\varepsilon})$, especificada pela forma paramétrica

$$\Sigma = \varphi_1 \boldsymbol{I}_n + \varphi_2 \boldsymbol{R}(\varphi_3),$$

em que $\boldsymbol{R}(\varphi_3) = [r_{ij}]$, é uma matriz simétrica, $n \times n$, com elementos da diagonal principal iguais a 1 e que depende do modelo teórico de variância adotado para modelar o semivariograma experimental. Os parâmetros que descrevem a depedência espacial são descritos a seguir e ilustrados na Figura 10.

- Efeito pepita (φ_1): teoricamente, $\gamma(0) = 0$, e, devido à continuidade, à medida que h se aproxima de zero, $\gamma(h)$ se aproxima de zero. Na prática, porém, à medida que h se aproxima de zero, $\gamma(h)$ se aproxima de um valor positivo, φ_1, ao qual denominamos efeito pepita. O que ocorre é que, na prática há, um valor mínimo de h entre amostras para o qual o valor $\gamma(h)$ pode ser quantificado. Quando esse valor é elevado, tem-se que há uma grande variabilidade à pequena escala. Há uma descontinuidade do semivariograma a uma escala não captada pelas amostras. Essa descontinuidade deve-se basicamente a dois fatores: à variabilidade de pequena escala não captada pela malha de amostragem e dos possíveis erros de medida. O termo efeito pepita deve-se, possivelmente, a um fato ligado ao processo de mineração: nas jazidas de ouro, é comum a ocorrência de pepitas (*nugget*), situação na qual acontece uma alternância de teores elevados de ouro (quando ocorrem pepitas), e baixos teores (quando as pepitas não estão presentes). Não existe relação entre o espaço físico e a formação da pepita. Assim, quando a variabilidade

acontece ao acaso, ou seja, não há relação entre a característica e o espaço físico, dizemos que ocorre um efeito pepita.

- Contribuição (φ_2): de um modelo geral as funções semivariâncias tendem assintóticamente para o valor $\varphi_1 + \varphi_2$, chamado de patamar. A partir do patamar o semivariograma se estabiliza. O patamar está relacionado com a variância total do processo.

- Alcance ($a = g(\varphi_3)$): é a distância a partir da qual os valores de $\gamma(h)$ param de crescer e são iguais ao patamar. É a distância dentro da qual as amostras apresentam autocorrelação espacial, isto é, o alcance mede a distância a partir da qual os valores de $Z(\boldsymbol{s})$ deixam de estar correlacionados. Nos modelos teóricos abordados nesta tese $a = c\varphi_3$, em que c é uma constante que depende do modelo. Observa-se que nos modelos teóricos apresentados a seguir, com excessão do modelo esférico, o alcance real é infinito, contudo, para fins de operações matemáticas, convenciona-se que nestes casos que o alcance é obtido a 95% do patamar, isto é, φ_3 é tal que $\gamma(\varphi_3) = 0,95\,(\varphi_1 + \varphi_2)$, como observado por Seidel e Oliveira (2013) e referência neles citada.

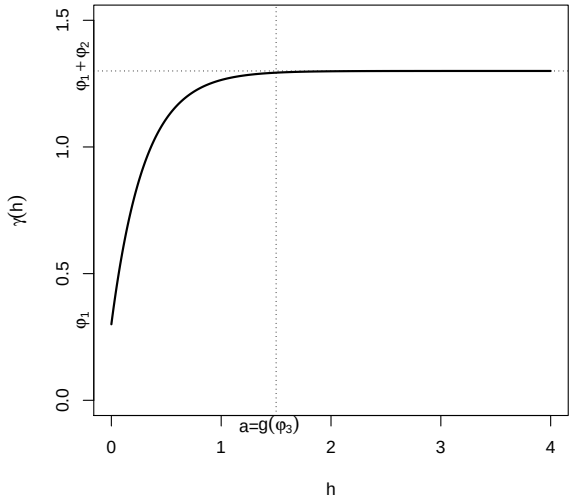


Figura 10 Ilustração dos parâmetros espaciais do modelo.

3.2.1.2 Modelos teóricos de semivariogramas

Os modelos espaciais utilizados para explicar a estrutura de dependencia espacial são divididos em modelos que possuem patamar (modelos transitivos) e modelos que não possuem patamar (modelos não transitivos) (MARDIA; MARSHALL, 1984; CRESSIE, 2015). Porém, apesar de existirem fenômenos em que o crescimento de $\gamma(h)$ não tende para um patamar, e os semivariogramas precisam ser ajustados por modelos sem patamar, os modelos com patamar cobrem as situações de dispersão de fenômenos espaciais nas Ciências da terra e do ambiente (SOARES, 2000). Apresentam-se, na sequência, os principais modelos de semivariogramas

transitivos (esférico, exponencial, gaussiano, exponencial potência e família Matérn), descritos como funções dos parâmetros φ_1, φ_2 e φ_3.

- Modelo esférico: neste modelo, a função de semivariância é dada por

$$\gamma(h) = \begin{cases} 0 \quad \text{se} \quad h = 0 \\ \varphi_1 + \varphi_2 \left(\frac{3}{2} \left(\frac{h}{\varphi_3} \right) - \frac{1}{2} \left(\frac{h}{\varphi_3} \right)^3 \right) \quad \text{se} \quad 0 < h \leq \varphi_3 \\ \varphi_1 + \varphi_2 \quad \text{se} \quad h > \varphi_3 \end{cases}$$

  Este modelo apresenta crescimento rápido junto à origem e o alcance é $a = \varphi_3$. A Figura 11 ilustra um semivariograma que apresenta modelo esférico.

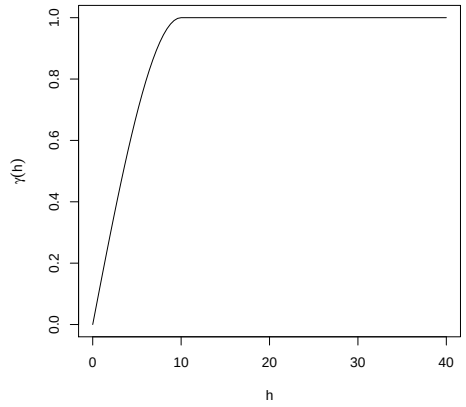


Figura 11 Semivariograma teórico: modelo esférico.

- Modelo exponencial: a função de semivariância, neste caso, é dada por

$$\gamma(h) = \begin{cases} 0 \quad \text{se} \quad h = 0 \\ \varphi_1 + \varphi_2 \left(1 - \exp\left(-\frac{h}{\varphi_3} \right) \right) \quad \text{se} \quad h > 0. \end{cases}$$

  Neste modelo, o alcance é atingido de forma assintótica e é dado por $a = 3\varphi_3$. O modelo exponencial também apresenta crescimento rápido junto à origem. A Figura 12 mostra um semivariograma que segue um modelo exponencial.

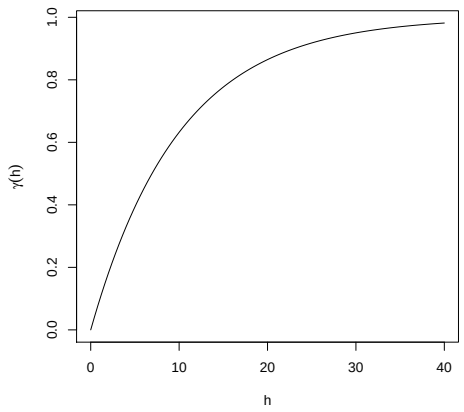


Figura 12 Semivariograma teórico: modelo exponencial.

- Modelo Gaussiano: os semivariogramas dos modelos esférico e exponencial apresentam crescimento relativamente rápido junto à origem. De acordo com Soares (2000), os fenômenos regulares e contínuos apresentam, tipicamente, um semivariograma com crescimento lento e comportamento parabólico próximo à origem, características do modelo Gaussiano. No modelo Gaussiano, a função semivariância é dada por

$$\gamma(h) = \begin{cases} 0 & \text{se} \quad h = 0 \\ \varphi_1 + \varphi_2 \left(1 - \exp\left(-(\frac{h}{\varphi_3})^2\right)\right) & \text{se} \quad h > 0 \end{cases}$$

  A Figura 13 mostra o comportamento de um semivariograma que segue o modelo Gaussiano.

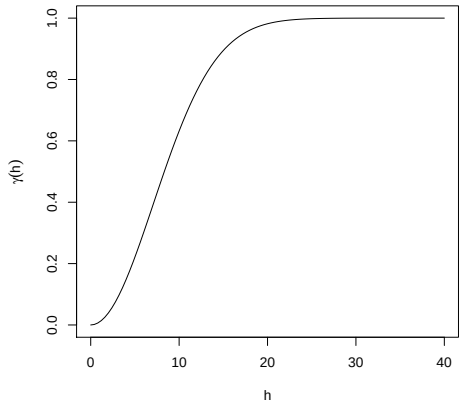


Figura 13 Semivariograma teórico: modelo gaussiano.

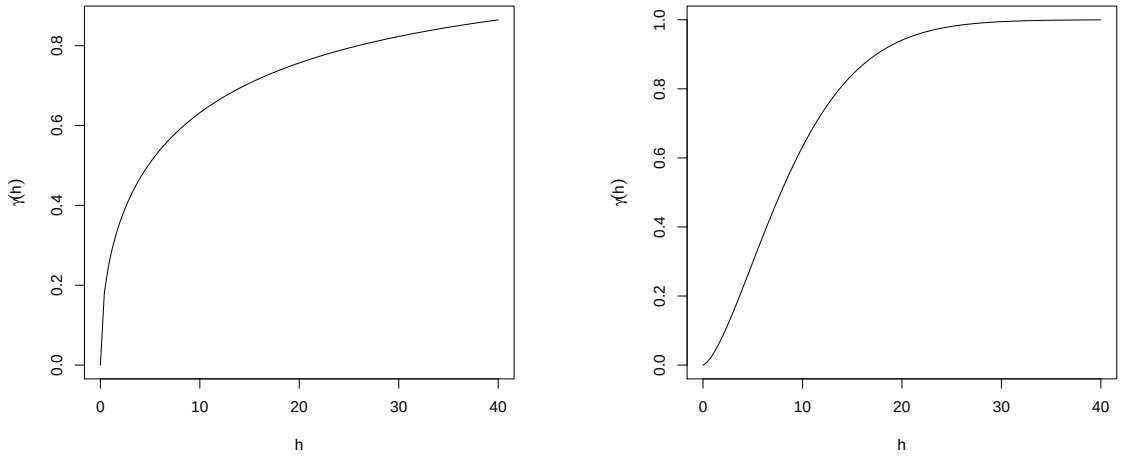
  Assim como o modelo exponencial, o modelo gaussiano atinge o patamar assintoticamente. Seu alcance é dado por $a = \sqrt{3}\varphi_3$.

- Modelo exponencial Potência: A família de modelos cujas funções semivariâncias são

dadas por

$$\gamma(h) = \begin{cases} 0 \quad \text{se} \quad h = 0 \\ \varphi_1 + \varphi_2 \left(1 - \exp\left(-(\frac{h}{\varphi_3})^p\right)\right) \quad \text{se} \quad h > 0 \end{cases}$$

em que $p \in \mathbb{R}$, para $0 < p \leq 2$. Pode-se observar que, para $p = 2$, o modelo exponencial potência é o modelo gaussiano e, para $p = 1$ o modelo exponencial. A relação entre o alcance prático e o parâmetro φ_3 depende do valor de p. A Figura 14 mostra o semivariograma para os casos $p = 0,5$ (a) e $p = 1,5$ (b).



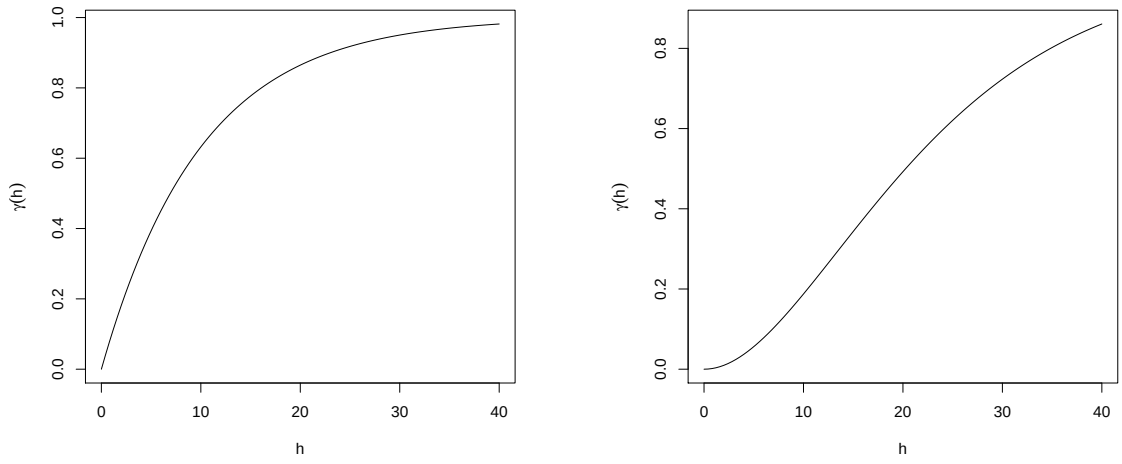
(a) (b)

Figura 14 Semivariogramas teóricos: modelo exponencial potência $p = 0,5$ (a) e $p = 1,5$ (b).

- Modelo família Matérn: Matérn (1986) apresentou uma função para modelar a dependência espacial que ficou conhecida como família Matérn, cuja expressão é

$$\gamma(h) = \begin{cases} 0 \quad \text{se} \quad h = 0 \\ \varphi_1 + \varphi_2 \left(1 - (2^{\delta-1}\Gamma(\delta))^{-1}(\frac{h}{\varphi})^{\delta} K_{\delta}(\frac{h}{\varphi_3})\right) \quad \text{se} \quad h > 0 \end{cases}$$

em que δ é uma constante positiva e $K_{\delta}(\cdot)$ é a função de Bessel modificada do terceiro tipo de ordem δ. No modelo família Matérn, o parâmetro δ é um parâmetro de forma, assim como p no modelo exponencial potência. Quando $\delta = 0,5$, tal modelo corresponde ao modelo exponencial, e quando δ fica muito grande ($\delta \to \infty$) o modelo se aproxima do modelo gaussiano. Assim, como para a família exponencial potência, a relação entre o alcance e o parâmetro φ_3 depende do valor de δ e quanto maior o valor de δ, maior é o valor do alcance. A Figura 15 mostra o semivariograma para os casos $\delta = 0,5$ e $\delta = 2$.



(a) (b)

Figura 15 Semivariograma teórico: modelo família Matérn $\delta = 0,5$ (a) e $\delta = 2$ (b).

3.2.2 Krigagem

Geralmente, no estudo da variabilidade espacial de determinada variável um dos interesses é predizer o valor desta variável em um ou mais pontos da região sob estudo. Existem vários métodos de interpolação utilizados para predição espacial e, segundo Gotway e Hartford (1996), a eficiência desses na predição de valores depende da natureza da variável, expressa por sua variância, assimetria, curtose, isotropia e do padrão de sua variabilidade espacial. Dentre tais métodos, existem os métodos clássicos que, apesar de considerarem a dependência espacial da variável, não consideram o padrão de variabilidade espacial da mesma, como é o caso do inverso do quadrado da distância (LANDIM, 2000), que calcula o valor da característica em cada posição s, a partir de uma média ponderada dos valores da variável nos pontos amostrados $\boldsymbol{s_i}$ e atribui pesos que dependem unicamente da distância entre s e $\boldsymbol{s_i}$. A krigagem é um método de interpolação bastante utilizado em estudos geoestatísticos. Tal método também utiliza médias ponderadas, mas a diferença entre a krigagem e os métodos de interpolação clássicos é a maneira como os pesos são atribuídos às diferentes amostras. Na interpolação por inverso do quadrado da distância, por exemplo, os pesos são $1/h^2$, em que h é a distância euclidiana entre o ponto observado e o ponto sob interpolação. Na krigagem, os pesos são determinados a partir de uma análise espacial, baseada no semivariograma experimental, isto é, o valor estimado para um ponto $\boldsymbol{s_0}$ é dado por

$$\hat{z}(\boldsymbol{s_0}) = \sum_{i=1}^{n} \lambda_i z(\boldsymbol{s_i}), \tag{3.4}$$

em que n é o número de amostras, $z(\boldsymbol{s_i})$ é o valor da variável no ponto amostrado $\boldsymbol{s_i}$ e λ_i é o peso calculado a partir do semivariograma experimental.

Os valores $\lambda_i,\ i = 1, \ldots, n$ são escolhidos de forma que o preditor seja o melhor preditor linear não viesado (*best linear unbiased predictor* - BLUP). Tais qualidades são garantidas

impondo que:

- a esperança matemática do erro é zero, isto é,

$$\mathrm{E}(Z(\boldsymbol{s}) - \hat{Z}(\boldsymbol{s})) = 0; \tag{3.5}$$

- a variância do erro é mínima, isto é, precisa-se minimizar

$$\mathrm{Var}(Z(\boldsymbol{s}) - \hat{Z}(\boldsymbol{s})) = \mathrm{E}((Z(\boldsymbol{s}) - \hat{Z}(\boldsymbol{s}))^2) - \mathrm{E}^2(Z(\boldsymbol{s}) - \hat{Z}(\boldsymbol{s})). \tag{3.6}$$

3.2.2.1 Krigagem ordinária

Ao se admitir que o processo estocástico é intrinsecamente estacionário, das equações (3.4) e (3.5), pode-se concluir que $\sum_{i=1}^{n} \lambda_i = 1$.

Além disso, o segundo termo do lado direito da equação (3.6) corresponde ao viés, como se quer um estimador não viesado, o viés deve ser zero e a variância do erro dada por

$$\mathrm{Var}(Z(\boldsymbol{s}) - \hat{Z}(\boldsymbol{s})) = \mathrm{E}((Z(\boldsymbol{s}) - \hat{Z}(\boldsymbol{s}))^2)$$

deve ser mínima. Após algumas manipulações algébricas (ISAAKS; SRIVASTAVA, 1989), chega-se à expressão que deve ser minimizada

$$\mathrm{Var}(Z(\boldsymbol{s}) - \hat{Z}(\boldsymbol{s})) = \mathrm{Var}(Z(\boldsymbol{s})) + \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j \mathrm{C}(Z(\boldsymbol{s_i}), Z(\boldsymbol{s_j})) - 2\sum_{i=1}^{n} \lambda_i \mathrm{C}(Z(\boldsymbol{s_i}), Z(\boldsymbol{s})). \tag{3.7}$$

Isaaks e Srivastava (1989) colocam que para minimizar (3.7), é necessário utilizar a técnica dos multiplicadores de Lagrange, que transformam um problema de minimização com restrição em um problema de minimização sem restrição. Após acrescentar o parâmetro de Lagrange, η, de maneira conveniente na equação (3.7), deriva-a com relação às $n+1$ variáveis $(\lambda_1, \ldots, \lambda_n, \eta)$, e igualam-se estas derivadas a zero. Assim, os λ_i's são as soluções do sistema de equações:

$$S: \begin{cases} \displaystyle\sum_{j=1}^{n} \lambda_j \mathrm{C}((Z(\boldsymbol{s_j}), Z(\boldsymbol{s_i})) + \eta = \mathrm{C}(Z(\boldsymbol{s_i}), Z(\boldsymbol{s})), 1 \leq i \leq n. \\ \displaystyle\sum_{i=1}^{n} \lambda_i = 1 \end{cases} \tag{3.8}$$

Utilizando-se a relação $\gamma(\boldsymbol{s_i}, \boldsymbol{s_j}) = \mathrm{Var}(Z(\boldsymbol{s})) - \mathrm{C}(Z(\boldsymbol{s_i}), Z(\boldsymbol{s_j}))$, em termos das semivariâncias, o sistema de equações dado em (3.8), conhecido como sistema de krigagem ordinária, pode ser escrito da seguinte forma

$$S: \begin{cases} \displaystyle\sum_{j=1}^{n} \gamma(\boldsymbol{s_1}, \boldsymbol{s_j}) - \eta = \gamma(\boldsymbol{s_1}, s), 1 \leq i \leq n. \\ \displaystyle\sum_{i=1}^{n} \lambda_i = 1 \end{cases}$$

ou na forma matricial, $\boldsymbol{\gamma\lambda} = \boldsymbol{A}$, o que implica $\boldsymbol{\lambda} = \boldsymbol{\gamma}^{-1}\boldsymbol{A}$, em que

$$\boldsymbol{\gamma} = \begin{pmatrix} \gamma(\boldsymbol{s_1}, \boldsymbol{s_1}) & \gamma(\boldsymbol{s_1}, \boldsymbol{s_2}) & \dots & \gamma(\boldsymbol{s_1}, \boldsymbol{s_n}) & 1 \\ \gamma(\boldsymbol{s_2}, \boldsymbol{s_1}) & \gamma(\boldsymbol{s_2}, \boldsymbol{s_2}) & \dots & \gamma(\boldsymbol{s_2}, \boldsymbol{s_n}) & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \gamma(\boldsymbol{s_n}, \boldsymbol{s_1}) & \gamma(\boldsymbol{s_n}, \boldsymbol{s_2}) & \dots & \gamma(\boldsymbol{s_n}, \boldsymbol{s_n}) & 1 \\ 1 & 1 & \dots & 1 & 0 \end{pmatrix},$$

$\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_n, \eta)^T$ e $\boldsymbol{A} = (\gamma(\boldsymbol{s_1}, s), \dots, \gamma(\boldsymbol{s_n}, \boldsymbol{s}), 1)^T$.

3.2.2.2 Krigagem com drift externo

Quando trabalha-se sob a hipótese de processo estocástico intrinsecamente estacionário, utiliza-se o conceito de krigagem ordinária, onde assume-se que $\mathrm{E}(Z(\boldsymbol{s})) = \mu$, com μ constante para todo s. No entanto, se uma informação auxiliar, correlacionada com a variável resposta, está disponível em todos os pontos da grade, regressão krigagem ou krigagem com um modelo de tendência, também conhecida como krigagem com drift externo, deve ser usada. Os três termos descrevem praticamente o mesmo método, fornecem as mesmas predições e diferem apenas em passos metodológicos (HENGL; HEUVELINK; STEIN, 2003).

Assim como em krigagem ordinária, o objetivo é encontrar $\lambda_i, \quad i = 1, \dots, n$, tal que $\widehat{Y}(\boldsymbol{s_0}) = \sum_{i=1}^{n} \lambda_i Y(\boldsymbol{s_i})$, com erro de predição não enviesado e variância mínima, em que $\boldsymbol{s_0} \in D \subset \mathbb{R}^n$, é um ponto não amostrado.

Para assegurar a condição de não enviesamento, é necessário e suficiente supor que

$$X_{0j} = \sum_{i=1}^{n} \lambda_i X_{ij}, \quad j = 0, \dots, p, \tag{3.9}$$

uma vez que

$$\begin{aligned} \mathrm{E}\left(\widehat{Y}(\boldsymbol{s_0}) - Y(\boldsymbol{s_0})\right) &= \left(\sum_{i=1}^{n} \lambda_i \mathrm{E}\left(Y(\boldsymbol{s_i})\right)\right) - \mathrm{E}\left(Y(\boldsymbol{s_0})\right) = \left(\sum_{i=1}^{n} \lambda_i \sum_{j=0}^{p} \beta_j X_{ij}\right) - \sum_{j=0}^{p} \beta_j X_{0j} \\ &= \sum_{j=0}^{p} \beta_j \left(\sum_{i=1}^{n} \lambda_i X_{ij} - X_{0j}\right). \end{aligned}$$

Para o estabelecimento de condições que minimizem a variância dos erros, primeiramente, é conveniente lembrar-se que $Y(s) = \mu(s) + \epsilon(s)$, em que $\mathrm{E}(\epsilon(\boldsymbol{s})) = 0$ e que a covariância entre $Y(\boldsymbol{s_i})$ e $Y(\boldsymbol{s_j})$ depende apenas da distância entre $\boldsymbol{s_i}$ e $\boldsymbol{s_j}$, isto é, $\mathrm{Cov}(\boldsymbol{s_i}, \boldsymbol{s_j}) = C(h_{ij})$, em que h_{ij} é a distância euclidiana entre $\boldsymbol{s_i}$ e $\boldsymbol{s_j}$. Desta forma, a variância da estimação pode ser escrita em função dos resíduos $\epsilon(\boldsymbol{s})$, pois,

$$\begin{aligned} \widehat{Y}(\boldsymbol{s_0}) - Y(\boldsymbol{s_0}) &= \left(\sum_{i=1}^{n} \lambda_i Y(\boldsymbol{s_i})\right) - Y(\boldsymbol{s_0}) = \left(\sum_{i=1}^{n} \lambda_i \left(\mu(\boldsymbol{s_i}) + \epsilon(\boldsymbol{s_i})\right)\right) - \left(\mu(\boldsymbol{s_0}) + \epsilon(\boldsymbol{s_0})\right) \\ &= \left(\sum_{i=1}^{n} \lambda_i \left(\mu(\boldsymbol{s_i})\right) - \mu(\boldsymbol{s_0})\right) + \left(\sum_{i=1}^{n} \lambda_i \left(\epsilon(\boldsymbol{s_i})\right) - \epsilon(\boldsymbol{s_0})\right). \end{aligned}$$

De onde segue que $\widehat{Y}(\boldsymbol{s_0}) - Y(\boldsymbol{s_0}) = \left(\sum_{i=1}^{n} \lambda_i \epsilon(\boldsymbol{s_i})\right) - \epsilon(\boldsymbol{s_0}) = \hat{\epsilon}(\boldsymbol{s_0}) - \epsilon(\boldsymbol{s_0})$. Logo, assim como em krigagem ordinária, tem-se que

$$\mathrm{Var}\left(\widehat{Y}(\boldsymbol{s_0}) - Y(\boldsymbol{s_0})\right) = \mathrm{E}\left(\left(\hat{\epsilon}(\boldsymbol{s_0}) - \epsilon(\boldsymbol{s_0})\right)^2\right) = \mathrm{C}(0) + \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j \mathrm{C}(\boldsymbol{s_i}, \boldsymbol{s_j}) - 2\sum_{i=1}^{n} \mathrm{C}(\boldsymbol{s_0}, \boldsymbol{s_i}). \tag{3.10}$$

O problema de minimização é idêntico ao da krigagem ordinária: derivar a equação (3.10) em relação a $\lambda_1, \ldots, \lambda_n$ e igualar as equações obtidas a zero. No entanto, pretende-se ainda que sejam cumpridas as condições estabelecidas pelas $p+1$ equações definidas em (3.9). Assim, acrescentam-se $p+1$ termos nulos na equação (3.10), os quais correspondem a estas condições, de forma que

$$\mathrm{Var}\left(\widehat{Y}(\boldsymbol{s_0}) - Y(\boldsymbol{s_0})\right) = \mathrm{C}(0) + \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j \mathrm{C}(\boldsymbol{s_i}, \boldsymbol{s_j}) - 2\sum_{i=1}^{n} \mathrm{C}(\boldsymbol{s_0}, \boldsymbol{s_i}) + 2\sum_{j=0}^{p} \eta_j \sum_{i=1}^{n} \lambda_i X_{ij} - X_{0j},$$

em que η_j são parâmetros de Lagrange.

Igualando a zero as $n+p+1$ derivadas em relação a $\lambda_i, i = 1, \ldots, n$ e $\eta_j, j = 0, \ldots, p$, obtém-se o seguinte sistema de equações lineares:

$$S = \begin{cases} \displaystyle\sum_{k=1}^{n} \lambda_k \mathrm{C}\left(\boldsymbol{s_i}, \boldsymbol{s_k}\right) + \sum_{k=0}^{p} \eta_k X_{ik} = \mathrm{C}\left(\boldsymbol{s_i}, \boldsymbol{s_0}\right), \quad i = 1, \ldots, n. \\ \\ \displaystyle\sum_{k=1}^{n} \lambda_k X_{kj} = X_{0j}, \quad j = 0, \ldots, p. \end{cases}$$

que, em forma matricial, pode ser escrito como $S : \boldsymbol{C\lambda} = \boldsymbol{A}$, em que $\boldsymbol{C}$ é a matriz, $(n+p+1) \times (n+p+1)$, dada por,

$$\boldsymbol{C} = \begin{pmatrix} \mathrm{C}\left(\boldsymbol{s_1}, \boldsymbol{s_1}\right) & \cdots & \mathrm{C}\left(\boldsymbol{s_1}, \boldsymbol{s_n}\right) & 1 & X_{11} & \cdots & X_{1p} \\ \mathrm{C}\left(\boldsymbol{s_2}, \boldsymbol{s_1}\right) & \cdots & \mathrm{C}\left(\boldsymbol{s_2}, \boldsymbol{s_n}\right) & 1 & X_{21} & \cdots & X_{2p} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathrm{C}\left(\boldsymbol{s_n}, \boldsymbol{s_1}\right) & \cdots & \mathrm{C}\left(\boldsymbol{s_n}, \boldsymbol{s_n}\right) & 1 & X_{n1} & \cdots & X_{np} \\ 1 & \cdots & 1 & 0 & 0 & \cdots & 0 \\ X_{11} & \cdots & X_{n1} & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ X_{p1} & \cdots & X_{np} & 0 & 0 & \cdots & 0 \end{pmatrix};$$

$\boldsymbol{\lambda}$ é o vetor, $(n+p+1)\times 1$, dado por $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n, \eta_0, \ldots, \eta_p)^\top$ e $\boldsymbol{A}$ é o vetor, $(n+p+1)\times 1$, dado por $\boldsymbol{A} = \left(\mathrm{C}\left(\boldsymbol{s_1}, \boldsymbol{s_0}\right), \ldots, \mathrm{C}\left(\boldsymbol{s_1}, \boldsymbol{s_0}\right), 1, X_{01}, \ldots, X_{0p}\right)$. Desta forma, os pesos $\lambda_i, i = 1, \ldots, n$ são dados por $\boldsymbol{\lambda} = \boldsymbol{C}^{-1}\boldsymbol{A}$.

A krigagem universal é um caso particular de krigagem com drift externo em que a média é uma função das coordenadas da grade amostral (HENGL; HEUVELINK; STEIN, 2003).

Um exemplo de krigagem universal é apresentado em De Bastiani et al. (2015).

3.2.3 Métodos de validação de modelos

Uma etapa importante do estudo de modelos estatísticos, seja para escolher o modelo mais adequado entre vários modelos possíveis, seja para certificar-se que um dado modelo ajusta bem os dados amostrais. Para a tomada de decisão sobre a adequação de um modelo, pode-se escolher utilizar alguns critérios dentre os quais destacamos: valor máximo da função de verossimilhança, critério de Akaike, critério Bayesiano, fator Bayes (BF), validação cruzada (CV) e *Q-Q plots*.

3.2.3.1 Valor máximo da função verossimilhança

Por esse critério o modelo que apresentar o maior valor para a função verossimilhança, equivalente ao maior valor para o logaritmo da função verossimilhança, é considerado o mais adequado. Os critérios de informação de Akaike (AIC) e Bayesiano de Schwarz (BIC) são baseados em uma penalização do logaritmo da função verossimilhança para modelos mais complexos, isto é, com um número maior de parâmetros. Partem do princípio que se dois modelos representam dados igualmente satisfatórios. Então, com o modelo mais simples, pode-se esperar melhor desempenho para a predição de novos dados. Em ambos os casos o melhor modelo é aquele que apresentar menor índice.

- Critério de informação de Akaike: o critério de Akaike (*Akaike information criterion*- AIC) (AKAIKE, 2002) é dado por dado por

$$\mathrm{AIC} = -2\ell(\widehat{\theta}) + 2d,$$

  em que $\ell(\widehat{\theta})$ é logaritmo da função verossimilhança para o parâmetro θ associado com o modelo, avaliado em $\theta = \widehat{\theta}$ e d é o número de parâmetros do modelo ajustado.

- Critério de informação Bayesiano de Schwarz é dado por

$$\mathrm{BIC} = -2\ell(\widehat{\theta}) + d\log(n),$$

  em que $\ell(\widehat{\theta})$ e d são dados como em AIC e n é o tamanho do conjunto de dados.

- Fator Bayes: o fator Bayes, denotado por B_{12}, permite comparar um modelo considerado correto (M_1,) com outro modelo (M_2), por meio de $2\log(B_{12}) \approx \mathrm{BIC}_{\mathrm{M}_2} - \mathrm{BIC}_{\mathrm{M}_1}$, em que $\mathrm{BIC}_{\mathrm{M}_i}$ denota o BIC associado ao modelo M_i, para $i = 1, 2$. O BF fornece um valor objetivo para quantificar o grau de superioridade de um modelo com respeito a outro. Uma interpretação do BF é fornecida na Tabela 4 (LEIVA et al., 2015e).

3.2.3.2 Validação cruzada

A validação cruzada consiste em retirar do conjunto de dados, temporariamente, o valor da amostra em certa localização. Um novo valor nesta localização é predito por krigagem,

Tabela 1 Interpretação de $2\log(B_{12})$ associado com o BF.

$2\log(B_{12})$	Evidência em favor de M_1
< 0	Negativa (M_2 é aceito)
$[0, 2)$	fraca
$[2, 6)$	Positiva
$[6, 10)$	forte
≥ 10	muito forte

usando as amostras restantes. Compara-se então o valor estimado com o valor amostrado, isto é, calcula-se o erro $\zeta_i = Y(\boldsymbol{s_i}) - \widehat{Y}(\boldsymbol{s_i})$, em que $Y(\boldsymbol{s_i})$, é o valor observado no ponto $\boldsymbol{s_i}$ e $\widehat{Y}(\boldsymbol{s_i})$, valor predito por krigagem ordinária no ponto $\boldsymbol{s_i}$, sem considerar a observação $Y(\boldsymbol{s_i})$. Tal procedimento é repetido para todas as amostras disponíveis e é conhecido como método de "deixar-um-fora"(ISAAKS; SRIVASTAVA, 1989). As medidas erro absoluto $EA = \sum_{i=1}^{n} |\zeta_i|$; erro médio $EM = \frac{1}{n}\sum_{i=1}^{n} \zeta_i$; erro médio reduzido $ER = \frac{1}{n}\sum_{i=1}^{n} \frac{\zeta_i}{\sigma(\widehat{Y}(\boldsymbol{s_i}))}$, em que $\sigma(\widehat{Y}(\boldsymbol{s_i}))$ é o desvio padrão da krigagem no ponto $\boldsymbol{s_i}$; desvio padrão dos erros médios $DP_{EM} = \sqrt{\mathrm{Var}(\boldsymbol{\zeta})}$, em que $\mathrm{Var}(\boldsymbol{\zeta})$ é a variância de $\boldsymbol{\zeta} = (\zeta_1, \ldots, \zeta_n)^\top$ e desvio padrão dos erros médios reduzidos dada por $DP_{ER} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} \frac{|\zeta_i|}{\sigma(\widehat{Y}(\boldsymbol{s_i}))}}$, podem ser usadas como instrumento para avaliar modelos. Espera-se que os erros de predição tenham média nula e variância mínima. Portanto, os valores de EM e ER mais próximos de zero, o menor valor de DP_{EM} e de EA e o valor de DP_{ER} mais próximo de um são critérios para escolha do melhor modelo ajustado (FARACO et al., 2008).

Segundo De Bastiani et al. (2015), é possível utilizar como medida de validação cruzada o coeficiente

$$VC = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{z(\boldsymbol{s_i}) - \hat{z}_{(i)}(\boldsymbol{s_i})}{1 - h_{ii}}\right)^2, i = 1, \cdots, n,$$

em que h_{ii} é o i-ésimo elemento da diagonal da matriz projeção $\boldsymbol{H}$, $\hat{z}_{(i)}(\boldsymbol{s_i}) = \boldsymbol{X}_{(i)}^\top \hat{\boldsymbol{\beta}}_{(i)}$, com $\boldsymbol{X}_{(i)}^\top$ a i-ésima coluna da matriz de covariáveis, $\boldsymbol{X}$, sem considerar a i-ésima observação e $\hat{\boldsymbol{\beta}}_{(i)}$ é o estimador ML de $\boldsymbol{\beta}$ sem considerar a i-ésima observação. A matriz projeção de um modelo $\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\epsilon}$ permite escrever os valores preditos como combinação linear dos valores observados, isto é, é a matriz $\boldsymbol{H}$ tal que $\widehat{\boldsymbol{y}} = \boldsymbol{Hy}$. Geometricamente, se as colunas de $\boldsymbol{X}$ representam pontos do espaço euclidiano $\mathbb{R}^n$, então os pontos de $\boldsymbol{X\beta}$ (combinações lineares dos vetores colunas) constituem um subespaço do $\mathbb{R}^n$ de dimensão p, e o vetor predito $\widehat{\boldsymbol{y}}$ é a projeção ortogonal de $\boldsymbol{y}$ neste subespaço. Por isso, tem como nome matriz projeção (HOAGLIN; WELSECH, 1978).

3.2.3.3 *Q-Q plots*

A verificação de que determinada distribuição se ajusta de maneira eficiente a um conjunto de dados e, portanto, pode ser usada para estudá-lo, é uma etapa importante do trabalho estatístico. Para verificação da normalidade, um método tradicionalmente utilizado é traçar um gráfico quantil-quantil (*Q-Q plot*) que compara os quantis empíricos com os quantis

teóricos da distribuição normal. Se os pontos estão dispostos ao longo da reta $y = x$, têm-se indícios de normalidade. Em particular, valores que se afastam consideravelmente da reta indicam casos discrepantes (LANGE; LITTLE; TAYLOR, 1989). Quando os dados não são supostamente normais, uma alternativa é conseguir dados normais a partir de transformações. Dentre as possíveis aproximações, tem-se a aproximação de Wilson-Hilferty (PAULSON, 1942; JOHNSON; KOTZ; BALAKRISHMAN, 1994) dada por

$$z_i = \frac{\left(\frac{\delta_i}{n}\right)^{1/3} - (1 - \frac{2}{9n})}{\left(\frac{2}{9n}\right)^{1/2}},$$

$i = 1, \ldots, n$, sendo δ_i uma variável aleatória com distribuição χ^2 com n graus de liberdade. E aproximação obtida pelos Resíduos Quantiles Padronizados definidos em Dunn & Smyth (1996) como

$$r_i = \Phi^{-1}\left(F(\delta_i)\right)$$

em que $\Phi(.)$ e $F(.)$ indicam, respectivamente, as f.d.a. das distribuições normal padrão e χ^2 com n graus de liberdade.

3.2.4 Medidas de similaridade entre mapas

A comparação visual de dois mapas é muito subjetiva e comparar mapas apenas visualmente pode fazer com que diferenças importantes entre eles passem desapercebidas. Uma comparação mais objetiva entre mapas pode ser realizada utilizando medidas de similaridade baseadas na matriz de erros. Na matriz de erros, os pixels do mapa de referência são quantificados nas colunas e os pixels do mapa modelo nas linhas, de modo que um elemento f_{ij} da matriz de erros fornece a quantidade de pixels que pertence à classe C_i do mapa modelo e à classe C_j do mapa de referência, para $i, j = 1, \ldots, m$, em que m é o número de classes de ambos os mapas. Cabe ressaltar que, para realizar a comparação entre dois mapas, é necessário que esses estejam particionados nas mesmas classes e que, neste trabalho, um pixel corresponde a um elemento do vetor de valores krigados. A Tabela 2 mostra uma matriz de erros genérica. Se f é o número total de pixels, o índice de acurácia global (AG) definido como $\mathrm{AG} = (\sum_{i=1}^{m} f_{ii})/f$ é uma medida baseada nos elementos da diagonal principal da matriz de erros, que representa os pixels que têm a mesma classificação em ambos os mapas. De acordo com Anderson et al. (1976), para ter mapas com uma similaridade aceitável, AG deve ser maior que 0,85. Outra medida de acurácia é o índice kappa definido como $\kappa = (f \sum_{i=1}^{m} f_{ii} - \sum_{i=1}^{m} f_{\cdot i} f_{i\cdot})/(f^2 - \sum_{i=1}^{m} f_{\cdot i} f_{i\cdot})$, em que $f_{\cdot i}$ é o número de pixels na classe C_i do mapa de referência e $f_{i\cdot}$ o número de pixels na classe C_i do mapa modelo. O índice utiliza todos os elementos da matriz de erros. De acordo com Krippendorff (2004), os mapas têm baixa similaridade se $\kappa < 0,67$, média similaridade se $0,67 \leq \kappa < 0,80$ e alta similaridade se $\kappa \geq 0,80$.

Tabela 2 Matriz de erros genéricos $m \times m$.

Pixels do mapa de referência						
Classes		C_1	C_2	$\cdots$	C_m	Total
Pixels do mapa modelo	C_1	f_{11}	f_{12}	$\cdots$	f_{1m}	$f_{1\cdot}$
	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
	C_m	f_{m1}	f_{m2}	$\cdots$	f_{mm}	$f_{m\cdot}$
Total		$f_{\cdot 1}$	$f_{\cdot 2}$	$\cdots$	$f_{\cdot m}$	f

3.2.5 Análise de diagnóstico

Outra etapa importante é a verificação da existência de observações que possam provocar alterações nos parâmetros estimados. Esta etapa é conhecida como análise de diagnóstico.

3.2.5.1 Influência global

Consiste em deletar uma observação particular do conjuntos de dados e observar o impacto dessa retirada na estimação dos parâmetros. Uma medida, proposta inicialmente para modelos lineares com resposta normal por Cook (1977), para mensurar a mudança dos parâmetros estimados quando uma observação é deletada, é dada por $\mathrm{LD}(\hat{\boldsymbol{\theta}}_{(i)}) = 2(\ell(\hat{\boldsymbol{\theta}}) - \ell(\hat{\boldsymbol{\theta}}_{(i)}))$, em que $\ell(\cdot)$ é o logaritmo da função verossimilhança e $\hat{\boldsymbol{\theta}}$ e $\hat{\boldsymbol{\theta}}_{(i)}$ são, respectivamente, as estimativas por ML de $\boldsymbol{\theta}$, considerando-se o conjunto de dados completo e o conjunto de dados sem a i-ésima observação. Tal medida ficou conhecida como distância de Cook e, desde que foi proposta tem sido bastante usada na identificação de pontos influentes de diversos tipos de modelos. Além disso, surgiram outras propostas para estudar a influência a partir da deleção de casos, baseadas na distância de Cook, como pode ser observado a seguir,

Utilizando expansão de Taylor até segunda ordem, em torno de $\hat{\boldsymbol{\theta}}$, e considerando que $\dot{\ell}_{(i)}(\hat{\boldsymbol{\theta}}) = 0$, em que $\ell_{(i)}(\cdot)$ é o logaritmo da função verossimilhança obtida após a exclusão da $i-$ésima observação, tem-se (PAN; FEI; FOSTER, 2014)

$$\mathrm{LD}(\hat{\boldsymbol{\theta}}_{(i)}) = \left(\hat{\boldsymbol{\theta}}_{(i)} - \hat{\boldsymbol{\theta}}\right)^{\top} \left(-\ddot{\ell}(\hat{\boldsymbol{\theta}})\right) \left(\hat{\boldsymbol{\theta}}_{(i)} - \hat{\boldsymbol{\theta}}\right), \tag{3.11}$$

em que, $\dot{\ell}_{(i)}(\boldsymbol{\theta}) = \frac{\partial \ell_{(i)}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ e $\ddot{\ell}_{(i)}(\boldsymbol{\theta}) = \frac{\partial \ell^2_{(i)}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^{\top}}$.

Quando se utiliza o estimador a um passo de Newton-Raphson (LEHMANN, 1999; PAN; FEI; FOSTER, 2014), obtém-se uma aproximação para $\hat{\boldsymbol{\theta}}_{(i)}$

$$\hat{\boldsymbol{\theta}}_{(i)} = \hat{\boldsymbol{\theta}} + \left(-\ddot{\ell}_{(i)}(\hat{\boldsymbol{\theta}})\right)^{-1} \dot{\ell}_{(i)}(\hat{\boldsymbol{\theta}}). \tag{3.12}$$

Ao se Substituir (3.12) na equação (3.11), tem-se

$$\mathrm{LD}^1(\hat{\boldsymbol{\theta}}_{(i)}) = (\dot{\ell}_{(i)}(\hat{\boldsymbol{\theta}}))^{\top}(-\ddot{\ell}_{(i)}(\hat{\boldsymbol{\theta}}))^{-1}(\dot{\ell}_{(i)}(\hat{\boldsymbol{\theta}})).$$

Pan, Fei e Foster (2014) mostraram que é possível substituir $\ddot{\ell}_{(i)}(\hat{\boldsymbol{\theta}})$ por $\ddot{\ell}(\hat{\boldsymbol{\theta}})$ ou ainda

por $I(\boldsymbol{\theta}) = \mathrm{E}(-\ddot{\ell}(\boldsymbol{\theta}))$, obtendo-se, respectivamente, as medidas de influência,

$$\begin{aligned} \mathrm{D}_i^1(\boldsymbol{\theta}) &= (\dot{\ell}_{(i)}(\hat{\boldsymbol{\theta}}))^\top (-\ddot{\ell}(\hat{\boldsymbol{\theta}}))^{-1} (\dot{\ell}_{(i)}(\hat{\boldsymbol{\theta}})) \\ \mathrm{D}_i(\boldsymbol{\theta}) &= (\dot{\ell}_{(i)}(\hat{\boldsymbol{\theta}}))^\top \boldsymbol{I}(\hat{\boldsymbol{\theta}})^{-1} (\dot{\ell}_{(i)}(\hat{\boldsymbol{\theta}})). \end{aligned}$$

A medida $\mathrm{D}_i^1(\boldsymbol{\theta})$ é preferível à medida $\mathrm{LD}^1(\hat{\boldsymbol{\theta}}_{(i)})$ porque reduz os cálculos computacionais. A medida $\mathrm{D}_i(\boldsymbol{\theta})$ pode permitir a decomposição da distância de Cook em diferentes componentes de interesse. Se o valor de $\mathrm{D}_i(\boldsymbol{\theta})$ é elevado, a observação i é, possivelmente, um ponto influente, embora não exista um consenso sobre quais valores devem ser considerados elevados.

3.2.5.2 Influência local

Um problema que pode ocorrer com a deleção individual de pontos é deixar de detectar pontos conjuntamente influentes. Uma metodologia que não exige a eliminação de observações e que, portanto, permite avaliar a influência conjunta, denominada influência local, consiste em estudar o comportamento dos parâmetros estimados mediante uma pequena pertubação nos dados e/ou nas suposições do modelo. Cook (1986) propôs avaliar a influência por meio do afastamento da função verossimilhança, $\mathrm{LD}(\boldsymbol{\omega}) = 2(\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widehat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}))$, em que $\widehat{\boldsymbol{\theta}}$ e $\widehat{\boldsymbol{\theta}}_\omega$ são, respectivamente, os estimadores ML de $\boldsymbol{\theta}$ no modelo proposto e no modelo perturbado por $\boldsymbol{\omega}$. Mais especificamente, Cook (1986) propôs medir a influência a partir do estudo da curvatura normal do gráfico de influência $\boldsymbol{\omega} \times \mathrm{LD}(\boldsymbol{\omega})$ em uma vizinhança do ponto de não pertubação ω_0, na direção de um vetor unitário $\boldsymbol{d}$. Tal autor, usando conceitos de geometria diferencial, mostrou que a curvatura normal na direção $\boldsymbol{d}$ assume a forma $C_{\boldsymbol{d}} = 2|\boldsymbol{d}^\top \boldsymbol{B} \boldsymbol{d}|$, em que $\boldsymbol{B} = -\boldsymbol{\Delta}\ddot{\boldsymbol{\ell}}(\widehat{\boldsymbol{\theta}})^{-1}\boldsymbol{\Delta}$, com $\ddot{\boldsymbol{\ell}}(\widehat{\boldsymbol{\theta}})$ a matriz Hessiana avaliada em $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$, e $\boldsymbol{\Delta} = \frac{\partial \ell^2(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\omega}^\top}$ a matriz de pertubação avaliada em $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ e $\boldsymbol{\omega} = \boldsymbol{\omega_0}$.

Além disso, Cook (1986) argumenta que uma direção importante a se considerar é a direção da curvatura normal máxima, $\boldsymbol{d} = \boldsymbol{d}_{max}$. A curvatura normal máxima, $C_{\boldsymbol{d}_{max}}$, corresponde ao maior autovalor (em valor absoluto) da matriz $\boldsymbol{B}$ e $\boldsymbol{d}_{max}$ é o autovetor associado a este autovalor. Assim, o gráfico de $|C_{\boldsymbol{d}_{max}}|$ contra a ordem das observações pode revelar os pontos com maior influência na vizinhança de $\mathrm{LD}(\boldsymbol{\omega}_0)$, que podem ser responsáveis por mudanças consideráveis nas estimativas dos parâmetros, sob pequenas perturbações nos dados. Outra direção importante a se considerar é $\boldsymbol{d} = \boldsymbol{e_i}$, em que $\boldsymbol{e_i}$ é o vetor básico do $\mathbb{R}^n$, cuja i-ésima coordenada é 1 e as demais são 0. Neste caso, a curvatura normal é dada por $C_i = 2|b_{ii}|$, em que b_{ii} é o i-ésimo elemento da matriz $\boldsymbol{B}$, $i = 1, \ldots, n$. Da mesma forma, o gráfico de C_i contra a ordem das observações pode ser utilizado para identificar pontos influentes.

Embora a curvatura normal de Cook (1986) seja muito usada, outras medidas para avaliar a influência local foram propostas. A curvatura normal conformal, $B_d = \frac{C_d}{\mathrm{tr}(\boldsymbol{B})}$, proposta por Poon e Poon (1999), cujo cálculo não requer muito mais esforço que o cálculo de C_d, possui como vantagens ser invariante sob reparametrizações conformais e ser uma medida normalizada, o que facilita estabelecer um ponto de corte. Como as curvaturas C_d e B_d diferem apenas por uma constante positiva, o autovetor $\boldsymbol{d}_{max}$ fornece também a curvatura conformal

máxima $B_{\boldsymbol{d}_{max}}$. Poon e Poon (1999) sugerem que as coordenadas do vetor $B_{l_{max}}$, tomadas em módulo, que apresentam valores maiores que $\frac{1}{\sqrt{n}}$ indicam pontos influentes. No que diz respeito às curvaturas normais conformais na direção dos vetores básicos, $B_i = B_{\boldsymbol{e_i}} = \frac{2|b_{ii}|}{\mathrm{tr}(\boldsymbol{B})}$. Poon e Poon (1999) sugerem que a i-ésima observação seja considerada influente se $B_i > 2\overline{B}$, em que $\overline{B}$ é a média aritmética das curvaturas normais conformais básicas. Zhu e Lee (2001) sugerem um ponto de corte que além de considerar a média das curvaturas normais conformais também leva em consideração a variação dessas curvaturas. Considerando um caso particular da proposta de Zhu e Lee (2001), tem-se que a i-ésima observação é influente quando $B_i > \overline{B} + 2\mathrm{SE}(B)$, em que $\mathrm{SE}(B)$ é o desvio padrão das curvaturas normais conformais básicas. Outra medida que pode ser utilizada como medida de influência local é apresentada na discusão de Cook (1986) por Tsai (1986). Formalmente definida por Billor e Loynes (1993) tal medida utiliza o afastamento pela função verossimilhança modificada, definido por $\mathrm{LD}^*(\boldsymbol{\omega}) = -2(\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widehat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}|\boldsymbol{\omega}))$, em que $\ell(\boldsymbol{\theta}_{\boldsymbol{\omega}}|\boldsymbol{\omega})$ é a verossimilhança perturbada por $\boldsymbol{\omega}$. A principal vantagem dessa medida é que, diferente do que ocorre com a medida proposta por Cook (1986), a primeira derivada não se anula no ponto de não pertubação, e então tais autores propõem medir a influência pelo estudo da inclinação do gráfico de influência $\boldsymbol{\omega} \times \mathrm{LD}^*(\boldsymbol{\omega})$ em uma vizinhança de $\boldsymbol{\omega_0}$, na direção de um vetor unitário $\boldsymbol{d}$. Em particular, a inclinação máxima, S_{max}, e a correspondente direção máxima, d^*_{max}, devem ser consideradas. A direção máxima de LD^* no ponto $\boldsymbol{\omega}_0$ é a direção do gradiente $\nabla LD^*(\boldsymbol{\omega}_0)$, e consequentemente a inclinação máxima é dada por (Lobato Júnior (2005), Leiva et al.(2015c)),

$$S_{max} = \|\nabla \mathrm{LD}^*(\boldsymbol{\omega}_0)\| = 2\left\|\frac{\partial \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \omega}\right\|$$

avaliada em $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ e $\boldsymbol{\omega} = \boldsymbol{\omega}_0$. A abordagem de Billor e Loynes (1993) permite que se tenham os valores individuais das coordenadas do gradiente, $d^*_{max} = \nabla LD^*(\boldsymbol{\omega}_0)$ e assim podem-se avaliar quais observações contribuem com a maior parcela para o valor da inclinação máxima. Deste modo, o gráfico das coordenadas de S_{max} contra a ordem das observações pode ser utilizado para identificar os pontos influentes.

3.2.5.3 Alavanca generalizada

O propósito da alavanca em modelos de regressão linear é medir a influência que observações individuais exercem sobre seus próprios valores preditos (COOK, WEISBERG, 1982). Wei, Hu e Fung (1998) apresentaram uma definição de alavanca que generaliza, para modelos não lineares, o conceito de alavanca utilizado para modelos de regressão linear. Segundo Wei, Hu e Fung (1998), a matriz alavanca generalizada $\boldsymbol{GL}(\theta) = \frac{\partial \widehat{Y}}{\partial Y^\top}$ tem a forma

$$\boldsymbol{GL}(\boldsymbol{\theta}) = \boldsymbol{D}_{\boldsymbol{\theta}} \left(-\ddot{\boldsymbol{\ell}}(\boldsymbol{\theta})\right)^{-1} \boldsymbol{L}_{\boldsymbol{\theta Y}},$$

em que $\boldsymbol{D}_{\boldsymbol{\theta}} = \frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\theta}^\top}$, com $\boldsymbol{\mu} = \boldsymbol{X}\boldsymbol{\beta}$ o valor esperado para Y; $-\ddot{\boldsymbol{\ell}}(\boldsymbol{\theta})$ é a matriz de informação observada e $\boldsymbol{L}_{\boldsymbol{\theta Y}} = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial Y^\top}$.

O elemento GL_{ij} da matriz $\boldsymbol{GL}(\hat{\boldsymbol{\theta}})$ é a taxa de variação instantânea do i-ésimo valor

predito com respeito ao j-ésimo valor observado. Assim, os elementos da diagonal principal da matriz $\boldsymbol{GL}(\hat{\boldsymbol{\theta}})$ que apresentam valores altos são considerados pontos de alavanca, isto é, pontos cujo valor observado exerce grande influência sobre o valor predito. Considera-se que a i-ésima resposta é potencialmente influente se $GL_{ii} > \overline{GL} + 2\mathrm{SE}(GL)$, em que $\overline{GL}$ e $\mathrm{sd}(GL)$ são, respectivamente, a média e o desvio padrão de $GL_{11}, GL_{22}, \ldots, GL_{nn}$ (DE BASTIANI et al., 2015).

3.3 Atributos químicos do solo

O objetivo desta Seção é apresentar um breve relato sobre os atributos químicos do solo, em especial sobre os nutrientes que serão utilizados nos estudos de casos abordados nos artigos que compõe este trabalho: fósforo (P), magnésio (Mg) e cálcio (Ca).

De uma forma minimalista, pode-se dizer que solos são o meio no qual as culturas desenvolvem-se para alimentar e abrigar o mundo. Embora não seja o único, a fertilidade do solo é um fator vital para a produtividade e a compreensão de seus princípios é essencial para uma produção eficiente das culturas e para a proteção ambiental (INSTITUTO DA POTASSA E FOSFATO, 1998). Segundo Oliveira et. al (1991) a análise química do solo é um dos sistemas mais usados no Brasil para avaliar a fertilidade do solo e os objetivos principais da análises de rotina de solos para fins de fertilidade são obter informações para:

- avaliar o estado de fertilidade de uma região;
- manter o nível de fertilidade de determinada área;
- predizer a probabilidade de se obter resposta lucrativa com o uso de corretivos e fertilizantes;
- fornecer base para recomendação da quantidade de calcário e fertilizantes a aplicar.

Dezesseis elementos químicos são essenciais para o crescimento e desenvolvimento das plantas: os não minerais (carbono, hidrogênio e oxigênio) os quais são encontrados na atmosfera e na água e participam da fotossíntese, e os minerais fornecidos pelo solo e complementados pelo uso de calcário e adubação. Os nutrientes minerais, por sua vez, são divididos em três grupos: nutrientes primários (nitrogênio, fósforo e potássio), os quais geralmente tornam-se deficientes no solo antes dos demais, pois as plantas os usam em quantidades relativamente grandes; e os nutrientes secundários (cálcio, magnésio e enxofre) e micronutrientes (boro, cloro, cobre, ferro, manganês, molibdênio e zinco). Os nutrientes secundários e os micronutrientes são geralmente menos deficientes e usados pelas plantas em quantidades menores, mas são tão importantes quanto os nutrientes primários. Portanto, para a adequada fertilidade do solo, as plantas precisam tê-los à disposição quando e onde são necessário (INSTITUTO DA POTASSA E FOSFATO, 1998). Cabe aqui citar a Lei do Mínimo (base da produtividade): "O rendimento de uma colheita é limitado pela ausência de qualquer um dos nutrientes essenciais, mesmo que todos os demais estejam disponíveis em quantidades adequadas".

3.3.1 Fósforo (P)

Ainda segundo Instituto da Potassa e Fosfato (1998), o fósforo (P) é essencial para o crescimento das plantas. A deficiência de fósforo pode ser mais limitante para a produção das culturas do que qualquer outra deficiência, toxicidade ou doenças. A planta precisa de fósforo para completar seu ciclo normal de produção. O fósforo atua na fotossíntese, na respiração, no armazenamento e na transferência de energia, na divisão celular e no crescimento das células. Promove a formação e o crescimento prematuro das raízes, melhora a qualidade de frutas, verduras e graníferas, sendo essencial para a formação das sementes. Aumenta a resistência aos rigores do inverno, favorece a resistência às doenças, acelera a maturidade e também está envolvido nas transferências dos códigos genéticos de uma geração para outra. Um grande problema da deficiência de fósforo no solo é que os sintomas visuais geralmente não são tão claros, sendo assim, é conveniente confirmar os indícios visuais com análises do solo e das plantas (RAIJ, 1991; INSTITUTO DA POTASSA E FOSFATO, 1998). De acordo com a classificação dada por COAMO-CODETEC (2001), a concentração de fósforo em solo argiloso é considerada como segue. Para o plantio de soja, é: (i) baixa para níveis inferiores a 3 mg/dm^3; (ii) média para níveis entre 3,1 e 6 mg/dm^3; (iii) alta para níveis entre 6,1 e 9 mg/dm^3; e (iv) muito alta para níveis superiores a 9 mg/dm^3. Para a plantação de milho, a concentração de fósforo é: (i) baixa para níveis inferiores a 2 mg/dm^3; (ii) média para níveis entre 2,1 e 4,5 mg/dm^3; (iii) alta para níveis entre 4,6 e 11 mg/dm^3; e (iv) muito alta para níveis superiores a 11 mg/dm^3.

3.3.2 Cálcio (Ca)

O cálcio dentro da planta estimula o desenvolvimento das raízes e das folhas; forma compostos que são parte das paredes celulares, reforçando a estrutura das plantas; ajuda a ativar vários sistemas enzimáticos; influencia a produção por diminuir a acidez do solo e é exigido em grande quantidade pelas bactérias fixadoras do nitrogênio (INSTITUTO DA POTASSA E FOSFATO, 1998).

3.3.3 Magnésio (Mg)

Segundo Wland (2007), o magnésio, tem como função predominante na planta, a atuação na molécula da clorofila e está envolvido na fixação de CO_2. Assim, em casos de baixa disponibilidade de magnésio nas folhas, a fixação de CO_2 fica comprometida. Durante a assimilação do CO_2 as moléculas de açúcar são formadas a partir da água e do dióxido de carbono, utilizando a energia solar. Esta é uma reação básica para a síntese de outros constituintes das plantas, tais como amido, proteína, gorduras e vitaminas. A deficiência de magnésio resulta em há dificuldade na translocação de carboidratos para a raiz e prejudica o desenvolvimento do sistema radicular que, por sua vez, reduzirá a absorção de outros nutrientes. A nutrição com magnésio, além dos critérios nutricionais, com reflexos na saúde humana e animal, confere qualidade ao produto final, como número, peso e conteúdo proteico dos grãos;

doçura, cor, sabor e maciez dos frutos. Assim, a presença em dose inadequada do magnésio no sistema solo-planta prejudica o crescimento e a vitalidade da planta. O magnésio também ajuda no metabolismo do fosfato, na respiração da planta e na ativação de vários sistemas enzimáticos (INSTITUTO DA POTASSA E FOSFATO, 1998). Além disso, a deficiência de magnésio no solo apresenta uma faixa de deficiência oculta, em que os sintomas não são visíveis, mas exercem forte impacto negativo sobre a produção e a qualidade das culturas, principais diretrizes na produção das culturas. Assim, a concentração de magnésio no solo deve ser atenciosamente monitorada (WLAND, 2007).

3.3.4 A relação Ca:Mg

Ainda segundo Wland (2007), os sistemas agrícolas estão expostos ás ações naturais e às atividades humanas. Dentre as ações naturais, a deficiência e o desequilíbrio entre nutrientes no solo são considerados os principais problemas. Ademais, além do conhecimento dos teores dos minérios do solo, pesquisas apontam para a necessidade de conhecer as relações entre os nutrientes presentes no solo. Uma relação bastante conhecida no meio agronômico, talvez a mais conhecida, é a relação cálcio e magnésio (Ca:Mg). Esta relação é importante por haver uma competição pelos sítios de adsorção (adesão de moléculas de um fluido (o adsorvido) a uma superfície sólida (o adsorvente)) no solo e na absorção pelas raízes. Como consequência, a presença excessiva de um pode prejudicar os processos de adsorção e absorção do outro e afetar o desenvolvimento das plantas. De acordo com Instituto da Potassa e Fosfato (1998), o desequilíbrio entre cálcio e o magnésio no solo pode acentuar a deficiência de magnésio. Quando a relação Ca:Mg torna-se muito alta, as plantas podem absorver menos magnésio, isto pode ocorrer, por exemplo, quando o agricultor usa somente calcário calcítico por muitos anos, em solos relativamente pobres em magnésio.

4 REFERÊNCIAS BIBLIOGRÁFICAS

ABRAMOWITZ, M.; STEGUN, I.A. **Handbook of Mathematical Functions.** New-York: Dover Press, 1972.

AKAIKE, H. Information theory and a extension of the maximum likelihood principle. **Journal of the Royal Statistical Society**, v. 51, p. 469-483, 2002.

ANDERSON, J.R.; HARDY, E.E.; ROACH, J.T.; WITMER, R.E. **A land use and land cover classification system for use with remote sensor data.** Washington: United States Government Printing Office, 1976.

ASSUMPÇÃO, R.A.B.; URIBE-OPAZO, M.A.; GALEA, M. Local influence for spatial analysis of soil physical properties and soybean yield using Student-t distribution. **Revista Brasileira de Ciência do Solo,** v. 35, p. 1917-1926, 2011.

ASSUMPÇÃO, R.A.B.; URIBE-OPAZO, M.A.; GALEA, M. Analysis of local influence in geostatistics using Student-t distribution. **Journal of Applied Statistics,** v. 41, p. 2323-2341, 2014.

BILLOR, N.; LOYNES, R.M. Local influence: a new approach. **Communications in Statistics-Theory and Methods,** v. 22, p. 1595-1611, 1993.

BILLOR, N.; LOYNES, R.M. An application of the local influence approach to ridge regression. **Journal of Applied Statistics,** v. 26, p. 177-183, 1999.

BIRNBAUM, Z.W.; SAUNDERS, S.C. A probabilistic interpretation of Miner's rule. **Journal of Applied Mechanics**, v. 16, p. 637-652, 1968.

BIRNBAUM, Z.W.; SAUNDERS, S.C. A new family of life distributions. **Journal of Applied Probability,** v. 6, p. 319-327, 1969a.

BIRNBAUM, Z.W.; SAUNDERS, S.C. Estimation for a family of life distributions with applications to fatigue. **Journal of Applied Probability,** v. 6, p. 328-347, 1969b.

BORSSOI, J.A.; URIBE-OPAZO, M.A.; GALEA, M. Técnicas de diagnósticos de influência local na análise espacial da produtividade da soja. **Engenharia Agrícola,** v. 31, p. 376-387, 2011.

COOPERATIVA AGROPECUÁRIA MOURÃOENSE LTDA.(COAMO)-COOPERATIVA CENTRAL DE DESENVOLVIMENTO TECNOLÓGICO E ECONÔMICO LTDA.(COODETEC). **Fertilidade do solo e nutrição de plantas.** Cascavel: Coamo-Coodetec, 2001.

COOK, R.D. Detection of influential observations in linear regressions. **Technometrics,** v. 19, p. 15-18, 1977.

COOK, R.D.; WEISBERG, S. **Residuals and Influence in Regression,** London: Chapman and Hall, 1982.

COOK, R.D. Assessment of local influence. **Journal of the Royal Statistics Society,** Série B, v. 48, p. 133-169, 1986.

COOK, R.D. Influence assessment. **Journal of Applied Statistics,** v. 14, p. 117-131, 1987.

CRESSIE, N.A.C. **Statistics for Spatial Data.** New-York: Jonh Wiley & Sons, 2015.

DE BASTIANI, F.; CYSNEIROS, A.H.M.A.; URIBE-OPAZO, M.A.; GALEA, M. Influence diagnostics in elliptical spatial linear models. **TEST,** v. 24, p. 322-340, 2015.

DÍAZ-GARCÍA, J. A.; LEIVA, V. A new family of life distributions based on the contoured elliptically distributions. **Journal of Statistical Planning and Inference,** v. 128, p. 445-457, 2005.

DIGGLE, P.J.; RIBEIRO JÚNIOR, P.J. **Model-Based Geoestatistics.** New York: Springer, 2007.

DUNN, P. K.; SMYTH, G.K. Randomized Quantile Residuals. **Journal of Computational and Graphical Statistics,** v. 5, p. 236-244, 1996.

EMPRESA BRASILEIRA DE INVESTIGAÇÃO AGROPECUÁRIA (Embrapa). **Sistema brasileiro de classificação de solo.** Rio de Janeiro: Empresa Brasileira de Investigação Agropecuária-Centro Nacional de Pesquisa de Solos (Embrapa-SPI), 2009.

FARACO; M.A.; URIBE-OPAZO, M.A.; SILVA, E.A.; JOHANN, J.A.; BORSSOI, J.A. Seleção de modelos de variabilidade especial para elaboração de mapas temáticos de atributos físicos do solo e produtividade da soja. **Revista Brasileira de Ciência do Solo**, v. 32, p. 463-476, 2008.

GOTWAY, C.A.; HARTFORD, A.H. Geostatistical methods for incorporating auxiliary information in the prediction of spatial variables. **Journal of agricultural biology and environmental statistics**, v. 1, p. 17-39, 1996.

GUEDES, L.P.C.; URIBE-OPAZO, M. A; JOHANN, J. A.; SOUZA, E. G. Anisotropia no estudo da variabilidade espacial de algumas variáveis químicas do solo. **Revista Brasileira de Ciência do Solo,** v. 32, p. 2217-2226, 2008.

GUEDES, L.P.C.; RIBEIRO JÚNIOR, P.J.; PIEDADE, S.M.S; URIBE-OPAZO, M.A. Optimization of spatial sample configurations using hybrid genetic algorithm and simulated annealing. **Chilean Journal of Statistics**, v. 2, p. 39-50, 2011.

HENGL, T.; HEUVELINK, G.B.M.; STEIN, A. Comparison of kriging with external drift and regression-kriging. Technical report. **International Institute for Geo-information Science and Earth Observation (ITC),** 2003. Disponível em: <http://www.itc.nl/library/Academic-output>. Acesso em 12 nov. 2015.

HENGL, T.; HEUVELINK, G.B.M.; STEIN, A. A generic framework for spatial prediction of soil variables based on regression-kriging. **Geoderma,** v. 120, p. 75-93, 2004.

HOAGLIN, D.C.; WELSCH, R.E. The hat matrix in regression and ANOVA. **The American Statistician,** v. 32, p. 17-22, 1978.

INSTITUTO DA POTASSA & FOSFATO. **Manual internacional de fertilidade do solo**. Tradução e adaptação de Alfredo Scheid Lopes. Piracicaba: Potafos, 1998.

ISAAKS, E.; SRIVASTAVA, R. **An Introduction to Applied Geostatistics.** New York: Oxford University Press, 1989.

JOHNSON, N.L.; KOTZ, S.; BALAKRISHNAN, N. **Continuous Univariate Distributions.** New York: Jonh Wiley & Sons, 1994. v.1

JOHNSON, N.L.; KOTZ, S.; BALAKRISHNAN, N. **Continuous Univariate Distributions.** New York: Jonh Wiley & Sons, 1995. v.2

JOURNEL, A.G.; HUIJBREGTS, C.J. **Mining Geostatistics**. London: Academic Press, 1978.

KRIGE, D.G. A statistical approach to some basic mine valuation problems on the Witwatersrand. **Journal of the Chemical, Metallurgical and Mining Society of South Africa,** v. 52, p. 119-139, 1951.

KRIPPENDORFF, K. **Content Analysis: An Introduction to its Methodology.** Thousand Oaks: Sage Publications, 2004.

KUNDU, D.; BALAKRISHNAN, N.; JAMALIZADEH, A. Bivariate Birnbaum-Saunders distribution and associated inference. **Journal of Multivariate Analysis,** v. 101, p. 113-125, 2010.

KUNDU, D.; BALAKRISHNAN, N.; JAMALIZADEH, A. Generalized multivariate Birnbaum-Saunders distributions and related inferential issues. **Journal of Multivariate Analysis,** v. 116, p. 230-244, 2013.

LANDIM, P.M.B. **Introdução aos métodos de estimação espacial para confecção de mapas**. Rio Claro: DGA, IGCE, UNESP, Lab. Geomatemática,Texto Didático 02, 2000. Disponível em <www.rc.unesp.br/igce/aplicada/textodi.html>. Acesso em: 02 out. 2015.

LANGE, K.L.; LITTLE, R.J.; TAYLOR, J.M. Robust statistical modeling using the t distribution. **Journal of the American Statistical Association,** v. 84, p. 881-896, 1989.

LANGE, K. **Numerical Analysis for Statisticians.** New York: Springer, 2000.

LEHMANN, E.L. **Elements of Large-Sample Theory.** New York: Springer-Verlag, 1999.

LEIVA, V.; BARROS, M.; PAULA, G.A.; GALEA, M. Influence diagnostics in log-Birnbaum-Saunders regression models with censored data. **Computational Statistics and Data Analysis,** v. 51, p. 5694-5707, 2007.

LEIVA, V.; BARROS, M.; PAULA, G.A.; SANHUEZA, A. Generalized Birnbaum-Saunders distributions applied to air pollutant concentration. **Environmetrics,** v. 19, p. 235-249, 2008.

LEIVA, V.; SANHUEZA, A.; ANGULO, J.M. A length-biased version of the Birnbaum-Saunders distribution with application in water quality. **Stochastic Environmental Research and Risk Assessment,** v. 23, p. 299-307, 2009.

LEIVA, V.; BARROS, M.; PAULA, G. A. **Generalized Birnbaum-Saunders Models using R**. Recife: XI Escola de Modelos de Regressão, 2009.

LEIVA, V.; VILCA, F.; BALAKRISHNAN, N.; SANHUEZA, A. A skewed Sinh-Normal distribution and its properties and application to air pollution. **Communications in Statistics-Theory and Methods,** v. 39, p. 426-443, 2010.

LEIVA, V.; ROJAS, E.; GALEA, M.; SANHUEZA, A. Diagnostics in Birnbaum-Saunders accelerated life models with an application to fatigue data. **Applied Stochastic Models in Business and Industry,** v. 30, p. 115-131, 2014c.

LEIVA, V. **The Birnbaum-Saunders Distribution**. New York: Academic Press, 2015.

LEIVA, V.; FERREIRA, M.; GOMES, M.I.; LILLO, C. Extreme value Birnbaum-Saunders regression models applied to environmental data. **Stochastic Environmental Research and Risk Assessment** (in press), 2015a. Disponível em <http://dx.doi.org/10.1007/s00477-015-1069-6, 2015a>. Acesso em: 08 out. 2015.

LEIVA, V.; MARCHANT, C.; RUGGERI, F.; SAULO, H. A criterion for environmental assessment using Birnbaum-Saunders attribute control charts. **Environmetrics** (in press), 2015b. Disponível em <http://dx.doi.org/10.1002/env.2349>. Acesso em: 08 out. 2015.

LEIVA, V.; LIU S.; SHI, L.; CYSNEIROS, F.J.A. Diagnostics in elliptical regression models with stochastic restrictions applied to econometrics. **Journal of Applied Statistics** (in press), 2015c. Disponível em <http://dx.doi.org/10.1080/02664763.2015>. Acesso em: 20 nov. 2015.

LEIVA, V.; SANTOS-NETO, M; CYSNEIROS, F.J.A.; BARROS, M. A methodology for stochastic inventory models based on a zero-adjusted Birnbaum-Saunders distribution. **Applied Stochastic Models in Business and Industry** (in press), 2015d. Disponível em:<http://dx.doi.org/10.1007/s10260-015-0329-4>. Acesso em: 20 nov. 2015.

LEIVA, V.; TEJO, M.; GUIRAUD, P.; SCHMACHTENBERG, O; ORIO, P.; MARMOLEJO-RAMOS, F. Modeling neural activity with cumulative damage distributions. **Biological Cybernetics,** v. 109, p. 421-433, 2015e.

LOBATO JÚNIOR, D. **Influêncial Local em Modelos de Regressão,** 2005, 106 f. Dissertação (Mestrado em Matemática). Universidade Federal de Campina Grande, Campina Grande.

MARDIA, K.V.; MARSHALL, R.J. Maximum likelihood estimation of models for residual covariance in spatial regression. **Biometrika,** v. 71, p. 135-146, 1984.

MATÉRN, B. Spatial Variation. **Lecture notes in Statistics.** New York: Springer, 1986.

MATHERON, G. Principles of geoestatistics. **Economic Geology**, v. 58, p. 1246-1266, 1963.

MOOD, A.M.; GRAYBILL, F.A.; BOES, D.C. **Introduction to the theory of statistics.** Singapore: McGraw-Hill Book, 1974.

OLIVEIRA, A.J.; GARRIDO, W.E.; ARAUJO, J.D.; LOURENÇO, S. (Org.). **Métodos de pesquisa em fertilidade do solo.** Brasília: EMBRAPA-SEA, 1991.

PAN, J.; FEI, Y.; FOSTER, P. Case-Delection Diagnostics for Linear Models. **Technometrics,** v. 56, p. 269-281, 2014.

PAULSON, E. An approximate normalization of the analysis of variance distribution. **The annais of Mathematical Statistics,** v.13, p.233-235, 1942.

POON, W.; POON, Y.S. Conformal normal curvature and assessment of local influence. **Royal Statistical Society,** v. 61, p. 51-61, 1999.

RAIJ, B.V. **Fertilidade do solo e adubação.** Piracicaba: Ceres, Potafos, 1991.

RIECK, J.R. **Statistical Analysis for the Birnbaum-Saunders fatigue life distribution.** 1989, 132 f. Dissertação (Mestrado em Estatística). Graduate School de Clemson University, Clemson.

RIECK, J.R.; NEDELMAN, J.R. A log-linear model for the Birnbaum-Saunders distribution. **Technometrics,** v. 33, p. 51-60, 1991.

SEIDEL, E.J.; OLIVEIRA, M.S. Proposal of a generalization for Exponential and Gaussian semivariogram models. **Semina: Ciências Exatas e Tecnológicas,** v. 34, p. 125-132, 2013.

SOARES, A. **Geoestatística para as ciências da terra e do ambiente.** Lisboa: IST Press, 2000.

TSAI, C.L. Discussion of assessment of local influence by R. D. Cook. **Journal of the Royal Statistics Society Series B,** v. 48, p. 165, 1986.

VILCA, F.; SANHUEZA, A.; LEIVA, V.; CHRISTAKOS, G. An extended Birnbaum-Saunders model and its application in the study of environmental quality in Santiago, Chile. **Stochastic Environmental Research and Risk Assessment,** v. 24, p. 771-782, 2010.

WEI, B.; HU, Y.Q.; FUNG, W.K. Generalized leverage and its applications. **Scandinavian Journal of Statistics,** v. 25, p. 25-37, 1998.

WLAND, T. Magnésio no solo e nas plantas. **Informações Agronômicas,** n. 117, p. 19-21, 2007.

ZHU, H.; LEE, S. Local Influence for incomplete-data models. **Royal Statistical Society,** v. 63, p. 111-126, 2001.

5 ARTIGOS

5.1 ARTIGO 1: Modelagem espacial Birnbaum-Saunders e diagnósticos aplicados a dados de engenharia agrícola

Resumo: Aplicações de modelos estatísticos para descrever a dependência espacial em dados georreferenciados são muito difundidas em diversas áreas dos conhecimento, incluindo ciências ambientais. A maioria dessas aplicações assumem que os dados seguem uma distribuição Gaussiana. Contudo, em muitas delas, a suposição de normalidade, e até mesmo uma suposição mais geral de simetria, não são apropriadas. Em aplicações não espaciais, quando a distribuição dos dados é unimodal e assimétrica positiva, a distribuição Birnbaum-Saunders se destaca. Assim, este artigo propõe um modelo log-linear espacial baseado na distribuição Birnbaum-Saunders. Os parâmetros do modelo são estimados pelo método da máxima verossimilhança. Diagnósticos de influência local são realizados para avaliar a sensibilidade dos estimadores às pertubações na variável-resposta. Como ilustração, o modelo proposto e seus diagnósticos são usados para analisar um conjunto de dados simulados e dados agrícolas reais, em que se considera a variabilidade espacial da concentração de Fósforo no solo, a qual é extremamente importante para o manejo agrícola.

Palavras-chave: análise de dados espaciais, distribuições assimétricas, influência local, método da máxima verossimilhança, modelo Matérn, não normalidade, simulação de Monte Carlo, *software* R.

5.1.1 Introdução

Modelos estatísticos espaciais lineares levam em consideração a dependência espacial de uma variável sobre o espaço usando dados georeferenciados. Esses modelos são essenciais em muitas áreas (KRIGE, 1951; MARDIA; MARSHALL, 1984; WALTER; GOTWAY, 2004; MILITINO et al., 2011). Estudos recentes incluem Borssoi et al. (2011); Uribe-Opazo, Borssoi e Galea (2012) e Grzegozewski et al. (2013). Todos esses trabalhos consideram a distribuição normal (ou gaussiana) para modelar a variabilidade espacial. Contudo, tal suposição nem sempre é apropriada (DAVIS, 1952; LANGE et al., 1989). Uma maneira de lidar com a não normalidade é utilizar transformações para alcançar a normalidade, pelo menos aproximada. Não obstante, trabalhando-se com transformações de dados, problemas, tais como a dificuldade de interpretar os resultados a partir da análise na escala original, podem ser introduzidos na modelagem (AZZALINI; CAPITANIO, 1999). Quando a suposição de normalidade não poder ser considerada, uma abordagem alternativa consiste em realizar a modelagem com alguma distribuição não normal adequada para os dados sob análise. Por exemplo, Assumpção, Uribe-Opazo e Galea (2014) conduziram um estudo geoestatístico usando a distribuição t-Student, que tem caudas mais pesadas do que a distribuição normal. De Bastiani et al.

(2015) estudaram modelagem espacial e diagnóstico baseados em distribuição da família elíptica (simétricas), que tem como membros as distribuições Gaussiana e t-Student. Contudo, modelos elípticos ainda não são apropriados se os dados seguem uma distribuição assimétrica.

Existem poucos trabalhos na literatura que investigam o uso de distribuições assimétricas para analisar dados espaciais. Entretanto, em situações não espaciais, várias distribuições têm sido propostas para modelar fenômenos que dão origem a dados assimétricos, tais como as distribuições Birnbaum-Saunders (BS), exponencial, gamma, log-normal e Weibull (JOHNSON et al. 1994, 1995). Em particular, a distribuição BS foi proposta para a modelagem de variáveis aleatórias provenientes de processos de fadiga por Birnbaum e Saunders (1969). Aplicações em ciências da terra da distribuição BS têm sido consideradas, por exemplo, por Leiva et al. (2008); Leiva, Sanhueza e Angulo (2009); Leiva et al. (2015a); Podlaski (2008); Vilca et al. (2010); Marchant et al. (2013) e Saulo et al. (2013). O que faz esta distribuição atrativa para a análise de dados assimétricos são suas propriedades e sua relação com a distribuição normal (JOHNSON; KOTZ; BALAKRISHNAN, 1995). Em constraste com a sua aplicação original para processos de fadiga, Leiva et al. (2015b) justificaram, usando argumentos teóricos baseados na lei de efeitos proporcionais, porque a distribuição BS é adequada para a modelagem de dados da terra e ambientais. Rieck e Nedelman (1991) definiram a relação entre a distribuição BS e sua versão logarítmica, chamada de distribuição log-BS. Eles usaram essa relação para propor um modelo BS de efeito fixo; enquanto Villegas, Paula e Leiva (2011) consideraram um modelo BS de efeito misto. Extensões multivariadas da distribuição BS e um modelo BS de efeito fixo foram apresentados por Kundu, Balakrishnan e Jamalizadeh (2013) e Marchant, Leiva e Cysneiros (2015), respectivamente. Uma abordagem para a modelagem BS espacial foi fornecida por Xia, Zeephongsekul e Packer (2011), que apresentaram uma metodologia baseada em processos semi-Markov para propor um modelo espaço-temporal para a movimentação de turistas. Os autores consideraram várias distribuições, incluindo o modelo BS. Até o momento, no entanto, modelos espaciais baseados na distribuição BS não têm sido estudados em exemplos práticos.

A identificação de casos que possam produzir mudanças substanciais nos parâmetros estimados é uma importante etapa em qualquer investigação estatística. A tarefa de detectar possíveis casos atípicos pode ser realizada pela eliminação, um-a-um, de casos do conjunto de dados e medindo os efeitos nos parâmetros estimados: esta técnica é conhecida como influência global (COOK, 1987). Outro método para a detecção de casos que possam ser potencialmente influentes, conhecido como influência local, foi proposto por Cook (1986). Esse método estuda o efeito de pequenas perturbações introduzidas nos modelos e/ou nos dados sobre as estimativas de máxima verossimilhança (ML). Diferentes esquemas de pertubação são considerados para avaliar a sensibilidade das estimativas ML dos parâmetros do modelo a tais perturbações. O método de influência local tem pelo menos duas vantagens sobre o método de influência global: tem menor custo computacional, especialmente quando o número de casos é grande, e permite a detecção de grupo de dados que exercem influência conjunta. Zhu et al. (2007) propuseram uma metodologia para escolher um esquema de perturbação que é apropriado para o modelo particular a ser considerado. Gimenez e Galea (2013) aplicaram o

método proposto por Zhu et al. (2007) em modelos heteroscedásticos com erros de medição funcional. Galea, Leiva e Paula (2004) aplicaram o método de influência local em modelos BS de efeito fixo, enquanto Leiva et al. (2014); Leiva et al. (2015c) e Liu et al. (2015) utilizaram ferramentas de diagnósticos em modelos de vida acelerada, em modelos de efeito fixo com restrições estocásticas e no modelo linear possivelmente heteroscedástico com restrições exatas. Em modelagem espacial, técnicas de diagnósticos foram discutidas para modelos Gaussianos por Militino, Palacius e Ugarte (2006) e Uribe-Opazo, Borssoi e Galea (2012), para modelos t-Student por Assumpção, Uribe-Opazo e Galea (2014), e para modelos elípticos por De Bastiani et al. (2015).

O principal objetivo deste artigo foi desenvolver um modelo espacial log-linear baseado na distribuição BS e obter os diagnósticos correspondentes a este modelo. Esta distribuição pode ser mais apropriada que a distribuição Gaussiana na análise de dados espaciais com comportamento assimétrico positivo. Estimadores ML dos parâmetros do modelo e ferramentas de diagnósticos de influência local são obtidos para o modelo espacial BS. Especificamente, na Seção 5.1.2, fornecem-se informações preliminares sobre as distribuições BS e log-BS uni e multivariadas e sobre modelagem espacial. Na Seção 5.1.3, o modelo espacial BS log-linear é formulado; realiza-se a estimação dos seus parâmetros usando o método ML e encontra-se um esquema de pertubação apropriado para a variável-resposta, a partir da metodologia proposta por Zhu et al. (2007). Na Seção 5.1.4, serão conduzidos dois estudos de simulação de Monte Carlo (MC) para avaliar a performance do correspondente estimador ML e das ferramentas de diagnósticos. A Seção 5.1.5 ilustra potenciais aplicações do modelo proposto e seus diagnósticos com dados reais provenientes da engenharia agrícola. Na Seção 5.1.6, apresentam-se algumas conclusões e possíveis trabalhos futuros. A álgebra detalhada é apresentada nos anexos.

5.1.2 Preliminares

5.1.2.1 A distribuição Birnbaum-Saunders

Se uma variável aleatória T segue uma distribuição BS com parâmetro de forma α e parâmetro de escala γ, usa-se a notação $T \sim \mathrm{BS}(\alpha, \gamma)$. A distribuição pode ser definida por sua função de distribuição acumulada (fda) dada por

$$F_T(t;\alpha,\gamma) = \Phi\left(\frac{1}{\alpha}\left(\sqrt{t/\gamma} - \sqrt{\gamma/t}\right)\right), \quad t>0, \alpha>0, \gamma>0, \tag{5.1}$$

em que $\Phi(\cdot)$ é a fda da distribuição normal padrão. Então, a função densidade de probabilidade (fdp) de T obtida de (5.1) é expressa como

$$f_T(t;\alpha,\gamma) = \frac{1}{2\alpha}\left(\sqrt{1/\gamma t} + \sqrt{\gamma/t^{\frac{3}{2}}}\right)\phi\left(\frac{1}{\alpha}\left(\sqrt{t/\gamma} - \sqrt{\gamma/t}\right)\right), \quad t>0, \alpha>0, \gamma>0. \tag{5.2}$$

em que $\phi(\cdot)$ é a fdp da normal padrão. Então, a fdp dada em (5.2) pode ser reescrita como

$$f_T\left(t;\alpha,\gamma\right) = \frac{\exp(\alpha^{-2})}{2\alpha\sqrt{2\pi\gamma}}\exp\left(-\frac{1}{2\alpha^2}\left(\frac{t}{\gamma}+\frac{\gamma}{t}\right)\right)t^{-\frac{3}{2}}\left(t+\gamma\right), \quad t>0, \alpha>0, \gamma>0, \tag{5.3}$$

Também pode-se dizer que uma variável aleatória contínua T tem uma distribuição BS com parâmetros $\alpha > 0$ e $\gamma > 0$, se, e somente se $Z = (1/\alpha)\big(\sqrt{T/\gamma} - \sqrt{\gamma/T}\big) \sim \mathrm{N}(0,1)$. Algumas propriedades da distribuição BS são apresentadas como segue. Se $T \sim \mathrm{BS}(\alpha,\gamma)$, então: (i) $\mathrm{E}(T) = \gamma(1+\alpha^2/2)$ e $\mathrm{Var}(T) = (\alpha\gamma)^2(1+5\alpha^2)$; (ii) se $b > 0$, então $bT \sim \mathrm{BS}(\alpha, b\gamma)$, isto é, a distribuição BS é fechada para a multiplicação por escalar; (iii) $1/T \sim \mathrm{BS}(\alpha, 1/\gamma)$, o que significa que a distribuição BS é fechada sobre reciprocidade; (iv) a mediana da distribuição de T é γ, que pode se obtida diretamente fazendo $q = 0,5$ na sua função quantil dada por $t(q;\alpha,\gamma) = F_T^{-1}(q;\alpha,\gamma) = \beta(\alpha z(q)/2 + \sqrt{(\alpha z(q)/2)^2+1}\,)^2$, para $0 < q < 1$, em que $z(q)$ é a função quantil da distribuição normal padrão; e (v) a distribuição BS é assimétrica positiva quando α cresce e aproximadamente simétrica, em torno de γ quando α tende a zero (veja Figura 16 (a)).

5.1.2.2 A distribuição log-Birnbaum-Saunders

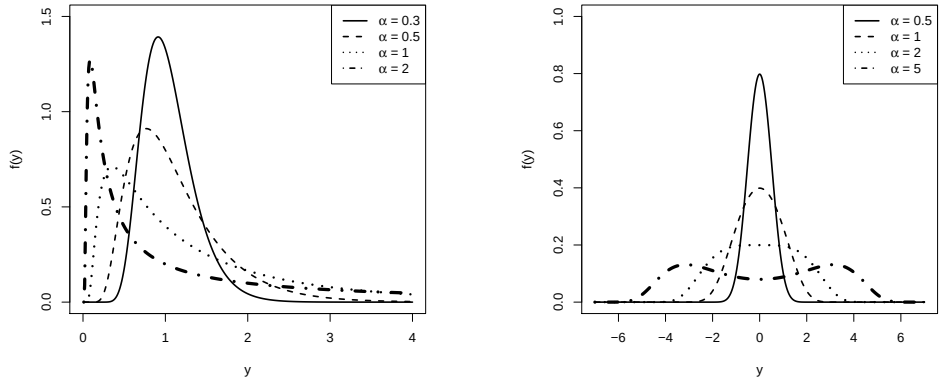
Uma variável aleatória contínua Y tem uma distribuição log-BS com parâmetro de forma $\alpha > 0$ e parâmetro de localização $\mu \in \mathbb{R}$, que é denotada por $\log\text{-}\mathrm{BS}(\alpha,\mu)$, se, e somente se $Z = (2/\alpha)\mathrm{senh}((Y-\mu)/2) \sim \mathrm{N}(0,1)$. Então, a fda de Y é dada por

$$F_Y(y;\alpha,\mu) = \Phi\left(\frac{2}{\alpha}\mathrm{senh}\left(\frac{y-\mu}{2}\right)\right), \quad -\infty < y, \mu < +\infty, \alpha > 0. \tag{5.4}$$

Consequentemente, de (5.4), a fdp de Y é obtida como

$$f_Y(y;\alpha,\mu) = \frac{1}{\alpha\sqrt{2\pi}}\cosh\left(\frac{y-\mu}{2}\right)\exp\left(-\frac{2}{\alpha^2}\mathrm{senh}^2\left(\frac{y-\mu}{2}\right)\right), \quad -\infty < y, \mu < +\infty, \alpha > 0. \tag{5.5}$$

Algumas propriedades da distribuição log-BS são apresentadas a seguir. Se $Y \sim \log\text{-}\mathrm{BS}(\alpha,\mu)$, então: (i) $T = \exp(Y) \sim \mathrm{BS}(\alpha,\gamma)$, o que significa que a fdp da distribuição log-BS dada em (5.5) pode ser obtida da fdp da normal padrão ou da fdp da BS definida em (5.3); (ii) $\mathrm{E}(Y) = \mu$; (iii) não existe forma fechada para a variância de Y, mas aproximação assintótica para a função geradora de momentos da distribuição log-BS, segue que, se $\alpha \to 0$, então $\mathrm{Var}(T) = \alpha^2 - \alpha^4/4$, e se $\alpha \to \infty$, então $\mathrm{Var}(T) = 4(\log^2(\sqrt{2}\alpha) + 2 - 2\log(\sqrt{2}\alpha))$; (iv) se $X = \pm Y + d$, então $X \sim \log\text{-}\mathrm{BS}(\alpha, \pm\mu + d)$; e (v) a distribuição log-BS é simétrica em torno de μ, unimodal para $\alpha \leq 2$ e bimodal para $\alpha > 2$ (Figura 16 (b)).



(a) (b)

Figura 16 Funções distribuição de probabilidade das distribuições BS$(\alpha, 1)$ (a) e log-BS$(\alpha, 0)$ (b) para o valor do parâmetro de forma α indicado.

Se um vetor aleatório $\boldsymbol{Y} = (Y_1, \ldots, Y_n)^\top$ segue uma distribuição log-BS n-variada, usa-se a notação $\boldsymbol{Y} \sim \log\text{-BS}_n(\boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\Sigma})$. Nesse caso, o vetor de parâmetros de forma é $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)^\top$, com $\alpha_i > 0$, o vetor de parâmetro de locação é $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)^\top$, com $-\infty < \mu_i < +\infty$, para $i = 1, \ldots, n$, e $\boldsymbol{\Sigma}$ é uma matriz positiva definida (não singular), $n \times n$, de parâmetros de escala. A fda de $\boldsymbol{Y}$ pode ser definida de (5.5) como

$$F_{\boldsymbol{Y}}(\boldsymbol{y}; \boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \Phi_n\left(\frac{2}{\alpha_1}\text{senh}\left(\frac{y_1 - \mu_1}{2}\right), \ldots, \frac{2}{\alpha_n}\text{senh}\left(\frac{y_n - \mu_n}{2}\right); \boldsymbol{\Sigma}\right), \quad \boldsymbol{y} \in \mathbb{R}^n, \tag{5.6}$$

em que $\Phi_n(\cdot; \boldsymbol{\Sigma})$ é a fda da distribuição normal padrão n-variada com matriz de variância e covariância $\boldsymbol{\Sigma}$. Portanto, a fdp de $\boldsymbol{Y}$ pode ser obtida de (5.6) como

$$f_{\boldsymbol{Y}}(\boldsymbol{y}; \boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \phi_n\left(\frac{2}{\alpha_1}\text{senh}\left(\frac{y_1 - \mu_1}{2}\right), \ldots, \frac{2}{\alpha_n}\text{senh}\left(\frac{y_n - \mu_n}{2}\right); \boldsymbol{\Sigma}\right) \prod_{i=1}^{n} \frac{1}{\alpha_i}\cosh\left(\frac{y_i - \mu_i}{2}\right), \tag{5.7}$$

para $\boldsymbol{y} \in \mathbb{R}^n$, em que $\phi_n(\cdot; \boldsymbol{\Sigma})$ é a f.d.p. da normal n-variada com matriz de variância e covariância $\boldsymbol{\Sigma}$.

5.1.2.3 Modelos espaciais

Considere um processo estocástico $\{Y(\boldsymbol{s}), \boldsymbol{s} \in D\}$, definido sobre uma região D, com $D \subset \mathbb{R}^2$, descrito pelo modelo espacial linear

$$Y(\boldsymbol{s}) = \mu(\boldsymbol{s}) + \varepsilon(\boldsymbol{s}), \quad \boldsymbol{s} \in D, \tag{5.8}$$

em que $\mu(\cdot)$ é uma função média e $\varepsilon(\cdot)$ é o erro do modelo. Esses erros têm média zero e variância comum σ^2, o que significa que $\text{E}(Y(\boldsymbol{s})) = \mu(s)$ e $\text{Var}(Y(\boldsymbol{s})) = \sigma^2$, para todo $\boldsymbol{s} \in D$.

O processo espacial $\{Y(\boldsymbol{s}), \boldsymbol{s} \in D\}$ é assumido estacionário, então, sua média é constante, isto é, $\mu(\boldsymbol{s}) = \mu$, para todo $\boldsymbol{s} \in D$. Além disso, o processo é assumido isotrópico e

sua função covariância depende somente da distância entre as localizações espaciais, isto é,

$$\mathrm{Cov}(Y(\boldsymbol{s_i}), Y(\boldsymbol{s_j})) = \mathrm{C}(\boldsymbol{s_i}, \boldsymbol{s_j}) = \mathrm{C}(h_{ij}), \quad \boldsymbol{s_i}, \boldsymbol{s_j} \in D, \tag{5.9}$$

em que $h_{ij} = \|\boldsymbol{s_i} - \boldsymbol{s_j}\|$ é a distância Euclidiana entre $\boldsymbol{s_i}$ e $\boldsymbol{s_j}$.

5.1.2.4 Covariância e modelos de variogramas

Agora suponha que n medições são coletadas em um conjunto de localizações espaciais conhecidas $\{\boldsymbol{s_1}, \ldots, \boldsymbol{s_n}\}$ e forneça o vetor aleatório n-variado $\boldsymbol{Y} = (Y_1, \ldots, Y_n)^\top$, em que $Y_i = Y(\boldsymbol{s_i})$, para $i = 1, \ldots, n$. A covariância entre todos os pares de variáveis aleatórias (Y_i, Y_j) é determinada por uma matriz escala, $n \times n$, $\boldsymbol{\Sigma} = [\sigma_{ij}]$, simétrica e positiva definida, em que $\sigma_{ij} = \mathrm{C}(h_{ij})$, com $\mathrm{C}(\cdot)$ dada em (5.9). Então, $\boldsymbol{\Sigma}$ define uma estrutura de dependência espacial para um processo estacionário e isotrópico por um vetor de parâmetros $\boldsymbol{\varphi} = (\varphi_1, \varphi_2, \varphi_3)^\top$, em que $\varphi_1 \geq 0, \varphi_2 \geq 0$ são parâmetros conhecidos como efeito pepita e contribuição, respectivamente. O parâmetro $\varphi_3 \geq 0$ está relacionado ao raio de dependência espacial $a = g(\varphi_3)$ (MARDIA; MARSHAL, 1984; URIBE-OPAZO; BORSSOI; GALEA, 2012). O efeito pepita está relacionado a um erro analítico, indicando uma variabilidade não explicável. Esta variabilidade pode ser atribuída a erros de medida ou a uma variabilidade não capturada devido à distância de amostragem utilizada (CAMBARDELLA et al., 1994). O efeito pepita pode servir também como um mecanismo de regularização em delineamentos espaciais para campos aleatórios (MÜLLER; STEHLÍK, 2010) para o caso Gaussiano. Assume-se uma forma paramétrica particular para a matriz escala dada por

$$\boldsymbol{\Sigma} = \varphi_1 \boldsymbol{I}_n + \varphi_2 \boldsymbol{R}, \tag{5.10}$$

em que $\boldsymbol{I}_n$ é a matriz identidade $n \times n$ e $\boldsymbol{R} = [r_{ij}]$ é uma matriz simétrica $n \times n$ com elementos da diagonal $r_{ii} = 1$, para $i = 1, \ldots, n$. Formas específicas para r_{ij} dadas por $r_{ij} = \sigma_{ij}/\varphi_2$, com $i \neq j$ e $\varphi_2 \neq 0$, definem o modelo usado para explicar a dependência espacial, sendo as formas mais comuns obtidas das famílias Matérn e exponencial potência (ISAAKS; SRIVASTAVA, 1989; DIGGLE; RIBEIRO JÚNIOR, 2007).

Na família de modelos Matérn, tem-se

$$r_{ij} = \begin{cases} 1, & i = j, \\ \frac{1}{2^{\delta-1}\Gamma(\delta)} \left(\frac{h_{ij}}{\varphi_3}\right)^{\delta} K_\delta \left(\frac{h_{ij}}{\varphi_3}\right), & i \neq j, \end{cases} \tag{5.11}$$

em que δ é um parâmetro de forma, $\Gamma(\cdot)$ é a função gamma usual e $K_\delta(\cdot)$ é a função de Bessel modificada do terceiro tipo de order δ. A partir de (5.11), tem-se

$$\sigma_{ij} = \begin{cases} \varphi_1 + \varphi_2, & i = j, \\ \frac{\varphi_2}{2^{\delta-1}\Gamma(\delta)} \left(\frac{h_{ij}}{\varphi_3}\right)^{\delta} K_\delta \left(\frac{h_{ij}}{\varphi_3}\right), & i \neq j. \end{cases} \tag{5.12}$$

Nos modelos da família exponêncial potência, para $i \neq j$, tem-se $r_{ij} = \exp(-(h_{ij}/\varphi_3)^p)$,

em que $0 < p \leq 2$ é um parâmetro de forma, o que implica em

$$\sigma_{ij} = \begin{cases} \varphi_1 + \varphi_2, & i = j, \\ \varphi_2 \exp\left(-\left(\frac{h_{ij}}{\varphi_3}\right)^p\right), & i \neq j. \end{cases} \tag{5.13}$$

Embora os modelos dados em (5.12) e (5.13) não tenham um alcance finito, o alcance a é definido como 95% do patamar. Os modelos exponencial e Gaussiano são membros da família exponencial potência quando $p = 1$ e $p = 2$, respectivamente, e também da família Matérn quando $\delta = 0,5$ e $\delta \to \infty$, respectivamente. Em processos estacionários, da função covariância dada em (5.9), é possível definir a função semivariância por

$$\gamma(h) = C(0) - C(h), \quad h > 0, \tag{5.14}$$

em que $C(0) = \varphi_1 + \varphi_2$ e $C(h)$ é especificada a partir de $\mathrm{C}(h_{ij}) = \sigma_{ij}$, para $h = h_{ij}$, com um membro apropriado da família Matérn ou exponencial potência dadas em (5.12) e (5.13), respectivamente. O gráfico de pontos $(h, \gamma(h))$ obtido da função semivariância dada em (5.14) é uma ferramenta útil em estatística espacial.

5.1.2.5 Interpolação Krigagem

Em análise geoestatística, um método frequentemente usado para interpolação é a Krigagem (KRIGE, 1951). A predição Krigagem é dada por uma combinação linear dos dados observados $\boldsymbol{y} = (y_1, \ldots, y_n)^\top$ definida como

$$\widehat{y}(\boldsymbol{s_0}) = \sum_{i=1}^{n} \lambda_i \, y_i, \tag{5.15}$$

em que $\widehat{y}(\boldsymbol{s_0})$ é o valor predito em uma nova localização $\boldsymbol{s_0}$ e $\lambda_1, \ldots, \lambda_n$ são pesos escolhidos para definir o melhor preditor linear não enviesado, que pode ser obtido quando se minimiza a variância do erro quanto aos pesos. É possível mostrar que, sobre suposições de normalidade, os valores de $\lambda_1, \ldots, \lambda_n$ são dados pela solução do sistema $\boldsymbol{C\lambda} = \boldsymbol{\Upsilon}$, em que

$$\boldsymbol{C} = \begin{pmatrix} \mathrm{C}(\boldsymbol{s_1}, \boldsymbol{s_1}) & \ldots & \mathrm{C}(\boldsymbol{s_1}, \boldsymbol{s_n}) & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \mathrm{C}(\boldsymbol{s_n}, \boldsymbol{s_1}) & \ldots & \mathrm{C}(\boldsymbol{s_n}, \boldsymbol{s_n}) & 1 \\ 1 & \ldots & 1 & 0 \end{pmatrix}; \boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n, \varrho)^\top; \boldsymbol{\Upsilon} = (\mathrm{C}(\boldsymbol{s_1}, \boldsymbol{s_0}), \ldots, \mathrm{C}(\boldsymbol{s_n}, \boldsymbol{s_0}), 1)^\top;$$

e ϱ é um multiplicador de Lagrange introduzido durante a minimização da variância dos erros para garantir o não enviesamento dos mesmos.

5.1.2.6 Medidas de acurácia

Os índices acurácia global (AG) e kappa (κ) podem ser usados para quantificar a similaridade entre dois mapas. Considere dois mapas (um chamado mapa de referência e outro chamado mapa modelo), ambos particionados nas mesmas m classes, denotadas por M_i, para

$i = 1, \ldots, m$. Adicionalmente, seja N_{ij}, para $i, j = 1, \ldots, m$, o número de pixels pertecentes à classe M_i do mapa modelo e à classe M_j do mapa de referência, e N é o número total de pixels em cada mapa. O índice AG é baseado nos pixels que pertencem às mesmas classes em ambos os mapas e é definido como $\mathrm{AG} = (\sum_{i=1}^{m} N_{ii})/N$. Os mapas modelo e de referência têm uma similaridade aceitável se o índice AG é maior que 0,85 (ANDERSON, 1976). O índice κ é baseado em todos os pixels (os que pertencem às mesmas classes e os que pertencem às classes diferentes) e é definido como $\kappa = (N \sum_{i=1}^{m} N_{ii} - \sum_{i=1}^{m} N_{\bullet i} N_{i\bullet})/(N^2 - \sum_{i=1}^{m} N_{\bullet i} N_{i\bullet})$, em que $N_{i\bullet} = \sum_{j=1}^{m} N_{ij}$ e $N_{\bullet i} = \sum_{j=1}^{m} N_{ji}$. Os mapas modelo e de referência possuem baixa similaridade se $\kappa < 0,67$, média similaridade se $0,67 \leq \kappa \leq 0,80$, e alta similaridade se $\kappa > 0,80$ (KRIPPENDORFF, 2004).

5.1.3 O modelo espacial Birnbaum-Saunders

5.1.3.1 Formulação do modelo

Seja $\{T(\boldsymbol{s}), \boldsymbol{s} \in D\}$ um processo estocástico definido sobre uma região D, com $D \subset \mathbb{R}^2$. Suponha que n medidas $\boldsymbol{T} = (T_1, \ldots, T_n)^\top$, em que $T_i = T(\boldsymbol{s_i})$, para $i = 1, \ldots, n$, são coletadas em um conjunto de localizações espaciais conhecidas $\{\boldsymbol{s_1}, \ldots, \boldsymbol{s_n}\}$. Considere um modelo espacial da forma

$$T_i = \exp(\mu_i)\, \eta_i, \quad i = 1, \ldots, n. \tag{5.16}$$

Assume-se estacionaridade, de modo que $\mu_i = \mu(\boldsymbol{s_i}) = \mu$, e o erro do modelo $\eta_i = \eta(\boldsymbol{s_i}) \sim \mathrm{BS}(\alpha, 1)$, para $i = 1, \ldots, n$. Então, $\exp(\mu)$ é a mediana do modelo. Note que o parâmetro de forma α é assumido constante para todas as localizações espaciais. Ao se aplicar a transformação logarítmica em (5.16),obtém-se um modelo BS espacial log-linear dado por

$$Y_i = \log(T_i) = \mu + \log(\eta_i) = \mu + \varepsilon_i, \quad i = 1, \ldots, n, \tag{5.17}$$

em que $\varepsilon_i = \log(\eta_i) \sim \text{log-BS}(\alpha, 0)$, para $i = 1, \ldots, n$. Note que o modelo BS espacial log-linear definido em (5.17) tem uma forma similar a do modelo dado em (5.8). Para facilitar a notação, o modelo dado em (5.17) pode ser escrito em forma matricial como

$$\boldsymbol{Y} = \mu \mathbf{1} + \boldsymbol{\varepsilon}, \tag{5.18}$$

com $\boldsymbol{Y} = (Y_1, \ldots, Y_n)^\top$, $\mathbf{1} = (1, \ldots, 1)^\top$ and $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_n)^\top$ sendo vetores $n \times 1$. Neste caso, $\boldsymbol{\varepsilon}$ é um processo estocástico espacial estacionário com vetor de médias $\mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}$. Suponha que a covariância entre todos os pares (Y_i, Y_j) é determinada pela matriz escala $\boldsymbol{\Sigma}$, $n \times n$, satisfazendo as condições dadas em (5.10), em que os elementos de $\boldsymbol{R}$ podem ser modelados com a estrutura Matérn dada em (5.11).

5.1.3.2 Estimação dos parâmetros

Seja $\boldsymbol{\theta} = (\alpha, \mu, \varphi_1, \varphi_2, \varphi_3)^\top$ o vetor de parâmetros desconhecidos do modelo espacial formulado em (5.18) para ser estimado. A função verossimilhança para $\boldsymbol{\theta}$, baseada nas

observações $\boldsymbol{y} = (y_1, \ldots, y_n)^\top$ de $\boldsymbol{Y}$, obtida de (5.7) é dada por

$$\mathcal{L}(\boldsymbol{\theta}) = \mathcal{L}(\boldsymbol{\theta}; \boldsymbol{y}) = \frac{1}{\alpha^n (2\pi)^{\frac{n}{2}} |\boldsymbol{\Sigma}|^{\frac{1}{2}}} \exp\left(-\frac{2}{\alpha^2}\boldsymbol{V}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{V}\right) \prod_{i=1}^{n} \cosh\left(\frac{y_i - \mu}{2}\right), \tag{5.19}$$

em que $\boldsymbol{V} = (V_1, \ldots, V_n)^\top$ é um vetor $n \times 1$ com elementos $V_i = \operatorname{senh}((y_i - \mu)/2)$, para $i = 1, \ldots, n$. O correspondente logaritmo da função verossimilhança para $\boldsymbol{\theta}$ obtida de (5.19) é então

$$\ell(\boldsymbol{\theta}) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\log(|\boldsymbol{\Sigma}|) - n\log(\alpha) - \frac{2}{\alpha^2}\boldsymbol{V}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{V} + \sum_{i=1}^{n} \log\left(\cosh\left(\frac{y_i - \mu}{2}\right)\right). \tag{5.20}$$

O método ML define o estimador $\widehat{\boldsymbol{\theta}}$ de $\boldsymbol{\theta}$ como o vetor que maximiza $\mathcal{L}(\boldsymbol{\theta})$, ou equivalentemente $\ell(\boldsymbol{\theta})$, sobre o espaço paramétrico de $\boldsymbol{\theta}$. Então,

$$\widehat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} \ell(\boldsymbol{\theta}). \tag{5.21}$$

Quando o valor em (5.21) está associado a um processo estacionário, ele pode ser obtido da solução do sistema homogêneo de equações formuladas a partir do vetor escore é dado por

$$\frac{\partial \ell(\boldsymbol{\theta})}{\partial \alpha} = 0, \frac{\partial \ell(\boldsymbol{\theta})}{\partial \mu} = 0, \frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\varphi}} = \boldsymbol{0}; \tag{5.22}$$

Veja detalhes do vetor escore no Apêndice I. Note que não existe solução analítica para o sistema de equações dados em (5.22). Então, o estimador ML $\widehat{\boldsymbol{\theta}}$ deve ser calculado utilizando um processo iterativo para solução de sistemas não lineares. Neste trabalho, o processo quasi-Newton Broyden-Fletcher-Goldfarb-Shanno (BFGS) (NOCEDAL; WRIGHT, 1999; LANGE, 2000) será utilizado pelas funções `optim` e `optimx` implementadas no *software* R (R TEAM, 2015). Os sinais do determinante da matriz Hessiana e de seus menores principais foram avaliados para assegurar que o máximo tinha sido encontrado.

Sobre condições usuais de regularidade, o estimador ML $\widehat{\boldsymbol{\theta}}$ é consistente para $\boldsymbol{\theta}$ e tem uma distribuição normal assintótica (DAGENAIS; DUFOUR, 1991). Então, quando $n \to \infty$,

$$\sqrt{n}\,(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{\mathrm{D}} \mathsf{N}_5(\boldsymbol{0}, \boldsymbol{J}(\boldsymbol{\theta})^{-1}), \tag{5.23}$$

em que $\boldsymbol{J}(\boldsymbol{\theta}) = \lim_{n\to\infty} \boldsymbol{I}(\boldsymbol{\theta})/n$, com $\boldsymbol{I}(\boldsymbol{\theta})$, sendo a matriz de informação esperada de Fisher dado no Apêndice III, e $\xrightarrow{\mathrm{D}}$ indicando convergência em distribuição. Intervalos de confiança assintóticos (ICs), de $100 \times [1-\zeta]\%$, para μ, α e φ_i, com $i = 1, 2, 3$, podem ser obtidos da propriedade de normalidade assintótica dada em (5.23) como

$$\begin{aligned}
\mathrm{CI}(\mu, [1-\zeta]\times 100\%) &= \left[\widehat{\mu} - z(1-\zeta/2)\widehat{\mathrm{SE}}(\widehat{\mu}), \widehat{\mu} + z(1-\zeta/2)\widehat{\mathrm{SE}}(\widehat{\mu})\right], \\
\mathrm{CI}(\theta, [1-\zeta]\times 100\%) &= \left[\exp\left(\widehat{\theta^*} - z(1-\zeta/2)\widehat{\mathrm{SE}}(\widehat{\theta^*})\right), \exp\left(\widehat{\theta^*} + z(1-\zeta/2)\widehat{\mathrm{SE}}(\widehat{\theta^*})\right)\right],
\end{aligned}$$

em que $\widehat{\theta^*}$ é o estimador ML de $\theta^* = \log(\theta)$, com $\theta = \alpha$ ou $\theta = \varphi_i$, para $i = 1, 2, 3$, e $\widehat{\mathrm{SE}}(\widehat{\theta^*})$ é o erro padrão estimado assintótico (SE) do estimador ML de θ^* (LEIVA et al., 2015d).

5.1.3.3 Influência local

Seja $\ell(\boldsymbol{\theta})$ o logaritmo da função verossimilhança para o vetor de parâmetros do modelo $\boldsymbol{\theta} = (\alpha, \mu, \varphi_1, \varphi_2, \varphi_3)^\top$ dado em (5.18), a qual é conhecida como log-verossimilhança não perturbada. Então, seja $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_n)^\top \in \Omega \subset \mathbb{R}^n$ um vetor perturbação, $n \times 1$, em que Ω é um conjunto aberto de perturbações. Seja $\ell(\boldsymbol{\theta}|\boldsymbol{\omega})$ o logaritmo da função verossimilhança perturbada por $\boldsymbol{\omega}$, chamada de log-verossimilhança perturbada, e $\widehat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}$ a estimativa ML de $\boldsymbol{\theta}$ obtida de $\ell(\boldsymbol{\theta}|\boldsymbol{\omega})$. Adicionalmente, seja $\boldsymbol{\omega}_0 \in \Omega$ um vetor de não perturbação, $n \times 1$, isto é, $\ell(\boldsymbol{\theta}|\boldsymbol{\omega}_0) = \ell(\boldsymbol{\theta})$. Suponha que $\ell(\boldsymbol{\theta}|\boldsymbol{\omega})$ seja contínua e duas vezes diferenciável em uma vizinhança de $(\widehat{\boldsymbol{\theta}}, \boldsymbol{\omega}_0)$. Comparando-se as estimativas ML dos parâmetros $\widehat{\boldsymbol{\theta}}$ e $\widehat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}$ a partir de influência local, é possivel investigar como tais estimativas são afetadas pela perturbação. Considere o afastamento da verossimilhança dado por

$$\mathrm{LD}(\boldsymbol{\omega}) = 2(\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widehat{\boldsymbol{\theta}}_{\boldsymbol{\omega}})), \tag{5.24}$$

que é utilizada para avaliar a influência da perturbação $\boldsymbol{\omega}$. Valores grandes de $\mathrm{LD}(\boldsymbol{\omega})$ em (5.24) indicam que $\widehat{\boldsymbol{\theta}}$ e $\widehat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}$ diferem consideravelmente entre si. O método estuda o comportamento local do gráfico de influência $a(\boldsymbol{\omega}) = (\boldsymbol{\omega}^\top, \mathrm{LD}(\boldsymbol{\omega}))^\top$ em torno de $\boldsymbol{\omega}_0$. Cook (1986) sugere investigar a direção de curvatura máxima, $C_{\max}$, da superfície $a(\boldsymbol{\omega})$. Para $\mathrm{LD}(\boldsymbol{\omega})$ dada em (5.24), $C_{\max} = \max_{||\boldsymbol{d}||=1} C_{\boldsymbol{d}}$, em que $C_{\boldsymbol{d}} = 2|\boldsymbol{d}^\top \boldsymbol{B} \boldsymbol{d}|$, com $\boldsymbol{B}$ sendo uma matriz $n \times n$ e $\boldsymbol{d}$ um vetor unitário. Para encontrar $C_{\max}$ e o vetor direção $\boldsymbol{d}_{\max}$, é necessário calcular $\boldsymbol{B} = -\boldsymbol{\Delta}\ddot{\boldsymbol{\ell}}(\widehat{\boldsymbol{\theta}})^{-1}\boldsymbol{\Delta}$, em que $\ddot{\boldsymbol{\ell}}(\widehat{\boldsymbol{\theta}})$ é a matriz Hessiana obtida de (5.19) avaliada em $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$; veja detalhes da matriz Hessiana no Apêndice II. Neste caso, $\boldsymbol{\Delta}$ é uma matriz $5 \times n$ obtida do logaritmo da função verossimilhança perturbada e dada por

$$\boldsymbol{\Delta} = \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\omega}^\top} = \begin{pmatrix} \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \alpha \partial \boldsymbol{\omega}^\top} \\ \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \mu \partial \boldsymbol{\omega}^\top} \\ \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \boldsymbol{\varphi} \partial \boldsymbol{\omega}^\top} \end{pmatrix} = \begin{pmatrix} \Delta_\alpha \\ \Delta_\mu \\ \boldsymbol{\Delta}_{\boldsymbol{\varphi}} \end{pmatrix}. \tag{5.25}$$

Note que (5.25) deve ser avaliada em $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ e $\boldsymbol{\omega} = \boldsymbol{\omega}_0$. Então, $\boldsymbol{d}_{\max}$ é um autovetor unitário associado ao máximo autovalor (em valor absoluto) $C_{\max}$ de $\boldsymbol{B}$. Um valor absoluto grande de um elemento de $\boldsymbol{d}_{\max}$ revela que o caso é possivelmente influente. Outras direções importantes correspondem aos vetores da base canônica $\boldsymbol{d} = \boldsymbol{e}_i$, para $i = 1, \ldots, n$, em que $\boldsymbol{e}_i$ é um vetor do $\mathbb{R}^n$ com um (1) na i-ésima posição e zeros (0) nas demais posições. Neste caso, a curvatura é dada por $C_i = 2|b_{ii}|$, onde b_{ii} é o (i,i) elemento da matriz $\boldsymbol{B}$, para $i = 1, \ldots, n$. O gráfico de $C_i = C_{\boldsymbol{d}_i}$ *versus* o índice i pode também ser usado para identificar casos influentes. Os gráficos de C_i e $|d_{\max_i}|$ são usados como medidas de diagnósticos de influência local na Seção 5.1.5. Embora não exista um consenso sobre um ponto de referência para determinar casos influentes, neste trabalho usa-se um valor análogo ao proposto por Zhu e Lee (2001), que considere o caso i como influente se $C_i > \overline{C} + 2\mathrm{SE}(C)$, para $i = 1, \ldots, n$, em que $\overline{C}$ e $\mathrm{SE}(C)$ denotam, respectivamente, a média das curvaturas normais e o correspondente desvio padrão amostral. Similarmente, no gráfico $|d_{\max_i}|$, o caso i é indicado como influente

se $|d_{\max_i}| > \overline{|\boldsymbol{d}_{\max}|} + 2\mathrm{SE}(|\boldsymbol{d}_{\max}|)$, para $i = 1, \ldots, n$, em que $\overline{|\boldsymbol{d}_{\max}|}$ e $\mathrm{SE}(|\boldsymbol{d}_{\max}|)$ denotam, respectivamente, a média dos elementos, em valor absoluto, do vetor $\boldsymbol{d}_{\max}$ e o correspondente desvio padrão amostral. Os valores de referência são utilizados no estudo de simulação e na aplicação realizados nas Seções 5.1.4 e 5.1.5, respectivamente.

5.1.3.4 Seleção do esquema de perturbação apropriado

Suponha que a perturbação na variável resposta é da forma $\boldsymbol{Y}_{\boldsymbol{\omega}}(\boldsymbol{s}) = \boldsymbol{Y}(\boldsymbol{s}) + \boldsymbol{A}\boldsymbol{\omega}$, em que $\boldsymbol{A}$ é uma matriz simétrica e não singular. Assim, $\boldsymbol{Y}_{\boldsymbol{\omega_0}}(\boldsymbol{s}) = \boldsymbol{Y}(\boldsymbol{s})$ e $Y_{\boldsymbol{\omega}}(\boldsymbol{s_i}) = Y(\boldsymbol{s_i}) + \boldsymbol{a}_i\,\boldsymbol{\omega}$, em que $\boldsymbol{a}_i$ é a $i-$ésima linha da matriz $\boldsymbol{A}$. Neste caso, o logaritmo da função verossimilhança perturbada é dada por

$$\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\log\left(|\boldsymbol{\Sigma}|\right) - n\log(\alpha) - \frac{2}{\alpha^2}\boldsymbol{V}_{\boldsymbol{\omega}}^{\top}\boldsymbol{\Sigma}^{-1}\boldsymbol{V}_{\boldsymbol{\omega}} + \sum_{i=1}^{n}\log\left(\cosh\left(\frac{y_i + \boldsymbol{a}_i\boldsymbol{\omega} - \mu}{2}\right)\right), \tag{5.26}$$

em que $\boldsymbol{V}_{\boldsymbol{\omega}} = (V_{\omega_1}, \ldots, V_{\omega_n})^{\top}$, com $V_{\omega_i} = \mathrm{senh}((y_i + \boldsymbol{a}_i\boldsymbol{\omega} - \mu)/2)$, para $i = 1, \ldots, n$. Então, o vetor escore correspondente obtido de (5.26) é dado por (veja detalhes no Apêndice IV)

$$\boldsymbol{U}(\boldsymbol{\omega}) = \frac{\partial\ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial\boldsymbol{\omega}} = -\frac{2}{\alpha^2}\boldsymbol{A}\boldsymbol{\Sigma}^{-1}\boldsymbol{V}_{\boldsymbol{\omega}} + \frac{1}{2}\boldsymbol{A}\boldsymbol{V}_{\boldsymbol{\omega}}. \tag{5.27}$$

Considere a variância do vetor escore dado em (5.27) como uma função do vetor perturbação $\boldsymbol{\omega}$, isto é, $\boldsymbol{G}(\boldsymbol{\omega}) = \mathrm{Var}\left(\boldsymbol{U}(\boldsymbol{\omega})\right) = \mathrm{E}\left(\boldsymbol{U}(\boldsymbol{\omega})\boldsymbol{U}^{\top}(\boldsymbol{\omega})\right)$, lembrando que $\mathrm{E}\left(\boldsymbol{U}(\boldsymbol{\omega})\right) = \boldsymbol{0}$. Para o modelo BS espacial log-linear, tem-se que (novamente veja detalhes no Apêndice IV)

$$\boldsymbol{G}(\boldsymbol{\omega}) = \boldsymbol{A}\left(\frac{1}{\alpha}\boldsymbol{\Sigma}^{-\frac{1}{2}} - \frac{\alpha}{4}\boldsymbol{\Sigma}^{\frac{1}{2}}\right)^2\boldsymbol{A}. \tag{5.28}$$

De acordo com Zhu et al. (2007), a perturbação $\boldsymbol{\omega}$ é apropriada se, e somente se $\boldsymbol{G}(\boldsymbol{\omega_0}) = c\,\boldsymbol{I}_n$, para $c > 0$, com $\boldsymbol{G}(\cdot)$ dada em (5.28). De modo geral, para uma matriz simétrica e não singular arbitrária $\boldsymbol{A}$, $\boldsymbol{G}(\boldsymbol{\omega}) \neq c\,\boldsymbol{I}_n$. No entanto, para a perturbação ser apropriada, $\boldsymbol{A}$ deve satisfazer a condição

$$\boldsymbol{A}\left(\frac{1}{\alpha}\boldsymbol{\Sigma}^{-\frac{1}{2}} - \frac{\alpha}{4}\boldsymbol{\Sigma}^{\frac{1}{2}}\right)^2\boldsymbol{A} = c\,\boldsymbol{I}_n, \tag{5.29}$$

para algum valor de $c > 0$. Considerando $c = 1$, imediatamente tem-se que $\boldsymbol{A} = ((1/\alpha)\boldsymbol{\Sigma}^{-\frac{1}{2}} - (\alpha/4)\boldsymbol{\Sigma}^{\frac{1}{2}})^{-1}$ satisfaz a condição dada em (5.29). Então, para o modelo BS espacial log-linear, um esquema de perturbação apropriado para a variável-resposta é dado por

$$\boldsymbol{Y}_{\boldsymbol{\omega}}(\boldsymbol{s}) = \boldsymbol{Y}(\boldsymbol{s}) + \left(\frac{1}{\alpha}\boldsymbol{\Sigma}^{-\frac{1}{2}} - \frac{\alpha}{4}\boldsymbol{\Sigma}^{\frac{1}{2}}\right)^{-1}\boldsymbol{\omega}. \tag{5.30}$$

Agora, considere a matriz de perturbação $\boldsymbol{\Delta}$ definida em (5.25), a perturbação dada em (5.30), e

$$\frac{\partial\ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial\boldsymbol{\omega}^{\top}} = -\frac{2}{\alpha^2}\boldsymbol{V}_{\boldsymbol{\omega}}^{\top}\boldsymbol{\Sigma}^{-1}\boldsymbol{A} + \frac{1}{2}\boldsymbol{V}_{\boldsymbol{\omega}}^{\top}\boldsymbol{A} = \boldsymbol{V}_{\boldsymbol{\omega}}^{\top}\left(-\frac{2}{\alpha^2}\boldsymbol{\Sigma}^{-1} + \frac{1}{2}\boldsymbol{I_n}\right)\boldsymbol{A}. \tag{5.31}$$

Então, os elementos de $\boldsymbol{\Delta}$ obtidos de (5.31) são expressos como

$$\Delta_{\mu} = \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \mu \partial \boldsymbol{\omega}^{\top}} = \mathbf{1}^{\top}\left(\frac{1}{\alpha^2}\boldsymbol{\Sigma}^{-1} - \frac{1}{4}\boldsymbol{I_n}\right)\boldsymbol{A}, \tag{5.32}$$

$$\begin{aligned}\Delta_{\alpha} &= \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \alpha \partial \boldsymbol{\omega}^{\top}} = \boldsymbol{D}^{\top}\left(\frac{-1}{\alpha^2}\boldsymbol{\Sigma}^{-1} + \frac{1}{4}\boldsymbol{I_n}\right)\boldsymbol{A}, \\ &\quad + \boldsymbol{V}_{\boldsymbol{\omega}}^{\top}\left(\frac{4}{\alpha^3}\boldsymbol{\Sigma}^{-1}\boldsymbol{A} + \left(-\frac{2}{\alpha^2}\boldsymbol{\Sigma}^{-1} + \frac{1}{2}\boldsymbol{I_n}\right)\boldsymbol{A}\left(\frac{1}{\alpha^2}\boldsymbol{\Sigma}^{-\frac{1}{2}} + \frac{1}{4}\boldsymbol{\Sigma}^{\frac{1}{2}}\right)\boldsymbol{A}\right),\end{aligned} \tag{5.33}$$

$$\boldsymbol{\Delta}_{\boldsymbol{\varphi}} = \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial^2 \boldsymbol{\varphi} \partial \boldsymbol{\omega}^{\top}} = \left(\frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \varphi_1 \partial \boldsymbol{\omega}^{\top}}, \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \varphi_2 \partial \boldsymbol{\omega}^{\top}}, \frac{\partial \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \varphi_3 \partial \boldsymbol{\omega}^{\top}}\right)^{\top}, \tag{5.34}$$

em que $\boldsymbol{D} = (d_1, \ldots, d_n)^{\top}$, com $d_i = \boldsymbol{l}_i\,\boldsymbol{\omega}$, para $i = 1, \ldots, n$, e $\boldsymbol{l}_i$ sendo a $i-$ésima linha da matriz $\boldsymbol{A}((1/\alpha^2)\boldsymbol{\Sigma}^{-\frac{1}{2}} + (1/4)\boldsymbol{\Sigma}^{\frac{1}{2}})\boldsymbol{A}$ e os elementos de $\boldsymbol{\Delta}_{\boldsymbol{\varphi}}$ dados em (5.34) são definidos como

$$\begin{aligned}\frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \varphi_i \partial \boldsymbol{\omega}^{\top}} &= \boldsymbol{M}^{\top}\left(-\frac{1}{\alpha^2}\boldsymbol{\Sigma}^{-1} + \frac{1}{4}\boldsymbol{I}_n\right)\boldsymbol{A} + \boldsymbol{V}_{\boldsymbol{\omega}}^{\top}\left(\frac{2}{\alpha^2}\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_i}\boldsymbol{\Sigma}^{-1}\right)\boldsymbol{A} \\ &\quad + \mathbf{1}^{\top}\left(-\frac{2}{\alpha^2}\boldsymbol{\Sigma}^{-1} + \frac{1}{2}\boldsymbol{I}_n\right)\boldsymbol{A}\left(\frac{1}{\alpha}\boldsymbol{\Sigma}^{-\frac{1}{2}}\frac{\partial \boldsymbol{\Sigma}^{\frac{1}{2}}}{\partial \varphi_i}\boldsymbol{\Sigma}^{-\frac{1}{2}} + \frac{\alpha}{4}\frac{\partial \boldsymbol{\Sigma}^{\frac{1}{2}}}{\partial \varphi_i}\right)\boldsymbol{A},\end{aligned}$$

em que $\boldsymbol{M} = (m_1, \ldots, m_n)^{\top}$, com $m_i = \boldsymbol{L}_i\,\boldsymbol{\omega}$, para $i = 1, \ldots, n$, e $\boldsymbol{L}_i$ sendo a $i-$ésima linha da matriz

$$\boldsymbol{A}\left(-\frac{1}{\alpha}\boldsymbol{\Sigma}^{-\frac{1}{2}}\frac{\partial \boldsymbol{\Sigma}^{\frac{1}{2}}}{\partial \varphi_i}\boldsymbol{\Sigma}^{-\frac{1}{2}} + \frac{\alpha}{4}\frac{\partial \boldsymbol{\Sigma}^{\frac{1}{2}}}{\partial \varphi_i}\right)\boldsymbol{A}.$$

A álgebra detalhada das expressões anteriores é apresentada no Apêndice V.

5.1.4 Estudos de simulação

Nesta Seção, foram conduzidos dois estudos de simulação, baseados em n pontos amostrais $\boldsymbol{y} = (y_1, \ldots, y_n)^{\top}$, gerados a partir de uma distribuição log-BS com parâmetro de forma $\alpha \in \{0,3; 1\}$, parâmetro de escala $\mu = 2$ e estrutura de dependência espacial descrita pelo modelo família Matérn para $\delta = 0,5$, $\varphi_1 \in \{0,5; 1\}$, $\varphi_2 \in \{0,5; 1\}$ e $\varphi_3 \in \{1; 1,5\}$. Assim, foi considerada uma grade regular com distância mínima entre pontos de uma unidade. Primeiro, avaliou-se o desempenho dos estimadores dos parâmetros do modelo para amostras de tamanhos pequeno e grande. Em seguida, avaliou-se o desempenho da metodologia proposta para detectar casos influentes.

5.1.4.1 Estudo I: Estimação ML

Para avaliar a eficiência do estimador $\widehat{\theta}_i$ empiricamente, usam-se o viés relativo absoluto (VRA) e a raiz do erro quadrático médio (REQM), definidos como seguem

$$\mathrm{VRA}(\widehat{\theta}_i) = \left|\frac{\bar{\theta}_i - \theta_i}{\theta_i}\right| \times 100, \quad \mathrm{REQM}(\widehat{\theta}_i) = \frac{1}{p}\sum_{j=1}^{p}(\widehat{\theta}_{ij} - \theta_i)^2,$$

em que $\bar{\theta}_i = (1/p)\sum_{j=1}^{p}\widehat{\theta}_{ij}$, com $\widehat{\theta}_{ij}$ sendo a estimativa ML de $\theta_i \in \{\alpha, \mu, \varphi_1, \varphi_2, \varphi_3\}$ na $j-$ésima réplica de Monte Carlo (MC). Note que $\widehat{\boldsymbol{\theta}} = \{\widehat{\alpha}, \widehat{\mu}, \widehat{\varphi}_1, \widehat{\varphi}_2, \widehat{\varphi}_3\}^{\top}$ é a estimativa ML de $\boldsymbol{\theta} = \{\alpha, \mu, \varphi_1, \varphi_2, \varphi_3\}^{\top}$. Considera-se $p = 500$ réplicas em cada caso. A Tabela 3 exibe os VRAs

e REQMs para os correspondentes estimadores ML. Note que, para $n = 100$, todos os VRAs são pequenos. Em particular, os VRAs dos estimadores ML do parâmetro α são menores que 10% em quinze dos dezesseis casos estudados, além disso em 50% desses casos, os VRAs são menores que 2%. Para os estimadores ML do parâmetro μ, os VRAs são menores que 1% em praticamente todos os casos. Para os estimadores ML dos parâmetros que descrevem a estrutura espacial, os VRAs são maiores que 10% em alguns casos, porém o máximo é 13%. No entanto, para $n = 36$, esse quadro é bem diferente, porque somente os VRAs dos estimadores ML do parâmetro μ permanecem baixos. Para os outros parâmetros, em vários cenários, os VRAs são elevados. Em geral, os REQMs são pequenos para os estimadores de α e μ, mas são moderados para os parâmetros espaciais. Como esperado, os valores aumentam para $n = 36$ em comparação com obtidos para $n = 100$. Tal resultado mostra a sensibilidade dos estimadores ML do modelo espacial BS para amostras pequenas.

Tabela 3 VRA (em $\%$) e REQM do estimador ML dos parâmetros e n indicados para dados simulados com o modelo espacial BS.

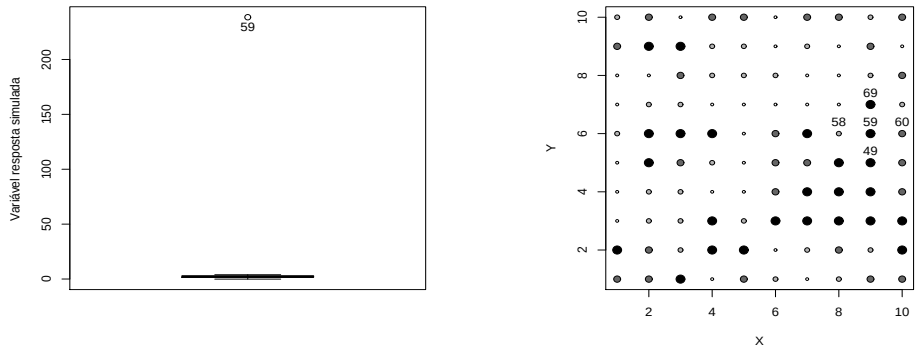
Parâmetro	Valor	VRA	REQM	Valor	VRA	REQM	Valor	VRA	REQM	Valor	VRA	REQM
	$n = 36$						$n = 100$					
α	$1,00$	$8,501$	$0,211$	$0,30$	$38,932$	$0,175$	$1,00$	$8,683$	$0,170$	$0,30$	$6,526$	$0,058$
μ	$2,00$	$0,377$	$0,425$	$2,00$	$1,010$	$0,339$	$2,00$	$0,48$	$0,245$	$2,00$	$0,102$	$0,070$
φ_1	$0,50$	$15,204$	$0,628$	$0,50$	$0,134$	$0,474$	$0,50$	$6,446$	$0,505$	$0,50$	$1,624$	$0,586$
φ_2	$1,00$	$2,879$	$0,549$	$1,00$	$12,064$	$0,546$	$1,00$	$3,702$	$0,492$	$1,00$	$9,129$	$0,200$
φ_3	$1,00$	$10,631$	$1,549$	$1,00$	$3,283$	$1,591$	$1,00$	$8,288$	$0,922$	$1,00$	$10,102$	$0,731$
α	$1,00$	$0,036$	$0,131$	$0,30$	$16,731$	$0,093$	$1,00$	$0,403$	$0,121$	$0,30$	$7,886$	$0,083$
μ	$2,00$	$0,782$	$0,172$	$2,00$	$0,310$	$0,076$	$2,00$	$0,298$	$0,310$	$2,00$	$0,124$	$0,051$
φ_1	$0,50$	$16,061$	$0,407$	$0,50$	$37,419$	$0,362$	$0,50$	$12,548$	$0,424$	$0,50$	$9,748$	$0,448$
φ_2	$0,50$	$20,495$	$0,357$	$0,50$	$9,631$	$0,285$	$0,50$	$4,074$	$0,579$	$0,50$	$8,521$	$0,418$
φ_3	$1,00$	$19,786$	$1,406$	$1,00$	$31,122$	$0,598$	$1,00$	$12,092$	$1,000$	$1,00$	$3,368$	$0,840$
α	$1,00$	$5,018$	$0,180$	$0,30$	$0,713$	$0,052$	$1,00$	$1,543$	$0,122$	$0,30$	$0,969$	$0,049$
μ	$2,00$	$0,569$	$0,196$	$2,00$	$0,128$	$0,088$	$2,00$	$0,094$	$0,308$	$2,00$	$0,056$	$0,054$
φ_1	$1,00$	$23,470$	$0,559$	$1,00$	$30,419$	$0,598$	$1,00$	$4,416$	$0,444$	$1,00$	$7,977$	$0,296$
φ_2	$0,50$	$69,708$	$0,812$	$0,50$	$72,644$	$1,077$	$0,50$	$3,011$	$0,712$	$0,50$	$12,575$	$0,594$
φ_3	$1,00$	$23,311$	$1,479$	$1,00$	$27,248$	$0,508$	$1,00$	$3,736$	$0,966$	$1,00$	$10,600$	$0,838$
α	$1,00$	$11,580$	$0,149$	$0,30$	$2,084$	$0,044$	$1,00$	$0,675$	$0,116$	$0,30$	$1,361$	$0,041$
μ	$2,00$	$1,073$	$0,347$	$2,00$	$0,044$	$0,095$	$2,00$	$0,261$	$0,219$	$2,00$	$0,143$	$0,065$
φ_1	$0,50$	$13,250$	$0,417$	$0,50$	$10,198$	$0,392$	$0,50$	$3,457$	$0,380$	$0,50$	$1,860$	$0,226$
φ_2	$0,50$	$47,626$	$0,420$	$0,50$	$7,230$	$0,409$	$0,50$	$9,557$	$0,600$	$0,50$	$2,624$	$0,265$
φ_3	$1,50$	$4,095$	$1,365$	$1,50$	$28,613$	$0,875$	$1,50$	$7,062$	$0,999$	$1,50$	$7,850$	$1,048$
α	$1,00$	$9,409$	$0,182$	$0,30$	$34,818$	$0,127$	$1,00$	$6,615$	$0,131$	$0,30$	$9,041$	$0,050$
μ	$2,00$	$0,832$	$0,424$	$2,00$	$0,135$	$0,112$	$2,00$	$0,351$	$0,262$	$2,00$	$0,141$	$0,071$
φ_1	$1,00$	$22,382$	$0,688$	$1,00$	$58,626$	$0,704$	$1,00$	$9,877$	$0,600$	$1,00$	$6,292$	$0,609$
φ_2	$1,00$	$16,801$	$0,698$	$1,00$	$32,248$	$0,494$	$1,00$	$11,614$	$0,649$	$1,00$	$10,743$	$0,399$
φ_3	$1,00$	$29,993$	$1,441$	$1,00$	$25,988$	$0,722$	$1,00$	$6,278$	$1,088$	$1,00$	$6,915$	$0,854$
α	$1,00$	$3,315$	$0,213$	$0,30$	$7,271$	$0,084$	$1,00$	$1,731$	$0,136$	$0,30$	$3,474$	$0,068$
μ	$2,00$	$1,291$	$0,464$	$2,00$	$0,208$	$0,131$	$2,00$	$1,147$	$0,312$	$2,00$	$0,029$	$0,088$
φ_1	$0,50$	$13,031$	$0,513$	$0,50$	$0,375$	$0,799$	$0,50$	$6,876$	$0,406$	$0,50$	$4,700$	$0,419$
φ_2	$1,00$	$8,044$	$0,385$	$1,00$	$11,701$	$0,383$	$1,00$	$1,374$	$0,477$	$1,00$	$7,779$	$0,324$
φ_3	$1,50$	$11,971$	$1,455$	$1,50$	$26,713$	$0,862$	$1,50$	$11,409$	$1,010$	$1,50$	$1,422$	$1,048$
α	$1,00$	$2,479$	$0,122$	$0,30$	$36,118$	$0,150$	$1,00$	$1,645$	$0,096$	$0,30$	$0,124$	$0,047$
μ	$2,00$	$1,012$	$0,254$	$2,00$	$0,016$	$0,110$	$2,00$	$0,310$	$0,339$	$2,00$	$0,101$	$0,067$
φ_1	$1,00$	$12,815$	$0,434$	$1,00$	$49,457$	$0,593$	$1,00$	$2,893$	$0,401$	$1,00$	$8,174$	$0,312$
φ_2	$0,50$	$44,539$	$0,610$	$0,50$	$51,302$	$0,840$	$0,50$	$9,810$	$0,648$	$0,50$	$7,033$	$0,827$
φ_3	$1,50$	$19,714$	$1,927$	$1,50$	$10,903$	$0,570$	$1,50$	$2,042$	$1,105$	$1,50$	$0,266$	$1,052$
α	$1,00$	$7,312$	$0,159$	$0,30$	$21,1904$	$0,094$	$1,00$	$6,414$	$0,176$	$0,30$	$10,473$	$0,051$
μ	$2,00$	$0,318$	$0,456$	$2,00$	$0,335$	$0,136$	$2,00$	$0,622$	$0,331$	$2,00$	$0,207$	$0,093$
φ_1	$1,00$	$9,491$	$0,691$	$1,00$	$42,096$	$0,615$	$1,00$	$1,245$	$0,466$	$1,00$	$3,463$	$0,560$
φ_2	$1,00$	$8,964$	$0,528$	$1,00$	$28,573$	$1,121$	$1,00$	$1,951$	$0,910$	$1,00$	$9,094$	$0,328$
φ_3	$1,50$	$24,874$	$1,753$	$1,50$	$27,398$	$0,883$	$1,50$	$10,009$	$1,138$	$1,50$	$13,259$	$0,928$

5.1.4.2 Estudo II: diagnósticos de influência

Nesta seção avalia-se a performance das ferramentas de diagnósticos para detectar casos influentes. A fim de realizar esta avaliação, geraram-se dois conjuntos de dados com a variável-resposta dada como em (5.18), com estrutura de matriz escala dada pelo modelo família Matérn, considerando-se o parâmetro de forma $\delta \in \{0,5; 1.0\}$ e $n = 100$. Após gerar um conjunto de dados $\boldsymbol{y} = (y_1, \ldots, y_n)^\top$, contaminou-se seu valor máximo com o objetivo de gerar um *outlier*. Tal contaminação é similar à utilizada por Ortega, Bolfarine e Paula (2003) para estudar a influência de uma perturbação em uma variável explicativa em modelos log-gamma generalizados de efeito fixo. Especificamente, considera-se a contaminação

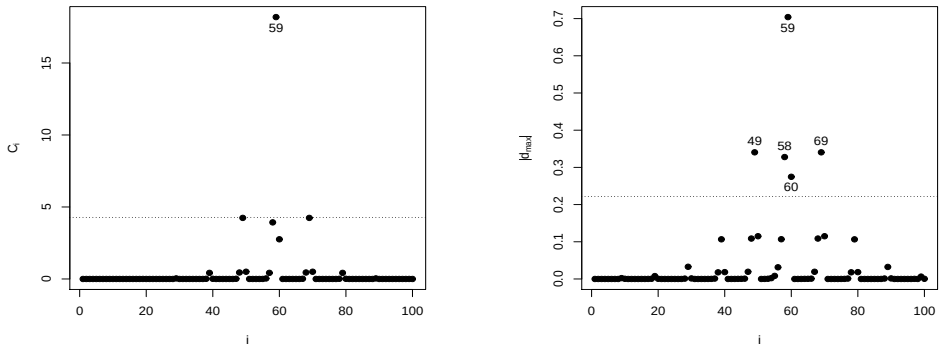
$$y^*_{\max} = y_{\max} + 10\sqrt{\boldsymbol{y}^\top \boldsymbol{y}}. \tag{5.35}$$

Para o primeiro conjunto de dados ($\delta = 0,5; n = 100$), o caso contaminado de acordo com (5.35) é o caso #59. A Figura 17 (a) mostra o *box plot* para este conjunto de dados, o qual identifica o caso #59 como um *outlier*. A Figura 17 (b) exibe a grade amostral, dividida por quartis, na qual é possível identificar a localização do *outlier*. Depois da contaminação, as estimativas dos parâmetros (obtidas usando o método BFGS) com seus desvios padrões estimados assintóticos (entre parênteses) são: $\widehat{\alpha} = 4,828(3,414 \times 10^{-1})$, $\widehat{\mu} = 71,188(1,309 \times 10)$, $\widehat{\varphi_1} = 480,101(8,989 \times 10^2)$, $\widehat{\varphi_2} = 200,090(8,999 \times 10^2)$ e $\widehat{\varphi_3} = 0,346(1,472 \times 10^{-3})$, resultando em um raio de dependência espacial estimado de $\widehat{a} = 1,038(4,416 \times 10^{-3})$. A Figura 18 mostra o gráfico de C_i (18 (a)) e $|d_{\max}|$ (18 (b)) para a estrutua de dependência espacial descrita pelo modelo família Matérn com $\delta = 0,5$, considerando a contaminação de acordo com (5.35). Este estudo de influência local mostra que o caso #59 é detectado como influente.



(a) (b)

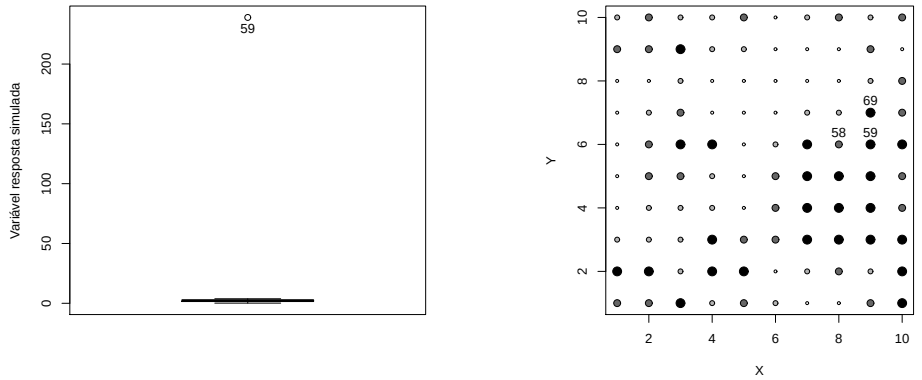
Figura 17 *Box plot* (a) e *post plot* (b) dos dados simulados para o modelo Matérn com $\delta = 0,5$.



(a) (b)

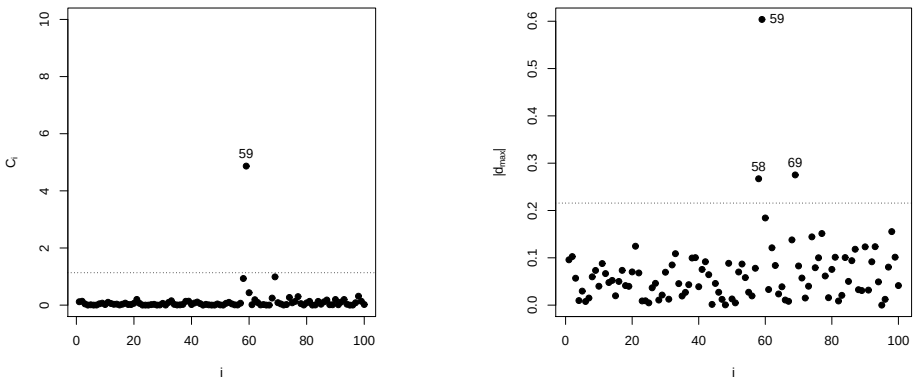
Figura 18 Gráficos C_i *versus* i (a) e $|d_{\max_i}|$*versus* i (b) para os dados simulados para o modelo Matérn com $\delta = 0,5$.

Similarmente ao primeiro conjunto de dados, para o segundo ($\delta = 1,0; n = 100$), o caso contaminado de acordo com (5.35) é o caso #59. A Figura 19 (a) mostra o *box plot* para esse segundo conjunto de dados, enquanto a Figura 19 (b) revela a correpondente grade amostral dividida em quartis. No segundo conjunto, depois da contaminação, os parâmetros estimados (obtidos pelo método BFGS) com seus SEs (entre parênteses) são: $\widehat{\alpha} = 3,525(0,249)$, $\widehat{\mu} = 5,199(2,980)$, $\widehat{\varphi_1} = 0,025(0,066)$, $\widehat{\varphi_2} = 3,538(0,845)$ e $\widehat{\varphi_3} = 2,084(0,002)$, e apresentam distância máxima de dependência espacial estimada de $\widehat{a} = 8,336(0,008)$. A Figura 20 mostra os gráficos de influência local que, como no primeiro caso, indicam o caso #59 como ponto influente. Portanto, a metodologia de diagnóstico proposta parece ser eficaz para a detecção de casos influentes.



(a) (b)

Figura 19 *Box plot* (a) e *Post plot* (b) dos dados simulados para o modelo Matérn com $\delta = 1$.



(a) (b)

Figura 20 Gráficos usando C_i *versus* i (a) e $|d_{\max}|$ *versus* i (b) dos dados simulados para o modelo Matérn com $\delta = 1$.

5.1.5 Aplicações numéricas

5.1.5.1 Descrição da área experimental e estudo das variáveis

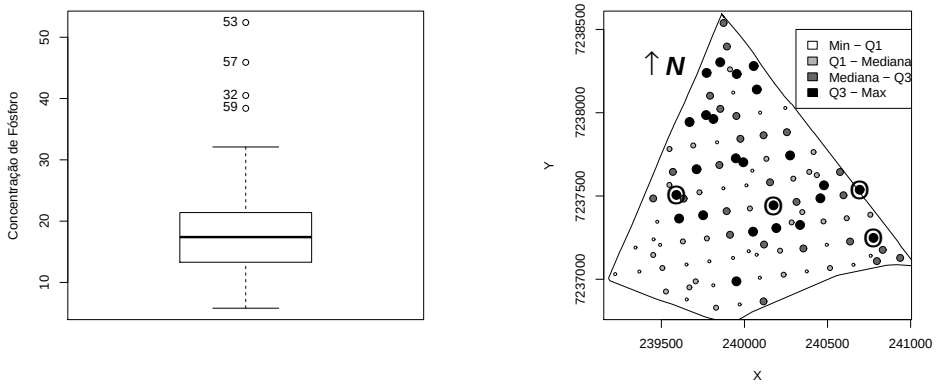
O estudo experimental foi conduzido durante o ano safra 2012/2013 em uma área comercial de produção de grãos de 167,35 ha de Cascavel, que está localizada na região Oeste do Estado do Paraná, Brasil. A área tem localização geográfica de, aproximadamente, 24,95^{o} S/53,57^{o} W, com uma altitude média de 650 m. O solo é classificado como LATOSSOLO VERMELHO Distroférrico argissólico (EMBRAPA, 2009). O clima da região é classificado como temperado mesotérmico e superúmido, tipo climático Cfa Köeppen e a temperatura anual média é de 21^{o}C.

Os locais de amostragem do solo foram definidos formando uma grade regular com distância entre pontos de aproximadamente 140 m. Em adição, localizações extras foram escolhidas aleatoriamente em toda região. Isso conduz a um total de $n = 102$ localizações, as quais foram georreferenciadas em coordenadas UTM usando um dispositivo GPS. Para a análise das propriedades químicas, quatro amostras de solo foram coletadas nas camadas de 0,0 a 0,2 m de profundidade, nas proximidades do ponto, e misturadas para produzir o elemento amostral representativo da parcela. As amostras compostas foram analisadas no laboratório da Cooperativa Central de Desenvolvimento Tecnológico e Econômico Ltda. (CODETEC, Brasil) para determinar as concentrações de macro e micronutrientes. A concentração do elemento fósforo foi escolhida como variável de estudo porque tem efeito positivo no crescimento e nutrição das plantas. Além disso, a variabilidade espacial do fósforo é extremamente importante para a gestão agrícola. De acordo com a classificação dada por COAMO-CODETEC (2001), a concentração de fósforo em solo argiloso é considerada como segue. Para o plantio de soja, é: (i) baixa para níveis inferiores a 3 mg/dm^3; (ii) média para níveis entre 3,1 e 6 mg/dm^3; (iii) alta para níveis entre 6,1 and 9 mg/dm^3; e (iv) muito alta para níveis superiores a 9 mg/dm^3. Para a plantação de milho, a concentração de fósforo é: (i) baixa para níveis inferiores a 2 mg/dm^3; (ii) média para níveis entre 2,1 e 4,5 mg/dm^3; (iii) alta para níveis entre 4,6 e 11 mg/dm^3; e (iv)

muito alta para níveis superiores a 11 mg/dm^3.

5.1.5.2 Análise exploratória e de dependência espacial

A análise exploratória de dados (AED) para a concentração de Fósforo foi dividida em duas: não espacial e espacial. Do ponto de vista não espacial, a média amostral da concentração de Fósforo no solo da área sob estudo é 18,11 mg/dm^3. Todavia, os correspondentes coeficientes amostrais de variação (CV), simetria (CS) e curtose (CK) são CV = 0,41 (41%), CS = 1,787 e CK = 5,104. Estas estatísticas descritivas indicam um grau razoável de homogeneidade em torno da média, uma assimetria positiva e um alto nível de curtose, tal como visualizado no *box plot* da Figura 21 (a). Note que quatro *outliers*, identificados como casos #32, #53, #57 e #59, são detectados. Os pontos circulados na Figura 21 (b) identificam os casos extremos, que estão localizados na parte inferior da região estudada. A AED não espacial é compatível com o uso da distribuição BS. Do ponto de vista espacial, uma análise padrão dos variogramas (omitida aqui), usando as direções 0^o, 45^o, 90^o e 135^o, mostra que os variogramas direcionais possuem comportamentos semelhantes até a distância de aproximadamente 900 m. Portanto, podemos assumir isotropia até essa distância.



(a) (b)

Figura 21 *Box plot* (a) e *post plot* (b) dos dados da concentração de Fósforo.

5.1.5.3 Estimação dos parâmetros e krigagem

A escolha do melhor modelo para descrever a estrutura de dependência espacial da concentração de fósforo considerou os critérios de validação cruzada e o valor máximo da função log-verossimilhanaça. O parâmetro δ, correspondente à ordem do modelo de variograma na família Matérn, não é estimado para evitar problemas de identificabilidade na estimação dos parâmetros da matriz de covariância. Assim, vários modelos de variograma baseados no modelo família Matérn foram ajustados e então o modelo com o menor valor de validação cruzada foi escolhido. Uma vez determinado δ, o melhor modelo Matérn foi usado para estimar $\boldsymbol{\theta}$ pelo método ML. Usando os critérios anteriormente descritos, o melhor modelo corresponde ao modelo família Matérn com parâmetro $\delta = 2,5$. Os parâmetros do modelo BS (com SE estimado

assintótico entre parênteses) são $\widehat{\alpha} = 0,997(3,521)$, $\widehat{\mu} = 2,807(0,082)$, $\widehat{\varphi}_1 = 0,134(0,946)$, $\widehat{\varphi}_2 = 0,020(0,142)$, $\widehat{\varphi}_3 = 177,940(0,0000014)$, e $\widehat{a} = 1,053(0,0000083)$; veja Tabela 6. Então, obtém-se o mapa espacial apresentado na Figura 23 (a) para a concentração de Fósforo no solo usando o modelo espacial BS ajustado obtido em (5.16) e o método de interpolação Krigagem ordinária descrito em (5.15). Observe que das estimativas ML obtidas e apresentadas acima, a matriz escala ajustada é $\widehat{\boldsymbol{\Sigma}} = 0,134\boldsymbol{I}_n + 0,020\widehat{\boldsymbol{R}}$, com $\widehat{\boldsymbol{R}}$ dada em (5.11), e $\delta = 2,5$. De acordo com a classificação fornecida em COAMO-CODETEC (2001) (Seção 5.1.5.1), note que a concentração de Fósforo é considerada adequada tanto para o plantio de soja como de milho.

5.1.5.4 Seleção do modelo

Nesta Seção comparam-se os modelos espaciais BS e Gaussiano de acordo com os critérios de informação de Akaike (AIC) e Bayesiano de Schwarz (BIC). Esses critérios são dados por $\mathrm{AIC} = -2\ell(\widehat{\theta}) + 2d$ e $\mathrm{BIC} = -2\ell(\widehat{\theta}) + d\log(n)$, em que $\ell(\widehat{\theta})$ é logaritmo da função verossimilhança para o parâmetro θ associado com o modelo avaliado em $\theta = \widehat{\theta}$, d é a dimensão do espaço paramêtrico, e n é o tamanho do conjunto de dados. Ambos critérios são baseados em uma penalização do logaritmo da função verossimilhança para modelos mais complexos, isto é, com mais parâmetros. Assim, um modelo cujos critérios de informação tem valores menores é melhor (FERREIRA; GOMES; LEIVA, 2012; LEIVA et al., 2015e) e referências neles citadas. Além dos critérios de informação AIC e BIC, o fator Bayes (BF) também pode ser usado para comparar os modelos espaciais BS e Gaussiano. O BF, denotado por B_{12}, permite comparar M_1 (modelo considerado como correto) com M_2 (modelo a ser comparado com M_1) a partir de $2\log(B_{12}) \approx \mathrm{BIC}_{\mathrm{M}_2} - \mathrm{BIC}_{\mathrm{M}_1}$, em que $\mathrm{BIC}_{\mathrm{M}_i}$ denota o BIC associado ao modelo M_i, para $i = 1, 2$. O BF fornece um valor objetivo para quantificar o grau de superioridade de um modelo com respeito a outro. Uma interpretação do BF é fornecida na Tabela 4 (LEIVA et al., 2015e). Então, de acordo com as Tabelas 4 e 5, detecta-se que o modelo BS é superior ao modelo Gaussiano com uma evidência muito forte em seu favor quando os dados não são transformados pelo logaritmo.

Tabela 4 Interpretação de $2\log(B_{12})$ associado com o BF.

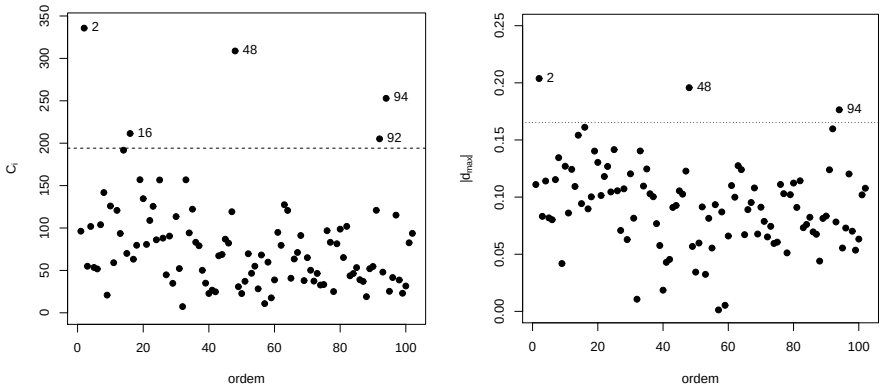
$2\log(B_{12})$	Evidência em favor de M_1
< 0	Negativa (M_2 é aceito)
$[0, 2)$	fraca
$[2, 6)$	Positiva
$[6, 10)$	forte
≥ 10	muito forte

Tabela 5 Comparação dos modelos BS e normal

Modelo	$\ell(\widehat{\theta})$	AIC	BIC	$2\log(B_{12})$
BS (dados transformados)	$-44,577$	$99,154$	$112,279$	–
Gaussiano (dados transformados)	$-43,890$	$95,780$	$106,280$	–
BS	$-332,576$	$675,152$	$688,276$	28,224
Gaussiano	$-349,000$	$706,000$	$716,500$	–

5.1.5.5 Diagnóstico de Influência

Realizou-se o diagnóstico de influência local para avaliar o efeito de casos atípicos no mapa espacial ajustado exibido na Figura 23 (a). Usando um esquema de perturbação na variável-resposta para a detecção de casos influentes, gráficos de C_i *versus* i e $|d_{\max}|$ *versus* i são considerados. Fica claro, a partir da Figura 22, que os casos #2, #48 e #94 são identificados como influentes por ambas as técnicas. Note que os três casos são diferentes dos *outliers* detectados pelo *box plot* na Figura 21 (a), que indica a relevância de usar o método de influência local ao invés de uma simples análise simples usando o *box plot*.



(a) (b)

Figura 22 Gráficos C_i *versus* i (a) e $|d_{\max}|$ *versus* i (b) para os dados de concentração de Fósforo.

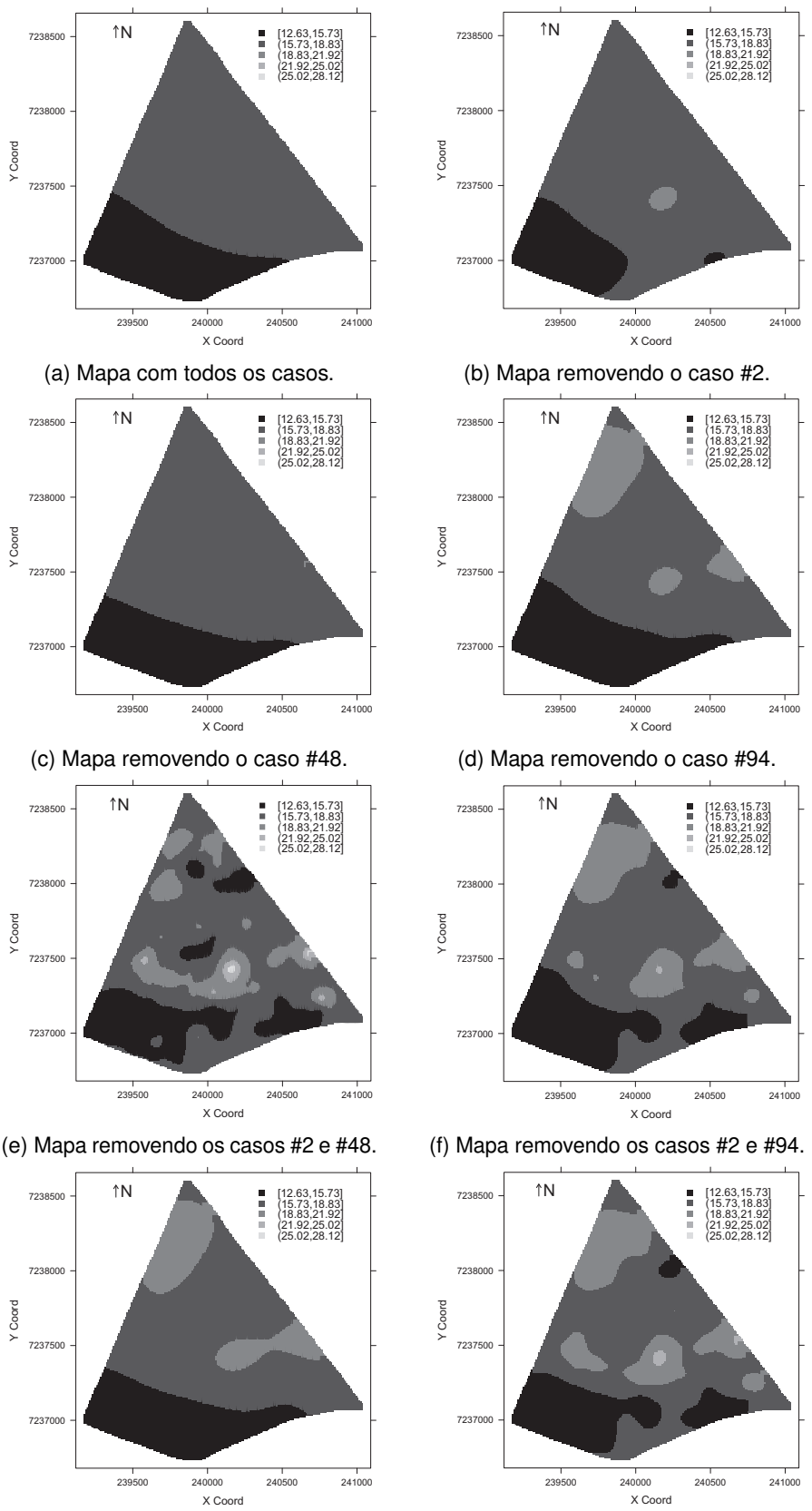
A Tabela 6 mostra os parâmetros estimados do modelo e os correspondentes SEs estimados entre parênteses, para vários modelos espaciais ajustados com o conjunto de dados completo e subconjuntos dele quando os casos influentes são removidos, individualmente ou em conjunto. Como é habitual no método de influência, uma vez que um ou mais casos influentes são identificados, os casos são removidos do conjunto de dados para investigar como sua remoção afeta a seleção do modelo, a estimação dos parâmetros e a construção dos mapas. Note que a remoção de casos influentes causa mudanças dramáticas no raio de dependência espacial estimado, $\widehat{a}$, especialmente quando os casos #32 e #48 são removidos em conjunto. Isto se deve a uma mudança considerável na estimativa de $\widehat{\varphi}_3$ e uma mudança no modelo de variograma escolhido para descrever a variabilidade espacial. A Tabela 6 mostra também os p-valores para hipóteses da forma H_0: $\theta_j = 0$ *versus* H_1: $\theta_j \neq 0$, em que θ_j é qualquer um dos parâmetros do vetor $\boldsymbol{\theta} = (\alpha, \mu, \varphi_1, \varphi_2, \varphi_3)^\top$. Para testar H_0, pode-se usar a estatística de teste $Z = \widehat{\theta}_j / \mathrm{SE}(\widehat{\theta}_j)$, em que $\widehat{\theta}_j$ é o estimador ML de θ_j e $\mathrm{SE}(\widehat{\theta}_j)$ o correspondente SE. Sabe-se que Z possui distribuição assintóticamente normal; veja resultados detalhados em (5.23). Dos p-valores apresentados nesta tabela, segue que, para todos os modelos, a hipótese H_0 é rejeitada a 5% significância para os parâmetros μ e φ_3, ao passo que H_0 não é rejeitada nos testes para os parâmetros α, φ_1 e φ_2. Note ainda que nenhuma mudança inferencial é

detectada ao serem retirados os casos influentes. Observe que, nos casos em que o espaço de parâmetros é o conjunto dos números reais positivos, precisa-se restringir H_1 a este conjunto, o que foi considerado em conformidade com os p-valores correspondentes. Note também que, para o modelo Matérn, o raio de dependência espacial a é uma função de φ_3, isto é, $a = c\varphi_3$, em que c é uma constante que depende de δ, de modo que $\widehat{a} = c\widehat{\varphi}_3$ e então $\mathrm{SE}(\widehat{a}) = c\mathrm{SE}(\widehat{\varphi}_3)$.

Tabela 6 Estimativas ML dos parâmetros estimados (com SE estimados assintóticos entre parênteses e p-valores entre colchetes), e valores dos índices GA e κ para o conjunto de dados de concentração de Fósforo com caso(s) excluídos indicados.

Caso(s) excluídos	Modelo	$\widehat{\alpha}$	$\widehat{\mu}$	$\widehat{\varphi_1}$	$\widehat{\varphi_2}$	$\widehat{\varphi_3}$	$\widehat{a}$	GA	κ
Nenhum	Matérn $\delta = 2,5$	0,997	2,807	0,134	0,020	177,940	1053,405	-	-
		(3,521)	(0,082)	(0,946)	(0,142)	(0,0000014)	(0,0000083)		
		[0,389]	[$< 0,001$]	[0,444]	[0,444]	[$< 0,001$]	[$< 0,001$]		
#2	Matérn $\delta = 2,5$	0,993	2,831	0,125	0,016	108,655	643,238	0,90	0,72
		(4,097)	(0,059)	(1,031)	(0,132)	(0,0000008)	(0,0000047)		
		[0,404]	[$< 0,001$]	[0,452]	[0,452]	[$< 0,001$]	[$< 0,001$]		
#48	Matérn $\delta = 2,5$	0,996	2,824	0,125	0,019	152,374	902,054	0,97	0,92
		(3,417)	(0,073)	(0,856)	(0,133)	(0,0000012)	(0,0000071)		
		[0,386]	[$< 0,001$]	[0,442]	[0,443]	[$< 0,001$]	[$< 0,001$]		
#94	Matérn $\delta = 1,0$	0,997	2,817	0,122	0,028	308,828	1235,312	0,84	0,69
		(3,417)	(0,073)	(0,856)	(0,133)	(0,0000012)	(0,0009092)		
		[0,374]	[$< 0,001$]	[0,436]	[0,437]	[$< 0,001$]	[$< 0,001$]		
#2, #48	Matérn $\delta = 0,5$	0,991	2,845	0,071	0,060	81,052	243,156	0,65	0,31
	(exponencial)	(6,149)	(0,046)	(0,882)	(0,746)	(0,0000884)	(0,0002652)		
		[0,436]	[$< 0,001$]	[0,468]	[0,469]	[$< 0,001$]	[$< 0,001$]		
#2, #94	Matérn $\delta = 0,5$	0,995	2,840	0,097	0,038	182,059	546,177	0,72	0,45
	(exponencial)	(3,308)	(0,063)	(0,644)	(0,254)	(0,0002974)	(0,0008922)		
		[0,382]	[$< 0,001$]	[0,440]	[0,441]	[$< 0,001$]	[$< 0,001$]		
#48, #94	Matérn $\delta = 2,5$	0,994	2,834	0,114	0,024	144,768	856,737	0,80	0,60
		(2,980)	(0,077)	(0,683)	(0,145)	(0,0000010)	(0,0000059)		
		[0,370]	[$< 0,001$]	[0,434]	[0,435]	[$< 0,001$]	[$< 0,001$]		
#2, #48, #94	Matérn $\delta = 0,5$	0,982	2,855	0,085	0,041	143,699	431,097	0,66	0,35
	(exponencial)	(3,442)	(0,056)	(0,597)	(0,287)	(0,0002051)	(0,0006153)		
		[0,388]	[$< 0,001$]	[0,444]	[0,443]	[$< 0,001$]	[$< 0,001$]		

A Figura 23 exibe mapas de contorno da concentração de fósforo no solo, produzidos a partir do modelo espacial BS e da interpolação por Krigagem ordinária. Os mapas foram criados baseados em dois cenários: (i) usando o conjunto de dados completos (mapa de referência) e (ii) removendo os casos influentes #2, #48 e #94, individualmente e em conjunto. Para a construção dos mapas, consideram-se cinco classes de mesmo tamanho, obtidas pela divisão em partes iguais do intervalo de variação da concentração de Fósforo estimada. Note que a retirada de casos individualmente não altera o mapa de forma expressiva. Contudo, a remoção conjunta de casos causa mudança considerável.



(a) Mapa com todos os casos.

(b) Mapa removendo o caso #2.

(c) Mapa removendo o caso #48.

(d) Mapa removendo o caso #94.

(e) Mapa removendo os casos #2 e #48.

(f) Mapa removendo os casos #2 e #94.

(g) Mapa removendo os casos #48 e #94.

(h) Mapa removendo os casos #2, #48 e #94.

Figura 23 Mapas de contorno com os efeitos da remoção do(s) caso(s) indicados para os dados de concentração de Fósforo.

Uma comparação mais objetiva dos mapas foi realizada usando os índices GA e κ

(Tabela 6). Então, de acordo com Anderson et al. (1976), quando os casos são removidos individualmente, os mapas são similares ao mapa de referência, apresentado na Figura 21 (a). No entanto, isso não ocorre quando os casos são removidos em conjunto. Quando se aplica a classificação de Krippendorff (2004) para o índice κ, é possível notar o seguinte: (i) o mapa que remove o caso influente #48 tem uma similaridade alta quando comparado ao mapa de referência; (ii) os mapas que removem os casos #2 e #94 individualmente têm uma similaridade média com o mapa de referência; e (iii) os outros mapas apresentam baixa similaridade com o mapa de referência. Tais conclusões sugerem que os casos são influentes apenas conjuntamente.

5.1.6 Conclusões e trabalhos futuros

Se propôs um novo modelo espacial log-linear baseado na distribuição Birnbaum-Saunders. Este é um modelo alternativo ao modelo espacial Gaussiano para descrever dados com estrutura de dependência espacial e, o mais importante, com uma distribuição assimétrica positiva. Estimativas por máxima verossimilhança dos parâmetros do modelo foram calculadas usando uma abordagem iterativa. Diagnósticos de influência locais e equações correspondentes para a perturbação mais adequada para o novo modelo foram obtidos. Avaliou-se a performance do processo de estimação e das ferramentas de diagnóstico por simulação. Para amostras grandes, a estimação e os diagnósticos apresentam boa performance. A abordagem proposta também foi usada para analisar dados reais da engenharia agrícola. Nesta aplicação, casos influentes foram detectados, e suas retiradas causaram mudança considerável no raio de dependência espacial e nos mapas. Mas, é importante notar que nem todos esses casos influentes foram identificados como *outliers*.

Alguns possíveis problemas a serem abordados em estudos futuros são os seguintes. Primeiro, uma vez que a distribuição Birnbaum-Saunders baseia-se na distribuição normal, uma versão de cauda mais pesada baseada, por exemplo, na distribuição t-Student pode ser considerada. Assim, é possível reduzir a influência de dados atípicos, que podem ter efeito prejudicial em mapas espaciais. Segundo, variáveis explanatórias podem ser consideradas na modelagem espacial, pois podem ajudar a melhorar a capacidade de predição do modelo. Terceiro, outros esquemas de perturbação podem ser considerados para avaliar a influência de pontos. Quarto, pode-se considerar mais de uma variável aleatória na modelagem espacial por estruturas multivariadas para a distribuição Birnbaum-Saunders. Os trabalhos sobre estes quatro temas estão atualmente em andamento e espera-se relatar algumas descobertas em artigos futuros.

5.1.7 Referências

ANDERSON, J.R.; HARDY, E.E.; ROACH, J.T.; WITMER, R.E. **A land use and land cover classification system for use with remote sensor data.** Washington: United States Government Printing Office, 1976.

ASSUMPÇÃO, R.A.B.; URIBE-OPAZO, M.A.; GALEA, M. Local influence for spatial analysis of soil physical properties and soybean yield using Student-t distribution. **Revista Brasileira de Ciência do Solo,** v. 35, p. 1917-1926, 2011.

ASSUMPÇÃO, R.A.B.; URIBE-OPAZO, M.A.; GALEA, M. Analysis of local influence in geostatistics using Student-t distribution. **Journal of Applied Statistics,** v. 41, p. 2323-2341, 2014.

AZZALINI, A.; CAPITANIO, A. Statistical applications of the multivariate skew-normal distribution. **Journal of the Royal Statistical Society B,** v. 61, p. 579-602, 1999.

BARROS, M.; PAULA, G.A.; LEIVA, V. An R implementation for generalized Birnbaum-Saunders distributions. **Computational Statistics and Data Analysis,** v. 53, p. 1511-1528, 2009.

BIRNBAUM, Z.W.; SAUNDERS, S.C. A new family of life distributions. **Journal of Applied Probability,** v. 6, p. 319-327, 1969.

BORSSOI, J.A.; DE BASTIANE, F.; URIBE-OPAZO, M.A.; GALEA, M. Local Influence of explanatory variables in Gaussian spatial linear models. **Chilean Journal of Statistics,** v. 2, p. 29-38, 2011.

CAMBARDELLA, C. A.; MOORMAN, T. B.; NOVAK, J. M.; PARKIN, T. B.; KARLEN, D. L.; TURCO, R.F.; KONOPKA, A. E. Field-scale variability of soil properties in Central Iowa Soils. **Soil Science Society of America Journal,** v. 58, p. 1501-1511, 1994.

COOPERATIVA AGROPECUÁRIA MOUÃOENSE LTDA.(COAMO)-COOPERATIVA CENTRAL DE DESENVOLVIMENTO TECNOLÓGICO E ECONÔMICO LTDA.(COODETEC). **Fertilidade do solo e nutrição de plantas.** Cascavel: Coamo-Coodetec, 2001.

COOK, R.D. Assessment of local influence. **Journal of the Royal Statistics Society B,** v. 48, p. 133-169, 1986.

COOK, R.D. Influence assessment. **Journal of Applied Statistics,** v. 14, p. 117-131, 1987.

CRESSIE, N.A.C. **Statistics for Spatial Data.** New-York: Jonh Wiley & Sons, 1993.

DAGENAIS, M.G.; DUFOUR, J.M. Invariance, nonlinear models and asymptotic tests. **Econometrica,** v. 59, p. 1601-1615, 1991.

DAVIS, D.J. An analysis of some failure data. **Journal of the American Statistical Association,** v. 47, p. 113-150, 1952.

DE BASTIANI, F.; CYSNEIROS, A.H.M.A.; URIBE-OPAZO, M.A.; GALEA, M. Influence diagnostics in elliptical spatial linear models. **TEST,** v. 24, p. 322-340, 2015.

DIGGLE, P.J.; RIBEIRO JÚNIOR, P.J. **Model-Based Geoestatistics.** New York: Springer, 2007.

EMPRESA BRASILEIRA DE INVESTIGAÇÃO AGROPECUÁRIA (Embrapa). **Sistema brasileiro de classificação de solo.** Rio de Janeiro: Empresa Brasileira de Investigação Agropecuária-Centro Nacional de Pesquisa de Solos (Embrapa-SPI), 2009.

FERREIRA, M.; GOMES, M.I.; LEIVA, V. On an extreme value version of the Birnbaum-Saunders distribution. **Revistat Statistical Journal,** v. 10, p. 181-210, 2012.

GALEA, M.; LEIVA, V.; PAULA, G.A. Influence diagnostics in log-Birnbaum-Saunders regression models. **Journal of Applied Statistics,** v. 31, p. 1049-1064, 2004.

GIMENEZ, P.; GALEA, M. Influence measures on corrected score estimators in functional heteroscedastic measurement error models. **Journal of Multivariate Analysis,** v. 114, p. 1-15, 2013.

GRZEGOZEWSKI, D.M.; URIBE-OPAZO, M. A.; DE BASTIANE, F.; GALEA, M. Local influence when fitting Gaussian spatial linear models: An agriculture application. **Ciência e Investigação Agrária,** v. 40, p. 235-252, 2013.

ISAAKS, E.; SRIVASTAVA, R. **An Introduction to Applied Geostatistics.** New York: Oxford University Press, 1989.

JOHNSON, N.L.; KOTZ, S.; BALAKRISHNAN, N. **Continuous Univariate Distributions.** New York: Jonh Wiley & Sons, 1994. v.1

JOHNSON, N.L.; KOTZ, S.; BALAKRISHNAN, N. **Continuous Univariate Distributions.** New York: Jonh Wiley & Sons, 1995. v.2

KENDRICK, D. **Stochastic Control for Economic Models.** New York: McGraw-Hill, 2002.

KRIGE, D.G. A statistical approach to some basic mine valuation problems on the Witwatersrand. **Journal of the Chemical, Metallurgical and Mining Society of South Africa,** v. 52, p. 119-139, 1951.

KRIPPENDORFF, K. **Content Analysis: An Introduction to its Methodology.** Thousand Oaks: Sage Publications, 2004.

KUNDU, D.; BALAKRISHNAN, N.; JAMALIZADEH, A. Generalized multivariate Birnbaum-Saunders distributions and related inferential issues. **Journal of Multivariate Analysis,** v. 116, p. 230-244, 2013.

LANGE, K.L.; LITTLE, R.J.; TAYLOR, J.M. Robust statistical modeling using the t distribution. **Journal of the American Statistical Association,** v. 84, p. 881-896, 1989.

LANGE, K. **Numerical Analysis for Statisticians.** New York: Springer, 2000.

LEIVA, V.; BARROS, M.; PAULA, G.A.; SANHUEZA, A. Generalized Birnbaum-Saunders distributions applied to air pollutant concentration. **Environmetrics,** v. 19, p. 235-249, 2008.

LEIVA, V.; SANHUEZA, A.; ANGULO, J.M. A length-biased version of the Birnbaum-Saunders distribution with application in water quality. **Stochastic Environmental Research and Risk Assessment,** v. 23, p. 299-307, 2009.

LEIVA, V.; ROJAS, E.; GALEA, M.; SANHUEZA, A. Diagnostics in Birnbaum-Saunders accelerated life models with an application to fatigue data. **Applied Stochastic Models in Business and Industry,** v. 30, p. 115-131, 2014.

LEIVA, V.; FERREIRA, M.; GOMES, M.I.; LILLO, C. Extreme value Birnbaum-Saunders regression models applied to environmental data. **Stochastic**

Environmental Research and Risk Assessment (in press)2015a. Disponível em <http://dx.doi.org/10.1007/s00477-015-1069-6, 2015a>. Acesso em: 08 out. 2015.

LEIVA, V.; MARCHANT, C.; RUGGERI, F.; SAULO, H. A criterion for environmental assessment using Birnbaum-Saunders attribute control charts. **Environmetrics** (in press), 2015b. Disponível em <http://dx.doi.org/10.1002/env.2349>. Acesso em: 08 out. 2015.

LEIVA V.; LIU, S.; SHI, L.; CYSNEIROS, F.J.A. Diagnostics in elliptical regression models with stochastic restrictions applied to econometrics. **Journal of Applied Statistics** (in press), 2015c. Disponível em <http://dx.doi.org/10.1080/02664763.2015>. Acesso em: 20 nov. 2015.

LEIVA, V.; SANTOS-NETO, M; CYSNEIROS, F.J.A.; BARROS, M. A methodology for stochastic inventory models based on a zero-adjusted Birnbaum-Saunders distribution. **Applied Stochastic Models in Business and Industry** (in press), 2015d. Disponível em:<http://dx.doi.org/10.1007/s10260-015-0329-4>. Acesso em: 20 nov. 2015.

LEIVA, V.; TEJO, M., GUIRAUD, P.; SCHMACHTENBERG, O; ORIO, P.; MARMOLEJO-RAMOS, F. Modeling neural activity with cumulative damage distributions. **Biological Cybernetics,** v.109, p.421-433, 2015e.

LIU, S.; LEIVA, V.; MA, T.; WELSH, A. Influence diagnostic analysis in the possibly heteroskedastic linear model with exact restrictions. **Statistical Methods and Applications** (in press), 2015. Disponível em: <http://dx.doi.org/10.1080/02664763.2015.1072140>. Acesso em 21 nov. 2015.

MARCHANT, C.; LEIVA, V.; CAVIERES, M.F.; SANHUEZA, A. Air contaminant statistical distributions with application to PM10 in Santiago, Chile. **Reviews of Environmental Contamination and Toxicology,** v. 223, p. 1-31, 2013.

MARCHANT, C.; LEIVA, V.; CYSNEIROS, F.J.A. A multivariate log-linear model for Birnbaum-Saunders distributions. **IEEE Transactions on Reliability** (in press), 2016. Disponível em: <http://dx.doi.org/10.1109/TR.2015.2499964>. Acesso em: 16 jan. 2016.

MARDIA, K.V.; MARSHALL, R.J. Maximum likelihood estimation of models for residual covariance in spatial regression. **Biometrika,** v. 71, p. 135-146, 1984.

MILITINO, A.F.; PALACIUS, M.B.; UGARTE, M.D. Outliers detection in multivariate spatial linear models. **Journal of Statistical Planning and Inference,** v. 136, p. 125-146, 2006.

MUIRHEAD, R.J. **Aspects of Multivariate Statistical Theory.** New York: Wiley & Sons, 1982.

MÜLLER, W.; STEHLÍK , M. Issues in the optimal design of computer simulation experiments. **Applied Stochastic Models in Business and Industry,** v. 25, p. 163-177, 2009.

MÜLLER, W.; STEHLÍK , M. Compound optimal spatial designs. **Environmetrics,** v. 21, p. 354-364, 2010.

NOCEDAL, J.; WRIGHT, S.J. **Numerical Optimization.** New York: Springer, 1999.

ORTEGA, E.M.M.; BOLFARINE, H.; PAULA, G.A. Influence diagnostics in generalized log-gamma regression models. **Computational Statistics and Data Analysis,** v. 42, p.

165-186, 2003.

PODLASKI, R. Characterization of diameter distribution data in near-natural forests using the Birnbaum-Saunders distribution. **Canadian Journal of Forest Research,** v. 38, p. 518-527, 2008.

R TEAM R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, 2015. Disponível em: <www.R-project.org>. Acesso em: 25 set. 2015.

RIECK, J.R.; NEDELMAN, J.R. A log-linear model for the Birnbaum-Saunders distribution. **Technometrics,** v. 33, p. 51-60, 1991.

SAULO, H.; LEIVA, V.; ZIEGELMANN, F.A.; MARCHANT, C. A nonparametric method for estimating asymmetric densities based on skewed Birnbaum-Saunders distributions applied to environmental data. **Stochastic Environmental Research and Risk Assessment,** v. 27, p. 1479-1491, 2013.

URIBE-OPAZO, M.A.; BORSSOI, J.A.; GALEA, M. Influence diagnostics in Gaussian spatial linear models. **Journal of Applied Statistics,** v. 39, p. 615-630, 2012.

VILCA, F.; SANHUEZA, A.; LEIVA, V.; CHRISTAKOS, G. An extended Birnbaum-Saunders model and its application in the study of environmental quality in Santiago, Chile. **Stochastic Environmental Research and Risk Assessment,** v. 24, p. 771-782, 2010.

VILLEGAS, C.; PAULA, G.A.; LEIVA, V. Birnbaum-Saunders mixed models for censored reliability data analysis. **IEEE Transactions on Reliability,** v. 60, p. 748-758, 2011.

WALLER, L.; GOTWAY, C. **Applied Spatial Statistics for Public Health Data.** New Jersey: Wiley & Sons, 2004.

XIA, J.; ZEEPHONGSEKUL, P.; PACKER, D. Spatial and temporal modelling of tourist movements using semi-Markov processes. **Tourism Management,** v. 32, p. 844-851, 2011.

ZHU, H.; LEE, S. Local Influence for incomplete-data models. **Journal of the Royal Statistical Society B,** v. 63, p. 111-126, 2001.

ZHU, H.; IBRAHIM, J.G.; LEE, S., ZHANG, H. Perturbation selection and influence measures in local influence analysis. **The Annals of Statistics,** v. 35, p. 2565-2588, 2007.

5.1.8 Apêndices

5.1.8.1 Apêndice I: O vetor escore do modelo BS log-linear espacial

Para o modelo BS log-linear espacial, o vetor escore é definido por

$$\boldsymbol{U}(\boldsymbol{\theta}) = \left(\frac{\partial \ell(\boldsymbol{\theta})}{\partial \alpha}, \frac{\partial \ell(\boldsymbol{\theta})}{\partial \mu}, \frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\varphi}^\top}\right)^\top = \left(U(\alpha), U(\mu), U(\varphi_1), U(\varphi_2), U(\varphi_3)\right)^\top.$$

Usando-se o logaritmo da função verossimilhança dado em (5.56), tem-se

$$\begin{aligned}
U(\alpha) &= \frac{\partial \ell(\boldsymbol{\theta})}{\partial \alpha} = \frac{\partial}{\partial \alpha}\left(-n\log(\alpha) - \frac{2}{\alpha^2}\boldsymbol{Y}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{Y}\right) = -\frac{n}{\alpha} + \frac{4}{\alpha^3}\boldsymbol{Y}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{Y},\\
U(\mu) &= \frac{\partial \ell(\boldsymbol{\theta})}{\partial \mu} = \frac{\partial}{\partial \mu}\left(-\frac{2}{\alpha^2}\boldsymbol{Y}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{Y} + \sum_{i=1}^{n}\log\left(\cosh\left(\frac{y_i-\mu}{2}\right)\right)\right)\\
&= \frac{2}{\alpha^2}\boldsymbol{U}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{Y} - \frac{1}{2}\sum_{i=1}^{n}\tanh\left(\frac{y_i-\mu}{2}\right),\\
U(\varphi_i) &= \frac{\partial \ell(\boldsymbol{\theta})}{\partial \varphi_i} = \frac{\partial}{\partial \varphi_i}\left(-\frac{1}{2}\log\left(|\boldsymbol{\Sigma}|\right) - \frac{2}{\alpha^2}\boldsymbol{Y}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{Y}\right)\\
&= -\frac{1}{2}\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\varphi_i}\right) + \frac{2}{\alpha^2}\boldsymbol{Y}^\top\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\varphi_i}\boldsymbol{\Sigma}^{-1}\boldsymbol{Y}, \quad j = 1,2,3.
\end{aligned}$$

em que $\boldsymbol{U} = (U_1, \ldots, U_n)^\top$, com $U_i = \cosh((y_i-\mu)/2)$, e $\boldsymbol{V} = (V_1, \ldots, V_n)^\top$ como dado em (5.19) com $V_i = \mathrm{senh}((y_i-\mu)/2)$, para $i = 1, \ldots, n$. Então, considerando-se que $\boldsymbol{\Sigma} = \varphi_1\boldsymbol{I}_n + \varphi_2\boldsymbol{R}$, tem-se

$$\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_1} = \boldsymbol{I}_n, \frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_2} = \boldsymbol{R}, \frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_3} = \varphi_2\frac{\partial\boldsymbol{R}}{\partial\varphi_3}.$$

Portanto,

$$\begin{aligned}
U(\varphi_1) &= -\frac{1}{2}\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\right) + \frac{2}{\alpha^2}\boldsymbol{Y}^\top\left(\boldsymbol{\Sigma}^{-1}\right)^2\boldsymbol{Y},\\
U(\varphi_2) &= -\frac{1}{2}\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\boldsymbol{R}\right) + \frac{2}{\alpha^2}\boldsymbol{Y}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{R}\boldsymbol{\Sigma}^{-1}\boldsymbol{Y},\\
U(\varphi_3) &= -\frac{\varphi_2}{2}\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{R}}{\partial\varphi_3}\right) + \frac{2\varphi_2}{\alpha^2}\boldsymbol{Y}^\top\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{R}}{\partial\varphi_3}\boldsymbol{\Sigma}^{-1}\boldsymbol{Y}.
\end{aligned}$$

A partir do modelo Matérn, dado em (5.11), para descrever a variabilidade espacial, tem-se que $\partial\boldsymbol{R}/\partial\varphi_3 = [\partial r_{ij}/\partial\varphi_3]$, em que $K'_\delta(u) = \partial K_\delta(u)/\partial u = -(1/2)(K_{\delta-1}(u) + K_{\delta+1}(u))$,

$$\frac{\partial r_{ij}}{\partial\varphi_3} = -\left(\frac{1}{\varphi_3}\right)\left(\delta\, r_{ij} + \frac{1}{2^{\delta-1}\mu(\delta)}\left(\frac{h_{ij}}{\varphi_3}\right)^{\delta+1}K'_\delta\left(\frac{h_{ij}}{\varphi_3}\right)\right), \quad i \neq j, i, j = 1, \ldots, n.$$

5.1.8.2 Apêndice II: A matriz de informação observada para o modelo BS log-linear espacial

A matriz de informação observada de Fisher para o modelo BS log-linear espacial é definida por $-\ddot{\boldsymbol{\ell}}(\boldsymbol{\theta})$ avaliada em $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$, em que $\ddot{\boldsymbol{\ell}}(\boldsymbol{\theta})$ é a matriz Hessiana dada por

$$\ddot{\boldsymbol{\ell}}(\boldsymbol{\theta}) = \begin{pmatrix}
\frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\alpha\partial\alpha} & \frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\alpha\partial\mu} & \frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\alpha\partial\boldsymbol{\varphi}^\top}\\
\frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\mu\partial\alpha} & \frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\mu\partial\mu} & \frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\mu\partial\boldsymbol{\varphi}^\top}\\
\frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\varphi}\partial\alpha} & \frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\varphi}\partial\mu} & \frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\varphi}\partial\boldsymbol{\varphi}^\top}
\end{pmatrix} = \begin{pmatrix}
\ddot{\ell}_{\alpha\alpha} & \ddot{\ell}_{\alpha\mu} & \ddot{\boldsymbol{\ell}}_{\alpha\boldsymbol{\varphi}}\\
\ddot{\ell}_{\mu\alpha} & \ddot{\ell}_{\mu\mu} & \ddot{\boldsymbol{\ell}}_{\mu\boldsymbol{\varphi}}\\
\ddot{\boldsymbol{\ell}}_{\boldsymbol{\varphi}\alpha} & \ddot{\boldsymbol{\ell}}_{\boldsymbol{\varphi}\mu} & \ddot{\boldsymbol{\ell}}_{\boldsymbol{\varphi}\boldsymbol{\varphi}}
\end{pmatrix},$$

com, para $\ddot{\ell}_{\alpha\mu} = \ddot{\ell}_{\mu\alpha}$,

$$\begin{aligned}
\ddot{\ell}_{\alpha\alpha} &= \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \alpha^2} = \frac{\partial}{\partial \alpha}\left(-\frac{n}{\alpha} + \frac{4}{\alpha^3}\boldsymbol{V}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{V}\right) = \frac{n}{\alpha^2} - \frac{12}{\alpha^4}\boldsymbol{V}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{V}, \\
\ddot{\ell}_{\mu\alpha} &= \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \mu \partial \alpha} = \frac{\partial}{\partial \mu}\left(\frac{4}{\alpha^3}\boldsymbol{V}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{V}\right) = -\frac{4}{\alpha^3}\boldsymbol{U}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{V},
\end{aligned}$$

$$\begin{aligned}
\ddot{\ell}_{\mu\mu} &= \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \mu^2} = \frac{\partial}{\partial \mu}\left(\frac{2}{\alpha^2}\boldsymbol{U}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{V} - \frac{1}{2}\sum_{i=1}^{n}\tanh\left(\frac{y_i - \mu}{2}\right)\right) \\
&= -\frac{1}{\alpha^2}\left(\boldsymbol{V}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{V} + \boldsymbol{U}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{U}\right) + \frac{1}{4}\sum_{i=1}^{n}\operatorname{sech}^2\left(\frac{y_i - \mu}{2}\right).
\end{aligned}$$

Note que $\ddot{\boldsymbol{\ell}}_{\boldsymbol{\varphi}\alpha} = \ddot{\boldsymbol{\ell}}_{\alpha\boldsymbol{\varphi}}^\top$ é um vetor 3×1 com elementos dados por

$$\ddot{\ell}_{\alpha\varphi_j} = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \alpha \partial \varphi_j} = \frac{\partial}{\partial \alpha}\left(-\frac{1}{2}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\varphi_i}\right) + \frac{2}{\alpha^2}\boldsymbol{V}^\top \boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\varphi_i}\boldsymbol{\Sigma}^{-1}\boldsymbol{V}\right) = -\frac{4}{\alpha^3}\boldsymbol{V}^\top\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_i}\boldsymbol{\Sigma}^{-1}\right)\boldsymbol{V},$$

para $j = 1, 2, 3$, em que $\partial\boldsymbol{\Sigma}/\partial\varphi_i$ é dado no Apêndice I. Então, com $\ddot{\ell}_{\alpha\varphi_1} = \ddot{\ell}_{\varphi_1\alpha}$, $\ddot{\ell}_{\alpha\varphi_2} = \ddot{\ell}_{\varphi_2\alpha}$ e $\ddot{\ell}_{\alpha\varphi_3} = \ddot{\ell}_{\varphi_3\alpha}$, tem-se

$$\ddot{\ell}_{\alpha\varphi_1} = -\frac{4}{\alpha^3}\boldsymbol{V}^\top\left(\boldsymbol{\Sigma}^{-1}\right)^2\boldsymbol{V}, \ddot{\ell}_{\alpha\varphi_2} = -\frac{4}{\alpha^3}\boldsymbol{V}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{R}\boldsymbol{\Sigma}^{-1}\boldsymbol{V}, \ddot{\ell}_{\alpha\varphi_3} = -\frac{4\varphi_2}{\alpha^3}\boldsymbol{V}^\top \boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{R}}{\partial \varphi_3}\boldsymbol{\Sigma}^{-1}\boldsymbol{V}.$$

Além disso, $\ddot{\boldsymbol{\ell}}_{\boldsymbol{\varphi}\mu} = \ddot{\boldsymbol{\ell}}_{\mu\boldsymbol{\varphi}}^\top$ é um vetor 3×1 com elementos

$$\ddot{\ell}_{\mu\varphi_j} = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \mu \partial \varphi_j} = \frac{\partial}{\partial \mu}\left(-\frac{1}{2}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\varphi_i}\right) + \frac{2}{\alpha^2}\boldsymbol{V}^\top \boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\varphi_i}\boldsymbol{\Sigma}^{-1}\boldsymbol{V}\right) = -\frac{2}{\alpha^2}\boldsymbol{U}^\top\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_j}\boldsymbol{\Sigma}^{-1}\right)\boldsymbol{V},$$

para $j = 1, 2, 3$. Então, com $\ddot{\ell}_{\mu\varphi_1} = \ddot{\ell}_{\varphi_1\mu}$, $\ddot{\ell}_{\mu\varphi_2} = \ddot{\ell}_{\varphi_2\mu}$ e $\ddot{\ell}_{\mu\varphi_3} = \ddot{\ell}_{\varphi_3\mu}$, tem-se

$$\ddot{\ell}_{\mu\varphi_1} = -\frac{2}{\alpha^2}\boldsymbol{U}^\top\left(\boldsymbol{\Sigma}^{-1}\right)^2\boldsymbol{V}, \ddot{\ell}_{\mu\varphi_2} = -\frac{2}{\alpha^2}\boldsymbol{U}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{R}\boldsymbol{\Sigma}^{-1}\boldsymbol{V}, \ddot{\ell}_{\mu\varphi_3} = -\frac{2\varphi_2}{\alpha^2}\boldsymbol{U}^\top \boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{R}}{\partial \varphi_3}\boldsymbol{\Sigma}^{-1}\boldsymbol{V}.$$

Adicionalmente, $\ddot{\boldsymbol{\ell}}_{\boldsymbol{\varphi}\boldsymbol{\varphi}} = (\ddot{\ell}_{\varphi_i\varphi_j})$ é uma matriz 3×3 simétrica com elementos dados por

$$\begin{aligned}
\ddot{\ell}_{\varphi_j\varphi_i} = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \varphi_j \varphi_i} &= -\frac{1}{2}\operatorname{tr}\left(-\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_j}\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_i} + \boldsymbol{\Sigma}^{-1}\frac{\partial^2 \boldsymbol{\Sigma}}{\partial \varphi_j \partial \varphi_i}\right) + \frac{2}{\alpha^2}\boldsymbol{V}^\top\left(-\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_j}\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\varphi_i}\boldsymbol{\Sigma}^{-1}\right)\boldsymbol{V} \\
&\quad + \frac{2}{\alpha^2}\boldsymbol{V}^\top\left(\boldsymbol{\Sigma}^{-1}\left(\frac{\partial^2 \boldsymbol{\Sigma}}{\partial \varphi_j \partial \varphi_i}\boldsymbol{\Sigma}^{-1} + \frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_i}\left(-\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_j}\boldsymbol{\Sigma}^{-1}\right)\right)\right)\boldsymbol{V}, \quad i, j = 1, 2, 3.
\end{aligned}$$

Então, como $\ddot{\ell}_{\varphi_1\varphi_2} = \ddot{\ell}_{\varphi_2\varphi_1}$, $\ddot{\ell}_{\varphi_1\varphi_3} = \ddot{\ell}_{\varphi_3\varphi_1}$ e $\ddot{\ell}_{\varphi_2\varphi_3} = \ddot{\ell}_{\varphi_3\varphi_2}$, tem-se

$$
\begin{aligned}
\ddot{\ell}_{\varphi_1\varphi_1} &= \frac{1}{2}\mathrm{tr}\left((\boldsymbol{\Sigma^{-1}})^2\right) - \frac{2}{\alpha^2}\boldsymbol{V}^\top\left((\boldsymbol{\Sigma^{-1}})^3 + (\boldsymbol{\Sigma^{-1}})^2\right)\boldsymbol{V},\\
\ddot{\ell}_{\varphi_1\varphi_2} &= \frac{1}{2}\mathrm{tr}\left(\boldsymbol{\Sigma^{-1}}\boldsymbol{R}\boldsymbol{\Sigma^{-1}}\right) - \frac{2}{\alpha^2}\boldsymbol{V}^\top\left(\boldsymbol{\Sigma^{-1}}\boldsymbol{R}(\boldsymbol{\Sigma^{-1}})^2 + \left(\boldsymbol{\Sigma^{-1}}\right)^2\boldsymbol{R}\boldsymbol{\Sigma}^{-1}\right)\boldsymbol{V},\\
\ddot{\ell}_{\varphi_1\varphi_3} &= \frac{\varphi_2}{2}\mathrm{tr}\left(\boldsymbol{\Sigma^{-1}}\frac{\partial\boldsymbol{R}}{\partial\varphi_3}\boldsymbol{\Sigma^{-1}}\right) - \frac{2\varphi_2}{\alpha^2}\boldsymbol{V}^\top\left(\boldsymbol{\Sigma^{-1}}\frac{\partial\boldsymbol{R}}{\partial\varphi_3}(\boldsymbol{\Sigma^{-1}})^2 + (\boldsymbol{\Sigma^{-1}})^2\frac{\partial\boldsymbol{R}}{\partial\varphi_3}\boldsymbol{\Sigma}^{-1}\right)\boldsymbol{V},\\
\ddot{\ell}_{\varphi_2\varphi_2} &= \frac{1}{2}\mathrm{tr}\left(\boldsymbol{\Sigma^{-1}}\boldsymbol{R}\boldsymbol{\Sigma^{-1}}\boldsymbol{R}\right) - \frac{4}{\alpha^2}\boldsymbol{V}^\top\left(\boldsymbol{\Sigma^{-1}}\boldsymbol{R}\boldsymbol{\Sigma^{-1}}\boldsymbol{R}\boldsymbol{\Sigma}^{-1}\right)\boldsymbol{V},\\
\ddot{\ell}_{\varphi_2\varphi_3} &= -\frac{1}{2}\mathrm{tr}\left(-\varphi_2\boldsymbol{\Sigma^{-1}}\frac{\partial\boldsymbol{R}}{\partial\varphi_3}\boldsymbol{\Sigma^{-1}}\boldsymbol{R} + \boldsymbol{\Sigma^{-1}}\frac{\partial\boldsymbol{R}}{\partial\varphi_3}\right) + \frac{2}{\alpha^2}\boldsymbol{V}^\top\left(-\varphi_2\boldsymbol{\Sigma^{-1}}\frac{\partial\boldsymbol{R}}{\partial\varphi_3}\boldsymbol{\Sigma^{-1}}\boldsymbol{R}\boldsymbol{\Sigma}^{-1}\right)\boldsymbol{V}\\
&\quad + \frac{2}{\alpha^2}\boldsymbol{V}^\top\boldsymbol{\Sigma^{-1}}\left(\frac{\partial\boldsymbol{R}}{\partial\varphi_3}\boldsymbol{\Sigma}^{-1} + \boldsymbol{R}\left(-\varphi_2\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{R}}{\partial\varphi_3}\boldsymbol{\Sigma}^{-1}\right)\right)\boldsymbol{V},\\
\ddot{\ell}_{\varphi_3\varphi_3} &= -\frac{\varphi_2}{2}\mathrm{tr}\left(-\varphi_2\boldsymbol{\Sigma^{-1}}\frac{\partial\boldsymbol{R}}{\partial\varphi_3}\boldsymbol{\Sigma^{-1}}\frac{\partial\boldsymbol{R}}{\partial\varphi_3} + \boldsymbol{\Sigma}^{-1}\frac{\partial^2\boldsymbol{R}}{\partial\varphi_3^2}\right) - \frac{2\varphi_2^2}{\alpha^2}\boldsymbol{V}^\top\left(\boldsymbol{\Sigma^{-1}}\frac{\partial\boldsymbol{R}}{\partial\varphi_3}\boldsymbol{\Sigma^{-1}}\frac{\partial\boldsymbol{R}}{\partial\varphi_3}\boldsymbol{\Sigma}^{-1}\right)\boldsymbol{V}\\
&\quad + \frac{2\varphi_2}{\alpha^2}\boldsymbol{V}^\top\boldsymbol{\Sigma}^{-1}\left(\frac{\partial^2\boldsymbol{R}}{\partial\varphi_3^2}\boldsymbol{\Sigma}^{-1} - \varphi_2\frac{\partial\boldsymbol{R}}{\partial\varphi_3}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{R}}{\partial\varphi_3}\boldsymbol{\Sigma}^{-1}\right)\boldsymbol{V}.
\end{aligned}
$$

5.1.8.3 Apêndice III: Matriz de informação esperada para o modelo BS log-linear espacial

A matriz de informação esperada de Fisher é $\boldsymbol{I}(\boldsymbol{\theta}) = \mathrm{E}[-\ddot{\boldsymbol{\ell}}(\boldsymbol{\theta})]$. Para o modelo BS log-linear espacial, esta matriz é dada por

$$
\boldsymbol{I}(\boldsymbol{\theta}) = \begin{pmatrix} I_{\alpha\alpha} & I_{\alpha\mu} & \boldsymbol{I}_{\alpha\boldsymbol{\varphi}} \\ I_{\mu\alpha} & I_{\mu\mu} & \boldsymbol{I}_{\mu\boldsymbol{\varphi}} \\ \boldsymbol{I}_{\boldsymbol{\varphi}\alpha} & \boldsymbol{I}_{\boldsymbol{\varphi}\mu} & \boldsymbol{I}_{\boldsymbol{\varphi}\boldsymbol{\varphi}} \end{pmatrix}.
$$

Uma vez que o erro do modelo $\boldsymbol{\varepsilon} \sim \log\text{-}\mathrm{BS}_n(\alpha\mathbf{1}, \mathbf{0}, \boldsymbol{\Sigma})$, tem-se que $(2/\alpha)\boldsymbol{V} \sim \mathrm{N}_n(\mathbf{0}, \boldsymbol{\Sigma})$, com $\boldsymbol{V}$ dado em (5.19). Então, $W = (2/\alpha)\boldsymbol{V}^\top\boldsymbol{\Sigma}^{-1}(2/\alpha)\boldsymbol{V} \sim \chi^2_n$, em que χ^2_n denota a distribuição qui-quadrado com n graus de liberdade, e então, $\mathrm{E}[W] = n$ (MUIRHEAD, 1982). Assim,

$$
I_{\alpha\alpha} = \mathrm{E}[-\ddot{\ell}_{\alpha\alpha}] = \mathrm{E}\left[-\frac{n}{\alpha^2} + \frac{12}{\alpha^4}\boldsymbol{V}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{V}\right] = \frac{2n}{\alpha^2}.
$$

Além disso, como $\mathrm{E}\left[\boldsymbol{X}^\top\boldsymbol{A}\boldsymbol{X}\right] = \left(\mathrm{E}\left[\boldsymbol{X}\right]\right)^\top\boldsymbol{A}\left(\mathrm{E}\left[\boldsymbol{X}\right]\right) + \mathrm{tr}\left(\boldsymbol{A}\boldsymbol{C}\right)$, em que $\boldsymbol{C}$ é a matriz de covariância de $\boldsymbol{X}$ (KENDRICK, 2002), $\boldsymbol{I}_{\boldsymbol{\varphi}\boldsymbol{\varphi}} = (I_{\varphi_i\varphi_j})$ é uma matriz simétrica 3×3 com elementos

$$
I_{\varphi_i\varphi_j} = \mathrm{E}[-\ddot{\ell}_{\varphi_i\varphi_j}] = \frac{1}{2}\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_i}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_j}\right), \quad i,j = 1,2,3.
$$

Além do mais, $\boldsymbol{I}_{\boldsymbol{\varphi}\alpha} = \boldsymbol{I}_{\alpha\boldsymbol{\varphi}}^\top = (I_{\varphi_1\alpha}, I_{\varphi_2\alpha}, I_{\varphi_3\alpha})^\top$ é um vetor 3×1 com elementos dados por

$$
I_{\varphi_i\alpha} = \mathrm{E}[-\ddot{\ell}_{\varphi_i\alpha}] = \frac{1}{\alpha}\mathrm{E}\left[\left(\frac{2}{\alpha}\boldsymbol{V}\right)^\top\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_i}\boldsymbol{\Sigma}^{-1}\right)\left(\frac{2}{\alpha}\boldsymbol{V}\right)\right] = \frac{1}{\alpha}\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_i}\right), \quad i = 1,2,3.
$$

em que $\mathrm{tr}(\boldsymbol{A})$ denota o traço da matriz $\boldsymbol{A}$. Para obterem-se os elementos $I_{\alpha\mu}, I_{\mu\mu}$ e $I_{\mu\boldsymbol{\varphi}}$, como $(y_i - \mu)/2$ é suficientemente pequeno tem-se $\cosh(\cdot) \approx 1$, usando a expansão da série de Taylor

para $\cosh(\cdot)$, de forma similar à usada por Rieck e Nedelman (1991) para $\mathrm{senh}(\cdot)$, tem-se

$$\begin{aligned} I_{\alpha\mu} &= I_{\mu\alpha} = \mathrm{E}\left[-\ddot{\ell}_{\alpha\mu}\right] = \mathrm{E}\left[\frac{4}{\alpha^3}\boldsymbol{U}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{V}\right] \approx \mathrm{E}\left[\frac{4}{\alpha^3}\mathbf{1}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{V}\right] = 0. \\ I_{\mu\mu} &= \mathrm{E}[-\ddot{\ell}_{\mu\mu}] \approx \frac{1}{4}\mathrm{E}\left[\left(\frac{2}{\alpha}\boldsymbol{V}\right)^\top\boldsymbol{\Sigma}^{-1}\left(\frac{2}{\alpha}\boldsymbol{V}\right)\right] + \mathrm{E}\left[\frac{1}{\alpha^2}\mathbf{1}^\top\boldsymbol{\Sigma}^{-1}\mathbf{1} - \frac{1}{4}\sum_{i=1}^{n}1\right] = \frac{1}{\alpha^2}\mathbf{1}^\top\boldsymbol{\Sigma}^{-1}\mathbf{1}. \end{aligned}$$

Além disso, $\boldsymbol{I}_{\boldsymbol{\varphi}\mu} = \boldsymbol{I}_{\mu\boldsymbol{\varphi}}^\top = (I_{\varphi_1\mu}, I_{\varphi_2\mu}, I_{\varphi_3\mu})^\top$ é um vetor 3×1 com elementos dados por

$$I_{\mu\varphi_i} = \mathrm{E}[-\ddot{\ell}_{\mu\varphi_i}] \approx \frac{1}{\alpha}\mathrm{E}\left[\left(\mathbf{1}^\top\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_i}\boldsymbol{\Sigma}^{-1}\right)\left(\frac{2}{\alpha}\boldsymbol{V}\right)\right] = 0, \quad i = 1,2,3.$$

5.1.8.4 Apêndice IV: Vetor escore $U(\omega)$ e matriz $G(\omega)$

O vetor escore usado no método de influência local é dado por $\boldsymbol{U}(\boldsymbol{\omega}) = \left(\partial\boldsymbol{V}_{\boldsymbol{\omega}}/\partial\boldsymbol{\omega}^\top\right)^\top\left(2\boldsymbol{\Sigma}^{-1}\boldsymbol{V}_{\boldsymbol{\omega}}\right) + (1/2)\boldsymbol{A}\boldsymbol{T}_{\boldsymbol{\omega}}$, em que $\boldsymbol{T}_{\boldsymbol{\omega}} = (\boldsymbol{T}_{\omega_1},\ldots,\boldsymbol{T}_{\omega_n})$, com $\boldsymbol{T}_{\omega_i} = \tanh((y_i + \boldsymbol{A}_i\boldsymbol{\omega} - \mu)/2)$, e $\boldsymbol{A}_i$ a i-ésima linha da matriz $\boldsymbol{A}$. Considerando $\cosh((y_i + \boldsymbol{a}_i\boldsymbol{\omega} - \mu)/2) \approx 1$, para $i = 1,\ldots,n$, tem-se

$$\left(\frac{\partial\boldsymbol{V}_{\boldsymbol{\omega}}}{\partial\boldsymbol{\omega}^\top}\right) = \frac{1}{2}\begin{pmatrix} \cosh\left(\frac{y_1+\boldsymbol{A_1}\boldsymbol{\omega}-\mu}{2}\right)\sigma_{11} & \cdots & \cosh\left(\frac{y_1+\boldsymbol{A_1}\boldsymbol{\omega}-\mu}{2}\right)\sigma_{1n} \\ \vdots & \ddots & \vdots \\ \cosh\left(\frac{y_n+\boldsymbol{A_n}\boldsymbol{\omega}-\mu}{2}\right)\sigma_{n1} & \cdots & \cosh\left(\frac{y_n+\boldsymbol{A_n}\boldsymbol{\omega}-\mu}{2}\right)\sigma_{nn} \end{pmatrix} = \frac{1}{2}\boldsymbol{A},$$

em que σ_{ij} é o (i,j) elemento da matriz $\boldsymbol{A}$, para $i,j = 1,\ldots,n$. Logo, como $\cosh((y_i + \boldsymbol{a}_i\boldsymbol{\omega} - \mu)/2) \approx 1$, com $i = 1,\ldots,n$, $\boldsymbol{T}_{\boldsymbol{\omega}}$ pode ser aproximado por $\boldsymbol{V}_{\boldsymbol{\omega}}$, de onde segue que $\boldsymbol{U}(\boldsymbol{\omega}) = (1/2)\boldsymbol{A}\boldsymbol{V}_{\boldsymbol{\omega}} - (2/\alpha^2)\boldsymbol{A}\boldsymbol{\Sigma}^{-1}\boldsymbol{V}_{\boldsymbol{\omega}}$. Então,

$$\begin{aligned} \boldsymbol{U}(\boldsymbol{\omega})\boldsymbol{U}^\top(\boldsymbol{\omega}) &= \left(\frac{1}{2}\boldsymbol{A}\boldsymbol{V}_{\boldsymbol{\omega}} - \frac{2}{\alpha^2}\boldsymbol{A}\boldsymbol{\Sigma}^{-1}\boldsymbol{V}_{\boldsymbol{\omega}}\right)\left(\frac{1}{2}\boldsymbol{A}\boldsymbol{V}_{\boldsymbol{\omega}} - \frac{2}{\alpha^2}\boldsymbol{A}\boldsymbol{\Sigma}^{-1}\boldsymbol{V}_{\boldsymbol{\omega}}\right)^\top \\ &= \frac{\alpha^2}{16}\boldsymbol{A}\left(\frac{2}{\alpha}\boldsymbol{V}_{\boldsymbol{\omega}}\right)\left(\frac{2}{\alpha}\boldsymbol{V}_{\boldsymbol{\omega}}\right)^\top\boldsymbol{A} - \frac{1}{4}\boldsymbol{A}\left(\frac{2}{\alpha}\boldsymbol{V}_{\boldsymbol{\omega}}\right)\left(\frac{2}{\alpha}\boldsymbol{V}_{\boldsymbol{\omega}}\right)^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{A} \\ &\quad -\frac{1}{4}\boldsymbol{A}\boldsymbol{\Sigma}^{-1}\left(\frac{2}{\alpha}\boldsymbol{V}_{\boldsymbol{\omega}}\right)\left(\frac{2}{\alpha}\boldsymbol{V}_{\boldsymbol{\omega}}\right)^\top\boldsymbol{A} + \frac{1}{\alpha^2}\boldsymbol{A}\boldsymbol{\Sigma}^{-1}\left(\frac{2}{\alpha}\boldsymbol{V}_{\boldsymbol{\omega}}\right)\left(\frac{2}{\alpha}\boldsymbol{V}_{\boldsymbol{\omega}}\right)^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{A}. \end{aligned}$$

Portanto, $\boldsymbol{G}(\boldsymbol{\omega}) = \mathrm{E}\left[\boldsymbol{U}(\boldsymbol{\omega})\boldsymbol{U}^\top(\boldsymbol{\omega})\right] = \boldsymbol{A}\left(\frac{\alpha}{4}\boldsymbol{\Sigma}^{(1/2)} - (1/\alpha)\boldsymbol{\Sigma}^{-(1/2)}\right)^2\boldsymbol{A}$. Para encontrar a perturbação apropriada, de acordo com a metodologia proposta por Zhu et al. (2007), é necessário encontrar $\boldsymbol{A}$, tal que $\boldsymbol{G}(\boldsymbol{\omega}) = c\boldsymbol{I}_n$, for $c > 0$. Considerando $c = 1$, então $\boldsymbol{A}$ deve satisfazer $((\alpha/4)\boldsymbol{\Sigma}^{(1/2)} - (1/\alpha)\boldsymbol{\Sigma}^{-\frac{1}{2}})^2 = (\boldsymbol{A}^{-1})^2$. Uma solução da equação acima é $\boldsymbol{A} = ((\alpha/4)\boldsymbol{\Sigma}^{\frac{1}{2}} - (1/\alpha)\boldsymbol{\Sigma}^{-\frac{1}{2}})^{-1}$. Então, $\tilde{\boldsymbol{\omega}} = ((\alpha/4)\boldsymbol{\Sigma}^{\frac{1}{2}} - (1/\alpha)\boldsymbol{\Sigma}^{-\frac{1}{2}})^{-1}\boldsymbol{\omega}$ é uma perturbação apropriada para o modelo BS log-linear espacial.

5.1.8.5 Apêndice V: A matriz perturbação

A matriz perturbação para o modelo BS log-linear espacial obtido de (5.56) é dada por

$$\Delta_\mu = \frac{\partial\ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial\mu\partial\boldsymbol{\omega}^\top} = \frac{\partial\boldsymbol{V}_{\boldsymbol{\omega}}^\top}{\partial\mu}\left(\frac{-2}{\alpha^2}\boldsymbol{\Sigma}^{-1} + \frac{1}{2}\boldsymbol{I_n}\right)\boldsymbol{A} = -\frac{1}{2}\boldsymbol{U}_{\boldsymbol{\omega}}^\top\left(\frac{-2}{\alpha^2}\boldsymbol{\Sigma}^{-1} + \frac{1}{2}\boldsymbol{I_n}\right)\boldsymbol{A},$$

em que $\boldsymbol{U_\omega} = (\boldsymbol{U}_{\omega_1}, \ldots, \boldsymbol{U}_{\omega_n})$, com $\boldsymbol{U}_{\omega_i} = \cosh((y_i + \boldsymbol{a_i\omega} - \mu)/2)$, para $i = 1, \ldots, n$. Então, os resultados apresentados em (5.32), (5.33) e (5.34) são obtidos considerando, $\cosh((y_i + \boldsymbol{a_i\omega} - \mu)/2) \approx 1$,

$$
\begin{aligned}
\Delta_\mu &= \frac{\partial \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \mu \partial \boldsymbol{\omega}^\top} \approx \mathbf{1}^\top \left(\frac{1}{\alpha^2}\boldsymbol{\Sigma}^{-1} - \frac{1}{4}\boldsymbol{I_n} \right) \boldsymbol{A}, \\
\Delta_\alpha &= \frac{\partial \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \alpha \partial \boldsymbol{\omega}^\top} = \frac{\partial \boldsymbol{V}_{\boldsymbol{\omega}}^\top}{\partial \alpha} \left(-\frac{2}{\alpha^2}\boldsymbol{\Sigma}^{-1} + \frac{1}{2}\boldsymbol{I_n} \right) \boldsymbol{A} + \boldsymbol{V}_{\boldsymbol{\omega}}^\top \frac{\partial}{\partial \alpha} \left[\left(\frac{-2}{\alpha^2}\boldsymbol{\Sigma}^{-1} + \frac{1}{2}\boldsymbol{I_n} \right) \boldsymbol{A} \right] \\
&= \boldsymbol{D}^\top \left(\frac{-1}{\alpha^2}\boldsymbol{\Sigma}^{-1} + \frac{1}{4}\boldsymbol{I_n} \right) \boldsymbol{A} + \left(\frac{4}{\alpha^3}\boldsymbol{\Sigma}^{-1}\boldsymbol{A} + \left(-\frac{2}{\alpha^2}\boldsymbol{\Sigma}^{-1} + \frac{1}{2}\boldsymbol{I_n} \right) \boldsymbol{A} \left(\frac{1}{\alpha^2}\boldsymbol{\Sigma}^{-\frac{1}{2}} + \frac{1}{4}\boldsymbol{\Sigma}^{\frac{1}{2}} \right) \boldsymbol{A} \right),
\end{aligned}
$$

em que $\boldsymbol{D} = (d_1, \ldots, d_n)^\top$, com $d_i = \boldsymbol{l}_i\omega$ e $\boldsymbol{l}_i$ sendo a i-ésima linha da matriz $\boldsymbol{A}\left(\frac{1}{\alpha^2}\boldsymbol{\Sigma}^{-\frac{1}{2}} + \frac{1}{4}\boldsymbol{\Sigma}^{\frac{1}{2}} \right) \boldsymbol{A}$. Em adição,

$$
\begin{aligned}
\Delta_{\varphi_i} &= \frac{\partial \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \boldsymbol{\varphi_i} \partial \boldsymbol{\omega}^\top} = \frac{\partial \boldsymbol{V}_{\boldsymbol{\omega}}^\top}{\partial \varphi_i} \left(-\frac{2}{\alpha^2}\boldsymbol{\Sigma}^{-1} + \frac{1}{2}\boldsymbol{I_n} \right) \boldsymbol{A} \\
&\quad + \boldsymbol{V}_{\boldsymbol{\omega}}^\top \left(\frac{2}{\alpha^2}\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_i}\boldsymbol{\Sigma}^{-1}\boldsymbol{A} + \left(-\frac{2}{\alpha^2}\boldsymbol{\Sigma}^{-1} + \frac{1}{2}\boldsymbol{I_n} \right) \frac{\partial \boldsymbol{A}}{\partial \varphi_i} \right) \\
&= \boldsymbol{M}^\top \left(\frac{-1}{\alpha^2}\boldsymbol{\Sigma}^{-1} + \frac{1}{4}\boldsymbol{I_n} \right) \boldsymbol{A} + \boldsymbol{V}_{\boldsymbol{\omega}}^\top \left(\frac{2}{\alpha^2}\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_i}\boldsymbol{\Sigma}^{-1} \right) \boldsymbol{A} \\
&\quad + \left(\frac{-2}{\alpha^2}\boldsymbol{\Sigma}^{-1} + \frac{1}{2}\boldsymbol{I_n} \right) \boldsymbol{A} \left(\frac{1}{\alpha}\boldsymbol{\Sigma}^{-\frac{1}{2}}\frac{\partial \boldsymbol{\Sigma}^{\frac{1}{2}}}{\partial \varphi_i}\boldsymbol{\Sigma}^{-\frac{1}{2}} + \frac{\alpha}{4}\frac{\partial \boldsymbol{\Sigma}^{\frac{1}{2}}}{\partial \varphi_i} \right) \boldsymbol{A},
\end{aligned}
$$

em que $\boldsymbol{M} = (m_1, \ldots, m_n)^\top$, com $m_i = \boldsymbol{L}_i\boldsymbol{\omega}$ e $\boldsymbol{L}_i$ sendo a i-ésima linha da matriz

$$
\boldsymbol{A} \left(-\frac{1}{\alpha}\boldsymbol{\Sigma}^{-\frac{1}{2}}\frac{\partial \boldsymbol{\Sigma}^{\frac{1}{2}}}{\partial \varphi_i}\boldsymbol{\Sigma}^{-\frac{1}{2}} + \frac{\alpha}{4}\frac{\partial \boldsymbol{\Sigma}^{\frac{1}{2}}}{\partial \varphi_i} \right) \boldsymbol{A}, \quad j = 1, \ldots, n, i = 1, 2, 3.
$$

Detalhes sobre $\partial \boldsymbol{\Sigma}^{\frac{1}{2}}/\partial \varphi_i$ podem ser encontrados em De Bastiani et al. (2015).

5.2 ARTIGO 2: Krigagem com *drift* externo em um modelo geoestatístico Birnbaum-Saunders

Resumo: A utilização de modelos estatísticos para descrever a dependência espacial em dados georreferenciados são comuns em muitas áreas, incluindo ciências ambientais. Muitas dessas aplicações assumem que os dados seguem uma distribuição Gaussiana. Contudo, em muitas delas, a suposição de normalidade e até mesmo uma suposição mais geral de simetria não são apropriadas. Em aplicações não espaciais, nos casos em que o conjunto de dados é unimodal e assimétrico positivo, a distribuição Birnbaum-Saunders tem se destacado. Este artigo propôs um modelo espacial log-linear e um estudo de interpolação por krigagem com *drift* externo, baseado na distribuição Birnbaum-Saunders. Os parâmetros do modelo foram estimados a partir do método da máxima verossimilhança. Como ilustração, o modelo proposto foi usado para analisar um conjunto de dados agrícolas provenientes de uma situação real. Tendo vista o fato que estudos das propriedades químicas do solo são extremamente importante para refinar práticas de gestão agrícola e para avaliar os efeitos da agricultura na qualidade ambiental; que nos estudos das propriedades químicas, além de analisar as propriedades de cada elemento individualmente, é importante considerar as relações entre esses elementos, e que uma das relações mais conhecida no meio agronômico é a relação Cálcio-Magnésio (Ca:Mg), no estudo com dados reais, a concentração de Magnésio no solo foi tomada como variável-resposta e a concentração de Cálcio no solo foi utilizada na explicação desta variável-resposta. Um estudo de gráficos quantil-quantil mostrou que o modelo espacial BS se ajustou bem ao conjunto de dados, e uma comparação entre os métodos de predição espacial krigagem com drift externo e krigagem ordinária mostrou que a krigagem com *drift* externo é mais eficiente na predição dos dados do que a krigagem ordinária.
Palavras-chave: Análise de dados espaciais, distribuições assimétricas, método da máxima verossimilhanaça, modelo Matérn, não normalidade, *software* R.

5.2.1 Introdução

O conhecimento da distribuição espacial de dados georrefenciados é de grande interesse de várias áreas tais como geografia, ecologia e engenharias (CRESSIE, 2015). Particularmente, modelos geoestatísticos são frequentemente usados em ciências ambientais (WEBSTER; OLIVER, 2009). Por exemplo, Cambardella et al. (1994) argumentaram que o conhecimento da distribuição espacial de atributos do solo é importante para o refinamento de práticas de manejo e para a avaliação dos efeitos da agricultura sobre a qualidade ambiental. Modelar a variabilidade espacial de dados georreferenciados permite estimar os parâmetros que definem a estrutura de dependência espacial. Esta variabilidade pode ser descrita por diferentes modelos de funções semivariâncias, os quais podem ser selecionados por validação cruzada, quando se comparam valores amostrados com os valores determinados pelo modelo geoestatístico (DE BASTIANI et al., 2015). Valores da variável sob estudo (variável-resposta) em localizações não amostradas são frequentemente interpolados em geoestatística pelo método krigagem (KRIGE,

1951; ISAAKS; SRIVASTAVA, 1989).

Com o objetivo de melhorar o poder preditivo, modelos espaciais podem considerar uma ou mais variáveis explanatórias (covariáveis) para descrever a variável-resposta. Por exemplo, Hu et al. (2015) estudaram a variabilidade espacial de arsênio no solo e a associação desse com o nitrogênio no solo em sistemas de produção intensiva. Valores da variável-resposta em localizações não amostradas também podem ser determinados por krigagem com *drift* externo, uma técnica de regressão espacial que combina krigagem e regressão (regressão-krigagem) (GOOVAERTS, 1999; SOARES, 2000). Um considerável número de autores relataram que a regressão-krigagem fornece o melhor estimador linear não viesado e é superior a outras técnicas de predição espacial (BISHOP; MCBRATNEY, 2001; HENGL; HEUVELINK; STEIN, 2004).

A maioria dos modelos espaciais propostos, até o momento, foram baseados em suposições de normalidade. Entretanto, essas suposições não satisfazem a diversos fenômenos. Assumpção, Uribe-Opazo e Galea (2014) conduziram um estudo espacial da produtividade da soja usando um modelo geoestatístico com uma variável-resposta com distribuição t-Student, que é uma distribuição simétrica. Pesquisas com dados georreferenciados usando distribuições assimétricas têm sido pouco exploradas na literatura. Em modelos não espaciais, as distribuições Birnbaum-Saunders (BS), exponencial, gamma, log-normal e Weibull têm sido frequentemente consideradas para descrever dados assimétricos (JOHNSON; KOTZ; BALAKRISHNAN, 1994, 1995; LEIVA, 2015). A distribuição BS está sendo amplamente estudada e aplicada, inclusive nas ciências ambientais (LEIVA et al., 2008, 2015; LEIVA; SANHUEZA; ANGULO, 1999; PODLASKI, 2008; FERREIRA; GOMES; LEIVA, 2012; MARCHANT et al., 2013; SAULO et al., 2013). Em contraste com sua aplicação original em processos de fadiga, Leiva et al. (2015a) justificaram porque o modelo BS é adequado para descrever dados ambientais usando argumentos teóricos baseados na lei dos efeitos proporcionais. Xia, Zeephongsekul e Packer (2011) apresentaram um método para modelos espaciais BS para estudar movimentos de turistas. Garcia-Papani et al. (2016) propuseram um modelo espacial e diagnósticos baseados na distribuição BS e os aplicaram aos dados de engenharia agrícola.

A estimação dos parâmetros do modelo log-linear espacial BS, baseado no método da máximo verossimilhança (ML), necessita de uma versão logarítma multivariada da distribuição BS (log-BS) (MARCHANT; LEIVA; CYSNEIROS, 2015; GARCIA-PAPANI et al., 2016). Até o momento, estudos envolvendo regressão-krigagem em modelos espaciais BS não foram considerados.

O principal objetivo deste artigo foi propor uma metodologia baseada em modelos espaciais BS considerando covariáveis. Especificamente, os parâmetros do modelo foram estimados pelo método ML. Adicionalmente, a estrutura de dependência espacial foi descrita pelos modelos da família Matérn, e o modelo mais adequado foi selecionado por validação cruzada. Além disso, valores da variável-resposta em localizações não amostradas foram determinados por regressão-krigagem.

O restante deste artigo é organizado como segue. A Seção 5.2.2 apresenta a definição e as principais propriedades das distribuições BS e log-BS multivariadas e a formulação do modelo BS espacial log-linear com *drift* externo. Na Seção 5.2.3, estimaram-se os parâmetros do modelo pelo método ML, discutiu-se a validação do modelo, e descreveu-se a técnica de interpolação espacial conhecida como krigagem com *drift* externo. Na Seção 5.2.4, ilustrou-se o modelo espacial por análise de dados experimentais. Na Seção 5.2.5, podem ser observadas as conclusões e ideias para trabalhos futuros. Os cálculos mais elaborados foram apresentados nos Apêndices.

5.2.2 Modelo espacial Birnbaum-Saunders

Nesta Seção, são apresentadas as distribuições BS e log-BS multivariadas. Além disso, formula-se um modelo espacial BS considerando a média como uma combinação linear de covariáveis.

5.2.2.1 Distribuições BS e log-BS multivariadas

Uma variável aleatória contínua $T_i > 0$ tem distribuição BS com parâmetros de forma ($\alpha_i > 0$) e escala ($\beta_i > 0$), denota-se por $T_i \sim \mathrm{BS}(\alpha_i, \beta_i)$, se (veja detalhes da distribuição BS em Leiva, (2015))

$$Z_i = \frac{1}{\alpha_i}\left(\sqrt{\frac{T_i}{\beta_i}} - \sqrt{\frac{\beta_i}{T_i}}\right) \sim \mathrm{N}(0,1), \quad \text{em que} \quad T_i = \beta_i\left(\frac{\alpha_i Z_i}{2} + \sqrt{\left(\frac{\alpha_i Z_i}{2}\right)^2 + 1}\right)^2, \quad i = 1, \ldots, n. \tag{5.36}$$

Seja o vetor aleatório $\boldsymbol{T} = (T_1, \ldots, T_n)^\top \in \mathbb{R}^n_+$ com distribuição BS n-variada, com vetor de forma $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)^\top \in \mathbb{R}^n_+$, vetor escala $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_n)^\top \in \mathbb{R}^n_+$ e matriz escala $\boldsymbol{\Sigma} \in \mathbb{R}^{n\times n}_+$, com $\mathrm{rk}(\boldsymbol{\Sigma}) = n$. A notação neste caso será $\boldsymbol{T} \sim \mathrm{BS}_n(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Sigma})$. Note que $\boldsymbol{\Sigma} = (\sigma_{ij})$ é a matriz, $n \times n$, de variância e covariância do vetor aleatório $\boldsymbol{Z} = (Z_1, \ldots, Z_n)^\top \sim \mathrm{N}_n(\boldsymbol{0}, \boldsymbol{\Sigma})$, com $\sigma_{ij} = 1$, se $i = j$, enquanto $\sigma_{ij} = \mathrm{Cov}(Z_i, Z_j)$, se $i \neq j$, isto é, a covariância entre Z_i e Z_j como dados em (5.36), para $i, j = 1, \ldots, n$. Aqui, $\boldsymbol{0}$ representa um vetor de zeros, $n \times 1$. Assim, se $\boldsymbol{T} \sim \mathrm{BS}_n(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Sigma})$, então sua função densidade de probabilidade (f.d.p.) é

$$f_{\boldsymbol{T}}(\boldsymbol{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Sigma}) = \phi_n\left(\frac{1}{\alpha_1}\left(\sqrt{\frac{t_1}{\beta_1}} - \sqrt{\frac{\beta_1}{t_1}}\right), \ldots, \frac{1}{\alpha_n}\left(\sqrt{\frac{t_n}{\beta_n}} - \sqrt{\frac{\beta_n}{t_n}}\right); \boldsymbol{\Sigma}\right) \times \tag{5.37}$$

$$\prod_{i=1}^{n} \frac{1}{2\alpha_i\beta_i}\left(\sqrt{\frac{\beta_i}{t_i}} + \sqrt{\frac{\beta_i^3}{t_i^3}}\right), \boldsymbol{t} \in \mathbb{R}^n_+, \tag{5.38}$$

em que $\phi_n(\cdot; \boldsymbol{\Sigma})$ é a f.d.p. da distribuição normal-padrão n-variada. O vetor aleatório $\boldsymbol{T}$ satisfaz as seguintes propriedades: (i) $\mathrm{Cov}(T_i, T_j) = \gamma_i\gamma_j\left(\left(2 + \frac{3\alpha_i^2}{2} + \frac{3\alpha_j^2}{2} + \frac{9\alpha_i^2\alpha_j^2}{2^5}\right)\rho_{ij} + \frac{\alpha_i^2\alpha^j}{2}\rho_{ij}^2 + \frac{3\alpha_i^2\alpha_j^2}{2^4}\rho_{ij}^3 + r(Z_i, Z_j)\right)$, com ρ_{ij} a covariância entre Z_i e Z_j e $r(Z_i, Z_j)$ é dado no Apêndice I, para $i, j = 1, \ldots, n$; (ii) considerem-se as partições

$$\boldsymbol{T} = (\boldsymbol{T_1}^\top, \boldsymbol{T_2}^\top)^\top, \quad \boldsymbol{\alpha} = (\boldsymbol{\alpha_1}^\top, \boldsymbol{\alpha_2}^\top)^\top, \quad \boldsymbol{\beta} = (\boldsymbol{\beta_1}^\top, \boldsymbol{\beta_2}^\top)^\top, \quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma_{11}} & \boldsymbol{\Sigma_{12}} \\ \boldsymbol{\Sigma_{21}} & \boldsymbol{\Sigma_{22}} \end{pmatrix},$$

em que $\boldsymbol{T_1}, \boldsymbol{\alpha_1}, \boldsymbol{\beta_1}$ são vetores $n_1 \times 1$, $\boldsymbol{T_2}, \boldsymbol{\alpha_2}, \boldsymbol{\beta_2}$ são vetores $n_2 \times 1$, $\boldsymbol{\Sigma_{11}}$ é uma matriz $n_1 \times n_1$, $\boldsymbol{\Sigma_{22}}$ é uma matriz $n_2 \times n_2$, e $\boldsymbol{\Sigma_{12}} = \boldsymbol{\Sigma_{21}}^\top$ é uma matriz $n_1 \times n_2$, com $n_1 + n_2 = n$; então, (i) $\boldsymbol{T}_l \sim \mathrm{BS}_{n_l}(\boldsymbol{\alpha}_l, \boldsymbol{\beta}_l; \boldsymbol{\Sigma}_{ll})$, para $l = 1, 2$; (iii) se $c > 0$, então $c\boldsymbol{T} \sim \mathrm{BS}_n(\boldsymbol{\alpha}, c\boldsymbol{\beta}; \boldsymbol{\Sigma})$ e $(c\boldsymbol{T_1}^\top, \boldsymbol{T_2}^\top)^\top \sim \mathrm{BS}_n(\boldsymbol{\alpha}, (c\boldsymbol{\beta_1}^\top, \boldsymbol{\beta_2}^\top)^\top; \boldsymbol{\Sigma})$, uma propriedade análoga se aplica a $(\boldsymbol{T_1}^\top, c\boldsymbol{T_2})^\top$; e (iv) considere a notação $\boldsymbol{A}^{-1} = (1/A_i)$, em que $\boldsymbol{A} = (A_i)$ é um vetor e A_i seus elementos, então $\boldsymbol{T}^{-1} = ((\boldsymbol{T_1}^{-1})^\top, (\boldsymbol{T_2}^{-1})^\top)^\top \sim \mathrm{BS}_n(\boldsymbol{\alpha}, \boldsymbol{\beta}^{-1}; \boldsymbol{\Sigma})$ e $((\boldsymbol{T_1}^{-1})^\top, \boldsymbol{T_2}^\top)^\top \sim \mathrm{BS}_n(\boldsymbol{\alpha}, ((\boldsymbol{\beta_1}^{-1})^\top, \boldsymbol{\beta_2}^\top)^\top; \boldsymbol{\Sigma}_1)^\top$, em que $\boldsymbol{\Sigma}_1 = \begin{pmatrix} \boldsymbol{\Sigma_{11}} & -\boldsymbol{\Sigma_{12}} \\ -\boldsymbol{\Sigma_{21}} & \boldsymbol{\Sigma_{22}} \end{pmatrix}$, propriedade análoga se aplica a $\boldsymbol{T_2}^{-1}$.

Sejam $\boldsymbol{T} = (T_1, \ldots, T_n)^\top \sim \mathrm{BS}_n(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Sigma})$, $Y_i = \log(T_i)$ e $\mu_i = \log(\beta_i)$, para $i = 1, \ldots, n$, com $-\infty < Y_i, \mu_i < +\infty$. Então, o vetor aleatório $\boldsymbol{Y} = (Y_1, \ldots, Y_n)^\top$ tem distribuição log-BS n-variada, com vetor de forma $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)^\top$, vetor de locação $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)^\top$, e matriz escala $\boldsymbol{\Sigma}$. Isto é indicado por $\boldsymbol{Y} \sim \text{log-BS}_n(\boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\Sigma})$, e sua pdf é

$$f_{\boldsymbol{Y}}(\boldsymbol{y}; \boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \phi_n\left(\frac{2}{\alpha_1}\mathrm{senh}\left(\frac{y_1 - \mu_1}{2}\right), \ldots, \frac{2}{\alpha_n}\mathrm{senh}\left(\frac{y_n - \mu_n}{2}\right); \boldsymbol{\Sigma}\right) \times \tag{5.39}$$

$$\prod_{i=1}^{n} \frac{1}{\alpha_i}\cosh\left(\frac{y_i - \mu_i}{2}\right), \boldsymbol{y} \in \mathbb{R}^n, \tag{5.40}$$

em que $\phi_n(\cdot; \boldsymbol{\Sigma})$ é dado em (5.37) e as seguintes propriedades são satisfeitas: (i) $Y_i \sim \text{log-BS}(\alpha_i, \mu_i)$, para $i = 1, \ldots, n$; (ii) se $\boldsymbol{a} = (a_1, \ldots, a_n)^\top$ é um vetor constante, $\boldsymbol{Y} + \boldsymbol{a} \sim \text{log-BS}_n(\boldsymbol{\alpha}, \boldsymbol{\mu} + \boldsymbol{a}; \boldsymbol{\Sigma})$; (iii) $\mathrm{E}(\boldsymbol{Y}) = \boldsymbol{\mu}$; (iv) se $\boldsymbol{\alpha} = \alpha \boldsymbol{1}_{n\times 1}^\top$ e $\boldsymbol{\mu} = \boldsymbol{0}_{n\times 1}$, $\mathrm{Cov}(\boldsymbol{Y}) \approx \alpha^2 \boldsymbol{\Sigma}$, com $\boldsymbol{1}_{n\times 1}$ sendo um vetor de uns, $n \times 1$; e (v) se $\boldsymbol{Y} \sim \text{log-BS}_n(\alpha \boldsymbol{1}_{n\times 1}, \boldsymbol{\mu}; \boldsymbol{\Sigma})$, então $((2/\alpha)\boldsymbol{V}^\top)\boldsymbol{\Sigma}^{-1}((2/\alpha)\boldsymbol{V})$ tem uma distribuição χ^2 com n graus de liberdade, em que $\boldsymbol{V} = (V_1, \ldots, V_n)^\top$, com $V_j = \mathrm{senh}((y_j - \boldsymbol{x}_j^\top \widehat{\boldsymbol{\beta}})/2)$, para $j = 1, \ldots, n$.

5.2.2.2 Formulação do modelo de regressão

Para modelar um conjunto de dados espacialmente correlacionados, considera-se um processo estocástico $\boldsymbol{T} = \{T(\boldsymbol{s}), \boldsymbol{s} \in D\}$, com $D \subset \mathbb{R}^2$. Assume-se que o processo estocástico $\boldsymbol{T}$ é estacionário e isotrópico. Adicionalmente, assume-se que para localizações espaciais conhecidas $\boldsymbol{s_i}$, com $i = 1, \ldots, n$, o valor resposta em $\boldsymbol{s_i}$, $T(\boldsymbol{s_i})$, é dado por

$$T(\boldsymbol{s_i}) = \exp(\mu(\boldsymbol{s_i}))\, \eta(\boldsymbol{s_i}), \quad i = 1, \ldots, n,$$

em que $\mu(\boldsymbol{s_i}) = \boldsymbol{x_i}^\top \boldsymbol{\beta}$, com $\boldsymbol{x_i} = (x_{i0}, \ldots, x_{ip})^\top$ um vetor $(p+1) \times 1$ com $x_{i0} = 1$, e $x_{ij} = x_j(\boldsymbol{s_i})$ para $j = 1, \ldots, p$, isto é, x_{ij} é o valor da covariável X_j na localização $\boldsymbol{s_i}$; $\boldsymbol{\beta} = (\beta_0, \ldots, \beta_p)^\top$ é um vetor de parâmetros desconhecidos a ser estimado; e $\eta(\boldsymbol{s_i})$ é o erro aleatório do modelo em $\boldsymbol{s_i}$. Supõe-se que $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_n)^\top \sim \mathrm{BS}_n(\alpha \boldsymbol{1}_{n\times 1}, \boldsymbol{1}_{n\times 1}, \boldsymbol{\Sigma})$, com $\alpha > 0$. O modelo espacial definido em (5.2.2.2) pode ser linearizado por aplicação de logaritmo da seguinte forma

$$Y(\boldsymbol{s_i}) = \log(T(\boldsymbol{s_i})) = \boldsymbol{x_i}^\top \boldsymbol{\beta} + \varepsilon(\boldsymbol{s_i}), \quad i = 1, \ldots, n,$$

em que $\varepsilon(\boldsymbol{s_i}) = \log(\eta(\boldsymbol{s_i}))$. Equivalentemente, em notação matricial, o modelo espacial log-linear BS definido em (5.2.2.2) pode ser escrito como

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{5.41}$$

em que $\boldsymbol{Y} = (Y(\boldsymbol{s_1}), \ldots, Y(s_n))^\top$ é o vetor das respostas, $n \times 1$, com $Y(\boldsymbol{s_i}) = \log(T(\boldsymbol{s_i}))$; $\boldsymbol{X} = \left(\boldsymbol{x_0}^\top, \ldots, \boldsymbol{x_n}^\top\right)^\top$ é uma matriz, $n \times (p+1)$, com os valores das covariáveis; e $\boldsymbol{\varepsilon} = (\varepsilon(\boldsymbol{s_1}), \ldots, \varepsilon(s_n))^\top$ é o vetor dos erros aleatórios do modelo espacial, $n \times 1$, com $\boldsymbol{\varepsilon} \sim \text{log-BS}_n(\alpha \mathbf{1}_{n\times 1}, \mathbf{0}_{n\times 1}; \boldsymbol{\Sigma})$ e, portanto, $\mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}_{n\times 1}$.

5.2.2.3 Formulação do modelo espacial

Assume-me que a dependência espacial é determinada por uma matriz escala, $n \times n$, $\boldsymbol{\Sigma}$, a qual é simétrica, não singular e positiva-definida, proporcional a $\mathrm{Cov}\,(\boldsymbol{\varepsilon})$. Além disso, assume-se que $\mathrm{Cov}\,(\varepsilon(\boldsymbol{s_i}), \varepsilon(\boldsymbol{s_j}))$ depende somente da distância Euclidiana entre as localizações $\boldsymbol{s_i}$ e $\boldsymbol{s_j}$, isto é, $\mathrm{Cov}\,(\varepsilon(\boldsymbol{s_i}), \varepsilon(\boldsymbol{s_j})) = \mathrm{C}(\boldsymbol{s_i}, \boldsymbol{s_j}) = \mathrm{C}\,(h_{ij})$, em $h_{ij} = \|\boldsymbol{s_i} - \boldsymbol{s_j}\|$. Assume-se ainda que a matriz escala tem a forma paramétrica

$$\boldsymbol{\Sigma} = \varphi_1 \boldsymbol{I}_n + \varphi_2 \boldsymbol{R}(\varphi_3),$$

em que φ_1 é o efeito pepita ou variância do erro, φ_2 é a contribuição ou variância da dispersão, φ_3 é uma função do alcance ou raio de dependência espacial do modelo, $\boldsymbol{I}_n$ é a matriz identidade $n \times n$ e $\boldsymbol{R} = (r_{ij})$ é uma matriz simétrica, $n \times n$, com elementos diagonais $r_{ii} = 1$, para $i = 1, \ldots, n$. Note que $\boldsymbol{R}$ depende do modelo teórico do variograma adotado para descrever a dependência espacial. Adotam-se neste estudo modelos família Matérn e então

$$r_{ij} = \begin{cases} 1, & i = j, \\ \frac{1}{2^{\delta-1}\Gamma(\delta)} \left(\frac{h_{ij}}{\varphi_3}\right)^{\delta} K_\delta \left(\frac{h_{ij}}{\varphi_3}\right), & i \neq j, \end{cases} \tag{5.42}$$

em que δ é um parâmetro de forma, $\Gamma(\cdot)$ é a função gamma usual e $K_\delta(\cdot)$ é a função modificada de Bessel do terceiro tipo de ordem δ. Assim, sob as especificações do modelo log-linear espacial BS, a função covariância é dada por

$$\mathrm{C}(h_{ij}) = \begin{cases} \alpha^2(\varphi_1 + \varphi_2), & i = j, \\ \frac{\alpha^2\varphi_2}{2^{\delta-1}\Gamma(\delta)} \left(\frac{h_{ij}}{\varphi_3}\right)^{\delta} K_\delta \left(\frac{h_{ij}}{\varphi_3}\right), & i \neq j. \end{cases}$$

Os modelos exponencial e Gaussiano são membros da família Matérn, quando $\delta = 0,5$ e $\delta \to \infty$, dados, respectivamente, por

$$\mathrm{C}(h_{ij}) = \begin{cases} \alpha^2(\varphi_1 + \varphi_2), & i = j, \\ \alpha^2\varphi_2 \exp(-(h_{ij}/\varphi_3)), & i \neq j, \end{cases}$$

e

$$\mathrm{C}(h_{ij}) = \begin{cases} \alpha^2(\varphi_1 + \varphi_2), & i = j, \\ \alpha^2\varphi_2 \exp(-(h_{ij}/\varphi_3)^2), & i \neq j. \end{cases}$$

5.2.3 Estimação, validação e predição

Nesta Seção, estimam-se os parâmetros do modelo espacial BS pelo método ML, discute-se a validação do modelo, e apresenta-se a krigagem com *drift* externo.

5.2.3.1 Estimação dos parâmetros

Devido a $\boldsymbol{\varepsilon} = \boldsymbol{Y} - \boldsymbol{X}^\top\boldsymbol{\beta} \sim \text{log-BS}_n\left(\alpha\mathbf{1}_{n\times 1}, \mathbf{0}_{n\times 1}; \boldsymbol{\Sigma}\right)$, o vetor de parâmetros $\boldsymbol{\theta} = \left(\boldsymbol{\beta}^\top, \boldsymbol{\varphi}^\top, \alpha\right)^\top$ do modelo espacial definido em (5.41) pode ser estimado ao se maximizar a função log-verossimilhança em relação a $\boldsymbol{\theta}$ baseado na f.d.p. dada em (5.39), a menos de uma constante, por

$$\ell(\boldsymbol{\theta}) = -\frac{1}{2}\log\left(|\boldsymbol{\Sigma}|\right) - n\log(\alpha) - \frac{2}{\alpha^2}\boldsymbol{V}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{V} + \sum_{i=1}^{n}\log\left(\cosh\left(\frac{y_i - \boldsymbol{x}_i^\top\boldsymbol{\beta}}{2}\right)\right), \tag{5.43}$$

em que $\boldsymbol{V} = (V_1, \ldots, V_n)^\top$ é um vetor $n \times 1$ com elementos $V_i = \text{senh}((y_i - \boldsymbol{x}_i^\top\boldsymbol{\beta})/2)$, para $i = 1, \ldots n$, e $\boldsymbol{\Sigma}$ é dada em (5.2.2.3).

Neste caso, o vetor escore $(p+5)\times 1$ é dado por

$$U\left(\boldsymbol{\theta}\right) = \left(\left(\frac{\partial\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\beta}}\right)^\top, \left(\frac{\partial\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\varphi}}\right)^\top, \frac{\partial\ell(\boldsymbol{\theta})}{\partial\alpha}\right)^\top = \left(U(\beta_0), \ldots, U(\beta_p), U(\varphi_1), U(\varphi_2), U(\varphi_3), U(\alpha)\right)^\top.$$

De (5.43), tem-se que

$$\begin{aligned}
U\left(\beta_j\right) &= \frac{\partial\ell(\boldsymbol{\theta})}{\partial\beta_j} = \frac{\partial}{\partial\beta_j}\left(-\frac{2}{\alpha^2}\boldsymbol{V}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{V} + \sum_{i=1}^{n}\log\left(\cosh\left(\frac{y_i - \boldsymbol{x}_i^\top\boldsymbol{\beta}}{2}\right)\right)\right) \\
&= \frac{2}{\alpha^2}\left(\boldsymbol{W}^\top \odot \boldsymbol{X}_j^\top\right)\boldsymbol{\Sigma}^{-1}\boldsymbol{V} - \frac{1}{2}\sum_{i=1}^{n}\tanh\left(\frac{y_i - \boldsymbol{x}_i^\top\boldsymbol{\beta}}{2}\right)x_{ij}, \quad j = 0, \ldots, p. \\
U\left(\varphi_j\right) &= \frac{\partial\ell(\boldsymbol{\theta})}{\partial\varphi_j} = \frac{\partial}{\partial\varphi_j}\left(-\frac{1}{2}\log\left(|\boldsymbol{\Sigma}|\right) - \frac{2}{\alpha^2}\boldsymbol{V}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{V}\right) \\
&= -\frac{1}{2}\text{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\varphi_j}\right) + \frac{2}{\alpha^2}\boldsymbol{V}^\top\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\varphi_j}\boldsymbol{\Sigma}^{-1}\boldsymbol{V}, \quad j = 1, 2, 3. \\
U\left(\alpha\right) &= \frac{\partial\ell(\boldsymbol{\theta})}{\partial\alpha} = \frac{\partial}{\partial\alpha}\left(-n\log(\alpha) - \frac{2}{\alpha^2}\boldsymbol{V}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{V}\right) = -\frac{n}{\alpha} + \frac{4}{\alpha^3}\boldsymbol{V}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{V}.
\end{aligned}$$

em que $\odot$ e $\text{tr}\left(\boldsymbol{A}\right)$ denotam o produto de Hadamard e o traço da matriz $\boldsymbol{A}$, respectivamente, $\boldsymbol{W} = (W_1, \ldots, W_n)^\top$, com $W_i = \cosh((y_i - \boldsymbol{x}_i^\top\boldsymbol{\beta})/2)$, para $i = 1, \ldots, n$, $\boldsymbol{V} = (V_1, \ldots, V_n)^\top$ é dado em (5.43), e $\boldsymbol{x}_j = (x_{1j}, \ldots, x_{nj})^\top$, para $j = 0, \ldots p$. Destaca-se que $x_{i0} = 1$, para $i = 1, \ldots, n$, e que

$$\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_1} = \boldsymbol{I}_n, \frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_2} = \boldsymbol{R}, \frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_3} = \varphi_2\frac{\partial\boldsymbol{R}}{\partial\varphi_3} = \varphi_2\boldsymbol{R}'.$$

Para o modelo Matérn dado em (5.42), $\boldsymbol{R}' = (\partial\boldsymbol{R}/\partial\varphi_3)$ tem elementos dados por

$$\frac{\partial r_{ij}}{\partial\varphi_3} = -\left(\frac{1}{\varphi_3}\right)\left(\delta r_{ij} + \frac{1}{2^{\delta-1}\Gamma(\delta)}\left(\frac{h_{ij}}{\varphi_3}\right)^{\delta+1}K_\delta'\left(\frac{h_{ij}}{\varphi_3}\right)\right), i \neq j, i, j = 1, \ldots, n,$$

com $K_\delta'(u) = -(1/2)(K_{\delta-1}(u) + K_{\delta+1}(u))$. Então, a fim de se obter $\widehat{\boldsymbol{\theta}}$, deve-se resolver o sistema não linear $\boldsymbol{U}\left(\boldsymbol{\theta}\right) = \mathbf{0}_{(p+5)\times 1}$. Como este sistema não tem solução analítica explícita,

o estimador $\widehat{\boldsymbol{\theta}}$ deve ser calculado usando um processo iterativo para solução de sistemas não lineares. Na aplicação da Seção 5.2.4, utilizou-se o método quase-Newton L-BFGS-B (NOCEDAL; WRIGHT, 1999; LANGE, 2001). Pode-se assintoticamente encontrar os erros padrões (SEs) dos estimadores dos parâmetros por meio da inversa da matriz de informação esperada, uma vez que

$$\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{D} \mathrm{N}_p(\mathbf{0}, \boldsymbol{J}(\boldsymbol{\theta})^{-1}),$$

quando $n \to \infty$, em que $\boldsymbol{J}(\boldsymbol{\theta}) = \lim_{n\to\infty}[1/n]\boldsymbol{I}(\boldsymbol{\theta})$, com $\boldsymbol{I}(\boldsymbol{\theta})$ sendo a matriz de informação esperada de Fisher e $\xrightarrow{D}$ denota convergência em distribuição. Para o modelo espacial BS, a matriz de informação esperada é a matriz bloco diagonal, $(p+5)\times(p+5)$, dada por

$$\boldsymbol{I}(\boldsymbol{\theta}) = \begin{pmatrix} \boldsymbol{I}_{\boldsymbol{\beta\beta}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{I}_{\boldsymbol{\varphi\varphi}} & \boldsymbol{I}_{\boldsymbol{\varphi}\alpha} \\ \mathbf{0} & \boldsymbol{I}_{\alpha\boldsymbol{\varphi}} & I_{\alpha\alpha} \end{pmatrix} = \begin{pmatrix} \boldsymbol{I}_{\boldsymbol{\beta\beta}} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{I}_{\boldsymbol{\psi\psi}} \end{pmatrix}$$

em que $\boldsymbol{I}_{\boldsymbol{\beta\beta}}$ denota a matriz de informação esperada para o vetor $\boldsymbol{\beta} = (\beta_0, \ldots, \beta_p)^\top$ e $\boldsymbol{I}_{\boldsymbol{\psi\psi}}$ é a matriz de informação esperada para $\boldsymbol{\psi} = \left(\boldsymbol{\varphi}^\top, \alpha\right)^\top$; veja detalhes no Apêndice III. Note que testes de hipóteses do tipo H_0: $\theta_i = 0$ *versus* H_1: $\theta_i \neq 0$, em que θ_i é qualquer elemento do vetor de parâmetros $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\varphi}^\top, \alpha)$, também podem ser conduzidos.

5.2.3.2 Validação do modelo

A propriedade (v) da distribuição $\log\text{-}\mathrm{BS}_n$ pode ser usada para validar o modelo em situações práticas como descrito a seguir. Seja $u_i = ((2/\alpha)\boldsymbol{v}_{(i)}^\top)\boldsymbol{\Sigma}^{-1}((2/\alpha)\boldsymbol{v}_{(i)})$, em que $\boldsymbol{v}_{(i)} = (v_{1(i)}, \ldots, v_{n(i)})^\top$, com $v_{j(i)} = \operatorname{senh}((y_j - \boldsymbol{x}_j^\top\widehat{\boldsymbol{\beta}}_{(i)})/2)$ e $\widehat{\boldsymbol{\beta}}_{(i)}$ a estimativa ML de $\boldsymbol{\beta}$ obtida usando os dados sem a observação i, para $i, j = 1, \ldots n$. Uma aproximação para $\widehat{\boldsymbol{\theta}}_{(i)}$ pode ser obtida por

$$\widehat{\boldsymbol{\theta}}_{(i)} = \widehat{\boldsymbol{\theta}} + (-\ddot{\ell}_{(i)}(\widehat{\boldsymbol{\theta}}))^{-1}\dot{\ell}_{(i)}(\widehat{\boldsymbol{\theta}}),$$

em que $\ddot{\ell}_{(i)}(\boldsymbol{\theta})$ e $\dot{\ell}_{(i)}(\boldsymbol{\theta})$ são a matriz Hessiana e o vetor escore associados com o modelo log-linear espacial BS. Sob a suposição de que $\boldsymbol{Y} \sim \mathrm{BS}_n(\alpha\mathbf{1}_{n\times1}, \mathbf{1}_{n\times1}; \boldsymbol{\Sigma})$, sabe-se que u_i apresenta distribuição χ^2 com $n-1$ graus de liberdade, para $i = 1, \ldots, n$. Usando a aproximação de Wilson-Hilferty (IBACACHE-PULGAR; PAULA; GALEA, 2014), tem-se que

$$Z_i = \frac{\left(\frac{u_i}{n-1}\right)^{1/3} - (1 - \frac{2}{9(n-1)})}{\left(\frac{2}{9(n-1)}\right)^{1/2}}, \quad i = 1, \ldots, n, \tag{5.44}$$

é uma variável aleatória com distribuição normal padrão. Então, um *Q-Q plot* dos z_i, definidos em (5.44), para $i = 1, \ldots, n$, pode ser usado para avaliar o ajuste do modelo. Em adição à aproximação de Wilson-Hilferty, o resíduo quantil randonizado definido por Dunn & Smyth (1996) também pode ser usado para avaliar o ajuste do modelo log-linear espacial BS. No caso deste modelo, estes resíduos são dados por

$$r_i = \Phi^{-1}\left(F(u_i)\right), \quad i = 1, \ldots, n, \tag{5.45}$$

em que $\Phi^{-1}(\cdot)$ é a função inversa da f.d.a. da distribuição N(0, 1) e $F(\cdot)$ é a f.d.a. da $\chi^2(n-1)$. Devido ao fato dos resíduos quantis randonizados terem distribuição normal padrão, um *Q-Q plot* dos r_i, definidos em (5.45), para $i = 1, \ldots, n$, pode ser empregado para avaliar o ajuste do modelo.

5.2.3.3 Predição espacial: Krigagem com *drift* externo

Um dos objetivos no estudo de dados espaciais é predizer a resposta em um ponto não amostrado da área estudada. A krigagem é um método de predição (interpolação) comumente usado em geoestatística. Quando supõe-se as médias do conjunto das variáveis aleatórias como uma combinação linear de variáveis auxiliares, tem-se um estimador, variante do estimador apresentado na krigagem ordinária, conhecido como Krigagem com *drift* externo (SOARES, 2000; HENGL; HEUVELINK; STEIN, 2003). A Krigagem com *drift* externo melhora a predição espacial baseada em padrões descritos por dados auxiliares (HENGL; HEUVELINK; STEIN, 2004). Neste caso, assim como na krigagem ordinária, a predição é dada por uma combinação linear dos dados observados $\boldsymbol{y} = (y_1, \ldots, y_n)^\top$, e os valores observados das covariáveis serão utilizados na determinação dos pesos λ_i, isto é,

$$\widehat{y}(s_0) = \sum_{i=1}^{n} \lambda_i \, y_i,$$

em que $\widehat{y}(s_0)$ é o valor predito em uma nova localização $s_0 \in D \subset \mathbb{R}^2$ e $\lambda_1, \ldots, \lambda_n$ são pesos escolhidos de forma que se possa obter o melhor preditor linear não viesado. Isto pode ser alcançado minimizando a variância dos erros com respeito aos pesos ao mesmo tempo ao se requerer que o preditor seja não viesado. Especificamente, temos que encontrar λ_i, para $i = 1, \ldots, n$, de forma que o erro seja não enviesado, isto é, $\mathrm{E}(\widehat{Y}(s_0) - Y(s_0)) = 0$ e a variância do erro minimizada. Para garantir a condição de não enviesamento, é necessário e suficiente assumir que

$$x_{0j} = \sum_{i=1}^{n} \lambda_i x_{ij}, \quad j = 1, \ldots, p, \tag{5.46}$$

uma vez que

$$\begin{aligned}
\mathrm{E}\left(\widehat{Y}(s_0) - Y(s_0)\right) &= \left(\sum_{i=1}^{n} \lambda_i \mathrm{E}\left(Y(\boldsymbol{s_i})\right)\right) - \mathrm{E}\left(Y(s_0)\right) \\
&= \left(\sum_{i=1}^{n} \lambda_i \sum_{j=0}^{p} \beta_j x_{ij}\right) - \sum_{j=0}^{p} \beta_j x_{0j} \\
&= \sum_{j=0}^{p} \beta_j \left(\sum_{i=1}^{n} \lambda_i x_{ij} - x_{0j}\right).
\end{aligned}$$

Para minimizar a variância do erro, primeiro, deve-se lembrar que $Y(\boldsymbol{s}) = \mu(\boldsymbol{s}) + \varepsilon(\boldsymbol{s})$, em que $\mathrm{E}(\varepsilon(\boldsymbol{s})) = 0$. Então, a variância do erro pode ser escrita em função dos resíduos dos dados por

$$\begin{aligned}\widehat{Y}(\boldsymbol{s_0}) - Y(\boldsymbol{s_0}) &= \left(\sum_{i=1}^{n} \lambda_i Y(\boldsymbol{s_i})\right) - Y(\boldsymbol{s_0}) = \left(\sum_{i=1}^{n} \lambda_i \left(\mu(\boldsymbol{s_i}) + \varepsilon(\boldsymbol{s_i})\right)\right) - \left(\mu(\boldsymbol{s_0}) + \varepsilon(\boldsymbol{s_0})\right) \\ &= \left(\sum_{i=1}^{n} \lambda_i \left(\mu(\boldsymbol{s_i})\right) - \mu(\boldsymbol{s_0})\right) + \left(\sum_{i=1}^{n} \lambda_i \left(\varepsilon(\boldsymbol{s_i})\right) - \varepsilon(\boldsymbol{s_0})\right).\end{aligned}$$

Consequentemente, segue que

$$\widehat{Y}(s_0) - Y(s_0) = \left(\sum_{i=1}^{n} \lambda_i \varepsilon(\boldsymbol{s_i})\right) - \varepsilon(s_0) = \widehat{\varepsilon}(s_0) - \varepsilon(s_0).$$

Portanto, como em krigagem ordinária, tem-se que

$$\mathrm{Var}\left(\widehat{Y}(s_0) - Y(s_0)\right) = \mathrm{E}\left(\left(\widehat{\varepsilon}(s_0) - \varepsilon(s_0)\right)^2\right) = \mathrm{C}(0) + \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j \mathrm{C}(\boldsymbol{s_i}, \boldsymbol{s_j}) - 2\sum_{i=1}^{n} \mathrm{C}(s_0, \boldsymbol{s_i}). \tag{5.47}$$

Então, o problema de minimização é idêntico ao de krigagem ordinária, isto é, deve-se diferenciar a equação (5.47) com respeito a λ_i, para $i = 1, \ldots, n$, e igualar essas derivadas a zero. Contudo, deseja-se que as condições estabelecidas pelas p equações definidas em (5.46) sejam satisfeitas. Assim, adicionam-se na equação (5.47) p termos nulos correspondentes a estas condições. Então, tem-se que

$$\mathrm{Var}\left(\widehat{Y}(s_0) - Y(s_0)\right) = \mathrm{C}(0) + \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j \mathrm{C}(\boldsymbol{s_i}, \boldsymbol{s_j}) - 2\sum_{i=1}^{n} \mathrm{C}(\boldsymbol{s_0}, \boldsymbol{s_i}) + 2\sum_{j=0}^{p} \eta_j \sum_{i=1}^{n} \lambda_i x_{ij} - x_{0j},$$

em que η_j são parâmetros de Lagrange. E quando as $n + p$ derivadas com respeito a λ_i, para $i = 1, \ldots, n$, e η_j, para $j = 1, \ldots, p$, são igualadas a zero tem-se o sistema de equações lineares

$$S = \begin{cases} \displaystyle\sum_{k=1}^{n} \lambda_k \mathrm{C}\left(\boldsymbol{s_i}, s_k\right) + \sum_{k=0}^{p} \eta_k x_{ik} = \mathrm{C}\left(\boldsymbol{s_i}, \boldsymbol{s_0}\right), \quad i = 1, \ldots, n. \\ \displaystyle\sum_{k=1}^{n} \lambda_k x_{kj} = x_{0j}, \quad j = 1, \ldots, p. \end{cases}$$

Em forma matricial, o sistema S dado em (5.2.3.3) pode ser escrito como $\boldsymbol{C\lambda} = \boldsymbol{A}$, em que

$$\boldsymbol{C} = \begin{pmatrix} \mathrm{C}\left(\boldsymbol{s_1}, \boldsymbol{s_1}\right) & \cdots & \mathrm{C}\left(\boldsymbol{s_1}, s_n\right) & 1 & x_{11} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathrm{C}\left(s_n, \boldsymbol{s_1}\right) & \cdots & \mathrm{C}\left(s_n, s_n\right) & 1 & x_{n1} & \cdots & x_{np} \\ 1 & \cdots & 1 & 0 & 0 & \cdots & 0 \\ x_{11} & \cdots & x_{n1} & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{p1} & \cdots & x_{np} & 0 & 0 & \cdots & 0 \end{pmatrix};$$

$\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n, \eta_1, \ldots, \eta_p)^\top$ e $\boldsymbol{A} = (\mathrm{C}\left(\boldsymbol{s_1}, \boldsymbol{s_0}\right), \ldots, \mathrm{C}\left(\boldsymbol{s_1}, \boldsymbol{s_0}\right), 1, x_{01}, \ldots, x_{0p})$. Então, os pesos λ_i, para $i = 1, \ldots, n$, são dados por $\boldsymbol{\lambda} = \boldsymbol{C}^{-1}\boldsymbol{A}$.

5.2.4 Aplicação

Nesta Seção, a metodologia desenvolvida foi ilustrada a partir de um conjunto de dados experimentais coletados pelo grupo de estatística espacial da Universidade Estadual do Oeste do Paraná.

5.2.4.1 Dados

No total foram amostrados 102 pontos (Figura 24(a)) de uma área comercial de aproximadamente 167 ha, localizada no município de Cascavel. Os pontos foram georreferenciados usando o sistema de coordenadas UTM. Desses pontos, 20 pontos (os marcados com $\oplus$ na Figura 24(a)) foram retirados propositalmente do conjunto de dados, de forma a trabalhar com um conjunto de dados com 82 pontos dispostos em uma grade regular.

A variável-resposta e a covariável escolhidas foram, respectivamente, concentração de Magnésio (Mg) no solo (cmolc/dm^{-3}) e concentração de Cálcio (Ca) no solo (cmolc/dm^{-3}), referentes ao ano safra 2014/2015. Os valores da variável resposta nos 20 pontos excluídos do conjunto de dados (indicados na Figura 24(a) por $\oplus$) foram utilizados para cumprir o primeiro objetivo da aplicação: predizer o valor da variável-resposta utilizando krigagem com *drift* externo e o modelo espacial BS; predizer o valor da variável-resposta utilizando krigagem ordinária e o modelo espacial BS, sem covariáveis e comparar esses dois métodos de predição utilizando os erros de estimação.

A escolha da variável-resposta e da covariável ocorreu por vários motivos. Primeiro, ainda que a fertilidade do solo não seja o único fator vital para a produtividade, a compreensão de seus princípios é essencial para uma produção eficiente das culturas e para a proteção ambiental (INSTITUTO DA POTASSA E FOSFATO, 1998) e a análise química do solo é um dos sistemas mais usados no Brasil para avaliar a fertilidade do solo (OLIVEIRA et. al, 1991). Particularmente, o Magnésio, segundo Oliveira et al. (1991), tem como função predominante na planta a atuação na molécula da clorofila, estando envolvido na fixação de CO_2. Assim, em casos de baixa disponibilidade de Magnésio nas folhas, a fixação de CO_2 fica comprometida. Na deficiência de Magnésio, há dificuldade na translocação de carboidratos para a raiz, a qual prejudica o desenvolvimento do sistema radicular que, por sua vez reduzirá, a absorção de outros nutrientes. A nutrição com Magnésio, além dos critérios nutricionais, confere qualidade ao produto final, como número, peso e conteúdo proteico dos grãos; doçura, cor, sabor e maciez dos frutos. Assim, a presença do Magnésio em dose inadequada no sistema solo-planta prejudica o crescimento e a vitalidade da planta. Além disso, a deficiência de Magnésio no solo apresenta uma faixa de deficiência oculta, em que os sintomas não são visíveis, mas exerce forte impacto negativo sobre a produção e a qualidade das culturas, principais diretrizes na produção das culturas. Assim, a concentração de Magnésio no solo deve ser atenciosamente monitorada (WLAND, 2007). Outro motivo foi o fato da variável Mg apresentar distribuição de probabilidade assimétrica positiva, condição necessária para a utilização do modelo BS.

Uma vez escolhida a variável-resposta, para a escolha da covariável, consideraram-se

que além do conhecimento dos teores dos minérios do solo, pesquisas apontam para a necessidade de conhecer as relações entre os nutrientes presentes no solo (WLAND, 2007). Uma relação bastante conhecida no meio agronômico, talvez a mais conhecida, é a relação Cálcio e Magnésio (Ca:Mg). Esta relação é importante por haver uma competição pelos sítios de adsorção (adesão de moléculas de um fluido (o adsorvido) a uma superfície sólida (o adsorvente)) no solo e na absorção pelas raízes. Como consequência, a presença excessiva de um pode prejudicar os processos de adsorção e absorção do outro e afetar o desenvolvimento das plantas. Segundo Instituto da Potassa e Fosfato (1998), quando a relação é muito alta, as plantas podem absorver menos Magnésio. Isto pode ocorrer, por exemplo, quando o agricultor usa somente calcário calcítico por muitos anos, em solos relativamente pobres em Magnésio. Finalmente, no caso estudado, as variáveis Mg e Ca apresentaram considerável correlação linear.

5.2.4.2 Análise exploratória dos dados

A Figura 24(b) mostra o *box plot* dos dados da concentração de Magnésio, o qual apresenta cinco *outliers*. A Figura 24(c) apresenta um mapa com os pontos divididos em quartis e a localização dos *outliers*. A Figura 24(d) mostra que não existe direção privilegiada, portanto, o processo estocástico pode ser considerado isotrópico.

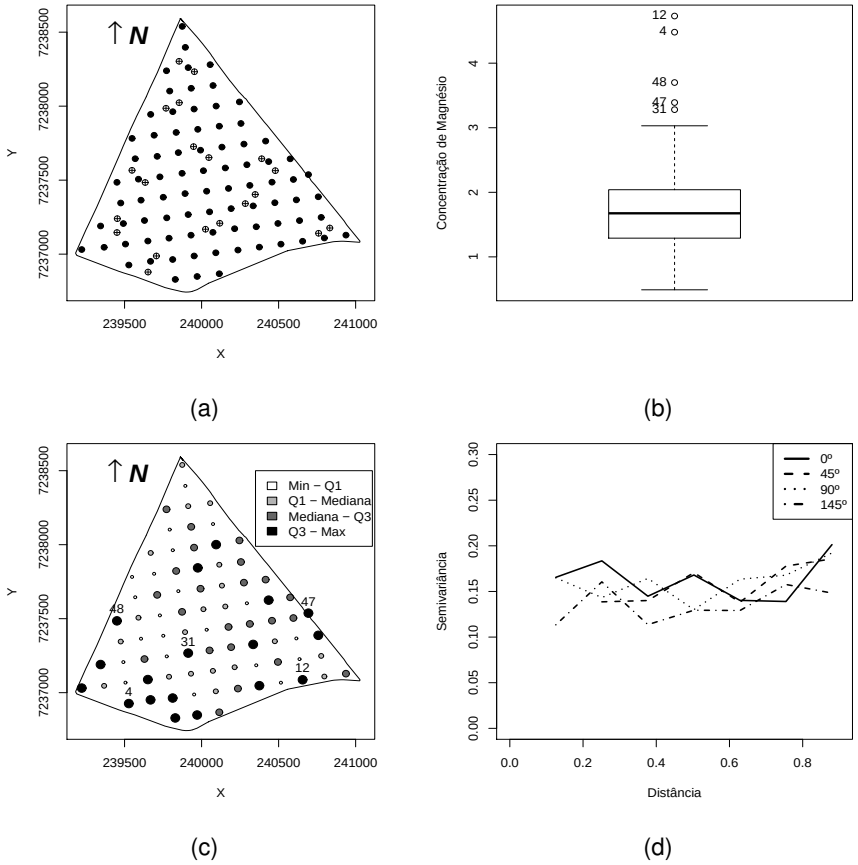


Figura 24 Grade amostral e pontos para serem preditos (a), *box plot* da concentração de Mg (b), pontos amostrais divididos em quartis e localização dos *outliers* (c) e semivariogramas direcionais (d).

5.2.4.3 Análise espacial do modelo

Consideram-se os modelos da família Matérn, dados em (5.42), para alguns valores do parâmetro de ordem $\delta > 0$ para estimar a estrutura de dependência espacial. Dentre todos os modelos considerados, o mais apropriado foi selecionado usando a medida de validação cruzada (VC)(DE BASTIANI et. al, 2015) definida por

$$\mathrm{CV} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{y(\boldsymbol{s_i}) - \widehat{y}_{(i)}(\boldsymbol{s_i})}{1-h_{ii}}\right)^2, \quad i = 1, \ldots, n,$$

em que h_{ii} é o $i-$ésimo elemento diagonal da matriz projeção $\boldsymbol{H}$, $\widehat{y}_{(i)}(\boldsymbol{s_i}) = \boldsymbol{x}_{(i)}^{\top}\widehat{\boldsymbol{\beta}}_{(i)}$, com $\boldsymbol{x}_{(i)}^{\top}$ é a $i-$ésima coluna da matriz de covariáveis, $\boldsymbol{X}$, sem a observação i, e $\widehat{\boldsymbol{\beta}}_{(i)}$ é a estimativa ML de $\boldsymbol{\beta}$ sem a observação i. A matriz projeção $\boldsymbol{H}$ para o modelo log-linear espacial BS é dada por

$$\boldsymbol{H} = \boldsymbol{X}\left(\boldsymbol{X}^{\top}\left(\frac{2}{\alpha^2}\boldsymbol{\Sigma}^{-1} - \frac{1}{2}\boldsymbol{I_n}\right)\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\top}\left(\frac{2}{\alpha^2}\boldsymbol{\Sigma}^{-1} - \frac{1}{2}\boldsymbol{I_n}\right),$$

e que deve ser avaliada em $\widehat{\boldsymbol{\theta}}$. Para ver detalhes sobre $\boldsymbol{H}$ consulte o Apêndice IV. A Tabela 7 mostra valores de VC para alguns valores do parâmetro de forma do modelo Matérn (δ), enquanto a Figura 25 exibe a correspondente representação gráfica. Note que VC é uma função

crescente de δ. Portanto, o melhor modelo é o membro da família Matérn com $\delta = 0,5$ (modelo exponencial).

Tabela 7 Valores da validação cruzada para o modelo de variabilidade espacial da família Matérn com δ indicado.

δ	CV	δ	CV
$0,50$	$2,030 \times 10^{-5}$	$2,75$	$2,421 \times 10^{-5}$
$0,75$	$2,137 \times 10^{-5}$	$3,00$	$2,425 \times 10^{-5}$
$1,00$	$2,194 \times 10^{-5}$	$3,25$	$2,430 \times 10^{-5}$
$1,25$	$2,247 \times 10^{-5}$	$3,50$	$2,436 \times 10^{-5}$
$1,50$	$2,306 \times 10^{-5}$	$3,75$	$2,438 \times 10^{-5}$
$1,75$	$2,360 \times 10^{-5}$	$4,00$	$2,448 \times 10^{-5}$
$2,00$	$2,379 \times 10^{-5}$	$4,25$	$2,451 \times 10^{-5}$
$2,25$	$2,401 \times 10^{-5}$	$4,50$	$2,470 \times 10^{-5}$
$2,50$	$2,416 \times 10^{-5}$	$4,75$	$2,475 \times 10^{-5}$

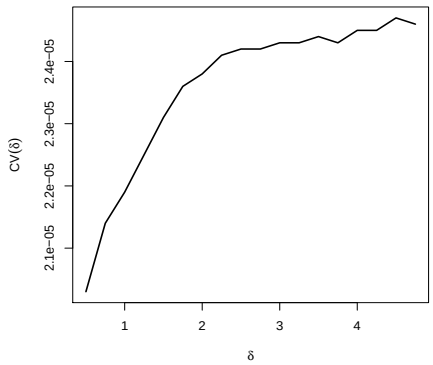


Figura 25 Gráfico da validação cruzada em função de δ para os dados de Mg.

Para o modelo escolhido, os parâmetros estimados pelo método ML (com seus correspondentes SEs entre parênteses e p-valores entre colchetes) são: $\widehat{\alpha} = 1,0168(0,9791)[0,1495]$,

$$\widehat{\beta_0} = -0,9595(0,1088)[> 0,999], \quad \widehat{\beta_1} = 0,2728(0,0146)[< 0,001],$$

$\widehat{\varphi_1} = 0,0270(0,0518)[0,3011]$, $\widehat{\varphi_2} = 0,0115(0,0225)[0,3046]$, $\widehat{\varphi_3} = 0,9436(0,0055)[< 0,001]$.

Então, o modelo espacial BS ajustado é dado por

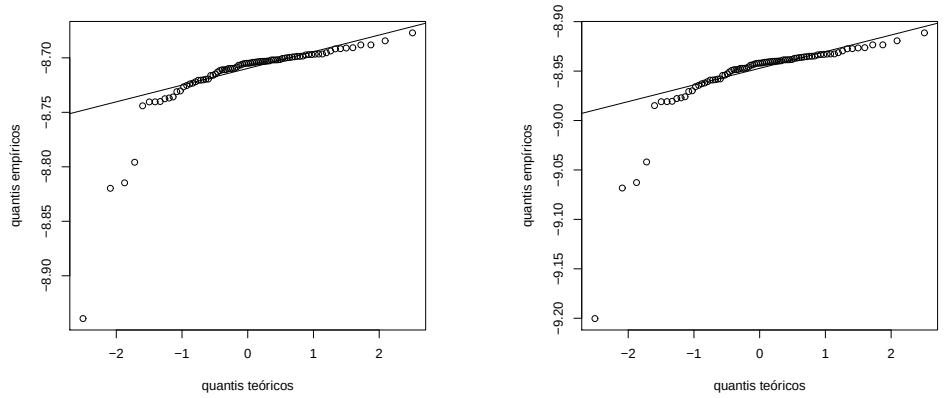
$$\widehat{\mathrm{Mg}}(\boldsymbol{s_i}) = \exp(-0,9595 + 0,2728\,\mathrm{Ca}(\boldsymbol{s_i})), \quad i = 1, \ldots, 82,$$

com $\widehat{\boldsymbol{\Sigma}} = 0,0270\boldsymbol{I}_{82} + 0,0115\widehat{\boldsymbol{R}}$, em que os elementos da matriz $\boldsymbol{R}$ são apresentados em (5.42) com $\delta = 0,5$. O raio de dependência espacial é maior que a distância máxima entre pontos da grade, isto é, todos os pontos da grade são espacialmente dependentes. Para os p-valores fornecidos acima, pode-se concluir que a hipótese H_0: $\theta_i = 0$, em que $\theta_i \in \{\alpha, \beta_0, \beta_1, \varphi_1, \varphi_2, \varphi_3\}$ é rejeitada a 5% de significância para os parâmetros β_1 e φ_3. Portanto, a covariável (concentração de Ca no solo) é significativa para a explicação da resposta

(concentração de Mg no solo).

5.2.4.4 Bondade de ajuste

Agora, utilizam-se os gráficos da distância de Mahalanobis transformada e dos resíduos quantis randonizados, definidos na Seção 5.2.3.2, para avaliar a adequação do modelo log-linear espacial BS no ajuste dos dados sob estudo. A Figura 26 exibe os *Q-Q plots* usando a distância transformada pela aproximação de Wilson-Hilferty e os resíduos quantis randonizados. Nota-se um razoável ajuste, uma vez que a maioria dos pontos estão dispostos ao longo da reta $y = x$, embora algumas observações possam ser consideradas *outliers*. A presença desses *outliers* sugere que um modelo log-linear espacial BS com caudas mais pesadas, tal como um modelo log-linear espacial BS t-Student, pode melhorar o ajuste do modelo aos dados.



(a) (b)

Figura 26 *Q-Q plots* usando a aproximação de Wilson-Hilferty (a) e resíduos quantis randonizados (b) para os dados de Mg.

5.2.4.5 Predição

Nesta Seção estimou-se a concentração de Mg no solo para as vinte ocalizações propositalmente excluídas do conjunto de dados. A estimação foi realizada de duas formas: por krigagem com drift externo e krigagem ordinária. Calcularam-se-se ainda os erros obtidos nos dois processos de estimação. A concentração de Mg amostrada e estimada por cada um dos métodos ($\widehat{\mathrm{Mg}}$) e os respectivos erros são apresentados na Tabela 8. Na estimação por krigagem com drift externo (KDE) encontraram-se erro quadrado médio 0,5559 e a variância do erro quadrado médio 0,5943; enquanto na estimação por krigagem ordinária foram encontrados o erro quadrado médio 0,6363 e a variância do erro quadrado médio 1,2789.

Tabela 8 Concentrações de Mg, amostradas e estimadas (cmolc/dm^{-3}).

nº do ponto	Mg	$\widehat{\mathrm{Mg}}_{KDE}$	$\widehat{\mathrm{Mg}}_{KO}$	Erro$_{KDE}$	Erro$_{KO}$
1	1,88	2,45	1,63	0,3249	0,0625
2	2,10	2,44	1,65	0,1156	0,2025
3	3,93	2,25	1,64	2,8224	5,2441
4	1,76	2,24	1,63	0,2304	0,0169
5	2,50	2,44	1,64	0,0036	0,7396
6	2,16	2,42	1,66	0,0676	0,2500
7	1,13	2,51	1,63	1,9044	0,2500
8	1,12	2,53	1,64	1,9881	0,2704
9	1,98	2,33	1,64	0,1225	0,1156
10	2,60	2,30	1,64	0,0900	0,9216
11	2,50	2,28	1,65	0,6084	0,0255
12	2,91	2,27	1,63	0,4096	1,6384
13	1,32	2,20	1,64	0,7744	0,1024
14	2,25	2,24	1,65	0,0001	0,3600
15	1,43	2,12	1,65	0,4761	0,0484
16	1,16	2,19	1,64	1,0609	0,2304
17	2,01	2,05	1,63	0,0016	0,1444
18	1,59	2,02	1,65	0,1849	0,0036
19	2,60	1,97	1,63	0,3969	0,9409
20	2,31	2,00	1,63	0,0961	0,4624

5.2.5 Conclusões

Foi proposto um modelo log-linear espacial com *drift* externo baseado na distribuição Birnbaum-Saunders. Este modelo mostrou-se adequado para ajustar um conjunto de dados com estrutura de dependência espacial com distribuição de probabilidade assimétrica positiva. As estimativas por máxima verossimilhança dos parâmetros do modelo foram calculadas de acordo com o método quase-Newton L-BFGS-B. A abordagem proposta foi aplicada em um estudo com dados reais (dados agrícolas). Em concordância com o relatado na literatura, a predição de dados por krigagem com *drift* externo se mostrou mais efeciente do que a predição por krigagem ordinária, uma vez que o primeiro método de estimação apresentou erro quadrado médio menor do que o erro quadrado médio da krigagem ordinária e a variância dos erros foi bem menor do que a variância dos erros da krigagem ordinária.

5.2.6 Trabalhos futuros

Alguns possíveis problemas a serem abordados em estudos futuros são os seguintes: primeiro, devido à distribuição Birnbaum-Saunders ser baseada na distribuição normal, uma versão baseada em uma distribuição de cauda mais pesada, por exemplo, na distribuição t-Student, pode ser considerada. Assim, pode-se reduzir a influência de dados atípicos, que podem ter efeito desfavorável em mapas espaciais. Adicionalmente, métodos de diagnósticos podem ser considerados para avaliar a influência de dados atípicos. Pode-se ainda considerar mais de uma variável aleatória na modelagem espacial a partir de estruturas multivariadas da

distribuição Birnbaum-Saunders.

5.2.7 Referências

ASSUMPÇÃO, R.A.B.; URIBE-OPAZO, M.A.; GALEA, M. Analysis of local influence in geostatistics using Student-t distribution. **Journal of Applied Statistics,** v. 41, p. 2323-2341, 2014.

BISHOP, T.; MCBRATNEY, A. A Comparison of predictions methods for creation of the creation of field-extent soil property maps. **Geoderma,** v. 103, p. 149-160, 2001.

CAMBARDELLA, C.A.; MOORMAN, T.B; NOVAK, J.M.; PARKIN, T.B.; KARLEN, D.L.; TURCO, R.F.; KONOPKA, A.E. Field-scale variability of soil properties in central Iowa soils. **Soil Science Society of America Journal,** v. 58, p. 1501-1511, 1994.

CRESSIE, N.A.C. **Statistics for Spatial Data.** New-York: Jonh Wiley & Sons, 2015.

DE BASTIANI, F.; CYSNEIROS, A.H.M.A.; URIBE-OPAZO, M.A.; GALEA, M. Influence diagnostics in elliptical spatial linear models. **TEST,** v. 24, p. 322-340, 2015.

DUNN, P. K.; SMYTH, G.K. Randomized Quantile Residuals. **Journal of Computational and Graphical Statistics,** v. 5, p. 236-244, 1996.

FERREIRA, M.; GOMES, M.I.; LEIVA, V. On an extreme value version of the Birnbaum-Saunders distribution. **REVSTAT-Statistical Journal,** v. 10, p. 181-210, 2012.

GARCIA-PAPANI, F.M.; URIBE-OPAZO; M.A., LEIVA, V.; AYKROYD, R. G. Birnbaum-Saunders spatial modelling and diagnostics applied to agricultural engineering data. **Stochastic Environmental Research and Risk Assessment,** (in press), 2016. Disponível em: <http://dx.doi.org/10.1007/s00477-015-1204-4>. Acesso em: 02 fev. 2016.

GOOVAERTS, P. Using elevation to aid the geostatistical mapping of rainfall erosivity. **Catena,** v. 34, p. 227-242, 1999.

HENGL, T.; HEUVELINK, G.B.M.; STEIN, A. Comparison of kriging with external drift and regression-kriging. Technical report. **International Institute for Geo-information Science and Earth Observation (ITC),** 2003. Disponível em: <http://www.itc.nl/library/Academic-output>. Acesso em: 14 set. 2015.

HENGL, T.; HEUVELINK, G.B.M.; STEIN, A. A generic framework for spatial prediction of soil variables based on regression-kriging. **Geoderma,** v. 120, p. 75-93, 2004.

HU, Y.; JIA, Z.; CHENG, J.; ZHAO, Z.; CHEN, F. Spatial variability of soil arsenic and its association with soil nitrogen in intensive farming systems. **Journal Soils Sediments, 2015.** Disponível em: <http://dx.doi.org/10.1007/s11368-015-1182-7>. Acesso em 12 nov. 2015.

IBACACHE-PULGAR, G.; PAULA, G.; GALEA, M. On influence diagnostics in elliptical multivariate regression models with equicorrelated random errors. **Statistical Modelling**, v. 16, p. 14-21, 2014.

INSTITUTO DA POTASSA & FOSFATO. **Manual internacional de fertilidade do solo**. Tradução e adaptação de Alfredo Scheid Lopes. Piracicaba: Potafos, 1998.

ISAAKS, E.; SRIVASTAVA, R. **An Introduction to Applied Geostatistics.** New York: Oxford University Press, 1989.

JOHNSON, N.L.; KOTZ, S.; BALAKRISHNAN, N. **Continuous Univariate Distributions.** New York: Jonh Wiley & Sons, 1994. v.1

JOHNSON, N.L.; KOTZ, S.; BALAKRISHNAN, N. **Continuous Univariate Distributions.** New York: Jonh Wiley & Sons, 1995. v.2

KRIGE, D.G. A statistical approach to some basic mine valuation problems on the Witwatersrand. **Journal of the Chemical, Metallurgical and Mining Society of South Africa,** v. 52, p. 119-139, 1951.

KUNDU, D.; BALAKRISHNAN, N.; JAMALIZADEH, A. Bivariate Birnbaum-Saunders distribution and associated inference. **Journal of Multivariate Analysis,** v. 101, p. 113-125, 2010.

LANGE, K.L.; LITTLE, R.J.; TAYLOR, J.M. Robust statistical modeling using the t distribution. **Journal of the American Statistical Association,** v. 84, p. 881-896, 1989.

LEITHOLD, L. **O cálculo com geometria analítica.** Tradução de Cyro de Carvalho Patarra. São Paulo: Harbra, 1994.

LEIVA, V.; BARROS, M.; PAULA, G.; SANHUEZA, A. Generalized Birnbaum-Saunders distribution applied to air pollutant concentration. **Environmetrics**, v. 19, p. 235-249, 2008.

LEIVA, V.; SANHUEZA, A.; ANGULO, J.M. A length-biased version of the Birnbaum-Saunders distribution with application in water quality. **Stochastic Environmental Research and Risk Assessment**, v. 23, p. 299-307, 2009.

LEIVA, V.; MARCHANT, C.; RUGGERI, F.; SAULO, H. A criterion for environmental assessment using Birnbaum-Saunders attribute control charts. **Environmetrics**, v. 26, p. 463-476, 2015.

LEIVA, V. **The Birnbaum-Saunders Distribution**. New York: Academic Press, 2015.

LEIVA, V.; FERREIRA, M.; GOMES, M.I.; LILLO, C. Extreme value Birnbaum-Saunders regression models applied to environmental data. **Stochastic Environmental Research and Risk Assessment** (in press), 2015. Disponível em <http://dx.doi.org/10.1007/s00477-015-1069-6>. Acesso em: 08 out. 2015.

MARCHANT, C.; LEIVA, V.; CAVIERES, M.F.; SANHUEZA, A. Air contaminant statistical distributions with application to PM10 in Santiago, Chile. **Reviews of Environmental Contamination and Toxicology,** v. 223, p. 1-31, 2013.

MARCHANT, C.; LEIVA, V.; CYSNEIROS, F.J.A. A multivariate log-linear model for Birnbaum-Saunders distributions. **IEEE Transactions on Reliability** (*in press*), 2015.

MAGNUS, J.R.; NEUDECKER, H. **Matrix Differential Calculus with applications in statistics and econometrics.** Chichester: John Wiley & Sons, 2007.

NOCEDAL, J.; WRIGHT, S.J. **Numerical Optimization.** New York: Springer, 1999.

OLIVEIRA, A.J.; GARRIDO, W.E.; ARAUJO, J.D.; LOURENÇO, S. (Org.). **Métodos de pesquisa em fertilidade do solo.** Brasília:EMBRAPA-SEA, 1991.

PODLASKI, R. Characterization of diameter distribution data in near-natural forests using the Birnbaum-Saunders distribution. **Canadian Journal of Forest Research,** v. 38, p. 518-527, 2008.

SAULO, H.; LEIVA, V.; ZIEGELMANN, F.A.; MARCHANT, C. A nonparametric method for estimating asymmetric densities based on skewed Birnbaum-Saunders distributions applied to environmental data. **Stochastic Environmental Research and Risk Assessment,** v. 27, p. 1479-1491, 2013.

SOARES, A. **Geoestatística para as ciências da terra e do ambiente.** Lisboa: IST Press, 2000.

WLAND, T. Magnésio no solo e nas plantas. **Informações Agronômicas,** n. 117, p. 19-21, 2007.

. WEBSTER, R.; OLIVER, M. **Geostatistics for Environmental Scientists.** Chichester: Jonh Wiley & Sons, 2009.

XIA, J.; ZEEPHONGSEKUL, P.; PACKER, D. Spatial and temporal modelling of tourist movements using semi-Markov processes. **Tourism Management,** v. 32, p. 844-851, 2011.

5.2.8 Apêndices

5.2.8.1 Apêndice I: Covariância da distribuição BS multivariada

A covariância entre T_i e T_j é dada por $\mathrm{Cov}(T_i, T_j) = \mathrm{E}(T_iT_j) - \mathrm{E}(T_i)\mathrm{E}(T_j),\ i, j = 1, 2, \cdots, n$. Tem-se ainda $T_k = \gamma_k \left(\frac{\alpha_k}{2} Z_k + \sqrt{\left(\frac{\alpha_k}{2} Z_k\right)^2 + 1} \right)^2,\ k = 1, 2, \cdots, n$. Portanto,

$$
\begin{aligned}
\frac{\mathrm{E}(T_iT_j)}{\gamma_i\gamma_j} &= \mathrm{E}\left(\left(\frac{\alpha_\mathrm{i}}{2}\mathrm{Z_i} + \sqrt{\left(\frac{\alpha_\mathrm{i}}{2}\mathrm{Z_i}\right)^2 + 1} \right)^2 \times \left(\frac{\alpha_\mathrm{j}}{2}\mathrm{Z_j} + \sqrt{\left(\frac{\alpha_\mathrm{j}}{2}\mathrm{Z_j}\right)^2 + 1} \right)^2 \right) \\
&= \mathrm{E}\left(\left(\frac{1}{2}\alpha_\mathrm{i}^2\mathrm{Z_i^2} + 1 + \alpha_\mathrm{i}\mathrm{Z_i}\sqrt{\left(\frac{\alpha_\mathrm{i}}{2}\mathrm{Z_i}\right)^2 + 1} \right) \times \left(\frac{1}{2}\alpha_\mathrm{j}^2\mathrm{Z_j^2} + 1 + \alpha_\mathrm{j}\mathrm{Z_j}\sqrt{\left(\frac{\alpha_\mathrm{j}}{2}\mathrm{Z_j}\right)^2 + 1} \right) \right) \\
&= 1 + \frac{\alpha_i^2\alpha_j^2}{4}\mathrm{E}(Z_i^2Z_j^2) + \frac{\alpha_i^2}{2}\mathrm{E}(Z_i^2) + \frac{\alpha_j^2}{2}\mathrm{E}(Z_j^2) + \alpha_j\mathrm{E}\left(Z_j\sqrt{\left(\frac{\alpha_j}{2}Z_j\right)^2 + 1} \right) \\
&\quad + \alpha_i\mathrm{E}\left(Z_i\sqrt{\left(\frac{\alpha_i}{2}Z_i\right)^2 + 1} \right) + \frac{\alpha_i^2\alpha_j}{2}\mathrm{E}\left(Z_i^2Z_j\sqrt{\left(\frac{\alpha_j}{2}Z_j\right)^2 + 1} \right) \\
&\quad + \frac{\alpha_j^2\alpha_i}{2}\mathrm{E}\left(Z_j^2Z_i\sqrt{\left(\frac{\alpha_i}{2}Z_i\right)^2 + 1} \right) + \alpha_i\alpha_j\mathrm{E}\left(Z_iZ_j\sqrt{\left(\frac{\alpha_j}{2}Z_j\right)^2 + 1}\sqrt{\left(\frac{\alpha_i}{2}Z_i\right)^2 + 1} \right).
\end{aligned}
$$

Sabe-se que a função geradora de momentos da distribuição normal padrão bivariada $m(t_i, t_j) = \exp\left\{ \frac{1}{2}\left(t_i^2 + 2t_it_j\rho + t_j^2 \right) \right\}$, e que para se obter o momento conjunto $\mathrm{E}(T_i^rT_j^s)$ diferencia-se $m(t_i, t_j)$ r vezes em relação a t_i e s vezes em relação a t_j e faz-se t_i e t_j iguais a zero. Realizando estes cálculos, encontra-se $\mathrm{E}(Z_k^2) = 1,\ k = 1, \cdots, n$ e $\mathrm{E}(Z_i^2Z_j^2) = 1 + 2\rho_{ij}^2,\ i, j = 1, \cdots, n$. Além disso, uma vez que $Z_k\sqrt{\left(\frac{\alpha_k}{2}Z_k\right)^2 + 1}$ é uma função ímpar, sabe-se que $\mathrm{E}\left(Z_k\sqrt{\left(\frac{\alpha_k}{2}Z_k\right)^2 + 1} \right) = \mathrm{E}\left(Z_k^2Z_k\sqrt{\left(\frac{\alpha_k}{2}Z_k\right)^2 + 1} \right) = 0, k = 1, \cdots, n$ (KUNDU; BALAKRISHNAN; JAMALIZADEH, 2010). Portanto,

$$
\frac{\mathrm{E}(T_iT_j)}{\gamma_i\gamma_j} = \left(1 + \frac{\alpha_i^2\alpha_j^2}{4}(1 + 2\rho_{ij}^2) + \frac{1}{2}\left(\alpha_i^2 + \alpha_j^2\right) + \alpha_i\alpha_j\mathrm{E}\left(Z_iZ_j\sqrt{\left(\frac{\alpha_j}{2}Z_j\right)^2 + 1}\sqrt{\left(\frac{\alpha_i}{2}Z_i\right)^2 + 1} \right) \right),\ i, j = 1, \cdots, n.
$$

De acordo com o Teorema Binomial, pode-se escrever

$$
Z_k\sqrt{\left(\frac{\alpha_k}{2}Z_k\right)^2 + 1} = Z_k + \frac{\alpha_k^2}{2^3}Z_k^3 + \sum_{p=2}^{\infty}(-1)^{p-1}\frac{1 \cdot 3 \cdot (2p-3)}{2^{3p}p!}\alpha_k^{2p}Z_k^{2p+1},\ k = 1, \cdots, n.
$$

E quando se utiliza que

$$
\mathrm{E}\left(Z_i^{2m+1}Z_j^{2n+1} \right) = \frac{(2m+1)!(2n+1)!}{2^m n} \sum_{p=0}^{\mathrm{min(m,n)}} \frac{(2\rho)^{2p+1}}{(m-p)!(n-p)!(2p+1)!},\ i, j = 1, \cdots, n.
$$

tem-se que,

$$
\mathrm{E}\left(Z_iZ_j\sqrt{\left(\frac{\alpha_j}{2}Z_j\right)^2 + 1}\sqrt{\left(\frac{\alpha_i}{2}Z_i\right)^2 + 1} \right) = \left(2 + \frac{3\alpha_j^2}{4} + \frac{3\alpha_i^2}{4} + \frac{9\alpha_i^2\alpha_j^2}{2^5} \right)\rho_{ij} + \frac{3\alpha_i^2\alpha_j^2}{2^4}\rho_{ij}^3 + r(Z_i, Z_j),
$$

em que $r(Z_i, Z_j)$ é dado por,

$$\begin{aligned} r(Z_i,Z_j) &= \sum_{p=2}^{\infty}(-1)^{p-1}\frac{1\cdot 3\cdots(2p-3)}{2^{3p}p!}\alpha_2^{2p}\mathrm{E}\left(Z_iZ_j^{2p+1}\right) \\ &+ \sum_{k=p}^{\infty}(-1)^{p-1}\frac{1\cdot 3\cdots(2p-3)}{2^{3p}p!}\alpha_i^{2p}\mathrm{E}\left(Z_iZ_j^{2p+1}\right) \\ &+ \frac{\alpha_i^2}{2^3}\sum_{p=2}^{\infty}(-1)^{p-1}\frac{1\cdot 3\cdots(2p-3)}{2^{3p}p!}\alpha_2^{2p}\mathrm{E}\left(Z_i^3Z_j^{2k+1}\right) \\ &+ \frac{\alpha_j^2}{2^3}\sum_{p=2}^{\infty}(-1)^{p-1}\frac{1\cdot 3\cdots(2p-3)}{2^{3p}p!}\alpha_i^{2p}\mathrm{E}\left(Z_j^3Z_i^{2p+1}\right) \\ &+ \sum_{p=2}^{\infty}\sum_{q=2}^{\infty}(-1)^{p+q-2}\frac{1\cdot 3\cdots(2p-3)}{2^{3p}p!}\frac{1\cdot 3\cdots(2q-3)}{2^{3q}q!}\alpha_i^{2p}\alpha_j^{2q}\mathrm{E}\left(Z_i^{2p+1}Z_j^{2q+1}\right). \end{aligned}$$

Desta forma, tem-se

$$\mathrm{Cov}(T_i,T_j) = \gamma_i\gamma_j\left(\left(2+\frac{3\alpha_i^2}{2}+\frac{3\alpha_j^2}{2}+\frac{9\alpha_i^2\alpha_j^2}{2^5}\right)\rho_{ij}+\frac{\alpha_i^2\alpha_j^2}{2}\rho_{ij}^2+\frac{3\alpha_i^2\alpha_j^2}{2^4}\rho_{ij}^3+r(Z_i,Z_j)\right).$$

5.2.8.2 Apêndice II: Matriz de informação observada de Fisher

A matriz de informação é definida por $-\ddot{\boldsymbol{\ell}}(\boldsymbol{\theta})$, avaliada em $\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}$, em que $\ddot{\boldsymbol{\ell}}(\boldsymbol{\theta})$ é a matriz Hessiana. Para o modelo log-linear espacial BS, apresentado em (5.41), a matriz Hessiana $(p+5)\times(p+5)$ é dada por

$$\ddot{\boldsymbol{\ell}}(\boldsymbol{\theta}) = \begin{pmatrix} \frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}^\top} & \frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\beta}\partial\boldsymbol{\varphi}^\top} & \frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\beta}\partial\alpha} \\ \frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\varphi}\partial\boldsymbol{\beta}^\top} & \frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\varphi}\partial\boldsymbol{\varphi}^\top} & \frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\varphi}\partial\alpha} \\ \frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\alpha\partial\boldsymbol{\beta}^\top} & \frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\alpha\partial\boldsymbol{\varphi}^\top} & \frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\alpha\partial\alpha} \end{pmatrix} = \begin{pmatrix} \ddot{\boldsymbol{\ell}}_{\boldsymbol{\beta}\boldsymbol{\beta}} & \ddot{\boldsymbol{\ell}}_{\boldsymbol{\beta}\boldsymbol{\varphi}} & \ddot{\boldsymbol{\ell}}_{\boldsymbol{\beta}\alpha} \\ \ddot{\boldsymbol{\ell}}_{\boldsymbol{\varphi}\boldsymbol{\beta}} & \ddot{\boldsymbol{\ell}}_{\boldsymbol{\varphi}\boldsymbol{\varphi}} & \ddot{\boldsymbol{\ell}}_{\boldsymbol{\varphi}\alpha} \\ \ddot{\boldsymbol{\ell}}_{\alpha\boldsymbol{\beta}} & \ddot{\boldsymbol{\ell}}_{\alpha\boldsymbol{\varphi}} & \ddot{\ell}_{\alpha\alpha} \end{pmatrix},$$

em que

$$\ddot{\ell}_{\alpha\alpha} = \frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\alpha\partial\alpha} = \frac{n}{\alpha^2}-\frac{12}{\alpha^4}\boldsymbol{V}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{V};$$

$\ddot{\boldsymbol{\ell}}_{\boldsymbol{\beta}\boldsymbol{\beta}}$ é uma matriz $p\times p$ com elementos dados por

$$\begin{aligned} \ddot{\ell}_{\beta_j\beta_l} = \frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\beta_j\partial\beta_l} &= -\frac{1}{\alpha^2}\left(\boldsymbol{V}^\top\odot\boldsymbol{X}_j^\top\odot\boldsymbol{X}_l^\top\right)\boldsymbol{\Sigma}^{-1}\boldsymbol{V}-\frac{1}{\alpha^2}\left(\boldsymbol{W}^\top\odot\boldsymbol{X}_j^\top\right)\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{W}\odot\boldsymbol{X}_l\right) \\ &+\frac{1}{4}\sum_{i=1}^{n}\mathrm{sech}^2\left(\frac{y_i-\boldsymbol{x}_i^\top\boldsymbol{\beta}}{2}\right)\boldsymbol{X}_{ij}\boldsymbol{X}_{il}, \quad j,l=1,\ldots,p; \end{aligned}$$

$\ddot{\boldsymbol{\ell}}_{\boldsymbol{\beta}\boldsymbol{\varphi}}=\ddot{\boldsymbol{\ell}}_{\boldsymbol{\varphi}\boldsymbol{\beta}}^\top$ é uma matriz $p\times 3$ com elementos dados por

$$\ddot{\ell}_{\beta_j\varphi_l} = \frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\beta_j\partial\varphi_l} = -\frac{2}{\alpha^2}\left(\boldsymbol{W}^\top\odot\boldsymbol{X}_j^\top\right)\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_l}\boldsymbol{\Sigma}^{-1}\boldsymbol{V}, \quad j=1,\ldots,p; \quad l=1,2,3;$$

e $\ddot{\boldsymbol{\ell}}_{\boldsymbol{\varphi\varphi}}$ é uma matriz 3×3 com elementos dados por

$$\begin{aligned}\ddot{\ell}_{\varphi_j\varphi_l} = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial\varphi_j\partial\varphi_l} &= -\frac{1}{2}\mathrm{tr}\left(-\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_l}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_j} + \boldsymbol{\Sigma}^{-1}\frac{\partial^2\boldsymbol{\Sigma}}{\partial\varphi_l\partial\varphi_j}\right) \\ &\quad +\frac{2}{\alpha^2}\boldsymbol{V}^\top\left(-\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_l}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\varphi_j}\boldsymbol{\Sigma}^{-1}\right. \\ &\quad \left.+\boldsymbol{\Sigma}^{-1}\left(\frac{\partial^2\boldsymbol{\Sigma}}{\partial\varphi_l\partial\varphi_j}\boldsymbol{\Sigma}^{-1} - \frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_j}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_l}\boldsymbol{\Sigma}^{-1}\right)\right)\boldsymbol{V}; \quad j,l = 1,2,3,\end{aligned}$$

em que

$$\frac{\partial^2\boldsymbol{\Sigma}}{\partial\varphi_j\partial\varphi_l} = \frac{\partial^2\boldsymbol{\Sigma}}{\partial\varphi_l\partial\varphi_j}, \frac{\partial^2\boldsymbol{\Sigma}}{\partial\varphi_1\partial\varphi_l} = \frac{\partial^2\boldsymbol{\Sigma}}{\partial\varphi_2\partial\varphi_2} = 0, \frac{\partial^2\boldsymbol{\Sigma}}{\partial\varphi_2\partial\varphi_3} = \boldsymbol{R}', \quad \frac{\partial^2\boldsymbol{\Sigma}}{\partial\varphi_3\partial\varphi_3} = \varphi_2\boldsymbol{R}'', \quad l = 1,2,3.$$

Para o modelo Matérn, os elementos da matriz $\boldsymbol{R}''$ são dados

$$\frac{\partial^2 r_{ij}}{\partial\varphi_3\partial\varphi_3} = \left(\frac{\delta(\delta+1)r_{ij}}{\varphi_3^2}\right) + \left(\frac{1}{\varphi_3^2 2^{\delta-1}\Gamma(\delta)}\right)\left(\frac{h_{ij}}{\varphi_3}\right)^{\delta+1}\left(2(\delta+1)K'_\delta\left(\frac{h_{ij}}{\varphi_3}\right) + \left(\frac{h_{ij}}{\varphi_3}\right)K''_\delta\left(\frac{h_{ij}}{\varphi_3}\right)\right),$$

para $i \neq j, i, j = 1, \ldots, n$, com

$$K'_\delta(u) = -(1/2)(K_{\delta-1}(u) + K_{\delta+1}(u)), K''_\delta(u) = (1/4)(K_{\delta-2}(u) + 2K_\delta(u) + K_{\delta+2}(u)).$$

Adicionalmente, $\ddot{\boldsymbol{\ell}}_{\boldsymbol{\beta}\alpha} = \ddot{\boldsymbol{\ell}}_{\alpha\boldsymbol{\beta}}^\top$ é um vetor $p \times 1$ com elementos dados por

$$\ddot{\ell}_{\beta_j\alpha} = \frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\beta_j\partial\alpha} = -\frac{4}{\alpha^3}\left(\boldsymbol{W}^\top \odot \boldsymbol{X}_j^\top\right)\boldsymbol{\Sigma}^{-1}\boldsymbol{V}; \quad j = 1, \ldots, p.$$

Além disso, $\ddot{\boldsymbol{\ell}}_{\boldsymbol{\varphi}\alpha} = \ddot{\boldsymbol{\ell}}_{\alpha\boldsymbol{\varphi}}^\top$ é um vetor, 3×1, com elementos dados por

$$\ddot{\ell}_{\varphi_j\alpha} = \frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\varphi_j\partial\alpha} = -\frac{4}{\alpha^3}\boldsymbol{V}^\top\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_j}\boldsymbol{\Sigma}^{-1}\right)\boldsymbol{V}; \quad j = 1,2,3.$$

5.2.8.3 Apêndice III: Matriz de informação esperada de Fisher

A correspondente matriz de informação esperada é dada por $\boldsymbol{I}(\boldsymbol{\theta}) = \mathrm{E}[-\ddot{\boldsymbol{\ell}}(\boldsymbol{\theta})]$, em que $\ddot{\boldsymbol{\ell}}(\boldsymbol{\theta})$ é a matriz de informação observada, apresentada no apêndice II. Então, para o modelo log-linear espacial BS, esta matriz $(p+5) \times (p+5)$ é dada por

$$\boldsymbol{I}(\boldsymbol{\theta}) = \begin{pmatrix} \boldsymbol{I}_{\boldsymbol{\beta\beta}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{I}_{\boldsymbol{\varphi\varphi}} & \boldsymbol{I}_{\boldsymbol{\varphi}\alpha} \\ \mathbf{0} & \boldsymbol{I}_{\alpha\boldsymbol{\varphi}} & I_{\alpha\alpha} \end{pmatrix} = \begin{pmatrix} \boldsymbol{I}_{\boldsymbol{\beta\beta}} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{I}_{\boldsymbol{\psi\psi}} \end{pmatrix},$$

em que

$$I_{\alpha\alpha} = \mathrm{E}[-\ddot{\ell}_{\alpha\alpha}] = \mathrm{E}\left(-\frac{n}{\alpha^2} + \frac{12}{\alpha^4}\boldsymbol{V}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{V}\right) = \frac{2n}{\alpha^2},$$

uma vez que $\left(\frac{2}{\alpha}\boldsymbol{V}\right)^{\top}\boldsymbol{\Sigma}^{-1}\left(\frac{2}{\alpha}\boldsymbol{V}\right) \sim \chi^2_n$; $\boldsymbol{I}_{\boldsymbol{\beta\beta}} = \mathrm{E}[-\ddot{\ell}_{\boldsymbol{\beta\beta}}]$ é uma matriz $(p+1)\times(p+1)$, com elementos expresso por

$$\begin{aligned} I_{\beta_j\beta_l} &= \mathrm{E}\left(\frac{1}{\alpha^2}\left(\boldsymbol{V}^{\top}\odot(\boldsymbol{X}_j\odot\boldsymbol{X}_l)^{\top}\right)\boldsymbol{\Sigma}^{-1}\boldsymbol{V}\right)+\mathrm{E}\left(\frac{1}{\alpha^2}\left(\boldsymbol{W}^{\top}\odot\boldsymbol{X}_j^{\top}\right)\boldsymbol{\Sigma}^{-1}(\boldsymbol{W}\odot\boldsymbol{X}_l)\right)\\ &\quad+\mathrm{E}\left(-\frac{1}{4}\sum_{i=1}^{n}\mathrm{sech}^2\left(\frac{y_i-\boldsymbol{x}_i^{\top}\boldsymbol{\beta}}{2}\right)\boldsymbol{X}_{ij}\boldsymbol{X}_{il}\right)\\ &= \frac{1}{4}\mathrm{E}\left(\left(\left(\frac{2}{\alpha}\boldsymbol{V}\right)^{\top}\odot(\boldsymbol{X}_j\odot\boldsymbol{X}_l)^{\top}\right)\boldsymbol{\Sigma}^{-1}\left(\frac{2}{\alpha}\boldsymbol{V}\right)\right)+\frac{1}{\alpha^2}\boldsymbol{X}_j^{\top}\boldsymbol{\Sigma}^{-1}\boldsymbol{X}_l-\frac{1}{4}\sum_{i=1}^{n}\boldsymbol{X}_{ij}\boldsymbol{X}_{il}\\ &= \frac{1}{4}\mathbf{1}^{\top}(\boldsymbol{X}_j\odot\boldsymbol{X}_l)+\frac{1}{\alpha^2}\boldsymbol{X}_j^{\top}\boldsymbol{\Sigma}^{-1}\boldsymbol{X}_l-\frac{1}{4}\sum_{i=1}^{n}\boldsymbol{X}_{ij}\boldsymbol{X}_{il}=\frac{1}{\alpha^2}\boldsymbol{X}_j^{\top}\boldsymbol{\Sigma}^{-1}\boldsymbol{X}_l; \end{aligned}$$

Ao se consider novamente que $\boldsymbol{W}\cong\mathbf{1}$, $\mathrm{sech}^2\left(\frac{y_i-\boldsymbol{x}_i^{\top}\boldsymbol{\beta}}{2}\right)\cong 1$ e usar propriedades do traço e do produto de Hadamard (MAGNUS; NEUDECKER, 2007), observa-se que $\boldsymbol{I}_{\boldsymbol{\beta\varphi}} = \mathrm{E}[-\ddot{\ell}_{\boldsymbol{\beta\varphi}}]$ é uma matriz $(p+1)\times 3$, com elementos

$$\boldsymbol{I}_{\beta_j\varphi_l} = \mathrm{E}\left(\frac{2}{\alpha^2}\left(\boldsymbol{W}^{\top}\odot\boldsymbol{X}_j^{\top}\right)\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_l}\boldsymbol{\Sigma}^{-1}\boldsymbol{V}\right)=\frac{1}{\alpha}\boldsymbol{X}_j^{\top}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_l}\boldsymbol{\Sigma}^{-1}\mathrm{E}\left(\frac{2}{\alpha}\boldsymbol{V}\right)=0;$$

$\boldsymbol{I}_{\boldsymbol{\varphi\varphi}} = \mathrm{E}[-\ddot{\ell}_{\boldsymbol{\varphi\varphi}}]$ é uma matriz simétrica 3×3, com elementos

$$\begin{aligned} I_{\varphi_j\varphi_l} &= \mathrm{E}\left(\frac{1}{2}\mathrm{tr}\left(-\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_l}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_j}+\boldsymbol{\Sigma}^{-1}\frac{\partial^2\boldsymbol{\Sigma}}{\partial\varphi_l\partial\varphi_j}\right)\right)+\mathrm{E}\left(\frac{2}{\alpha^2}\boldsymbol{V}^{\top}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_l}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\varphi_j}\boldsymbol{\Sigma}^{-1}\right)\boldsymbol{V}\right)\\ &\quad-\mathrm{E}\left(\frac{2}{\alpha^2}\boldsymbol{V}^{\top}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial^2\boldsymbol{\Sigma}}{\partial\varphi_l\partial\varphi_j}\boldsymbol{\Sigma}^{-1}\right)\boldsymbol{V}\right)+\mathrm{E}\left(\frac{2}{\alpha^2}\boldsymbol{V}^{\top}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_j}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_l}\boldsymbol{\Sigma}^{-1}\right)\boldsymbol{V}\right)\\ &= \frac{1}{2}\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_j}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_l}\right),\quad j,l=1,2,3; \end{aligned}$$

$\boldsymbol{I}_{\boldsymbol{\beta}\alpha} = \mathrm{E}[-\ddot{\ell}_{\boldsymbol{\beta}\alpha}]$ é um vetor $p\times 1$, com elementos dados por

$$\boldsymbol{I}_{\beta_j\alpha} = \mathrm{E}\left(\frac{4}{\alpha^3}\left(\boldsymbol{W}^{\top}\odot\boldsymbol{X}_j^{\top}\right)\boldsymbol{\Sigma}^{-1}\boldsymbol{V}\right)\cong\mathrm{E}\left(\frac{4}{\alpha^3}\boldsymbol{X}_j^{\top}\boldsymbol{\Sigma}^{-1}\boldsymbol{V}\right)=0,$$

já que $\left(\frac{2}{\alpha}\boldsymbol{V}\right)\sim\mathsf{N}_n(\mathbf{0},\boldsymbol{\Sigma})$; $\boldsymbol{I}_{\boldsymbol{\varphi}\alpha} = \mathrm{E}[-\ddot{\ell}_{\boldsymbol{\varphi}\alpha}]$ é um vetor 3×1, com elementos dados por

$$I_{\varphi_j\alpha} = \mathrm{E}\left(\frac{4}{\alpha^3}\boldsymbol{V}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_j}\boldsymbol{\Sigma}^{-1}\right)\boldsymbol{V}\right)=\frac{1}{\alpha}\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_j}\right),j=1,2,3,$$

em que $\mathrm{tr}(\boldsymbol{A})$ denota o traço da matriz $\boldsymbol{A}$; $\boldsymbol{I}_{\boldsymbol{\varphi\beta}} = (\boldsymbol{I}_{\boldsymbol{\beta\varphi}})^{\top}$; $\boldsymbol{I}_{\alpha\boldsymbol{\beta}} = (\boldsymbol{I}_{\boldsymbol{\beta}\alpha})^{\top}$; e $\boldsymbol{I}_{\alpha\boldsymbol{\varphi}} = (\boldsymbol{I}_{\boldsymbol{\varphi}\alpha})^{\top}$.

5.2.8.4 Apêndice IV: Matriz projeção para o modelo log-linear espacial BS

Considere o modelo log-linear espacial BS definido na equação (5.41). A resposta estimada é dada por $\widehat{\boldsymbol{Y}} = \boldsymbol{X}\widehat{\boldsymbol{\beta}}$. A fim de se determinar a estimativa ML de $\boldsymbol{\beta}$, tem-se que

$$U(\beta_j)=\frac{\partial\ell(\boldsymbol{\theta})}{\partial\beta_j}=\frac{2}{\alpha^2}\left(\boldsymbol{W}^{\top}\odot\boldsymbol{X}_j^{\top}\right)\boldsymbol{\Sigma}^{-1}\boldsymbol{V}-\frac{1}{2}\sum_{i=1}^{n}\tanh\left(\frac{y_i-\boldsymbol{x}_i^{\top}\boldsymbol{\beta}}{2}\right)\boldsymbol{X}_j,\quad j=1,\ldots,p,$$

em que $\boldsymbol{W} = (W_1,\ldots,W_n)^{\top}$, com $W_i = \cosh((y_i-\boldsymbol{x}_i^{\top}\boldsymbol{\beta})/2)$, para $i=1,\ldots,n$; $\boldsymbol{X}_j = (x_{1j},\ldots,x_{nj})^{\top}$, para $j=1,\ldots,p$. Usando o primeiro termo da expansão da série de Taylor,

tem-se que $\boldsymbol{W} \approx \mathbf{1}_{n\times 1}$ e

$$\tanh\left(\frac{y_i - \boldsymbol{x}_i^\top \boldsymbol{\beta}}{2}\right) \approx \operatorname{senh}\left(\frac{y_i - \boldsymbol{x}_i^\top \boldsymbol{\beta}}{2}\right).$$

Então, segue diretamente que

$$\boldsymbol{U}(\boldsymbol{\beta}) = \frac{2}{\alpha^2}\boldsymbol{X}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{V} - \frac{1}{2}\boldsymbol{X}^\top \boldsymbol{V}.$$

Novamente, usando-se expansão por série Taylor, agora para a função $\operatorname{senh}(\cdot)$, tem-se $\boldsymbol{V} \approx (1/2)(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta})$, e então pode-se escrever

$$\boldsymbol{U}(\boldsymbol{\beta}) = \boldsymbol{X}^\top \left(\frac{1}{\alpha^2}\boldsymbol{\Sigma}^{-1} - \frac{1}{4}\boldsymbol{I_n}\right)(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta}).$$

Portanto, a estimativa ML de $\boldsymbol{\beta}$ é dada por

$$\widehat{\boldsymbol{\beta}} = \left(\boldsymbol{X}^\top \left(\frac{1}{\alpha^2}\boldsymbol{\Sigma}^{-1} - \frac{1}{4}\boldsymbol{I_n}\right)\boldsymbol{X}\right)^{-1} \boldsymbol{X}^\top \left(\frac{1}{\alpha^2}\boldsymbol{\Sigma}^{-1} - \frac{1}{4}\boldsymbol{I_n}\right)\boldsymbol{Y}.$$

Consequentemente,

$$\widehat{\boldsymbol{Y}} = \boldsymbol{X}\left(\boldsymbol{X}^\top \left(\frac{1}{\alpha^2}\boldsymbol{\Sigma}^{-1} - \frac{1}{4}\boldsymbol{I_n}\right)\boldsymbol{X}\right)^{-1} \boldsymbol{X}^\top \left(\frac{1}{\alpha^2}\boldsymbol{\Sigma}^{-1} - \frac{1}{4}\boldsymbol{I_n}\right)\boldsymbol{Y}.$$

Então, a matriz projeção para o modelo log-linear espacial BS é dada por

$$\boldsymbol{H} = \boldsymbol{X}\left(\boldsymbol{X}^\top \left(\frac{1}{\alpha^2}\boldsymbol{\Sigma}^{-1} - \frac{1}{4}\boldsymbol{I_n}\right)\boldsymbol{X}\right)^{-1} \boldsymbol{X}^\top \left(\frac{1}{\alpha^2}\boldsymbol{\Sigma}^{-1} - \frac{1}{4}\boldsymbol{I_n}\right)$$

avaliada em $\widehat{\boldsymbol{\theta}}$.

5.3 ARTIGO 3: Diagnósticos de influência em um modelo log-linear Birnbaum-Saunders

Resumo: Uma etapa importante durante o estudo de modelos estatísticos é a análise de diagnósticos. Neste trabalho, foi realizada uma análise de diagnóstico em um modelo log-linear espacial Birnbaum-Saunders para avaliar a sensibilidade do estimador de máxima verossimilhança a pontos atípicos. A análise de diagnóstico foi baseada em técnicas gráficas de influência global, construídas a partir da técnica usual; apresentada por Cook (1986) e de técnicas para a análise de influência local, que perturbam a variável-resposta ou uma das covariáveis. A metodologia discutida no trabalho foi aplicada a um conjunto de dados reais. Do ponto de vista teórico, conclui-se que o trabalho contribui para a teoria geoestatística, uma vez que apresenta o desenvolvimento de métodos de diagnósticos para modelos espaciais log-lineares Birnbaum-Saunders, que possibilitam modelar a dependência espacial de dados experimentais, estritamente positivos e que apresentam evidências de distribuição assimétrica positiva. Na prática, a análise de diagnóstico por diferente técnicas apontam, basicamente, os mesmos pontos influentes. Alguns desses pontos também foram apontados como *outliers*, outros não, concordando com resultados obtidos em outros trabalhos que abordam estudo de influência em dados espacialmente correlacionados.
Palavras-chave: Alavanca generalizada, influência global, influência local, perturbação de Zhu, verossimilhança modificada.

5.3.1 Introdução

Uma etapa importante no estudo de modelos estatísticos é o diagnóstico de influência, isto é, verificar a existência de observações que interferem nas estimativas dos parâmetros do modelo. Para avaliar a influência de observações individuais sobre a estimação dos parâmetros do modelo, Cook (1977) propõe, para modelos lineares com resposta normal, uma técnica de deleção individual. Esta técnica de influência global ficou conhecida como distância de Cook e posteriormente foi desenvolvida para outros modelos. Por exemplo, Ross (1987) discutiu a deleção de casos em modelos de regressão não lineares e Galea, Riquelme e Paula (2000) investigaram a metodologia em modelos elípticos multivariados. Para modelos espaciais, podem-se citar os trabalhos de Christensen, Johnson e Pearson (1992, 1993). Além disso, surgiram outras propostas para estudar a influência a partir da deleção de casos, muitas delas baseadas na distância de Cook, como as propostas de Amaral, Flores e Cysneiros (2013) e Pan, Fei e Foster (2014). Por se tratar de uma proposta de deleção individual, um problema que ocorre com a técnica de influência global é deixar de detectar pontos que possivelmente exerçam influência conjunta.

Uma metodologia que permite avaliar a influência conjunta das observações, denominada influência local, foi proposta por Cook (1986). Tal metodologia propõe avaliar a influência das observações após uma pertubação no modelo e/ou nos dados utilizando curvatura normal. Desde a publicação de Cook (1986), muitos trabalhos utilizaram a influência local para

realizar o diagnóstico de influência. Dentre os trabalhos que utilizaram a pertubação aditiva, pode-se citar Paula (1993), que realizou um estudo de influência local em modelos gaussianos lineares com restrições nos parâmetros; Galea, Paula e Uribe-Opazo (2003) que apresentaram estudos de influência local em modelos de contornos elípticos. Leiva et al. (2007) e Leiva et al. (2014c) também realizaram um estudo de influência local em modelos Birnbaum-Saunders. Na área de estatística espacial, Cerioli e Riani (1999) e Uribe-Opazo, Borssoi e Galea (2012) estudaram a influência local em modelos espaciais gaussianos e Assumpção, Uribe-Opazo e Galea (2014) abordaram a metodologia em modelos espaciais t-Student. Saindo da classe das pertubações aditivas, Zhu et al. (2007) propõem uma metodologia para encontrar a pertubação mais adequada para um modelo específico, enquanto Gimenez e Galea (2013) aplicaram a metodologia proposta por Zhu et al. (2007) para realizar um estudo de influência local em modelos heteroscedásticos com erros de medição funcional. Na área espacial De Bastiani et al. (2015) e Garcia-Papani et al. (2016) encontraram a pertubação mais adequada e a utilizaram para a realização de estudos em modelos baseados em distribuições da família das elípticas e na distribuição BS, respectivamente.

Apesar da curvatura normal proposta por Cook (1986) ter sido amplamente utilizada, outras ferramentas foram propostas para realizar o estudo de influência local. Por exemplo, a curvatura normal conformal proposta por Poon e Poon (1999) e a medida apresentada por Tsai (1986) e formalizada por Billor e Loynes (1993). A curvatura normal conformal é equivalente à curvatura normal de Cook (1986), porém, é mais eficiente por ser normalizada e invariante sob reparametrizações adequadas. Zhu e Lee (2001) utilizaram a curvatura normal conformal para realizar um estudo em modelos de dados incompletos. Na área espacial, De Bastiani et al. (2015) utilizaram curvatura conformal para estudar influência local em modelos espaciais com resposta na família de distribuição elípticas. A medida definida por Billor e Loynes (1993) utiliza como medida de influência o afastamento da função verossimilhança modificada, dado por $\mathrm{LD}^*(\boldsymbol{\omega}) = -2(\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widehat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}|\boldsymbol{\omega}))$. Como a primeira derivada de LD^* não se anula no ponto de não pertubação, diferentemente do que ocorre com a medida proposta por Cook (1986), tais autores propõem medir a influência pela inclinação do gráfico de influência. Billor e Loynes (1999) utilizaram a medida que definiram anteriormente para realizar um estudo de influência local em modelos de regressão *ridge*. Cadigan e Farrell (2002) definiram uma medida de influência local generalizada, da qual a medida proposta por Billor e Loynes (1993) é um caso particular e Leiva et al. (2015) fizeram um estudo utilizando esta medida de influência local alternativa em modelos de regressão elípticos com restrições estocásticas.

Assim, o objetivo deste artigo foi realizar um estudo de diagnóstico de influência no modelo log-linear espacial, cuja média é uma combinação linear de covariáveis, baseado na distribuição Birnbaum-Saunders (BS). Mais especificamente, foi realizado um estudo de influência global utilizando a técnica de deleção individual proposta por Pan, Fei e Foster (2014), o qual faz uma adaptação na proposta de Cook (1977); um estudo de influência local, com pertubações nas covariáveis e na variável-resposta, considerando a pertubação mais adequada para o modelo de acordo com a proposta de Zhu et al. (2007). A influência das observações foi

quantificada pela curvatura normal conformal e pela medida dada pela inclinação do gráfico de influência proposta por Tsai (1986) e Billor e Loynes (1993). Adicionalmente, foi realizado um estudo de alavanca generalizada (WEI; HU e FUNG, 1982), para avaliar a influência da resposta observada sobre o correspondente valor ajustado. A metodologia desenvolvida foi aplicada em um estudo com dados experimentais.

A organização do artigo ocorreu da seguinte forma: Na Seção 5.3.2, apresentam-se o modelo log-linear espacial BS e a estimação dos parâmetros por máxima verossimilhança (ML). Na Seção 5.3.3, são revistos alguns conceitos relacionados à análise de diagnósticos e apresenta-se um estudo para a seleção do esquema de pertubação mais apropriado, de acordo com a metodologia proposta por Zhu et al. (2007). Na Seção 5.3.4, aplica-se a metodologia desenvolvida em um estudo com dados experimentais. E, finalmente, na Seção 5.3.5, apresentam-se as conclusões.

5.3.2 Modelo espacial Birnbaum-Saunders

Nesta Seção, formula-se um modelo log-linear espacial BS considerando a média como uma combinação linear de covariáveis e apresenta-se a estimação dos parâmetros.

5.3.2.1 Formulação do modelo

Para modelar um conjunto de dados espacialmente correlacionados, considera-se um processo estocástico $\{T(\boldsymbol{s}), \boldsymbol{s} \in D\}$, definido sobre uma região $D \subset \mathbb{R}^2$. Supõe-se que o processo estocástico seja estacionário e isotrópico, e que para localizações espaciais conhecidas $\boldsymbol{s_i}, i = 1, \ldots, n$, o valor da variável-resposta na posição $\boldsymbol{s_i}$, $T(\boldsymbol{s_i})$, seja dado por

$$T(\boldsymbol{s_i}) = \exp(\mu(\boldsymbol{s_i})\, \eta(\boldsymbol{s_i})), \quad i = 1, \ldots, n, \tag{5.48}$$

em que, $\mu(\boldsymbol{s_i}) = \boldsymbol{X_i}^\top \boldsymbol{\beta}$, com $\boldsymbol{X_i} = (X_{i0}, \ldots, X_{ip})^\top$, $(p+1) \times 1$, com $X_{i0} = 1 \quad i = 1, \ldots, n$; $X_{ij} = X_j(\boldsymbol{s_i}), \quad j = 1, \ldots, p$, isto é, X_{ij} é o valor da covariável X_j na localização $\boldsymbol{s_i}$; $\boldsymbol{\beta} = (\beta_0, \ldots, \beta_p)^\top$, é o vetor de parâmetros desconhecidos, $(p+1) \times 1$, ; e $\eta(\boldsymbol{s_i})$ é o erro aleatório em $\boldsymbol{s_i}$. Supõe-se que $\boldsymbol{\eta} = (\eta(s_1), \ldots, \eta(s_n)) \sim \mathrm{BS}_n(\alpha \mathbf{1}, \mathbf{1}; \boldsymbol{\Sigma}), \quad \alpha > 0$.

O modelo espacial definido na equação (5.48) pode ser linearizado aplicando-se a transformação logaritmo. Assim,

$$Y(\boldsymbol{s_i}) = \log(T(\boldsymbol{s_i})) = \boldsymbol{X_i}^\top \boldsymbol{\beta} + \varepsilon(\boldsymbol{s_i}), \quad i = 1, \ldots, n, \tag{5.49}$$

em que $\varepsilon(\boldsymbol{s_i}) = \log(\eta(\boldsymbol{s_i}))$.

Equivalentemente, em notação matricial, o modelo espacial log-linear BS, definido na equação (5.49), pode ser escrito da seguinte forma,

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{5.50}$$

em que $\boldsymbol{Y} = (Y_1, \ldots, Y_n)^\top, n \times 1, \quad Y_i = \log(T_i)$ é o vetor de variáveis resposta;
$\boldsymbol{X} = \left(\boldsymbol{X_1}^\top, \ldots, \boldsymbol{X_n}^\top\right)^\top, n \times (p+1)$ é a matriz das covariáveis e $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_n)^\top, n \times 1$, é o vetor

dos erros aleatórios do processo estacionário. Tem-se que $\boldsymbol{\varepsilon} \sim \log\text{-BS}_n(\alpha\mathbf{1}, \mathbf{0}; \boldsymbol{\Sigma})$, portanto, $\mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}$.

Assume-se ainda que a dependência espacial é determinada pela matriz escala $\boldsymbol{\Sigma}, n \times n$, simétrica, não singular e positiva definida, proporcional a $\mathrm{Cov}(\boldsymbol{\varepsilon})$, e que $\mathrm{Cov}(\varepsilon(\boldsymbol{s_i}), \varepsilon(\boldsymbol{s_j}))$ depende apenas da distância euclidiana entre as localizações $\boldsymbol{s_i}$ e $\boldsymbol{s_j}$, isto é, $\mathrm{Cov}(\varepsilon(\boldsymbol{s_i}), \varepsilon(\boldsymbol{s_j})) = \mathrm{C}(h_{ij})$, em que $h_{ij} = \|\boldsymbol{s_i} - \boldsymbol{s_j}\|$. Além disso, assume-se que $\mathrm{C}(\boldsymbol{s_i}, \boldsymbol{s_j})$ é uma função do vetor de parâmetros $\boldsymbol{\varphi} = (\varphi_1, \varphi_2, \varphi_3)^\top$, dada por

$$\boldsymbol{\Sigma} = \varphi_1 \boldsymbol{I}_n + \varphi_2 \boldsymbol{R}(\varphi_3), \tag{5.51}$$

em que, φ_1 é o efeito pepita ou erro de variância; φ_2 é a contribuição ou variância de dispersão; φ_3 é função do alcance ou raio de dependência espacial do modelo; $\boldsymbol{I}_n$ é a matriz identidade $n \times n$ e $\boldsymbol{R}(\varphi_3) = [r_{ij}]$ é uma matriz simétrica, $n \times n$, com elementos da diagonal principal iguais a 1. $\boldsymbol{R}(\varphi_3)$ depende do modelo teórico de variância adotado para descrever a dependência espacial, por exemplo, nos modelos família Matérn, que adotaremos neste trabalho, r_{ij} é dado por

$$r_{ij} = \begin{cases} 1, & i = j, \\ \frac{1}{2^{\delta-1}\Gamma(\delta)} \left(\frac{h_{ij}}{\varphi_3}\right)^{\delta} K_\delta \left(\frac{h_{ij}}{\varphi_3}\right), & i \neq j, \end{cases} \tag{5.52}$$

em que δ é um parâmetro de forma e $\Gamma(\cdot), K_\delta(\cdot)$ são, respectivamente, as funções gamma e Bessel modificada de terceiro tipo de ordem δ.

Consequentemente, sob as especificações do modelo BS, a covariância é dada por

$$C(h_{ij}) = \begin{cases} \alpha^2(\varphi_1 + \varphi_2), & i = j, \\ \frac{\alpha^2\varphi_2}{2^{\delta-1}\Gamma(\delta)} \left(\frac{h_{ij}}{\varphi_3}\right)^{\delta} K_\delta \left(\frac{h_{ij}}{\varphi_3}\right), & i \neq j \end{cases}$$

Os modelos teóricos das funções covariâncias, exponêncial e gaussiano são escritos respectivamente por

$$C(h_{ij}) = \begin{cases} \alpha^2(\varphi_1 + \varphi_2), & i = j, \\ \alpha^2\varphi_2 \exp(-(h_{ij}/\varphi_3)), & i \neq j, \end{cases}$$

e

$$C(h_{ij}) = \begin{cases} \alpha^2(\varphi_1 + \varphi_2), & i = j, \\ \alpha^2\varphi_2 \exp(-(h_{ij}/\varphi_3)^2), & i \neq j, \end{cases}$$

são casos particulares da família de modelos Matérn, quando $\delta = 0,5$ e $\delta \to \infty$, respectivamente.

5.3.2.2 Estimação dos parâmetros

Neste artigo, os parâmetros do modelo foram estimados pelo método ML e como o sistema derivado da função escore não possui solução analítica, a solução do mesmo se dará por um processo iterativo para solução de sistemas não lineares.

O vetor de parâmetros desconhecidos do modelo espacial definido em (5.50), $\boldsymbol{\theta} = \left(\boldsymbol{\beta}^\top, \boldsymbol{\varphi}^\top, \alpha\right)^\top$, pode ser estimado maximizando-se o logaritmo da função verossimilhança. Uma vez que $\boldsymbol{\varepsilon} = \left(\boldsymbol{Y} - \boldsymbol{X}^\top\boldsymbol{\beta}\right) \sim \text{log-BS}_n\left(\alpha\mathbf{1}, \mathbf{0}; \boldsymbol{\Sigma}\right)$, o logaritmo da função verossimilhança para $\boldsymbol{\theta}$, a menos de uma constante, é dado por

$$\ell(\boldsymbol{\theta}) = -\frac{1}{2}\log\left(|\boldsymbol{\Sigma}|\right) - n\log(\alpha) - \frac{2}{\alpha^2}\boldsymbol{V}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{V} + \sum_{i=1}^{n}\log\left(\cosh\left(\frac{y_i - \boldsymbol{X}_i^\top\boldsymbol{\beta}}{2}\right)\right), \tag{5.53}$$

em que, $\boldsymbol{V} = (V_1, \ldots, V_n)^\top$ é um vetor $n \times 1$ com elementos $V_i = \text{senh}((y_i - \boldsymbol{X}_i^\top\boldsymbol{\beta})/2),\ i = 1, \ldots n$, e $\boldsymbol{\Sigma} = \varphi_1\boldsymbol{I}_n + \varphi_2\boldsymbol{R}\left(\varphi_3\right)$, como em (5.51).

Para o modelo log-linear espacial BS, o vetor escore, $(p+5) \times 1$, é dado por

$$U\left(\boldsymbol{\theta}\right) = \left(\left(\frac{\partial\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\beta}}\right)^\top, \left(\frac{\partial\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\varphi}}\right)^\top, \frac{\partial\ell(\boldsymbol{\theta})}{\partial\alpha}\right)^\top = \left(U(\beta_0), \ldots, U(\beta_p), U(\varphi_1), U(\varphi_2), U(\varphi_3), U(\alpha)\right)^\top.$$

Pela equação (5.53), tem-se

$$\begin{aligned}
U\left(\beta_j\right) &= \frac{\partial\ell(\boldsymbol{\theta})}{\partial\beta_j} = \frac{\partial}{\partial\beta_j}\left(-\frac{2}{\alpha^2}\boldsymbol{V}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{V} + \sum_{i=1}^{n}\log\left(\cosh\left(\frac{y_i - \boldsymbol{X}_i^\top\boldsymbol{\beta}}{2}\right)\right)\right) \\
&= \frac{2}{\alpha^2}\left(\boldsymbol{W}^\top \odot \boldsymbol{X}_{\cdot j}^\top\right)\boldsymbol{\Sigma}^{-1}\boldsymbol{V} - \frac{1}{2}\sum_{i=1}^{n}\tanh\left(\frac{y_i - \boldsymbol{X}_i^\top\boldsymbol{\beta}}{2}\right)X_{ij}, \quad j = 0, \ldots, p; \\
U\left(\varphi_j\right) &= \frac{\partial\ell(\boldsymbol{\theta})}{\partial\varphi_j} = \frac{\partial}{\partial\varphi_j}\left(-\frac{1}{2}\log\left(|\boldsymbol{\Sigma}|\right) - \frac{2}{\alpha^2}\boldsymbol{V}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{V}\right) \\
&= -\frac{1}{2}\text{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\varphi_j}\right) + \frac{2}{\alpha^2}\boldsymbol{V}^\top\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\varphi_j}\boldsymbol{\Sigma}^{-1}\boldsymbol{V}, \quad j = 1, 2, 3; \\
U\left(\alpha\right) &= \frac{\partial\ell(\boldsymbol{\theta})}{\partial\alpha} = \frac{\partial}{\partial\alpha}\left(-n\log(\alpha) - \frac{2}{\alpha^2}\boldsymbol{V}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{V}\right) = -\frac{n}{\alpha} + \frac{4}{\alpha^3}\boldsymbol{V}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{V}.
\end{aligned}$$

em que, $\odot$ denota o produto de Hadamard; $\text{tr}\left(\boldsymbol{A}\right)$ denota o traço da matriz $\boldsymbol{A}$; $\boldsymbol{W} = (W_1, \ldots, W_n)^\top$, com $W_i = \cosh((y_i - \boldsymbol{X}_i^\top\boldsymbol{\beta})/2)$, para $i = 1, \ldots, n$ e $\boldsymbol{V} = (V_1, \ldots, V_n)^\top$ como dado em (5.53); e $\boldsymbol{X}_{\cdot j} = \left(X_{1j}, \ldots, X_{nj}\right)^\top$, para $j = 0, \ldots p$. Cabe lembrar que $X_{i0} = 1$, para $i = 1, \ldots, n$ e que $\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_1} = \boldsymbol{I}_n$, $\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_2} = \boldsymbol{R}$, $\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_3} = \varphi_2\frac{\partial\boldsymbol{R}}{\partial\varphi_3}$. Para o modelo Matérn dado em (5.3.2.1) $\partial\boldsymbol{R}/\partial\varphi_3 = [\partial r_{ij}/\partial\varphi_3]$, em que

$$\frac{\partial r_{ij}}{\partial\varphi_3} = -\left(\frac{1}{\varphi_3}\right)\left(\delta r_{ij} + \frac{1}{2^{\delta-1}\Gamma(\delta)}\left(\frac{h_{ij}}{\varphi_3}\right)^{\delta+1}K_\delta'\left(\frac{h_{ij}}{\varphi_3}\right)\right) \tag{5.54}$$

para $i \neq j, i, j = 1, \ldots, n$, com $K_\delta'(u) = -(1/2)(K_{\delta-1}(u) + K_{\delta+1}(u))$.

Assim, a fim de se obter a estimativa $\widehat{\boldsymbol{\theta}}$ de $\boldsymbol{\theta}$, deve-se solucionar o sistema não linear $\boldsymbol{U}\left(\boldsymbol{\theta}\right) = \mathbf{0}$. Como esse sistema não apresenta solução análitica explicita, $\widehat{\boldsymbol{\theta}}$ deve ser obtido por um processo iterativo para obtenção de soluções de sistemas não lineares (NOCEDAL; WRIGHT, 1999; LANGE, 2000). Podem-se também encontrar, assintoticamente, os desvios padrões dos parâmetros estimados, e ainda realizar teste de hipóteses do tipo $H_0 : \theta_i = 0$ *versus* $H_1 : \theta_i \neq 0$, em que θ_i é qualquer um dos parâmetros do vetor de parâmetros $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\varphi}^\top, \alpha)^\top$, utilizando a inversa da matriz de informação de Fisher, uma vez que $\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{D} \text{N}_{(p+1)}(\mathbf{0}, \boldsymbol{J}(\boldsymbol{\theta})^{-1})$,

quando $n \to \infty$, com $\boldsymbol{J}(\boldsymbol{\theta}) = \lim_{n\to\infty}[1/n]\boldsymbol{I}(\boldsymbol{\theta})$, em que $\boldsymbol{I}(\boldsymbol{\theta})$ é a matriz de informação esperada de Fisher e $\xrightarrow{D}$ denota convergência da distribuição. Para o modelo log-linear espacial BS, a matriz de informação esperada é a matriz $(p+5)\times(p+5)$, bloco diagonal, dada por

$$\boldsymbol{I}(\boldsymbol{\theta}) = \begin{pmatrix} \boldsymbol{I}_{\boldsymbol{\beta\beta}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{I}_{\boldsymbol{\varphi\varphi}} & \boldsymbol{I}_{\boldsymbol{\varphi}\alpha} \\ \mathbf{0} & \boldsymbol{I}_{\alpha\boldsymbol{\varphi}} & I_{\alpha\alpha} \end{pmatrix} = \begin{pmatrix} \boldsymbol{I}_{\boldsymbol{\beta\beta}} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{I}_{\boldsymbol{\psi\psi}} \end{pmatrix} \tag{5.55}$$

em que $\boldsymbol{I}_{\boldsymbol{\beta\beta}}$ denota a matriz de informação esperada para o vetor $\boldsymbol{\beta} = (\beta_0, \ldots, \beta_p)^\top$ e $\boldsymbol{I}_{\boldsymbol{\psi\psi}}$ é a matriz de informação esperada para $\boldsymbol{\psi} = \left(\boldsymbol{\varphi}^\top, \alpha\right)^\top$. Veja detalhes do vetor escore e da matriz de informação esperada nos Apêndices I e II.

5.3.3 Diagnósticos de Influência

A realização de diagnósticos de influência é uma etapa importante da análise estatística. Excluir uma observação do conjunto de dados é, provavelmente, a maneira mais natural de se detectar sua influência no ajuste do modelo. Cook (1977) propõe uma metodologia baseada na deleção de casos para estudar a influência de uma observação particular sobre a estimativa dos parâmetros do modelo. Além da metodologia proposta Cook (1977) surgiram outras propostas que generalizaram tal metodologia, como as propostas de Pan, Fei e Foster (2014).

Um problema que pode ocorrer com a deleção individual de pontos é a não detecção de pontos conjuntamente influentes. Uma metodologia que propõe avaliar a influência conjunta das observações, denominada influência local, foi proposta por Cook (1986). Além da metodologia proposta por Cook (1986), surgiram outras propostas para realizar estudo de influência local, por exemplo, a medida apresentada por Tsai (1986) e definida por Billor e Loynes (1993), a qual utiliza o afastamento da verossimilhança modificada e a curvatura normal conformal, proposta por Poon e Poon (1999).

Destaca-se ainda a alavanca generalizada proposta por Wei, Hu e Fung (1998) para avaliar a influência da resposta observada sobre o correspondente valor ajustado.

Nesta Seção, apresenta-se a medida baseada em eliminação de casos abordada por Pan, Fei e Foster (2014), que faz adaptação na proposta de Cook (1977); o método de influência local, utilizando as medidas propostas por Poon e Poon (1999) e Billor e Loynes (1993) e o método "alavanca generalizada"proposto por Wei, Hu e Fung (1998), aplicados ao estudo do modelo log-linear espacial BS.

5.3.3.1 Influência global

Cook (1977), propôs para mensurar a mudança dos parâmetros estimados quando uma observação é deletada do conjunto de dados, a medida $\mathrm{LD}(\hat{\boldsymbol{\theta}}_{(i)}) = 2(\ell(\hat{\boldsymbol{\theta}}) - \ell(\hat{\boldsymbol{\theta}}_{(i)}))$, em que $\ell(\cdot)$ é o logaritmo da função verossimilhança; $\hat{\boldsymbol{\theta}}$ e $\hat{\boldsymbol{\theta}}_{(i)}$ são, respectivamente, as estimativas por ML de $\boldsymbol{\theta}$, considerando o conjuntos de dados completo e o conjunto de dados sem a i-ésima observação.

A partir da distância de Cook, utilizando expansão de Taylor em torno de $\hat{\boldsymbol{\theta}}$ e o estimador a um passo de Newton-Raphson, Pan, Fei e Foster (2014) obtiveram uma nova medida de influência global dada por

$$\mathrm{LD}^1(\hat{\boldsymbol{\theta}}_{(i)}) = (\dot{\ell}_{(i)}(\hat{\boldsymbol{\theta}}))^\top(-\ddot{\ell}_{(i)}(\hat{\boldsymbol{\theta}}))^{-1}(\dot{\ell}_{(i)}(\hat{\boldsymbol{\theta}})),$$

em que $\ell_{(i)}(\cdot)$ é o logaritmo da função verossimilhança obtida após a exclusão da $i-$ésima observação, $\dot{\ell}_{(i)}(\boldsymbol{\theta}) = \frac{\partial \ell_{(i)}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ e $\ddot{\ell}_{(i)}(\boldsymbol{\theta}) = \frac{\partial \ell^2_{(i)}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}^\top}$.

Pan, Fei e Foster (2014) mostraram ainda que é possível substituir $\ddot{\ell}_{(i)}(\hat{\boldsymbol{\theta}})$ por $\ddot{\ell}(\hat{\boldsymbol{\theta}})$ ou ainda por $I(\boldsymbol{\theta}) = \mathrm{E}(-\ddot{\ell}(\boldsymbol{\theta}))$, obtendo-se, respectivamente, as medidas de influência

$$\begin{aligned}
\mathrm{D}^1_i(\boldsymbol{\theta}) &= (\dot{\ell}_{(i)}(\hat{\boldsymbol{\theta}}))^\top(-\ddot{\ell}(\hat{\boldsymbol{\theta}}))^{-1}(\dot{\ell}_{(i)}(\hat{\boldsymbol{\theta}})) \\
\mathrm{D}_i(\boldsymbol{\theta}) &= (\dot{\ell}_{(i)}(\hat{\boldsymbol{\theta}}))^\top \boldsymbol{I}(\hat{\boldsymbol{\theta}})^{-1}(\dot{\ell}_{(i)}(\hat{\boldsymbol{\theta}})).
\end{aligned}$$

A medida $\mathrm{D}^1_i(\boldsymbol{\theta})$, é preferível à medida $\mathrm{LD}^1(\hat{\boldsymbol{\theta}}_{(i)})$, porque reduz os cálculos computacionais. A medida $\mathrm{D}_i(\boldsymbol{\theta})$ é preferível pois pode permitir a decomposição da distância de Cook em diferentes componentes de interesse. No caso do modelo log-linear BS, a matriz $\boldsymbol{I}(\boldsymbol{\theta})$ é bloco diagonal (veja equação 5.55) e permite calcular a distância de Cook para o vetor $\boldsymbol{\beta}$ e o vetor $\boldsymbol{\psi}$, dado por $\boldsymbol{\psi}^\top = \left(\boldsymbol{\varphi}^\top, \alpha\right)$, isto é,

$$\begin{aligned}
\mathrm{D}_i(\boldsymbol{\theta}) &= (\dot{\ell}_{(i)}(\hat{\boldsymbol{\theta}}))^\top \boldsymbol{I}(\hat{\boldsymbol{\theta}})^{-1}(\dot{\ell}_{(i)}(\hat{\boldsymbol{\theta}})) \\
&= (\dot{\ell}_{(i)}(\hat{\boldsymbol{\beta}}))^\top \boldsymbol{I}(\hat{\boldsymbol{\beta}})^{-1}(\dot{\ell}_{(i)}(\hat{\boldsymbol{\beta}})) + (\dot{\ell}_{(i)}(\hat{\boldsymbol{\psi}}))^\top \boldsymbol{I}(\hat{\boldsymbol{\psi}})^{-1}(\dot{\ell}_{(i)}(\hat{\boldsymbol{\psi}})) \\
&= \mathrm{D}_i(\boldsymbol{\beta}) + \mathrm{D}_i(\boldsymbol{\psi}).
\end{aligned}$$

em que, $\hat{\boldsymbol{\beta}}$ e $\hat{\boldsymbol{\psi}}$ são os estimadores ML para o vetor $\boldsymbol{\beta}$ e o vetor $\boldsymbol{\psi} = \left(\boldsymbol{\varphi}^\top, \alpha\right)^\top$; $\dot{\ell}_{(i)}(\hat{\boldsymbol{\beta}})$; $\dot{\ell}_{(i)}(\hat{\boldsymbol{\psi}})$ são subvetores de $\dot{\ell}_{(i)}(\hat{\boldsymbol{\theta}})$ referentes aos vetores $\boldsymbol{\beta}$ e $\boldsymbol{\psi}$, respectivamente; e $\boldsymbol{I}(\boldsymbol{\beta})$ e $\boldsymbol{I}(\boldsymbol{\psi})$ são os blocos da matriz de Fisher referentes aos vetores $\boldsymbol{\beta}$ e $\boldsymbol{\psi}$.

Se o valor de $\mathrm{D}_i(\boldsymbol{\theta})$ é alto, a observação i é, possivelmente, um ponto influente. Não existe um consenso sobre quais valores devem ser considerados altos. Cook e Weisberg (1982) colocam que a definição exata de alto depende do contexto. Um valor alto de $\mathrm{D}_i(\boldsymbol{\beta})$ indica que a observação i é um possível ponto influente na estimativa do vetor $\boldsymbol{\beta}$, e um valor alto de $\mathrm{D}_i(\boldsymbol{\psi})$ indica que a observação i é um ponto influente na estimativa do vetor $\boldsymbol{\psi}$.

5.3.3.2 Influência local

Influência local consiste em estudar o comportamento dos parâmetros estimados mediante uma pequena pertubação nos dados ou nas suposições do modelo. Cook (1986) propõe medir a influência pelo estudo da curvatura normal do gráfico de influência $\boldsymbol{\omega} \times \mathrm{LD}(\boldsymbol{\omega})$ em uma vizinhança de ω_0, na direção de um vetor unitário $\boldsymbol{d}$, em que $\mathrm{LD}(\boldsymbol{\omega}) = 2(\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widehat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}))$, com $\widehat{\boldsymbol{\theta}}$ e $\widehat{\boldsymbol{\theta}}_{\omega}$ respectivamente, os estimadores ML de $\boldsymbol{\theta}$ no modelo proposto e no modelo perturbado por $\boldsymbol{\omega}$. Tal autor mostrou que a curvatura normal na direção $\boldsymbol{d}$ assume a forma $C_{\boldsymbol{d}} = 2|\boldsymbol{d}^\top \boldsymbol{B}\boldsymbol{d}|$, em que $\boldsymbol{B} = -\boldsymbol{\Delta}\ddot{\boldsymbol{\ell}}(\widehat{\boldsymbol{\theta}})^{-1}\boldsymbol{\Delta}$, com $\ddot{\boldsymbol{\ell}}(\widehat{\boldsymbol{\theta}})$ a matriz Hessiana avaliada em $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$, e $\boldsymbol{\Delta} = \frac{\partial \ell^2(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \boldsymbol{\theta}\partial \boldsymbol{\omega}^\top}$ a

matriz de pertubação avaliada em $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ e $\boldsymbol{\omega} = \boldsymbol{\omega_0}$. Veja detalhes da matriz Hessiana e da matriz de pertubação nos Apêndices I e III, respectivamente.

Segundo Cook (1986), uma direção importante a se considerar é a direção da curvatura normal máxima, $\boldsymbol{d} = \boldsymbol{d}_{max}$. A curvatura normal máxima, $C_{\boldsymbol{d}_{max}}$, correponde ao maior autovalor (em valor absoluto) da matriz $\boldsymbol{B}$ e $\boldsymbol{d}_{max}$ é o autovetor associado a esse autovalor. O gráfico de $|C_{\boldsymbol{d}_{max}}|$ contra a ordem das observações pode revelar os pontos com maior influência na vizinhança de $\mathrm{LD}(\boldsymbol{\omega}_0)$. Outra direção importante é $\boldsymbol{d} = \boldsymbol{e_i}$, em que $\boldsymbol{e_i}$ é o vetor básico do $\mathbb{R}^n$, cuja i-ésima coordenada é 1 e as demais são 0. Neste caso, a curvatura normal é dada por $C_i = 2|b_{ii}|$, em que b_{ii} é o i-ésimo elemento da matriz $\boldsymbol{B}$, $i = 1, \ldots, n$. Assim como no caso anterior, o gráfico de C_i contra a ordem das observações pode ser utilizado para identificar pontos influentes.

Embora a curvatura normal de Cook (1986) seja muito usada, outras medidas para avaliar a influência local foram propostas, como a curvatura normal conformal, proposta por Poon e Poon (1999) e a medida de influência local apresentada na discusão de Cook (1986) por Tsai (1986) e formalmente definida por Billor e Loynes (1993), que utiliza o afastamento pela função verossimilhança modificada, definido por $\mathrm{LD}^*(\boldsymbol{\omega}) = -2(\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widehat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}|\boldsymbol{\omega}))$, em que $\ell(\boldsymbol{\theta}_{\boldsymbol{\omega}}|\boldsymbol{\omega})$ é a verossimilhança perturbada por ω.

A curvatura normal conformal, $B_d = \frac{C_d}{\mathrm{tr}(\boldsymbol{B})}$, possui como vantagens os fatos de ser invariante sob reparametrizações conformais e de ser uma medida normalizada, o que facilita estabelecer um ponto de corte. O autovetor $\boldsymbol{d}_{max}$ fornece também a curvatura conformal máxima $B_{\boldsymbol{d}_{max}}$. Poon e Poon (1999) sugerem que as coordenadas do vetor $B_{d_{max}}$, tomadas em módulo, que apresentam valores maiores que $\frac{1}{\sqrt{n}}$ indicam pontos influentes. No que diz respeito às curvaturas normais conformais na direção dos vetores básicos, $B_i = B_{\boldsymbol{e_i}} = \frac{2|b_{ii}|}{\mathrm{tr}(\boldsymbol{B})}$. Poon e Poon (1999) sugerem que a i-ésima observação seja considerada influente se $B_i > 2\overline{B}$, em que $\overline{B}$ é a média aritmética das curvaturas normais conformais básicas. Zhu e Lee (2001) sugerem um ponto de corte que além de considerar a média das curvaturas normais conformais também leva em consideração a variação dessas curvaturas. Considerando um caso particular da proposta de Zhu e Lee (2001), tem-se que a i-ésima observação é influente quando $B_i > \overline{B} + 2\mathrm{SE}(B)$, em que $\mathrm{SE}(B)$ é o desvio padrão das curvaturas normais conformais básicas.

A medida que se utiliza do afastamento pela função verossimilhança modificada, definido por $\mathrm{LD}^*(\boldsymbol{\omega}) = -2(\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widehat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}|\boldsymbol{\omega}))$, em que $\ell(\boldsymbol{\theta}_{\boldsymbol{\omega}}|\boldsymbol{\omega})$ é a verossimilhança perturbada por $\boldsymbol{\omega}$, apresenta como principal vantagem o fato da primeira derivada não se anular no ponto de não pertubação, e então permite medir a influência por estudo da inclinação do gráfico de influência $\boldsymbol{\omega} \times \mathrm{LD}^*(\boldsymbol{\omega})$ em uma vizinhança de $\boldsymbol{\omega_0}$, na direção de um vetor unitário $\boldsymbol{d}$. Em particular, a inclinação máxima, S_{max}, e a correspondente direção máxima, d^*_{max}, devem ser consideradas. A direção máxima de LD^* no ponto $\boldsymbol{\omega}_0$ é a direção do gradiente $\nabla LD^*(\boldsymbol{\omega}_0)$ e, consequentemente, a inclinação máxima é dada por (Lobato Júnior (2005), Leiva et al.(2015c)),

$$S_{max} = \|\nabla \mathrm{LD}^*(\boldsymbol{\omega}_0)\| = 2\left\|\frac{\partial \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \omega}\right\|$$

avaliada em $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ e $\boldsymbol{\omega} = \boldsymbol{\omega}_0$. Tal abordagem permite que se tenham os valores individuais das coordenadas do gradiente, $d^*_{max} = \nabla LD^*(\boldsymbol{\omega}_0)$ e assim pode-se avaliar quais observações contribuem com a maior parcela para o valor da inclinação máxima. Deste modo, o gráfico das coordenadas de S_{max} contra a ordem das observações pode ser utilizado para identificar os pontos influentes.

Neste trabalho, foi realizado um estudo de influência local, para o modelo espacial log-linear BS, considerando pertubação na variável-resposta e em uma das variáveis explicativas. O esquema de pertubação utilizado considerou a perturbação mais adequada, segundo a metodologia proposta por Zhu et al. (2007). Como medidas de influência local, foram utilizadas a curvatura normal conformal e a medida que utiliza a inclinação máxima. A matriz de pertubação para cada um dos casos é apresentada no Apêndice III.

Perturbação na variável-resposta

Seja a perturbação na variável resposta dada por

$$\boldsymbol{Y}_{\boldsymbol{\omega}}(\boldsymbol{s}) = \boldsymbol{Y}(\boldsymbol{s}) + \boldsymbol{A}\boldsymbol{\omega},$$

em que $\boldsymbol{A}$ é uma matriz simétrica, não singular e $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_n)^\top \in \mathbb{R}^n$ é o vetor de perturbações. Note que $\boldsymbol{\omega_0} = \boldsymbol{0}, n \times 1$, é o vetor de não perturbação.

O logaritmo da função verossimilhança perturbada é dado por

$$\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = -\frac{1}{2}\log\left(|\boldsymbol{\Sigma}|\right) - n\log(\alpha) - \frac{2}{\alpha^2}\boldsymbol{V}_{\boldsymbol{\omega}}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{V}_{\boldsymbol{\omega}} + \sum_{i=1}^{n}\log\left(\cosh\left(\frac{y_i + \boldsymbol{A}_i\boldsymbol{\omega} - \boldsymbol{X}_i^\top\boldsymbol{\beta}}{2}\right)\right). \quad (5.56)$$

em que $\boldsymbol{V}_{\boldsymbol{\omega}} = (V_{\omega_1}, \ldots, V_{\omega_n})^\top$ com $V_{\omega_i} = \operatorname{senh}\left(\frac{y_i + \boldsymbol{A}_i\boldsymbol{\omega} - \boldsymbol{X}_i^\top\boldsymbol{\beta}}{2}\right)$.

Segundo Zhu et al. (2007), a perturbação $\boldsymbol{\omega}$ é apropriada se, e somente se $\boldsymbol{G}(\boldsymbol{\omega_0}) = c\boldsymbol{I}_n$, em que $c \in \mathbb{R}, c > 0$; $\boldsymbol{I}_n$ é a matriz identidade de ordem n e $\boldsymbol{G}(\boldsymbol{\omega}) = \mathrm{E}\left(\boldsymbol{U}\left(\boldsymbol{\omega}\right)\boldsymbol{U}^\top\left(\boldsymbol{\omega}\right)\right)$, com $\boldsymbol{U}(\boldsymbol{\omega}) = \frac{\partial\ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial\boldsymbol{\omega}}$.

Para o modelo espacial log-linear BS tem-se que $\boldsymbol{U}\left(\boldsymbol{\omega}\right) = -\frac{2}{\alpha^2}\boldsymbol{A}\boldsymbol{\Sigma}^{-1}\boldsymbol{V}_{\boldsymbol{\omega}} + \frac{1}{2}\boldsymbol{A}\boldsymbol{V}_{\boldsymbol{\omega}}$ e, consequentemente, $G(\boldsymbol{\omega}) = \boldsymbol{A}\left(\frac{\alpha}{4}\boldsymbol{\Sigma}^{1/2} - \frac{1}{\alpha}\boldsymbol{\Sigma}^{-1/2}\right)^2\boldsymbol{A}$. Desta forma, para o modelo definido na Seção 5.3.2.1, um esquema de perturbação apropriado para a variável-resposta é dado por

$$\boldsymbol{Y}_{\boldsymbol{\omega}}(\boldsymbol{s}) = \boldsymbol{Y}(\boldsymbol{s}) + \left(\frac{\alpha}{4}\boldsymbol{\Sigma}^{1/2} - \frac{1}{\alpha}\boldsymbol{\Sigma}^{-1/2}\right)^{-1}\boldsymbol{\omega}.$$

Perturbação em uma covariável

Sem perda de generalidade, pode-se supor que a covariável perturbada é a covariável $\boldsymbol{X}_1$. As demais covariáveis não sofrem perturbação, isto é, tem-se um esquema de

perturbação nas covariáveis dado por

$$\begin{aligned} \boldsymbol{X_{1\omega}}(\boldsymbol{s}) &= \boldsymbol{X}_1(\boldsymbol{s}) + \boldsymbol{A\omega}; \\ \boldsymbol{X_{i\omega}}(\boldsymbol{s}) &= \boldsymbol{X}_i(\boldsymbol{s}), \quad i = 2, \ldots, n. \end{aligned}$$

em que $\boldsymbol{A}, \boldsymbol{\omega} \in \mathbb{R}^n$ e $\boldsymbol{\omega_0} = \boldsymbol{0}$ são como no item anterior. Neste esquema de perturbação, o logaritmo da função verossimilhança do modelo perturbado é dado por

$$\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = -\frac{1}{2}\log\left(|\boldsymbol{\Sigma}|\right) - n\log(\alpha) - \frac{2}{\alpha^2}\boldsymbol{V}_{\boldsymbol{\omega}}^{\top}\boldsymbol{\Sigma}^{-1}\boldsymbol{V}_{\boldsymbol{\omega}} + \sum_{i=1}^{n}\log\left(\cosh\left(\frac{y_i - \boldsymbol{X}_{i\boldsymbol{\omega}}^{\top}\boldsymbol{\beta}}{2}\right)\right).$$

em que $\boldsymbol{V}_{\boldsymbol{\omega}} = \left(V_{\omega_1}, \ldots, V_{\omega_n}\right)^{\top}, n \times 1$, com $V_{\omega_i} = \operatorname{senh}\left(\frac{y_i - \boldsymbol{X}_{i\boldsymbol{\omega}}^{\top}\boldsymbol{\beta}}{2}\right), i = 1, \ldots n.$

Desta forma, tem-se (veja Apêndice II), $U(\boldsymbol{\omega}) = \frac{\partial \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \omega} = \frac{2\beta_1}{\alpha^2}\boldsymbol{A\Sigma}^{-1}\boldsymbol{V}_{\boldsymbol{\omega}} - \frac{\beta_1}{2}\boldsymbol{AV}_{\boldsymbol{\omega}}$, e $\boldsymbol{G}(\boldsymbol{\omega}) = \mathrm{E}\left(U\left(\boldsymbol{\omega}\right)U\left(\boldsymbol{\omega}\right)^{\top}\right) = \beta_1^2\boldsymbol{A}\left(\frac{\alpha}{4}\boldsymbol{\Sigma}^{\frac{1}{2}} - \frac{1}{\alpha}\boldsymbol{\Sigma}^{-\frac{1}{2}}\right)^2\boldsymbol{A}$.

Assim, para atender à proposta de Zhu et al. (2007), deve-se considerar, assim como no caso de perturbação na variável- resposta, $\boldsymbol{A} = \left(\frac{\alpha}{4}\boldsymbol{\Sigma}^{\frac{1}{2}} - \frac{1}{\alpha}\boldsymbol{\Sigma}^{-\frac{1}{2}}\right)^{-1}$. Isto é, o esquema de perturbação mais adequado para perturbação em uma das covariáveis é

$$\boldsymbol{X}_{\boldsymbol{\omega}_1}(\boldsymbol{s}) = \boldsymbol{X}_1(\boldsymbol{s}) + \left(\frac{\alpha}{4}\boldsymbol{\Sigma}^{\frac{1}{2}} - \frac{1}{\alpha}\boldsymbol{\Sigma}^{-\frac{1}{2}}\right)^{-1}\boldsymbol{\omega}.$$

5.3.3.3 Alavanca generalizada

A alavanca generalizada foi proposta por Wei, Hu e Fung (1998) com o objetivo de avaliar a influência da resposta observada sobre o correspondente valor ajustado. Segundo Wei, Hu e Fung (1998), a matriz alavanca generalizada $\boldsymbol{GL}(\theta) = \frac{\partial \widehat{Y}}{\partial Y^{\top}}$ tem a forma

$$\boldsymbol{GL}(\boldsymbol{\theta}) = \boldsymbol{D}_{\boldsymbol{\theta}}\left(-\ddot{\boldsymbol{\ell}}(\boldsymbol{\theta})\right)^{-1}\boldsymbol{L}_{\boldsymbol{\theta Y}},$$

em que $\boldsymbol{D}_{\boldsymbol{\theta}} = \frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\theta}^{\top}}$, com $\boldsymbol{\mu} = \boldsymbol{X\beta}$ o valor esperado para Y; $-\ddot{\boldsymbol{\ell}}(\boldsymbol{\theta})$ é a matriz de informação observada e $\boldsymbol{L}_{\boldsymbol{\theta Y}} = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial Y^{\top}}$.

O elemento GL_{ii} da matriz $\boldsymbol{GL}(\hat{\boldsymbol{\theta}})$ que apresentar valores maiores que $\overline{GL} + 2\mathrm{SE}(GL)$, em que $\overline{GL}$ e $\mathrm{SE}(GL)$ são, respectivamente, a média e o desvio padrão de $GL_{11}, GL_{22}, \ldots, GL_{nn}$, são considerados pontos de alavanca, isto é, pontos cujo valor observado exerce grande influência sobre o valor predito (DE BASTIANE et al.,2015).

No caso do modelo espacial log-linear BS, tem-se que

$$\boldsymbol{X} = \begin{pmatrix} 1 & X_{11} & \ldots & X_{1p} \\ 1 & X_{21} & \ldots & X_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & \ldots & X_{np} \end{pmatrix}; \boldsymbol{\theta} = \left(\boldsymbol{\beta}^{\top}, \boldsymbol{\varphi}^{\top}, \alpha\right)^{\top},$$

em que $\boldsymbol{\beta} = \left(\beta_0, \ldots, \beta_p\right)^{\top}$ e $\boldsymbol{\varphi} = \left(\varphi_1, \varphi_2, \varphi_3\right)^{\top}$, de onde segue que $\boldsymbol{D}_{\boldsymbol{\theta}}$ é a matriz, $n \times (p+5)$,

dada por

$$\boldsymbol{D_\theta} = \begin{pmatrix} 1 & X_{11} & \dots & X_{1p} & 0 & 0 & 0 & 0 \\ 1 & X_{21} & \dots & X_{2p} & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & X_{n1} & \dots & X_{np} & 0 & 0 & 0 & 0 \end{pmatrix} = (\boldsymbol{X}, \boldsymbol{0})$$

e $\boldsymbol{L_{\theta Y}} = (\boldsymbol{L_{\beta Y}}, \boldsymbol{L_{\varphi Y}}, \boldsymbol{L_{\alpha Y}})^\top$, em que $\boldsymbol{L_{\beta Y}}$ é a matriz, $(p+1) \times n$, cujos elementos são dados por

$$\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \beta_j \partial y_i} = \frac{1}{\alpha^2}\left(\left(\boldsymbol{s_i}^\top \odot \boldsymbol{X_{.j}}^\top\right)\boldsymbol{\Sigma}^{-1}\boldsymbol{V} + \left(\boldsymbol{U}^\top \odot \boldsymbol{X_{.j}}\right)\boldsymbol{\Sigma}^{-1}\boldsymbol{C_i}\right) - \frac{1}{4}\operatorname{sech}^2\left(\frac{y_i - \boldsymbol{X_i}^\top\boldsymbol{\beta}}{2}\right)X_{ij},$$

em $\boldsymbol{S_i}$ e $\boldsymbol{C_i}$ são vetores, $n \times 1$, com $\operatorname{senh}\left(\frac{y_i - \boldsymbol{X_i}^\top\boldsymbol{\beta}}{2}\right)$ e $\cosh\left(\frac{y_i - \boldsymbol{X_i}^\top\boldsymbol{\beta}}{2}\right)$, respectivamente, na posição i e 0 nas demais posições, $i = 1, \dots, n; \quad j = 0, \dots, p$.
$\boldsymbol{L_{\varphi Y}}$ é a matriz, $3 \times n$, cujos elementos são dados por

$$\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \varphi_j \partial y_i} = \frac{2}{\alpha^2}\boldsymbol{C_i}^\top\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_j}\boldsymbol{\Sigma}^{-1}\boldsymbol{V}, \quad i = 1, \dots, n; \quad j = 1, 2, 3;$$

$\boldsymbol{L_{\alpha Y}}$ é o vetor, $1 \times n$, cujos elementos são dados por

$$\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \alpha \partial y_i} = \frac{4}{\alpha^3}\boldsymbol{C_i}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{V}, \quad i = 1, \dots, n.$$

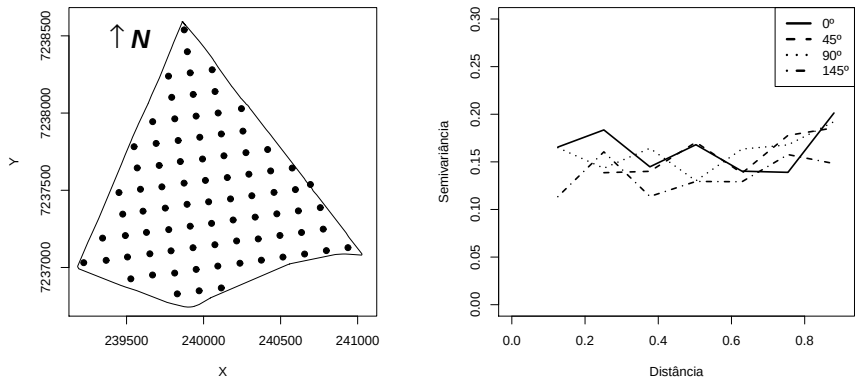
5.3.4 Aplicação

5.3.4.1 Descrição, análise exploratória e análise espacial dos dados

Nesta seção, a metodologia desenvolvida no trabalho é ilustrada usando um conjunto de dados com 82 pontos, referentes ao ano agrícola 2014/2015 de uma área comercial, de aproximadamente 167 ha, localizada no município de Cascavel, no Oeste do Estado do Paraná, Brasil. A variável-resposta é a concentração de Magnésio no solo $(Mg)(cmolc\ dm^{-3})$ e a variável auxiliar é a concentração de Cálcio no solo $(Ca)(cmolc\ dm^{-3})$. Os pontos (X,Y) foram georefenciados utilizando o sistema de coordenadas UTM e a grade amostral pode ser observada na Figura 27 (a). A Figura 27 (b) mostra que não existe uma direção privilegiada, portanto, o processo estocástico pode ser considerado isotrópico.

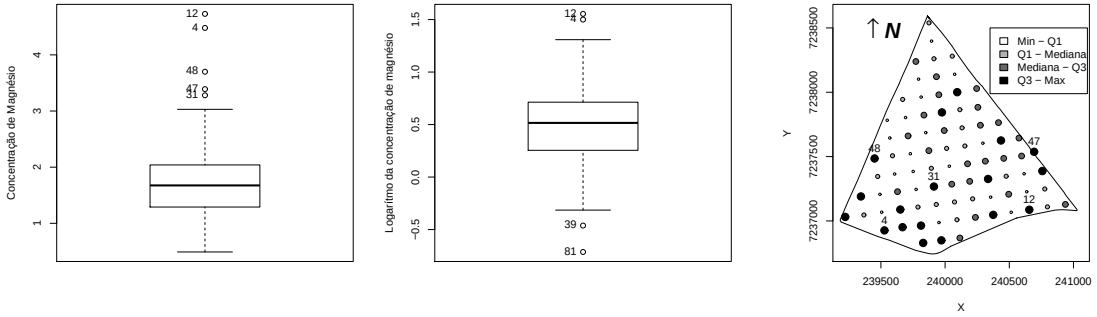
Os dados da concentração de Magnésio apresentam cinco *outliers* (Figura 28(a)). Após a linearização (obtida ao se aplicar o logaritmo nos dados) somente duas dessas observações continuam sendo apontadas como *outliers*, e outras duas observações destacam-se como *outliers* apenas para o logaritmo dos dados (Figura 28(b)). A Figura 28(c) exibe os pontos divididos por quartis bem como a localização dos *outliers*.

Para estimar a estrutura de dependência espacial, considerou-se o modelo da família Matérn, apresentado na equação (5.52) para $\delta = 0,5$ (modelo exponencial), escolhido por validação cruzada. Um estudo dos gráficos *Q-Q plots* com dados transformados mostrou que os dados se ajustam de modo razoável ao conjunto de dados. Para o modelo escolhido,



(a) (b)

Figura 27 Grade amostral (a) e semivariogramas direcionais (b).



(a) (b) (c)

Figura 28 *Box plots* da concentração de Magnésio (a) e do logaritmo da concentração de Magnésio (b), e localização dos *outliers* na grade amostral (c).

os parâmetros estimados por ML, com os respectivos erros padrões (apresentados entre parênteses) e p-valores (apresentados entre colchetes) são $\hat{\alpha} = 1,0168(0,9791)[0,1495]$, $\hat{\beta}_0 = -0,9595(0,1088)[> 0,999]$, $\hat{\beta}_1 = 0,2728(0,0146)[< 0,001]$, $\hat{\varphi}_1 = 0,0270(0,0518)[0,3011]$, $\hat{\varphi}_2 = 0,0115(0,0225)[0,3046]$ e $\hat{\varphi}_3 = 0,9436(0,0055)[< 0,001]$. Então, tem-se que

$$\hat{Mg}(\boldsymbol{s_i}) = \exp(-0,9595 + 0,2728Ca(\boldsymbol{s_i}))\eta(\boldsymbol{s_i}),$$

com $\eta(\boldsymbol{s_i}) \sim \mathrm{BS}(1,0168, 1)$, e matriz escala dada por $\boldsymbol{\Sigma} = 0,0270\boldsymbol{I}_{82} + 0,0115\boldsymbol{R}$, com $\boldsymbol{R}$ dado em (5.54), com $\delta = 0,5$.

5.3.4.2 Diagnósticos

A Figura 29 mostra os pontos que possivelmente influenciam a estimação dos parâmetros. Note que os pontos #12, #15 e #27 influenciam tanto a estimação do vetor de efeito fixo $\boldsymbol{\beta} = (\beta_0, \beta_1)^\top$, quanto a estimação do vetor $\boldsymbol{\psi} = (\boldsymbol{\varphi}^\top, \alpha)^\top$, enquanto a observação #19 é possivelmente influente apenas na estimação de $\boldsymbol{\beta}$, e as observações #62, #67 e #80 são influentes na estimação de $\boldsymbol{\psi}$. Dos possíveis pontos influentes o ponto #12 foi o único detectado

como um *outlier*.

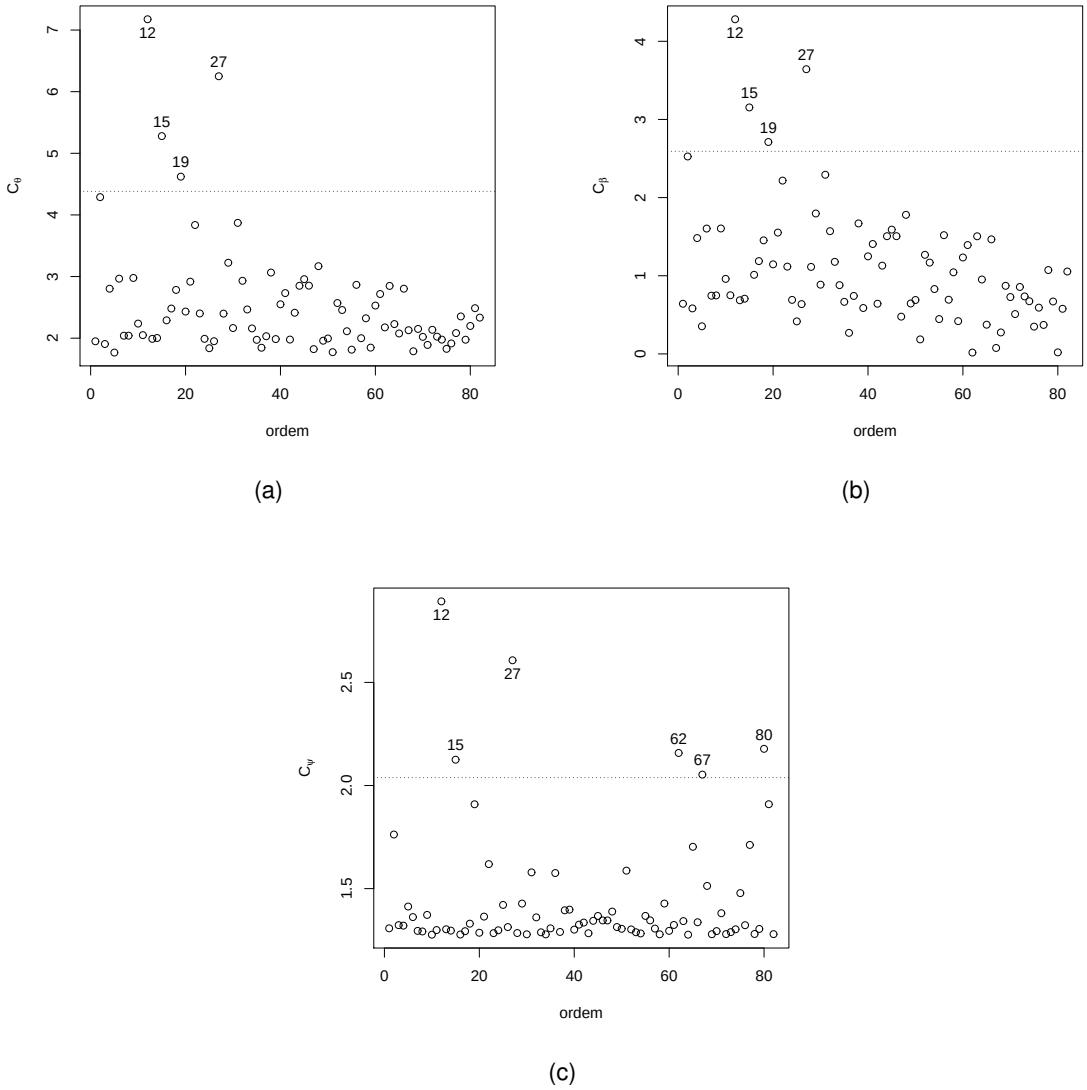
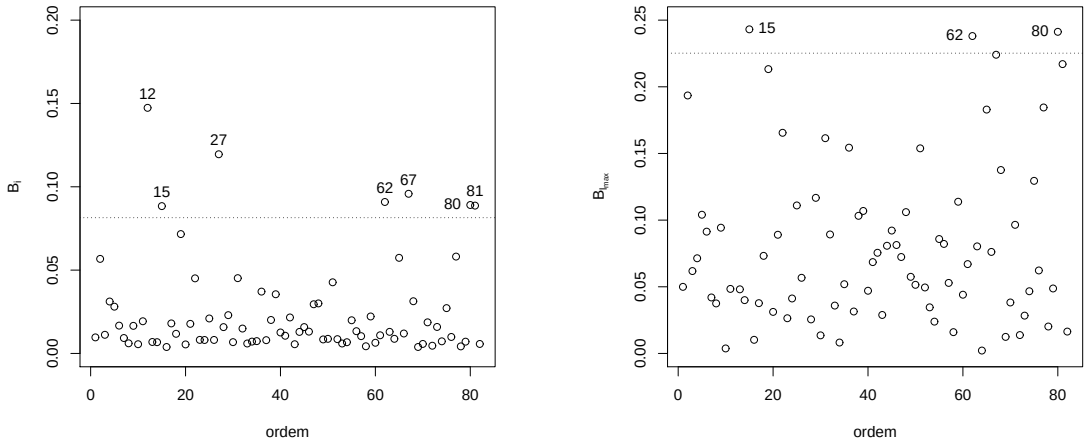


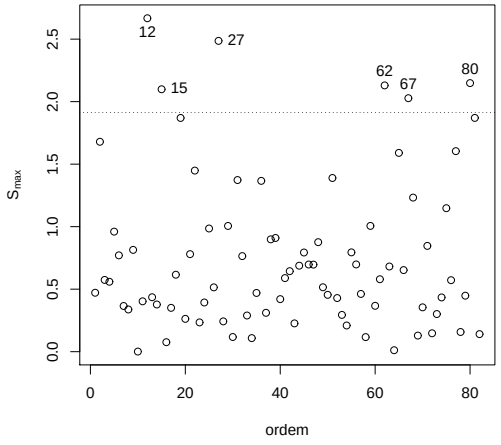
Figura 29 Distância de Cook para o vetor de parâmetros $\boldsymbol{\theta}$ (a), para o vetor $\boldsymbol{\beta}$ (b) e para o vetor $\boldsymbol{\psi}$ (c).

Os gráficos de influência local considerando-se a perturbação na variável-resposta, são apresentados na Figura 30. Enquanto a Figura 31 apresenta os gráficos de influência local quando perturba-se a covariável. Os gráficos $B_{d_{max}}$ e S_{max} apontaram as observações #12, #15, #27, #62, #67 e #80 como influentes na variável resposta. O gráfico $B_{l_{max}}$ identifica ainda a observação #81. Já o gráfico B_i identifica apenas as observações #15, #62 e #80. Quando a pertubação ocorre na covariável os gráficos $B_{d_{max}}$ e S_{max} identificam como possíveis observações influentes as mesmas observações #12, #15, #27, #62, #67 e #80 e B_i não aponta nenhuma observação como influente.



(a)

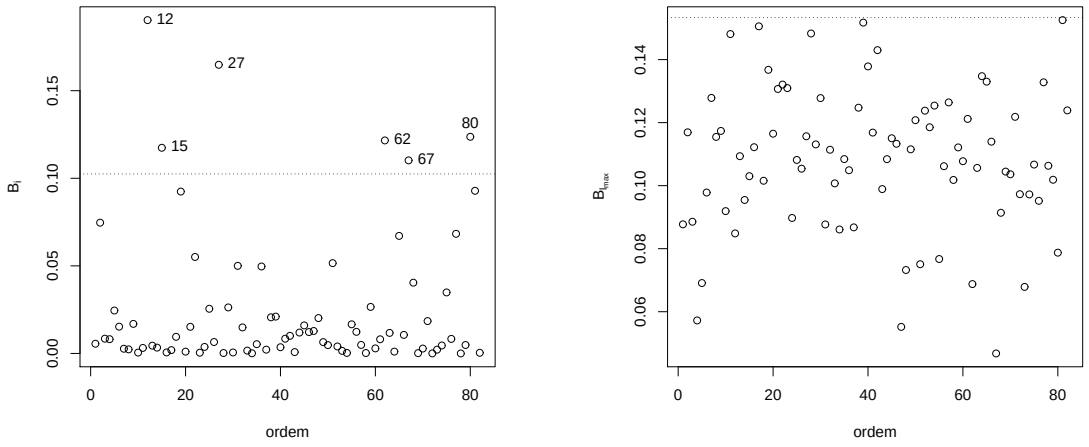
(b)



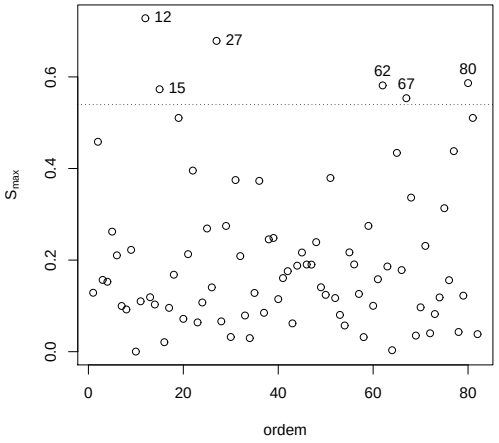
(c)

Figura 30 Gráficos de influência local, B_i *versus* ordem (a), $B_{d_{max}}$ *versus* ordem (b), e S_{max} *versus*ordem (c), pertubando a variável resposta.

Quanto às observações que influenciam seus próprios valores preditos (pontos de avalavanca), destacam-se as observações #4, #39, #47, #67 e #81 (Figura 32). Dessas observações, #4, #39 e #81 são *outliers* para a variável-resposta linearizada $(\log(Mg))$ e #47 foi apontada como *outlier* para a variável Mg (Figura 28). Destacam-se que a observação #12 identificada como *outlier* é apontada como influente em praticamente todos os gráficos de influência. Destacam-se ainda, que existem observações apontadas como influentes que não foram apontadas como *outliers*, assim como acontece em outros estudos, como no estudo realizado por Assumpção, Uribe-Opazo e Galea (2014).



(a) (b)



(c)

Figura 31 Gráficos de influência local, B_i *versus* ordem (a), $B_{d_{max}}$ *versus* ordem (b) e S_{max} *versus* ordem (c), pertubando a covariável.

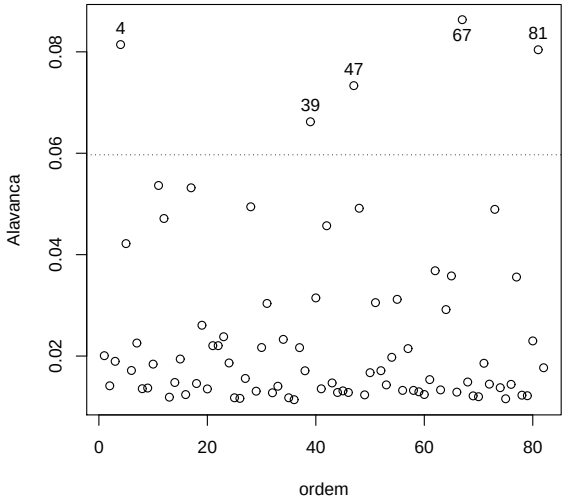


Figura 32 Alavanca generalizada para os dados de concentração de magnésio.

5.3.5 Conclusões

Do ponto de vista teórico, a principal contribuição para a teoria geoestatística foi o desenvolvimento de métodos de diagnósticos para modelos espaciais log-lineares Birnbaum-Saunders. Tais modelos possibilitam modelar a dependência espacial de dados experimentais, estritamente positivos, que apresentam evidências de distribuição assimétrica positiva, e é um modelo alternativo aos usuais modelos espaciais, baseados na distribuição Gaussiana. Do ponto de vista prático, a análise de diagnóstico, baseada em técnicas gráficas, mostra que as técnicas de influência global e local apontam basicamente os mesmos pontos como possíveis pontos influentes. Dentre as observações que exercem maior influência (#12 e #27), a observação #12 é um *outlier*, enquanto #27 não. Isso também é evidente em outros trabalhos, tais como os conduzidos por Assumpção, Uribe-Opazo e Galea (2014) e De Bastiani et al. (2015). onstatação reforça o fato de que, em estatística espacial, um ponto influente não é necessariamente um *outlier*, bem como um *outlier* não é necessariamente um ponto influente. Em estatística espacial, o conceito de influência depende, além do valor da variável, do valor da variável em localizações vizinhas. No caso específico da observação #27, pode-se observar na Figura 33 que esta observação apresenta localizações vizinhas com valores no primeiro quartil.

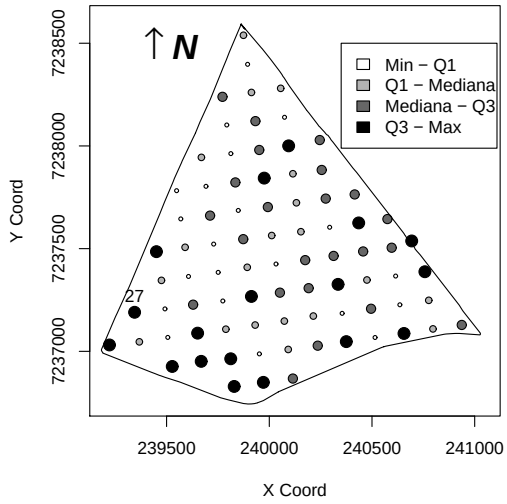


Figura 33 Grade amostral com destaque para a localização da observação #27.

5.3.6 Trabalhos futuros

Como trabalhm eliminação dos pontos influentes a fim de visualizar de que forma a eliminação de pontos influentes afeta tais mapas. Adicionalmente, pretende-se realizar estudo de influência em modelos espaciais BS generalizados, isto é, modelos BS baseados em distribuições com caudas mais pesadas.

5.3.7 Referências

AMARAL, G.J.A; FLORES, O.P.R.; CYSNEIROS, F.J.A. Graphical and Numerical Methods for Detecting Influential Observations in Complex Bingham Data. **Communications in Statistics -**

Simulation and Computation, v. 42, p. 1801-1814, 2013.

ASSUMPÇÃO, R.A.B.; URIBE-OPAZO, M.A.; GALEA, M. Analysis of local influence in geostatistics using Student-t distribution. **Journal of Applied Statistics,** v. 41, p. 2323-2341, 2014.

BILLOR, N.; LOYNES, R.M. Local influence: a new approach. **Communications in Statistics-Theory and Methods,** v. 22, p. 1595-1611, 1993.

BILLOR, N.; LOYNES, R.M. An application of the local influence approach to ridge regression. **Journal of Applied Statistics,** v. 26, p. 177-183, 1999.

CADIGAN, N.G.; FARRELL, P.J. Generalized local influence with applications to fish stock cohord analysis. **Journal of the Royal Statistical Society Series C,** v. 51, p. 469-483, 2002.

CERIOLI, A.; RIANI, M. The Ordering of Spatial Data and the Detection of Multiple Outliers. **Journal of Computational and Graphical Statistics,** v. 8, p. 239-258, 1999.

CHRISTENSEN, R.; JOHNSON, W.; PEARSON, L.M. Prediction diagnostics for spatial linear models. **Biometrika,** v. 79, p. 583-591, 1992.

CHRISTENSEN, R.; JOHNSON, W.; PEARSON, L.M. Covariance function diagnostics for spatial linear models. **Mathematical Geology,** v. 25, p. 145-160, 1993.

COOK, R.D. Detection of influential observations in linear regressions. **Technometrics,** v. 19, p. 15-18, 1977.

COOK, R.D.; WEISBERG, S. **Residuals and Influence in Regression,** Londres: Chapman and Hall, 1982.

COOK, R.D. Assessment of local influence. **Journal of the Royal Statistics Society B,** v. 48, p. 133-169, 1986.

DE BASTIANI, F.; CYSNEIROS; A.H.M.A.; URIBE-OPAZO, M.A.; GALEA, M. Influence diagnostics in elliptical spatial linear models. **TEST,** v. 24, p. 322-340, 2015.

GALEA, M.; RIQUELME, M.; PAULA, G.A. Diagnostic methods in elliptical linear regression models. **Brazilian Journal of Probability and Statistics,** v. 14, p. 167-184, 2000.

GALEA, M.; PAULA, G.A.; URIBE-OPAZO, M. On influence diagnostic in univariate elliptical linear regression models. **Statistical Papers,** v. 44, p. 23-45, 2003.

GARCIA-PAPANI, F.M.; URIBE-OPAZO, M.A.; LEIVA, V.; AYKROYD, R. G. Birnbaum-Saunders spatial modelling and diagnostics applied to agricultural engineering data. **Stochastic Environmental Research and Risk Assessment,** (in press), 2016. Disponível em: <://dx.doi.org/10.1007/s00477-015-1204-4>. Acesso em: 02 fev. 2016.

GIMENEZ, P.; GALEA, M. Influence measures on corrected score estimators in functional heteroscedastic measurement error models. **Journal of Multivariate Analysis,** v. 114, p. 1-15, 2013.

LEIVA, V.; BARROS, M.; PAULA, G.A.; GALEA, M. Influence diagnostics in log-Birnbaum-Saunders regression models with censored data. **Computational Statistics and Data Analysis,** v. 51, p. 5694-5707, 2007.

LEIVA, V.; ROJAS, E.; GALEA, M.; SANHUEZA, A. Diagnostics in Birnbaum-Saunders accelerated life models with an application to fatigue data. **Applied Stochastic Models in Business and Industry,** v. 30, p. 115-131, 2014.

LEIVA, V.; LIU S.; SHI, L.; CYSNEIROS, F.J.A. Diagnostics in elliptical regression models with stochastic restrictions applied to econometrics. **Journal of Applied Statistics** (in press), 2015. Disponível em: <http://dx.doi.org/10.1080/02664763.2015>. Acesso em: 20 nov. 2015.

PAN, J.; FEI, Y.; FOSTER, P. Case-Delection Diagnostics for Linear Models. **Technometrics,** v. 56, p. 269-281, 2014.

PAULA, G.A. Assessing local influence in restricted regression moldels. **Computational Statistics and Data Analysis,** v. 16, p. 63-79, 1993.

POON, W.; POON, Y.S. Conformal normal curvature and assessment of local influence. **Royal Statistical Society,** v. 61, p. 51-61, 1999.

ROSS, W.H. The geometry of case deletion and the assessment of influence in nonlinear regression. **Canadian Journal of Statistics,** v. 15, p. 91-103, 1987.

TSAI, C.L. Discussion of assessment of local influence by R. D. Cook. **Journal of the Royal Statistics Society Series B,** v. 48, p. 165, 1986.

URIBE-OPAZO, M.A.; BORSSOI, J.A.; GALEA, M. Influence diagnostics in Gaussian spatial linear models. **Journal of Applied Statistics,** v. 39, p. 615-630, 2012.

WEI, B.; HU, Y.Q.; FUNG, W.K. Generalized leverage and its applications. **Scandinavian Journal of Statistics,** v. 25, p. 25-37, 1998.

ZHU, H.; LEE, S. Local Influence for incomplete-data models. **Royal Statistical Society,** v. 63, p. 111-126, 2001.

ZHU, H.; IBRAHIM, J.G.; LEE, S.; ZHANG, H. Perturbation selection and influence measures in local influence analysis. **The Annals of Statistics,** v. 35, p. 2565-2588, 2007.

5.3.8 Apêndices

5.3.8.1 Apêndice I: Matrizes de informação observada e esperada para o modelo espacial log-linear BS

A matriz de informação observada é definida por $-\ddot{\boldsymbol{\ell}}(\boldsymbol{\theta})$, avaliada em $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$, em que $\ddot{\boldsymbol{\ell}}(\boldsymbol{\theta})$ é a matriz Hessiana. Para o modelo espacial log-linear BS, apresentado em (5.49), a matriz Hessiana, $(p+5) \times (p+5)$, é dada por

$$\ddot{\boldsymbol{\ell}}(\boldsymbol{\theta}) = \begin{pmatrix} \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} & \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\varphi}^\top} & \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \alpha} \\ \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\varphi} \partial \boldsymbol{\beta}^\top} & \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\varphi} \partial \boldsymbol{\varphi}^\top} & \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\varphi} \partial \alpha} \\ \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \alpha \partial \boldsymbol{\beta}^\top} & \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \alpha \partial \boldsymbol{\varphi}^\top} & \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \alpha \partial \alpha} \end{pmatrix} = \begin{pmatrix} \ddot{\boldsymbol{\ell}}_{\boldsymbol{\beta}\boldsymbol{\beta}} & \ddot{\boldsymbol{\ell}}_{\boldsymbol{\beta}\boldsymbol{\varphi}} & \ddot{\boldsymbol{\ell}}_{\boldsymbol{\beta}\alpha} \\ \ddot{\boldsymbol{\ell}}_{\boldsymbol{\varphi}\boldsymbol{\beta}} & \ddot{\boldsymbol{\ell}}_{\boldsymbol{\varphi}\boldsymbol{\varphi}} & \ddot{\boldsymbol{\ell}}_{\boldsymbol{\varphi}\alpha} \\ \ddot{\boldsymbol{\ell}}_{\alpha\boldsymbol{\beta}} & \ddot{\boldsymbol{\ell}}_{\alpha\boldsymbol{\varphi}} & \ddot{\boldsymbol{\ell}}_{\alpha\alpha} \end{pmatrix},$$

em que, $\ddot{\boldsymbol{\ell}}_{\boldsymbol{\beta}\boldsymbol{\beta}}$ é uma matriz, $(p+1) \times (p+1)$, com elementos dados por

$$\begin{aligned} \ddot{\ell}_{\beta_j \beta_l} &= \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \beta_j \partial \beta_l} = -\frac{1}{\alpha^2} \left(\boldsymbol{V}^\top \odot \boldsymbol{X}_{\cdot j}^\top \odot \boldsymbol{X}_{\cdot l}^\top \right) \boldsymbol{\Sigma}^{-1} \boldsymbol{V} - \frac{1}{\alpha^2} \left(\boldsymbol{W}^\top \odot \boldsymbol{X}_{\cdot j}^\top \right) \boldsymbol{\Sigma}^{-1} \left(\boldsymbol{W} \odot \boldsymbol{X}_{\cdot l} \right) \\ &+ \frac{1}{4} \sum_{i=1}^{n} \operatorname{sech}^2 \left(\frac{y_i - \boldsymbol{X}_i^\top \boldsymbol{\beta}}{2} \right) \boldsymbol{X}_{ij} \boldsymbol{X}_{il}; \quad j, l = 0 \ldots p; \end{aligned}$$

$\ddot{\boldsymbol{\ell}}_{\boldsymbol{\beta}\boldsymbol{\varphi}} = \ddot{\boldsymbol{\ell}}_{\boldsymbol{\varphi}\boldsymbol{\beta}}^\top$ é uma matriz, $(p+1) \times 3$ com elementos dados por

$$\ddot{\ell}_{\beta_j \varphi_l} = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \beta_j \partial \varphi_l} = -\frac{2}{\alpha^2} \left(\boldsymbol{W}^\top \odot \boldsymbol{X}_{\cdot j}^\top \right) \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_l} \boldsymbol{\Sigma}^{-1} \boldsymbol{V}; \quad j = 1, \ldots, p; \quad l = 1, 2, 3;$$

$\ddot{\boldsymbol{\ell}}_{\boldsymbol{\beta}\alpha} = \ddot{\boldsymbol{\ell}}_{\alpha\boldsymbol{\beta}}^\top$ é um vetor, $(p+1) \times 1$, com elementos dados por

$$\ddot{\ell}_{\beta_j \alpha} = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \beta_j \partial \alpha} = -\frac{4}{\alpha^3} \left(\boldsymbol{W}^\top \odot \boldsymbol{X}_{\cdot j}^\top \right) \boldsymbol{\Sigma}^{-1} \boldsymbol{V}; \quad j = 0, \ldots, p;$$

$\ddot{\boldsymbol{\ell}}_{\boldsymbol{\varphi}\boldsymbol{\varphi}}$ é uma matriz, 3×3, com elementos dados por

$$\begin{aligned} \ddot{\ell}_{\varphi_j \varphi_l} &= \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \varphi_j \partial \varphi_l} = -\frac{1}{2} \mathsf{tr} \left(-\boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_l} \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_j} + \boldsymbol{\Sigma}^{-1} \frac{\partial^2 \boldsymbol{\Sigma}}{\partial \varphi_l \partial \varphi_j} \right) \\ &+ \frac{2}{\alpha^2} \boldsymbol{V}^\top \left(-\boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_l} \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\varphi_j} \boldsymbol{\Sigma}^{-1} + \boldsymbol{\Sigma}^{-1} \left(\frac{\partial^2 \boldsymbol{\Sigma}}{\partial \varphi_l \partial \varphi_j} \boldsymbol{\Sigma}^{-1} - \frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_j} \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_l} \boldsymbol{\Sigma}^{-1} \right) \right) \boldsymbol{V}; j, l = 1, 2, 3. \end{aligned}$$

$\ddot{\boldsymbol{\ell}}_{\boldsymbol{\varphi}\alpha} = \ddot{\boldsymbol{\ell}}_{\alpha\boldsymbol{\varphi}}^\top$ é um vetor, 3×1, com elementos dados por

$$\ddot{\ell}_{\varphi_j \alpha} = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \varphi_j \partial \alpha} = -\frac{4}{\alpha^3} \boldsymbol{V}^\top \left(\boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_j} \boldsymbol{\Sigma}^{-1} \right) \boldsymbol{V}; \quad j = 1, 2, 3.$$

Finalmente, $\ddot{\ell}_{\alpha\alpha} = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \alpha \partial \alpha} = \frac{n}{\alpha^2} - \frac{12}{\alpha^4} \boldsymbol{V}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{V}$.

A matriz de informação esperada é dada por $\boldsymbol{I}(\boldsymbol{\theta}) = \mathrm{E}[-\ddot{\boldsymbol{\ell}}(\boldsymbol{\theta})]$. Assim, para o modelo espacial log-linear, essa é uma matriz $(p+5) \times (p+5)$, dada por

$$\boldsymbol{I}(\boldsymbol{\theta}) = \begin{pmatrix} \boldsymbol{I}_{\boldsymbol{\beta}\boldsymbol{\beta}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{I}_{\boldsymbol{\varphi}\boldsymbol{\varphi}} & \boldsymbol{I}_{\boldsymbol{\varphi}\alpha} \\ \mathbf{0} & \boldsymbol{I}_{\alpha\boldsymbol{\varphi}} & I_{\alpha\alpha} \end{pmatrix} = \begin{pmatrix} \boldsymbol{I}_{\boldsymbol{\beta}\boldsymbol{\beta}} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{I}_{\boldsymbol{\psi}\boldsymbol{\psi}} \end{pmatrix}$$

em que, $\boldsymbol{I}_{\boldsymbol{\beta\beta}} = \mathrm{E}[-\ddot{\ell}_{\boldsymbol{\beta\beta}}]$ é uma matriz, $(p+1)\times(p+1)$, com elementos dados por

$$\begin{aligned} I_{\beta_j\beta_l} &= \mathrm{E}\left(\frac{1}{\alpha^2}\left(\boldsymbol{V}^\top \odot (\boldsymbol{X}_{\cdot j}\odot \boldsymbol{X}_{\cdot l})^\top\right)\boldsymbol{\Sigma}^{-1}\boldsymbol{V}\right) + \mathrm{E}\left(\frac{1}{\alpha^2}\left(\boldsymbol{W}^\top \odot \boldsymbol{X}_{\cdot j}^\top\right)\boldsymbol{\Sigma}^{-1}(\boldsymbol{W}\odot \boldsymbol{X}_{\cdot l})\right) \\ &+ \mathrm{E}\left(-\frac{1}{4}\sum_{i=1}^{n}\mathrm{sech}^2\left(\frac{y_i-\boldsymbol{X}_i^\top\boldsymbol{\beta}}{2}\right)\boldsymbol{X}_{ij}\boldsymbol{X}_{il}\right) \\ &= \frac{1}{4}\mathrm{E}\left(\left(\left(\frac{2}{\alpha}\boldsymbol{V}\right)^\top \odot (\boldsymbol{X}_{\cdot j}\odot \boldsymbol{X}_{\cdot l})^\top\right)\boldsymbol{\Sigma}^{-1}\left(\frac{2}{\alpha}\boldsymbol{V}\right)\right) + \frac{1}{\alpha^2}\boldsymbol{X}_{\cdot j}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{X}_{\cdot l} - \frac{1}{4}\sum_{i=1}^{n}\boldsymbol{X}_{ij}\boldsymbol{X}_{il} \\ &= \frac{1}{4}\mathbf{1}^\top(\boldsymbol{X}_{\cdot j}\odot \boldsymbol{X}_{\cdot l}) + \frac{1}{\alpha^2}\boldsymbol{X}_{\cdot j}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{X}_{\cdot l} - \frac{1}{4}\sum_{i=1}^{n}\boldsymbol{X}_{ij}\boldsymbol{X}_{il} = \frac{1}{\alpha^2}\boldsymbol{X}_{\cdot j}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{X}_{\cdot l}, \end{aligned}$$

uma vez que $\boldsymbol{W} \cong \mathbf{1},\ \mathrm{sech}^2\left(\frac{y_i-\boldsymbol{X}_i^\top\boldsymbol{\beta}}{2}\right) \cong 1$ e por propriedades de traço e de produto de Hadamard;

$\boldsymbol{I}_{\boldsymbol{\beta\varphi}} = \mathrm{E}[-\ddot{\ell}_{\boldsymbol{\beta\varphi}}]$ é uma matriz de ordem $(p+1)\times 3$, com elementos

$$\boldsymbol{I}_{\beta_j\varphi_l} = \mathrm{E}\left(\frac{2}{\alpha^2}\left(\boldsymbol{W}^\top \odot \boldsymbol{X}_{\cdot j}^\top\right)\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_l}\boldsymbol{\Sigma}^{-1}\boldsymbol{V}\right) = \frac{1}{\alpha}\boldsymbol{X}_{\cdot j}^\top\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_l}\boldsymbol{\Sigma}^{-1}\mathsf{E}\left(\frac{2}{\alpha}\boldsymbol{V}\right) = 0.$$

Adicionalmente, $\boldsymbol{I}_{\boldsymbol{\beta}\alpha} = \mathrm{E}[-\ddot{\ell}_{\boldsymbol{\beta}\alpha}]$ é um vetor, $(p+1)\times 1$, em que os elementos são dados por

$$\boldsymbol{I}_{\beta_j\alpha} = \mathrm{E}\left(\frac{4}{\alpha^3}\left(\boldsymbol{W}^\top \odot \boldsymbol{X}_{\cdot j}^\top\right)\boldsymbol{\Sigma}^{-1}\boldsymbol{V}\right) \cong \mathsf{E}\left(\frac{4}{\alpha^3}\boldsymbol{X}_{\cdot j}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{V}\right) = 0,$$

uma vez que $\left(\frac{2}{\alpha}\boldsymbol{V}\right) \sim \mathsf{N}_n(\mathbf{0},\boldsymbol{\Sigma})$.

$\boldsymbol{I}_{\boldsymbol{\varphi\varphi}} = \mathrm{E}[-\ddot{\ell}_{\boldsymbol{\varphi\varphi}}]$ é uma matriz simétrica, 3×3, que tem como elementos

$$\begin{aligned} I_{\varphi_j\varphi_l} &= \mathsf{E}\left(\frac{1}{2}\mathrm{tr}\left(-\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_l}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_j} + \boldsymbol{\Sigma}^{-1}\frac{\partial^2\boldsymbol{\Sigma}}{\partial\varphi_l\partial\varphi_j}\right)\right) + \mathrm{E}\left(\frac{2}{\alpha^2}\boldsymbol{V}^\top\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_l}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\varphi_j}\boldsymbol{\Sigma}^{-1}\right)\boldsymbol{V}\right) \\ &- \mathrm{E}\left(\frac{2}{\alpha^2}\boldsymbol{V}^\top\left(\boldsymbol{\Sigma}^{-1}\frac{\partial^2\boldsymbol{\Sigma}}{\partial\varphi_l\partial\varphi_j}\boldsymbol{\Sigma}^{-1}\right)\boldsymbol{V}\right) + \mathrm{E}\left(\frac{2}{\alpha^2}\boldsymbol{V}^\top\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_j}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_l}\boldsymbol{\Sigma}^{-1}\right)\boldsymbol{V}\right) \\ &= \frac{1}{2}\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_j}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_l}\right);\ j,l=1,2,3. \end{aligned}$$

$\boldsymbol{I}_{\boldsymbol{\varphi}\alpha} = \mathrm{E}[-\ddot{\ell}_{\boldsymbol{\varphi}\alpha}]$ é um vetor, 3×1, com elementos dados por,

$$I_{\varphi_j\alpha} = \mathrm{E}\left(\frac{4}{\alpha^3}\boldsymbol{V}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_j}\boldsymbol{\Sigma}^{-1}\right)\boldsymbol{V}\right) = \frac{1}{\alpha}\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_j}\right), j=1,2,3.$$

Finalmente, $I_{\alpha\alpha} = \mathrm{E}[-\ddot{\ell}_{\alpha\alpha}] = \mathrm{E}\left(-\frac{n}{\alpha^2} + \frac{12}{\alpha^4}\boldsymbol{V}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{V}\right) = \frac{2n}{\alpha^2}$, uma vez que $\left(\frac{2}{\alpha}\boldsymbol{V}\right)^\top\boldsymbol{\Sigma}^{-1}\left(\frac{2}{\alpha}\boldsymbol{V}\right) \sim \chi^2_n$, em que χ^2_n representa a distribuição qui-quadrado com n graus de liberdade.

$\boldsymbol{I}_{\boldsymbol{\varphi\beta}}$; $\boldsymbol{I}_{\alpha\boldsymbol{\beta}}$ e $\boldsymbol{I}_{\alpha\boldsymbol{\varphi}}$ são, respectivamente, dadas por $(\boldsymbol{I}_{\boldsymbol{\beta\varphi}})^\top$; $(\boldsymbol{I}_{\boldsymbol{\beta}\alpha})^\top$ e $(\boldsymbol{I}_{\boldsymbol{\varphi}\alpha})^\top$.

5.3.8.2 Apêndice II: Vetor escore $U(\omega)$

Para a pertubação na variável-resposta, $U(\omega) = \frac{\partial\ell(\theta|\omega)}{\partial\omega}$ pode ser obtido com cálculos similares ao realizados em Garcia-Papani et al. (2016). Para o esquema de pertubação em uma das covariáveis tem-se

$$\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = -\frac{1}{2}\log(|\boldsymbol{\Sigma}|) - n\log(\alpha) - \frac{2}{\alpha^2}\boldsymbol{V}_{\boldsymbol{\omega}}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{V}_{\boldsymbol{\omega}} + \sum_{i=1}^{n}\log\left(\cosh\left(\frac{y_i-\boldsymbol{X}_i^\top\boldsymbol{\beta}-\beta_1\boldsymbol{A}_i\boldsymbol{\omega}}{2}\right)\right).$$

em que $\boldsymbol{V_\omega} = (V_{\omega_1}, \ldots, V_{\omega_n})^\top$ é um vetor $n \times 1$, com $V_{\omega_i} = \operatorname{senh}\left(\frac{y_i - \boldsymbol{X}_i^\top\boldsymbol{\beta} - \beta_1 \boldsymbol{A}_i\boldsymbol{\omega}}{2}\right), i = 1, \ldots n$, e $\boldsymbol{A}_i$ a $i-$ésima linha da matriz $\boldsymbol{A}$.

Desta forma, tem-se

$$\frac{\partial \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \omega} = \frac{\partial}{\partial \omega}\left(-\frac{2}{\alpha^2}\boldsymbol{V}_{\boldsymbol{\omega}}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{V}_{\boldsymbol{\omega}}\right) + \frac{\partial}{\partial \omega}\left(\sum_{i=1}^{n} \log\left(\cosh\left(\frac{y_i - \boldsymbol{X}_i^\top\boldsymbol{\beta} - \beta_1\boldsymbol{A}_i\boldsymbol{\omega}}{2}\right)\right)\right).$$

Além disso, tem-se que

$$\begin{aligned}\frac{\partial}{\partial \omega}\left(-\frac{2}{\alpha^2}\boldsymbol{V}_{\boldsymbol{\omega}}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{V}_{\boldsymbol{\omega}}\right) &= -\frac{2}{\alpha^2}\left(\frac{\partial \boldsymbol{V}_{\boldsymbol{\omega}}}{\partial \boldsymbol{\omega}^\top}\right)^\top \left(2\boldsymbol{\Sigma}^{-1}\boldsymbol{V}_{\boldsymbol{\omega}}\right) \\ &\approx -\frac{2}{\alpha^2}\left(-\frac{\beta_1}{2}\boldsymbol{A}^\top\right)\left(2\boldsymbol{\Sigma}^{-1}\boldsymbol{V}_{\boldsymbol{\omega}}\right) = \frac{2\beta_1}{\alpha^2}\boldsymbol{A}\boldsymbol{\Sigma}^{-1}\boldsymbol{V}_{\boldsymbol{\omega}} \text{ e,}\end{aligned}$$

$$\frac{\partial}{\partial \omega}\left(\sum_{i=1}^{n} \log\left(\cosh\left(\frac{y_i - \boldsymbol{X}_{i\boldsymbol{\omega}}^\top\boldsymbol{\beta}}{2}\right)\right)\right) = -\frac{\beta_1}{2}\boldsymbol{A}\boldsymbol{T}_{\boldsymbol{\omega}} \approx -\frac{\beta_1}{2}\boldsymbol{A}\boldsymbol{V}_{\boldsymbol{\omega}}.$$

em que $\boldsymbol{T_\omega} = (T_{\omega_1}, \ldots, T_{\omega_n})^\top$ é um vetor, $n \times 1$, com $T_{\omega_i} = \tanh\left(\frac{y_i - \boldsymbol{X}_i^\top\boldsymbol{\beta} - \beta_1\boldsymbol{A}_i\boldsymbol{\omega}}{2}\right), i = 1, \ldots n.$

Portanto,

$$\frac{\partial \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \omega} = \frac{2\beta_1}{\alpha^2}\boldsymbol{A}\boldsymbol{\Sigma}^{-1}\boldsymbol{V}_{\boldsymbol{\omega}} - \frac{\beta_1}{2}\boldsymbol{A}\boldsymbol{V}_{\boldsymbol{\omega}} = \beta_1\boldsymbol{A}\left(\frac{2}{\alpha^2}\boldsymbol{\Sigma}^{-1} - \frac{1}{2}\boldsymbol{I_n}\right)\boldsymbol{V}_{\boldsymbol{\omega}}.$$

5.3.8.3 Apêndice III: Matriz de pertubação para o modelo espacial log-linear BS

Pertubação na variável-resposta

Para o modelo espacial log-linear BS abordado neste capítulo tem-se que

$$\frac{\partial \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\top} = \boldsymbol{V}_{\boldsymbol{\omega}}^\top\left(-\frac{2}{\alpha^2}\boldsymbol{\Sigma}^{-1} + \frac{1}{2}\boldsymbol{I_n}\right)\boldsymbol{A},$$

em que, $\boldsymbol{I_n}$ é a matriz identidade de ordem n, $\boldsymbol{V_\omega} = (\boldsymbol{V_{\omega_1}}, \ldots, \boldsymbol{V_{\omega_n}})$, com elementos dados por $\boldsymbol{V_{\omega_i}} = \operatorname{senh}\left(\frac{Y_i + \boldsymbol{A_i\omega} - \boldsymbol{X_i}^\top\boldsymbol{\beta}}{2}\right), \quad i = 1, \ldots, n$, e $\boldsymbol{A_i}$ é a i-ésima linha da matriz $\boldsymbol{A}$.

Assim, a matriz de perturbação, $(p+5) \times n$, é dada por $\boldsymbol{\Delta} = \frac{\partial \ell^2(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \boldsymbol{\theta}\partial \boldsymbol{\omega}^\top} = (\boldsymbol{\Delta_\beta}, \boldsymbol{\Delta_\varphi}, \Delta_\alpha)^\top$, com

$$\boldsymbol{\Delta_\beta} = \frac{\partial \ell^2(\boldsymbol{\theta}|\boldsymbol{\omega}}{\partial \boldsymbol{\beta}\partial \boldsymbol{\omega}^\top} = \left(\frac{\partial \ell^2(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \beta_0 \partial \boldsymbol{\omega}^\top}, \ldots, \frac{\partial \ell^2(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \beta_p \partial \boldsymbol{\omega}^\top}\right)^\top,$$

em que,

$$\frac{\partial \ell^2(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \boldsymbol{\beta_j}\partial \boldsymbol{\omega}^\top} = \frac{\partial \boldsymbol{V}_{\boldsymbol{\omega}}^\top}{\partial \beta_j}\left(-\frac{2}{\alpha^2}\boldsymbol{\Sigma}^{-1} + \frac{1}{2}\boldsymbol{I_n}\right)\boldsymbol{A} = -\frac{1}{2}\boldsymbol{W}_{\boldsymbol{\omega}}^\top \odot \boldsymbol{X}_{\boldsymbol{\cdot j}}^\top\left(-\frac{2}{\alpha^2}\boldsymbol{\Sigma}^{-1} + \frac{1}{2}\boldsymbol{I_n}\right)\boldsymbol{A}, \quad j = 0, \ldots, p$$

e $\boldsymbol{W_\omega} = (W_{\omega_1}, \ldots, W_{\omega_n})^\top$, com $W_{\omega_i} = \cosh\left(\frac{y_i + \boldsymbol{A_i\omega} - \boldsymbol{X_i}^\top\boldsymbol{\beta}}{2}\right)$. Uma vez que $W_{\omega_i} = \cosh\left(\frac{y_i + \boldsymbol{A_i\omega} - \boldsymbol{X_i}^\top\boldsymbol{\beta}}{2}\right) \cong 1$ tem-se que $\frac{\partial \ell^2(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \boldsymbol{\beta_j}\partial \boldsymbol{\omega}^\top} = \boldsymbol{X}_{\boldsymbol{\cdot j}}^\top\left(\frac{1}{\alpha^2}\boldsymbol{\Sigma}^{-1} - \frac{1}{4}\boldsymbol{I_n}\right)\boldsymbol{A}$.

$$\boldsymbol{\Delta_\varphi} = \frac{\partial \ell^2(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \boldsymbol{\varphi}\partial \boldsymbol{\omega}^\top} = \left(\frac{\partial \ell^2(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \varphi_1 \partial \boldsymbol{\omega}^\top}, \frac{\partial \ell^2(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \varphi_2 \partial \boldsymbol{\omega}^\top}, \frac{\partial \ell^2(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \varphi_3 \partial \boldsymbol{\omega}^\top}\right)^\top,$$

em que,

$$\begin{aligned}\frac{\partial \ell^2(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial\varphi_i\partial\boldsymbol{\omega}^\top} &= \frac{\partial \boldsymbol{V}_{\boldsymbol{\omega}}^\top}{\partial\varphi_i}\left(-\frac{2}{\alpha^2}\boldsymbol{\Sigma}^{-1}+\frac{1}{2}\boldsymbol{I}_n\right)\boldsymbol{A}+\boldsymbol{V}_{\boldsymbol{\omega}}^\top\left(\frac{2}{\alpha^2}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_i}\boldsymbol{\Sigma}^{-1}\boldsymbol{A}+\left(-\frac{2}{\alpha^2}\boldsymbol{\Sigma}^{-1}+\frac{1}{2}\boldsymbol{I}_n\right)\frac{\partial\boldsymbol{A}}{\partial\varphi_i}\right)\\ &= \boldsymbol{M}^\top\left(-\frac{1}{\alpha^2}\boldsymbol{\Sigma}^{-1}+\frac{1}{4}\boldsymbol{I_n}\right)\boldsymbol{A}+\boldsymbol{V}_{\boldsymbol{\omega}}^\top\left(\frac{2}{\alpha^2}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_i}\boldsymbol{\Sigma}^{-1}\right)\boldsymbol{A}\\ &+ \boldsymbol{V}_{\boldsymbol{\omega}}^\top\left(-\frac{2}{\alpha^2}\boldsymbol{\Sigma}^{-1}+\frac{1}{2}\boldsymbol{I_n}\right)\boldsymbol{A}\left(\frac{1}{\alpha}\boldsymbol{\Sigma}^{-1/2}\frac{\partial\boldsymbol{\Sigma}^{1/2}}{\partial\varphi_i}\boldsymbol{\Sigma}^{-1/2}+\frac{\alpha}{4}\frac{\partial\boldsymbol{\Sigma}^{1/2}}{\partial\varphi_i}\right)\boldsymbol{A},\ i=1,2,3,\end{aligned}$$

em que $\boldsymbol{M}=(M_1,\ldots,M_n)^\top$, com $M_j=\boldsymbol{l_j\omega}$ e $\boldsymbol{l_j}$ a j-ésima linha da matriz

$$\boldsymbol{A}\left(-\frac{1}{\alpha}\boldsymbol{\Sigma}^{-1/2}\frac{\partial\boldsymbol{\Sigma}^{1/2}}{\partial\varphi_i}\boldsymbol{\Sigma}^{-1/2}+\frac{\alpha}{4}\frac{\partial\boldsymbol{\Sigma}^{1/2}}{\partial\varphi_i}\right)\boldsymbol{A},\ j=1,\ldots,n.$$

$$\begin{aligned}\Delta_\alpha &= \frac{\partial\ell^2(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial\alpha\partial\boldsymbol{\omega}^\top}=\frac{\partial\boldsymbol{V}_{\boldsymbol{\omega}}^\top}{\partial\alpha}\left(-\frac{2}{\alpha^2}\boldsymbol{\Sigma}^{-1}+\frac{1}{2}\boldsymbol{I}_n\right)\boldsymbol{A}+\boldsymbol{V}_{\boldsymbol{\omega}}^\top\left(\frac{4}{\alpha^3}\boldsymbol{\Sigma}^{-1}\boldsymbol{A}+\left(-\frac{2}{\alpha^2}\boldsymbol{\Sigma}^{-1}+\frac{1}{2}\boldsymbol{I}_n\right)\frac{\partial\boldsymbol{A}}{\partial\alpha}\right)\\ &= \boldsymbol{D}^\top\left(\frac{-1}{\alpha^2}\boldsymbol{\Sigma}^{-1}+\frac{1}{4}\boldsymbol{I_n}\right)\boldsymbol{A}+\boldsymbol{V}_{\boldsymbol{\omega}}^\top\left(\frac{4}{\alpha^3}\boldsymbol{\Sigma}^{-1}\boldsymbol{A}+\left(-\frac{2}{\alpha^2}\boldsymbol{\Sigma}^{-1}+\frac{1}{2}\boldsymbol{I_n}\right)\boldsymbol{A}\left(\frac{1}{\alpha^2}\boldsymbol{\Sigma}^{-1/2}+\frac{1}{4}\boldsymbol{\Sigma}^{1/2}\right)\boldsymbol{A}\right),\end{aligned}$$

em que, $\boldsymbol{D}=(D_1,\ldots,D_n)^\top$, com $D_j=\boldsymbol{L_j\omega}$ e $\boldsymbol{L_j}$ a j-ésima linha da matriz

$$\boldsymbol{A}\left(\frac{1}{\alpha^2}\boldsymbol{\Sigma}^{-1/2}+\frac{1}{4}\boldsymbol{\Sigma}^{1/2}\right)\boldsymbol{A},\quad j=1,\ldots,n.$$

Pertubação em uma covariável

Neste caso, $\frac{\partial\ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial\boldsymbol{\omega}^\top}=\beta_1\boldsymbol{V}_{\boldsymbol{\omega}}^\top\left(\frac{2}{\alpha^2}\boldsymbol{\Sigma}^{-1}-\frac{1}{2}\boldsymbol{I}_n\right)\boldsymbol{A}$, em que $\boldsymbol{V_\omega}=(\boldsymbol{V_{\omega_1}},\ldots,\boldsymbol{V_{\omega_n}})^\top$, com $\boldsymbol{V_{\omega_i}}=\operatorname{senh}\left(\frac{y_i-\boldsymbol{X}_i^\top\boldsymbol{\beta}-\beta_1\boldsymbol{A}_i\boldsymbol{\omega}}{2}\right),i=1,\ldots,n$ e $\boldsymbol{A}=\left(\frac{1}{\alpha}\boldsymbol{\Sigma}^{-\frac{1}{2}}-\frac{\alpha}{4}\boldsymbol{\Sigma}^{\frac{1}{2}}\right)^{-1}$.

Assim, as derivadas de segunda ordem que compõem a matriz $\boldsymbol{\Delta}$ são dadas por

$$\boldsymbol{\Delta_\beta}=\frac{\partial\ell^2(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial\boldsymbol{\beta}\partial\boldsymbol{\omega}^\top}=\left(\frac{\partial\ell^2(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial\beta_0\partial\boldsymbol{\omega}^\top},\ldots,\frac{\partial\ell^2(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial\beta_p\partial\boldsymbol{\omega}^\top}\right)^\top,$$

Em que

$$\begin{aligned}\frac{\partial\ell^2(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial\beta_1\partial\omega^\top} &= \boldsymbol{V}_{\boldsymbol{\omega}}^\top\left(\frac{2}{\alpha^2}\boldsymbol{\Sigma}^{-1}-\frac{1}{2}\boldsymbol{I}_n\right)\boldsymbol{A}+\beta_1\frac{\partial\boldsymbol{V}_{\boldsymbol{\omega}}^\top}{\partial\beta_1}\left(\frac{2}{\alpha^2}\boldsymbol{\Sigma}^{-1}-\frac{1}{2}\boldsymbol{I}_n\right)\boldsymbol{A}\\ &\approx \boldsymbol{V}_{\boldsymbol{\omega}}^\top\left(\frac{2}{\alpha^2}\boldsymbol{\Sigma}^{-1}-\frac{1}{2}\boldsymbol{I}_n\right)\boldsymbol{A}-\frac{\beta_1}{2}\left(\boldsymbol{X_1}+\boldsymbol{A\omega}\right)\left(\frac{2}{\alpha^2}\boldsymbol{\Sigma}^{-1}-\frac{1}{2}\boldsymbol{I}_n\right)\boldsymbol{A}\\ &= \left(\boldsymbol{V}_{\boldsymbol{\omega}}^\top-\frac{\beta_1}{2}\left(\boldsymbol{X_1}+\boldsymbol{A\omega}\right)\right)\left(\frac{2}{\alpha^2}\boldsymbol{\Sigma}^{-1}-\frac{1}{2}\boldsymbol{I}_n\right)\boldsymbol{A};\\ \frac{\partial\ell^2(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial\beta_j\partial\omega^\top} &= \beta_1\frac{\partial\boldsymbol{V}_{\boldsymbol{\omega}}^\top}{\partial\beta_j}\left(\frac{2}{\alpha^2}\boldsymbol{\Sigma}^{-1}-\frac{1}{2}\boldsymbol{I}_n\right)\boldsymbol{A}\\ &= \beta_1\left(U_{\omega_1}\left(-X_{1j}\right),\ldots,U_{\omega_n}\left(-X_{nj}\right)\right)\left(\frac{2}{\alpha^2}\boldsymbol{\Sigma}^{-1}-\frac{1}{2}\boldsymbol{I}_n\right)\boldsymbol{A}\\ &\approx -\beta_1\boldsymbol{X}_{\cdot j}^\top\left(\frac{2}{\alpha^2}\boldsymbol{\Sigma}^{-1}-\frac{1}{2}\boldsymbol{I}_n\right)\boldsymbol{A},\quad j=0,\ldots,p;j\neq 1\end{aligned}$$

Cabe lembrar que $\boldsymbol{X}_{.0} = \mathbf{1}$.

$$\boldsymbol{\Delta}_{\boldsymbol{\varphi}} = \frac{\partial \ell^2(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \boldsymbol{\varphi} \partial \boldsymbol{\omega}^\top} = \left(\frac{\partial \ell^2(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \varphi_1 \partial \boldsymbol{\omega}^\top}, \frac{\partial \ell^2(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \varphi_2 \partial \boldsymbol{\omega}^\top}, \frac{\partial \ell^2(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \varphi_3 \partial \boldsymbol{\omega}^\top} \right)^\top,$$

em que, para $j = 1, 2, 3$ tem-se,

$$\begin{aligned}
\frac{\partial \ell^2(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \varphi_j \partial \omega^\top} &= \beta_1 \frac{\partial \boldsymbol{V}_{\boldsymbol{\omega}}^\top}{\partial \varphi_j} \left(\frac{2}{\alpha^2} \boldsymbol{\Sigma}^{-1} - \frac{1}{2} \boldsymbol{I}_n \right) \boldsymbol{A} + \beta_1 \boldsymbol{V}_{\boldsymbol{\omega}}^\top \left(-\frac{2}{\alpha^2} \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_j} \boldsymbol{\Sigma}^{-1} \boldsymbol{A} + \left(\frac{2}{\alpha^2} \boldsymbol{\Sigma}^{-1} - \frac{1}{2} \boldsymbol{I}_n \right) \frac{\partial \boldsymbol{A}}{\partial \varphi_j} \right) \\
&\approx -\frac{\beta_1^2}{2} \left(\frac{\partial \boldsymbol{A}}{\partial \varphi_j} \boldsymbol{\omega} \right)^\top \left(\frac{2}{\alpha^2} \boldsymbol{\Sigma}^{-1} - \frac{1}{2} \boldsymbol{I}_n \right) \boldsymbol{A} \\
&+ \beta_1 \boldsymbol{V}_{\boldsymbol{\omega}}^\top \left(\frac{2}{\alpha^2} \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_j} \boldsymbol{\Sigma}^{-1} \boldsymbol{A} + \left(\frac{2}{\alpha^2} \boldsymbol{\Sigma}^{-1} - \frac{1}{2} \boldsymbol{I}_n \right) \frac{\partial \boldsymbol{A}}{\partial \varphi_j} \right) \\
&= -\frac{\beta_1^2}{2} \boldsymbol{\omega}^\top \boldsymbol{A} \left(\frac{1}{\alpha} \boldsymbol{\Sigma}^{-\frac{1}{2}} \frac{\partial \boldsymbol{\Sigma}^{\frac{1}{2}}}{\partial \varphi_j} \boldsymbol{\Sigma}^{-\frac{1}{2}} + \frac{\alpha}{4} \frac{\partial \boldsymbol{\Sigma}^{\frac{1}{2}}}{\partial \varphi_j} \right) \boldsymbol{A} \left(\frac{2}{\alpha^2} \boldsymbol{\Sigma}^{-1} - \frac{1}{2} \boldsymbol{I}_n \right) \boldsymbol{A} \\
&+ \beta_1 \boldsymbol{V}_{\boldsymbol{\omega}}^\top \left(\frac{2}{\alpha^2} \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_j} \boldsymbol{\Sigma}^{-1} \boldsymbol{A} \left(\frac{2}{\alpha^2} \boldsymbol{\Sigma}^{-1} - \frac{1}{2} \boldsymbol{I}_n \right) \boldsymbol{A} \left(\frac{1}{\alpha} \boldsymbol{\Sigma}^{-\frac{1}{2}} \frac{\partial \boldsymbol{\Sigma}^{\frac{1}{2}}}{\partial \varphi_j} \boldsymbol{\Sigma}^{-\frac{1}{2}} + \frac{\alpha}{4} \frac{\partial \boldsymbol{\Sigma}^{\frac{1}{2}}}{\partial \varphi_j} \boldsymbol{A} \right) \right).
\end{aligned}$$

$$\begin{aligned}
\Delta_\alpha &= \frac{\partial \ell^2(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \alpha \partial \omega^\top} = \beta_1 \frac{\partial \boldsymbol{V}_{\boldsymbol{\omega}}^\top}{\partial \alpha} \left(\frac{2}{\alpha^2} \boldsymbol{\Sigma}^{-1} - \frac{1}{2} \boldsymbol{I}_n \right) \boldsymbol{A} + \beta_1 \boldsymbol{V}_{\boldsymbol{\omega}}^\top \left(-\frac{4}{\alpha^3} \boldsymbol{\Sigma}^{-1} \boldsymbol{A} + \left(\frac{2}{\alpha^2} \boldsymbol{\Sigma}^{-1} - \frac{1}{2} \boldsymbol{I}_n \right) \frac{\partial \boldsymbol{A}}{\partial \alpha} \right) \\
&\approx -\frac{\beta_1^2}{2} \left(\frac{\partial \boldsymbol{A}}{\partial \alpha} \boldsymbol{\omega} \right)^\top \left(\frac{2}{\alpha^2} \boldsymbol{\Sigma}^{-1} - \frac{1}{2} \boldsymbol{I}_n \right) \boldsymbol{A} + \beta_1 \boldsymbol{V}_{\boldsymbol{\omega}}^\top \left(-\frac{4}{\alpha^3} \boldsymbol{\Sigma}^{-1} \boldsymbol{A} + \left(\frac{2}{\alpha^2} \boldsymbol{\Sigma}^{-1} - \frac{1}{2} \boldsymbol{I}_n \right) \frac{\partial \boldsymbol{A}}{\partial \alpha} \right) \\
&= -\frac{\beta_1^2}{2} \boldsymbol{\omega}^\top \boldsymbol{A} \left(\frac{-1}{\alpha^2} \boldsymbol{\Sigma}^{\frac{-1}{2}} - \frac{1}{4} \boldsymbol{\Sigma}^{\frac{1}{2}} \right) \boldsymbol{A} \left(\frac{2}{\alpha^2} \boldsymbol{\Sigma}^{-1} - \frac{1}{2} \boldsymbol{I}_n \right) \boldsymbol{A} \\
&\quad + \beta_1 \boldsymbol{V}_{\boldsymbol{\omega}}^\top \left(\frac{-4}{\alpha^3} \boldsymbol{\Sigma}^{-1} \boldsymbol{A} + \left(\frac{2}{\alpha^2} \boldsymbol{\Sigma}^{-1} - \frac{1}{2} \boldsymbol{I}_n \right) \boldsymbol{A} \left(\frac{-1}{\alpha^2} \boldsymbol{\Sigma}^{\frac{-1}{2}} - \frac{1}{4} \boldsymbol{\Sigma}^{\frac{1}{2}} \right) \boldsymbol{A} \right) \\
&= \left(-\frac{\beta_1^2}{2} \boldsymbol{\omega}^\top \boldsymbol{A} + \beta_1 \boldsymbol{V}_{\boldsymbol{\omega}}^\top \right) \boldsymbol{A} \left(\frac{-1}{\alpha^2} \boldsymbol{\Sigma}^{\frac{-1}{2}} - \frac{1}{4} \boldsymbol{\Sigma}^{\frac{1}{2}} \right) \boldsymbol{A} \left(\frac{2}{\alpha^2} \boldsymbol{\Sigma}^{-1} - \frac{1}{2} \boldsymbol{I}_n \right) - \frac{4\beta_1}{\alpha^3} \boldsymbol{V}_{\boldsymbol{\omega}}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{A}.
\end{aligned}$$

6 CONSIDERAÇÕES FINAIS

Os resultados desta tese foram apresentados em forma de artigos. Apresenta-se aqui as principais contribuições verificadas durante o desenvolvimento do trabalho:

- O trabalho contribui com a geoestatística uma vez que propôs um novo modelo espacial, baseado na distribuição Birnbaum-Saunders, para o estudos de dados com distribuição amostral assimétrica positiva. Em um primeiro momento, foi proposto um modelo espacial sem a utilização de covariáveis (Artigo 1) e depois propõe-se um modelo espacial BS utilizando-se um número finito de covariáveis (Artigo 2). Em ambos os casos, foram desenvolvidas técnicas para a estimação dos parâmetros do modelo por máxima verossimilhança;

- O trabalhou apresentou a definição da distribuição log-Birnbaum-Saunders n-variada e uma descrição detalhada das principais propriedades desta distribuição;

- No trabalho foram desenvolvidas técnicas, global e local, de diagnósticos para o modelo BS. No estudo de influência local, a perturbação mais adequada para o modelo espacial BS foi determinada;

- Avaliou-se a performance do processo de estimação e das ferramentas de diagnóstico utilizando simulação. O processo de estimação apresentou boa performance para amostras de tamanho ($n = 100$). A simulação para estudo de pontos influentes se mostrou bastante satisfatória.

- Na análise de dados, casos influentes foram detectados e um estudo de similaridade entre o mapa original e mapas construidos a partir da retirada de pontos influentes foi realizado e mostrou mudanças consideráveis.

- A predição de valores foi realizada por krigagem com drift externo e krigagem ordinária. Em um estudo de caso, a predição por krigagem com drift externo apresentou melhor qualidade, em conformidade com o relatado pela literatura.

Além das contribuições apresentadas acima, surgiram ao longo do desenvolvimento da tese alguns possíveis problemas a serem abordados em estudos futuros:

- Propor e abordar um modelo espacial BS baseado em uma distribuição com caudas mais pesadas, como a distribuição t-Student, com o objetivo de reduzir a influência de dados atípicos;

- Utilizar outros esquemas de perturbação para avaliar a influência de pontos e

- Utilizar o modelo proposto para realizar outros estudos na área agrícola.